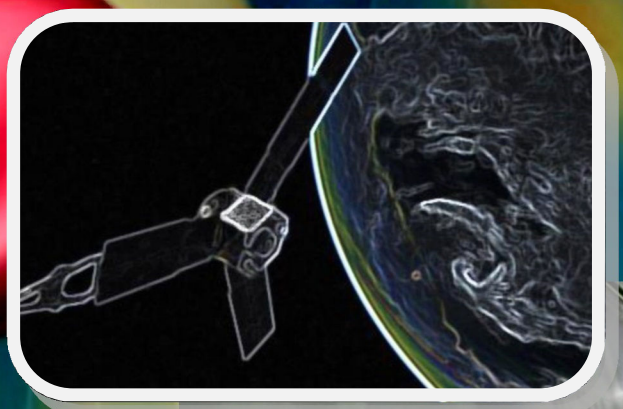
MODELLING FOR ENGINEERING AND HUMAN BEHAVIOUR 2015

Instituto Universitario de Matemática Multidisciplinar



L. Jódar, L. Acedo and J. C. Cortés (Editors)

Instituto Universitario de Matemática Multidisciplinar

im² Instituto de Matemática Multidisciplinar

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

GENERALITAT VALENCIANA

MODELLING FOR ENGINEERING, & HUMAN BEHAVIOUR 2015

Instituto Universitario de Matemática Multidisciplinar
Universitat Politècnica de València
Valencia 46022, SPAIN

Edited by
Lucas Jódar, Luis Acedo and Juan Carlos Cortés
Instituto Universitario de Matemática Multidisciplinar
Universitat Politècnica de València
I.S.B.N.: 978-84-608-5355-8

CONTENTS

Modelling the flyby anomaly in a Whiteheadian theory of gravity

L. Acedo *

Instituto Universitario de Matemática Multidisciplinar,

Building 8G, Door C, Second Floor,

Universitat Politècnica de Valéncia, 46022 Valencia, Spain

November 30, 2015

1 Introduction

Back in 1990 NASA engineers noticed that the Galileo spacecraft flyby of the Earth exhibited an anomalous behaviour. The Deep Space Network records the Doppler frequency shifts of radio signals as a measured of the velocity of these spacecraft. After substracting all possible sources of perturbations (including the Sun, the Moon, other planets, etc) a residual postencounter increase in the frequency of radiosignals persisted. This can be interpreted as an increase of 3.92 mm/sec in the final velocity of the Galileo spacecraft in this particular flyby [1]. Later on, a similar anomaly showed up in subsequent flybys performed by the Galileo spacecraft in 1992, NEAR in 1998, Cassini in 1999 and the Rosetta and Messenger spacecraft in 2005. It is expected that the recent Juno flyby whose perigee was achieved on September, 10^{th}, 2013 would also exhibit the anomaly but, apparently, the data is still not analyzed [2].

Among the conventional effects considered to explain the anomaly, some authors have calculated the order of magnitude of atmospheric friction, tides, charge and magnetic moment of the spacecraft, Earth albedo and Solar wind

*e-mail: luiacrod@imm.upv.es

[3]. Detailed calculations of the effects of the zonal harmonics in the improved Earth gravity models [2] and the gravitomagnetic effect predicted by General Relativity [4] have also proved insufficient to explain the anomalous increases or decreases of asymptotic velocities. Possible explanations based upon new ideas have also been considered in recent literature such as the interaction of the spacecraft with an halo of dark matter surrounding the Earth [5], a strong gravitomagnetic field generated by the Earth and following the celestial parallels [6], anomalous couplings to the gravitational potential vector of linearized General Relativity [7] or a modification of the classical Newtonian potential of a spherical planet by an offset of the gravitational centre [8]. All these unconventional ideas are highly speculative and they are not supported by a theoretical framework.

Consequently, after 25 years of research this problem is still unsolved and it has even been featured in the list of unsolved problems in physics provided by Wikipedia [9].

In this work we have considered a recently proposed extension of Whitehead's theory of gravity [10, 11] as a possible avenue to explain the flyby anomaly.

2 The extended Whitehead-Bel theory

The famous philosopher and mathematician A. N. Whitehead was also a recognized author in the field of theoretical physics. Although not so well-known as his work on philosophy of mathematics he proposed an alternative to Einstein's theory of General Relativity in 1922 in a book entitled *The Principle of Relativity* [12]. The basic concept in Whitehead's theory is the null vector joining the retarded position of the source of gravity and the test particle:

$$L^\alpha = x^\alpha - \hat{x}^\alpha \ . \tag{1}$$

In terms of the source's four-velocity, u^α and the four-vector $l^\alpha = 1/u_\alpha L^\alpha L^\alpha$ we define the following symmetric covariant tensor:

$$g_{\mu\nu} = \eta_{\mu\nu} - \frac{2Gm}{c^2 r} l_\mu l_\nu \ , \tag{2}$$

where $\eta_{\mu\nu}$ is the diagonal Minkowski's metric: $\eta_{00} = -1$, $\eta_{11} = \eta_{22} = \eta_{33} = 1$, $\eta_{\mu\nu} = 0$, $\mu \neq \nu$. Here, $g_{\mu\nu}$ plays the role of the metric tensor in the curved spacetime General Relativity (GR) but Whitehead's considered the

background spacetime as Minkowsikian. With these tools we can recover the predictions for the classical tests of GR [13].

Recently, Bel has proposed an extension of Whitehead's theory based upon the more general symmetric covariant tensor we can build in terms of linear combinations using the four-velocity, u^μ, the retarded position vector, l^μ and Minkowski's metric tensor, $\eta_{\mu\nu}$ as follows:

$$g_{\mu\nu} = \eta_{\mu\nu} + \frac{1}{r}\left(A_0 u_\mu u_\nu + A_1 \eta_{\mu\nu} + A_2\left(l_\mu u_\nu + l_\nu u_\mu\right) - A_3 l_\mu l_\nu\right) \ , \qquad (3)$$

where A_0, ..., A_3 are constants determined by two conditions:(i) Consistency with Einstein's vacuum equations and (ii) The Newtonian limit. This yields $A_0 = 2(A_1 - A_2$ and $A_1 - A_3 = 2GM$. G being the gravitational constant and M the mass of the source [10, 11, 14].

3 Results and Conclusions

The equations of motion for the model described in the previous section can be obtained from the geodesics [14]. In the limit of low velocities (the source move at speed small compared with the speed of light) we obtain that the total force exerted upon a test particle at a distance D from the center of the Earth and with declination angle δ is given by:

$$F^1 = -\frac{Gm}{D^2}\cos\delta \ , \qquad (4)$$

$$F^2 = -\frac{Gm}{D^2}\left(\frac{\omega R^2}{5Dc}\right)\cos\delta \ , \qquad (5)$$

$$F^3 = -\frac{Gm}{D^2}\sin\delta \ , \qquad (6)$$

where R is the Earth's radius, m its mass, c the speed of light and ω the angular velocity corresponding to Earth's rotation around its axis. The components in Eqs. (4) and (6) are give the classical Newtonian force but there is an extra component in Eq. (5) not predicted by GR or Newtonian theory corresponding to a circulating field of force which follows the celestial parallels.

Although the Newtonian forces acting upon a rigid spherical body cannot modify its angular momentum, it can be shown that the circulating field in Eq. (5) allows for a transfer of energy and angular momentum from the Earth to the spacecraft or viceversa providing a mechanism to explain the reported flyby anomalies. For a quantitative analysis of the predictions compared with the observations for six flybys the interested reader is referred to Ref. [14].

References

[1] J. D. Anderson, J. K. Campbell, J. E. Ekelund, J. Ellis, J. F. Jordan, Anomalous orbital-energy changes observed during spacecraft flybys of the Earth, Phys. Rev. Lett. 100 (2008) id. 091102.

[2] L. Iorio, A flyby anomaly for Juno ?. Not from standard physics, Adv. Space Res. 54(11) (2014) 2441-45. arXiv:1311.4218

[3] C. Lämmerzahl, O. Preuss, H. Dittus, Is the physics of the Solar System really understood ? in Lasers, Clocks and Drag-Free Control, Astrophys. and Space Science Library 349 (2008) 75-101. arXiv:gr-qc/0604052

[4] L. Iorio, The effect of General Relativity on hyperbolic orbits and its applicatioin to the flyby anomaly, Scholarly Research Exchange 2009, id. 807695, 1-8.

[5] S. L. Adler, Can the flyby anomaly be attributed to earth-bound dark matter ?, Phys. Rev. D 79 (2009) 023505. arXiv:0805.2895v4

[6] L. Acedo, The flyby anomaly: A case for strong gravitomagnetism ?, Adv. Space Res. 54 (2014) 788-796. arXiv:1505.06884

[7] M. J. Pinheiro, The flyby anomaly and the effect of a topological torsion current, Phys. Lett. A 378 (2014) 3007-11.arXiv:1404.1101

[8] K. Wilhelm, B. N. Dwivedi, Anomalous Earth flybys of spacecraft, Astrophys. Space Sci. 358(18) (2015) 1-8.

[9] List of unsolved problems in physics, Wikipedia, https://en.wikipedia.org/wiki/List_of_unsolved_problems_in_physics (accessed November 18$^{\text{th}}$, 2015).

[10] Ll. Bel, A look inside the theory of the linear approximation, arXiv:gr-qc/0605057v3, 2007.

[11] Ll. Bel, A new look inside the theory of the linear approximation: Gravity assists and Flybys, unpublished paper. Available online at http://www.lluisbel.com/upload/OnHold/FlyBys.pdf, 2015.

[12] A. N. Whitehead, *The Principle of Relativity*, Cosimo, Inc., New York, 2007.

[13] A. J. Coleman, Whitehead's principle of Relativity, arXiv:physics/0505027v2, 2005.

[14] L. Acedo, The flyby anomaly in an extended Whitehead's theory, Galaxies 3, 2015, 113-128.

Supply schedule in intermittent water supply based on quantitative and qualitative criteria

A. E. Ilaya-Ayza *, R. Pérez-García,
J. Izquierdo, and J. Benítez

Fluing-Instituto de Matemática Multidisciplinar (IMM)

Universitat Politècnica de València,

Camino de Vera SN, pc: 46015, Valencia, Spain

November 30, 2015

1 Introduction

Intermittent water supply (IWS) is a form of access to water in which the water is supplied only a few hours per day. This leads to problems of equity of supply [1, 2], infrastructure deterioration [3] and water quality [4]. Despite these problems, IWS is very common in developing countries [5], where continuous supply is very difficult due to insufficient funding, physical scarcity or mismanagement [6].

Generally, water companies that manage IWS have reduced economic resources. Therefore, proposals to improve the system performance must be based on alternatives that involve reduced human and economic resources.

If the network is not sectorised, IWS is simultaneous for all users. When the network is sectorised, sectors have non-simultaneous delivery schedules.

One problem of IWS is the peak flow occurring at certain times of day. This value is usually greater than the peak flow in a system with continuous water supply (CWS). This reduces pressure and flow at the ends or high points and causes inequity in the supply and complaints from users.

*e-mail: amilay@upv.es

Reorganizing supply schedules based on qualitative and quantitative technical criteria reduces the peak flow and consequently improves pressures.

Water company experts' opinion is very important in the optimization process. Thus, pairwise comparison matrices defined in the Analytic Hierarchy Process (AHP) [7] method are used to quantify this opinion.

2 The Proposed Methodology

In an IWS the pattern of supply tends to be constant by the presence of household deposits. Large flows occur in the first minutes but they are reduced at the end of the delivery period. The flow variation is not large, ranging from 20% to 30%. Therefore, it can be simplified by calculating the average volume per delivery period (Vs) [8] or supply blocks.

$$Vs_j = \frac{\text{daily volume supplied to sectors}_\text{j}}{\text{number of supply hours}}.$$

The optimization process assigns a new schedule to these supply blocks based on the defined technical requirements. In this paper we use four criteria, three quantitative and one qualitative.

Pressure (C1): sectors with lower operating pressure are those that can change their schedule, and users with lower pressure accept the measures taken to improve their conditions of service. Number of users (C2): the more users the sector have, the lower the possibility of modifying its supply schedule. The aim is to improve the service and reduce the number of people affected by the change. Supply hours (C3): the sectors that have fewer supply hours have greater flexibility in assigning a new schedule.

The qualitative criterion is related to the ease of operation of the sector (C4). It depends on various factors: the availability of sectorisation valves, their proper performance, their accessibility, the working difficulty for operators, the mainly manual operation, complaints from the users, and others. Thus, this qualitative variable is consulted to experts from the water company with the AHP methodology. The weight of this variable in every sector is the normalized geometric mean of the eigenvectors of the pairwise comparison matrices of each expert. Experts are also consulted to establish the weight of each criterion. Then the weight associated to number of users (wu_j), pressure (wp_j), supply hours (wh_j), and ease of operation (wo_j) for each i sector are calculated. These weights are used in the objective function.

We use integer linear programming (LP) with binary variables. As a result of the optimization process we obtain an $m \times n$ matrix of binary variables $V = (v_{kl})$, where m is the number of hours of a day, and n is the total number of sectors. A value of 1 in the matrix indicates the start of the supply period of the sector.

To reallocate supply schedules in a block (all supply period), we propose to use the $m \times m$ schedule square matrix $U^{(j)} = u_{kl}$, where

$$u_{kl} = \begin{cases} 1 & k = 1, \ldots, m \ \wedge \ l = k, k+1, \ldots, k+h-1, \\ 1 & \text{if } k+h-1 > m, \text{ then } l = 1, 2, \ldots, k+h-1-m, \\ 0 & \text{otherwise.} \end{cases}$$

to calculate supply schedule vectors for each sector j:

$$X_j = \left(\sum_{i=1}^{m} v_{ij} u_{i1}^{(j)}, \sum_{i=1}^{m} v_{ij} u_{i2}^{(j)}, \ldots, \sum_{i=1}^{m} v_{ij} u_{im}^{(j)} \right) = (x_{1j}, x_{2j}, \ldots, x_{mj}).$$

Multiplying the hourly volume Vs_j by the supply schedule vector X_j correspondingly, we obtain the volume delivered to each sector in the corresponding period represented by vector B_j.

$$B_j = Vs_j \cdot (x_{1j}, x_{2j}, \ldots, x_{mj}) = (b_{1j}, b_{2j}, \ldots, b_{mj}).$$

To start with an initial priority in the current schedule for each sector, we propose an $m \times n$ matrix $S = (s_{ij})$ where $s_{ij} = 1$ for an hour with water supply and $s_{ij} = 0$ for an hour without water supply. We must include in the objective function the information of volume entering the tank per hour. The normalized value is t_i. The hour in which more water enters the tank is prioritized. There is also the possibility of sectors operating in cascade. Therefore, we consider the supply hour s_{ix} and supply volume x_{iy} of the sector in cascade.

These elements allow us to configure the optimization problem.

$$\text{Maximize } \sum_{i=1}^{m} \left(\sum_{j=1}^{n} (wp_j + wu_j + wh_j + wo_j + s_{ij}) x_{ij} + t_i \sum_{j=1}^{n} x_{ij} + s_{ix} \cdot x_{iy} \right)$$

$$\text{Subject to } \sum_{i=1}^{m} v_{ij} = 1 \ \forall \ j; \quad \sum_{j=1}^{n} x_{kj} \leq SS \ \forall \ k; \quad \sum_{i=1}^{n} b_{ki} \leq VS \ \forall \ k.$$

The first constraint forces one supply period per sector. In the second constraint the number of sectors that can work simultaneously, SS, is limited.

The output volume of the tank, VS, is the most important constraint in the optimization process. This value allows us to rearrange the supply schedule of each sector and to find scenarios with better service conditions.

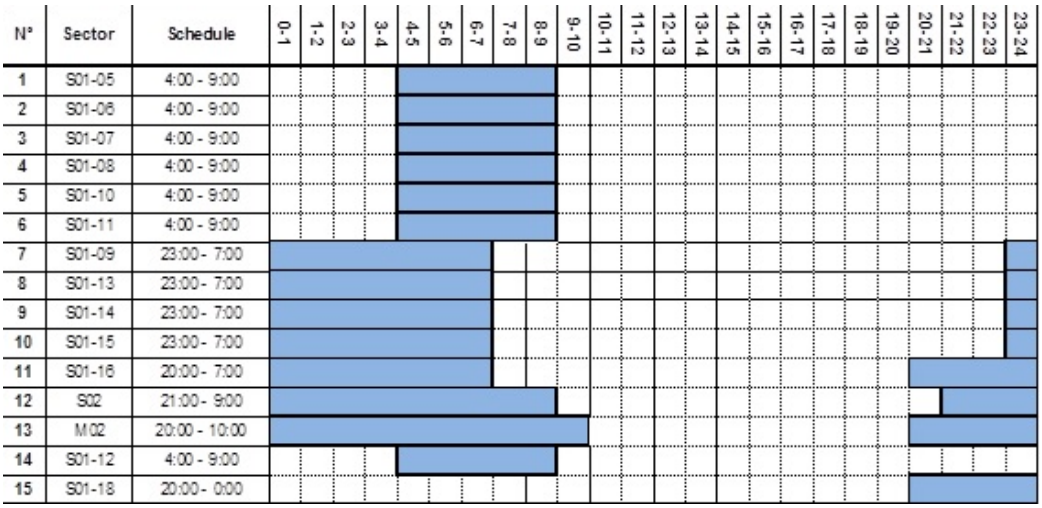


Figure 1: Supply schedule of each sector, Oruro south subsystem

3 Example of implementation

The proposed methodology is applied to the south subsystem of the supply network of Oruro city (Bolivia). This network is configured by fifteen sectors with a single feed point (tank). Daily volume supplied, current supply schedule (Figure 1), number of supply hours, pressure, number of users of each of the sectors and the volume entering the tank each hour (Figure 2) are known. The optimization process selects sectors that must change their schedules and moves them to an optimal schedule based on quantitative and qualitative criteria. Thereby, we reduce the peak flow that is characteristic of IWS systems. An example of a restriction of 350 m^3/h reconfiguration schedule is shown in Figure 3. Pressures in sectors improve after reducing the peak flow and changing some supply schedules.

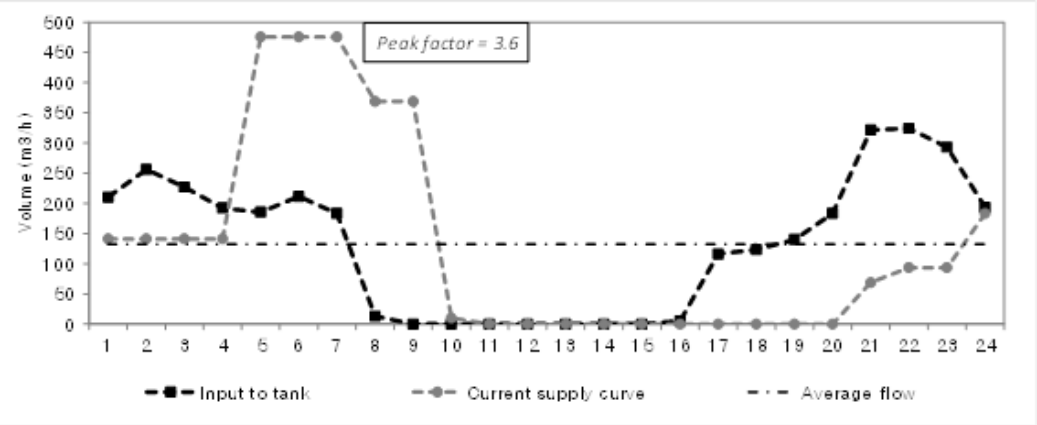


Figure 2: Volume entering the tank and IWS current supply curve

4 Conclusion

Supply schedule management does not seek to perpetuate IWS. It is intended as a short-term technical management solution, which seeks to improve the conditions of service and therefore improve quality of life. It is also a useful tool for a gradual transition from IWS to CWS.

Solutions to make IWS systems work better don't necessarily require the construction of new infrastructure. IWS systems with insufficient funding have few economic resources. Solutions must be found based on existing infrastructures and demanding minimal human and economic resources.

As a part of the LP problem, we propose the use of schedule matrices, which play an important role in the solution. They allow change the sector schedule in blocks. Therefore, they are very useful for this kind of problems.

References

[1] K. Vairavamoorthy and K. Elango. Guidelines for the design and control of intermittent water distribution systems, *Waterlines*, 21: 19–21, 2002.

[2] K. Vairavamoorthy, S. D. Gorantiwar, and A. Pathiranaa. Managing urban water supplies in developing countries; Climate change and water scarcity scenarios, *Physics and Chemistry of the Earth*, 33: 330–339, 2008.

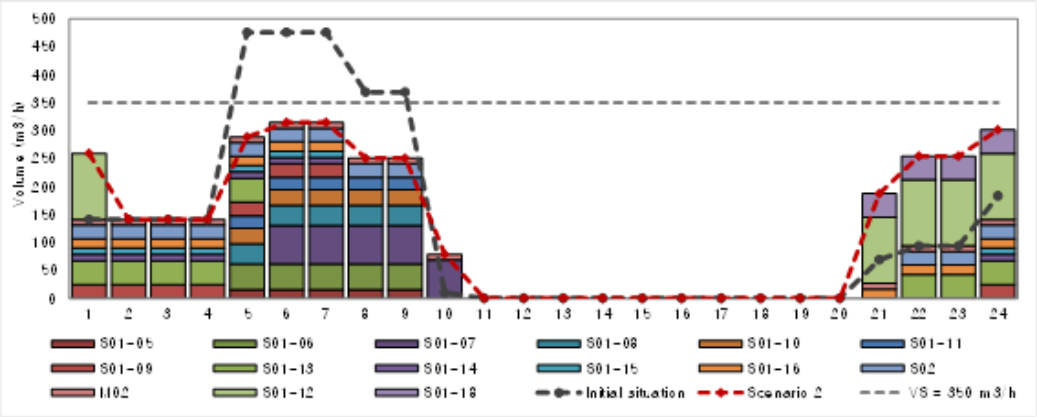


Figure 3: Optimizing supply schedules with constraint $VS = 350$ m^3/h

[3] B. Charalambous. The Effects of Intermittent Supply on Water Distribution Networks, *Water Loss 2012 Conference*, Manila, Philippines, 2012.

[4] E. Kumpel and K. L. Nelson, Comparing microbial water quality in an intermittent and continuous piped water supply, *Water Research*, 47: 5176–5188, 2013.

[5] C. Van den Berg and A. Danilenko, The IBNET Water Supply and Sanitation Performance Blue Book. Washington D.C., The World Bank, 2011.

[6] N. Totsuka, N. Trifunovic, and K. Vairavamoorthy, Intermittent urban water supply under water starving situations, *30th WEDC International Conference*, Vientiane, Lao, 2004.

[7] T. L. Saaty and L. Vargas, Models, Methods, Concepts & Applications of the Analytic Hierarchy Process. New York, Springer, 2012.

[8] J. A. Cabrera-Béjar and V. G. Tzatchkov, Inexpensive Modeling of Intermittent Service Water Distribution Networks, *World Environmental and Water Resources Congress*, 2009.

Internal Lubricant Content in inhalation Capsules

G.Ayala[♭*], F. Díez[†], M.T.Gassó[‡] and B.E. Jones[♮]

(♭) Department of Statistics and O.R. University of Valencia

Avda. Vicent Andrés Estellés, 1, 46100-Burjasot (Spain),

(†) Qualicaps Europe,

Qualicaps Europe S.A.U. 28108 Spain,

(‡) Institute of Multidisciplinar Mathematics and Department of Mathematics

Polythecnical University of Valencia, Camino de Vera s/n 46022. Valencia (Spain),

(♮) School of Pharmacy & Pharmaceutical Sciences, Cardiff University, Cardiff, UK.

November 30, 2015

1 Introduction

Hard capsules are manufactured in a continuous process on large automatic machines, see Figure 1. They are formed on stainless steel mould pins mounted in-line onto metal strips (bars). There are different sets of bars to make the caps and bodies of each size of capsule. Groups of bars are dipped in to a temperature controlled container, called a dip pan, containing a warm aqueous solution of the polymer, either gelatin or hypromellose (HPMC). Films are formed on the mould pins most commonly by a gelation process that relies on the temperature difference between the cold pin and the hot solution. This is an inherent property of gelatin solutions and HPMC solutions are formulated to gel by the addition of a network former such as carrageenan and potassium chloride as a promoter [2]. The bars are raised out of the dip-pan and are rotated end over end to improve the film distribution on the pins as they are transferred from the lower level of the machine to the upper one. At this point the films have set and are no longer mobile. Groups of bars are moved by hydraulic pushers through a series of drying kilns, which use large volumes of controlled humidity and temperature to dry the films. At the end of the upper level the bars are transferred to the lower level and

*guillermo.ayala@uv.es;Fdiez@qualicaps.es; mgasso@mat.upv.es,JBE2@cardiff.ac.uk

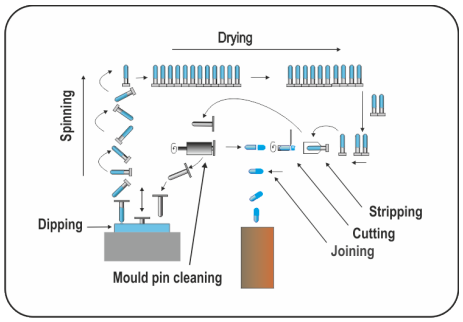


Figure 1: Capsules manufacturing (From Qualicaps Europe).

are moved back to the front-end of the machine. When the pins emerged from the kilns they are dried to a level of >16.0%, which is just above the upper level of the standard moisture content specification. These dried films adhere strongly to the pins. The next part of the process is to strip them from the pins using metal jaws. The internal lubricant content (ILC) is a critical factor enabling this to occur without capsule damage. If insufficient is used the capsule shells will split during removal. Pairs of bars, one cap and one body are selected from each side of the machine and enter into the automatic section. The lubricant is a propriety mixture pharmaceutical grade excipients and is different for each capsule manufacturer and their compositions are registered in the companies Drug Master File. Lubricant is loaded into a pump, the flow rate from which can be adjusted using a pressure valve. The lubricant is applied to a circular foam roller that transfers a sufficient quantity to the pins as they pass underneath. The pin bars are moved towards the centre of the machine and the pins are inserted into rotating circular tubes lined with a felt pad. These clean the pins and spread the lubricant evenly over their surface. These pads are changed at regular intervals to avoid a build-up and saturation with the lubricant [1, 2, 3, 4, 5].

Several papers have described the influence of ILC on aerosolization [6, 7]. The reference [7] showed that there is an optimum ILC range to obtain good powder release from capsules as measured by their emitted dose and fine particle fraction [7]. They suggested that the effect could be related to the roughness of capsule internal surface.

2 Results

The experts have evaluated the most important factors in the process control. Our study have two main goals. Firstly, the evaluation of which of these factors are important i.e. their effects are relevant for the response variable, internal lubricant content (ILC from now on). Our experimental design has two categorical factors (pump flow and pin location) and one numeric covariable (time) with a sample size of 432 corresponding to three levels for each experimental factors, 24

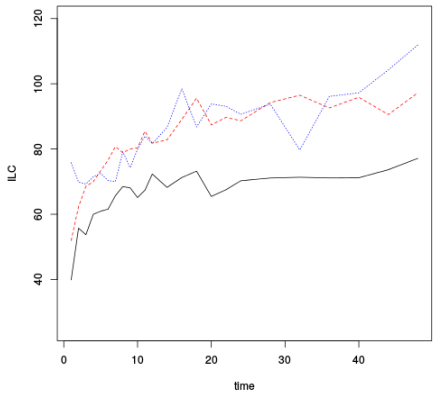


Figure 2: Observed means for different pump flows: low (solid line), medium (dashed line) and high (dotted line).

instants for the covariable time and two replications. The statistical analysis was perfomed using the software environment R. We evaluate if there are significant factors or if there exists any relevant interaction between them. All pairwise comparisons of means and variances for each categorical factor have been performed and the results appears in table 1.

Table 1: Comparison of means (t-test) and variances(F-test)

		Dif.	95% CI	p_1	Ratio	95% CI	p_2
Pump flow	1-2	-16.66	[-20.14,-13.18]	0.00	0.61	[0.44,0.85]	0.00
	2-3	-1.69	[-5.56,2.17]	0.39	1.00	[0.72,1.40]	0.98
	1-3	-18.35	[-21.83,-14.88]	0.00	0.62	[0.45,0.86]	0.00
Location	Bar4-Pi1	-9.43	[-13.12,-5.73]	0.00	0.74	[0.53,1.03]	0.07
	Pi1-Pi28	-2.32	[-6.47, 1.83]	0.27	0.84	[0.60,1.17]	0.29
	Bar4-Pi28	-11.75	[-15.64,-7.85]	0.00	0.62	[0.45,0.86]	0.00

A full factorial linear model have been fitted. A variable selection was applied using and stepwise approach in order to minimize the Akaike information criterium. The final selected model has only the main effects (for pump flow and location) and two interaction, pump flow with time and pump flow with location.

Secondly, a functional data approach has been used. The response is now the observed ILC along time and the predictors are pump flow and location. The mean functions for different pump flows (respectively locations) were compared using a functional anova. Significant differences are observed for pump flow factor but not for location.

Acknowledgements

This research was supported by Qualicaps Europe, S.A.U.; Spanish DGI Grant no. MTM2014-58159-P (M.T. Gassó) and G. Ayala (DPI2013-47279-C2-1-R).

References

[1] B.E. Jones. Quali-V®-I: a new key for dry powder inhalers. *Drug Delivery Technology*, 3(6):52–57, 2003.

[2] B.E. Jones. *Manufacture and properties of two- piece hard capsules*, pages 79–100. In: Podczeck, F., Jones, B.E. (Eds), Pharmaceutical Press, London, 2nd edition, 2004.

[3] B.E. Jones. The evolution of DPI capsules. *Inhalation*, 2(6):20–23, 2008.

[4] S. Nagata. Advantages to HPMC capsules. A new generation's hard capsule. *Drug Deliv. Technol.*, 2:32–42, 2002.

[5] T. Ogura, Y. Furuya, and S. Matsuura. HPMC capsules , an alternative to Gelatin. *Pharm. Technol. Eur.*, 10:32–42, 1998.

[6] S. Saim and S.T. Horhota. Process for overcoming drug retention in hard gelatin inhalation capsules. *Drug development and industrial pharmacy*, 28: 641–654, 2002. doi: 10.1081/DDC-120003855.

[7] I.Y. Saleem, F. Diez, B.E. Jones, N. Kayali, and L. Polo. Investigation on the aerosol performance of dry powder inhalation hypromellose capsules with different lubricant levels. *International Journal of Pharmaceutics*, 492(1-2), 258 – 263, 2015. doi: http://dx.doi.org/10.1016/j.ijpharm.2015.07.034.

Microscopic and macroscopic models for gas leak detection

Fidel Aznar[♭] [*], Mar Pujol[♭], and Ramón Rizo[♭]

(♭) Department of Computer Science and Artificial Intelligence.

University of Alicante, San Vicent del Raspeig, Alicante (E-03080). Spain

November 30, 2015

1 Introduction

Fracking for natural gas and oil extraction significant leaks of methane, a potent greenhouse gas. This extraction technique proliferate across world, and scientists are raising questions about whether millions of gallons of contaminated drilling fluids could be threatening water supplies, affecting human health and increasing greenhouse effects.

Currently, many companies are using methane-leak detection tools, such as infrared cameras, that are too labor-intensive and fail to find many sources. Researchers are developing more sensitive methane sensors for leak detection that utilize cavity ring-down spectroscopy (CRDS). These sensors can discern between oil-and-gas-related methane emissions and those from biogenic sources, such as cattle. A principle advantage of the sensor is its simple design, which allows for a lighter weight and less expensive, less complicated systems. This makes the sensor suitable for large-scale deployment in both stationary systems and unmanned aerial vehicle and drone applications.

One of these applications is presented in [1], where a methane sensor mounted on a remotecontrolled aircraft is developed. The aircraft was used to quantify emission rates from well pads and a compressor station in Texas,

[*]Contact author mail: fidel@dccia.ua.es. This work has been supported by the Spanish Ministerio de Economia y Competitividad, project TIN2013-40982-R

USA. In another paper presented in [2], a drone have been used for detecting fugitive methane emissions. Althought currently there are several applications of UAV for methane detection we have not found specific swarm behaviours for this task, which is specially relevant for large area navigation and data recollection.

2 Microscopic behaviour

In this paper we will present a swarm behaviour, which monitors gas leaks by using a swarm of homogeneous UAV. Thus, this system is fully distributed, scalable and highly fault tolerable. Our main goal, once the behaviour is designed, is to determine the ability of the swarm to locate, converge and follow a gas leak. Therefore, a macroscopic model to predict the global behaviour of the swarm and to verify its performance will be specified. More specifically, we propose a homogeneous behaviour, executed by all agents, consisting of two states. Initially, drones look for any trace of the leak in the environment. Once the leak is detected, the drone will head to it. Then, the drone will try to stay in its perimeter. The initial state is *Discover*, since we initially assume that the position of the leak is unknown. The transition from *Discover* state is performed when the agent's methane sensor detects a leak. In this case, the new state will be *Gas*.

$$\begin{aligned} \boldsymbol{v}_D(t) = &\quad \boldsymbol{v}_D(t-1) + \boldsymbol{rand} \cdot \mu_1 \\ \boldsymbol{v}_G(t) = &\quad \left\| \sum_{s \in S} \left((\boldsymbol{pos}(s) - \boldsymbol{pos}(rob)) \cdot s \right) \right\| + \\ &\left\| \sum_{i=1}^{|R|} (\boldsymbol{pos}(r_i) - \boldsymbol{pos}(rob)) \right\| + \boldsymbol{rand} \cdot \mu_2 \end{aligned}$$

Where v_D and v_G define the velocity to be executed in the *Discover* and *Gas* behaviours respectively. μ parameters define the importance of random perturbations for discovering and action executions using a gaussian random generator with mean 0 and variance 1. S is a set that contains the most intense readings for our methane sensor in a period of 1 minute. R is a set that contains all the swarm drones.

3 Macroscopic behaviour

In this work, we consider the framework proposed in [4] in order to obtain the probability distribution of the swarm position for any time t. This will enable us to predict, in great detail, the behaviour of the overall system. As described by [4], once the microscopic behaviour has been defined, the global behaviour of the system can be calculated using the Fokker-Planck equation. This equation provides a method to statistically model a swarm of robots based on modelling techniques of multi-particle systems from the field of quantum physics. From a Langevin equation, that represents the behaviour of a single particle, the Fokker-Planck equation is derived for all the system.

$$\frac{\partial \rho(\boldsymbol{r},t)}{\partial t} = -\nabla\left(\boldsymbol{A}(\boldsymbol{r},t)\rho(\boldsymbol{r},t)\right) + \frac{1}{2}Q\nabla^2\left(B^2(\boldsymbol{r},t)\rho(\boldsymbol{r},t)\right) \tag{1}$$

As we have already seen in [3], the FokkerPlanck equation implements the necessary abstraction of microscopic details as described above and treats rapidly-changing parameters such as noise. It is important to underline that the equation is still exact if this noise is generated by a Gaussian process, that is, if it is fully determined by the first two moments. It gives the temporal evolution of the probability density describing the positions of the agents.

A function determines the displacement of the swarm. A depends primarily on a vector representing the directional information. A potential field P is commonly used to define it. In our case, we need to establish a function that takes into account the following things based on the proposed microscopic model: the random motion states of the robot; the probability that a movement of an agent fails in its execution (e.g. due to a collision); and the potential field where the robots move. Although it is possible to model a probability distribution for each state, as our microscopic model has no interaction between agents (except purely physical, as collisions) and the behaviour of states is relatively simple, the macroscopic behaviour of the swarm can be comprised in a single distribution.

$$\boldsymbol{A} = \gamma\frac{\nabla m_{leak}(\boldsymbol{r},t)}{||\nabla m_{leak}(\boldsymbol{r},t)||}\boldsymbol{v} \tag{2}$$

Where γ is a normalization term and m_{leak} is a potential field obtanied from the expected leak evolution through the environment.

Function B describes the nondeterministic motion and, therefore, it takes into account the random motion of agents. Two forces, that must be considered, take part in the microscopic behaviour. On the one hand, some influences derived from agents that are on *Discover* and *Gas* states. On the other hand, the behaviour itself causes that the environment has areas with a higher density of agents. In these areas the probability of collision can be increased depending on the density of agents at a given time.

4 Experimentation

The simulation of the microscopic model has been developed using the software MASON. We use a swarm of 100 agents randomly distributed by the environment that moves uniformly at 60km/h. The simulation uses small size drones ($< 3m^2$) and a environment size of $1km^2$.

We have carried out 200 random leak tests through the environment to verify the correct operation of our model at local level. Initially, we check the convergence of the swarm for a single static leak. We obtain that all the robots stay within the gas leak or in their perimeter (distance to the leak < 10 meters) 98% of the tests. Next, using a gas leak simulation tools we verify de convergence of the swarm with several active leaks (that change and evolve each simulated time step) with another 200 tests. For this experimentation we simulate variable winds from 0 to 5m/s. In this case we found that 78% of the times one or more drones are able to detect and follow the leak.

5 Conclusions

Taking into account the properties of methane emissions we have presented a microscopic model for a swarm of drones, capable of monitoring these leaks properly. We have also provided a macroscopical model, based on Fokker-Planck equations that can be used to predict the overall evolution of the swarm. This mathematical model will calculate the likelihood of an agent to be placed in a position at a given time. The experimental results presented show the proper evolution of our system and their ability to detect and follow a simulated gas leak.

References

[1] Amir Khan, David Schaefer, Lei Tao, David J. Miller, Kang Sun, Mark A. Zondlo, William A. Harrison, Bryan Roscoe and David J. Lary. *Low Power Greenhouse Gas Sensors for Unmanned Aerial Vehicles.* Remote Sens. 2012, 4, 1355-1368.

[2] Kathryn McKain, Adrian Down, Steve M. Raciti, John Budney, Lucy R. Hutyra, and Cody Floerchinger, Scott C. Herndon, Thomas Nehrkorn, Mark S. Zahniser, Robert B. Jackson, and others. *Methane emissions from natural gas infrastructure and use in the urban region of Boston, Massachusetts.* Proceedings of the National Academy of Sciences, 2015, 12 (7), 1941-1946.

[3] Heiko Hamann. *Space-Time Continuous Models of Swarm Robotic Systems: Supporting Global-to-Local Programming*, volume 9. Springer, 2010.

[4] Heiko Hamann and Heinz Wörn. A framework of space–time continuous models for algorithm design in swarm robotics. *Swarm Intelligence*, 2(2-4):209–239, 2008.

Modelling the survival of the Spanish construction SME in a crisis environment

Isabel Barrachina; Elena De la Poza.

Centro de Ingeniería Económica

Universitat Politècnica de València

Abstract

The construction sector headed the Spanish economy during the blooming period 1998-2007; its direct contribution to the Spanish economic growth amounts over the 20% in annual terms.

The turmoil of the financial sector impacted the Spanish construction companies, mainly due to their high indebtedness levels combined with the credit shortage from the banking system to units, sellers and buyers. As a result a relevant number of construction companies declared in bankruptcy producing a dominos effect impacting the whole Spanish economy. In this context, taking into account financial-economic information we classify the small medium sized construction companies by their financial distress levels (solvent; on risk of default).

Then, after identifying those companies which survived to the economic recession we model the processes of survival of the Spanish SME to find out possible financial-economic patterns that explain their continuity and to propose guidelines to strategy decision makers.

Keywords: survival; financial distress; construction sector; SME.

1. Introduction

The construction sector in Spain experienced a large expansion since 2000 becoming the main engine of the Spanish economy, (ICE, 2009).

There was an increase in housing prices and unprecedented growth of mortgage debt.

The global economic crisis (2008) caused a decline in mortgage approvals granted by banks it led to a clear increase of developers and construction companies declared bankrupt. However, some companies achieve to survive the financial downsize.

The primary goal of this study is to identify the strategies followed by the Spanish construction companies on default risk at the beginning of the economic recession that managed to survive. Also, the specific objectives are:

- Develop a descriptive analysis of Spanish construction companies and its trend throughout the period of crisis [2008, 2013]

-Classify the enterprises in 2008 according to their economic-financial risk

-Analyze the survival capability of companies.

2. Methods

2.1. Sources of information

Employing as primary source of information the database SABI, we selected those companies whose ownership form was a corporation, N= 125,599.

Following, we selected those companies only belonging to the building sector according the CNAE (code 412), N=87,565. Then, we reduced our sample selecting only those Spanish companies whose situation was active, default or bankruptcy at the starting period of study (2008), N=3,185.

Then, our sample was reduced to N=2,444 taking into account only those companies which a maximum level of annual incomes for the period 2008-2014 was 50 million Euros. After that, we identified those firms that were operating since 2000, (N=2,098).

Finally, those companies with missing data for the period of study (2008-2013) were removed from our sample; As a result our sample was composed by N=1,434 companies.

2.2 Methodology

2.2.1 Principal component analysis (PCA)

To identify the economic and financial dimensions of the company that mainly differentiate them, we apply the PCA that allow us to build artificial variables from the original information. This technique is employed for reducing the volume of information by avoiding any possible correlation between the chosen variables.

The ratios employed in the PCA:

Solvency Ratio (SR): Equity capital x 100 / total assets.
Working capital needs
General Liquidity (GL): Current assets/current liabilities.
(WCP) Working capital/ total assets
(ROA) Return on assets
(ROE) Return on equity (Net income/ Shareholder's equity)
ROCE Return on capital employed: (EBIT/ Capital employed).
RSF Return on Shareholders Funds
Indebtedness' percentage (IP) : total liabilities/capital employed
Gearing: Long term liabilities / Equity
Interest cover: Margin profit/Financial expenses

2.2.2. Multiple discriminant analysis (MDA)

In order to find the mathematical model that better discriminate the companies into two groups (default risk companies at year (n+1) versus no- risk companies at year (n+1)) by their financial ratios estimates at year n. Hypothesis: the companies are on default's risk when their solvency ratio is lower than 1.5 (Amat, 2008).

2.2.3. Survival analysis

The reason to employ survival analysis is the necessity to take into account the variable time (since other regression techniques such as the logistic regression does not) when studying the behavior of companies. In particular, we want to identify the time that takes for a company declare in default as a result of its economic-financial trend.

We chose Kaplan-Meier, a non-parametric estimate of the survival function. This methodology is commonly used to describe survivorship of study population/s by a intuitive graphical presentation.

This method calculates the probability of survival every time a company suffers the event of bankruptcy, from the number of companies that are subject to bankruptcy due to the crisis. The factor that differentiates one business from others is the initial classification of companies according to their forecasted risk of default next year

3. Results

3.1 Descriptive Analysis and PCA results

At 2008, the starting year of the study, the 85% firms of our sample were functioning (Total N=1,434). While at the end of the period of study (2013), the number of companies operating was amounted to 1,210 which implies a drop of 15.6%. Derived from the application of the PCA technique the variables ROE, IP, LP and WCP were those with greater discriminant power, (Cohen *et al*., 2003)

3.2. MDA Results

Table 1 shows the standardized canonical discriminant Function coefficients with the variables that better discriminate our sample into two groups: default companies at (n+1) versus those that are not. The analysis diagnoses correctly the 77,8% companies.

Table 1.

Variables	Standardized Canonical discriminant Function coefficients	Canonical discriminant function coefficients
ROE	0,060	0,002
IP	0,898	0,043
LP	-0,017	-0,001
WCP	-0,760	-2,815
Constant		-1,388

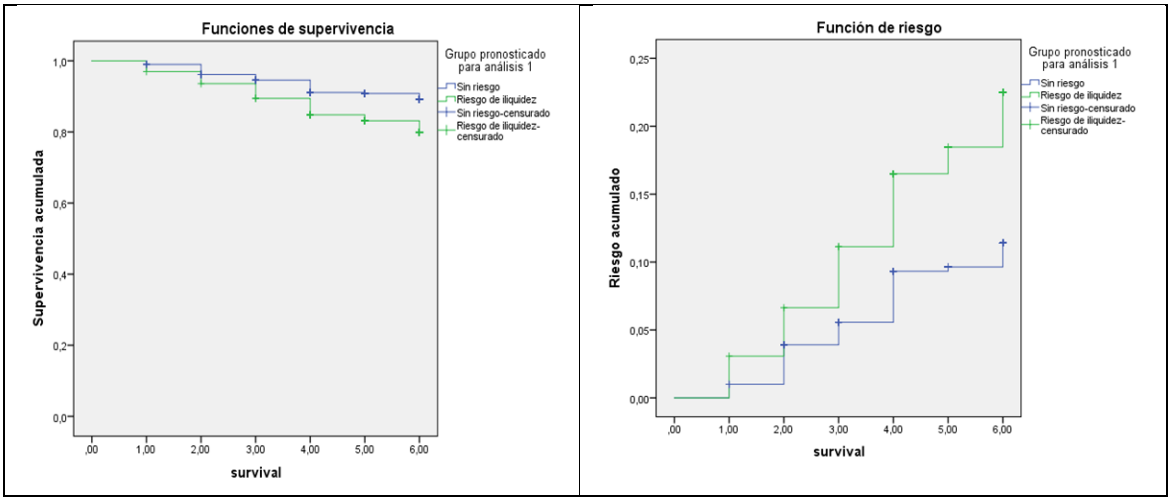
Table 1 shows the IP and the WCP contribute greater than the rest of variables to the discriminant function, (Altman, 1968).

3.3. Survival Analysis

In survival analysis the total number of companies is 1,267 of which 55.72 % had no default risk forecast while 44 % did. Also, 10.48% of companies without risk failed during the crisis, while the 19.61% of companies on risk declared in bankruptcy.

Figure 1 shows how those companies that did not show risk of default initially get better chance of survival, the difference increases from the third year, when companies without default risk do not declare in bankruptcy while others do.

Figure 1



4. Conclusions

The Spanish building sector has been severely damaged by the crisis, as many companies have closed and all have suffered as reflected in the evolution of its economic and financial ratios.

The MDA shows the variables that more influence on the firm's default risk are the working capital firms' percentage.

The MDA is a useful tool to predict the companies on risk to failure since classifies correctly about the 92% of companies of the sample.

Throughout this study it has been determined that the chances of survival of companies in a crisis environment decrease if the initial situation is of risk especially from the third year of crisis environment.

During the monitoring period 10.48% SME that initially were not classified as risk of default declared in bankruptcy especially during the first three years, and 19.61% of companies classified as default risk.

The difference between the chances of survival between the two groups of firms increases from the third year. 50% of the companies that survive the crisis whose initial situation was risky do in a bad financial economic situation, while the other half significantly improve their ratios.

Preliminary results show that its main strategy seems to be to improve the state of liquidity through debt restructuring and the decrease in financial expenses due to lower interest rates.

References

Altman, Edward I. (1968): "Financial ratios, discriminant analysis and the prediction of corporate

Amat, Oriol (2008): "Análisis de Estados Financieros. Fundamentos y aplicaciones"

Información Comercial Española, ICE: Revista de economía, ISSN 0019-977X, Nº 850, 2009 (Ejemplar dedicado a: La primera crisis global: procesos, consecuencias, medidas) , págs. 23-40

Cohen, J., Cohen, P. C., West, S. G., & Aiken, L. S. (2003). (3rd Ed.) Applied multiple regression/correlation analysis for the behavioral sciences. Mahwah, NJ.: Lawrence Erlbaum.

SABI Database [Last access: July 2015].

An algorithm for quasi-linear control problems in the economics of renewable resources: The steady state and end state for the infinite and long-term horizon.

L. Bayón[1]*, P.J. García-Nieto[1], R. García-Rubio[2], J.A. Otero[1] and C. Tasis[1]

(1) Department of Mathematics, University of Oviedo, Spain

(2) Department of Business & Administration, University of Salamanca, Spain

1 Introduction

This paper presents the problem of finding the optimal harvesting strategy (see [1], [2] and [3]), maximizing the expected present value of total revenues. The problem is formulated as an optimal control problem [4]. Combining the techniques of Pontryagin's Maximum Principle and the shooting method, an algorithm has been developed that is not affected by the values of the parameter. The algorithm is able to solve conventional problems as well as cases in which the optimal solution is shown to be bang-bang with singular arcs. In addition, we present a result that characterizes the optimal steady-state in infinite-horizon, autonomous models (except in the discount factor) and does not require the solution of the dynamic optimization problem. We also present a result that, under certain additional conditions, allows us to know a priori the final state solution when the optimization interval is finite. Finally, several numerical examples are presented to illustrate the different possibilities of the method.

*e-mail: bayon@uniovi.es

2 Statement of the problem

For the study of the economics of a renewable resource [2], we shall first see the pattern of biological growth of the resource. In this paper, we consider the growth function for a population of some species of fish. We assume that this fishery has a *intrinsic growth rate* denoted by r, which represents the difference between the population's birth and natural mortality rates. Let us assume that the *population stock* is x, and the *rate of change of the population is* $\dot{x}$. A commonly used functional form is the *simple logistic function*:

$$\dot{x}(t) = f_l(x) = rx(t)\left(1 - \frac{x(t)}{k}\right) \tag{1}$$

where k is the carrying capacity of the species. In this paper, and in line with [3], we model the dynamics of the fish stock biomass (x) when human harvesting is included in the problem, in the more general form as:

$$\dot{x}(t) = f_l(x) - h(t) \tag{2}$$

where $h(t)$, the rate of biomass harvest, will be considered as a independent variable. Let us now see how to model the cost functions.

Let $\pi(x, h)$ be the instantaneous net revenue from the harvest of the stock biomass:

$$\pi(x, h) = p(h)h - c(x, h) \tag{3}$$

where $p(h)$ is the inverse demand function and $c(x, h)$, the cost function associated with the harvest. The functional forms for the demand and cost functions adopted in this paper are:

$$p(h) = p_0 - p_1 h \tag{4}$$

$$c(x, h) = \frac{ch^\alpha}{x} \tag{5}$$

where h represents landings of fish and p_0 and p_1 are coefficients. Substituting (4) and (5) in (3), the profit function is:

$$\pi(x, h) = p_0 h - p_1 h^2 - \frac{ch^\alpha}{x} \tag{6}$$

where the meaning of the parameters is: p_0 is the price of the stock, p_1 is the strength of demand, c is the cost of exploitation and α is the harvest cost parameter.

Our model of renewable resource exploitation is an open-access fishery model, in which each firm takes the market price of landed fish as given. The firm's objective is to maximize profits from the harvest schedule over an infinite time horizon, subject to the dynamic constraint equation (2) and other natural and policy restrictions that involve limits (or bounds) for the harvest, $h(t)$, and stock, $x(t)$. Hence, our objective is to maximize profit from the harvest schedule over an infinite time horizon:

$$\max_{h(t)} \int_0^\infty \pi(x,h)e^{-\delta t}dt = \max_{h(t)} \int_0^\infty (p_0 h - p_1 h^2 - \frac{ch^\alpha}{x})e^{-\delta t}dt \tag{7}$$

subject to:

$$\dot{x}(t) = f_l(x) - h(t);\ x(0) = x_0 \tag{8}$$

$$h(t) \in H(t);\ x \in [0,k] \tag{9}$$

where $\delta > 0$ is the discount rate, i.e. the marginal returns on capital for the company, and x_0 is the initial stock level.

As can be seen, the stated problem (7), (8), (9) is one of Optimal Control (OC) that presents a number of noteworthy features. First, the optimization interval is infinite. Second, the time t is not explicitly present in the problem (time-autonomous problem), except in the discount factor. Third, we impose constraints on the control and, fourth, it constitutes a problem which is quasi-linear when real values are considered for the parameters.

3 Optimization Algorithm

Faced with the complication of having to use different techniques when the functional is linear or nonlinear in the control variable, the contribution of our method is that it is valid in cases ranging between quasi-linearity and singular arcs. We have used the combined techniques of Pontryagin's Maximum Principle (PMP) [4] and the shooting method to build this optimization algorithm. If we denote by $\mathbb{Y}_x(t)$ the *coordination function*:

$$\mathbb{Y}_x(t) = -\frac{F_u}{f_u} \cdot e^{\int_0^t f_x ds} + \int_0^t F_x \cdot e^{\int_0^s f_x dz} ds \tag{10}$$

the theoretical development carried out allows us to present a necessary maximum condition.

Theorem 1. A necessary maximum condition

Let u^* be the optimal control, let $x^* \in \widehat{C}^1$ be a solution of the above problem. Then there exists a constant $K \in \mathbb{R}$ such that:

$$\begin{array}{lll} \textit{If } u_{\min} < u^* < u_{\max} & \Longrightarrow & \mathbb{Y}_{x^*}(t) = K \\ \textit{If } u^* = u_{\max} & \Longrightarrow & \mathbb{Y}_{x^*}(t) \leq K \\ \textit{If } u^* = u_{\min} & \Longrightarrow & \mathbb{Y}_{x^*}(t) \geq K \end{array} \tag{11}$$

Thus, the problem consists in finding for each K the function x_K that satisfies: $x_K(0) = x_0$, the conditions of Theorem 1 and, from among these functions, the one that satisfies the transversality condition:

$$\lim_{t\to\infty} \lambda(t) = 0 \tag{12}$$

The algorithm consists of two fundamental steps:

Step 1) *The construction of x_K.* The construction of x_K can be performed using a discretized version of the *coordination equation:* $\mathbb{Y}_x(t) = K$. For each K, we construct the x_K, using this equation and when the values obtained do not obey the constraints, we force the solution to belong to the boundary until the moment established by conditions of Theorem 1.

Step 2) *The calculation of the optimal K.* The calculation of the optimal K could be achieved by means of an adaptation of the shooting method. Varying the coordination constant, K, we search for the extremal that fulfils the second boundary condition (12). Starting out from two values for the coordination constant, K: $K_{\min}$ and $K_{\max}$ and using a conventional method such as the secant method, our algorithm converges satisfactorily.

4 Steady-state solution

For time-autonomous problems, where the time, t, is not explicitly present in the problem, except in the discount factor, the optimal solution is time invariant in the long term and converges to an equilibrium state. The method developed in [5] characterizes the optimal steady-state in single-state, infinite-horizon problems, by means of a simple function of the state variable, called the *evolution function.*

The method consider the one-dimensional, infinite-horizon problems of

the form:

$$\max_{u(t)} J = \int_0^\infty G(x(t), u(t)) e^{-\delta t} dt \tag{13}$$

$$\dot{x}(t) = f\left(x(t), u(t)\right), \quad x(0) = x_0 \tag{14}$$

For a steady-state solution, $u = R(x)$, the evolution function, is defined by:

$$L(x) = \delta \left(\frac{G_u(x, R(x))}{f_u(x, R(x))} + \dot{W}(x) \right) \tag{15}$$

with:

$$W(x) = \frac{1}{\delta} G(x, R(x)) \tag{16}$$

The function $L(x)$ serves to formulate the following necessary condition for the location of the optimal steady state x_s:

$$L(x_s) = 0 \tag{17}$$

To the best of our knowledge, the problem stated in this paper has never been addressed using this approach. The necessary condition (17), for our problem, is:

$$L(x_s) = \delta \left(-\pi_h(x_s, f_l(x_s)) + \dot{W}(x_s) \right) = 0 \tag{18}$$

The above equation can be solved for x_s, allowing us to obtain the harvest equilibrium, h_s. We can thus know the steady-state solution a priori, without solving the dynamic problem.

5 Long-term horizon: the end state

Inspired by the previous section, in this section we present a new result that allows us to calculate a priori the final value that is reached when the optimization interval is not infinite (once again, without the need to solve the dynamic problem). To do so, we shall use the same model as above, but with an optimization interval $[0, T]$, assuming that it is long enough for the steady state to be reached in its development. In order to obtain the result, we need to make an additional assumption: the system must be autonomous, and hence we must consider $\delta = 0$. Let us consider the following problem:

$$\max \int_0^T F\left(x(t), \dot{x}(t)\right) dt = \max_{h(t)} \int_0^T \pi(x, h) dt \tag{19}$$

Based on the Euler's equation, that can be rewritten for autonomous systems as follows:

$$F - \dot{x}F_{\dot{x}} = cte \tag{20}$$

and given that the solution for the steady state, (x_s, h_s) (with $\dot{x} = 0$), may be known a priori by means of the method explained in the previous section, the value of the constant, cte, present in (20) can be obtained straightforwardly:

$$cte = F(x_s) = \pi(x_s, h_s) \tag{21}$$

If we now consider the final moment, T, the two following conditions must be simultaneously verified at that moment, given that the end state is free:

$$\begin{cases} F_{\dot{x}}(x(T), h(T)) = 0 \\ F(x(T), h(T)) = cte \end{cases} \tag{22}$$

The first is the transversality condition corresponding to the free end state, and the second the simplified Euler equation (20). Simply solving this system, the end state $(x(T), h(T))$ can be obtained straightforwardly.

References

[1] C.W. Clark, Mathematical Bioeconomics: The optimal management of renewable resources, Wiley, New York, 1990.

[2] R. Perman, Y. Ma, J. McGilvray and M. Common, Natural Resource & Environmental Economics, Pearson Education, Harlow, England, 2003.

[3] S. Agnarsson, R. Arnason, K. Johannsdottir, L. Ravn-Jonsen, L.K. Sandal, S.I. Steinshamn and N. Vestergaard, Comparative evaluation of the fisheries policies in Denmark, Iceland and Norway: multispecies and stochastic issues, SNF Report No 25/07, Bergen, Norway, 2008.

[4] L.S. Pontryagin, Mathematical Theory of Optimal Processes (Classics of Soviet Mathematics), CRC Press, 1987.

[5] Y. Tsur and A. Zemel, The infinite horizon dynamic optimization problem revisited: A simple method to determine equilibrium states, *European Journal of Operational Research*, 131(3): 482-490, 2001.

A polynomial expansion method based on Helmholtz equation for the Neutron Diffusion Equation discretized by the Finite Volume Method

A. Bernal[‡] [*], J.E. Roman[†], R. Miró[‡], and G. Verdú[‡]

(‡) Instituto de Seguridad Industrial, Radiofísica y Medioambiental(ISIRYM),

Universitat Politècnica de València, Camí de Vera s/n,

(†) Departamento de Sistemas Informáticos y Computación,

Universitat Politècnica de València, Camí de Vera s/n.

November 30, 2015

1 Introduction

The solution of the Neutron Diffusion Equation (NDE) is the easiest way to determine the spatial distribution of the neutron flux in nuclear reactors. Although this equation is a simplification of the neutron transport equation using the Fick's Law [1], it is an integro-differential equation depending on temporal and spatial terms. In reactor physics, one habitually eliminates the temporal dependence for solving the steady state by transforming the equation into an eigenvalue problem. In spite of this, geometrical discretization and numerical methods are required to solve the spatial derivatives of the NDE.

The Finite Volume Method (FVM) can be easily applied to unstructured meshes and is typically used in the transport equations [2]. Moreover, the

[*]e-mail: abernal@iqn.upv.es

FVM is feasible and suitable to be applied to the NDE [3].

One can obtain accurate results with typical algorithms of the FVM applied to NDE for fine meshes, but they require high computational resources [3]. In addition, a polynomial expansion method can be used to obtain accurate results for coarse meshes to accelerate the calculation [4]. By means of this method, the neutron flux is expanded in each cell of the discretized geometry, as a sum of a finite set of basic polynomial terms, which are assigned previously and their constant coefficients are determined by solving the eigenvalue problem by means of SLEPc [5]. SLEPc is a library for solving eigenvalue problems in which the associated matrices are large and sparse, such as those arising after the discretization of partial differential equations.

In this paper, another set of polynomial terms, which are obtained by applying the Helmholtz equation to the 3D neutron flux expansion, are used to obtain better results in coarser meshes to accelerate the calculation. The outline of the paper is as follows. Section 2 presents the methodology used. Section 3 describes the reactors used and shows the results. Section 4 contains the conclusions about the results.

2 Methodology

The method follows an iterative process for the eigenvalue ($\boldsymbol{k}$) :

1. Calculation of the Helmholtz coefficient (λ), for each cell, depending on ($\boldsymbol{k}$) and the coefficients of the NDE.
2. Calculation of the Helmholtz coefficients for separation of variables: λ_x , λ_y , λ_z.
3. Update the polynomial expansion of the neutron flux depending on λ_x , λ_y , λ_z.
4. Calculate the volume and surface averaged values of the polynomials and the surface averaged values of the gradient of the polynomials.
5. Solve the eigenvalue problem of the NDE to obtain $\boldsymbol{k}$.

3 Results

The eigenvalue ($\boldsymbol{k}$) and the power (P) will be the evaluated variables. The power in each cell i is a weighted sum of the neutron flux in i and is defined in Equation 1. The power is normalized to accomplish the Mean Power (MP) equals unity. The mean power is a weighted sum of the power in each cell with not null power, which is defined in Equation 2.

Relative errors are used to evaluate the main variables: Eigenvalue Error (EE) and Power Error (PE) , defined in Equations 3 and 4. With the aim of reducing the extension of this paper, the Mean Power Error (MPE) will be used instead of the Power Error for each cell, which is calculated as a weighted sum of the power of all the cells and is defined in Equation 5.

$$P_i = \left(\Sigma^i_{f,1} \cdot \phi_{1,i} + \Sigma^i_{f,2} \cdot \phi_{2,i}\right) \cdot constant \tag{1}$$

$$MP = \frac{\sum_{i=1}^{N} P_i \cdot V_i}{\sum_{i=1}^{N} V_i} \tag{2}$$

$$EE(pcm) = \frac{|\boldsymbol{k} - \boldsymbol{k}_{ref}|}{\boldsymbol{k}_{ref}} \cdot 10^5 \tag{3}$$

$$PE_i(\%) = \frac{|P_i - P_{i_{ref}}|}{P_{i_{ref}}} \cdot 100 \tag{4}$$

$$MPE(\%) = \frac{\sum_{i=1}^{N} PE_i \cdot |P_i| \cdot V_i}{\sum_{i=1}^{N} |P_i| \cdot V_i} \tag{5}$$

Two reactors were simulated, a homogeneous and a heterogeneous one. The results of this method (HELMHOLTZ) will be compared with the results of the polynomial expansion method with simple polynomials (PEM) [4].

3.1 Homogeneous reactor

This reactor is a parallelepiped of the following dimensions: 99 cm x 60 cm x 180 cm. It is composed of only one material. Boundary conditions of zero flux have been imposed and the calculation has been performed in a structured mesh of 3x3x6. Since this reactor has a simple geometry and is composed of only one material, this case has analytical solution and it was

Table 1: Results of the Homogeneous reactor

	HELMHOLTZ	**PEM**
Number of iterations	4	1
Time (s)	0.300	0.163
EE (pcm)	66.00	716.64
MPE (%)	0.00	0.00

the reference. The reference eigenvalue is 0.99339.

The results are shown in Table 1. One can see that EE decreases substantially, obtainining accurate results,that is, EE below 100 pcm. As regards the computational time, it increases due to the increase of the number of iterations, but is low enough.

3.2 Langenbuch reactor

It is a heterogeneous reactor composed of four materials. A quarter of this reactor is shown in Figure 1. Boundary conditions of zero flux have been imposed and the calculation of the full reactor has been performed in a structured mesh of 11x11x10. The reference solution is calculated with the code PARCS, which is a well-known neutron diffusion code in nuclear industry. The reference eigenvalue calculated is 0.9950.

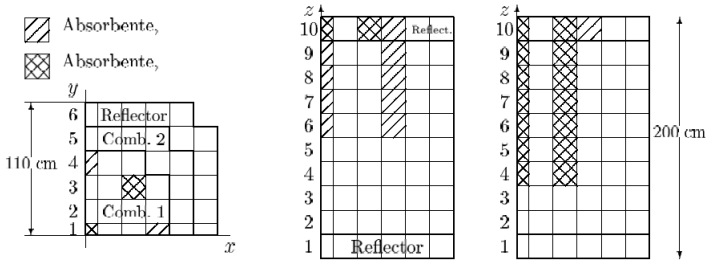


Figure 1: Cross sections of a quarter of Langenbuch reactor

The results are shown in Table 2. One can appreciate that both errors decrease, but this decrease is not substantial. However, if one pays attention to tha axial power errors, which are shown in Table 3, one will see that the error decreases more than 1 percent for the second and nineth axial levels. Moreover, in Tables 4 and 5, which display the radial power errors for each method, the diferences of the error are even higher.

Table 2: Results of Langenbuch reactor

	HELMHOLTZ	**PEM**
Number of iterations	3	1
Time (s)	2.178	0.795
EE (pcm)	71.22	92.89
MPE (%)	1.29	1.97

Table 3: Axial power relative error (%) of Langenbuch reactor

Axial level	**HELMHOLTZ**	**PEM**
9	2.63	3.92
8	0.50	0.64
7	0.31	0.50
6	0.67	0.98
5	0.66	0.99
4	0.43	0.64
3	0.25	0.36
2	2.26	3.55

Table 4: Radial power relative error (%) for HELMHOLTZ

	2.2577	2.5730	1.4016	0.9965	1.4016	2.5730	2.2577	
2.2577	0.7609	0.5927	0.1016	0.4642	0.1016	0.5927	0.7609	2.2577
2.5730	0.5927	1.0867	0.8261	0.6286	0.8261	1.0867	0.5927	2.5730
1.4016	0.1016	0.8261	1.2683	1.2096	1.2683	0.8261	0.1016	1.4016
0.9965	0.4642	0.6286	1.2096	1.8977	1.2096	0.6286	0.4642	0.9965
1.4016	0.1016	0.8261	1.2683	1.2096	1.2683	0.8261	0.1016	1.4016
2.5730	0.5927	1.0867	0.8261	0.6286	0.8261	1.0867	0.5927	2.5730
2.2577	0.7609	0.5927	0.1016	0.4642	0.1016	0.5927	0.7609	2.2577
	2.2577	2.5730	1.4016	0.9965	1.4016	2.5730	2.2577	

Table 5: Radial power relative error (%) for PEM

	4.0137	3.6148	2.3314	1.9122	2.3314	3.6148	4.0137	
4.0137	1.0887	0.5303	0.2679	0.6644	0.2679	0.5303	1.0887	4.0137
3.6148	0.5303	1.3972	1.2793	1.1314	1.2793	1.3972	0.5303	3.6148
2.3314	0.2679	1.2793	1.8541	1.8393	1.8541	1.2793	0.2679	2.3314
1.9122	0.6644	1.1314	1.8393	2.5471	1.8393	1.1314	0.6644	1.9122
2.3314	0.2679	1.2793	1.8541	1.8393	1.8541	1.2793	0.2679	2.3314
3.6148	0.5303	1.3972	1.2793	1.1314	1.2793	1.3972	0.5303	3.6148
4.0137	1.0887	0.5303	0.2679	0.6644	0.2679	0.5303	1.0887	4.0137
	4.0137	3.6148	2.3314	1.9122	2.3314	3.6148	4.0137	

4 Conclusions

A new polynomial expansion method based on Helmholtz equation has been developed to discretize the Neutron Diffusion Equation by means of the Finite Volume Method. This method provides accurate results for coarse meshes and consequently low computational time.

However, this method is not appropriate for fine meshes, because the computational time will be increased due to the iterations of $\boldsymbol{k}$. In addition, only the results for the first eigenvalue and eigenvector have been shown, because this method provides bad results for the calculation of several eigenvalues, due to the wrong estimation of the coefficients of the Helmholtz equation for each spatial variable x, y, z.

References

[1] W.M. Stacey, Nuclear Reactor Physics. New York, John Wiley & Sons, 2001.

[2] K.A. Hoffmann and S.T. Chiang, Computational Fluis Dynamics, vol.2. Wichita, Engineering Education System, 2000.

[3] Álvaro Bernal, Rafael Miró, Damián Ginestar and Gumersindo Verdú. Resolution of the Generalized Eigenvalue Problem in the Neutron Diffusion Equation Discretized by the Finite Volume Method, *Abstract and Applied Analysis*, Volume(2014):1–15, 2014.

[4] A. Bernal, J.E. Román, R. Miró, D. Ginestar and G. Verdú. An inter-cells polynomial expansion method for the steady-state 2 energy-group neutron diffusion equation discretized by the Finite Volume Method, *Mathematical Modelling in Engineering and Human Behaviour 2014. 16th Edition of the Mathematical Modelling Conference Series at the Institute for Multidisciplinary Mathematics*, Valencia, September 3th-5th, 2014.

[5] V. Hernández, J.E. Román and V. Vidal. SLEPc: a scalable and flexible toolkit for the solution of eigenvalue problems, *ACM Transactions on Mathematical Software*, Volume(31):351–362, 2005.

On the matrix Hill's equation and its applications to engineering models

P. Bader[♭], S. Blanes[†] [*],
E. Ponsoda[†], and M. Seydaoglu [‡]

(♭) Department of Mathematics and Statistics, La Trobe University, 3086 Bundoora VIC, Australia,

(†) Instituto de Matemática Multidisciplinar, Universitat Politècnica de València. Spain,

(‡) Department of Mathematics, Faculty of Art and Science, Muş Alparslan University, 49100, Muş, Turkey.

1 Introduction

The study of the potential of a charged particle moving in the electric field of a quadrupole, without considering the effects of the induced magnetic field, leads to the equations of motion

$$\begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix}'' = \begin{bmatrix} -\frac{2e}{md}\left(V_0 + V_1\cos(t)\right) & 0 & 0 \\ 0 & \frac{2e}{md}\left(V_0 + V_1\cos(t)\right) & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix};$$

where $x(t)'' \equiv \dfrac{d^2}{dt^2}x(t)$, e is the charge of the particle, m the mass, d is the minimum distance from the electrode to the z-axis (the direction in which the particle is traveling), and $V(t) = V_0 + V_1\cos(t)$ is the varying in time voltage, see [6].

The uncoupled equations for x_i, $i = 1, 2$, are the well-known Mathieu equations

$$x''(t) + (\alpha - \beta\cos(t))\, x(t) = 0\,;$$

[*]e-mail: serblaza@imm.upv.es

where α and β are appropriate constant parameters.

From [7], the motion of the particles can be regulated by V_0 and V_1 such that the stability of the equation is guaranteed only in a certain regions on the $\alpha - \beta$ plane, and this fact allows to filter particles which is the basis of the quadrupole mass spectrometry.

Now, let us consider the more general problem which is a matrix version of the so called Hill's equation

$$x''(t) + f(t)x(t) = 0\,, \tag{1}$$

where $t \in \mathbb{R}$, $x(t) \in \mathbb{C}^r$ and $f(t)$ is a $r \times r$ matrix valued function with period T.

Hill's equation has many applications in practical periodically variable systems like the study of quadrupole mass filter and quadrupole devices, spatially linear electric fields, dynamic buckling of structures, electrons in crystal lattices, waves in periodic media, etc. see [5]. In [4], the theory of Floquet is applied in order to study the equation (1). In most practical cases $f^T = f$, and the fundamental matrix solution is a symplectic matrix. This property plays a fundamental role for the stability of the system. Our goal is to build new efficient symplectic methods based on Magnus expansions. The new methods are closely related to commutator-free methods [1] but they show a superior performance for this problem. We check the efficiency of the methods obtained through their application into some numerical examples.

2 Numerical integration for one period

Let us consider eq. (1) and the equivalent first order system

$$Z'(t) = A(t)Z(t)\,; \quad Z(t) = \begin{bmatrix} x(t) \\ x'(t) \end{bmatrix}\,, \quad A(t) = \begin{bmatrix} 0_{r\times r} & I_{r\times r} \\ -f(t) & 0_{r\times r} \end{bmatrix}\,, \tag{2}$$

with $Z(0) = Z_0 \in \mathbb{C}^{r\times r}$, and where $I_{r\times r}$ denotes the identity matrix. Let $\Phi(t,0)$ be the fundamental matrix solution of (2), then

$$\Phi'(t,0) = A(t)\Phi(t,0)\,, \quad \Phi(0,0) = I_{2r\times 2r}\,. \tag{3}$$

Then, if we denote by $\Phi(T) = \Phi(T,0)$, the solution of (3) after a period, Floquet theory using $f(t) = f(t+T)$ tells us that $\Phi(t+T,t) = \Phi(T)$ and consequently

$$Z(T) = \Phi(T)Z_0\,, \quad Z(2T) = \Phi(T)Z(T) = \Phi^2(T)Z_0\,, \ \ldots\ , Z(nT) = \Phi^n(T)Z_0.$$

This implies that the system is stable if the moduli of every eigenvalue of $\Phi(T)$ are less than one.

If $f^T = f$ then A defined by (2), belongs to the symplectic Lie algebra, i.e. $A^T J + J A = 0$ and Φ is a symplectic matrix, i.e. $\Phi^T J \Phi = J$, where J is the fundamental symplectic matrix (i.e. $J = A(t)$ for $f(t) = I$). The $2r$ eigenvalues of $\Phi(T)$ occur in reciprocal pairs $\lambda_k, \hat{\lambda}_k$, i.e., $\lambda_k \hat{\lambda}_k = 1,\ k = 1, \ldots, r$ and for stable systems all of them lie in the unit circle. For the numerical integration of (3), this property is not preserved in general by standard methods, like Runge-Kutta schemes, and the numerical values of the eigenvalues will, in general, not lie in the unit circle. Symplectic methods on the other hand preserve the reciprocal structure which is expected to be advantageous. Amongst them, we consider Magnus integrators [3] to solve (2)

$$Z(t) = \exp\left(\Omega(t)\right) Z_0\,, \quad \Omega(t) = \sum_{k=1}^{\infty} \Omega_k(t)\,, \tag{4}$$

where the first terms in the expansion are given by

$$\Omega_1(t) = \int_{t_0}^{t} A(t_1)\, dt_1\,, \quad \Omega_2(t) = \frac{1}{2} \int_{t_0}^{t} \int_{t_0}^{t_1} \left[A(t_1), A(t_2)\right]\, dt_2\, dt_1\,, \ \ldots$$

and where $[P, Q] = PQ - QP$ is the matrix commutator of P and Q. However, using commutators in (4) reduces the sparsity of the matrix $A(t)$. In fact, the computational cost to compute the symplectic matrix

$$E_1 = \exp \begin{bmatrix} A & B \\ C & -A^T \end{bmatrix},$$

with $B^T = B, C^T = C$, is considerably higher than the evaluation of the symplectic exponential matrices, E_2, E_3, given by

$$E_2 = \exp \begin{bmatrix} 0 & I \\ C & 0 \end{bmatrix}, \qquad \text{or} \qquad E_3 = \exp \begin{bmatrix} 0 & 0 \\ C & 0 \end{bmatrix} = \begin{bmatrix} I & 0 \\ C & I \end{bmatrix}. \tag{5}$$

If we denote the computational cost ($\mathrm{cost(E_i)}$) to compute E_i in terms of basic operations we find that

$$\mathrm{cost(E_3)} < \mathrm{cost(E_2)} \ll \mathrm{cost(E_1)}.$$

For this reason, we will look for composition methods that involve only the cheap (symplectic) exponentials with structures as in (5), and this requires a deep analysis of the Lie algebra associated with this problem.

We look for sixth-order methods based on the Magnus expansion, and we approximate $\Omega(t)$ defined by (4) in the graded free algebra generated by

$$\{\alpha_1, \ldots, \alpha_s\}\,, \qquad \alpha_{i+1} = \frac{h^{i+1}}{i!}\ \frac{d^i A(t)}{dt^i}\bigg|_{t=t_0+\frac{h}{2}}\,,$$

therefore $\alpha_i = \mathcal{O}(h^i)$. Up to sixth order, it suffices to consider only the algebra generated by α_1, α_2 and α_3 which leads to

$$\Omega^{[6]} = \alpha_1 + \frac{1}{12}\alpha_3 - \frac{1}{12}[12] + \frac{1}{240}[23] + \frac{1}{360}[113] - \frac{1}{240}[212] + \frac{1}{720}[1112] \quad (6)$$

where $[ij\ldots kl]$ represents the nested commutator $[\alpha_i, [\alpha_j, [\ldots, [\alpha_k, \alpha_l]\cdots]]]$. In order to avoid to compute the exponential of commutators in (6), one can use commutator-free Magnus integrators [2] (if $x_{1,k} \neq 0$)

$$\exp\left(\Omega^{[6]}\right) = \prod_{i=1}^{s} \exp\left(\sum_{k=1}^{3} x_{i,k}\,\alpha_k\right) = \prod_{i=1}^{s} \exp\left(\gamma_k h \begin{bmatrix} 0 & I \\ C_k & 0 \end{bmatrix}\right), \quad (7)$$

where $x_{i,k}$ are coefficients to be determined with $x_{1,k} = \gamma_k$, and C_k are linear combinations of $f(t)$ evaluated at a quadrature rule. For sixth-order methods, $s \geq 5$ is required and negative coefficients γ_k are involved. Notice the structural advantage for the generators

$$\alpha_1 = h \begin{bmatrix} 0 & I \\ M & 0 \end{bmatrix}, \quad \alpha_2 = h^2 \begin{bmatrix} 0 & 0 \\ N & 0 \end{bmatrix}, \quad \alpha_3 = h^3 \begin{bmatrix} 0 & 0 \\ P & 0 \end{bmatrix}.$$

The exponentials of α_2 and α_3 are trivial to compute since they are nilpotent matrices of order two: $\exp(\alpha_i) = I + \alpha_i\,, \quad i = 2,\, 3\,;$ while the exponential of α_1 has a lower computational cost than for a full matrix. In addition, we observe that $[\alpha_2, \alpha_3] = [23] = 0$ and the element $[212]$ has the same matrix structure as α_2 or α_3.

3 New efficient numerical methods

With an appropiate decomposition of the matrix exponential (7) in a product of matrix exponentials, we take advantage of the sparse structure of the matrices $\alpha_i\,,\ i = 1, 2, 3\,$. For example

$$\exp(x_1\alpha_1 + x_2\alpha_2 + x_3\alpha_3)\,\exp(x_4\alpha_1 + x_5\alpha_3 + x_6\,[\alpha_2, \alpha_1, \alpha_2]) \times \exp(x_1\alpha_1 - x_2\alpha_2 + x_3\alpha_3)\,, \quad (8)$$

corresponds to the composition

$$\exp\left(\gamma_3 h \begin{bmatrix} 0 & I \\ C_3 & 0 \end{bmatrix}\right) \exp\left(\gamma_2 h \begin{bmatrix} 0 & I \\ C_2 & 0 \end{bmatrix}\right) \exp\left(\gamma_1 h \begin{bmatrix} 0 & I \\ C_1 & 0 \end{bmatrix}\right),$$

where $C_i,\ i = 1, 2, 3$ are simple combinations of $f(t)$ evaluated in a set of quadrature points and $\gamma_1 = \gamma_2 = x_1,\ \gamma_2 = x_4$. There is only one solution with $\gamma_i > 0,\ i = 1, 2, 3$.

Another method proposed is

$$\begin{aligned} &\exp\left(x_1\alpha_2 + x_2\alpha_3 + x_3\left[\alpha_2, \alpha_1, \alpha_2\right]\right) \exp\left(x_4\alpha_1 + x_5\alpha_2 + x_6\alpha_3\right) \\ &\quad \times \exp\left(x_4\alpha_1 - x_5\alpha_2 + x_6\alpha_3\right) \exp\left(-x_1\alpha_2 + x_2\alpha_3 + x_3\left[\alpha_2, \alpha_1, \alpha_2\right]\right), \end{aligned} \quad (9)$$

that corresponds to the composition

$$\exp\left(\begin{bmatrix} 0 & 0 \\ C_4 & 0 \end{bmatrix}\right) \exp\left(\frac{h}{2} \begin{bmatrix} 0 & I \\ C_3 & 0 \end{bmatrix}\right) \exp\left(\frac{h}{2} \begin{bmatrix} 0 & I \\ C_2 & 0 \end{bmatrix}\right) \exp\left(\begin{bmatrix} 0 & 0 \\ C_1 & 0 \end{bmatrix}\right),$$

and only one real solution exists.

In order to test the performance of the numerical methods, we consider the Mathieu's equation

$$x'' + \left(w^2 + \epsilon\cos(t)\right) x = 0\,,$$

where we take $w = 5$, $\epsilon = 1$ and integrate for $t \in [0, 2\pi]$. We take as the *exact solution* a solution obtained numerically to sufficiently high accuracy. In Figure 1, the error in the 2-norm is computed at the final time $t = \pi$, where N.EVAL denotes the number of evaluations of the flow associated to $A(t)$ defined by (2), RK6 is the standard Runge-Kutta method of sixth order, Magnus4_2 is a commutator-free Magnus method of order 4, Magnus6_5 corresponds with the sixth-order commutator-free method (7) with 5 exponentials, and Magnus6_3 and Magnus6_4 are the numerical new methods (8) and (9) respectively.

References

[1] P. Bader, S. Blanes, F. Casas & E. Ponsoda, Efficient numerical integration of Nth-order non-autonomous linear differential equations. *J. Comp. Appl. Math.* **291**, pp. 380–390, 2016.

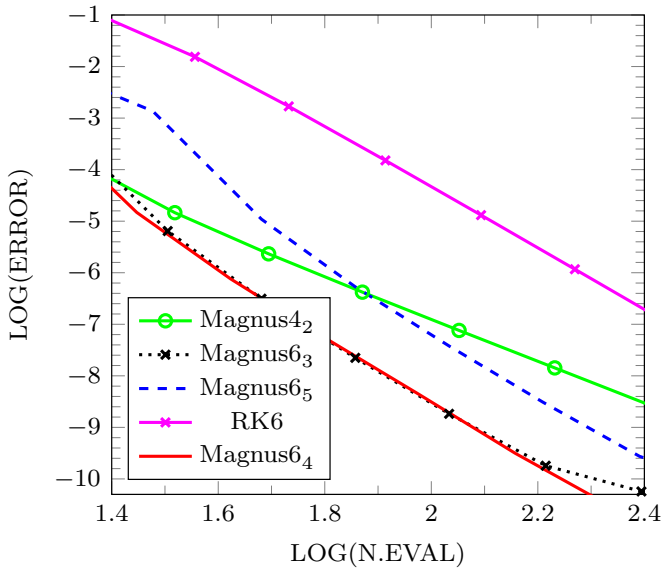


Figure 1: Error at the final time $t = \pi$

[2] S. Blanes & P.C. Moan, Fourth- and sixth-order commutator-free Magnus integrators for linear and non-linear dynamical systems, *Appl. Numer. Math.*, **56**, pp. 1519-1537, 2006.

[3] S. Blanes & E. Ponsoda, Time-averaging and exponential integrators for non-homogeneous linear IVPs and BVPs. *Appl. Num. Math.* **62**, pp. 875–894, 2012.

[4] R. Denk, On the Floquet exponents of Hill's equation systems. *Math. Nachr.* **172**, pp. 87–94, 1995.

[5] J. A. Richards, Analysis of periodically time-varying systems. Springer-Verlag. New York. 1983.

[6] L. Ruby, Applications of the Mathieu equation. *Am. J. Phys.* **64**(1), pp. 39–44, 1996.

[7] G. Teschl, Ordinary differential equations and dynamical systems. Graduate studies in mathematics, Vol. 140. AMS. U.S. 2012.

Adaptability of the acoustics of a room by varying the height of the acoustic ceiling

P. J. Blázquez[♭] *and L. Orcos[‡]

(♭) Universidad Internacional de La Rioja (UNIR),

Av. Gran Vía Rey Juan Carlos I, 41

26002, Logroño, La Rioja, Spain.

(‡) Universidad Nacional de Educación a Distancia (UNED),

Calle Barriocepo, 34

26001 Logroño, La Rioja, Spain.

November 15, 2015

1 Introduction

Many problems from Applied Sciences including engineering can be solved by means of finding the solutions to equations using Mathematical Modelling [1, 2]. For example, dynamic systems are mathematically modeled by difference or differential equations, and their solutions usually represent the states of the systems.

Currently, several mathematical models for the analysis of reverberation times are used. Two of the most commonly applied models are the Sabine [3] and Eyring [4] as they can be used as methods for reverberation times calculation [5, 6, 7, 8, 9] or as a comparative referencence for the application of new experimental methods [10, 11]. The application of these two mathematical models helps us to analyze and characterize the type of room with which we work, allowing us to estimate the behavior of reverberation times inside the room [5]. Both Sabine and Eyring Models, have served and

*e-mail: pedrojesus.blazquez@unir.net

are still worth as a baseline in order to calculate the reverberation time in different types of enclosures, fact that makes them to be taken into account as mathematical models in the study of the behavior of the halls [12].

The rest of the paper is organized as follows: in Section 2 we present the problem studied, in Section 3 two models are presented and in Section 4 some tests and simulations are shown. Finally, the conclusions drawn to this study are presented in the concluding Section 5.

2 Problem description

In this paper we analyze the application of the Sabine and Eyring mathematical models in three prototypes of acoustic rooms. We will observe if the results obtained satisfy the desired conditions for the acoustic characteristics of the different studied rooms. We will analyze the results of the reverberation times, by applying the Sabine and Eyring models, to prototypes of acoustic rooms, which will be raised without any acoustic treatment. The only difference between these prototypes will be the variation in height of certain sections of the roof. What we want to show is that, by analyzing the results obtained, when applying these two mathematical models to the prototypes raised, we can reach that the fact of installing a variable-conditioning ceiling with acoustic materials, can let us have different rooms types in the same one, being able to adapt the acoustic conditions, just varying the height of the ceiling, optimizing the structural and acoustic characteristics of the room.

3 Mathematical Models

Applying both mathematical models, we can calculate reverberation times of our prototypes. These mathematical models help us to analyze and characterize the kind of room where we are, as we can estimate their behavior through the results of reverberation times.

3.1 Sabine model

Sabine model is based on the existence of a diffuse field. The type of energy is specular, which fills the room with constant energy density, this means that the absorbent material is evenly distributed. To apply this model it has

to be assumed that the average absorption coefficient is very low, less than 0.3. Even so, this theory is applied to any asymmetrical distribution of the absorbent material in a room and for any value of α [4].

The mathematical model attends to the following expression:

$$Rt = \frac{0.161 \times V}{\bar{\alpha} \times S}$$

where

$$\bar{\alpha} = \frac{\sum_n \alpha_n \times S_n}{S}$$

so,

$$Rt = \frac{0.161 \times V}{\sum_n \alpha_n \times S_n}$$

being

- V = Room volume in cubic meters (m^3).
- S = Total room surface (m^2).
- $\bar{\alpha}$ = Average absorption coefficient.
- S_n = Material absorption surface (m^2).
- α_n = Material absorption coefficient.

3.2 Eyring model

The Eyring model is based on the assumption of a diffuse field, that is to say, it estimates that the absorption produced in the room is uniform. This model is considered more accurate for average absorption coefficients greater than 0.3. When the value of the average absorption coefficient is very low, less than 0.3, the Eyring model tends to be the Sabine one [4].

The mathematical model attends to the following expression:

$$Rt = \frac{0.161 \times V}{S \times (\ln(1 - \bar{\alpha}))}$$

where

$$\bar{\alpha} = \frac{\sum_n \alpha_n \times S_n}{S}$$

so,

$$Rt = \frac{0.161 \times V}{S \times (\ln(1 - \frac{\sum_n \alpha_n \times S_n}{S}))}$$

being

- V = Room volume in cubic meters (m^3).
- S = Total room surface (m^2).
- $\bar{\alpha}$ = Average absorption coefficient.
- S_n = Material absorption surface (m^2).
- α_n = Material absorption coefficient.

4 Tests and simulations

In this section we will present the results of the application of the Sabine and Eyring models to the prototypes. These results will help us to analyze the installation of a variable ceiling within a room.

In Figure 1 we show the results obtained in the simulation of prototypes without any acoustic treatment (7.0m Prototype, 7.05m Prototype 7.1m Prototype) letting us reflect on what it might mean the fact of changing the ceiling of our prototype. To carry out this proposal, we will apply both mathematical models, in order to analyze the reverberation times inside the prototypes. Then we will represent the results of reverberation times graphically and we will show the differences obtained in the two variations (from 0 to 0.5 meters and from 0.5 meters to 1 meter) for the applied models.

5 Conclusions

We realize that the application of Sabine and Eyring mathematical models to this work acoustic Prototypes helps us to analyze their behavior as we vary the height of the sections of deployment of our roof.

Analyzing the results of the application of the mathematical models (see Figure 1), we can see that the reduction of the reverberation time is considerable. This fact indicates that the variation of the roof is a positive factor for

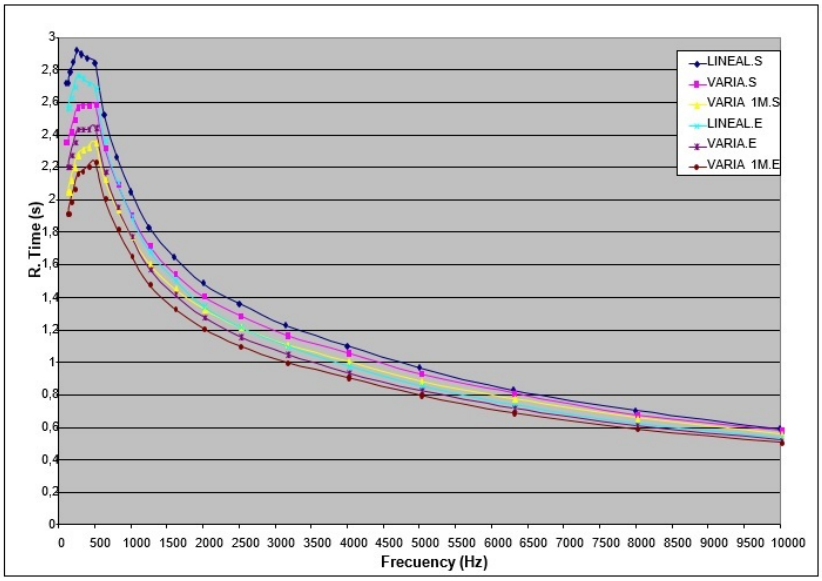


Figure 1: Graph of reverberation times for varying ceiling without treatment. Sabine and Eyring models.

the conditioning and the optimization of the rooms as we obtained favorable results when Prototypes are raised without acoustic treatment.

One factor to consider is that the application of the models to our prototypes will help us identify the change in ceiling height needed for a given type of event, which will allow us to decide what conditions are the optimal for our room.

References

[1] I. K. Argyros, D. González, Local Convergence for an Improved Jarratt-type Method in Banach Space. *International Journal of Artificial Intelligence and Interactive Multimedia*, 3 (Special Issue on Teaching Mathematics Using New and Classic Tools). 2015.

[2] I. K. Argyros, Á. A. Magreñán, On the convergence of an optimal fourth-order family of methods and its dynamics, *Applied Mathematics and Computation* 252(1), 336–346 (2015).

[3] W.C. Sabine, Dover Pub., 1900.

[4] C. F. Eyring, Reverberation time in dead rooms. *Journal of the Acoustical Society of America*, 1930.

[5] N. Nishihara, T. Hidaka T, L. L. Beranek, Mechanism of sound absorption by seated audience in halls. *Journal of the Acoustical Society of America*, 2001.

[6] H. Arau-Puchades, Improved Reverberation Formula. *Revista de acústica*, 1998.

[7] G. Moreira, A. Ramírez, Simulación Computacional de Recintos Arquitectónicos y Modelos Matemáticos V/S Condiciones Reales, *Congreso Internacional de Acústica y Audio Profesional*, 2011.

[8] J. Segura, R. Sanchis, J. Amarshi, E. A. Navarro, Condiciones acústicas de la cámara anecóica electromagnética de la Universitat de València. *Sociedad Española de Acústica*, 2001.

[9] L. Leo, Analysis of Sabine and Eyring equations and their application to concert hall audience and chair absorption. *Journal Acoustics*, 2006.

[10] N. Bastián, R. Flores, Comparación entre Modelación del Software EASE y Mediciones in situ. *Sonac*, 2012.

[11] C.E Boschi, A.F González, Comparación del valor teórico y la medición del TR60 mediante el método de la detonación. *EnIDI. Mendoza. Argentina*, 2008.

[12] H. Arau-Puchades, Revisión del tiempo de reverberación: El método de cálculo, la precisión y la ley. *Congreso nacional de acústica*, 2010.

Water pump scheduling Optimization using Agent Swarm Optimization

B. Brentan[‡] [*], I. Montalvo[†], E. Luvizotto Jr.[‡] J. Izquierdo [⋆] and Rafael Pérez-García[⋆]

(♭)Computational Hydraulic Laboratory, University of Campinas,

Saturnino de Brito, 951, Campinas, Sao Paulo, Brazil,

(†) IngeniousWare GmbH,

Bahnhofstrae 4a, 76137, Karlsruhe, Germany,

(‡) Fluing - IMM, Universitat Politècnica de València,

Camino de Vera, s/n, 5C bajo, Valencia, Spain.

November 30, 2015

1 Introduction

Utility companies must be able to supply water with the quality and quantity required by the consumers at the lowest energy consumption [**?**]. To this purpose, water supply management must be based on crucial decisions about maneuvers of pumps and valves, thus reducing energy consumption and water losses. Optimized rules for water pumping, such as start and stop operations or speed variations bring better control of the system costs [**?**, **?**]. Studies about pump operation, including both scheduling and rotation speed change decisions, have gained ground with the improvement of computational resources and the development of hydraulic models able to simulate the hydraulic behavior of water supply networks. And most recently, the application of optimization methods linked to such models has achieved smarter and safer operational and strategic decisions [**?**].

[*]e-mail: brentan@fec.unicamp.br

However, to focus on engineering problem optimization operational constraints to ensure the applicability of solutions in practice must be considered. One way to manage those constraints is by using multi-objective evolutionary optimization considering contraints as additional objectives to be fulfilled. This idea has been applied to water supply systems for optimal design including reliability issues, sensor placement, calibration models and operational rules. Considering a multi-objective approach and several constraints linked to energy saving, we present an optimization based on Agent Swarm Optimization (ASO). Our solution takes into account not only pressure limits as constraints, but also limits on tank levels, water age, and the number of pump maneuvers. As a result, applicable pump scheduling and efficient operation of variable speed pumps are obtained.

2 Optimization Problem

The daily fluctuation of water demand and energy prices, the possibility of managing tank levels, the operational limits on pressure, and the need to control water quality improve the ability of the energy saving problem to make decisions about pump status and speed. Keeping in mind the new paradigms for the cities, the main objective of finding optimal pump rules is the reduction of energy consumption, which can be written according to

$$Min(C_e) = \sum_{t=1}^{T} \sum_{i=1}^{n} \frac{Q_{t,i} \cdot H_{t,i} \cdot \gamma}{\eta_{t,i}}, \tag{1}$$

where C_e is the total cost, T is the total analysis time, usually 24 hours, $Q_{t,i}$ is the flow at time step t for pump i, $H_{t,i}$ is the head and $\eta_{t,i}$ is the efficiency of pump i at time step t.

To find the lowest cost means to find the best speed for each pump in the system for each time step. So the decision vector $\boldsymbol{v}$ has components representing the relative speed of each pump at each time step. The relative speed showing the percentage of full speed is used. However, the use of low speed can bring operational problems and should be limited.

Pump operation is not only constrained by energy saving, but also by normative and quality maintenance of distributed water parameters. While single objective algorithms treat constraints with penalty functions leaving discretionary the choice of multiplier parameters, multi-objective algorithms

can treat constraints as objectives. In this way, objective functions are added to the problem, corresponding to problem constraints.

So, the evaluation of the operational constraints is free of the empiricism inherent to the frequently arbitrary choice of penalty parameters. However the growing number of constraints requires of a robust algorithm to solve with quality and agility the optimization problem. The following equations present the constraints we take into account in this work.

$$Min(C_{p,min}) = \sum_{node=1}^{N_{total}} (p_{min} - p_{node,t}), \tag{2}$$

$$Min(C_{p,max}) = \sum_{node=1}^{N_{total}} (p_{max} - p_{node,t}), \tag{3}$$

$$Min(C_m) = \sum_{i=1}^{n} m_{T,i} - m^{\star}, \tag{4}$$

$$Min(C_l) = \sum_{k=1}^{N_{tanks}} (l_{f,k} - l_{i,k}), \tag{5}$$

$$Min(C_w) = \sum_{node=1}^{N_{total}} (w_{node} - w_s), \tag{6}$$

where p_{min} is the minimal pressure required by the water system, p_{max} is the maximal pressure, $p_{node,t}$ is the pressure at time step t at junction $node$, $m_{T,i}$ is the number of start and stop operations at pump i, $m_{\star}$ is the maximal pump operation, N_{tanks} is the total number of tanks, $l_{f,k}$ and $l_{i,k}$ are the final and initial levels at tank k, w_{node} is the water age at a junction, and w_s is the standard water age required by the system to maintain quality.

3 Agent Swarm Optimzation - ASO

Agent Swarm Optimization (ASO) [**?**] is a generalization of Particle Swarm Optimization (PSO). The interaction among the agents enables the generation of swarms, structures at an abstract viewpoint, which exhibit emergent behaviors and interact with other agents. The algorithm increases the quality of solutions because over the time the agents work in a distributed way

and each swarm follows some rules that enable to increase or decrease at real time the agents number. The presence of a commom framework where various algorithms work is one advantage of ASO, which allows to gather some populations based algorithms, such as PSO, GA, ACO, and the introduction of new agents to solve the problem to create a robust search method.

Population based methods generally use comparisons to determine the movement of each individual in the search space. At each iteration an important point, called singular point, is determined by the best value at each objective function. An Euclidian distance between an agent and the singular point is determined. The agent with less distance is chosen as the best point. As many objectives are not comparable because have different scales, etc., distance comparison is done by relative values, inside the range determined between maximum and minimum values. Another important point of ASO is the ability to divide the search space. Sometimes the interesting point is not so close to the singular point in the Pareto front. In these cases, this search space separation, with the addition of new agents to exploit this region, can generate a larger Pareto front and offer most interesting points.

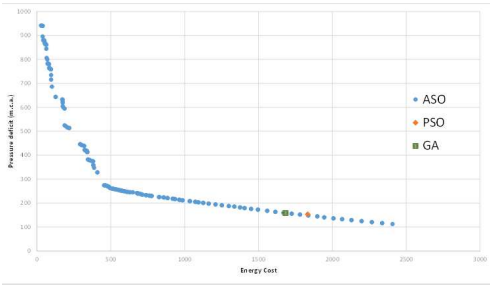
4 Case Study

The application of ASO to find optimal pump scheduling was done for a small network, which allows a result comparison with other techniques presented in [**?**]. The choice of this network is based on the possibility of comparing single and multi-objective optimization, and two different population based algorithms with ASO. The network has two pumps, two tanks, 11 pipelines and 6 demand nodes.

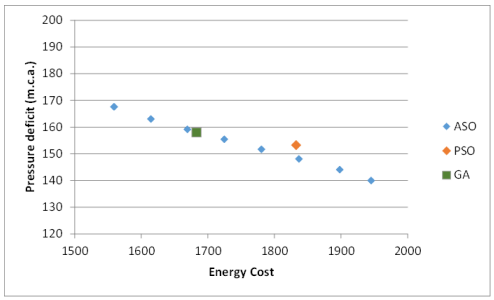
The management horizon used in this work is 24 hours, taking into account the water demand and energy cost profiles. For each hour of the day a relative speed, always greater than 0.5, has to be determined. This condition was adopted because relative speeds less than 0.5 can bring operation problems such as, vibration, cavitation and wear of machines. As minimum pressure 10wcm was adopted, according to the Brazilian technical norm. However in this network, even with full speed work, some points have pressure below this value. So, the Pareto front does not present a null value for deficit pressure. To select the maximum pressure, in this case 50wcm, also the Brazilian technical norm was used.

With these conditions, ASO worked with decision variables inside the

interval [0,1] and generated the following Pareto Front. Figure 1a shows the full Pareto front while figure 1b presents a detail around the comparison region, together with the solutions for the same problem obtained with single objective PSO and GA.



(a) Full Pareto front



(b) Hybrid model results

Figure 1: (a) Full Pareto front (b) Detail of Pareto front at comparison region

An interesting point of this result is the filling and emptying of the reservoir. During the hours that energy has lower price, the tanks fill, leaving harder work for pumps. When the energy price is higher, the tanks empty and the work of pumps is reduced, thus saving energy while still supplying the demand. The evolution of pressure at a specific node shows a critical period where the pressure does not comply with the minimum value.

Table 1: Comparison among three different solution methods for energy cost and pressure deficit

Method	Energy Cost	Pressure deficit
PSO	1832.72	153.21
GA	1683.25	157.76
ASO	1780.40	151.67

5 Conclusions

Scheduling pump optimization is a frequent operation problem for water companies. The difficult to manage pumps taking into account demand and energy price fluctuation need new techniques to find optimal operation rules.

This work presents an application of ASO, a multi objective optimization algorithm, thus eliminating the use of penalty functions. The Pareto front compared with previous GA and PSO results define an operational zone where the energy price decreases with a small increase of pressure deficit.

ASO is a generalization of PSO, treating swarms as agents and using intelligent sel-adaptive parameters to find optimal values. The case study presented here shows the efficiency of the method to build the Pareto front, reducing 25% of energy cost, while supplying demand and pressure nearby the minimum values required.

MATHEMATICAL MODELLING OF HEALTH CARE EXPENDITURE ADJUSTED BY MORBIDITY

Vicent Caballer-Tarazona[a], Natividad Guadalajara-Olmeda[a], David Vivas-Consuelo[a] and Antonio Clemente-Collado[b]

a) Research Centre for Health Economics and Management, Universitat Politècnica de València, Cno de Vera s/n, Valencia, Spain
b) Denia Hospital. Marina Salud. Denia. Spain

1. Introduction

Econometric modelling of health care expenditure is a highly useful technique, among other things for assigning capitation budgets, predicting health care use in the population and improving planning and management of medical services. Traditionally, socio-demographic variables (sex, income, social status) have been the most frequently used to explain the factors that have an impact on health care expenditure. However, in recent years various tools have appeared with the purpose of classifying the population into different homogeneous groups of expenditure according to clinical risk (1,2).

To carry out this study, the system of patient classification we have used is CRG (3), the system employed in the Valencia Community (*Comunidad Valenciana*) by the Directorate of Pharmacy and Health Care Products (*Dirección de Farmacia y Productos Sanitarios*). CRGs group population into 9 mutually exclusive health statuses according to diagnostics from contacts with the health care system at all levels (pharmacy consumption, hospital admissions, medical consultations, etc.). These health states are: 1) Healthy; 2) History of significant acute disease; 3) Single minor chronic disease; 4) Minor chronic diseases; 5) Single dominant chronic disease; 6) Chronic disease in 2 or more organ systems; 7) Dominant chronic disease in 3 or more organ systems; 8) Dominant neoplasms, metastases and complications; and 9) Extreme healthcare needs (Catastrophic conditions).

Within each health status, each patient may be further classified into 6 states of severity, according to the extent of the clinical condition.

However, to model health care expenditure we must bear in mind that there is positive asymmetry in the distribution, with many registers equal to 0 and with a small proportion of the public representing high cost, which causes heteroscedasticity when applying the traditional OLS models(4). To compensate for these anomalies, Duan (5) proposes applying a corrective coefficient, called a Smearing Estimator, to the lineal regression. Nevertheless, other authors propose using Generalized Linear Models (GLM) as a more appropriate tool to explain health expenditure (6). The GLM allow explanation of the result of an endogenous variable when there is heteroscedasticity and the data distribution is not normal.

The aim of this study is to model the total health care costs of patients from a single health care district by applying different mathematical approaches.

2. Methodology

Data and Setting:

Data was collected for a total of 156,811 inhabitants from the Denia Health Department for January to December 2013. The information concerning outpatient pharmaceutical expenditure, primary health care contacts and CRG classification was extracted using the corporate systems of the Valencia Community Regional Health Ministry (*Conselleria de Sanidad de la Comunidad Valenciana*): SIP, SIA, CRC and GAIA.

Variables:

A database indicating the cost per patient was designed, which included pharmaceutical and medical expenditure in the Denia Health Department. Each patient was placed in their corresponding CRG according to their morbidity.

Modelling:

Both the multivariate linear regression logarithmic (MLR) and the generalized linear model (GLM) were applied with health expenditure being the dependent variable and CRG health status and demographic variables (age and sex) as the independent variables.

Various modellings were carried out which included demographic and morbidity variables and those with the greatest level of explanation according to the coefficient of determination were chosen.

In both methodologies the logarithmic transformation of the dependent variable was carried out to increase the goodness of fit and avoid problems of heteroscedasticity.

The log regression model takes the form:

$$LnY = Ln\beta_0 + Ln\beta_1 * x_1 + Ln\beta_i * x_i + \varepsilon_i \quad (1)$$

Where:

- LnY= Naperian logarithm of Health Expenditure
- LnB_i =Parameter measuring the incidence of Group i on dependent variable in logarithmic terms.
- X_i= Each of the clinical risk groups and severity level.
- ε_i = Smearing Estimator

Values and a severity level are obtained for each group through the regression coefficients. Next, the logarithm of each group of patients is undone applying the exponential, after which it is multiplied by the Smearing Estimator (mean of the exponential of the regression residuals). Lastly, 1 is subtracted from the resulting value, as prior to carrying out the regression 1 was added to all cost values to avoid the loss of data equal to 0 in the logarithmic transformation.

The GLM Model takes the form

$$\eta i = \beta 0 + \beta 1 x 1 i + ... + \beta p x p i \quad (2)$$

Where:

- LnY= Health Expenditure
- LnB_i =Parameter measuring the incidence of Group i on dependent variable
- X_i= Each of the clinical risk groups and severity levels and two functions
- a link function that describes how the mean, $E(Yi) = \mu i$, depends on the linear predictor

$$g(\mu i) = \eta i \; I$$

- a variance function that describes how the variance, var(Yi) depends on the mean

$$var(Yi) = \varphi V(\mu)$$

where the dispersion parameter φ is a constant.

The value of the coefficient was transformed, undoing the logarithm to calculate the predicted values for the model.

Goodness of fit for both models was compared estimating the correlation coefficient between the real and predicted values for both models.

3. Results:

The total health care expenditure on the patients covered by the Denia health department (156,811 people) in 2013 was 154,114,807.85 Euros. 18.96% of the population had no contact with the health system, and therefore had zero cost. 76.93% had a cost less than the mean, of which 43.55% registered a cost of less than 100 Euros.

On the other hand, on classifying the department population into the 9 CRG health statuses it was observed that 54.63% of the population was classified as healthy, with a consumption representing 13.34% of the total. The health statuses with the greatest burden on the health system are those of group 5, single dominant or moderate chronic disease, and group 6, significant chronic disease in multiple organ systems. Group 6 supposes 33.36% of total department expenditure for a population that encompasses only 10.02% of the total. Patients in group 5 incurred 25.87% of the total expenditure, while being 15.85% of the population. The total consumption of health care resources for these two groups was 58.23%, while their demographic weight of 25.86% of the population is relatively small for their volume of expenditure.

Table 1. Percentage of population and expenditure according to CRG health status and severity level.

		Severity level							
Health status		0	1	2	3	4	5	6	TOTAL
1 Healthy	Health Expenditure	2,560,803							20,560,803
	Population	85,668							85,668
	Average	**240.01**							**240.01**
2 History of significant acute disease	Health Expenditure	6,222,537							6,222,537
	Population	6,142							6,142
	Average	**1,013.11**							**1,013.11**
3 Single minor chronic disease	Health Expenditure		9,370,832	1,246,942					10,617,774
	Population		14,805	805					15,610
	Average		**632.95**	**1,549.00**					**680.19**
4 Minor chronic disease in multiple organ systems	Health Expenditure		3,809,336	1,835,676	1,610.372	229,838			7,485,222
	Population		4.088	1.425	988	106			6,607
	Average		**931.83**	**1,288.19**	**1,629.93**	**2,168.29**			**1,132.92**
5 Single dominant or moderate chronic disease	Health Expenditure		24.982.953	8.948.721	3.949.288	879.551	1.033.506	71.547	39,865,567
	Population		18.364	4.495	1.483	183	313	12	24,850
	Average		**1,360.43**	**1,990.82**	**2,663.04**	**4,806.29**	**3,301.94**	**5,962.24**	**1,604.25**
6 Chronic disease in 2 or more organ systems	Health Expenditure		19.267.750	12.508.098	8.937.757	5.919.641	2.862.799	379.055	49,875,100
	Population		8.244	3.606	2.137	1.202	480	37	15,706
	Average		**2,337.18**	**3,468.69**	**4,182.39**	**4,924.83**	**5,964.16**	**10,244.73**	**3,175.54**
7 Dominant chronic disease in 3 or more organ systems	Health Expenditure		1,137,179	1,065,017	2,640,730	779,800	381,437	112,301	6,116,463
	Population		292	219	423	96	39	10	1,079
	Average		**3,894.45**	**4,863.09**	**6,242.86**	**8,122.91**	**9,780.43**	**11,230.08**	**5,668.64**
8 Dominant and metastatic malignances	Health Expenditure		763,859	2,404,672	2,360,788	929,477	234,478		6,693,274
	Population		107	249	226	81	23		686
	Average		**7,138.87**	**9,657.32**	**10,445.96**	**11,475.03**	**10,194.70**		**9,756.96**
9 Catastrophic conditions	Health Expenditure		245,664	1,731,729	2,245,323	1,240,186	829,240	385,927	6,678,068
	Population		59	221	87	59	28	9	463
	Average		**4,163.80**	**7,835.88**	**25,808.31**	**21,020.09**	**29,615.70**	**42,880.79**	**14,423.47**
TOTAL	Health Expenditure	26,783,340	59,577,574	29,740,855	21,744.,57	9,978,493	5,341,459	948,830	154,114,808
	Population	91,810	45,959	11,020	5,344	1,727	883	68	156,811
	Average	**291.73**	**1,296.32**	**2,698.81**	**4,068.91**	**5,777.93**	**6,049.22**	**13,953.38**	**982.81**

Severity levels, obviously, also affect health expenditure considerably, as appreciated in Table 1.

The results of the OLS models are presented in Table 2. The model with the greatest coefficient of determination was number 5, which combined the CRG health statuses, age and sex. Nevertheless, to develop a more practical management tool, we chose model 3, which only includes the CRG health status and severity level, but has a more simple implementation, although a slightly inferior coefficient of determination than model 5.

Table 2. Linear Regression Coefficients

Variables	**Model 1** Age range	**Model 2** CRG Health Status	**Model 3** CRG Health Status, and Severity level	**Model 4** CRG Health Status,age & sex	**Model 5** CRG Health Status,severity level, age and sex
Constant	4.124	4.124	3.38	3.74	3.76
Age				-0.02	-0.02
Sex (Female)				0.32	0.32
Age 0-1	1.749				
Age 2-14	0.481				
Age 15-24	-0.082				
Age 25-34	-0.04				
Age 45-54	0.486				
Age 55-64	1.166				
Age 65-74	1.664				
Age 75-84	2.421				
Age 85 +	2.628				
Health Status 2		2.824	2.82	2.84	2.84
Health Status 3		2.369	2.34	2.54	2.52
Health Status 4		3.199	2.95	3.50	3.25
Health Status 5		3.288	3.12	3.61	3.44
Health Status 6		4.22	3.88	4.74	4.39
Health Status 7		4.906	4.36	5.56	4.98
Health Status 8		4.991	4.39	5.45	4.81
Health Status 9		5.631	5.01	5.95	5.28
Severity level 2			0.54		0.58
Severity level 3			0.76		0.84
Severity level 4			0.96		1.09
Severity level 5			1.14		1.31
Severity level 6			1.57		1.67
F	1836.893	12966.873	8141.80	11044.67	7544.34
R^2	**0.095**	**0.398**	**0.403**	**0.413**	**0.419**

In the same way the coefficients for the various models were calculated through the GLM Table 3.

Table 3. Generalised Linear Models

Variables	**Model 1** Age range	**Model 2** CRG Health Status	**Model 3** CRG Health Status and Severity level	**Model 4** CRG Health Status,age & sex	**Model 5** CRG Health Status,severity level, age and sex
Constant	7.854	5.485	5.485	5.543	5.547
Age				-0.005	-0.006
Sex (Female)				0.201	0.206
Age 0-1	-0.726				
Age 2-14	-0.926				
Age 15-24	1.236				
Age 25-34	3.347				
Age 35-44	4.251				
Age 45-54	3.300				
Age 55-64	1.923				
Age 65-74	1.018				
Age 75-84	0.322				
Age 85 +	0.000				
Health Status 2		1.437	1.437	1.453	1.454
Health Status 3		1.039	0.997	1.129	1.093
Health Status 4		1.549	1.320	1.678	1.453
Health Status 5		1.896	1.728	2.058	1.891
Health Status 6		2.579	2.249	2.801	2.467
Health Status 7		3.158	2.665	3.433	2.936
Health Status 8		3.701	3.203	3.902	3.399
Health Status 9		4.092	3.446	4.297	3.633
Severity level 2			0.432		0.429
Severity level 3			0.634		0.650
Severity level 4			0.850		0.879
Severity level 5			0.947		1.002
Severity level 6			1.492		1.491
Deviance	1266380.53	451444.74	447127.05	443531.69	443531.69
Deviance climbing	1266380.53	451444.74	202345.12	202122.06	202122.06
Pearson Chi-square	189988.55	1180231.57	1144768.90	1178770.93	1178770.93
Pearson chi-square scaling	189988.55	1180231.57	518059.47	537178.32	537178.32
Maximum likelihood	-1555099.04	-1147631.14	-1098785.81	-1097969.44	-1097969.44
Akaike information criterion (AIC)	3110218.08	2295280.29	2197601.62	2195972.88	2195972.88
Finite sample corrected AIC (AICC)	3110218.08	2295280.29	2197601.62	2195972.88	2195972.88
Bayesian Information Criterion (BIC)	3110317.71	2295369.95	2197751.06	2196142.25	2196142.25
Consistent AIC (CAIC)	3110327.71	2295378.95	2197766.06	2196159.25	2196159.25

After carrying out the linear regression relating the value estimated by the models and the real value, the coefficient of determination was calculated (Table 4). The proposed values for the GLM present the best R2 of the three methodologies.

Table 4. Coefficient of determination for the linear regression between real and estimated value.

	LM	GLM	Adjusted Patients CRG
Squared correlation	0.247	0.255	0.214

4. Conclusion

By means of the CRG, it is possible to establish a Case Mix system appropriate for defining homogeneous consumption patterns, thus improving the efficiency of health care organizations.

Furthermore, this Case Mix system may be useful to predict the allocation of health care resources for the population according to their CRG health status.

5. References

1. Ellis RP, Pope GC, Iezzoni L, Ayanian JZ, Bates DW, Burstin H, et al. Diagnosis-based risk adjustment for Medicare capitation payments. Health Care Financ Rev. 1996;17(3):101–28.

2. Starfield B, Weiner J, Mumford L, Steinwachs D. Ambulatory care groups: a categorization of diagnoses for research and management. Health Serv Res. 1991;26(1):53–74.

3. Hughes JS, MD; Averill RF, MS; Eisenhandler J, PhD; Goldfield NI, MD; Muldoon, John MHA; Neff, John M. MD§; Gay JCM. Clinical Risk Groups (CRGs): A Classification System for Risk-Adjusted Capitation-Based Payment and Health Care Management. Med Care. 2004;42(1):81–90.

4. Jones AM. Models for health care. 2010;(January). Available from: http://www.york.ac.uk/media/economics/documents/herc/wp/10_01.pdf

5. Duan N. Smearing estimate - a nonparametric retransformation method. J Am Stat Assoc. 1983;78(383):605–10.

6. Hill E, Miller E. Health expenditure estimation and functional form: applications of the generalized gamma and extended estimating equations models. Health Econ. 2010 May;19(5):608-27

CALCULATED FORECAST FOR TECHNICAL OBSOLESCENCE IN COMPUTERISED TOMOGRAPHY EQUIPMENT.

Francisco Reyes-Santia[a], Vicent Caballer-Tarazona[b], Fernándo Gómez-López[c], Javier Rivero de Aguilar-Ameneiro[c] and David Vivas-Consuelo[b]

a) Facultad de Ciencias Empresariales y Turismo. Universidad de Vigo. Spain
b) Research Centre for Health Economics and Management, Universitat Politècnica de València, Caminoo de Vera s/n, Valencia, Spain
c) FIS-Instituto de Salud Carlos III. Madrid. Spain

1. Introduction

To estimate the useful life of Computerised Tomography Equipment (CT), with the aim of planning a budget the purchase of equipment and for the renewal of CT equipment within an already installed base so as to, therefore, maintain and ensure quality in providing an imaging diagnostic level in the National Health System.

Technical obsolescence is an increasingly decisive factor in the development and maintenance of all kind of health care equipment, particularly in medical imaging [1]

The design of a prediction model provides an advance warning of the appearance of any technology leap that involves technological obsolescence in Computerised Tomography technology in use.

The starting data is made up by the different Computerised Tomography models commercialised since 1974 and that have provided a technology leap in CT technology.

It is necessary to know if the data distribution fits into a normal curve, because if this is not so the data will be transformed. A main components analysis in this methodology has allowed for a reduction in the number of variables on the survey-file in Computerised Tomography technology and facilitates subsequent work without a significant loss of information. The Log Binomial Regression Model has enabled probability calculations on answers (technology leap) to the different levels of stimuli (changes in variables, temporary development, detection system, imaging resolution and equipment power). Using a Discriminant analysis, the objective has been to estimate, based on time, the chances of a technological leap occurring [2] [3].

2. Methodology

To estimate the technological life of Computerised Tomography equipment (CT) with the aim of planning a purchasing budget to obtain equipment and renew the already installed base of technology equipment (CT) and therefore maintain and ensure quality in providing an imaging diagnostic level in the National Health System.

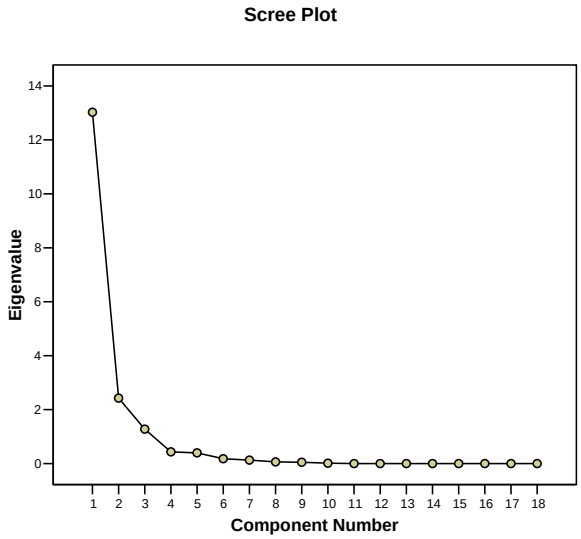
The design of a prediction model provides an advance warning of the appearance of any technology leap that involves technological obsolescence in Computerised Tomography Technology in use.

3. Results:

Main Components Analysis

The strategy followed has been to find a solution that enables us to explain the maximum percentage of variance and an acceptable parsimony of the model and to reduce the 18 descriptive CAT model variables; to set the explained percentage of variance in 92.9%, which is fulfilled in the first six factors in the present study and represented in graphic 1.

Graphic 1



A rotation of factors has been done and the varimax orthogonal rotation method has been used (axis turn orthogonally, in the same angle), which is intended to minimise the number of variables with high saturations in the same factor. The aim of this method is to increase the highest saturations in a factor, while decreasing the lowest for the factor to be easily interpretable.

The importance of each factor is evaluated considering the proportion of variance explained by the factor after rotation.

The 18 evaluated technical parameters in Computerised Tomography Technology have been grouped in three main components: *Detection System* (detector material, reconstruction matrices, reconstruction time, chest and abdomen scans, CT number range (HU), Processor, Maximum tube power (kW), scans ways) which explains 72.4% of the variance; *Imaging Resolution* (cutting thickness range, hard disk, selected kV, fine beam, thick beam) which explains 13.55% of the variance and *Equipment Power* (X-ray generator, number of data by turn or image, digital subtraction) which explains 7.1% of the variance.

Analysis Logit

The characteristics shown by the design of a provision model in a technology leap demands the econometric specification is carried out through carefully chosen models, the endogenous variable to model is a categorical variable with various response alternatives. Within this modelling typology, Logit methodology adjusts to this aim using logistics as an adjusting function. The use of this function ensures that the estimation result can be interpreted as the occurrence probability of each endogenous variable alternative because the estimated values are always between the variation range 0-1.

Within the Logit modelling models of dichotomous answer and multiple answer models are distinguished, according to whether the endogenous variable to model has two or more answer alternatives with different specification in both cases. For a single case of dichotomous variable, in which there are only two possible answer alternatives (there is a technology leap or not), the endogenous variable is usually coded with 1 to represent the occurrence of the studied event or 0 to represent the non-occurrence.

The explanatory variables used in the model are the 3 main components previously calculated. The prediction model obtains a successful percentage of 76.9%. Table 2 presents the results of the Logit model for the current research, showing its estimated coefficient β for the equation variables, typical error of β (S.T.), the Wald statistic, the degrees of freedom (df) and the p-value for the significance of the estimated coefficient (Sig.), the reason for the estimated advantages (Exp(β)) and the confidence interval for exp(β) at 95%.

Table 1.

VARIABLES	β	Std. Error	Wald	df	Sig.	Exp(B)
Intercept	-6,717	3,249	4,275	1	0,039	
DETECTION	-3,975	1,901	4,372	1	0,037	0,019
RESOLUTION	-3,766	1,814	4,312	1	0,038	0,023
POWER	-2,461	0,998	6,076	1	0,014	0,085
YEARS	0,431	0,194	4,904	1	0,027	1,538

The logistic regression allows us to approximate an assessment on the influence of each main component with the passing of time, the implementation of a technology leap,

with its significant influence with positive signs of temporary evolution (0.430), and with a negative sign for the main component the detection system(-3.974), image resolution (-3.766) and equipment power (-2.460), as well as a technology leap expectative depending on the independent variables influence for temporary evolution (1.538) and main components, detection system (0.018) image resolution (0.023) and equipment power (0.085).

Discriminant Analysis

The monitoring periods of this type of analysis are almost always different given that the CT models incorporate into the study at different periods of the observation, in which case, the last in doing so have been observed during a lesser period than those that entered at the beginning. The failure time of each model is measured from the start date of each model's study. Every model has real time, this corresponds to its incorporation date until its last observation and a time "t" which represents time (in years, months, days and so on), the time in which each model was monitored.

The Discriminant methodology has been used so that the intervals are determined by the event occurrence, that is to say, the survival probability is calculated every time an event happens. The conditional probability of survival, this being the probability of being a survivor at the end of an interval which is conditioned by the fact that each model is already a survivor at the start of the interval, is calculated from the exact number of risk cases when the event occurs. This assumes that the instantaneous rate is zero during the interval between two events.

The Discriminant analysis can be used to estimate the probability of changing the CT model over a determined period of time and the factors influencing.

As in the Log Binomial, the explanatory variables used in the model are the 3 components calculated. The prediction model obtains a lower percentage of success than the Log Binomial, around 66.7%.

Table 2 presents the results of the Discriminant analysis, showing the estimated coefficient for the equation variables. The Canonical correlation shows a result of 0.481, meaning a good discrimination for the function. Moreover, the Lambda Wilkx shows a value of 0.769 and a p-value of 0.05.

Table 2.

Variables	Coefficients
DETECTION	1,822
RESOLUTION	2,250
POWER	1,147
YEARS	-2,888

The most important factor in influencing the change of technology seems to be the image resolution followed by the detection system and a negative sing for temporary evolution.

4. Conclusion Remarks

The results of the present project will enable advance knowledge of the expectations of technological change in CT technology, allowing an advance in investment planning for this technology, for acquiring and installing the equipment in hospitals where this type of technology still does not exist and for the renewal of technology bases already installed.

The results of the present research can be validated by the application of other medical technologies, thus amplifying the impact of this research.

All this will have to be reflected by an investment budget plan for technology acquisition.

References

[1] E. Society, “Renewal of radiological equipment.,” *Insights Imaging*, vol. 5, no. 5, pp. 543–6, 2014.

[2] P. Sandborn, “Forecasting Technology and Part Obsolescence,” no. November, pp. 1–13, 2010.

[3] N. Joshi, F. Stahnisch, and T. Noseworthy, *Reassessment of Health Technologies: obsolescence and waste*, no. December. 2009.

Heuristic Framework to Jointly Calibrate Leakage and Pressure in Water Distribution Systems

E. Campbell[‡] *, B. Brentan[†], J. Izquierdo [⋆] and Rafael Pérez-García[⋆]

(‡) Berliner Wasserbetriebe, Anstalt des ffentlichen Rechts

10864 Berlin, Amtsgericht Charlottenburg, HRA 30951 B, Germany

(†) Computational Hydraulic Laboratory, University of Campinas,

Saturnino de Brito, 951, Campinas, São Paulo, Brazil,

(⋆) Fluing - IMM, Universitat Politècnica de València,

Camino de Vera, s/n, 5C bajo, Valencia, Spain.

November 30, 2015

1 Introduction

Usually, calibration of water distribution systems (WDSs) is performed by just changing pipe roughness values and water losses are not usually considered. Furthermore, a point highlighted by [4] is pipe grouping, which can bring several errors in calibration processes if flow distribution is not suitably balanced within the groups.

A common way to add water losses in a calibration problem is to associate nodes with emitter coefficients through flow balances. First, a global coefficient responsible to comply with daily demand balance is found. Then, a multiplier of hourly water demand is changed interactively to find some hourly flow balance. However, considering the physical correlation of water

*e-mail: enrique.campbell@bwb.de

losses and pressure this approach can bring some errors. If losses are fixed and pipeline roughness is changed, pressure change and flow balance may fail. Also, if roughness is fixed first, the emitter coefficients are fully conditioned by this distribution. All in all, it is obvious that a joint adjustment of pipeline roughness and node emitter coefficients is a more appropriate approach.

Taking into account optimization features of the calibration problem, several methods have been proposed for fast and accurate solution, such as linear and non-linear methods, and recently, bio-inspired algorithms. Among the most known bio-inspired methods, Genetic Algorithms (GAs) have been applied to solve various optimization problems in WDSs.

Bearing in mind the difficulty to calibrate roughness and emitter coefficients, we propose a joint calibration of these parameters using GAs for optimization applied to an objective function designed by its features and ability to facility GA convergence. Clustering methods and self-organized maps for grouping pipes and nodes are used as a way to reduce the dimension of the problem.

2 Optimization problem

2.1 Problem definition

Calibration of WDSs may be regarded as a minimization of the error between measured and simulated hydraulic parameters, subject to nodal pressure and pipeline flow contraints. In this work, we consider group roughness values, demand pattern factors, and emitter group coefficients as decision variables.

To obtain an accurate model, a right decision must be made on the objective function, even when the problem is solved by bio-inspired algorithms. The absolute error and the relative absolute error, [5], are applied in several works as objective functions. Also the square error is widely applied, and [2] uses this function to monitor nodal pressure.

However, the sole use of pressure as the monitored variable is not always able to bring the required accuracy for the model, and pipeline flows have to be considered to reduce uncertainty and improve the results. In this way, a sum of absolute pressure, flow and tank level errors is proposed by [3]. Since this approach takes into account more than one hydraulic parameter, the differences among variable scales can lead to convergence problems.

Considering the difficult of obtaining a good objective function able to

describe correctly the problem and to be solved without too many convergence problems we test a set of objective functions to compare not only the model final accuracy, but also the computational effort needed to find the sought results. To evaluate those functions we use hypothesis tests based on the assumption that the mean value of measured and simulated values do not diverge significantly. If this hypothesis is not validated, the respective objective function will not be accepted.

Our statistical hypothesis test shows that an objective function can be accepted, if none of the tested mean values (in our case, six pressure values and one flow value) shows a probability value lower than certain threshold. Also the coefficient correlation (r^2) among measured and simulated values is used to compare the model accuracy. However all the considered functions presented a high value of this parameter, close to 1, corresponding to high ability of finding good results.

Still, an important factor in bio-inspired optimization is the convergence time. The search can be hampered by local minima which may slow the search. These local minima are closely linked to the objective function. Therefore, we propose a quality index (QI) to evaluate not only the accuracy of the model, but also the computational effort. For an objective function o, QI is defined by

$$QI_o = \frac{\prod_{e=1}^{E} r_e^2}{\frac{iter_o}{min(\boldsymbol{Iter})}}, \tag{1}$$

where r_e^2 is the correlation coefficient for control element e in the studied network, having E control elements, $iter_o$ is the total iteration required to stop the optimization method using objective function o, and $\boldsymbol{Iter}$ is a vector with the number of iterations for each objective function.

The best value for parameter QI is found by composing dimensionless absolute errors using maximum and minimum flow and pressure:

$$Minf(Q,H) = \sum_{k=1}^{p} |\frac{Q_k - Q_k^m}{Max(\boldsymbol{Q})}| + \sum_{n=1}^{N} |\frac{p_n - p_n^k}{Max(\boldsymbol{p})}| \tag{2}$$

$$Minf(Q,H) = \sum_{k=1}^{p} |\frac{Q_k - Q_k^m}{Min(Q_k)}| + \sum_{n=1}^{N} |\frac{p_n - p_n^k}{Min(p_k)}| \tag{3}$$

where Q_k is the simulated flow through pipe k, Q_k^m is the modeled flow for the same pipe, $\boldsymbol{Q}$ is a vector containing all measured flows, p_n is the pressure at node n, p_n^m is the pressure modeled at the same node, and $\boldsymbol{p}$ is a vector containing all measured pressures.

Regarding the optimization algorithm, this work uses a version of GA developed by Palisade linked with MS Excel, using the demand pattern factor, the number of pipeline groups, and the number of node groups as the length of the decision vector. An iterative process of hydraulic simulations and optimization algorithm runs is done, which evaluates the error for 60 chromosomes.

3 Node and pipeline clustering

Keeping in mind the importance of the correct choice of the objective function, another fundamental process involved with the calibration problem is the correct grouping of pipelines and nodes, especially for large WDSs. The variation of roughness coefficients and water losses depend on a set of variables. We claim that the use of clustering methods can help divide the elements into homogeneous groups. A cluster is defined as a conglomerate of objects that has similar features.

3.1 Self organized maps - SOMs

SOMs are a type of unsupervised neural networks, with unconnected nodes that are fully connected to the input. Frequently, SOMs are used to represent multidimensional databases through a two-dimensional structure. Thus SOMs can be used as a clustering a technique.

The first step to build a SOM is to define its topology. This means to define the number of neurons (models), which depend on the implicit structure of the database. Usually this structure is not known and the test of several topologies can help define the size of the map. The main idea is to find an uniform distribution of the database multidimensional vectors among the neurons. To this purpose a distance among the elements is used as an indicator of similarity. The smaller the distance among elements associated to the same neuron, the better the topology [1].

Once defined the topology of a SOM, it is important to define some training parameters. Parameters help associate each neuron with an artificial

vector called codebook that is responsible for the comparison among the multidimensional vectors to be mapped by the SOM.

During the training process, the data set is presented to the SOM a number of times. Each time the distance among codebooks is measured and the neuron whose vector has the smallest distance is defined as its Best Matching Unit (BMU). So the neuron vector is changed to make it most similar to the BMU and further apply some change to nearby BMU neurons. This step is responsible to ensure the neighborhood feature of SOM, allowing the grouping of similar vectors. This feature is used for the application of SOM as a clustering tool.

In hydraulic models, the elements can be characterized by physical or topological variables such as roughness, length or diameter of pipes and elevation, base demand, emitter coefficient or coordinates of nodes. As roughness and emitter coefficients are the interest variables, they can not be used in clustering criteria. So, age, material, and pipe centrality are used as pipeline group criteria, and shortest path value, centrality and coordinates for grouping nodes, following the proposal by [1].

Once the SOM has been generated, pipes and nodes are subjected to a clustering process. Each class is represented by its codebook. Hierarchical clustering is applied. After this process, the silhouette index is evaluated to determine the combination that maximizes this index. Figure 1 shows the final partition of nodes and pipelines.

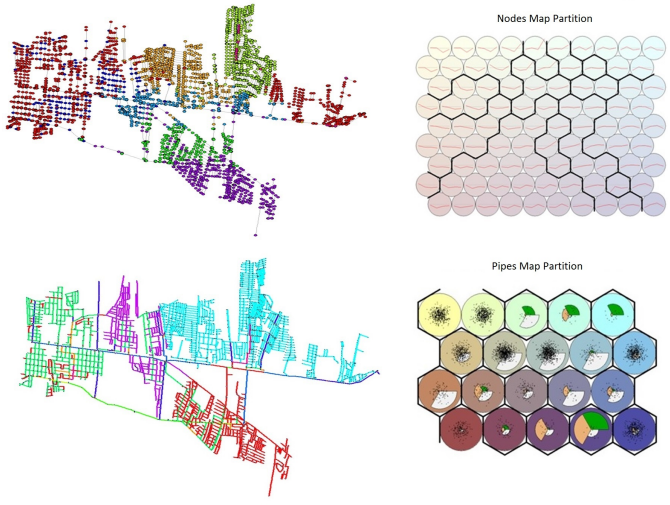


Figure 1: Partition of nodes and pipelines

4 Conclusions

The calibration of a WDS model that considers not only roughness but also leaks through emitter coefficients, and hourly demand factors is not easy and requires suitable optimization to define right values. The errors and difficulties of the process derive not only from adequate grouping processes but also from suitability of the objective function that will describe the problem. We present an approach, to jointly calibrate these three parameters using a bio-inspired genetic algorithm. To define the decision variables, we applied self organizing maps that adequately divide pipelines and nodes into groups. An objective function using absolute error weighting of maximal and minimal values of control parameters (flow and pressure) is used, once the quality index of this function is the best.

This methodology presents computational time saving, finding high correlation parameters between measured and simulated variables, and reducing the number of variables, once an SOM is able to define representative groups.

References

[1] E. Campbell, J. Izquierdo, I. Montalvo, A. Ilaya-Ayza, R. Pérez-García, and M. Tavera. A flexible methodology to sectorize water supply networks based on social network theory concepts and multi-objective optimization. 2015.

[2] T. Koppel and A. Vassiljev. Calibration of a model of an operational water distribution system containing pipes of different age. *Advances in Engineering Software*, 40(8):659–664, 2009.

[3] K. Lansey, W. El-Shorbagy, I. Ahmed, J. Araujo, and C. Haan. Calibration assessment and data collection for water distribution networks. *Journal of Hydraulic Engineering*, 127(4):270–279, 2001.

[4] R. W. Meier and B. D. Barkdoll. Sampling design for network model calibration using genetic algorithms. *Journal of Water Resources Planning and Management*, 126(4):245–250, 2000.

[5] L. E. Ormsbee. Implicit network calibration. *Journal of Water Resources Planning and Management*, 115(2):243–257, 1989.

A novel framework to sectorize water supply networks considering uncertainties

Enrique Campbell[♭] [*], Joaquín Izquierdo[†], Rafael Pérez-García[†], and Idel Motalvo[‡]

(♭) Berliner Wasserbetriebe

Neue Jüdenstraße 1, 10179, Berlin, Deutschland.

(†) Universitat Politècnica de València

Camino de Vera SN, 46022, Valencia, España.

(‡) IngeniousWare GmbH

Bahnhofstraße 4a, 76137, Karlsruhe, Deutschland.

November 30, 2015

1 Introduction

Water Supply Network (WSN) sectorization entails the partial isolation of areas within the networks by closing pipes and setting a flowmeter in a single (or few) feed lines. It has some drawbacks, namely, the economic investment in boundary valves and flowmeters and the pressure reduction due to increase of headloss. Whilst pressure reduction can translate into leakage reduction, it can also lead to supply outage. A good sectorization layout must offset these aspects; therefore, an optimal distribution of boundary valves and flowmeters must be defined. In the methodology here proposed, the network is treated as a social network, and a community detection algorithm is implemented to define the sectors, then, the arrangement entrances/boundaries is defined for each sector using Genetic algorithms (GA) + Monte Carlo Simulation (MCS).

[*]e-mail: enrique.campbell@bwb.de

2 Community detection algorithm

WSNs can be represented as graphs, which allows to implement algorithms derived from the graph theory domain. In the method proposed by Campbell *et al.* (2015) to define the trunk network, pipes are ranked according to their role in the supply of the entire network. For each node, the Shortest Paths, to all reachable nodes are calculated. The result is arranged in a square matrix, from where the Accumulated Shortest Path Value (ASPV) of each node is calculated and then, the values are transferred to pipes. This way, pipes connected to the sources are expected to have the greatest values of ASPV, whilst the extreme pipes are expected to have an ASPV equals 1. This is used as a criterion to segregate the trunk network from the distribution network. The resulting distribution network is treated as a Social Network, and the Multilevel Community Detection Algorithm (Blondel *et al.*, 2008) is implemented over it. The partition of the network with the highest value of Modularity (Newman, 2006) index can generate extremely small communities. Therefore, the communities in this partition are recursively merged using a merging process here proposed (see pseudocode below), to ensure that all sectors comply with a series of pre-established limits. Such limits can include, pipe length, number of connections, or a combination of several characteristics.

Pseudocode:
All pipes connecting communities are selected and put in a Set of Candidate Pipes (SCP); (**1**) Given pipe m, if $IN_m \in C_i$ & $FN_m \in C_j \longrightarrow m \in$ SCP; from the CP, extract the subgroup of pipes scp_i from which their corresponding communities do not exceed a pre-established limit for a given feature; (**2**) Given $scp_i \in$ SCP, where, for every m, L_i & $L_j < L_{max}$; for every m, if $L_i + L_j < L_{max} \longrightarrow i + j = i \cap j$; Steps 1-2 are repeated until no new pipes enter the scp_i.

Here m = pipe; IN_m = initial node of pipe m; FN_m = final node of pipe m; C = community; L = the characteristic used as a criterion.

Finally, the trunk network is re-coupled with the distribution network and the pipes connecting each sector with the first one are included among the SCP. Each pipe inside the mentioned set can be defined either as boundary valve or as sector entrance.

3 Entrance-Closing valves definition

Between the unavoidable leakage volume and the total leakage volume that can be found in a typical WSN, there is an economic management level, which corresponds to the level where the cost of a new repair exceeds the benefits from the savings. One way to estimate such economical leakage level is through the assessment of the SRELL (Short Run Economic Leakage Level) based on a combination of the BABE (Background and Burst Estimates) concept (Thornton *et al.*, 2008), and the FAVAD (Fixed and Variable Area Discharge) theory (May, 1994). A very important result derived from this estimation is the URELL (Lambert and Lalonde, 2005), which corresponds to the volume of unreported leaks that is feasible to let occur within one year. The method to estimate the SRELL includes an equation to calculate the Optimal Frequency of Inspection (OFI) of the WSN, which is based on three components: CI - Cost of intervention ($); CW - Variable cost of water ($ /m^3); RR -Rate of Rise (m^3/day/year).

The calculation of the OFI can be used to predict the benefits of pressure management in a WSN. According to the FAVAD theory, a pressure reduction entails a reduction in the leakage flow (background, detectable and reported). Furthermore by reducing the pressure, both, the RR, and the time required to perform a complete inspection of the WSN, is reduced, which translates into a reduction of the annual budget for inspection, the SRELL and the URELL. Further benefits derived from reducing pressure include: reduction of burst frequency, as described in Thornton & Lambert (2006), and a reduction of the domestic consumption. When the flow in the network is reduced, the energy is expected to vary correspondingly. To calculate the total cost of energy, the energy consumption of each hour (kWh), in which pumping is used, is multiplied by the cost of energy ($ / kWh) (according to an energy price pattern). The resulting hourly costs are added to obtain the daily energy consumption.

3.1 Objective function

In a WSN of hundred kilometres of pipe length, the number of pipes that can be set as boundary pipes or as sector entrances and the number of possible combinations entrances/boundary valves can be very large. The more entrances a given sector has, the lower is the negative impact of sectorization over the nodal pressure, but more difficult is to detect leakage events. On the other hand, the greater is the number of entrances, the more costly the sectorization becomes. In this paper, a GA is implemented to maximize the net benefit of sectorization. The status (closed/open) of the pipes, delimiting the sectors (candidate pipes), is established as decision variables (every candidate pipe's status is set as decision variable). The

equation below shows the objective function to maximize.

$$\begin{aligned} Max f(x) = A + B + C + D + E + F \\ \pm G - A^* - B^* - C^* \end{aligned} \tag{1}$$

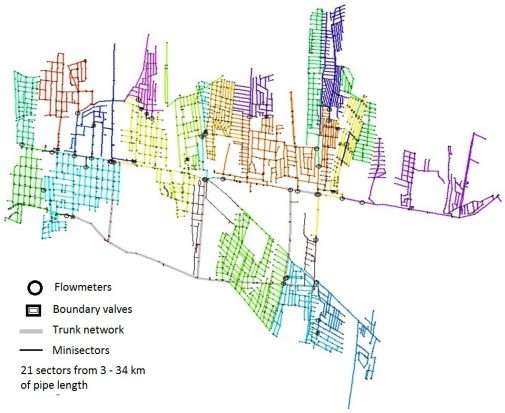
Subject to: $I_r < {I_r}^{max}$; $P_{min} < P_{min}^{req}$; $0 <$ A + B $< (A+B)^{max}$. Where $A =$ Saving by reducing background leakages (volume)(\$/year); $B =$ Saving by reducing reported bursts (volume)(\$/year); $C =$ Saving by reducing the number of pipes to repair (reported bursts)(\$/year); $D =$ Saving by reducing the number of pipes to repair (unreported bursts) (\$/year); $E =$ Saving by reducing the internal domestic consumption (volume)(\$/year); $F =$ Saving by reducing the external domestic consumption (volume)(\$/year); $G =$ Saving/Expenditure by reducing/increasing the energy consumption(\$/year); $A^* =$ Amortized cost of valves and flowmeters (\$/year); $B^* =$ Compensation cost for pressure deficit (\$/year); $C^* =$ Maintenance cost of flowmeters and valves (\$/year); I_r and ${I_r}^{reqs} =$ Variation of the resilience index (current and maximum allowed); P_{min} and ${P_{min}}^{reqs} =$ Maximum and minimum required pressure; $A + B$ and $(A+B)^{max} =$ Budget for the purchase of flowmeters and boundary valves and threshold budget.

As explained above, implementing pressure management based on sectorization has as advantage the capacity to detect new leakage events when they occur. The ability to detect and locate new leakage events will depend on two characteristics of each sector, namely, pipe length and number of connections and in the technical capacity of the water operator. The larger a given sector is and the more entrances it has, the more difficult becomes the detection and location of a new leakage event. The possible values of leakage event detection may be established as single values (fixed values) and then be used in a classical optimization model, which means that at each iteration, the volume of leakage that is saved due to the sectorization implementation is calculated. However, the prediction of the percentage of leakage that can be detected has a high degree of uncertainty; therefore, the prediction of such percentage can be better represented as a range of probabilistic values. Leakage rates are expected to be different for each sector (sectors located in older areas or areas with lower maintenance, are expected to have greater leakage rates than sectors located in newer areas with more continuous maintenance. The distribution of the expected leakage rates among sectors must be performed by the technical staff of the water utility. Once the detection probability curve for each of sector is established, the optimization process is initialized; however, differently from classical optimization, for every iteration, an MCS is carried out. In this simulation, values are probabilistically sampled from the aforementioned probability distribution. For every iteration, the net profit is re-computed. At the end, the average of all the iterations is calculated. The solutions in which the

simulation results meet all the restrictions are marked as valid, and are used as feedback to generate new solutions. The rest of the solutions are discarded.

4 Example of Implementation

To exemplify the methodology, it is implemented over a fraction of the WSN of Managua city, Capital of Nicaragua. First, the supply trunk was identified, then, a set of 21 sectors, with pipe length between 3 and 34 km and with their respective candidate pipes (47 in total) were established. From the 21 sectors, six had one single entrance and the rest a set of 2-4 candidate pipes. The maximum difference of elevation allowed between the nodes of each sector was set at 20 meters. As result of the optimization of entrances/boundary valves, 35 flowmeters and 12 valves were allocated. The annual net profit (savings - expenditures) resulted 72.397 $/year. If only background leakage reduction is considered among the benefits, the project would be infeasible.



5 Conclusion

In this paper we put in place a framework to account the benefits of sectorization, beyond only considering the benefits in terms of background leakage reduction. As it is shown in the implementation example, by considering benefits as: savings by reducing reported leakage; savings by reducing unreported detectable leakage; savings by reducing the number of pipes to repair (reported and unreported leakage) and savings due to reduction of domestic consumption (internal and external), the balance Savings/Investment can be extremely different than when only background

leaks are taken into consideration. In the presented example, the net profit obtained by implementing the proposed methodology is 72.397 $/year, significantly above the result obtained by only considering the reduction of background leakages and energy, in which case, the project is infeasible. This work introduces novel concepts to the sectorization research field, namely, Social Network Community Detection algorithms and MCS. With the first one, along with a merging process proposed in this work, it is possible to define sectors in WSNs depending of a trunk network. With MCS it is possible to take into consideration the inherent uncertainty in the prediction of new leakage events detection within the process of optimization of the entrances/boundaries valves. This is the first sectorization research work that entails an analysis of the effect of sectorization layouts over the detection of leakage events, and therefore we hope that the same could provide new guidelines for researchers and professionals involved in the WSN sectorization research field.

References

[1] Blondel, V., Guillaume, J.L, Lambiotte, R. & Lefebvre, E. Fast unfolding of communities in large networks *Journal of Statistical Mechanics: Theory and Experiment*, 10, pp. 10008, 2008.

[2] Campbell, E., Izquierdo, J., Montalvo, I., Ilaya-Ayza, A., Pérez-García, R. & Tavera, M. A Flexible methodology to sectorize water supply networks based on social network theory concepts and on multi-objective optimization. *Journal of Hydroinformatics*, 2015 (in press).

[3] Newman, M.Finding community structure in networks using the eigenvectors of ma-trices. *Physical Review E*, 74(3), 2006.

[4] May, J. Leakage, pressure and control. In BICS International Conference on Leakage Control Investigation in Underground Assets, 1994, London, UK

[5] Lambert, A. & Lalonde, A.Using practical predictions of economic intervention frequency to calculate short run economic leakage level, with or without pressure management. Proceedings of IWA Specialised Conference 'Leakage 2005', Halifax, 2005, Nova Scotia, Canada.

[6] Thornton, J. & Lambert, A. Recent advances in understanding pressure: burst frequency relationships. Update note (PowerPoint) for water loss task force members, 2008. From thornton@water-audit.com

Analysis of a class of discrete SIR epidemic model

B. Cantó*, C. Coll and E. Sánchez

Instituto de Matemática Multidisciplinar

Universitat Politècnica de València, Camí de Vera s/n 46022 Valencia Spain

Deterministic models are represented by differential equations or difference equations and therefore are associated equilibrium points. In this work, we develop a model with discrete age scale and discrete-time to analyze an epidemic with three states where each state is also subdivided in three compartments depending on the age. The associated model assumes that the chain of infection and recovery is described as happening in discrete-time steps, where each step is the interval between an individual acquiring infection and passing it on to the next infected. Thus, the SIR model is described by a discrete-time nonlinear system and it has unknown parameters which depend on the considered epidemic. This model is denoted by $SIR(\mathbf{p})$ and it is a structured model, where each state is subdivided in: susceptible individuals S_i, infected individuals I_i and recovered individuals R_i at the *i-th* age compartment, for $i = 1, 2, \ldots, m$. The size of the population remains constant to N and an entry of new individuals proportional $\beta(k)N$ state is added to S_1. The parameters involved in the model are as follows: *Survival rate*: p_i, q_i and r_i are the rate of survival individuals of S_i, I_i and R_i, respectively. *Change of compartment without changing the variable*: σ_i, μ_i and ν_i are the rate of individuals becoming S_i / S_{i+1}, I_i / I_{i+1} and R_i / R_{i+1}. *Change of compartment with changing the variable:* ϵ_i and δ_i are the rate of susceptible and infected individuals who pass from compartment i to $i+1$. Finally, *the exposition rate and the recovered rate* at time k are $\alpha_i \dfrac{I_i(k)}{N}$ and γ_i, respectively.

*e-mail: bcanto@mat.upv.es

Then the model is of the form $x(k+1) = f(x(k), \mathbf{p}), \quad k \in \mathbb{Z}$ where $x(k) \in \mathbb{R}^{n\times n}$ is the state vector, $\mathbf{p}$ is the parameter vector and $f : \mathbb{R}^n \to \mathbb{R}^n$ is a continuously differentiable function in $\mathbb{R}^n$.

In the modeling process is important to know the behavior of the system solution in a neighborhood of an equilibrium point. This allows us to deduce whether the disease will disappear, will be endemic or epidemic will grow to generate a pandemic. In this way, we will use a linear approximation $SIR_L(\mathbf{p})$. First, the existence of the equilibrium points is studied. Since the population size remains constant for all k

$$\beta(k)N = N - \sum_{i=1}^{3} p_i S_i(k) - \sum_{i=1}^{3} q_i I_i(k) - \sum_{i=1}^{3} r_i R_i(k),$$

we obtain the disease-free equilibrium point $P_f = (S_1^f, 0, 0, S_2^f, 0, 0, S_3^f, 0, 0)$ given by

$$S_1^f = \frac{N(1-p_2+\sigma_2)(1-p_3)}{\varpi}; \quad S_2^f = \frac{\sigma_1 N(1-p_3)}{\varpi}; \quad S_3^f = \frac{\sigma_1\sigma_2 N}{\varpi},$$

where $\varpi = (1-p_2+\sigma_2+\sigma_1)(1-p_3)+\sigma_2\sigma_1$.

If we linearize around this point we obtain a linear approximation of the initial system $SIR(\mathbf{p})$ given by $x(k+1) = A(\mathbf{p})x(k) + b$, being $x(k) = \begin{pmatrix} \bar{S}(k) \\ \bar{I}(k) \\ \bar{R}(k) \end{pmatrix}$, $A(\mathbf{p})$ the coefficient matrix that contains all the parameters of the system and $b = (N\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0)^T$.

In this work, the identifiability property of our model is analyzed in order to confirm the possible identifiable parameters of the model from a given data set. This analysis identifies the parameters or the relationship between parameters which are accessible from an input-output pair. In the last years, a great variety of works have been developed for the identifiability of several mathematical models [1, 2] and [3].

Thus, a model is identifiable if the relationship between the set of possible parameter values and the set of possible input-output behaviors is one-to-one. That is, given two parameter vectors $\mathbf{p}$, $\bar{\mathbf{p}}$ the equation $x^{\mathbf{p}}(k) = x^{\bar{\mathbf{p}}}(k)$ has only the solution $\mathbf{p} = \bar{\mathbf{p}}$.

From the solution of the system $SIR_L(\mathbf{p})$, $x(k) = A^k(\mathbf{p})x(0) + \sum_{i=0}^{k-1} A^i(\mathbf{p})b$, and using different initial conditions $x(0)$, we can identify all the parameters

except p_1. The non-identifiability of p_1 is therefore due only to the condition that the population is constant N at each time.

After solving the identifiability problem we obtain a model designed to represent theoretical concepts, and to serve as bases for devising quantitative relations and for predicting properties. By solving the identifiability problem we guarantee that the model parameters can be estimated under ideal conditions. In the next step, we will need to design a process to estimate the parameters that can be identified.

The usual estimation methods include the projection algorithm, gradient algorithm and least squares algorithm, [4, 5]. In the estimation process will assume that for a given system the value of the parameters are not known, but we have recorded inputs and outputs over a time interval. So, consider $p_i = p$, $q_i = q$, $r_i = r$, σ_i, μ_i and ν_i, $i = 1, 2, 3$, and ϵ_i and δ_i, $i = 1, 2$, are known and the collection of observed dates given by

$$ob(k) = (S_1(k)\ S_2(k)\ S_3(k)\ I_1(k)\ I_2(k)\ I_3(k)R_1(k)\ R_2(k)\ R_3(k))^T,$$

$k = 1,\ 2, \ldots, K,\ K \geq 1$. We will estimate relevant epidemiological parameters as the transmission rates and the recovery rates, that is, $\mathbf{p} = (f_1\ f_2\ f_3\ \gamma_1\ \gamma_2\ \gamma_3)^T$.

So, we rewrite the initial system as follows

$$x(k+1) = M(k)\mathbf{p} + N(k)$$

being matrices $M(k)$ and $N(k)$ obtained after the reorganization of the linear system $SIR_L(\mathbf{p})$.

Denote by $e(i) = (ob(i) - x(i))^T(ob(i) - x(i))$, $i = 1, \ldots, K$, and define $e_K = \text{col}(e(i))_{i=0}^{K-1}$, $d_K = \text{col}(d(i))_{i=1}^{K} = \text{col}(ob(i) - N(i-1))_{i=1}^{K}$, and $H_K = \text{col}(M(i))_{i=0}^{K-1}$.

From K observed dataset we want estimate the value of $\mathbf{p}$, that is, we can find the parameter vector which minimizes to the quadratic cost function

$$J(K) = \frac{1}{2}\sum_{i=1}^{K} e(i)^T e(i) = \frac{1}{2}e_K^T e_K = \frac{1}{2}(d_K - H_K\mathbf{p})^T(d_K - H_K\mathbf{p}).$$

And we show that if there exists k_0 such that $I_i(k_0) \neq 0,\ i = 1, 2, 3$, then the solution from K observations with $K > k_0$ is $\mathbf{p}_K = (H_k^T H_K)^{-1} H_K^T d_K$.

In addition, if there exists k_0 such that $I_i(k_0) \neq 0,\ i = 1, 2, 3$ then

$$\text{if } \|d_K\| < \frac{1}{\rho(S_K^{-1})\sqrt{\rho(S_K)}} \text{ then } \|\mathbf{p}_K\| < 1,$$

being $\mathbf{p}_K = S_K^{-1} H_K^T d_K$ the parameter which minimizes the problem associated with K observation data.

If we add one observation $ob(K+1)$, using a discrete-time variable system $\mathbf{p}_{K+1} = A_K \mathbf{p}_K + B_K,\ K \geq 1$ where $A_K = S_{K+1}^{-1} S_K$ and $B_K = S_{K+1}^{-1} M(K)^T d(K+1)$, being $S_K = H_K^T H_K$, we get convergence conditions guaranteeing the succession of obtained parameters.

Thus, assume that there exists k_0 such that $I_i(k_0) \neq 0,\ i = 1, 2, 3,\ K > k_0$, and consider the observation data such that

$$\|ob(K+1) - x(K+1)\| < \frac{\epsilon}{\rho(S_K^{-1})\rho(M^T(K)M(K))},$$

for some $\epsilon > 0$. Then

$$\|\mathbf{p}_{K+1} - \mathbf{p}_K\| < \epsilon.$$

Acknowledgements: This research was partially supported by Ministerio de Economía y Competitividad under grant MTM2013-43678-P.

References

[1] Ben-Zvi A., McLellan P.J. and McAuley K.B., Identifiability of Linear Time-Invariant Differential-Algebraic systems. 2. The Differential-Algebraic Approach, *Ind. Eng. Chem. Res.* 43(5): 1251–1259, 2004.

[2] Cantó B., Coll C. and Sánchez E, Parameter identification of a Class of Economical Models, *Discrete Dynamics in Nature and Society* ID 408346, 1-12, 2010.

[3] Cantó B., Coll C. and Sánchez E, Identifiability of a class of discretized linear partial differential algebraic equations, *Math. Problems Eng.* 1-12, 2011.

[4] Ding F., Liu P.X. and Liu G., Multiinnovatiovation Least-Squares Identification for System Modeling, *IEEE TRansaction on Systems,Man, and Cybernetics-Part B: Cybernetics* 18(3): 767–778, 2010.

[5] Malik M.B.and Salman M., State-space least mean square, *Digital Signal Processing* 18: 334-345, 2008.

Modeling Plant Virus Propagation with Delays

Mark Jackson [♭] [*], Benito M. Chen-Charpentier [♭] [†]

(♭) Department of Mathematics, University of Texas at Arlington, Arlington, TX 76019-0408,

November 30, 2015

1 Introduction

Plants are essential to not only man's existence, but to every species on Earth. Sometimes, plants become infected with a disease. There are many different ways that a plant may contract a disease. One of which is bacterial. For example, *Magnaporthe oryzae* is a bacteria that causes rice blast which can cause rice production to decrease up to 90 percent [4]. Also, a plant may become infected with a fungal disease. One such fungus is *Botrytis cinerea*, and it destroys the fruits the plant produces [10]. In this paper, we will be interested in plant viruses and their dynamics. In order to work with the viruses, we must first understand how they replicate. For a virus to replicate, it must invade a healthy cell and use the cell's DNA or RNA to reproduce. The infected cell bursts and several copies of the virus exit.The new virus particles infect other cells. The viruses continue this process until there are no more healthy cells to invade. Different virus processes have been widely studied. See for example [2], [12], [11], and [17].

Plant viruses cause many diseases some of which that affect many plants all over the world. For instance, the *Citrus tristeza* virus once wiped out millions of trees in Brazil [7]. But in order for an infected plant to infect another, a virus from an infected plant must come in contact with a healthy plant. This may happen in several different ways. For example, a field worker might contact the juices of an infected plant and contact a healthy

[*]e-mail:mark.jackson@mavs.uta.edu

[†]e-mail:bmchen@uta.edu

plant. Or maybe the interaction is more organic in that an infected plant might have its juices fall directly onto a plant underneath it. Another, and the most common, way for the transference is by an insect vector. Insect vectors transmit more than 70 percent of all known plant viruses [7]. Many vectors that transport these viruses include aphids, whiteflies, leafhoppers, etc.

There are many ways that plant viruses interact with the vectors, but in this paper we are concerned with circulative, persistent transmission. This transmission works in the following way. The vectors consume sap from an infected host through their stylets. The viruses in the sap enter the salivary glands, circulates within the vector, and then causes infection. This process can take up to a few hours or up to a few days depending on the insect-vector interaction. The vector will hold the infection for the rest of it's life. When the infected vector contacts a healthy plant, some virus particles leave the vector and invade the plant [15, 16, 13]. Once the virus has circulated and propagated throughout the plant, the plant may use defense mechanisms to combat the virus. On example is by antiviral RNA silencing, a process by which slicing or translation repression of viruses occurs [8].

Many physical and biological processes (gestation, maturation, reproduction, infection) take time to complete. In the case of a viral infection, it takes time for a virus to invade a cell, reproduce, and spread in order throughout the host. This process time is a delay time. Processes with delay times can be modeled using delay differential equations (DDE). Delay times can change the dynamics of the model. For example, delays may change the solutions, cause discontinuities in the derivative, introduce oscillations, affect uniqueness, or change the stability. Despite the complications and numerical difficulties, results are more realistic from the biological and physical points of view. The effects of the delay times are highly coupled with the parameters of the model. See for example [6, 9].

Although there are many models that describe in interaction between vectors and humans, there are not as many that describe the relationship between plants and vectors. In this paper, we construct a model assuming that the virus gets transmitted by plant and insect contact, and is modeled using Holling type II [1], since insects can only bite a limited number of plants. In [15] a similar model is presented that also consider that infection can be transmitted from plant to plant. We do not consider such transmission because in [16] it is said it is not common.

2 Mathematical Model

In this model we do not consider a specific virus, plant, or insect vector, but we make the following assumptions.

There are three populations of plants: Susceptible, S, healthy but subject to be infected by the virus, Infective, I, already infected by the virus, and Recovered, R. Each of these variables describe their respective population at time, t. The total number of plants will be denoted by the fixed positive constant K, $K = S + I + R$. It is reasonable to assume K is fixed, because when a plant dies by the virus or natural death in farms, it is usually replaced with a new healthy plant.

For the insect vectors, there will be two populations : Susceptible, X, and Infective, Y. Each of which describe the populations at time, t, as well. The total number of insects will be denoted by the constant, N, N=X+Y. Also, the rate at which the insects enter the system, by birth or immigration, is constant. There is no vertical transmission of the virus, and vectors cannot transmit the virus to another vector. In addition, vectors do not get killed by the virus nor do they defend against it. The vector will keep the virus for it's lifespan and does not recover. The infective insects do not get sick from the virus, they are just carriers.

As far as the interaction between the insects and the plants, an infected insect vector can only infect a susceptible plant. The only way for the vector to become infected is through coming in contact with an infected plant. The interaction between vector and plant is of predator-prey Holling type 2 [1].

The differential equations for this model are

$$\begin{aligned}
\frac{dS}{dt} &= \mu K + dI - \mu S - \frac{\beta Y}{1+\alpha Y} S \\
\frac{dI}{dt} &= \frac{\beta Y}{1+\alpha Y} S - (d + \mu + \gamma) I \\
\frac{dR}{dt} &= \gamma I - \mu R \\
\frac{dX}{dt} &= \Lambda - \frac{\beta_1 I X}{1+\alpha_1 I} - mX \\
\frac{dY}{dt} &= \frac{\beta_1 I X}{1+\alpha_1 I} - mY,
\end{aligned}$$

where μ is the natural death rate of the plants, d is the death rate of infected

plants due to the disease, α is the saturation constant of plants due to vectors, β is the infection rate of plants due to vectors, γ is the recovery rate of plants, Λ is the replenishing rate of vectors (birth and/or immigration) β_1 is infection rate of vectors due to plants, α_1 is the saturation constant of vectors due to plants and m is the natural death rate of vectors.

To get a more realistic model, we consider two delays, τ_1, which is time it takes a plant to become infected after contagion and τ_2, the time it takes a vector to become infected after contagion. The model with the two discrete delays, assuming that the total populations of plants and vectors are constant so R and X can be eliminated, is

$$\begin{aligned}
\frac{dS}{dt} &= \mu(K - S) - \frac{\beta Y(t-\tau_1)}{1+\alpha Y(t-\tau_1)} S(t-\tau_1) + dI \\
\frac{dI}{dt} &= \frac{\beta Y(t-\tau_1)}{1+\alpha Y(t-\tau_1)} S(t-\tau_1) - \omega I \\
\frac{dY}{dt} &= \frac{\beta_1 I(t-\tau_2)}{1+\alpha_1 I(t-\tau_2)} (\frac{\Lambda}{m} - Y(t-\tau_2)) - mY.
\end{aligned}$$

The delay differential equations are solved numerically [3, 14] and analyzed numerically using *dde-biftool* [5].

3 Discussion

We presented two plant virus propagation models, one with no delays and the other with two delays. In the case where the basic reproduction number was less than 1, we notice that the introduction of delays introduces significant changes in the solution for the susceptible, infected, and recovered plants, including a longer time to approach approach the disease free equilibrium point and oscillations. For the insect populations, there is a smaller change in the solutions. When the basic reproduction number was greater than 1, however, the endemic steady state becomes stable and the disease free becomes unstable and this also corroborated by the numerical calculations. The natural death rate of the vector was chosen as the continuation parameter. We changed the value of the natural death rate of the vector, since it is one value that can be modified by the use of pesticides, predators or other vector control means. Furthermore, the results show that it is not

necessary to eliminate the vector completely, only to increase its death rate to control an epidemic.

The study of the stability of both equilibrium points of the delay differential equation system, has to be done numerically. We used the code *dde-biftool* [5], which also allows the variation of some parameters and will approximate the location of bifurcation points.

We ran simulations for different values of the delay times, and similar results were obtained. In particular, we increased the delay time for the virus-insect interaction. The solutions were slightly different but the disease free equilibrium was stable for $R_0 < 1$ and unstable when $R_0 > 1$. Likewise, the endemic equilibrium was stable for $R_0 > 1$ and unstable when $R_0 < 1$.

In addition, there is a Hopf bifurcation of the disease free branch at $m = .2974$ when $\tau_1 = 24, \tau_2 = 1$, but the some populations at the bifurcated branch are negative, so the branch is unrealistic. The same is true for the Hopf bifurcation point on the endemic equilibrium branch. So there are no periodic solutions.

Delayed model is more realistic because it takes into account the time between the release of a factor and its absorption and its effect.

Although a specific plant, insect vector, and virus was not considered, the parameters can be modified to fit a particular situation.

References

[1] L. Allen, An Introduction to Mathematical Biology. Upper Saddle River, New Jersey, Pearson-Prentice Hall, 2007.

[2] R. M. Anderson and R. M. May, Infectious Diseases of Humans. Oxford, Oxford University Press, 1991.

[3] A. Bellen and M. Zennaro, Numerical Methods for Delay Differential equations. New York, Oxford University Press, 2005.

[4] R. K. Ellur et al., Improvement of Basmati rice varieties for resistance to blast and bacterial blight diseases using marker assisted backcross breeding, *Plant Science* Volume 242: 330–341, 2015.

[5] K. Engelborghs et al., DDE-BIFTOOL: a Matlab package for bifurcation analysis of delay differential equations. Leuven, Belgium, Department of Computer Sciences, Free University, 2001.

[6] T. Erneux, Applied Delay Differential Equations. New York, Springer, 2009.

[7] A. Fereres, Insect vectors as drivers of plant virus emergence, *Current Opinion in Virology*, Volume 10: 42–46, 2015.

[8] B. Ghoshal and H. Sanfacon, Symptom recovery in virus-infected plants: Revisiting the role of RNA silencing mechanisms, *Virology*, Volumes 479–480: 167–179, 2015.

[9] Y. Kuang, Delay Equations with Applications to Population Dynamics. San Diego, Academic Press, 1993.

[10] S.V. Leontopoulosa et al. Effect of Different Formulations of Polyphenolic Compounds Obtained from OMWW on the Growth of Several Fungal Plant and Food Borne Pathogens. Studies in vitro and in vivo. *Agriculture and Agricultural Science Procedia*, Volume 4: 327-337, 2015.

[11] M.A. Novak, Virus Dynamics: Mathematical Models of Immunology and Virology. New York, Oxford University Press, 2000.

[12] A.S. Perelson, D.E. Kirschner and R. De Boer. Dynamics of HIV Infection of $CD4^+$ T-cells, *Math. Biosci.* Volume 114: 81–125, 1993.

[13] Y. Robert and D. Bourdin, Virus and Virus-like Diseases of Potatoes and Production of Seed-Potatoes. Dordrecht, Springer Science, 2001.

[14] L. F. Shampine, I. Gladwell, and S. Thompson, Solving ODEs with MATLAB., Cambridge, Cambridge University Press, 2003.

[15] R. Shi, H. Zhao, and S. Tang. Global Dynamic Analysis of a Vector-Borne Plant Disease Model, *Advances in Difference Equations*, Volume 2014:59, 2014.

[16] A. Whitfield et al. Insect vector-mediated tranamission of plant viruses, *Virology*, Volumes 479–480: 278–289, 2015.

[17] D. Wodarz, Killer Cell Dynamics: Mathematical and Computational Approaches to Immunology. New York, Springer, 2007.

A Resource-constrained scheduling problem to optimize the use of resources in a distribution center with genetic algorithms.

David Cipres[♭] [*], Lorena Polo[♭], and Pablo Artaso[♭]

(♭) Aragon Institute of Technology (ITAINNOVA),

María Luna 7-8, 50018 Zaragoza, Spain

November 30, 2015

1 Introduction

The complexity of the planning process in a distribution center is high due to the large number of variables and possible combinations that may affect in the development process. Different trends in manufacturing and distribution have made the order picking process more and more important and complex. These new trends cause high variability in the workload, on the other hand the available resources are limited and they have to be planned in advance. Sometimes the available resources are adjusted to the company needs; other times available resources are below what is needed and could lead to problem in the overall operation.

Although the distribution centers have a key role in the success, or failure, in supply chains [Frazelle, 2002] at the design stage, there is no systematic or scientific approach to the physical design of ware-houses [Goetschalckx et al., 2002, Baker, 2008]. In the absence of a defined and accepted methodology warehouse designers develop their own methods [Oxley, 1994]. In the literature review there are different examples using meta heuristics in the transport and warehousing domain like [Escuín et al., 2012, Sicilia et al., 2015, Royo

[*]e-mail: dcipres@itainnova.es

et al., 2015]. But still there is an enormous gap between the published warehouse research and the practice of warehouse design and operations [Gu et al., 2007]. For a typical warehouse, the cost of order picking is estimated to be as much as 55% of the total warehouse operating expense [Tompkins et al., 2010].

2 Problem description

To solve this problem we propose a method that fits the complexity of the picking process with high variability and integrated with other warehousing operations. In this paper we focus in the resource allocation to tasks in a distribution center. Scheduling and Resource Allocation are traditionally considered as very complex business problems. The challenge is how to schedule operational resources according to the operational flow in a distribution center in specified time intervals. The main variables are the resources assigned to tasks and work intervals. To solve this assignation problem, we have implemented a genetic algorithm to get a fast and feasible solution. In addition, outputs are compared with the results obtained by a linear programming model (LP) solved by an optimization library. The objective from this model is to define a decision-making system for planning task and resource assignment with the required service level in the distribution centers. The main constrains identified are that we have to maintain the service level (Truck schedule time table). There is a limitation in the use of resources (operators, forklifts, docks). And finally we have to maintain the precedence sequence among the tasks.

2.1 Model description

The complexity of this problem could be high, it is a NP-Hard problem. For little instances, it could be solved with standard lineal programming techniques like Simplex Method. In this case the dimension of the problem is big, the number of scheduled task in one day is bigger than 300, the number of operators could be 50. The genetic algorithm is a heuristic, it is routinely used to generate useful solutions to optimization and search problems with no size limitation. The Chromosome representation in our algorithm has five genes: Orders (O), Tasks (T), tEams (E), time Period (P) and Quantity of resources (Q).

Phase	Description
Initialization	A first theoretical solution is computed taking only precedence constraints and temporal restrictions of the tasks into account.
Selection	The selection of the individuals who are going to reproduce.
Cross-Over	The Cross-Over operation mix the genes from two parents in one generation, and creates two new child called offspring.
Mutation	An order which belongs to a solution can be mutated in three distinct ways
Evaluation	Schedule performance is measured in terms of resource availability and length of the order execution.

Table 1: Genetic Algorithm phases

One order is composed by tasks. Each task from one Order has to be assigned to a Team with skill to perform the task in one period of time. The quantity of resources assigned is defined by Q. The algorithm schedules tasks from one initial plan and evolves considering precedence and resource constraints. The selected algorithm is a Steady-State Genetic algorithm because overlaps populations. The size of the population remains constant. Then, once the offspring has been added, the population is truncated by randomly removing two chromosomes from the ten per cent worst evaluated ones. The population size used in the problems depends on the nature of the problem. In this case the population size is 50.

Tasks have precedence relationships between them so so each task can not start until its predecessors have completely finished. In addition, there are temporal restrictions. That is to say that some tasks can be blocked, other can be delayed and other can be brought forward. In addition, there are several resources types with their own temporal restrictions and task-resource dependencies.

Variable	Level	Value
CrossOver	0	50
	1	90
Mutation	0	10
	1	70
Selection	0	Elitism
	1	Roulette Wheel

Table 2: Values from variables

3 Results

In the experiment, the effects of the variables in the fitness of the function are shown. In fig 1 this relationship is represented. The slope of the line in each graph represents the influence of this variable in the objective function. As it is shown, the variable Selection has little influence. The fitness function has almost the same result with the low value (0 Elitism) as with the high value (1 Roulette Wheel). On the other hand the variable mutation has a big impact in the fitness function. The slope from the line is high. The evaluation of the fitness function with the low value of Mutation (10 %) is about 18 and with the high value (70 %) the value obtained is almost 6. With the Cross operator, the influence is intermediate bigger than Selection but lower than mutation. In this case the low level of the Cross operator increase the performance from the algorithm.

4 Conclusions

The research from this papers shows there is a gap between practitioners and researchers in the distribution center operations domain. Some problems could be resolved with models developed from other industries. Scheduling problems are complex, but with the appropriate hypothesis in a distribution center they could be simplified. The genetic algorithm is a useful technique to solve the assignation problem, but with different parameters that affects to the fitness function. In the screening test performed over three parameters it is shown that the mutation operator has a big impact in the evolution of the fitness of the solution. A high value in the mutation operation achieve better results in the objective function. The Selection operator has little

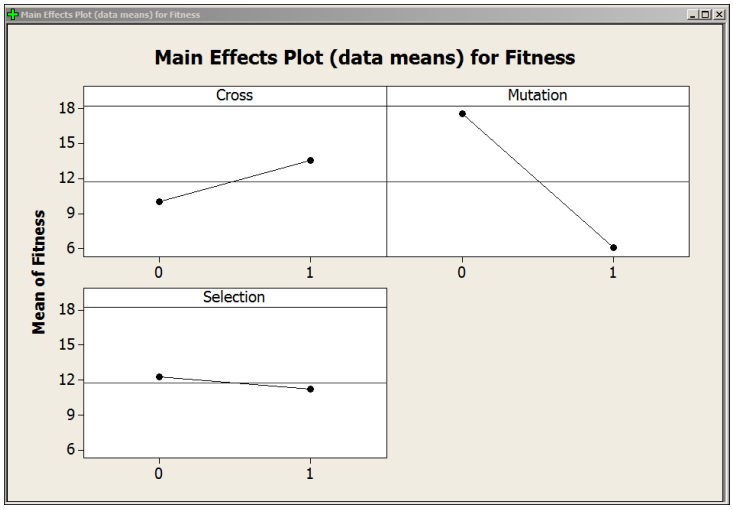


Figure 1: Fitness evolution in each experiment

influence in the performance in the algorithm. The Cross operator has a medium effect, the lower the better performance.

Acknowledgement

The dissemination of this work has been partly financed by the FSE Operative Programme for Aragon (2007-2013).

References

[Baker, 2008] Baker, P. (2008). *The Role, Design and Operation of Distribution Centres in Agile Supply Chains*. Cranfield University.

[Escuín et al., 2012] Escuín, D., Millán, C., and Larrodé, E. (2012). Modelization of time-dependent urban freight problems by using a multiple number of distribution centers. *Networks and Spatial Economics*, 12(3):321 336.

[Frazelle, 2002] Frazelle, E. (2002). *World-class warehousing and material handling*. McGraw-Hill New York.

[Goetschalckx et al., 2002] Goetschalckx, M., Vidal, C. J., and Dogan, K. (2002). Modeling and design of global logistics systems: A review of integrated strategic and tactical models and design algorithms. *European Journal of Operational Research*, 143(1):1–18.

[Gu et al., 2007] Gu, J., Goetschalckx, M., and Mcginnis, L. (2007). Research on warehouse operation: A comprehensive review. *European Journal of Operational Research*, 177(1):1–21.

[Oxley, 1994] Oxley, J. (1994). Avoiding inferior design: The design of a warehouse is not a simple matter and the consequences of inadequate warehouse planning cannot be overstated. *Storage handling distribution*, 38:28.

[Rowley, 2000] Rowley, J. (2000). The principles of warehouse design. *The Institute of Logistics & Transport, Corby*.

[Royo et al., 2015] Royo, B., Fraile, A., Larrodé, E., and Muerza, V. (2015). Route planning for a mixed delivery system in long distance transportation and comparison with pure delivery systems. *Journal of Computational and Applied Mathematics*.

[Sicilia et al., 2015] Sicilia, J. A., Quemada, C., Royo, B., and Escuín, D. (2015). An optimization algorithm for solving the rich vehicle routing problem based on variable neighborhood search and tabu search metaheuristics. *Journal of Computational and Applied Mathematics*.

[Tompkins et al., 2010] Tompkins, J., White, J., Bozer, Y., and Tanchoco, J. (2010). *Facilities Planning*. Wiley.

Dynamical tools for better understanding the stability of high-order methods for solving nonlinear problems *

A. Cordero[♭] †, A.A. Magreñán[‡], and Juan R. Torregrosa[♭]

(♭) Instituto de Matemáticas Multidisciplinar,

Universitàt Politecnica de València, Camino de Vera, s/n, 46022-Valencia, Spain

(‡) Universidad Internacional de La Rioja, Logroño, Spain

November 30, 2015

1 Introduction

In many branches of Science and Technology it is necessary to solve different kinds of nonlinear equations or systems $F(x) = 0$, where $F : X \to Y$, being X and Y Banach spaces. The best known iterative scheme is Newton's method

$$x^{(k+1)} = x^{(k)} - [F'(x^{(k)})]^{-1} F(x^{(k)}), \quad k = 0, 1, \ldots$$

but Traub's scheme increases the order of convergence of Newton's one, without a complex iterative formula

$$\begin{array}{rcl} y^{(k)} & = & x^{(k)} - [F'(x^{(k)})]^{-1} F(x^{(k)}), \\ x^{(k+1)} & = & y^{(k)} - [F'(x^{(k)})]^{-1} F(y^{(k)}), \quad k = 0, 1, \ldots \end{array} \tag{1}$$

where $F'(x)$ denotes the Fréchet derivative of F. This scheme can be successfully used, with third-order convergence, on nonlinear problems.

*This research was partially supported by Ministerio de Economía y Competitividad MTM2014-52016-C2-{1, 2}-P.

†e-mail: acordero@mat.upv.es

In the last years, the use of tools from Complex Dynamics has allowed the researchers in this area of Numerical Analysis to deep in the understanding of the stability of iterative schemes (see, for example, [1–6]). The analysis, in these terms, of the rational function R associated to the iterative procedure applied on quadratic polynomials, gives us valuable information about its role on the convergence's dependence on initial estimations, the size and shape of convergence regions and even on a possible convergence to fixed points that are not solution of the problems to be solved or to attracting cycles. Moreover, if a parametric family is studied under this point of view, the most stable elements of the class can be chosen, by means of an appropriated use of the parameter plane.

In this paper, we begin a dynamical analysis of the rational operator associated to Traub's method on cubic polynomials. Stable and pathological behaviors are obtained depending on the polynomial.

In order to introduce the complex dynamical concepts used in this manuscript, we refer to [7]. Indeed, we need that nonlinear function f is defined on Riemann sphere $\hat{\mathbb{C}}$, as ∞ becomes one more point to be taken into account.

The following results assure us that, if our aim is to analyze the stability of Traub's method on cubic polynomials, it is enough to study its behavior on $p(z) = (z-1)(z-r)(z+1)$, as the dynamics are equivalent (see the Scaling Theorem at [8]), that is, a conjugacy preserves fixed and periodic points as well as their character and basins of attraction.

Theorem 1 ([9]) *Let $q(z)$ be any cubic polynomial with simple roots. Then, it can be parametrized by means of an affine map to $p(z) = (z-1)(z-r)(z+1)$, $r \in \mathbb{C}$. This map induces a conjugacy between $R_q(z)$ and $R_p(z)$.*

The rest of the paper is organized as follows: in the following section, we analyze the dynamics of Traub's method on cubic polynomials. The calculus of the fixed points and their stability are presented. Moreover, some dynamical planes, corresponding to stable or unstable behavior, are shown.

2 Dynamics of Traub's scheme on cubic polynomials

In the following, we apply the fixed point operator associated to Traub's scheme (1) on polynomial $p(z) = (z-1)(z-r)(z+1)$. Then, the following

rational operator appears, depending on both $z \in \hat{\mathbb{C}}$ and $r \in \mathbb{C}$,

$$\begin{aligned} R_p(z,r) &= \frac{4(25r^2-9)z^7+20r(1-2r^2)z^6+2r(19-24r^2)z^4}{(2rz-3z^2+1)^4} \\ &+\frac{12r^2(r^2-5)z^3+12r(2r^2-1)z^2+2r^2(4-r^2)z}{(2rz-3z^2+1)^4} \\ &+\frac{6(r^4+8r^2+1)z^5-111rz^8+r+46z^9}{(2rz-3z^2+1)^4}. \end{aligned} \tag{2}$$

Let us remark that the rational function $R_p(z,r)$ is simpler for specific values of parameter r, forcing common roots in the polynomials of numerator and denominator of $R_p(z,r)$. In fact, they are double roots of $p(z)$, $r = \pm 1$.

Respect the fixed points of operator $R_p(z,r)$, the third-order of convergence of the original iterative scheme induces the superattracting character of the roots, as fixed points. However, there exist some strange fixed points of $R_p(z,r)$, whose character is analyzed in the following results. Firstly, we analyze the role of the infinite as a fixed point: $z = \infty$ corresponds to the divergence of the iterative method and it can be checked that it is a fixed point of $R_p(z,r)$. In fact, its character is repulsive as it happens in Newton's scheme.

Lemma 1 *The strange fixed points of $R_p(z,r)$ are the roots of $s(z) = -1 + (-8r+2r^3)z+(11-22r^2)z^2+(54r-10r^3)z^3+(-37+46r^2)z^4-70rz^5+35z^6$, that will be denoted by $s_i(r)$, $i = 1,2,\ldots,6$. Then, there exist six different strange fixed points except in cases $r = \pm 1$, when the operator is simpler and there exist only three strange fixed points.*

- *$s_1(r)$ and $s_6(r)$ are repulsive for all $r \in \mathbb{C}$*
- *$s_i(r)$, $i = 2,3,4,5$ can be attractive (even superattractive) in different small areas of the complex plane.*

Moreover, $s_2(r)$ is superattracting for $r = -3.06574$, $r = 2.93636$ and $r = 0.0161684$, meanwhile $s_5(r)$ is superattracting for $r = 3.06574$, $r = -2.93636$ and $r = -0.0161684$.

In fact, the regions of the complex plane where some of these strange fixed points become attractive are showed in Figures 1 and 2. In Figure 1, the stability function of $s_2(r)$, $St2(r) = |R_p'(s_2(r),r)|$, is plotted in those regions of $\mathbb{C}$ where it takes values lower than one, that is, where $s_2(r)$ is

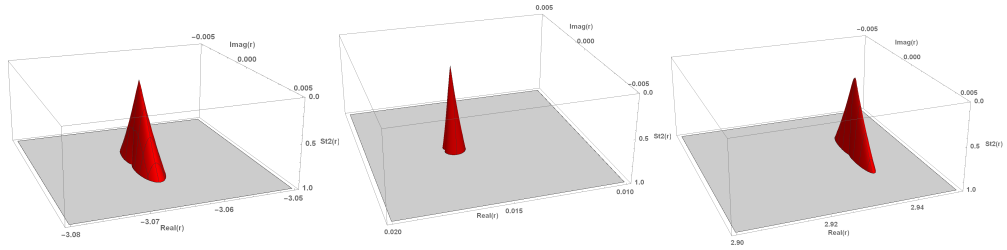


Figure 1: Stability regions of $s_2(r)$

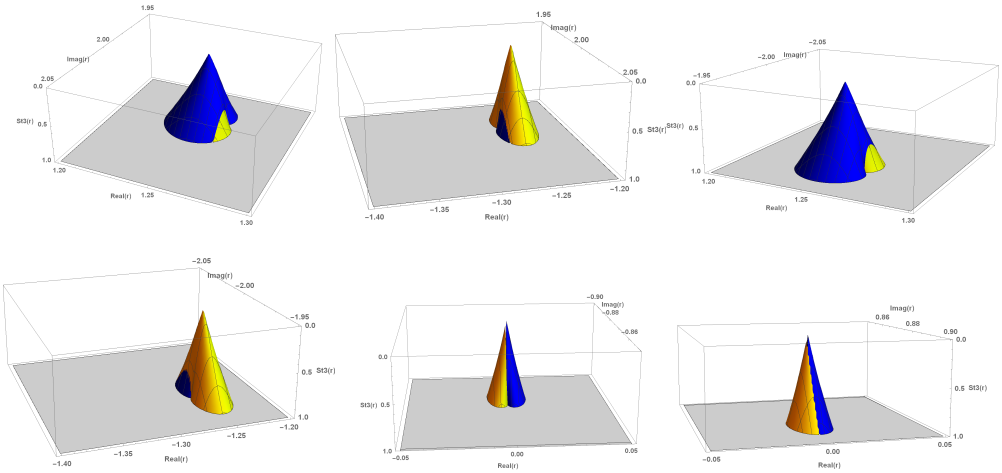
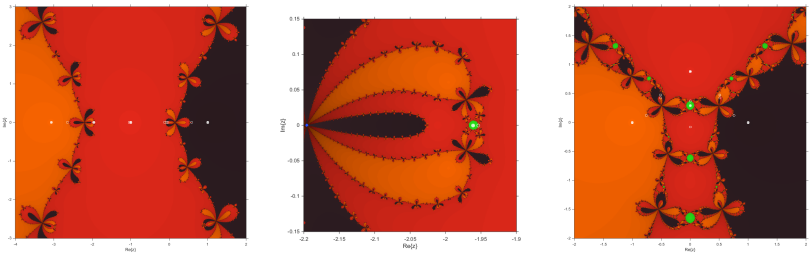


Figure 2: Stability regions of $s_3(r)$ and $s_4(r)$

attractive or superattractive. The behavior of $s_5(r)$ is similar but symmetric respect the imaginary axis. The stability of strange fixed points $s_3(r)$ and $s_4(r)$ is complementary, as it can be seen at Figure 2. In it, we show as the loci where $St3(r) = |R'_p(s_3(r), r)| < 1$ and $St4(r) = |R'_p(s_4(r), r)| < 1$ are complementary subsets of small cardioids. When a value of r is taken in these regions, only one of them ($s_3(r)$ or $s_4(r)$) will be attracting, existing a bifurcation curve (the intersection between dark and clear regions in Figure 2 in each case defined by those values of r where the stability of both points change simultaneously.

Some particular cases showing this kind of behavior can be observed in Figure 3, where one strange fixed point is superattracting. In spite of this, their basins of attraction are very small and frequently it is necessary to get

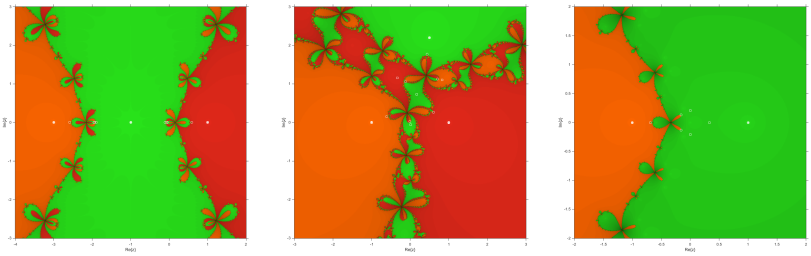
closer in order to see them (Figure 3(b) is a detail of Figure 3(a)). Also we



(a) $r = -3.06574$ (b) Detail (c) $r = 0.88115i$

Figure 3: Dynamical planes with unstable behavior

can find values of r where the behavior of the rational function is stable (see Figure 4). In fact, this stable behavior is the most usual, as the regions of instability are very small, as it has been stated previously.



(a) $r = -3$ (b) $r = 0.5 + 2.2i$ (c) $r = 1$

Figure 4: Dynamical planes with stable behavior

In this work, we have shown that third-order Traub's scheme has a very stable behavior on cubic polynomials. The only anomalies appear as very small basins of attracting strange fixed points. Moreover, some attracting periodic orbits of different periods have been found.

References

[1] Amat S., Busquier S., Bermúdez C., Plaza S. On two families of high order Newton type methods. *Appl. Math. Lett.*, 25: 2209–2217, 2012.

[2] Chicharro F., Cordero A., Torregrosa J.R. Drawing dynamical and parameter planes of iterative families and methods. *The Sci. World J.*, 2013: Article ID 780153, 2013.

[3] Cordero A., García-Maimó J., Torregrosa J.R., Vassileva M.P., Vindel P. Chaos in King's iterative family. *Appl. Math. Lett.*, 26: 842–848, 2013.

[4] Cordero A., Torregrosa J.R., Vindel P. Study of the dynamics of third-order iterative methods on quadratic polynomials. *Intern. J. Comput. Math.*, 89(13-14): 1826–1836, 2012.

[5] Magreñán Á. A. Different anomalies in a Jarratt family of iterative root-finding methods. *App. Math. Comput.*, 233: 29–38, 2014.

[6] Neta B., Chun C., Scott M. Basins of attraction for optimal eighth order methods to find simple roots of nonlinear equation. *App. Math. Comput.*, 227: 567–592, 2014.

[7] Blanchard P. Complex Analytic Dynamics on the Riemann Sphere. *Bull. AMS*, 11(1): 85–141, 1984.

[8] Amat S., Busquier S., Plaza S. Review of some iterative root-finding methods from a dynamical point of view. *Scientia*, 10: 3–35, 2004.

[9] Roberts G., Horgan-Kobelski J. Newton's versus Halley's method: a dynamical systems approach. *Intern. J. Bifur. Chaos*, 14: 3459–3475, 2004.

New Taylor algorithm for the matrix cosine

J. Ibáñez[‡,*], J. Sastre[⋆], P. Alonso[♭], J. Peinado[‡] and E. Defez[†]

[♭] DSIC, [‡] I3M, [⋆] ITEAM, [†] IMM.

Universitat Politècnica de València, Camino de Vera s/n, 46022, Valencia, España.

November 30, 2015

1 Introduction

Many engineering processes are described by second order differential equations, whose solution is given in terms of the trigonometric matrix functions sine and cosine. Examples arise in the spatially semi-discretization of the wave equation, or in mechanical systems without damping, where their solutions can be expressed in terms of integrals involving the matrix sine and cosine [1]. Several state-of-the-art algorithms have been provided recently [2, 3] for computing these matrix functions. In this work we present a sequential algorithm based on Taylor series that use Theorem 1 from [3] for computing matrix trigonometric functions.

Throughout this paper $\mathbb{C}^{n\times n}$ denotes the set of complex matrices of size $n \times n$, I the identity matrix for this set, $\rho(X)$ the spectral radius of matrix X, and $\mathbb{N}$ the set of positive integers. In this paper we use the 1-norm to compute the actual norms. Section 2 presents the Taylor algorithm, Section 3 deals with numerical tests and gives some conclusions.

*e-mail: jjibanez@dsic.upv.es. This work has been supported by the Spanish Ministerio de Economía y Competitividad grant TIN2014-59294-P

2 Algorithm for computing the matrix cosine

The matrix cosine can be defined for all $A \in \mathbb{C}^{n\times n}$ by $\cos(A) = \sum_{i=0}^{\infty} \frac{(-1)^i A^{2i}}{(2i)!}$, and let

$$T_{2m}(A) = \sum_{i=0}^{m} \frac{(-1)^i B^i}{(2i)!} = P_m(B) = \sum_{i=0}^{m} p_i B^i \ , \tag{1}$$

be the Taylor approximation of order $2m$ of $\cos(A)$, where $B = A^2$. Since Taylor series is accurate only near the origin, in algorithms that use this approximation the norm of matrix B is reduced by scaling the matrix, computing a Taylor or Padé approximation and recovering the approximation of $\cos(A)$ by means of the double angle formula $\cos(2X) = 2\cos^2(X) - I$. Algorithm 1 of [3] presents a general algorithm for computing the matrix cosine based on these ideas, where we will use the same considerations as in [3, Sec. 2] for computing the Taylor approximation $P_m(4^{-s}B) = T_{2m}(2^{-s}A)$ by Paterson-Stockmeyer method.

The error analysis for Taylor approximation of the matrix cosine, similar to that on Sec. 2.2 of [2] for Padé approximation, yields analogous results making it very restrictive. Instead of that we use an analysis based on the backward error of the matrix exponential computation by Taylor algorithm from [4], similar to that for Padé Sec. 2.3 of [2], overcoming the difficulty that the backward error for Taylor approximation is not odd as that for Padé, see [5]. This analysis yields Table 1 of new values for parameter Θ_{m_k} to substitute the values from [3, Table 2]. Let m_k be the maximum order allowed. If $||B|| \leqslant \Theta_{m_{k_0}}$ for some $m_{k_0} \leqslant m_k$, then the scaling of matrix B is not necessary and the order m_{k_0} is selected. Otherwise we scale the matrix B by selecting a positive integer s such that $4^{-s}\beta_{m_k} \leq \Theta_{m_k}$, where β_{m_k} can be obtained from [3, Th. 1] as

$$\beta_{m_k} = \max\left\{\left\|B^k\right\|^{\frac{1}{k}} :\ k \geq m_k\right\} = \max\left\{\left\|B^k\right\|^{\frac{1}{k}} :\ m_k \leq k \leq 2m_k - 1\right\}.$$

Analogously to [6, p. 374], we take for β_{m_k} the approximation

$$\bar{\beta}_{m_k} = \max\left\{\|B^{m_k}\|^{\frac{1}{m_k}}, \left\|B^{m_k+1}\right\|^{\frac{1}{m_k+1}}\right\}, \tag{2}$$

and then $s = \left\lceil \frac{1}{2}\log_2\left(\frac{\bar{\beta}_{m_k}}{\Theta_{m_k}}\right)\right\rceil$. Approximation (2) is justified since $\left\|B^k\right\|^{\frac{1}{k}} \to \rho(B)$ as $k \to \infty$, and then for the majority of matrices the values $\left\|B^k\right\|^{\frac{1}{k}}$ tend

to be decreasing and tend to have less variations for higher matrix powers. Algorithm 1 computes the matrix cosine based on the ideas above with the objective of simplicity for a parallel implementation [5].

Table 1: New values of Θ_{m_k} for [3, Table 2] and matrix powers B^2, B^3,$\cdots$, $B^{q_{m_k}}$ used to compute (1) by Paterson-Stockmeyer method.

k	m_k	q_k	Θ_{m_k}	k	m_k	q_k	Θ_{m_k}
1	1	1	6.661338018806219e-16	5	9	3	1.189983654063290
2	2	2	1.154075612730971e-07	6	12	4	4.924177884630485
3	4	2	2.491236564385514e-03	7	16	4	16.06054585896760
4	6	3	8.976968236812591e-02	8	20	5	35.62660483639449

Algorithm 1 Given a matrix $A \in \mathbb{C}^{n\times n}$ and m_k, this algorithm computes $C = \cos(A)$ by (1) with m lower than or equal to m_k.

1: $B = A^2$
2: **if** $||B|| \leqslant \Theta_{m_k}$ **then** ▷ (see Table 1 for the values Θ_{m_k})
3: Compute the first positive integer m_{k_0} such that $\|B\| \leq \Theta_{m_{k_0}}$
4: Compute powers $B^2, \cdots, B^{q_{k_0}}$ (see Table 1 for values q_k)
5: Compute $C = P_{m_{k_0}}(B)$ from (1) by Paterson-Stockmeyer method using the powers from Step 4
6: **else**
7: Compute the powers $B^2, \cdots, B^{q_k}$
8: Compute $\bar{\beta}_{m_k}$ from (2)
9: $s = \left\lceil \frac{1}{2}\log_2\left(\frac{\bar{\beta}_{m_k}}{\Theta_{m_k}}\right)\right\rceil$
10: Compute $C = P_{m_k}(4^{-s}B)$ from (1) by Paterson-Stockmeyer method, scaling and using the powers from Step 7
11: **for** $i = 1 : s$ **do**
12: $C = 2C^2 - I$
13: **end for**
14: **end if**

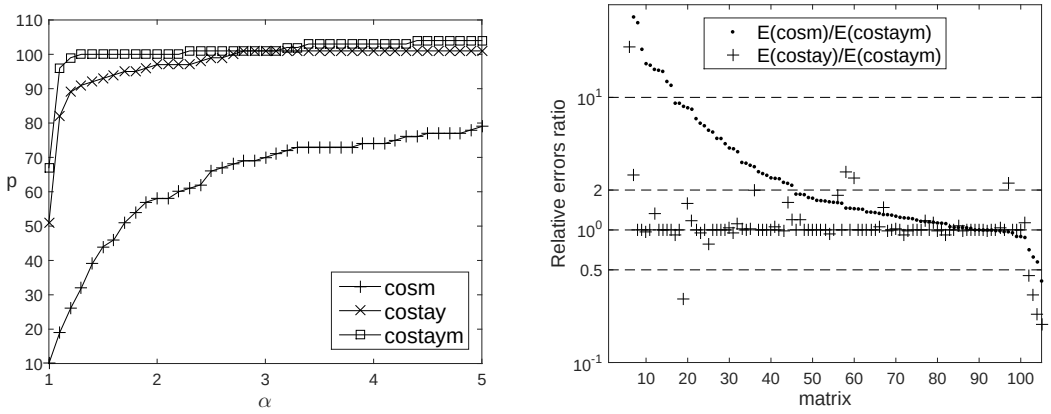
By using the fact that $\sin(A) = \cos(A - \frac{\pi}{2}I)$, Algorithm 1 can be easily used to compute the matrix sine. The computational cost of Algorithm 1 is $2k_0n^3$ flops provided $||B|| \leqslant \Theta_{m_k}$, or $2(k+s)n^3$ flops if $||B|| > \Theta_{m_k}$. The storage cost is $(2+q_{k0})n^2$ if $||B|| \leqslant \Theta_{m_k}$, and $(2+q_k)n^2$ otherwise.

3 Implementation and numerical experiments

In this section we compare `costaym`, a MATLAB implementation of Algorithm 1, with `costay`, based on Taylor approximation [3] (`http://personales.upv.es/~jorsasma/software/costay.m`), and `cosm`, based on Padé approximants [2, Alg. 4.2] (`http://github.com/sdrelton/cosm_sinm`). In tests we used MATLAB(R20014b) running on an Intel Core 2 Duo processor at 3.00 GHz with 4 GB main memory and 105 matrices: 10 diagonalizable 128×128 matrices, with 1-norms increasing from 2.50 to 25.06. 10 non diagonalizable Jordan block 128×128 matrices with eigenvalues whose algebraic multiplicity vary between 1 and 128 and 1-norms varying from 5.27 to 21.97. Forty three 128×128 matrices from the function `matrix` of the Matrix Computation Toolbox [7]. 14 matrices with dimensions lower or equal to 128 from the Eigtool MATLAB package [8], and 28 matrices from the matrix function literature. The "exact" matrix cosine was computed exactly for the first two sets of matrices, and by using MATLAB symbolic versions of Padé and Taylor algorithms with 4096 decimal digit arithmetic and several orders m and scaling parameters s higher than the ones used by `cosm` and `costaym`, respectively, in the other matrices. The relative differences between both Padé and Taylor approximations for these matrices were between $1.36\,10^{-23}$ and $5.52\ 10^{-25}$, except for two matrices where the relative differences were $7.29\,10^{-10}$ and $3.08\,10^{-07}$, but both `costaym` and `cosm` gave relative errors with respect to these "exact" values of orders between 10^{27} and 10^{43}, so there was no point in increasing the accuracy of these two "exact" values. The algorithm accuracy was tested by computing the relative error

$$\mathrm{E} = \frac{\|\cos(A) - \tilde{Y}\|_1}{\|\cos(A)\|_1},$$

where $\tilde{Y}$ is the computed solution and $\cos(A)$ is the exact solution. To compare the relative errors of the functions we plotted the performance profile and the ratio of relative errors of `cosm`, `costay` and `costaym` (with $m_k = 16$). In the performance profile (Fig. 1a), the α coordinate varies between 1 and 5 in steps equal to 0.1, and the p coordinate is the probability that the considered algorithm has a relative error lower than or equal to α-times the smallest error over all the methods. Fig. 1b shows the ratio of relative errors of `cosm` and `costay` both with `costaym`, in decreasing order of the ratio with `cosm` (the same order was used in Fig. 2). The Matrices 1 to 5 do not appear



(a) Perfomance profile. (b) Ratio of relative errors.

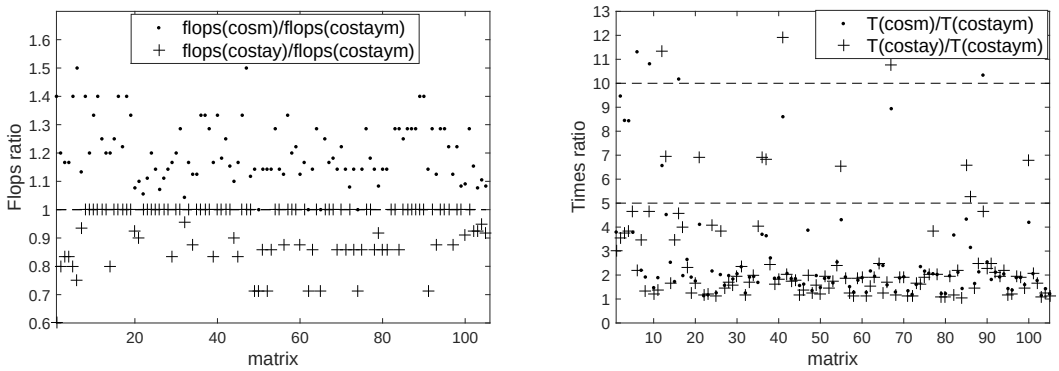
Figure 1: Accuracy tests.

in Fig. 1b since `costaym` error for them is 0. Fig. 1 shows that `costaym` was the most accurate function in tests, followed by `costay`. Figures 2a and 2b show the relative number of flops and the relative execution times (mean of 100 executions). The best results in terms of computational cost was `costay` (flop ratio) and `costaym` (execution time ratio). This discrepancy may be due because `costaym` uses far fewer times one estimator based on [9] for computing the 1-norm of matrix powers (computational cost $O(n^2)$) than the one used by `costay` and the size of the matrices is not so large so that the matrix products ($O(n^3)$) are the main term in cost. We have found that for the Toolbox matrices of dimension 1000 the execution time of costaym is slightly greater than the one of costay.

References

[1] S. Serbin, Rational approximations of trigonometric matrices with application to second-order systems of differential equations, Appl. Math. Comput. 5 (1) (1979) 75–92.

[2] A. H. Al-Mohy, N. J. Higham, S. D. Relton, New algorithms for computing the matrix sine and cosine separately or simultaneously, SIAM J. Sci. Comput. 37 (1) (2015) A456–A487.



(a) Ratio of flops.

(b) Ratio of execution times.

Figure 2: Computational cost tests.

[3] J. Sastre, J. Ibáñez, P. Ruiz, E. Defez, Efficient computation of the matrix cosine, Appl. Math. Comput. 219 (2013) 7575–7585.

[4] J. Sastre, J. J. Ibáñez, E. Defez, P. A. Ruiz, Accurate matrix exponential computation to solve coupled differential models in engineering, Math. Comput. Model. 54 (2011) 1835–1840.

[5] P. Alonso, J. Ibáñez, J. Sastre, J. Peinado, E. Defez, Efficient and accurate algorithms for computing matrix trigonometric functions, Submitted to J. Comput. Appl. Math. in Nov. 2015.

[6] P. Ruiz, J. Sastre, J. Ibáñez, E. Defez, High perfomance computing of the matrix exponential, J. Comput. Appl. Math. 291 (2016) 370–379.

[7] N. J. Higham, The Test Matrix Toolbox for MATLAB, Numerical Analysis Report No. 237, Manchester, England (Dec. 1993).

[8] T. G. Wright, Eigtool, version 2.1 (2009).
URL `web.comlab.ox.ac.uk/pseudospectra/eigtool`.

[9] N. J. Higham, Fortran codes for estimating the one-norm of a real or complex matrix, with applications to condition estimation, ACM Trans. Math. Softw. 14 (4) (1988) 381–396.

Modelling the spread of gender violence in Spain

Sonia Barreda[1]; Elena De la Poza*[2]; Lucas Jódar[1]

[1] Instituto Universtario de Matemática Multidisciplinar. Universitat Politècnica de València
[2] Centro de Ingeniería Económica. Departamento de Economía y Ciencias Sociales. Universitat Politècnica de València
* Corresponding author email: elpopla@esp.upv.es

Abstract

793 (September 2015) women were assassinated by their significant other in Spain since 2003 up until now. Even when the number of gender victims is comparable to the number of victims by traffic or terrorism, the Government has not assigned equal amount of resources to reduce the magnitude of this hidden social problem. In this paper we model the spread of this social disease, dividing the Spanish male population aged among the interval [16, 70] in mutually exclusive compartments according to degrees of gender abuse. Our results show a worsening of the trend of last years. However, recommendations are given to reduce the propagation of this problem.

1. Introduction

Gender violence is an old and universal problem, happening worldwide. The problem is inversely proportional to the level of egalitarianism of the country.
It is not until recent times that has been institutionally identified and condemned as a social urgency (1979).
The intimate violence problem is mainly hidden, due to several reasons such as:

- The low acceptance by the victims,
- It takes place in the core of the couple,
- It produces social shame, a feeling of emotional failure experienced by women,
- The strength of a "machismo" culture combined with the common victim behaviour: to forgive and forget lower levels of violence which by itself does not solve any IPV situation, all the contrary promotes its increase.

The magnitude of the problem is such that 793 (September 4th 2015) women were assassinated by their significant other or ex-couple in Spain since 2003 up until now; but only 7.2% of the murdered women had reported their fear and previous IPV to police.

Even when the number of gender victims is comparable to the number of victims by terrorism, the Government has not assigned equal amount of resources to reduce the magnitude of this hidden social problem.
The primary goal of this work consist of modelling the spread of this social disease to quantify the number of intimate partner aggressors aged [16, 74] in Spain during the period 2012-2017. Also, to provide recommendations (individual and Public Authorities) to reduce and prevent the magnitude of this social epidemic.

2. Methods

2.1 Population of study

At January 2012, male population aged [16, 74] in Spain amounted to 16,794,613 men. Then, the first step consists of splitting the population of potential men abusers in Spain aged among the interval [16, 74] in four compartments:
Gender Egalitarian men; regular man; emotional-psychological abuser and physical aggressor.

S: Gender Egalitarian man defined as men free of sexist micro-aggressions, (Altrocchi and Crosby, 1989)
N: Regular men defined as men influenced by the general sexist environment and they may practice low intensity sexist micro-aggression.
AS: Emotional-Psychological abuser, men who try to make their partners feel bad about themselves or they attack their partner's self-esteem, (Stith, 2011).
AF: Physical aggressor men who use physical force against their partner including forced sexual activity, (Stith, 2011).

S(n)=1,713,051, meant to be the 30% of the male population aged [16,40] at semester n=1. Indeed the 10.20% of the total male population, MacroSurvey, Spanish Ministry of Gender Issues, 2011).
N(n)= 8,901,145 represented the 53% of the total male population
AS(n)= 6,046,061 represented the 36% of the total male population, (Meil, 2012).
AF(n)= 134,359 amounted to the 0.8% of the total male population, (Spanish Ministry of Justice)

Following we identify the drivers that lead men to transit from one subpopulation to another. By modelling the transit coefficients we can build the mathematical model that allow us to predict the number of IPV aged [16, 74] in Spain in 2017.

2.2. Transition Coefficients' Modelling

Firstly, the factors that affect the transition between subpopulations were identified as follows:

-Demographics (birth & death rate and emigration)
-Women's role in the couple
-Men's alcohol and drug consumption
-Jealousy & wrong romantic stereotypes
-Economic stress (Long-term unemployment)
-Contagion effect (previous experiences of abuse)

In fact, Demographic, Economic, Sociological and Technological factors are taken into account.

Then, the transit coefficients are modelled and estimated:

$\beta_1(n)= \beta_1$=egalitarian transit coefficient (egalitarian couple influence, from N subpopulation to S)
$\beta_2(n)=\beta_2$= jealousy + drugs/alcohol consumption in a non-egalitarian relationship transit from N subpopulation to AS

$\beta_3(n)=\beta_3$ =drugs/alcohol consumption in a stable relationship transit from AS subpopulation to AF
$\beta_4(n)=\beta_4$=LT unemployment consumers of drugs/alcohol in a broken relationship from AS subpopulation to AF
$\beta_5(n)=\beta_5$=break up of a relationship with a non-egalitarian woman from S subpopulation to N
$\beta_6(n)=\beta_6$ =recovery transit due to a new relationship and therapy from AF subpopulation to N
$\beta_7(n)=\beta_7$= psychopathic rate transit from N subpopulation to AS
$\beta_8(n)=\beta_8$=Technology stacking transit from N subpopulation to AS
$\gamma_2(n)$= recovery transit by stable egalitarian couple from AS subpopulation to N

Following, it is explained how the coefficients are built.

- $\beta_1(n)=\beta_1$= egalitarian transit coefficient

The proportion of men aged [16,35] who are not jealous and their couple (stable relationship) is an egalitarian woman, β1=0.35*0.30*0.30= 0.03/4=0.0075.

- $\beta_2(n)=\beta_2$=jealousy + drugs/alcohol consumption in a non-egalitarian relationship transit

Assuming the hypothesis that amounts the 70% males suffer from jealousy, $\beta_2(n)$ measures the proportion of jealous men consuming drugs and/or alcohol in a stable relationship with a non-egalitarian woman. $\beta_2(n)=[1/2]*(0.7*(0.05+0,0E))*0.7=0,1225+0.035E$, where E is a perturbation coefficient due to the difficulty to measure the proportion of consumers of drugs and alcohol. We assume the variation range of E is 2% centred about 5%

- $\beta_3(n)=\beta_3$= drugs/alcohol abuse in a stable relationship transit.

It is estimated as $\beta_3=0,105/2=0,0525$. It reflects the proportion of men aggressors consumers of drugs and/or alcohol in a stable relationship,(Márkez et al, 2004)

- $\beta_4(n)$= Long term unemployment and consumers of drugs/alcohol abuse breaking the relationship.

It is estimated as $\beta_4(n)=[1/2]*[(0.5*0.27*0.02)+(0.5*(0.05+0.0E))]=0.026+0.025E$
It measures the proportion of long-term unemployed men consumers of drugs and alcohol whose relationship breaks (Márkez et al, 2004)

- $\beta_5(n)=\beta_5$=break up of a relationship with a non-egalitarian woman.

Estimated as $\beta_5=[1/2]*(0.7*0.174*0.5)=0.03045$. It measures the proportion of men in a relationship with a non-egalitarian woman experiencing the break of the relationship, (Survey of Health and Sexual Habits & Ethology)

- $\beta_6(n)=\beta_6$=recovery transit due to a new relationship and therapy

Estimated as $\beta_6=[1/2]*[(0,5*0.3)+(0.5*0.7*0.005)]=0.08375$. It measures the proportion of separated or divorced men engaging in a new relationship with an egalitarian woman plus the proportion of separated men engaging in a new relationship with a non-egalitarian women who attends to therapy (Stith, 1991)

- $\beta_7(n)$=psychopathic rate transit (Genetic).

$\beta_7(n)$=1%. It is the proportion of psychopathic men, (1%)=0.01. (Garrido, 2008).

- $\beta_8(n)=\beta_{8=}$ Technology stacking transit

Built as $\beta_8=1/2*[0,5*0,7*0,5*0,7]= 0,06125$. It measures the proportion of jealous men who are aged in the interval [16, 40] in a relationship with a non-egalitarian woman who uses technology as the channel for executing or exercising control aggression to their couples. (Logan et al., 2006)

- $\gamma2(n)$= recovery transit by stable egalitarian couple

It is measured as $\gamma2(n)=1/2*[0.35*0.3*1]=0.0649$. It represents the proportion of men aged in the interval [16,35] who starts a sta*ble* relationship (at least 1 year) with an egalitarian woman.

Finally, the demographic factors are defined as follows:
Incomers in the model due the male population becoming 16 years old, (INE). We assume these proportions remain constant for the short period of study. This coefficient is denoted as $\alpha_i(n-32)$.
Those who exit the model, there are two types, the population becoming 75 years old but also deceased population between [16,74].
In both cases, we assume these amounts remain constant for the short period of study considered. This coefficient is denoted as $D_i,\ 1 \le i \le 4$.

Finally, the emigration due to economic reasons (Izquierdo et al., 2015). We assume the amount of emigrants is sized according to the 2012 proportion of each subpopulation.

$$g_{1ri(n)} = \begin{cases} 113203*ri;\ 1 \le i \le 4 \quad n=0;\ 1\ (2012) \\ 131204*ri;\ 1 \le i \le 4 \quad n=2,\ 3\ \ (2013) \\ 49736*ri;\ 1 \le i \le 4 \quad\ \ n=4,\ 5\ \ (2014) \\ 0 \qquad\qquad\qquad\qquad 11 \ge n \ge 6\ \ (2015,\ 2016,\ 2017) \end{cases}$$

where r_i is the proportion of each subpopulation estimated at Jan 2012.

2.3 Model

The model is expressed as follows:

$$S(n+1) = S(n) + b_1 N(n) + a_1(n-32) - D_1 - g_{1ri}(n) - b_5 S(n)$$

$$N(n+1) = N(n) - b_1 N(n) + a_2(n-32) - D_2 - g_{1r2}(n) + b_5 S(n) + b_6 AF(n) + g_2 AS(n) - b_2 N(n) + T$$

$$AS(n+1) = AS(n) + a_3(n-32) - D_3 - g_{1r3}(n) - (b_3 + b_4) AS(n) - g_2 AS(n) + b_2 N(n) - T$$

$$AF(n+1) = AF(n) + a_4(n-32) - D_4 - g_{1r4}(n) + (b_3 + b_4) AS(n) - b_6 AF(n)$$

Figure 1 draws the transits between subpopulations:

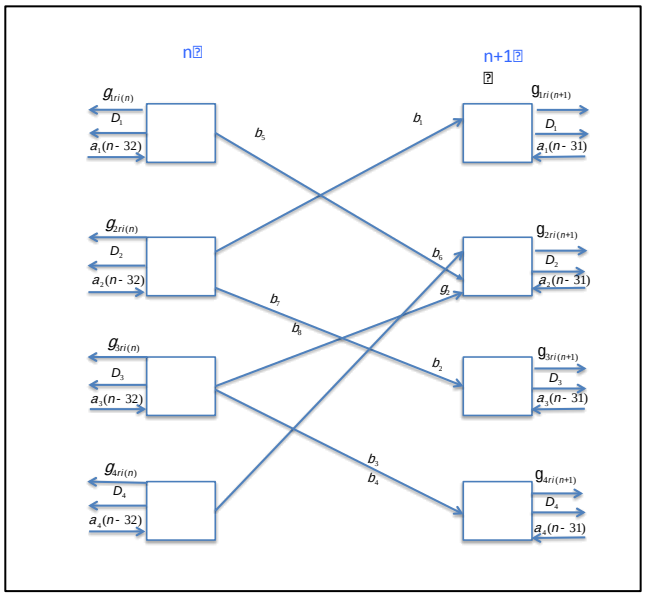
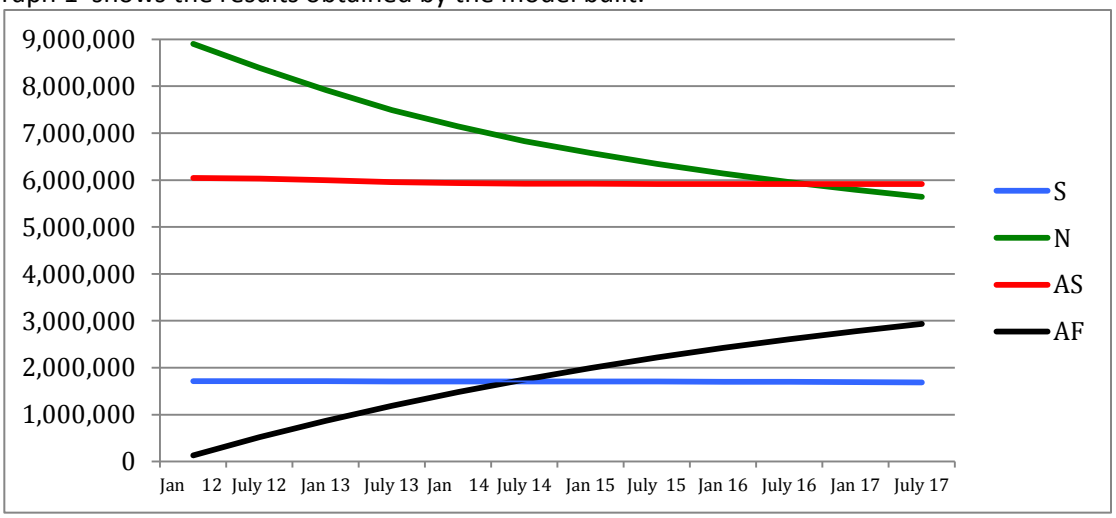


Figure 1. Dynamic Block Diagram

3. Results

Graph 1 shows the results obtained by the model built.



Graph 1. Forecast of the subpopulations.

4. Conclusions

The model quantifies the future population of aggressors (psychological and physical) taking into account the dynamic factors such as demographic, economic and socio-cultural (alcohol, drugs, jealousy, marital separations).

An important underlying consequence of the study shows the prevalent role played by the woman in each potential or real event of domestic violence, up the point that the main recommendation to overcome this dramatic problem is the active decision of women breaking up the relationship at the early stages of the psychological abuse.

The model also surges the hidden population of aggressors that is not quantified by the official statistics, which show the problem when it is irreparable.

The study is susceptible to be applied to any other geographical area where data become available as well as the period of study that can be also changed although it is important to take into account that the longer period the results obtained are less reliable.

References

J. Altrocchi and R.A. Crosby, Sex Roles, June 1989, Volume 20, Issue 11-12, pp 639-648
Clarifying and measuring the concept of traditional vs. egalitarian roles in marriages

MJ Diaz-Aguado; R Martinez. (2014), "Types of Adolescent Male Dating Violence Against Women, Self-Esteem, and Justification of Dominance and Aggression"*J Interpers Violence*, October 17, 2014

V. Garrido, (2013), *El psicópata*. Algar, 2000

M. Izquierdo, J. F. Jimeno and A. Lacuesta, (2015), "Spain: From Inmigration to Emigration?", Working paper N1503, Spanish Central Bank, Spain, Eurosystem.

TK. Logan, J Cole, L Shannon. "Partner Stalking: How Women Respond, Cope, and Survive", Springer: New York, United States, 2006.

I. Márkez, C. Romera, C. Merino, X. Arana, M. Calvo, A. Peleteiro, M. Poo, (2002), Violencia doméstica, consumo de sustancias y otras circunstancias concurrentes. ¿El derecho versus derecho a la salud? *Revista de la Asociación Española Neuropsiquiatría*, *XXII* (83), 125-133.

G. Meil, (2011), Análisis de la Encuesta sobre percepción social de la violencia de género. Ministerio de Sanidad, Servicio Sociales e Igualdad. Delegación del Gobierno para la Violencia de Género.

C. Southworth, J. Finn, S. Dawson, C. Fraser and S. Tucker, (2007), "Intimate Partner Violence, Technology, and Stalking", *Violence Against Women*, 13; 842. DOI: 10.1177/1077801207302045

Spanish Statistiscs Institute, www.ine.es

S. M. Stith, E.E. McCollum, K.H.Rosen, (2011), Couples Therapy for Domestic Violence, *American Psychological Association*, Washington, DC.

Mathematical modelling of radicalization processes based on the example of right-wing extremism in Germany

Tanja Deutsch[1], Matthias Ehrhardt[1,]

[a] *Lehrstuhl für Angewandte Mathematik und Numerische Analysis, Fachbereich C – Mathematik und Naturwissenschaften, Bergische Universität Wuppertal, Gaußstrasse 20, 42119 Wuppertal, Germany*

1. Introduction

Which social and economical factors encourage right-wing extremism? How do cultural values change by the influence of an extreme behavior? Extreme behavior is produced by a small group, but affects a large amount of the whole population. Which social principles determine right-wing extremism? On the one hand modelling in social science is an anthropological question: the central assumption is the synthetical theorie of evolution. The Human beeing is part of the genetic process in nature and is seen as a cultural animal. The evolution of cultural values has similarities with the synthetical theorie of evolution and culture is a result of the same process, which affects all creatures.[1]

On the other hand the spread of right-wing extremism is modeled by an epidemiological model. Manifest and organised right-wing extremism in Germany is integrated in a broad enviroment of latent right-wing extremism. To describe the ideological landscape of Germany in a mathematical way it is necessary to devide the German population into subgroups. In April 2002 Oliver Decker and Elmar Brähler from Leipzig University undertook a respresentive study about "right-wing extremism attitudes in Germany" for the first time. They created a questionnaire, which consists of six topics with three questions each. These six "dimensions" of right-wing extremism are as follows.

1. **Endorsement of a right-wing authoritarian dictatorship**
2. **Chauvinism**
3. **Xenophobia**
4. **Antisemitism**
5. **Social darwinism**
6. **Trivialisation of National Socialism**

All 18 statements were valued by the survey participants with the help of a Likert scale (strongly disagree, mainly disagree, neither agree nor disagree, mainly agree, strongly agree). This leads to five subgroups each question. 2004 the survey was repeated and as from 2006 realised by the Friedrich Ebert Foundation every two years. For each survey representative areas in Germany are

*Corresponding author

[1] see R. Boyd,
Culture and the Evolutionary Process,
University Of Chicago Press, second edition, 15th June 1988, page 1-4

determined by a sample-point-selection. The subgroubs as follows are thus representive of the German society.
For defining social subgroups respective the right-wing extremism in Germany we calculate the mean value derived from the three questions each "dimension". For the years 2002 to 2014 follows for each "dimension" and each subgroup one specific value and we obtain a classification of the German society respective right-wing extremism over a period of 13 years:

- **G(t)** Number of **opponents**, who disagree with the statements of the respective "dimension".
- **N(t)** Number of people, who have a **negative** attitude towards the statements of the respective "dimension" and mainly disagree with them.
- **U(t)** Number of **undecided**, who neither agree nor disagree with the statements of the respective "dimension".
- **S(t)** Number of **semi fanatics**, who mainly agree with the statements of the respective "dimension".
- **R(t)** Number of **radicals**, who agree with the statements of the respective "dimension".

2. The ETA-model

A paper[2] written by Francisco J. Santonja, Ana C. Tarazona and Rafael J. Villanueva published in 2008 describes in a mathematcial way the spread of an extreme ideology by the example of ETA in the bask country. In this model the size of the total population is not constant, but depends on intrinsic and extrinsic influencing variables. There is a birth rate and mortality rate, which have some proportional influence on the subgroups. $\Lambda(t)$ is the number of births and $\Phi(t)$ the number of deaths at time t. Extrinsic the number of immigrants $\Gamma(t)$ and the number of emigrants $\Sigma(t)$ influence proportional to their sizes subgroups G and N. In this context parameters α_1 and α_2 describe due to social climate the relationship between immigration and emigration at time t. Analogous to the model built in Valencia we model the spread of right-wing extremism in Germany. The dynamics between the subgroups are described by transition terms. It is particularly interesting how subgroup U behaves, because it is the fragile part of society, which is susceptible to radicalization.

After scaling the equations, we obtain the following system of ordinary differential equations.

$$g'(t) = (\sigma(t) - \gamma(t))g(t) + \alpha_2\gamma(t) - k\beta_1 g(t)n(t) - \alpha_1\sigma(t), \tag{1}$$

$$n'(t) = (\sigma(t) - \gamma(t))n(t) + (1-\alpha_2)\gamma(t) - \beta_1 n(t)u(t) - (1-\alpha_1)\sigma(t), \tag{2}$$

$$u'(t) = (\sigma(t) - \gamma(t))u(t) + \beta_1 n(t)u(t) + \beta_2 s(t)u(t), \tag{3}$$

[2] see F. J. Santonja, A. C. Tarazona und R. J. Villanueva,
A mathematical model of the pressure of an extreme ideology on a society,
Elsevier, August 2008

$$s'(t) = (\sigma(t) - \gamma(t))s(t) - \beta_2 s(t)u(t) + l\beta_2 r(t)s(t), \tag{4}$$

$$r'(t) = (\sigma(t) - \gamma(t))r(t) - l\beta_2 r(t)s(t). \tag{5}$$

The scaled system of differential equations is a non-autonomous system, because the rates depend on time. They are also non-linear, because they contain squared contact-terms.

We obtain the following optimal parameters for each "dimension".

Table 1: The optimal parameters of systems (1)-(5) in all six "dimensions".

	1	2	3	4	5	6
β_1	-0,2105	-0,2517	-0,1329	-0,1914	-0,2310	-0,1747
β_2	0,1854	0,2081	0,0767	0,1003	0,1272	0,099
k	0,2836	0,2024	0,0303	0,1633	0,4105	0,2765
l	1	1	0,0467	1	0	0,335
α_1	0	0	0	0	0	0
α_2	1	1	1	1	1	1

With the computed optimal parameters we get a first simulation. It represents the "natural" evolution of the German population:

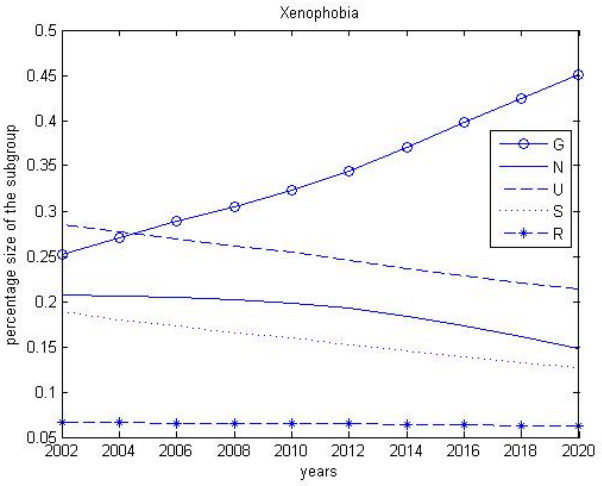


Figure 1: Simulation in "dimension" xenophobia from 2002 to 2020 using the optimal parameters.

3. The Castillo-Chavez-Song-model

Carlos Castillo-Chavez and Baojun Song deal with the question, which social landscapes support the existence of fanatic subcultures[3]. Especially the effects of "core" (ultra) fanatics on the spread of extreme ideologies is analysed. Therefore the population is divided in a non-core population $A := G+N$, which is used as a recruitment pool, and a core population C, which is the radical part of society. We define C, as $C := U + S + R$. The core population is relatively small and organized hierarchically respective size and radicalization:

Subgroup U is the biggest part and the most susceptible one. People in this subgroup are not radical yet, but fragile and susceptible to radicalization. Semi fanatics, who are partly convinced by right-wing extremism, are in subgroup S. Subgroup R is in general the smallest part of society and contains of right-wing extremists. As a whole it is $T = A + C = G + N + U + S + R$.

Because in this model the core population is particularly interesting, scaling the associated differential equations leads to following scaled system:

$$\frac{dU}{dt} = \mu_1(1 - C(t))C(t) - \mu_2 U(t)\frac{S(t) + R(t)}{C(t)} - \gamma_1 U(t), \tag{6}$$

$$\frac{dS}{dt} = \mu_2 U(t)\frac{S(t) + R(t)}{C(t)} - \mu_3 S(t)\frac{R(t)}{C(t)} - \gamma_2 S(t), \tag{7}$$

$$\frac{dR}{dt} = \mu_3 S(t)\frac{R(t)}{C(t)} - \gamma_3 R(t). \tag{8}$$

Analogous to the numerical implementation of the ETA-model, we implement the Castillo-Chavez-Song-model in Matlab and obtain the following optimal parameters.

Table 2: The optimal parameters of the Castillo-Chavez-Song-model (6)-(8) in all six "dimensions".

	1	2	3	4	5	6
μ_1	0,5	-0,0296	0,0088	0,8609	0,9207	0,9837
μ_2	-0,1009	-0,0904	0,0819	-0,0793	-0,0880	0,1051
μ_3	-0,2311	-0,2937	0,1096	-0,1177	0,2742	0,9667
γ_1	0,5985	0,0078	0,0040	0,8744	0,9967	0,9963
γ_2	0,0218	0,0245	0,0444	0	0	0,0003
γ_3	0,0196	0	0,0248	0,0206	0,0515	0,1442

The "natural" evolution of the German population from 2002 to 2020 with respect to right-wing extremism is displayed by a simulation with the optimal parameters.

[3] see C. Castillo-Chavez und B. Song,
Models for the Transmission Dynamics of Fanatic Behaviors,
Bioterrorism: Mathematical Modeling Applications in Homeland Security
Frontiers in Applied Mathematics, Book 29, 2003

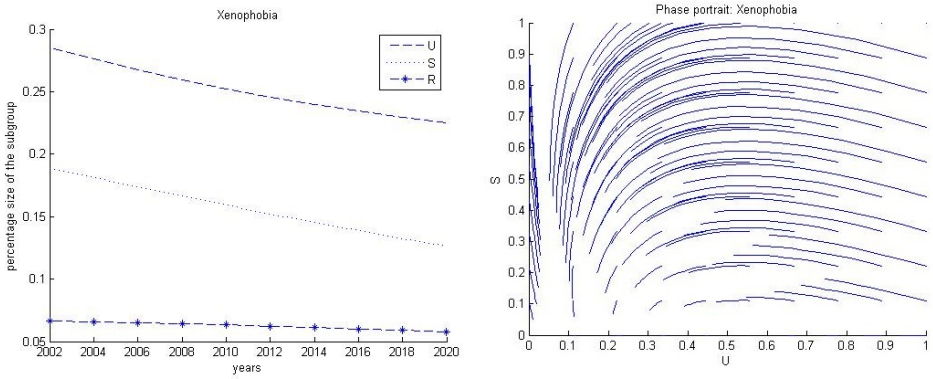


Figure 2: Solutions and phase portrait of the simulation with the optimal parameters and data from the "dimension" xenophobia. All subgroups decrease continuously. The subgroup of undecided people decreases to 0,225. The subgroup of semi fanatics decrease to 0,125 and the subgroup of radicals decrease to 0,05. The global attractor is (0,0).

In some further simulation we change some parameters to simulate a radicalization of society. The aim is to show the functioning of the model.

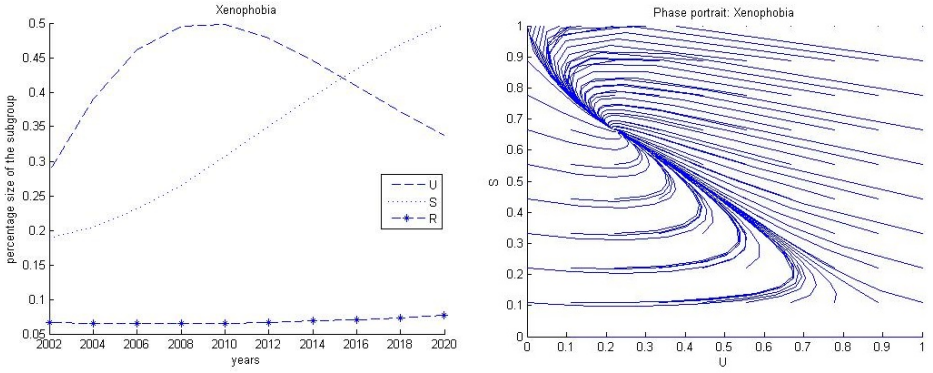


Figure 3: Solutions and phase portrait of the simulation with $\mu_1 = 0,3$, $\mu_2 = 0,2$ and $\mu_3 = 0,1$ in the "dimension" xenophobia. The subgroup of undecided people increases strongly, then decreases strongly. The subgroup of semi fanatics increases continuously. The subgroup of radicals increases slightly. The global attractor is (0,19578; 0,686).

4. Conclusion

The six different "dimensions" of right-wing extremism have complex analysis facilities. The classification of society respective right-wing extremism leads to two different models each "dimension", which at their hearts consist of a system of differential equations. Comparing the models with each other, it is clear that contrary to the ETA-model Castillo-Chavez and Song model a

circulation. Individuals, which come from a core subgroup and go via "recovery" γ_i, $i = 1;2;3$, back to the moderate part of society, become part of the recruitment pool again. γ_1 is in the most "dimensions" of right-wing extremism the highest one, whereby a remarkable small value ($\gamma_1 = 0,004$) occurs in "dimension" xenophobia. This implies that the decimation of subgroup U is significantly due to a transition to A. In "dimension" xenophobia subgroup U could be more susceptible to radicalization than to recovery. This conclusion is further suggested by $\mu_2 = 0,0819$, which is a positive value. When γ_1 is the highest parameter, it can be derived that a radicalization is stopped after the transition from A to U by an early movement back to the moderate part of society. Just a few individuals become more radical. Besides xenophobia, the "dimension" chauvinism makes an exception. In this "dimensions" μ_3 is according to amount the highest parameter. Because it is $\gamma_3 = 0$ at the same time, the decimation of subgroup R is explained by a transition from R to S.
In the ETA-model the dynamics occurring are linear $G \leftrightarrow N \leftrightarrow U \leftrightarrow S \leftrightarrow R$, which implies a step-by-step radicalization or moderation. Explicitly a seperat immigration and emigration rate for subgroups G and N are considered. For each "dimension" of right-wing extremism we get $\alpha_1 = 0$ and $\alpha_2 = 1$. The model gives the impression that the social climate in Germany causes an immigration to subgroup G and an emigration from subgroup N. The parameter l in "dimension" xenophobia is relatively small, which implies a slight dynamic $R \rightarrow S$. In the "dimensions" endorsement of a right-wing authoriterian dictatorship, chauvinism and antisemitism is $l = 1$, which leads to equipollent transitions $U \leftrightarrow S$ and $S \leftrightarrow R$. In all "dimensions" is $\beta_1 < 0$ and at the same time $\beta_2 > 0$, whereby according to amount β_1 is higher than β_2. This we interpret as the central dynamic

$$G \leftarrow N \leftarrow U \leftarrow S \leftarrow R.$$

Especially in "dimension" social darwinism k is relatively high, which implies a strong transition from N to G.
By calculating thresholds and attractors Castillo-Chavez and Song show that the most effective approach for the eradication of the right-wing extremism comes from sufficient effort to limit recruitment into the radical core group C: The control threshold is $L_1 = \frac{\mu_1}{\gamma_1}$. One way to bring this quantity below 1 is to reduce the value of μ_1. This reduction corresponds to an increase in the resistance of the general population A to "advances" from the core. The elimination of the fanatic population R is critically important. $L_3 = \frac{\mu_3}{\gamma_3} < 1$ implies that the fanatic population will crash regardless of its size. Since the value of γ_3 is actually a tiny number or, equivalently, since the residence time $\frac{1}{\gamma_3}$ is long, then it is quite unlikely that L_3 could be made less than 1. This scenario leads to an increase of R.

Applying the Structural Equation Model to Co-creation in Degree Programs in Ecuadorian Universities

Ribes-Giner, G.[♭] , Pantoja Díaz, O. [†] [*], and Peralt, A. [‡]

(♭) Universidad Politécnica de Valencia,

España,

(†) Escuela Politécnica Nacional,

Ecuador,

(‡) European University of Valencia,

España,

November 30, 2015

1 Abstract

Although a number of studies have been carried out on co-creation in innovative marketing, only a few apply this approach to higher education. The aim of this paper is to study the links between participation, communication, co-creation and satisfaction in undergraduate programs, in order to validate a model that could be applied in this field to enhance strategic management. In this approach the student is considered the cornerstone of the collaboration and a valuable stakeholder in the academic process. The literature was reviewed for topics supporting the proposed model and Structural Equation Modeling was applied to validate it. The results demonstrate that co-creation is an innovative approach that guarantees higher levels of satisfaction in the university context.

[*]e-mail: odette.pantoja@epn.edu.ec

2 Introduction

Co-creation is related to co-creation and co-production and is designed to involve clients as active participants in different steps of creating services or products. The co-creation approach has been studied by various authors [5, 9]. This tendency has been analyzed in different contexts, including information technology [4], the footwear sector [15], tourism, travel services [6] and financial services [2]. In the field of higher education, [17]; [16]; and [12] focused on postgraduates programs and found co-creation had a positive impact and generated student loyalty.

The objective of the present study is to analyze the different relationships that emerge when co-creation is applied to undergraduate students. This research was expected to allow us to validate or reject a proposed co-creation model, designed to be adopted as a strategic management tool in order to give universities a competitive advantage.

Adopting co-creation as a plausible alternative makes it possible to connect with students from the initial phase of idea generation [9]. In this way, the university obtains enhanced information on what students actually want [3] and from the students' perspective it improves their perception, trust, satisfaction and loyalty.

The study firstly focuses on a review of the literature, in order to propose a solid model based on co-creation. The principal constructs analyzed are communication, participation, co-creation and satisfaction. Exploratory and confirmatory factor analysis and structural equation modeling were the statistical tools implemented to validate the model, and undergraduate programs were used as the case study.

3 The Conceptual Model

This section delves into the theoretical foundation of the proposed co-creation model, with participation, communication, co-creation and satisfaction as the principal constructs.

3.1 Communication vs Participation

Several studies, e.g. [1, 8] have shown the relationship between communication and participation.

3.2 Participation vs Co-Creation

[2] and [5] comment on the direct and positive effect of participation on co-creation.

3.3 Communication vs Co-Creation

[11, 10, 13] have shown that communication with the customer is vital for successful co-creation.

3.4 Co-Creation vs Satisfaction

Satisfaction is one of the principal results of co-creation. This relationship has been studied by authors such as [14, 6] and [7].

3.5 The Proposed Co-creation Model

The proposed model is shown in Figure 1, and includes all the aforementioned relationships between communication and participation, participation and co-creation, communication and co-creation and co-creation and satisfaction.

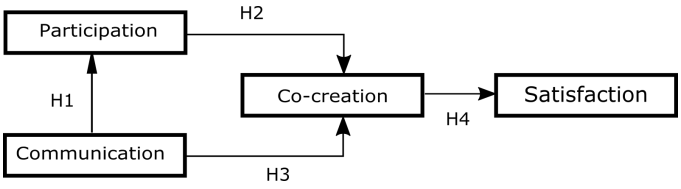


Figure 1: The Co-creation Research Model.

4 Methodology and Results

A questionnaire was given to 395 Ecuadorian undergraduate students from 11 different universities, using a seven-item Likert scale.

The exploratory factor analysis was applied to validate the instruments used in the SPSS program. The Varimax Rotation and the Maximum Likelihood extraction method with 4 fixed components were used to detect problems and found four items in the first iteration. In the second iteration the

conflicting items were excluded and a Cronbach's alpha of about 0.906 was obtained (over 0.7 is good). The explained variance was about 64.4%, the KMO value was 0.910 (higher than 0.5), and the Bartlett test gave p=.000.

A confirmatory factor analysis was carried out with the objective of exploring the associations between items and constructs, using SPSS AMOS (Analysis of Moment Structures) program. The convergent validity was verified, since all the constructs had an average variance extracted (AVE) above 0.5. The composite reliability of the constructs also showed values higher than 0.7, and the factor loadings for communication, participation, co-creation and satisfaction were higher than 0.5. The estimated coefficients of each item were all significant (t-value > 2.0). The discriminant validity was checked through the square roots of the AVEs, in which all the constructs had higher values than the inter-construct correlations.

The final analysis was by the commonly used structural equation modeling (SEM), to find the relationships between the constructs. Different indicators were studied to determine whether the model had a good fit. The comparative fix index (CFI) had a good value of 0.96 (over 0.95), and the adjusted goodness-of-fit index (AGFI) also showed a good value of 0.903 (>0.8). The root mean square residual (RMR) was 0.076 (<0.09), the normative fit index (NFI) was 0.948 and the root mean square error of approximation (RMSEA) was 0.077 (<0.08). The squared multiple correlation of co-creation showed that 62% (a high value) of this element is explained by the direct effect of participation and the direct and indirect effects of communication. In the case of participation, half of its variance (46%) is explained by the direct impact of communication and 68% of the satisfaction variance is explained by the direct effect of co-creation.

Also studied was mediation by participation in the relationship between communication and co-creation, where a poor but significant partial mediation was detected.

These results led us to conclude that the four relationships studied have significant and positive impacts, thus proving that the four hypotheses related to the undergraduate programs context were correct:

- Communication has a positive and significant impact above co-creation (γ =0.62, p-value <0.001).
- Communication has a positive and significant impact on participation (γ=0.68, p-value < 0.001).

- Participation has a positive and significant effect on co-creation with a lesser impact (β=0.22, p-value = 0.003).
- Co-creation has a significant and the highest impact on satisfaction (β=0.83, p-value < 0.001).

References

[1] James C Anderson and James A Narus. A model of distributor firm and manufacturer firm working partnerships. *the Journal of Marketing*, pages 42–58, 1990.

[2] Seigyoung Auh, Simon J Bell, Colin S McLeod, and Eric Shih. Co-production and customer loyalty in financial services. *Journal of retailing*, 83(3):359–370, 2007.

[3] Lance A Bettencourt, Stephen W Brown, and Nancy J Sirianni. The secret to true service innovation. *Business Horizons*, 56(1):13–22, 2013.

[4] Ja-Shen Chen, Hung-Tai Tsou, and Russell KH Ching. Co-production and its effects on service innovation. *Industrial Marketing Management*, 40(8):1331–1346, 2011.

[5] Michael Etgar. A descriptive model of the consumer co-production process. *Journal of the Academy of Marketing Science*, 36(1):97–108, July 2007.

[6] Ursula S Grissemann and Nicola E Stokburger-Sauer. Customer co-creation of travel services: The role of company support and customer satisfaction with the co-creation performance. *Tourism Management*, 33(6):1483–1492, 2012.

[7] Christian Grönroos. Service logic revisited: who creates value? and who co-creates? *European Business Review*, 20(4):298–314, 2008.

[8] Thomas Kohler, Kurt Matzler, and Johann Füller. Avatar-based innovation: Using virtual worlds for real-world innovation. *Technovation*, 29(6):395–407, 2009.

[9] Per Kristensson, Jonas Matthing, and Niklas Johansson. Key strategies for the successful involvement of customers in the co-creation of new technology-based services. *International Journal of Service Industry Management*, 19(4):474–491, 2008.

[10] Anders Lundkvist and Ali Yakhlef. Customer involvement in new service development: a conversational approach. *Managing Service Quality*, 14(2/3):249–257, 2004.

[11] Albert M Muñiz Jr and Hope Jensen Schau. How to inspire value-laden collaborative consumer-generated content. *Business Horizons*, 54(3):209–217, 2011.

[12] Agustín Peralt Rillo and Gabriela Ribes-Giner. Una orientación proactiva hacia el mercado para los programas de Postgrado. *Dirección y Organización*, 50:37–47, 2013.

[13] C.K. Prahalad and Venkat Ramaswamy. Co-creating unique value with customers. *Strategy & Leadership*, 32(3):4–9, January 2004.

[14] Coimbatore Krishna Prahalad and Venkat Ramaswamy. *The future of competition: Co-creating unique value with customers*. Harvard Business Press, 2013.

[15] Venkat Ramaswamy. Co-creating value through customers' experiences: the Nike case. *Strategy & Leadership*, 36(5):9–14, 2008.

[16] Gabriela Ribes-Giner, A. Peralt Rillo, and Ismael Moya Clemente. Co-creation innovation model for masters programs in the universities. *Innovation and Teaching Technologies*, 117, 2014.

[17] Gabriela Ribes-Giner and Agustín Peralt Rillo. Structural equation modeling of co-creation and its influence on the student's satisfaction and loyalty towards university. *Journal of Computational and Applied Mathematics*, 2015.

Front-fixing Transformation for Regime Switching Model of American Options

V.N. Egorova[♭] [*], R. Company[♭], and L. Jódar[♭]

([♭]) Universitat Politécnica de Valencia,

Camino de Vera s/n, 46022 Valencia, Spain

November 30, 2015

1 Introduction

We consider the American put option on the asset $S_t = S$ with strike price E and maturity $T < \infty$. Let $V_i(S, \tau)$ denote the option price functions, where $\tau = T - t$ denotes the time to maturity, the asset price S and the regime $\alpha_t = i$. Then, $V_i(S.\tau)$, $1 \leq i \leq I$, satisfy the following free boundary problem:

$$\frac{\partial V_i}{\partial \tau} = \frac{\sigma_i^2}{2} S^2 \frac{\partial^2 V_i}{\partial S^2} + r_i S \frac{\partial V_i}{\partial S} - r_i V_i + \sum_{l \neq i} q_{il}(V_l - V_i), \quad S > S_i^*(\tau), 0 < \tau \leq T, \tag{1}$$

where $S_i^*(\tau)$ denote optimal stopping boundaries of the option. Initial conditions are

$$V_i(S, 0) = \max(E - S, 0), \quad S_i^*(0) = E, \quad i = 1, ..., I. \tag{2}$$

[*]e-mail: egorova.vn@gmail.xom

Boundary conditions for $i = 1, .., I$ are as follows

$$\lim_{S\to\infty} V_i(S,\tau) = 0, \tag{3}$$

$$V_i(S_i^*(\tau),\tau) = E - S_i^*(\tau), \tag{4}$$

$$\frac{\partial V_i}{\partial S}(S_i^*(\tau),\tau) = -1. \tag{5}$$

2 Front-Fixing transformation

Based on the transformation used by the authors in [4], [1] for the case of just one equation, let us consider the multivariable transformation

$$x^i = \ln \frac{S}{S_i^*(\tau)}, \quad 1 \le i \le I. \tag{6}$$

Note that the new variables x^i lie in the fixed positive real line. Price V_i of i-th regime involved in i-th equation of the system and i-th free boundary are related by the dimensionless transformation

$$P_i(x^i,\tau) = \frac{V_i(S,\tau)}{E}, \quad X_i(\tau) = \frac{S_i^*(\tau)}{E}, \quad 1 \le i \le I. \tag{7}$$

Value of option l-th regime appearing in i-th coupled equation, $l \neq i$, becomes

$$P_{l,i}(x^i,\tau) = \frac{V_l(S,\tau)}{E}. \tag{8}$$

Since from (7), $\frac{V_l(S,\tau)}{E} = P_l(x^l,\tau)$ and taking into account transformation (6) for indexes i and l one gets that

$$P_{l,i}(x^i,\tau) = P_l(x^l,\tau), \tag{9}$$

and it occurs when the variables are related by the equation

$$x^l = x^i + \ln \frac{X_i(\tau)}{X_l(\tau)}, \quad 1 \le i, l \le I. \tag{10}$$

From (6) - (9) the problem (1) - (5) for $1 \le i \le I$ takes a new form:

$$\begin{aligned} &\frac{\partial P_i}{\partial \tau}(x^i,\tau) = \frac{\sigma_i^2}{2}\frac{\partial^2 P_i}{\partial (x^i)^2}(x^i,\tau) + \left(r_i - \frac{\sigma_i^2}{2} + \frac{X_i'(\tau)}{X_i(\tau)}\right)\frac{\partial P_i}{\partial x^i}(x^i,\tau) \\ &-r_i P_i(x^i,\tau) + \sum_{l\neq i} q_{il}(P_{l,i}(x^i,\tau) - P_i(x^i,\tau)) = 0,\ x^i > 0,\ 0 < \tau \le T, \end{aligned} \tag{11}$$

with initial and boundary conditions

$$P_i(x^i, 0) = \max(1 - e^{x^i}, 0) = 0, \quad X_i(0) = 1, \tag{12}$$

$$P_i(0, \tau) = 1 - X_i(\tau), \quad \frac{\partial P_i}{\partial x^i}(0, \tau) = -X_i(\tau), \tag{13}$$

$$\lim_{x^i \to \infty} P_i(x^i, \tau) = 0. \tag{14}$$

Note that from equation (10) x^l could be negative if $X_l(\tau) > X_i(\tau)$ and this means that due to the equation (6) $S < S_l^*(\tau)$, and in this case the value of the option at l-th regime agrees with the payoff, i.e.

$$P_{l,i}(x^i, \tau) = P_l(x^l, \tau) = 1 - X_l(\tau)e^{x^l}, \quad x^l \leq 0. \tag{15}$$

3 Numerical scheme construction

A numerical solution has to be found on infinite domain $[0; \infty) \times [0; T]$ for all regimes. It is sufficient to take the numerical domain for the transformed problem (11)-(14) as $[0; x_{\max}]$, $x_{\max} = 3$. The computational domain is covered by an uniform grid with common step sizes $h = \frac{x_{\max}}{M}$ and $k = \frac{T}{N}$. Nodes of the grid are denoted as follows

$$x_j = jh,\ 0 \leq j \leq M; \quad \tau^n = nk,\ 0 \leq n \leq N. \tag{16}$$

Let us denote $u_{i,j}^n \approx P_i(x_j, \tau^n)$ the approximation of P_i in i-th equation at mesh point $(x^i = x_j, \tau = \tau^n)$ and $\tilde{u}_{l_i,j}^n \approx P_{l,i}(x_j, \tau^n)$ be the approximation of P_l in i-th equation evaluated at the point $(x^i = x_j, \tau = \tau^n)$. The discretization of the transformed optimal stopping boundary is denoted by $X_i^n \approx X_i(\tau^n)$.

Let us denote

$$\tilde{u}_{l_i,j}^n \approx P_{l,i}(x_j, \tau^n) = P_l\left(x_j + \ln\frac{X_i^n}{X_l^n}, \tau^n\right), \tag{17}$$

the values are obtained by linear interpolation of values $u_{l,j}^n$ at the point $x_j + \ln\frac{X_i^n}{X_l^n}$ known from the previous time level n,

$$\tilde{u}_{l_i,j}^n = \begin{cases} 1 - X_i^n e^{x_j}, & x_j < -\ln\frac{X_i^n}{X_l^n}; \\ \alpha_{l,j}^n u_{l,j_0}^n + \beta_{l,j}^n u_{l,j_0+1}^n, & -\ln\frac{X_i^n}{X_l^n} \leq x_j \leq x_{\max} - \ln\frac{X_i^n}{X_l^n}; \\ 0, & x_j > x_{\max} - \ln\frac{X_i^n}{X_l^n}. \end{cases} \tag{18}$$

Note that in the first situation of (18), $x_j < \ln \frac{X_i^n}{X_l^n}$, means that in the original variables $S < S_l^*(\tau^n)$ where the option price is payoff value. In the second case we use the linear interpolation. Finally, in the last case we assign to $\tilde{u}_{l_i,j}^n = 0$ due to condition (14).

Denoting constants

$$a_i = \frac{\sigma_i^2}{2}\frac{k}{h^2} - \left(r_i - \frac{\sigma_i^2}{2}\right)\frac{k}{2h}, \tag{19}$$

$$b_i = 1 - \sigma_i^2 \frac{k}{h^2} - (r_i - q_{ii})k, \tag{20}$$

$$c_i = \frac{\sigma_i^2}{2}\frac{k}{h^2} + \left(r_i - \frac{\sigma_i^2}{2}\right)\frac{k}{2h}, \tag{21}$$

the explicit finite difference scheme can be presented for $j = 1, .., M-1$, $i = 1, .., I$, $n = 0, .., N-1$ as follows

$$u_{i,j}^{n+1} = a_i u_{i,j-1}^n + b_i u_{i,j}^n + c_i u_{i,j+1}^n + \frac{X_i^{n+1} - X_i^n}{2hX_i^n}\left(u_{i,j+1}^n - u_{i,j-1}^n\right) + k\sum_{l\neq i} q_{il}\tilde{u}_{l_i,j}^n. \tag{22}$$

From the boundary conditions (**??**), (14) we have

$$u_{i,0}^{n+1} = 1 - X_i^{n+1}, \quad u_M^{n+1} = 0. \tag{23}$$

Boundary condition (13) can be discretized by using the second order one-side-difference approximation :

$$\frac{-3u_{i,0}^{n+1} + 4u_{i,1}^{n+1} - u_{i,2}^{n+1}}{2h} + X_i^{n+1} = 0. \tag{24}$$

The unknown optimal stopping boundary can be derived from (22), (23) and (24), $X_i^{n+1} = \frac{\xi_i^n}{\eta_i^n}$, where

$$\begin{aligned}\xi_i^n &= 3 - 4a_i u_{i,0}^n - (4b_i - a_i)u_{i,1}^n - (4c_i - b_i)u_{i,2}^n + c_i u_{i,3}^n \\ &\quad + \frac{4(u_{i,2}^n - u_{i,0}^n) - (u_{i,3}^n - u_{i,1}^n)}{2h} - k\left(4\Sigma_1 - \Sigma_2\right), \end{aligned} \tag{25}$$

$$\eta_i^n = 3 + 2h + \frac{4(u_{i,2}^n - u_{i,0}^n) - (u_{i,3}^n - u_{i,1}^n)}{2hX_i^n}, \tag{26}$$

and $\Sigma_j = \sum_{l\neq i} q_{il}\tilde{u}_{l_i,j}^n$.

4 Results and discussion

Theorem 4.1 *With previous notation the scheme (22) is conditionally stable under the constraint*

$$k \leq \min_{1\leq R\leq I} \left(\frac{h^2}{\sigma_R^2 + (r_R - q_{R,R})h^2}, \frac{2r_R}{\left(r_R - \frac{\sigma_R^2}{2}\right)^2 + (r_R - q_{R,R})\sigma_R^2} \right). \quad (27)$$

Theorem 4.2 *Assuming that the solution of the PDE problem (11)-(14) admits two times continuous partial derivative with respect to time and up to order four with respect to space, the numerical solution computed by the scheme (22) with (24) is consistent with the equation (11) and boundary condition (13) of the second order in space and the first order in time.*

As a numerical example let us consider an American Put option in 2-regime switching model with the parameters (see Example 1 in [2]):

$$\boldsymbol{r} = \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} = \begin{pmatrix} 0.1 \\ 0.05 \end{pmatrix},\ \boldsymbol{\sigma} = \begin{pmatrix} \sigma_1 \\ \sigma_2 \end{pmatrix} = \begin{pmatrix} 0.8 \\ 0.3 \end{pmatrix},\ Q = \begin{pmatrix} -6 & 6 \\ 9 & -9 \end{pmatrix},\ T = 1,\ E = 9. \quad (28)$$

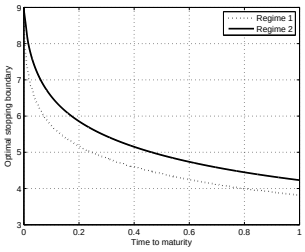


Figure 1: Optimal stopping boundary for regime 1 and regime 2 (stability condition is fulfilled).

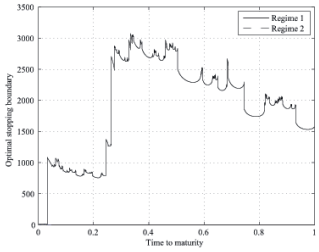


Figure 2: Optimal stopping boundary for regime 1 and regime 2 (stability condition is broken).

In order to compare the solution with penalty and lattice methods described in [2], Table 1 contains option prices for different values of asset price S computed by: our proposed front-fixing explicit method (FF-expl), the exponential time differencing Crank-Nicolson scheme (ETD-CN) and the binomial tree approach developed in [3] (Tree).

	Regime 1			Regime 2		
S	FF-expl	ETD-CN	Tree	FF-expl	ETD-CN	Tree
9.0	1.9713	1.9756	1.9722	1.8817	1.8859	1.8819
9.5	1.8049	1.8089	1.8058	1.7141	1.7181	1.7143
10.5	1.5177	1.5213	1.5186	1.4265	1.4301	1.4267
12.0	1.1796	1.1825	1.1803	1.0915	1.0945	1.0916

Table 1: Comparison of American put option prices in a two-regime model.

Acknowledgements

This work has been partially supported by the European Union in the FP7-PEOPLE-2012-ITN program under Grant Agreement Number 304617 (FP7 Marie Curie Action, Project Multi-ITN STRIKE-Novel Methods in Computational Finance) and the Ministerio de Economía y Competitividad Spanish grant MTM2013-41765-P.

References

[1] R. Company, V. N. Egorova, and L. Jódar (2014).*Solving American Option Pricing models by the front fixing method: numerical analysis and computing*. In: *Abstract and Applied Analysis*, vol. 2014, Article ID 146745, 9 pages.

[2] A.Q.M. Khaliq, B. Kleefeld and R.H. Liu (2013). *Solving complex PDE systems for pricing American options with regime-switching by efficient exponential time differencing schemes*. In: *Numerical Methods for Partial Differential Equations* 29(1), 320–336.

[3] R. H. Liu (2010). *Regime-switching recombining tree for option pricing*. In: *Int J Theor Appl Finance* 13, 479–499.

[4] L. Wu, Y.-K. Kwok, A Front-Fixing method for the Valuation of American Option. *The Journal of Financial Engineering*, 6(2) (1997) 83–97.

Positive numerical solution of two asset jump-diffusion partial-integro differential models

M. Fakharany[♭] *, R. Company[♭], and L. Jódar[♭]

(♭) Instituto de Matemática Multidisciplinar, Universitat Politècnica de València,

Camino de Vera s/n, 46022 Valencia, Spain.

November 30, 2015

1 Introduction

In this paper a finite difference scheme for the partial-integro differential equation (PIDE) for two assets under jump-diffusion model is proposed. First, in order to avoid numerical drawbacks of the mixed derivative term, a suitable transformation of variables removes this term. Second, the central finite difference approximation in space is implemented for the differential part of the new PIDE, while the integral part is approximated using a 2D Gauss-Hermite quadrature reducing the computational cost.
The stochastic differential equations for a two-dimensional Merton jump diffusion model are given by

$$\frac{dS_i(t)}{S_i(t)} = (r - q_i - \lambda\kappa_i)dt + \sigma_i dW_i + (e^{J_i} - 1)dZ(t), \quad i = 1, 2, \qquad (1)$$

where S_i, q_i, σ_i, $i = 1, 2$ are the two assets prices, asset dividend yields, asset volatilities respectively, r is the risk free interest and W_i are standard Brownian motions correlated by $\rho \in (-1, 1)$. J_1 and J_2 are the jump sizes correlated by $\rho_J \in (-1, 1)$, κ_i represent the expected relative jump sizes

*e-mail: fakharany@aucegypt.edu

($\kappa_i = E[e^{J_i} - 1]$), Z and λ are the poisson process and its jump intensity. Based on Itô calculus, the corresponding PIDE for the unknown option price $U(x_1, x_2, \tau)$ takes the form

$$\frac{\partial U}{\partial \tau} = \frac{\sigma_1^2}{2}\frac{\partial^2 U}{\partial x_1^2} + \rho\sigma_1\sigma_2\frac{\partial^2 U}{\partial x_1 \partial x_2} + \frac{\sigma_2^2}{2}\frac{\partial^2 U}{\partial x_2^2} + \left(r - q_1 - \lambda\kappa_1 - \frac{\sigma_1^2}{2}\right)\frac{\partial U}{\partial x_1}$$

$$+ \left(r - q_2 - \lambda\kappa_2 - \frac{\sigma_2^2}{2}\right)\frac{\partial U}{\partial x_2} - (r+\lambda)U + \lambda\int_{\mathbb{R}^2} U(x_1+\eta_1, x_2+\eta_2)g(\eta_1,\eta_2)d\eta_1 d\eta_2, \tag{2}$$

where $(x_1, x_2) = (\ln S_1/E, \ln S_2/E)$, E is the strike price, $\tau = T - t$ is the time to maturity and $g(\eta_1, \eta_2)$ is the probability density function of a bivariate normal distribution, given by

$$g(\eta_1,\eta_2) = \frac{\exp\left[-\frac{1}{2(1-\rho_J^2)}\left(\left(\frac{\eta_1-\mu_1}{\hat{\sigma}_1}\right)^2 - \frac{2\rho_J(\eta_1-\mu_1)(\eta_2-\mu_2)}{\hat{\sigma}_1\hat{\sigma}_2} + \left(\frac{\eta_2-\mu_2}{\hat{\sigma}_2}\right)^2\right)\right]}{2\pi\hat{\sigma}_1\hat{\sigma}_2\sqrt{1-\rho_J^2}}, \tag{3}$$

such that μ_1, μ_2, $\hat{\sigma}_1$ and $\hat{\sigma}_2$ are the means and standard deviations of the jumps J_1 and J_2 respectively. There are several kinds of two assets problem depending on the nature of the payoff. Here we focus on options of the best and worst types. In this case, the option gives a holder the right to receive the maximum or minimum of the two underlying assets at maturity. The payoff $f(x_1, x_2)$ for put on minimum [1, 3] and the boundary conditions are given by

$$f(x_1, x_2) = E\max(1 - \min(e^{x_1}, e^{x_2}), 0) \tag{4}$$

$$\begin{gathered}\lim_{x_1\to-\infty} U(x_1, x_2, \tau) = Ee^{-r\tau}, \;\; \lim_{x_2\to-\infty} U(x_1, x_2, \tau) = Ee^{-r\tau}, \\ \lim_{x_1\to\infty} \partial_{x_1} U(x_1, x_2, \tau) = 0, \;\; \lim_{x_2\to\infty} \partial_{x_2} U(x_1, x_2, \tau) = 0.\end{gathered} \tag{5}$$

2 The problem transformation and numerical scheme

From numerical analysis view, the presence of the mixed derivative in the PIDE leads to several problems in the numerical scheme performance such as delaying the convergence and poor accuracy. Moreover, the discretization of this term requires nine or seven constrained stencil points which increases the computational cost. Here this term is removed using the characteristic

equation as in [4, 5] since $\Delta = \sigma_1^2\sigma_2^2(\rho^2-1) < 0$. Consequently, the PIDE (2) is of elliptic type, then by using the canonical transformation [6], the suitable change or variables are given by

$$y_1 = \frac{\sigma_2\tilde{\rho}}{\sigma_1}x_1,\ \ y_2 = x_2 - \frac{\sigma_2\rho}{\sigma_1}x_1 = x_2 - \tilde{m}y_1,\ \tilde{m} = \frac{\rho}{\tilde{\rho}}, V(y_1,y_2,\tau) = \exp((r+\lambda)\tau)U(x_1,x_2,\tau), \tag{6}$$

and (2) is transformed into

$$\begin{array}{c} \frac{\partial V}{\partial \tau} = \frac{\sigma_2^2\tilde{\rho}^2}{2}\left(\frac{\partial^2 V}{\partial y_1^2} + \frac{\partial^2 V}{\partial y_2^2}\right) + a_1\frac{\partial V}{\partial y_1} + a_2\frac{\partial V}{\partial y_2} \\ + \frac{\sigma_1\lambda}{\sigma_2\tilde{\rho}}\int_{\mathbb{R}^2} V(\phi_1,\phi_2,\tau)g(\frac{\sigma_1}{\sigma_2\tilde{\rho}}(\phi_1-y_1), \phi_2 - y_2 + \tilde{m}(\phi_1-y_1))d\phi_1 d\phi_2, \end{array} \tag{7}$$

where

$$\begin{array}{l} a_1 = \frac{\tilde{\rho}\sigma_2}{\sigma_1}(r - q_1 - \lambda k_1 - \frac{\sigma_1^2}{2}), \\ a_2 = \left((1-\frac{\rho\sigma_2}{\sigma_1})r - (q_2 - \frac{\rho\sigma_2}{\sigma_1}q_1) - \lambda\kappa_2 + \frac{\rho\sigma_2}{\sigma_1}\lambda\kappa_1 - \frac{\sigma_2^2}{2} + \frac{\rho\sigma_1\sigma_2}{2}\right), \\ \phi_1 = y_1 + \frac{\sigma_2\tilde{\rho}}{\sigma_1}\eta_1,\ \ \phi_2 = y_2 - \frac{\sigma_2\rho}{\sigma_1}\eta_1 + \eta_2. \end{array} \tag{8}$$

For the associated boundary conditions for Put option on minimum, one gets

$$\begin{array}{c} \lim_{y_1\to-\infty} V(y_1,y_2,\tau) = Ee^{\lambda\tau},\ \ \lim_{y_2\to-\infty} V(y_1,y_2,\tau) = Ee^{\lambda\tau} \\ \lim_{y_1\to\infty} \partial_{y_1} V(y_1,y_2,\tau) = 0,\ \ \lim_{y_2\to\infty} \partial_{y_2} V(y_1,y_2,\tau) = 0. \end{array} \tag{9}$$

Now we ready to set up an efficient explicit numerical scheme for the problem (7)-(9). First, let us consider a rectangular domain in x_1x_2-plane with boundaries $x_1 \in [a,b]$ and $x_2 \in [c,d]$. Under the transformation (6), the rectangular domain is converted to a rhomboid domain in y_1y_2-plane. Let N_1 and N_2 represent the number of the discretization points in y_1 and y_2 directions, $h_{x_1} = (b-a)/N_1$ and $h_{x_2} = (d-c)/N_2$ are the stepsizes in x_1x_2-plane, h_1 and h_2 are the stepsizes in y_1y_2-plane which are obtained by h_{x_1} and h_{x_2} such that $h_1 = \frac{\sigma_2}{\sigma_1}\tilde{\rho}h_{x_1}$, $h_2 = h_{x_2}$. Hence $y_{1,i} = y_{1,0} + ih_1$, $0 \le i \le N_1$, $y_{1,0} = \frac{\sigma_2}{\sigma_1}\tilde{\rho}a$, we discretize y_2 using two indices i and j; $y_{2,j}^i = \hat{y}_{i,0} + jh_2$ where $\hat{y}_{i,0} = c - \frac{\sigma_2}{\sigma_1}\rho(a+ih_{x_1})$, $0 \le i \le N_1$, since we have oblique lines and $\tau^n = nk$, $0 \le n \le N_\tau$, $k = T/N_\tau$. Let $V_{i,j}^n$ be the approximation of $V(y_{1,i}, y_{2,j}^i, \tau^n)$, the central finite difference approximations are used for the first and second spatial derivatives while the explicit forward discretization is implemented to approximate the first time-derivative of V. Finally, the integral part is approximated using 2D Gauss-Hermite quadrature formula

[7], then the corresponding finite difference scheme is given by

$$V_{i,j}^{n+1} = \alpha_1 V_{i,j}^n + \alpha_2 V_{i+1,j}^n + \alpha_3 V_{i-1,j}^n + \alpha_4 V_{i,j+1}^n + \alpha_5 V_{i,j-1}^n + \frac{k\lambda\sigma_1}{\sigma_2\tilde{\rho}} I_{i,j}^n, \quad (10)$$

where

$$\alpha_1 = 1 - k\sigma_2^2\tilde{\rho}^2(\tfrac{1}{h_1^2} + \tfrac{1}{h_2^2}), \quad \alpha_2 = \tfrac{k}{2h_1}(\tfrac{\sigma_2^2\tilde{\rho}^2}{h_1} + a_1), \quad \alpha_3 = \tfrac{k}{2h_1}(\tfrac{\sigma_2^2\tilde{\rho}^2}{h_1} - a_1),$$
$$\alpha_4 = \tfrac{k}{2h_2}(\tfrac{\sigma_2^2\tilde{\rho}^2}{h_2} + a_2), \quad \alpha_5 = \tfrac{k}{2h_2}(\tfrac{\sigma_2^2\tilde{\rho}^2}{h_2} - a_2), \quad (11)$$

$$I_{i,j}^n = \sum_{\ell=1}^{L}\sum_{m=1}^{M} \omega_\ell\omega_m g(\frac{\sigma_1}{\sigma_2\tilde{\rho}}(\phi_{1,\ell}-y_{1,i}), \phi_{2,m}-y_{2,j}^i+\tilde{m}(\phi_{1,\ell}-y_{1,i}))\exp[\phi_{1,\ell}^2+\phi_{2,m}^2]V^n(\phi_{1,\ell},\phi_{2,m}), \quad (12)$$

such that ϕ_m and ϕ_ℓ are the roots of Hermite polynomials of degrees M and L respectively. ω_m and ω_ℓ are the corresponding weighting values.
Since $(\phi_{1,\ell}, \phi_{2,m})$ are not necessary points of the grid, then in order to match the discretization of the integral part with the differential part, the bivariate interpolation four point approximation [7] is used. Hence the approximation for $V_{\ell,m}^n$ is given by

$$V_{\ell,m}^n \approx \hat{\delta}_{i_\ell,2}(\delta_{j_m,2}V_{i_\ell,j_m}^n + \delta_{j_m,1}V_{i_\ell,j_m+1}^n) + \hat{\delta}_{i_\ell,1}(\delta_{j_m,3}V_{i_\ell+1,j_m+1}^n + \delta_{j_m,4}V_{i_\ell+1,j_m}^n), \quad (13)$$

where

$$\hat{\delta}_{i_\ell,1} = \tfrac{\phi_{1,\ell}-y_{1,i_\ell}}{h_1}, \quad \hat{\delta}_{i_\ell,2} = \tfrac{y_{1,i_\ell+1}-\phi_{1,\ell}}{h_1}, \quad \delta_{j_m,1} = \tfrac{\phi_{2,m}-y_{2,j_m}^{i_\ell}}{h_2},$$
$$\delta_{j_m,2} = \tfrac{y_{2,j_m+1}^{i_\ell}-\phi_{2,m}}{h_2}, \quad \delta_{j_m,3} = \tfrac{\phi_{2,m}-y_{2,j_m}^{i_\ell+1}}{h_2}, \quad \delta_{j_m,4} = \tfrac{y_{2,j_m+1}^{i_\ell+1}-\phi_{2,m}}{h_2}.$$

Finally the discretization for initial and boundary conditions are

$$V_{i,j}^0 = f(y_{1,i}, y_{2,j}^i),\ 0 \le i \le N_1,\ 0 \le j \le N_2,$$
$$V_{0,j}^n = V_{i,0}^n = Ee^{\lambda\tau^n},\ 0 \le i \le N_1,\ 0 \le j \le N_2,\ n \ge 1,$$
$$V_{N_1,j+1}^n = V_{N_1,j}^n,\ 0 \le j \le N_2 - 1,\ V_{i+1,N_2}^n = V_{i,N_2}^n,\ 0 \le i \le N_1 - 1.$$

3 Example

In order to show the efficiency of the proposed scheme let us consider the an European put option with parameters $T = 1$, $E = 100$, $r = 0.05$, $q_1 = q_2 = 0$, $\sigma_1 = 0.12$, $\sigma_2 = 0.15$, $\rho = 0.3$, $\lambda = 0.6$, $\mu_1 = -0.1$,

		$\hat{S}_1$			$\hat{S}_2$			$\hat{S}_3$		
	(N_1, N_2, N_τ)	**RMSRE**	Ratio	CPU (sec)	**RMSRE**	Ratio	CPU (sec)	**RMSRE**	Ratio	CPU (sec)
$L = M = 3$	(60,30,50)	1.993e-2	–	0.15	1.689e-2	–	0.15	2.263e-2	–	0.15
	(120,60,100)	9.271e-3	2.15	1.32	9.108e-3	1.85	1.32	9.464e-3	2.39	1.32
	(240,120,200)	1.965e-3	4.72	6.52	1.944e-3	4.68	6.52	1.829e-3	5.17	6.52
	(480,240,400)	6.174e-4	3.18	10.85	5.979e-4	3.25	10.85	5.569e-4	3.28	10.85
$L = M = 5$	(60,30,50)	8.376e-3	–	0.27	7.758e-3	–	0.27	1.054e-3	–	0.27
	(120,60,100)	3.006e-3	2.78	1.86	3.165e-3	2.45	1.86	3.386e-3	3.11	1.86
	(240,120,200)	6.199e-4	4.85	7.15	6.752e-4	4.69	7.15	6.422e-4	5.27	7.15
	(480,240,400)	2.402e-4	2.58	12.54	2.335e-4	2.89	12.54	2.609e-4	2.46	12.54

Table 1: The **RMSRE** for European put option on the minimum of two assets for several grids.

$\mu_2 = 0.1$, $\hat{\sigma}_1 = 0.17$, $\hat{\sigma}_2 = 0.13$, $\rho_J = -0.2$ and the boundaries x_1x_2-plane are $x_1, x_2 \in [-3, 3]$. The root mean square relative error (**RMSRE**) for $S_1, S_2 = \{90, 100, 110\}$ is calculated for $L = M = 3$ and 5. The reference values are in [1], here the **RMSRE** is obtained for three groups $\hat{S}_1 = \{(90, 90), (90, 100), (90, 110)\}$, $\hat{S}_2 = \{(100, 90), (100, 100), (100, 110)\}$ and $\hat{S}_3 = \{(110, 90), (110, 100), (110, 110)\}$, Table 1 reports the associated **RMSRE**, ratio and CPU time for several grids.

It is easy to show that the positivity of the scheme is guaranteed under the following conditions:

$$k < \frac{h_1^2 h_2^2}{\sigma_2^2 \tilde{\rho}^2 (h_1^2 + h_2^2)}, \quad h_1 < \frac{\sigma_2^2 \tilde{\rho}^2}{|a_1|}, \quad h_2 < \frac{\sigma_2^2 \tilde{\rho}^2}{|a_2|}. \tag{14}$$

Acknowledgements

This work has been partially supported by the European Union in the FP7-PEOPLE-2012-ITN program under Grant Agreement Number 304617 (FP7 Marie Curie Action, Project Multi-ITN STRIKE-Novel Methods in Computational Finance).

References

[1] Clift S. S., and Forsyth P. A. Numerical solution of two asset jump diffusion models for option valuation *Applied Mathematical Mathematics*, Volume(58): 743–782, 2008.

[2] Rambeerich N., Tangman D. Y., Lollchund M.R., and Bhuruth M. High-order computational methods for option valuation under multifactor models *European Journal of Operational Research*, Volume(224): 219–226,2013.

[3] D. J. Duffy, Finite Difference Methods in Financial Engineering: a Partial Differential Approach. England, John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, 2006.

[4] Company R., Jódar L., Fakharany M., and Casabán M.-C. Removing the correlation term in the option pricing Heston model: numerical analysis and computing *Abstract and Applied Analysis*, Volume (2013): 1–11.

[5] Fakharany M., Company R., and Jódar L. Positive finite difference schemes for partial integro-differential option pricing model *Applied Mathematics and Computation*, Volume(249): 320–332.

[6] S. J. Farlow, Partial differential equations for scientists and engineers. New York, Dover Publications Inc., 1993.

[7] M. Abramowitz and I. A. Stegun. Handbook of mathematical functions: with formulas, graphs, and mathematical tables. Dover Books on Mathematics, 1961.

Factors Affecting the Choice Modal of Transportation in an Urban Environment. Hierarchical Multi-Criteria Analysis

A. Fraile[♭] [*], J.A. Sicilia [§], E. Larrodé[♭] and B. Royo[†]

(♭) Department of Mechanical Engineering, Universidad de Zaragoza, Spain.

(§) Universidad Internacional de La Rioja (UNIR), Logroño, La Rioja, Spain.

(†) Instituto Tecnológico de Aragón, Zaragoza, Spain.

November 30, 2015

1 Introduction

In recent decades there have been profound social, economic and technological changes have led to a new model of urban mobility. That model, which tends to be implemented globally, is characterized by the increase in average distances travelled, changes in the grounds for the displacements and changes in the location of production activities. In modern economies is essential that suitable transport that enables the population mobility and the consequent accessibility to services. However, its current configuration is causing strong negative externalities and generates much of the problems of environmental, social and energy sustainability. The changes required by the current unsustainable configuration should be based on a new vision of transportation and the development of theoretical and empirical criteria that enable the creation of sustainable transport systems. Such a system must be based on the principle of multimodal transport.

The purpose of this paper is to analyze the most important factors of urban mobility in its current configuration, analyze the negative externalities

[*]e-mail: afrailep@unizar.es; juanantonio.sicilia@unir.net

caused by the transportation, and develop the principles necessary to develop a model in which the modal distribution is optimized in a city in a way that is sustainable, both in terms of passengers and freight.

Previously, in other works as [2, 3, 4] using different methodologies, the factors affecting modal choice in different countries are analyzed, success stories in sustainable mobility are described and traffic optimization indicators are described.

To do this, in the Decision Theory, it has been selected Analytic Hierarchy Process (AHP) technique to support different stages of the proposed methodology. This technique allows by building a hierarchical model, efficiently and graphically, organizes information about a problem, break it down and analyze it by parts in different matrix, visualize the effects of changes in levels and synthesize.

2 Objectives

The main objectives sought in this article are:

- Analyze key factors of urban mobility in the current situation of cities, taking into account all the elements involved in mobility in urban environments (technological advances, new regulations, influence the economic attitude, changes in urban planning,...) and study their behavior.

- From here, develop the principles necessary to develop a model to aid decision-making for action, in order to meet three overall objectives, customer satisfaction, sustainable growth and energy and environmental efficiency.

Similarly, secondary objectives are:

- Develop a model of analysis to study mobility in an urban environment taking into account the influence of factors affecting transport for best modal split.

- Justify, within the theory of decision, the selection of AHP technique to support various stages of the proposed methodology; Building the analysis model, in collaboration with a group of experts in the field; Define all elements of the model (criteria, sub-criteria, attributes and

alternatives), and develop the AHP. Also, an analysis of sensitivity and consistency of the model.

3 Methodology

The proposed methodology is based on the AHP, method appearing through Professor Thomas L. Saaty, 1977, in the Journal of Mathematical Psychology [5]. The AHP is a method of organizing information and reasoning used in making decisions. The AHP contributes to solving complex problems structuring a hierarchy of criteria, stakes and results, extracting trials to develop priorities. In short, the AHP is a general approach to defining problems, set priorities and make decisions. The AHP involves all aspects of the decision-making process, and modeling the problem through a hierarchical structure and uses a priority scale to synthesize and deliver judgments ordering or ranking of the alternatives according to the weights obtained (priorities).

To promote consistency of judgments Saaty proposes the fundamental scale that bears his name [5]. This scale is structured according to the intensity of importance of some judgments about others with the values 1, 3, 5, 7 or 9, from low to high importance.

To determine the best decision, in a generic way, the AHP method requires follow these steps: define the problem; choose the actors; structuring the decision problem in order to build a model of hierarchy; select the feasible alternatives; build the hierarchical model; login judgments; summary of results; and validation of the decision. Therefore, the system of the AHP of the establishment of a goal, criteria and alternative structure, as you can see in Figure 1.

The priority of the alternatives with respect to the goal, which was the objective sought is obtained according to the objective function:

$$W_{A_i}^{MI} = \sum_{j=1}^{M} \sum_{k=1}^{M} W_{A_i}^{C_{jk}} W_{C_{jk}}^{C_j} W_{C_j}^{MI} \tag{1}$$

To learn how priority values are obtained from the assessments carried out by the actors in the AHP, it must repeat the following three steps as many times as there are elements influenced in the hierarchy:

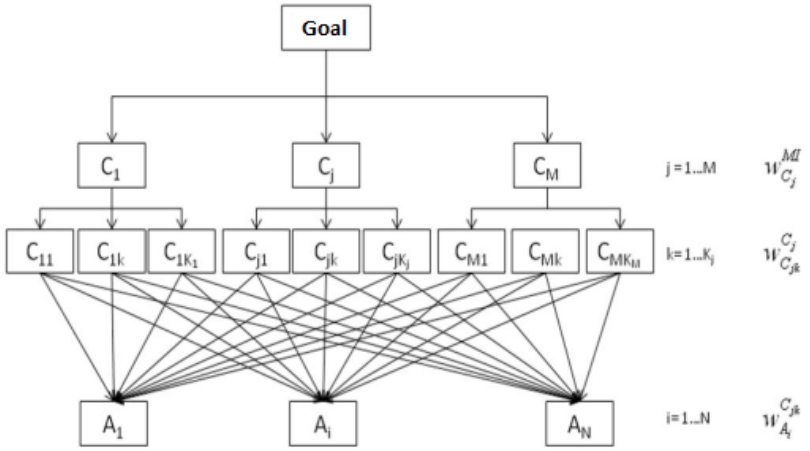


Figure 1: Structure of AHP (goal, criteria and alternatives)

1) It is building a Reciprocal Matrix Binary Comparisons $[A]$ (of order n, which equals the number of elements that are assessing their influence) from valuations $a_{ij}(i = 1, ..., n; j = 1, ..., n)$ made by the evaluator actor and / or decision maker. For this, each a_{ij} corresponds to a numeric value of the fundamental Saaty scale.

$$[A] = \begin{vmatrix} 1 & a_{12} & ... & a_{1n} \\ \frac{1}{a_{12}} & 1 & ... & a_{2n} \\ ... & ... & ... & ... \\ \frac{1}{a_{1n}} & \frac{1}{a_{2n}} & ... & 1 \end{vmatrix}$$

2) The next step is the calculation of local priorities. The mathematical procedure is followed in obtaining the principal eigenvector method to the right [5]. This method, based on the Perron-Frobenius theorem provides local priorities solving the system of equations:

$$AW = \lambda_{max} W \tag{2}$$

where $A = (a_{ij})$ is the Reciprocal Matrix Binary Comparisons, λ_{max} the principal eigenvalue of A, y $W = (W_1, W_2, \ldots, W_n)$ the vector of local priorities ratio scale measures and normalized to have unity. In this

case, normalization has applied the so called distributive ($\sum_j W_j = 1$). In practice, the solution W is obtained (power method) by raising the matrix judgments on a sufficiently large power, adding rows and normalizing these values by dividing the sum of each row for the total amount. The process ends when the difference between two consecutive powers is small. The priority vector w represents the relative importance of the criteria or sub compared in each pairwise comparison matrices.

3) For the analysis of consistency, it is estimated the maximum eigenvalue associated λ_{max} with $[A]$ and the consistency index (CI) is obtained:

$$CI = \frac{\lambda_{max} - n}{n - 1} \tag{3}$$

where n is the orden of the matrix.

If the reason for inconsistency (CR), where the random index (RI) is obtained from Table 1, is less than 10% then states that the valuations are consistent and the values $W_i (i = 1, ..., n)$ of the eigenvector W are accepted as good. If not, the decision maker is asked to redo the valuations of the Reciprocal Matrix Binary Comparisons:

$$CR = \frac{CI}{RI} \tag{4}$$

Matrix size	1	2	3	4	5	6	7	8	9	10
RI	0	0	0.525	0.882	1.115	1.252	1.341	1.404	1.452	1.484

Table 1: Table of random values (Saaty) [1]

Finally in this paper, the AHP model structure is composed of the goal, 2 criterias (C_i), 13 sub-criterias (SC_{ij}), 24 attributes (Q_i) y 6 alternatives A_i.

- Goal: Efficiency in the modal split in urban environments.
- C_1: Actors; C_2: Factors.

- SC_{11}: Users; SC_{12}: Regulator and administrator; SC_{13}: Urban population centres of attraction; SC_{14}: Suppliers of technological innovation; SC_{15}: Business mobility and distribution services; SC_{21}: Accessibility; SC_{22}: Quality of service; SC_{23}: Comfort;SC_{24} Efficiency; SC_{25}: Security; SC_{26}: Technology; SC_{27}: Environmental impact; SC_{28}: Organization of the urban environment.

- Q_1: Design of Access roads suitable to use; Q_2: Good organization of circulation; Q_3: Existence of stops proportion to population density; Q_4: Location appropriate car park; Q_5: Balanced distribution of loading and unloading areas; Q_6: Sufficient capacity; Q_7: Low journey; Q_8: High frequency; Q_9: High rate flexibility; Q_{10}: Comfort; Q_{11}: Good computer interactive service; Q_{12}: Low path length; Q_{13}: Difficult terrain; Q_{14}: Bad weather; Q_{15}: Low costs in use; Q_{16}: Low energy consumption; Q_{17}: Low accident; Q_{18}: Adequate traffic control; Q_{19}: Technology vehicular updated; Q_{20}: Existence of information and communications technology; Q_{21}: Low emissions; Q_{22}: Complete rules-use regulation; Q_{23}: Existence of mobility plan; Q_{24}: Complex morphology.

- A_1: Pedestrian mobility; A_2: Transportation by private vehicle; A_3: Collective transport; A_4: Goods transport vehicle; A_5: Mobility nonmotorized vehicle; A_6: New modes of mobility.

In short, it structured decision making from the most general to the most specific thanks to the knowledge of the expert group. These experts have also been responsible for analyzing the model through 505 judgments. In this phase of the methodology incorporates judgments based on the information obtained or the perception of the actors in the process. Therefore, once it has well defined and structured hierarchy, the operations described continues after systematic: the reciprocal matrix of binary relationships is obtained, the reason inconsistency and, if this is less than 10%, eigenvector accept the right of the Reciprocal Matrix Binary Comparisons as a measure of local priority.

4 Conclusions

Progress has been made in the use of techniques covered in the Decision Theory for modeling complex systems. In particular, it discussed a new

approach of using the AHP which allows taking into account the influence of factors interacting in deciding which mode of transport used for each shift and achieve the objective of meeting the needs of the user, sustainable growth of the city and improving the energy and environmental efficiency, which results in a certain modal split. This way you can have a holistic view of behavior in urban environments.

There are few jobs that advance the use of this technique of analysis for large cases such as transport problems and less focusing the analysis model from a global perspective in which agents interact with each other as is done in this project, unlike the widespread use of AHP to solve decision-making problems.

It has been proposed a method that through the results of the application of AHP allows assessment of the modal split in different settings, and can serve as a comparison for future decision-making, ie, has established a framework of comparison.

References

[1] Aguarón, J., Moreno-Jiménez, J.M. The geometric consistency index: Approximated thresholds. *European Journal of Operational Research.*, Volume 147 (1): 137–145, 2003.

[2] Buehler, R. Determinants of transport mode choice: a comparison of Germany and the USA. *Journal of Transport Geography*, 2011.

[3] Litman, T. Sustainable transportation indicators. A recommended research program for developing sustainable transportation indicators and data. *Transportation Research Board Annual Meeting*, 2009.

[4] Murphy, E. Excess commuting and modal choice. *Transportation research - Part A. Policy and Practice*, Volume 43 (8): 735-743, 2009.

[5] Saaty, T.L. The Analytic Hierarchy Process. New York: McGraw-Hill, 1980.

[6] Saaty, T.L. The legitimacy of rank reversal. ,*Omega*, Volume 12: 513-516, 1984.

Managing dependence in Flowgraphs models. An application to Reliability Engineering

C. Santamaría[†] *, B. García–Mora[†], G. Rubio[†] and R. Pérez–Ocón[‡]

(†) Instituto de Matemática Multidisciplinar,

Universitat Politècnica de València.

(‡) Departamento de Estadística e Investigación Operativa,

Universidad de Granada.

November 30, 2015

1 Introduction

Multi-state stochastic processes are a convenient framework for modeling reliability problems in engineering and the statistical flowgraph [1] approach is an eficient tool to perform the task. Flowgraphs methodology is specifically suited for semi-Markov processes. In fact it was originally developed to model total system waiting times for semi-Markov processes. The semi-Markov assumption implies independence among waiting time distributions. However, there are situations that need relax this assumption. Within the framework of the flowgraph methodology dependency management is not achieved with generality, only in a few special cases. A successful approach is required when the conditional independence assumption for waiting times does not hold. The Markovian Arrival Process (MAP) has the relevant property of dependence between consecutive interarrival times in a process with multiple events, what constitutes it an effective tool for modeling the dependence in multi-state processes. In the following sections we develop the approach, focusing on a device that undergoes two types of failure.

*e-mail: crisanna@imm.upv.es

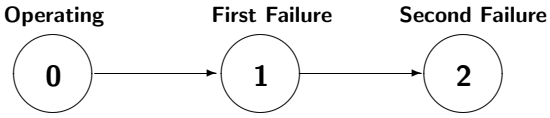


Figure 1: Three states Flowgraph Model: failures in the two component.

2 The Mathematical Flowgraph Models

The Flowgraph Model represents a multistate model with directed line segments or *transitions* (branches) connecting the *states* (nodes). Figure 1 represents a three states flowgraph model of a system with two components where the state 0 is the beginning of the process (the system is operating), state1 represents the first failure of the first component and the state 2 represents the failure of the second component fails (the system is off). The step-by-step process of setting up a Flowgraph Model and solving for quantities of interest consists of:

1. *Establishing states graph in the model, which are connected by* branches *or transitions*. Every transition $i \to j$ has a transition probability p_{ij} and a waiting time distribution $F_{ij}(x)$, the cumulative distribution function of the time x spent in state i, given that a transition to j occurs.

2. *Decide on a probability model family $F_{ij}(x)$ in each branch $i \to j$.* In this case we use a linear combination of three Erlang distributions [2], that they are a particular case of a type–phase distribution [3].This combination will be represented by

$$F_{ij}(x) = p_{1ij}G_{1ij}(x) + p_{2ij}G_{2ij}(x) + p_{3ij}G_{3ij}(x)$$

   where $G_{1ij}(t)$, $G_{2ij}(t)$, $G_{3ij}(t)$ are the three Erlang distributions and $p_{1ij} + p_{2ij} + p_{3ij} = 1$, $p_{mij} > 0$, $m = 1, 2, 3$ in each transition. Then the representation of $F_{ij}(t)$ as a phase-type distribution is (α, T) with α an initial probability vector and T a matrix of transitions inside of a generator matrix Q.

3. *Characterizing the transition $i \to j$ by means of a* transmittance. This consists of the product of p_{ij} and an integral transform of $F_{ij}(x)$, that

is $p_{ij}\,\mathcal{L}(F_{ij})$. This integral transform is the Laplace transform of a phase–type distribution with expression

$$\mathcal{L}_{ij} = \mathcal{L}(F_{ij}) = \alpha_{m+1} + \alpha(sI - T)^{-1}T^0, \text{ for } Re(s) > 0$$

where T is the transition matrix of the mixture of three Erlangs distribution.

4. *Obtention of the Laplace transform between the two states of interest.* Having calculated the *transmittances* $p_{ij}\mathcal{L}_{ij}$ for each transition $i \to j$, we have to reduce the flowgraph model to a single transmittance between the states 0 and 2 in the Figure 1. For it we use the Manson's rules [4] where the transmittance of transitions in series $(0 \longrightarrow 1 \longrightarrow 2)$ is the product of the series transmittances.

$$\mathcal{L}_{02} * (s) = p_{01}p_{12}\mathcal{L}_{01}(s)\mathcal{L}_{12}(s)$$

5. *Invert the Laplace transformation* $\mathcal{L}_{02} * (s)$ in order to obtain the distribution function (CDF) of the state of interest: *state 2*. We use an inversion algorithm called EULER [5].

3 Markovian Arrival Process

Within the framework of the Flowgraph models, the semi-Markov assumption implies independence between the two interarrival times of each transition $i \to j$ and we need to consider dependence between the two transitions. In this case Phase-type distributions can be extended to describe correlated inter-event times and the resulting models are the *Markovian Arrival Processes (MAPs)* [3].

A Markovian Arrival Process (π, D_0, D_1) is an irreducible Markov chain with a finite state space S, an initial vector π and a generator matrix Q which can be represented as $Q = D_0 + D_1$ where,

- $D_1 \geq 0$, $D_1 \neq 0$.
- $D_0(i, j) \geq 0$ for $i \neq j$.
- (π, D_0) is a phase-type distribution.

The joint density function of a MAP generating k consecutive events with inter-event times x_i, $i = 1, 2, \ldots, k$ is given by

$$f(x_1, x_2, \ldots, x_k) = \pi e^{D_0 x_1} D_1 e^{D_0 x_2} D_1 \ldots e^{D_0 x_k} D_1 \rho$$

where ρ represents a column matrix of ones of an appropriate dimension.

The idea is to perform a MAP (π, D_0, D_1) for modeling the transition of interest $0 \to 1 \to 2$ using the joint density function for the two dependent interarrival times x_1 and x_2

$$f(x_1, x_2) = \pi e^{D_0 x_1} D_1 e^{D_0 x_2} D_1 \rho$$

where

- (π, D_0) is a phase-type distribution. To build the matrix D_0 of a phase–type distribution we use the following theorem [3]:

  Theorem 3.1 *If $F_1(\cdot)$ and $F_2(\cdot)$ are both continuous PH-distributions with representations (α, T) and (β, S) of orders m and n respectively, then their convolution $F(\cdot) = (F_1 * F_2)(\cdot)$ is a PH-distribution with representation (γ, L), given by*

  $$\gamma = (\alpha, \alpha_{m+1}\beta) \quad \text{and} \quad L = \left[\begin{array}{c|c} T & T^0\beta \\ \hline 0 & S \end{array}\right]$$

  In our approach D_0 is L obtained from the convolution of the functions F_{01} and F_{12}, each one of them are the resulting mixtures of three Erlangs in each transition shown in the previous section. The probability vector π was set heuristically.

- D_1 represents a correction of D_0. The matrix D_1 with free parameters is fitted by maximizing the likelihood function for censored data [6] in the transitions $0 \to 1 \to 2$. The censored data likelihood at state j is

  $$\mathcal{L}_j(\beta_j|\mathcal{D}_j) = \left[\prod_{k=1}^{K_j} \prod_{i \in O_{jk}} \prod_{h=1}^{n_{jk,\,i}} p_{jk,\,ih}\, f_{jk}(t_{jk,\,ih}; \beta_{jk,\,ih})\right] \times \left[\prod_{i \in O_{jk}} (1 - F_j(t^*_{j,\,i}))\right]$$

  where $t^*_{j,\,i}$ is the observed censoring time in state j for observation i and F_j is the cumulative distribution function (CDF) corresponding to the mixture density $f_j(t^*_{j,\,i}) = \sum_{k=1}^{K_j} p_{jk,\,ih} f_{jk}(t^*_{j,\,i}; \beta_{jk,\,ih})$.

4 Aplication system of two components

We performed a simulation with dependent interarrival times (x_2 depends onx_1) for the system with two components represented in the Figure 1. Firstly the data are fitted by means of a Flowgraph approach and secondly we use the Markovian Arrival Process approach. In the Figure 2 we can conclude that in case of dependent interarrival times the Markovian Arrival Process is required.

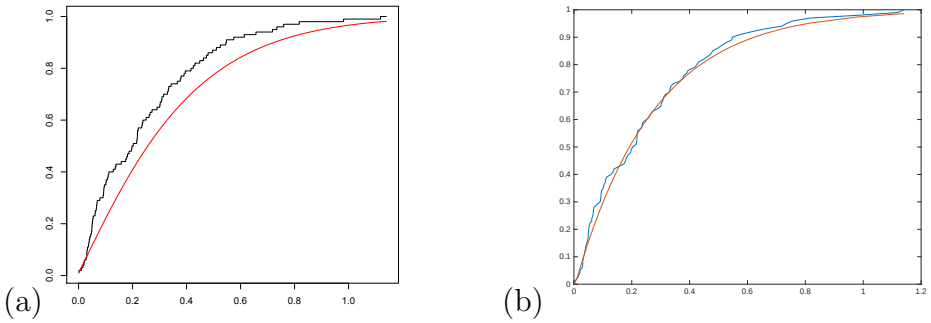


Figure 2: CDF function model (*smooth line*) and empirical survival function (*step function*) in the transition 02 for: (a) standard flowgraph with dependence and (b) Markovian Arrival Process.

References

[1] Huzurbazar, A. Flowgraph Models for Multistate Time-To-Event Data. New York, Wiley, 2005.

[2] Pérez–Ocón, R. Modeling lifetimes using phase-type distributions, in: Risk, Reliability and Societal Safety. *Proceedings of the European Safety and Reliability Conference, ESREL*, 1–5, 2007.

[3] Neuts M.F. Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach. John Hopkins University Press, 1981.

[4] Mason, S.J. Feedback theory–some properties of signal flowgraphs. *Proc IRE*, (44):920–926, 1956.

[5] Collins, D.H. and Huzurbazar, A.V. Prognostic models based on statistical Flowgraphs. *Appl. Stochastic Models Bus. Ind.*, (28):141–151, 2012.

[6] Huzurbazar, A. and Williams, B. Incorporating Covariates in Flowgraph Models: Applications to Recurrent Event Data. *Technometrics*, 52(2):198-208, 2010.

Hybrid wavelet support vector machine and artificial bee colony for predicting the cyanotoxin content from experimental cyanobacteria concentrations in the Trasona reservoir: A case study in Northern Spain

P.J. García Nieto[♭] [*], E. García-Gonzalo[♭],
J.R. Alonso Fernández[†], C. Díaz Muñiz[†]

(♭) Department of Mathematics, University of Oviedo,
Faculty of Sciences, C/ Calvo Sotelo s/n, 33007 Oviedo, Spain.

(†) Cantabrian Basin Authority, Spanish Ministry of Agriculture, Food and Environment,
Plaza de España 2, 33007 Oviedo, Spain.

November 30, 2015

Abstract

This study presents a novel hybrid algorithm, support vector machines with *Mexican hat wavelet kernel function* (wavelet SVMs) in combination with the *artificial bee colony* (ABC) technique, for predicting the cyanotoxin content from cyanobacterial concentrations determined experimentally in the Trasona reservoir. This optimization technique involves kernel parameter setting in the SVM training procedure, which significantly influences the regression accuracy. The agreement of the wavelet ABC–SVM–based model with experimental data confirmed its good performance. Indeed, a coefficient of determination equal to 0.91 was obtained. Finally, conclusions of this innovative research work are exposed.

[*]e-mail: lato@orion.ciencias.uniovi.es

Keywords Support vector machines (SVMs); Wavelet kernel; Artificial bee colony (ABC); Cyanotoxins; Regression analysis

1 Introduction

Cyanobacteria, also known as blue-green algae, is a phylum of bacteria that obtain their energy through photosynthesis. Indeed, cyanobacteria are aquatic and photosynthetic, that is, they live in the water, and can manufacture their own food. Sometimes, they become hazardous due to their uncontrolled growth giving place to the formation of extensive harmful algal blooms (HABs) [1]. In addition, some cyanobacteria produce toxins called *cyanotoxins* [1]. In this way, the association of toxicity with such blooms has frequently led to the closure of recreational waters when blooms are observed.

To fix ideas, the objective of this study is to evaluate the application of wavelet kernel support vector machines (SVMs) [2] in combination with the Artificial Bee Colony (ABC) technique [3] to identify cyanotoxins in the Trasona reservoir (Principality of Asturias, Northern Spain). In recent years, the combination of wavelet theories and SVMs has drawn considerable attention owing to its high predictive ability for a wide range of applications and better performance than other traditional learning machines. In order to carry out the optimization mechanism corresponding to the kernel optimal hyperparameters setting in the SVM training, the artificial bee colony (ABC) technique was used here with success. The artificial bee colony (ABC) technique [3] is an optimization algorithm based on the intelligent foraging behaviour of honey bee swarm.

2 Materials and methods

2.1 Experimental dataset

The dataset used for the SVM analysis were collected over 5 years (2006–2010) from several samples in Trasona reservoir. The total number of data processed was about 151 values (see Appendix A). Furthermore, we have taken into account the two dominant species of the cyanobacteria community in this research work: *M. aeruginosa* and *W. naegeliana*. It should be noted that the main goal of this study is to obtain the dependence relationship of cyanotoxins (output variable), expressed in micrograms per liter, as a

function of the following eight biological and fifteen physical–chemical input variables (see Appendix A).

2.2 Support vector machine (SVM) method

In machine learning, support vector machines (SVMs) are a set of related supervised learning methods used for classification and regression [4]. The SVMs were originally developed for classification, and were later generalized to solve regression problems. This last method is called *support vector regression* (SVR). The model produced by SVR depends on a subset of the training data, because the cost function for building the model ignores any training data that are close (within a threshold ε) to the model prediction. When the regression SVM is applied to non–linear separable data, it is necessary to use the *kernel trick* . The reason that this kernel trick is useful is that there are many regression problems that are not linearly regressable in the space of the inputs $\mathbf{x}$, which might be in a higher dimensionality feature space given a suitable mapping $\mathbf{x} \to \psi(\mathbf{x})$ [4].

2.3 Wavelet kernel and wavelet SVMs

Wavelet kernel SVMs refer to the wavelet kernel that combines the wavelet technique with SVMs [2]. In this sense, the translation invariant wavelet kernel so–called Mexican hat wavelet kernel is [2]:

$$\psi(\mathbf{x}) = \frac{2}{\sqrt[4]{9\pi}} \cdot \left(1 - \mathbf{x}^2\right) \cdot \exp\left(-\frac{\mathbf{x}^2}{2}\right)$$

Given the above mother wavelet, the corresponding wavelet kernel function is [2]:

$$K(\mathbf{x}, \mathbf{x}') = \prod_{j=1}^{N} \frac{2}{\sqrt[4]{9\pi}} \left(1 - \frac{(\mathbf{x}_j - \mathbf{x'}_j)^2}{a^2}\right) \cdot \exp\left(-\frac{(\mathbf{x}_j - \mathbf{x'}_j)^2}{2a^2}\right)$$

The decision function of wavelet SVMs for regression can be expressed as [2]:

$$f(\mathbf{x}) = \sum_{i=1}^{l} (\alpha_i - \alpha_i^*) \prod_{j=1}^{N} \psi\left(\frac{\mathbf{x}_j - \mathbf{x}_{ij}}{a}\right) + b$$

Specifically, the Mexican hat wavelet kernel has been used in this research work due to its similar shape to the HABs.

2.4 The artificial bee colony (ABC) algorithm

The algorithm Artificial Bee Colony (ABC) is an evolutionary optimization algorithm inspired in the behaviour of the bees foraging food sources [3]. Indeed, in the ABC technique, the colony consists of three groups of bees: employed bees, onlookers and scouts. It is assumed that there is only one artificial employed bee for each food source. In other words, the number of employed bees in the colony is equal to the number of food sources around the hive. Employed bees go to their food source and come back to hive and dance on this area. The employed bee whose food source has been abandoned becomes a scout and starts to search for finding a new food source. Onlookers watch the dances of employed bees and choose food sources depending on dances.

2.5 The goodness–of–fit of this approach

It is important to select the model that best fits the experimental data. The criterion based on the coefficient of determination (R^2) was considered here [5]. A coefficient of determination value of 1.0 indicates that the regression curve fits the data perfectly.

3 Analysis of results and discussion

The biological and physical-chemical input variables considered in this work are shown in Appendix A. The total number of predicting variables used to build the hybrid wavelet kernel ABC–SVM–based model was 24. As it can be observed one of the variables is formed by the product of the variable *M. aeruginosa* multiplied by the variable *W. naegeliana* due to the coexistence of these two species of cyanobacteria in order to reproduce their dynamics without interference from external factor. This mathematical formulation adds a multiplicative additional term to account for the two species' interactions according to a more realistic mathematical modelling in Biology [1]. This kind of interaction (synergistic interaction) will be explained later in a more detail. All the input biological variables are measured in biovolume (cubic millimeters per liter) and the output variable (cyanotoxins) in micrograms per liter.

Additionally, it is well known that the SVM techniques are strongly dependent on the SVM hyperparameters: the regularization factor C; the hy-

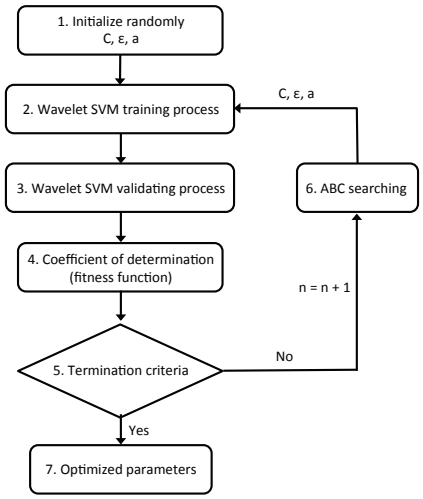


Figure 1: Flowchart of the new hybrid ABC–SVM–based model with Mexican hat wavelet kernel function.

perparameter ε that defines the $\varepsilon-$insensitive tube (allowable error); and finally a that represents the kernel parameter if a Mexican hat wavelet kernel is chosen. We have chosen ABC as a suitable, efficient and simple method for tuning the SVR parameters and a novel hybrid ABC–SVM–based model was applied to predict the cyanotoxin content (output variable) from the other twenty four remaining variables (input variables) in the Trasona reservoir [2–4]. As an example, Fig. 1 shows the flowchart of this new hybrid wavelet kernel ABC–SVM–based model developed in this paper. Cross validation was the standard technique used here for finding the real coefficient of determination (R^2) [5, 6]. Specifically, 10-fold cross-validation was used.

The regression modeling has been performed with SVR–ε using the LIBSVM library [7] and the parameters have been optimized with the ABC technique using the ABC version for MATLAB supplied by Karaboga [3]. The ABC searches for the best C, a and ε parameters by comparing the error in every iteration. Search space is organized in three dimensions, one for each parameter. Main fitness factor is the coefficient of determination (R^2).

Finally, this research work was able to predict the presence of cyanobacteria blooms from 2006 to 2010 in agreement to the actual cyanobacteria blooms observed using the ABC–SVM–based model with great accurateness

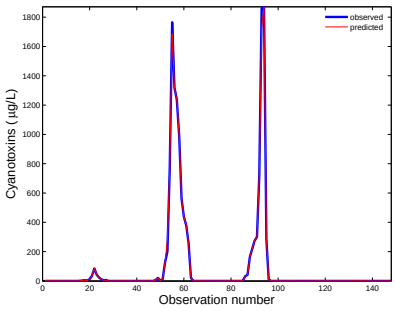


Figure 2: Comparison among the cyanotoxin contents with three blooms of cyanobacteria in the Trasona reservoir from 2006 to 2010 observed and predicted by the ABC–SVM–based models for the Mexican hat wavelet kernel ($R^2 = 0.910$).

and success. Fig. 2 shows the comparison on the Trasona reservoir from 2006 to 2010 between the three blooms of cyanobacteria observed and predicted by the ABC–SVM–based model for the Mexican hat wavelet kernel. Obviously, these results coincide again with the outcome criterion of 'goodness of fit' (R^2) so that the ABC–SVM–based model with a with Mexican hat wavelet kernel function has been the best fitting.

4 Conclusions

Based on the experimental and numerical results, the main findings of this research work can be summarized as follows:

- Firstly, the optimal hyperparameters obtained using a hybrid ABC–SVM–model with a Mexican hat wavelet kernel function in the Trasona reservoir were: $C = 55.9446$, $\varepsilon = 1.0076 \times 10^{-7}$ and $a = 1.4250$.
- Secondly, a high coefficient of determination equal to 0.910 was obtained when this hybrid ABC–SVM–based model with a Mexican hat wavelet kernel function was applied to the experimental dataset. Indeed, the predicted results for this model have been proven to be consistent with the historical dataset of observed actual cyanotoxin concentrations from 2006 to 2010 (see Fig. 2).

- Finally, the synergistic variable *Microcystis aeruginosa*$\times$*Woronichinia naegeliana* and pH values could be considered the most influential parameters in the prediction of cyanotoxin contents in the Trasona reservoir, respectively.

Appendix A. Supplementary data

Supplementary data associated with this article can be found at
`https://www.dropbox.com/s/byvcyy2smpwwv12/Trasona_reservoir_data_sc.xls?dl=0`.

References

[1] J. Huisman, H.C.P. Matthijs, P.M. Visser, Harmful Cyanobacteria, New York, Springer, 2010.

[2] A. Widodo, B.-S. Yang, Wavelet support vector machine for induction machine fault diagnosis based on transient current signal, *Expert Syst. Appl.* 35 (1–2):307–316, 2008.

[3] D. Karaboga, B. Basturk, A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm, *J. Global Optim.* 39(3):459–171, 2007.

[4] I. Steinwart, A. Christmann, Support Vector Machines, New York, Springer, 2008.

[5] D. Freedman, R. Pisani, R. Purves, Statistics, New York, W.W. Norton & Company, 2007.

[6] R. Picard, D. Cook, Cross-validation of regression models, *J. Am. Stat. Assoc.* 79(387):575–583, 1984.

[7] C.-C. Chang, C.-J. Lin, LIBSVM: a library for support vector machines, *ACM T. Int. Syst. Technol.* 2:1–27, 2011.

Valuation of commodity derivatives under jump-diffusion processes*

L. Gómez-Valle†, Z. Habibilashkary‡, and J. Martínez-Rodríguez§

Facultad de Ciencias Económicas y Empresariales, Universidad de Valladolid,

Avenida del Valle Esgueva, 6, 47010-Valladolid, Spain

November 30, 2015

1 Introduction

The estimation of the market prices of risk is an open question in the jump-diffusion derivative literature when a closed-form solution for the future pricing problem is not known. In this paper, we obtain some results that relate the drifts and jump intensities of the risk-neutral processes with future and spot prices. These results provide an original procedure to estimate the risk-neutral drifts and jump intensities. These functions are not observable but their estimation is necessary for pricing commodity derivatives. Moreover, this new approach avoids the estimation of the physical drift as well as the market prices of risk in order to price commodity futures. Finally, an application to NYMEX (New York Mercantile Exchange) data is illustrated.

*L. Gómez-Valle and J. Martínez-Rodríguez were supported in part by the GIR Optimización Dinámica, Finanzas Matemáticas y Utilidad Recursiva of the University of Valladolid and the projects and MTM2014-56022-C2-2-P of the Spanish Ministerio de Economía y Competitividad and European FEDER Funds and VA191U13 of the Consejería de Educación, JCyL.

†e-mail: lourdes@eco.uva.es

‡e-mail: ziba.habibilashkary@alumnos.uva.es

§e-mail: julia@eco.uva.es

2 The valuation model

In this section, we present a two-factor jump-diffusion model of commodity future prices. The first factor is the spot price S, and the second factor is the instantaneous convenience yield δ. Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a probability space equipped with a filtration $\mathcal{F}$ satisfying the usual conditions, [6]. For simplicity and tractability and as usual in the literature, we assume that the distribution of the jump size under $\mathcal{Q}$ measure, risk-neutral probability measure, is known and equal to the distribution under $\mathcal{P}$ measure. That is, we assume that all risk premium related to jump is artificially absorbed by the change in the intensity of jump under the physical measure to $\lambda_S^{\mathcal{Q}}$ under risk-neutral measure, see [4]. The factors of the model are assumed to follow these joint jump-diffusion stochastic processes:

$$dS = \left(\mu_S - \sigma_S \theta^{W_S}\right) dt + \sigma_S dW_S^{\mathcal{Q}} + Y d\tilde{N}_S^{\mathcal{Q}}, \tag{1}$$

$$d\delta = \left(\mu_\delta - \sigma_\delta \theta^{W_\delta}\right) dt + \sigma_\delta dW_\delta^{\mathcal{Q}}, \tag{2}$$

where μ_S and μ_δ are the drifts under $\mathcal{P}$ measure, σ_S and σ_δ the volatilities, Y is the jump amplitude and it is a random variable which follows a normal distribution $N(0, \sigma_Y)$. Moreover, $W_S^{\mathcal{Q}}$ and $W_\delta^{\mathcal{Q}}$ are the Wiener processes. The market prices of risk of Wiener processes are $\theta^{W_S}(S, \delta)$ and $\theta^{W_\delta}(S, \delta)$ and $\tilde{N}_S^{\mathcal{Q}}$ represents the compensated Poisson process, under $\mathcal{Q}$ measure, the risk-neutral measure, with intensity $\lambda_S^{\mathcal{Q}}(S, \delta) = \lambda_S(S, \delta)\theta^{N_S}(S, \delta)$. We assume that the increments to standard Brownian motions are correlated with:

$$dW_S^{\mathcal{Q}} dW_\delta^{\mathcal{Q}} = \rho dt,$$

and $dW_S^{\mathcal{Q}}$ is assumed to be independent of $d\tilde{N}_S^{\mathcal{Q}}$, which means that the diffusion and jump components are independent of each other. We suppose that all functions satisfy enough technical regularity conditions: see [5]. Under the above assumptions, a commodity future price at time t with maturity at time T, $t \leq T$, can be expressed as $F(t, S, \delta; T)$ and at maturity is

$$F(T, S, \delta; T) = S. \tag{3}$$

We also assume that the price of a future can be expressed by

$$F(t, S, \delta; T) = E^{\mathcal{Q}}[S(T)|\mathcal{F}_t], \tag{4}$$

where $E^{\mathcal{Q}}$ denotes the conditional expectation under the $\mathcal{Q}$ measure. The future price (4) is the solution of the following partial integro-differential equation

$$\begin{aligned} F_t &+ \left(\mu_S - \sigma_S\theta^{W_S}\right)F_S + \left(\mu_\delta - \sigma_\delta\theta^{W_\delta}\right)F_\delta + \frac{1}{2}\sigma_S^2 F_{SS} + \frac{1}{2}\sigma_\delta^2 F_{\delta\delta} + \rho\sigma_S\sigma_\delta F_{S\delta} \\ &+ \lambda_S^{\mathcal{Q}} E_{Y_S}^{\mathcal{Q}}[F(t, S + J_S, \delta) - F(t, S, \delta)] = 0 \end{aligned} \tag{5}$$

with terminal condition (3), see the version of Feynman-Kac Dynkin lemma for jump-diffusion processes in [3].

3 Exact results and approximations

In this section, we propose a new approach for estimating the functions of the risk-neutral jump-diffusion stochastic factors directly from data in the markets. Then, as we know all the coefficients of the partial integro-differential equation (5), we can price the future prices. Therefore, it is not necessary to estimate the market prices of risk for pricing commodity futures.

Theorem 1 *Let $F(t, S, \delta; T)$ be a solution to (5) subject to (3), and S follows a jump-diffusion process given by (1) and δ follows a diffusion process given by (2), then:*

$$\frac{\partial F}{\partial T}|_{T=t} = \left(\mu_S - \sigma_S\theta^{W_S}\right)(t), \tag{6}$$

$$\frac{\partial (SF)}{\partial T}|_{T=t} = \left(2S\frac{\partial F}{\partial T}|_{T=t} + \sigma_S^2 + \lambda_S^{\mathcal{Q}} E_{Y_S}^{\mathcal{Q}}[J_S^2]\right)(t), \tag{7}$$

$$\frac{\partial (\delta F)}{\partial T}|_{T=t} = \left(\delta\frac{\partial F}{\partial T}|_{T=t} + S(\mu_\delta - \sigma_\delta\theta^{W_\delta} + \rho\sigma_S\sigma_\delta\right)(t). \tag{8}$$

Parallel results for jump-diffusion interest rate models can be found in [2].

4 Empirical application

In order to show how the approach in Section 3 can be implemented, we will price natural gas futures with data from the NYMEX. Natural gas spot and future prices were obtained from the Energy Information Administration

Table 1: Measures of error, MAE, RMSE and PRMSE, for the out of sample period of time, January 2015-April 2015.

	RMSE	PRMSE	MAE
F1	1.929×10^{-1}	6.919	1.448×10^{-1}
F6	2.567×10^{-1}	8.815	2.271×10^{-1}
F9	1.666×10^{-1}	5.375	1.290×10^{-1}
F12	1.858×10^{-1}	5.678	1.509×10^{-1}
F18	1.711×10^{-1}	5.323	1.399×10^{-1}
F24	3.218×10^{-1}	9.001	2.892×10^{-1}
F36	4.229×10^{-1}	11.261	3.865×10^{-1}
F44	1.933×10^{-1}	5.334	1.711×10^{-1}

of the U.S. Department of Energy (E.I.A. database) and Quandl platform. The sample period covers from January 1997 to April 2015. We also consider future prices with maturities equal to $1, 2, 3, 4, 6, 9, 12, 18, 24, 36$ and 44 months. We use data from January 1997 to December 2014 for estimating the risk-neutral functions and we keep data from January to April 2015 in order to evaluate the results of our approach.

As it is well known in the literature, the convenience yield is not observed in the markets. Therefore, we approximate it as in (2), as usual in the literature.

In order to analyse the behaviour of our model, we use the root mean square error (RMSE), the percentage root mean square error (PRMSE) and the mean square error (MAE) for the out of sample period of time as measures of error:

$$\begin{aligned} RMSE &= \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(F_t-\hat{F}_t\right)^2}, \\ PRMSE &= \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(\frac{F_t-\hat{F}_t}{F_t}\right)^2}, \\ MAE &= \frac{1}{n}\sum_{t=1}^{n}|F_t-\hat{F}_t|, \end{aligned}$$

where n is the number of observations, F_t is the market future price, and $\hat{F}_t$

is the predicted future price by our jump-diffusion model.

Table 1 shows these three measures of error for the future prices along the out of sample period of time for different maturities. Overall, the three measures get the highest level for futures with maturities of 24 and 36 months, i.e., for long maturities.

References

[1] Gibson, R. and Schwartz, E.S. Stochastic convenience yield and the pricing of oil contingent claims, *The Journal of Finance*, Volume 45 (3): 959–976, 1990.

[2] Gómez-Valle, L. and Martínez-Rodríguez, J. Estimation of risk-neutral processes in single-factor jump- diffusion interest rate models, *Journal of Computational and Applied Mathematics*, Volume 291: 48–57, 2016.

[3] Hanson F.B., Applied Stochastic Processes and Control for Jump-Diffusions. Philadelphia, SIAM, 2007.

[4] Nawalkha, S.K., Beliaeva, N., and Soto, G. , Dynamic Term Structure Modelling: The Fixed Income Valuation Course. New Jersey, John Wiley & Sons, Inc, 2007.

[5] Øksendal, B., and Sulem, A., Applied Stochastic Control of Jump Diffusions, Berlin, Heidelberg, Springer-Verlag, 2007.

[6] Runggaldier, W.J., Jump-diffusion models, in: S.T. Rachev (Ed.), Handbook of Hevauy Tayled Distributions in Finance. Universitat Karisruhe, Karisruhe, North Holland, Germany, 169–209, 2003.

Applying a graph model for the Spanish Public University System

Antonio Hervas [♭] [*], Andreu Jimenez[†], Pedro-Pablo Soriano [†], Roberto Capilla [†], Jesus Peinado[†], Joan Guardia [‡], and Maribel Pero [‡]

(♭) Universitat Politecnica de Valencia,

Instituto Universitario de Matematica Multidisciplinar ,

(†) Universitat Politecnica de Valencia ,

(‡) Universitat de Barcelona,

Institut de Recerca en Cervell, Cognicio i Conducta (IR3C).

November 30, 2015

1 Introduction

In the process of accessing the Spanish Public University System (S.U.P.E.), students applying for admission to several degrees in an orderly manner, the "system" assigns each applicant a degree and an university based on the criteria established by the laws or is on a waiting list. Thus traffic occurs between students who have requested a degree and is allocated to another less desired. This paper presents a process that allows us to model the S.U.P.E. access system using a graph, allowing us to identify some properties of the degrees and consequently to analyze the abstract system performance.

2 Building the graph

In the S.U.P.E., after finishing the High School students must pass an exam, and then, according a polynomial formula they must apply for a degree.

[*]e-mail: ahervas@mat.imm.es

They must select in an ordered way their options. The regional government assigns the positions according to their own criteria. The number of students depends on regions, in Catalonia over 50.000 per year, in Valencia 25.000.

So, student applies for Degree A in first position, degree B in second place, another degree C in third place, and successively. We construct a graph whose vertices are the titles that are given in every university of the region. Edge (A, B) shows that students who have applied for the grade A as their first option have been enrolled in degree B. We assign weights based on the percentage of students accessing to their second, third, fourth and fifth option. In Figure 1 we can see the built graph. The size of the vertices indicates the number of vacancies offered to students, the edges indicate the direction of flow of students, and the weight of the edges indicates the value of the flow according to the criteria discussed above.

3 Graph structure

Once obtained the graph, the first thing we did was to study the connected components.In some cases it seemed to make sense. In other cases, not at all, i.e. in engineering appear certain geographic aspects.So, we study the graph over the point of view of communities. This concept is related with the high density of connections in the graph, [1], [3], [2].

In a graph, a community compose a set of vertices that are highly interrelated, meaning that there are many edges between them. In contrast, there are few edges that connect vertices community with the rest of the graph. Or put another way, there is a high density of connections within each community and a low density of connections between communities. The reason for using this technique is given by the fact that: *"Community structure methods normally assume that the network of interest divides naturally into subgroups and the experimenter's job is to find those groups. The number and size of the groups are thus determined by the network itself and not by the experimenter."*, [2].

Given a network, and proposed a division of the network into communities, modularity is a property that indicates how good is this division. Modularity, is a function which evaluates the goodness of partitions of a graph into clusters.[1], [4].

It is considered a good division one that gets communities have a large number of edges within the small communities face a number of inter-communal

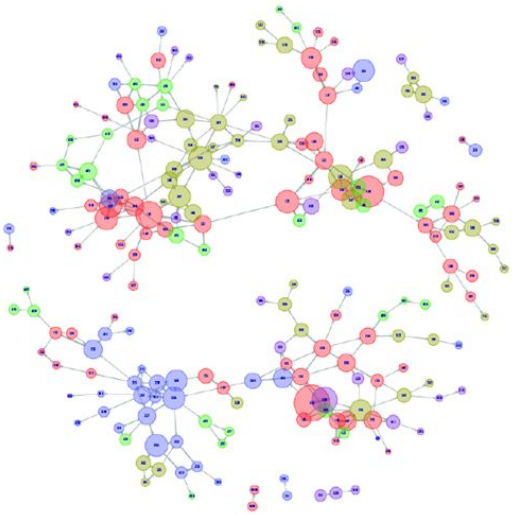


Figure 1: The graph for a system with 250 degrees and 25.000 students.

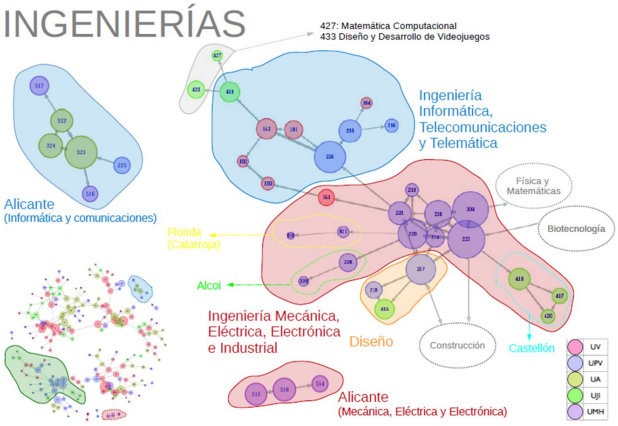


Figure 2: Connected component of the graph corresponding to the area of Engineering.

edges (vertices incidents these edges belong to different communities).There are several algorithms that allow us to get the communities on a graph, and show us modularity, but they offer different results depending on the criteria for group vertices used. Also most of them apply only to undirected graphs.[1], [3], [5], [6] and [7].

We apply all the algorithms to our graph. Even the community looks similar, really the results are very different. See Figure 3.

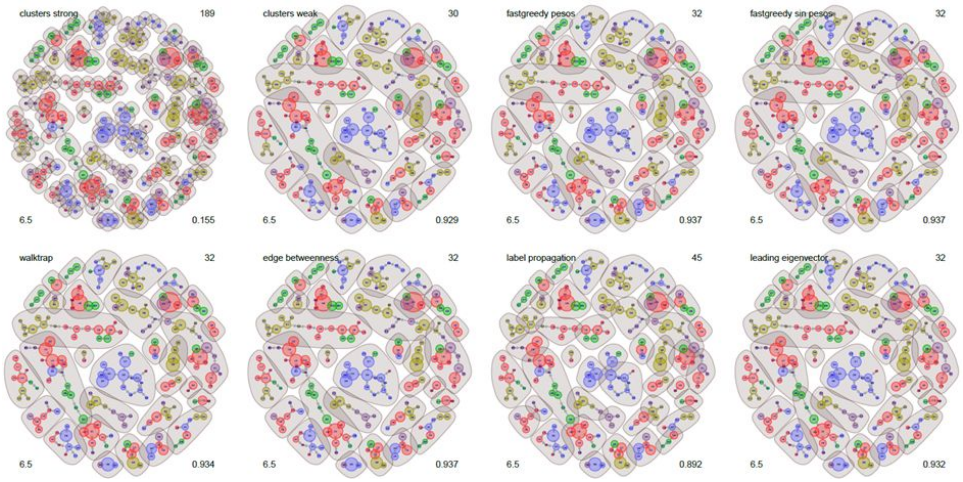


Figure 3: Community graph under different algorithms.

4 Proposed Algorithm.

So, we propose an algorithm that finds communities in which all vertices can reach the same vertex. The main idea of operating algorithm is as follows:

- Order the vertices for the number of other vertices can be reached from each one, m
- In the event that m was reached by two vertices, then orders the vertices for the number of vertices that can achieve
- Walk the ordered sequence of vertices for each vertex v marked as members of the same community to those from which one can reach the vertex v, and that does not yet belong to any community.

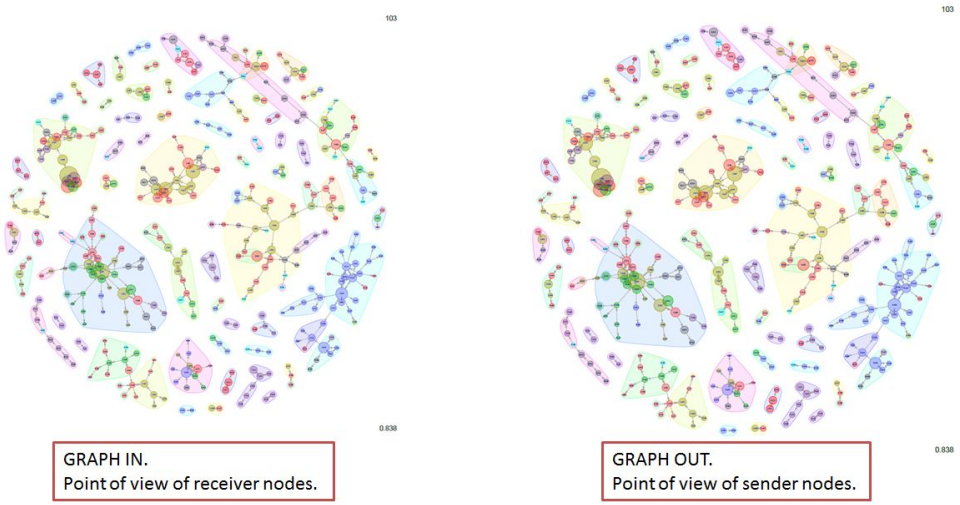


Figure 4: Graph In and Graph Out.

- Each community is established around the vertex from which fewer vertices are reached and to reach the most vertices.

Thus we obtain two approaches: GRAPH IN, from the standpoint of receiver vertex and GRAPH OUT, from the viewpoint of the sender vertex.

With this algorithm we obtain a more accurate vision of the students selection process, and this allows us to to strike up new lines of works once studied and verified the algorithm.

References

[1] S. Fortunato, Community detection in graphs, *Physics Reports*, vol. 486, no. 3,pp. 75-174, 2010.

[2] Newman, M. E. Modularity and community structure in networks. Proceedings of the National Academy of Sciences, 103(23), 8577-8582. (2006).

[3] M. Girvan and M. E. Newman Community structure in social and biological networks *Proceedings of the National Academy of Sciences*, vol. 99, no. 12, pp. 7821.7826, 2002.

[4] M. E. Newman and M. Girvan. Finding and evaluating community structure in networks *Physical review E*, vol. 69, no. 2, p. 026113, 2004.

[5] A. Clauset, M. E. Newman, and C. Moore Finding community structure in very large networks *Physical review E*, vol. 70, no. 6, p. 066111, 2004.

[6] U. N. Raghavan, R. Albert, and S. Kumara. Near linear time algorithm to detect community structures in large-scale networks. *Physical Review E*, vol. 76, no. 3,p. 036106, 2007.

[7] M. E. Newman. Finding community structure in networks using the eigenvectors of matrices *Physical review E*, vol. 74, no. 3, p. 036104, 2006.

New iterative splitting methods for partial differential equations*

J. Geiser♭, J.L. Hueso†, and E. Martínez‡

(♭) Dept. of Electrical Engineering and Information Technology,

Ruhr-University of Bochum, Germany,

(†) Instituto de Matemática Multidisciplinar,

Universitat Politècnica de València, Spain,

(‡) Instituto de Matemática Pura y Aplicada,

Universitat Politècnica de València, Spain.

November 30, 2015

1 Introduction

We propose some modifications in schemes of iterative splitting techniques for partial differential equations and introduce the parallel version of some algorithms. Theoretical results related to the order of the iterative splitting for these schemes are obtained. In the numerical experiments we compare the obtained results by applying iterative methods to approximate the solutions of the nonlinear systems obtained from the discretization of the splitting techniques.

Iterative splitting schemes are used to solve nonlinear systems obtained from ordinary differential equations or spatial discretized partial differential equations. In their original scheme, one applies a Picard-iterative technique

*This work has been supported by Ministerio de Ciencia e Innovación de España MTM2014-52016-C2-02.

†e-mail: jlhueso@mat.upv.es

to solve the nonlinear systems, see [2] and [3]. The drawback of such Picard's techniques is that they are a first order scheme, see [4].

In this work, we are motivated to develop new nonlinear solvers, which a are modification of the iterative splitting schemes in [1] and [2]. Our improvement is finding higher order nonlinear schemes by using Newton's method, which is embedded in the splitting methods. Such novel schemes are more accurate, of higher order and they accelerate the traditional solver schemes. Further, we derive parallel versions of each scheme to obtain an up to date nonlinear solver for large scale computations, see [5].

2 Splitting techniques

A great variety of natural phenomenons can be described by an ordinary differential equation or a partial differential equation, the solution of which not always can be obtained by analytical methods. In fact, in the majority of cases, it is much more practical using numerical methods in order to approximate the solution.

Specifically, we deal with a particular form for PDE that can be expressed in the following form:

$$\frac{\partial u(v,t)}{\partial t} = A(u(v,t))u(v,t) + B(u(v,t))u(v,t) + g(v,t), \tag{1}$$

where $t \in [0,T]$, the initial condition $u(v,0) = u_0(v)$, and $A(u), B(u)$ are operators, in a Banach space X, involving only spatial derivatives of u, while $g(v,t)$ is an exterior perturbation.

In this cases, the splitting techniques can be used with the aim of substitute the original problem into a sequence of simpler problems, when the size of the problem is big, or maybe we need solve the problem taking into account physical properties of some part of the equation, always with the final objective of building efficient methods with the usual properties of accuracy and stability.

The traditional method is the sequential operator splitting but nowadays iterative splitting is being the objective of different studies, [6],[7]. In all cases we discretize the time interval $[0,T]$ in N subintervals $[t^n, t^{n+1}], n = 1, 2, \ldots, N$ and solve different problems consecutively in each of these subintervals. Now, our aim is to introduce some modifications in the recent iterative splitting schemes defined in [1], and to present the parallel version of these algorithms.

2.1 New schemes for iterative splitting

Let us consider algorithm (5.13) of [1], that is written as follows:

$$\begin{aligned}\frac{d\tilde{u}_i(t)}{dt} &= A(u_{i-1}(t))\tilde{u}_i(t) + B(u_{i-1}(t))u_{i-1}(t), \text{ with } \tilde{u}_i(t^n) = u(t^n),\\ \frac{du_i(t)}{dt} &= A(\tilde{u}_i(t))\tilde{u}_i(t) + B(\tilde{u}_i(t))u_i(t), \text{ with } u_i(t^n) = u(t^n),\end{aligned}$$

We propose a modification based on the idea of using the solution of first equation $\tilde{u}_i(t)$ updating operators $A(u)$ and $B(u)$ in the second equation of the same iteration, that is, substituting $A(u_{i-1})$ and $B(u_{i-1})$ by $A(\tilde{u}_i)$ and $B(\tilde{u}_i)$. Moreover we consider all possible variations for combining the nonlinear operators, having four different schemes.

2.1.1 Linear Serial Iterative Splitting (LS)

$$\frac{d\tilde{u}_i(t)}{dt} = A(u_{i-1}(t))\tilde{u}_i(t) + B(u_{i-1}(t))u_{i-1}(t), \text{ with } \tilde{u}_i(t^n) = u(t^n), \tag{2}$$

$$\frac{du_i(t)}{dt} = A(\tilde{u}_i(t))\tilde{u}_i(t) + B(\tilde{u}_i(t))u_i(t), \text{ with } u_i(t^n) = u(t^n), \tag{3}$$

with $i = 1, 2, \ldots, m$, and $u_0(t)$ is a fixed starting function that verifies the initial condition $u_0(0) = u_0$. Once the m iterations have been performed, we take the approximated solution $u(t^{n+1}) = u_m(t^{n+1})$.

2.1.2 Linear Parallel Iterative Splitting (LP)

$$\frac{d\tilde{u}_i(t)}{dt} = A(u_{i-1}(t))\tilde{u}_i(t) + B(u_{i-1}(t))u_{i-1}(t), \text{ with } \tilde{u}_i(t^n) = c^n, \tag{4}$$

$$\frac{d\tilde{\tilde{u}}_i(t)}{dt} = A(u_{i-1}(t))u_{i-1}(t) + B(u_{i-1}(t))\tilde{\tilde{u}}_i(t), \text{ with } \tilde{\tilde{u}}_i(t^n) = u(t^n), \tag{5}$$

and the next iteration is

$$u_i(t) = \frac{\tilde{u}_i(t) + \tilde{\tilde{u}}_i(t)}{2}$$

with $i = 1, 2, \ldots, m$, and $u_0(t)$ is a fixed starting function that verifies the initial condition $u_0(0) = u_0$. Once the m iterations have been performed, we take the approximated solution $u(t^{n+1}) = u_m(t^{n+1})$.

2.1.3 Quasilinear Serial Iterative Splitting (QS)

$$\frac{d\tilde{u}_i(t)}{dt} = A(\tilde{u}_i(t))\tilde{u}_i(t) + B(u_{i-1}(t))u_{i-1}(t), \text{ with } \tilde{u}_i(t^n) = u(t^n),$$
$$\frac{du_i(t)}{dt} = A(\tilde{u}_i(t))\tilde{u}_i(t) + B(u_i(t))u_i(t), \text{ with } u_i(t^n) = u(t^n),$$

with $i = 1, 2, \ldots, m$, and $u_0(t)$ is a fixed starting function that verifies the initial condition $u_0(t) = 0$. Once the m iterations have been performed, we take the approximated solution $u(t^{n+1}) = u_m(t^{n+1})$.

2.1.4 Quasilinear Parallel Iterative Splitting (QP)

On the other hand, we can choose a parallel version, which is given as:

$$\frac{d\tilde{u}_i(t)}{dt} = A(\tilde{u}_i(t))\tilde{u}_i(t) + B(u_{i-1}(t))u_{i-1}(t), \text{ with } \tilde{u}_i(t^n) = u(t^n),$$
$$\frac{d\tilde{\tilde{u}}_i(t)}{dt} = A(u_{i-1}(t))u_{i-1}(t) + B(\tilde{\tilde{u}}_i(t))\tilde{\tilde{u}}_i(t), \text{ with } \tilde{\tilde{u}}_i(t^n) = u(t^n),$$

and the next iteration is

$$u_i(t) = \frac{\tilde{u}_i(t) + \tilde{\tilde{u}}_i(t)}{2}$$

with $i = 1, 2, \ldots, m$, and $u_0(t)$ is a fixed starting function that verifies the initial condition $u_0(t) = 0$.

Once the m iterations have been performed, we take the approximated solution $u(t^{n+1}) = u_m(t^{n+1})$.

3 Theoretical results

In this section we obtain the convergence order for the different algorithms and the expressions for the local error.

Theorem 3.1. *We assume to have bounded nonlinear operators A and B, while given as $||A(u)|| \leq A_{max}$ and $||B(u)|| \leq B_{max}$ for all $u \in X$.*

Then the numerical errors after m iterations are given for the different schemes as:

1. *Serial version, the numerical error is given as:*

$$||e_m(\tau)|| \leq A_{max}^m B_{max}^m O(\tau^{2m}),$$

2. *Parallel version, the numerical error is given as:*

$$||e_m(\tau)|| \leq (\frac{A_{max}}{2} + \frac{B_{max}}{2})^m O(\tau^m).$$

4 Numerical experience

We compare different splitting schemes applied to two examples, the mixed convection-diffusion and Burgers' equation and a momentum equation that models a viscous flow. The differential equations in each splitting interval are solved by using the back-Euler-Newton algorithm.

It is worth mentioning that one step of Newton's method suffices to reach an approximate solution of the implicit Euler method in each splitting interval. Additional steps do not produce a significant error reduction. Thus, in the numerical examples, only one Newton's step is performed.

References

[1] J. Geiser. Modified Jacobian Newton Iterative Method: Theory and Applications. *Mathematical problems in Engineering*, Hindawi Publishing Corporation, (2009), Article ID 307298.

[2] J. Geiser. *Iterative Splitting Methods for Differential Equations.* Chapman & Hall/CRC Numerical Analysis and Scientific Computing Series, ed. Magoules and Lai, 2011.

[3] O. Nevanlinna. Remarks on Picard-Lindelöf Iteration: Part I, *BIT* 29 (1989) 328–346.

[4] C.T. Kelly. *Iterative Methods for Linear and Nonlinear Equations.* Frontiers in Applied Mathematics, SIAM, Philadelphia, PA, 1995.

[5] M.J. Gander. *50 Years of Time Parallel Time Integration*, T. Carraro, M. Geiger, S. Körkel, R. Rannacher, to appear in Multiple Shooting and Time Domain Decomposition, Springer Verlag, 2015.

[6] I. Farago and J. Geiser. *Iterative Operator-Splitting methods for Linear Problems*, Preprint No. 1043 of Weierstrass Institute for Applied Analysis and Stochastics, Berlin, Germany, 2005.

[7] J. Geiser. Iterative Operator-Splitting Methods with higher order Time-Integration Methods and Applications for Parabolic Partial Differential Equations, *Journal of Computational and Applied Mathematics* 217, (2008), 227–242.

Real-time water demand forecasting using support vector machine and adaptive Fourier series

B. Brentan[⋆] [,*], E. Luvizotto Jr.[⋆], M. Herrera [†] J. Izquierdo [◦] and R. Pérez-García[◦]

(⋆)Computaional Hydraulic Laboratory, University of Campinas, Brazil

(†) EDEn - Dept. of Architecture and Civil Eng., University of Bath, UK

(◦) Fluing - IMM, Universitat Politècnica de València, Spain

November 30, 2015

1 Introduction

Both safe operations of water supply systems (WDSs) and design of new WDSs require water demand forecasting. For a monitored WDS, the use of data can provide deep knowledge on water demand and become a powerful management tool to improve system efficiency.

ARIMA models have been traditionally applied for modeling water demand. However this approach usually considers linear correlations among variables. This hypothesis not always help develop a model able to make predictions with required accuracy, thus harming control processes. Recent works propose the use of artificial intelligence and machine learn tools to model the non-linearity among the variables. [**?**] compare various predictive methods applied to hourly water demand forecasting, suggesting support vector machine as one of the best models. The arrival of new data can become obsolete in off-line models and the application to operation modes requires great improvement of accuracy, which can be obtained by on-line models.

*brentan@fec.unicamp.br:

The main feature of on-line models is their ability to improve the accuracy by recalculation the whole process each time new data is embedded. Some on-line methods are proposed in the literature such as sliding windows methodologies, which use kernel regression with fast parameter optimisation, hybrid model with ARIMA and ANN working together, applied to dayly demand forecast. A similar approach is used in [**?**] modeling water demand intervened (e.g. by open/close valve manoeuvres).

Based on [**?**, **?**], our work uses support vectro regression (SVR), running for short-term water demand forecasting. Built on this model, an on-line process based on Fourier time series is launched to improve the predictions. The error associated to the SVR model is investigated and using adaptive Fourier series (AFSs), the prediction made by SVR is adjusted by error prediction using AFS. Also an optimal training cycle is defined using an efficiency parameter, which is critical to update The SVR off-line model.

2 Hybrid model: Support Vector Regression - Adaptive Fourier Series

2.1 Support Vector Regression (SVR)

Kernel-based learning methods use an implicit mapping ϕ in a high dimensional (feature) space and convert the non-linear relations into linear ones. The learning then takes place in the feature space, while the learning algorithm can be expressed so that the data points only appear inside dot products with other points, readily calculated via the kernel. A specify margin ε is the key characteristic of SVR, within which we are willing to accept errors in the sample data without they affecting quality prediction. The SVR predictor is defined by those points or vectors which lie outside the region formed by the band of size $\pm\varepsilon$ around the regression. These vectors are the so-called *support vectors*. The goal is to find a function

$$\hat{f}(\boldsymbol{x}) = \langle \boldsymbol{w}, \phi(\boldsymbol{x}) \rangle + b, \tag{1}$$

that at most deviates ε from the observed output, y_i and, at the same time, minimizes the so-called model complexity, which depends only on the support vectors. This method of tolerating errors is said to be ε-*insensitive* [**?**]. To complete this approach, *slack variables* are included in the ε-regression

fostering the chances to achieve better predictions in a more relaxed space.

2.2 Adaptive Fourier Series (AFS)

The Fourier series set of equations presented here is based on [**?**], where trigonometric adjustment is applied to data coming from both equally and not-equally spaced measurements. Taking equally spaced values of t in the period of interest, and normalizing times to the interval $[0\ ,\ 2\pi]$, the error e between the real deviation d and the Fourier value at time step t_i is written

$$e = \sum_{i=0}^{N-1} \{f(t_i) - [a_0 + \sum_{j=1}^{M} a_j \cdot cos(jt_i) + \sum_{j=1}^{M} b_j \cdot sin(jt_i)]\}^2. \qquad (2)$$

where M is the length of the Fourier polynomial, and a_0, a_j and b_j are the adjustable Fourier coefficients.

Applying, the least square method to minimize e, taking into account obvious orthogonality conditions, it is possible to obtain each adjustable coefficient of the series.

2.3 Hybrid model

The hybrid model can be fast enough to respond to abrupt changes in water demand conditions. Using SVR for standard prediction, we propose an additive layer, the Fourier layer, to adjust the deviation. First, a calibration process is required to define the SVR parameters:

- Parameter C, which gives a trade-off between model complexity and the amount up to which deviations larger than ε are tolerated, also responsible of the robustness of the regression model.
- Parameter ε, which regulates the radius of the ε-tube around the regression function.

To train the model, a mesh is created with pairs of ϵ and C, with corresponding ranges: $0.05 \leq \epsilon \leq 0.9$ and $1 \leq C \leq 1500$. A Grid Search Method is applied to tune the parameters, in which each pair represent a possible of solution.

Once tuned, the model is running and the deviations d between predicted and observe demand are computed. Deviation is typically larger at demand

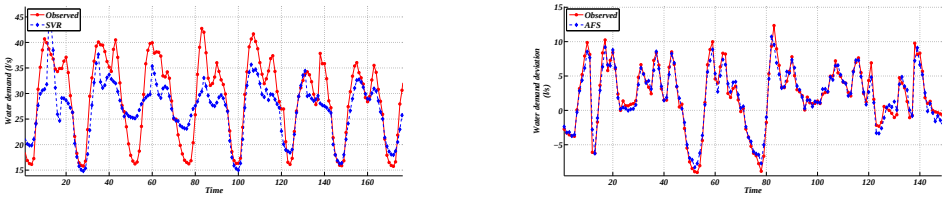
peaks and, as expected, has periodical behavior.The final value of water demand is obtained by correcting the SVR value y_{SVR} with the predicted error by the AFS model d_{AFS}. An optimal cycle of model regeneration is presented as a novelty. The off-line component model can become an obsolete structure, once the training data is far from the prediction time. However, a continuous update have a high computational cost and is deemed not necessary. This cycle is determined by controlling both model accuracy and total CPU time.

3 Experimental Study

This study uses water demand data collected from a real district metered area (DMA) in Franca, Brazil, corresponding to metered data at the DMA's inlet every 20 minutes from May 2012 until December 2013. Previous studies found in the literature use variables taking into account weather and calendar information for generating models to predict water demand [**?**, **?**, **?**]. Our SVR model uses rain, temperature, humidity, and wind velocity as the most important physical variables involved in water demand forecasting and uses weekday, hour of day, month or year and holiday occurrence as calendar variables.

The best training parameter values for C and ϵ found by Grid Search are 50 and 0.05, respectively. The predicted demand using the best validated model is presented in Figure 1a, which confirms the largest deviations occur at the maxima and minima. Finally, the SVR model prediction with corrected deviation via the AFS model is presented in Figure 1b. Statistical evaluation of error shows the decrease of error from 12.91l/s to 3.45l/s when applying the AFS adjustment and the increase of correlation coefficient from 0.745 to 0.974, pointing at clear improvement of the quality results obtained by the AFS-SVR hybrid model proposed.

The off-line model structure needs to be periodically updated through new data. Thus, determining an optimal updating cycle is of paramount importance. Let T the total CPU time spent to run the calibrated hybrid model corresponding to SVR prediction and ASF deviation adjustment. We define the efficiency as the trade-off between accuracy and computational cost, and can be written as the relation between the training data size N_{tr} and the product $e \times T$. The optimum training cycle we found was after 3000 register, corresponding to 125 days.



(a) SVR model results (b) Hybrid model results

Figure 1: (a) deviation between real and SVR predicted demand and (b) AFS approximation of this deviation

4 Conclusions

This article presents a hybrid model for hourly water demand forecasting in WDSs. The model builds over an off-line Support Vector Regression model, constituting a base forecasting, an an on-line Adaptive Fourier Series adjustment, responsible for correcting the SVR deviation. SVR accounts for physical and calendar information necessary for water demand forecasting. However, it is not able to capture well the extremes and, as a result, the accuracy diminishes at the demand peaks. The use of Adaptive Fourier Series aggregates to the SVR model a way to update the prediction in near-real time by correcting the demand predicted by the off-line base model.

A simple way to determine the optimum training data size for the off-line model is presented. This cycle can help water companies to organize interruptions of the model activity for update with new data. Updating the model is important since new data supplement the model with new correlations among demand and input variables.

Effects of the obesity on optimal control schedules of chemotherapy on a cancerous tumor

S.E. Delgadillo[♭] [*], R.A. Ku-Carrillo[†], and B.M. Chen-Charpentier[‡]

(♭, †)Departamento de Matemáticas y Física, Universidad Autónoma de Aguascalientes

Av. Universidad 940, Cd. Universitaria, Aguascalientes, Mexico

(‡) Department of Mathematics, University of Texas at Arlington,

Arlington, TX 76019, USA

November 30, 2015

1 Introduction

In this work, we propose and analyze a cancer-obesity model for the growth of a tumor where chemotherapy schedules are obtained using optimal control techniques. Our aim is to study how the obesity can affect different tumor growth scenarios when chemotherapy is applied. This was motivated by the reported relationship between obesity and cancer in several experimental studies such as [1, 2, 3, 4]. Among the cancers that present this relationship are colorectal cancer, breast cancer in postmenopausal women, endometrium cancer, renal cancer, and esophagus cancer [5].

Our model incorporates an equation for stored fat in the organism by means of a logistic equation. It is well known that this equation includes a coefficient to model the carrying capacity which could be related to the fat stored in the organism. In order to get some insights about this problem, this model proposes a mathematical interaction between the stored fat in the

[*]e-mail: sedelgad@correo.uaa.mx

organism and the cancerous tumor. We consider that an increment of fat in the organism increases the growth rate of the tumor. Another aspect that our model expects to capture is the dynamics of the immune system cells, particularly the T8 cells related to the regress of small tumor. In this case, the immune response is a natural control that kills tumor cells in order to keep a healthy organism.

A literature review of the models of obesity and cancer shows different approaches. Close related models to ours are those which take into account the immune system and the tumor cells populations such as [7, 8], Kirschner and Panetta [9], and Bellomo [10]. Even though this list is not exhaustive, it illustrates the wide variety of immune cancer models. Our model proposal is based on De Pillis and Radunskaya model, reported in [7, 8] which also models the chemotherapy treatment using control theory.

2 The cancer-obesity model with chemotherapy

The model that we study in this work is the following:

$$\begin{aligned}
\dot{I} &= s + \frac{\rho IT}{\alpha + \mu F + T} - c_1 IT - d_1 I - a_1(1 - e^{-u})I \\
\dot{T} &= r_1 T(1 - b_1 T) - c_2 IT - c_3 TN + c_5 TF - a_2(1 - e^{-u})T \\
\dot{N} &= r_2 N(1 - b_2 N) - c_4 TN - a_3(1 - e^{-u})N \\
\dot{F} &= r_3 F(1 - b_3 F) - c_6 FT - a_4(1 - e^{-u})F, \\
\dot{u} &= v - d_2 u
\end{aligned} \tag{1}$$

where $I(t)$ denotes the number of immune cells at time t, $T(t)$ is the number of tumor cells at time t, $N(t)$ the number of normal cells at time t, $F(t)$ is the fat stored in adipocytes at time t, and $u(t)$ is the drug in the organism. The function $v(t)$ models the application protocol of chemotherapy. The parameters associated to the immune system are s, the basal response, and ρ, the immune response stimulated by the cancer cells. The constants r_1, r_2 and r_3 are the growth rates for the tumor cells, the normal cells and the fat, respectively. The competition terms are the constants c_1, c_2, c_3, c_4 between the different populations and c_5 is the contribution of the fat to the cancerous cells growth. Finally, the constants a_1, a_2, a_3, and a_4 model the kill effectiveness of the drug on I, T, N and F populations, respectively.

The equilibrium analysis of model (1) without the presence of chemotherapy is important to understand the dynamics of the system, i.e. when $u(t) = 0$ for all t. Let us mention, that the model has been normalized in such a way that b_2.

The equilibrium points can be classified depending on whether there is presence of fat ($F = 1/b_3$) or not ($F = 0$). The following equilibrium points arise in the model (1) with the condition, $F = 1/b_3$ obtained by solving the associated nullclines system.

- *Dead:* $PO_1 = (\frac{s}{d_1}, 0, 0, \frac{1}{b_3})$. Only immune cells and fat are present. It is easy to verify that this equilibrium point is always unstable.
- *Total cancer invasion*: $PO_2 = (I, T, 0, \frac{1}{b_3})$ represents a case of death because there are not normal cells. The stability of the equilibrium point changes depending on the parameter values.
- *Tumor-free*: $PO_3 = (\frac{s}{d_1}, 0, \frac{1}{b_2}, \frac{1}{b_3})$ represents the case of health because the tumor density is zero. This point is stable when $r_1 + \frac{c_5}{b_3} < \frac{c_2 s}{d_1} + \frac{c_3}{b_2}$. The stability condition depends on the tumor growth rate, the interaction coefficients and the basal source rate of the immune system.
- *Coexistence*: $PO_4 = (I, T, N, \frac{1}{b_3})$. In these equilibria all types of cells coexist and there can be zero, one, two or three points depending on the parameter values. The stability might vary depending on the parameter values.

For the case when there is not presence of fat $F = 0$ we found the same type of equilibria, but they are all unstable because the growth rate of the fat is positive, $r_3 > 0$.

3 Optimal control application

Our optimal control problem consists in finding a function $v(t)$ such as the cancer cell population T hits a minimum at a given time t_f while the density of normal cells is greater than 0.75 for all t. We look for minimizing the objective function:

$$J(t_f) = T(t_f),$$

restricted to keeping the normal cells population above 0.75 :

$$N(t) \geq 0.75, \quad 0 \leq t \leq t_f.$$

For simplicity, we denote the state variables by $(I, T, N, F, u) = (x_1, x_2, x_3, x_4, x_5)$. If t_f is the final time, the optimal control problem will be:

$$\text{Min} \quad J(x, v) = x_2(t_f),$$

subject to the equation system (1) and the following constrain:

$$k(x, t, v) = x_3(t) - 0.75 \geq 0, \quad 0 \leq t \leq t_f. \tag{2}$$

The associated hamiltonian for the optimal control problem is:

$$H = p_1(\dot{x}_1) + p_2(\dot{x}_2) + p_3(\dot{x}_3) + p_4(\dot{x}_4) + p_5(\dot{x}_5) + \eta k,$$

where the functions $\dot{p}_i = -\frac{\partial H}{\partial x_i}$, or explicitly:

$$\begin{aligned}
\dot{p}_1 &= -p_1 \left(\frac{\rho x_2}{\alpha + x_2 + \mu x_4} - c_1 x_2 - d_1 - a_1(1 - e^{-x_5}) \right) + p_2 c_2 x_2 \\
\dot{p}_2 &= -p_1 \left(\frac{(\alpha + \mu x_4)\rho x_1}{(\alpha + x_2 + \mu x_4)^2} - c_1 x_1 \right) - p_2(r_1 - 2r_1 b_1 x_2 - c_2 x_1 - c_3 x_3 + c_5 x_4 \\
&\quad - a_2(1 - e^{-x_5})) + p_3 c_4 x_3 + p_4 c_6 x_4 \\
\dot{p}_3 &= p_2 c_3 x_2 - p_3 \left(r_2 - 2r_2 b_2 x_3 - c_4 x_2 - a_3(1 - e^{-x_5}) \right) - \eta(t) \\
\dot{p}_4 &= p_1 \left(\frac{\rho x_1 x_2 \mu}{(\alpha + x_2 + \mu x_4)^2} \right) - p_2 c_2 x_2 - p_4 \left(r_3 - 2r_3 b_3 x_4 - c_6 x_2 - a_4(1 - e^{-x_5}) \right) \\
\dot{p}_5 &= e^{-x_5} \left(p_1 a_1 x_1 + p_2 a_2 x_2 + p_3 a_3 x_3 + p_4 a_4 x_4 \right) + p_5 d_2,
\end{aligned} \tag{3}$$

where $\eta(t) \leq 0$ with $\eta(t)k(t) = 0$.

$$\eta(t) = \begin{cases} -1 & \text{if} \quad x_3(t) \leq 0.75 \\ 0 & \text{other} \quad \text{case.} \end{cases} \tag{4}$$

The boundary values for the co-state variables are:

$$p_1(t_f) = 0, \quad p_2(t_f) = 1, \quad p_3(t_f) = 0, \quad p_4(t_f) = 0, \quad p_5(t_f) = 0.$$

Finally, the control equation is given by,

$$\frac{\partial H}{\partial v} = p_5 = 0,$$

for $0 \leq v(t) \leq 1$. Bang-bang type solutions are compatible with this system,

$$v(t) = \begin{cases} 0 & p5 > 0 \\ 1 & p5 < 0. \end{cases} \tag{5}$$

We carried out numerical experiments to determine the optimal chemotherapy schedules using parameters reported in the literature.

4 Conclusions

A mathematical model for the effect of obesity on the tumor growth, and consequently, on the optimal control chemotherapy schedules has been discussed in this paper. The numerical experiments showed that the effects of losing weight can be adjuvant to the chemotherapy treatment, and also, that weight gain can cause a lesser reduction of the tumor when is applied chemotherapy. Specific examples of such effects were presented in paper.

References

[1] Rubin, H., Promotion and selection by serum growth factors drive field cancerization, which is anticipated in vivo by type 2 diabetes and obesity, *Proceedings of the National Academy of Sciences*, 110(34): 13927-13931, 2013.

[2] Ehsanipour, E.A. and Sheng, X. and Behan, J.W. and Wang, X. and Butturini, A. and Avramis, V.I. and Mittelman, S.D., Adipocytes cause leukemia cell resistance to L-asparaginase via release of glutamine, *Cancer research*, 73(10): 2998-3006, 2013.

[3] Taubes, G., Unraveling the obesity-cancer connection, *Science*, 335(6064): 28-32, 2012.

[4] Hursting, S.D., Minireview: The Year in Obesity and Cancer, *Molecular Endocrinology* 26(12): 1961-1966, 2012.

[5] Schwab, M., Encyclopedia of cancer, Berlin Heidelberg, Springer, 2016.

[6] Zhang, X., Wu, W.and Yu, J., Obesity and Cancer. Zurich Springer International Publishing, 2016.

[7] De Pillis, L.G. and Radunskaya, A., A mathematical tumor model with immune resistance and drug therapy: an optimal control approach, *Computational and Mathematical Methods in Medicine*, 3(2): 79-100, 2001.

[8] De Pillis, L.G. and Radunskaya, A., The dynamics of an optimally controlled tumor model: A case study, *Mathematical and Computer Modelling*, 37(11):1221-1244, 2003.

[9] Kirschner, D. and Panetta, J.C., *Journal of mathematical biology* 37(3): 235-252, 1998.

[10] Bellomo, N. and Li, N.K. and Maini, P.K., On the foundations of cancer modelling: selected topics, speculations, and perspectives, *Mathematical Models and Methods in Applied Sciences*, 18(04): 593–646, 2008.

Convergence results for an optimal fourth-order family of methods presented by Sharma

I. K. Argyros[♭], A. Cordero[†], Á. A. Magreñán[‡] *and J. R. Torregrosa[†]

(♭) Cameron University,

73505 Lawton, OK, USA

(†) Universitat Politècnica de València,

46022 València, Spain

(‡) Universidad Internacional de La Rioja (UNIR),

Av. Gran Vía Rey Juan Carlos I, 41, 26002 Logroño, La Rioja, Spain.

November 15, 2015

1 Introduction

In this study, we are concerned with the problem of approximating a solution x^* of the nonlinear equation

$$F(x) = 0 \tag{1}$$

where F is a differentiable operator defined on a subset Ω of X with values in Y. Using mathematical modelling, many problems in computational sciences and other disciplines can be brought in a form like (1). The solutions of these equations (1) can rarely be found in closed form. Therefore solutions of these equations (1) are approximated by iterative methods.

In this work, we consider the following family of methods defined by

*e-mail: alberto.magrenan@unir.net

Sharma in [6] for each $n = 0, 1, 2, \ldots$ by

$$\begin{aligned} y_n &= x_n - \frac{2}{3}\frac{F(x_n)}{F'(x_n)} \\ z_n &= x_n - \frac{F'(x_n) + 3F'(y_n)}{-2F'(x_n) + 6F'(y_n)}\frac{F(x_n)}{F'(x_n)} \\ x_{n+1} &= x_n - \frac{F'(x_n) + aF'(y_n)}{bF'(x_n) + cF'(y_n)}\frac{F(z_n)}{F'(x_n)} \end{aligned} \tag{2}$$

where a, b, c are parameters.

2 Convergence result

We shall study the local convergence analysis of method defined for each $x = 0, 1, 2 \ldots$

$$\begin{aligned} y_n &= x_n - \xi F'(x_n)^{-1} F(x_n) \\ z_n &= x_n - \lambda A_n^{-1} F'(x_n)^{-1} F(x_n) \\ x_{n+1} &= z_n - \mu B_n^{-1} F'(x_n)^{-1} F(z_n) \end{aligned} \tag{3}$$

where x_0 is an initial point, ξ, λ, μ are parameters,

$$A_n = 2\left(F'(x_n) + 3F'(y_n)\right)^{-1}\left(F'(x_n) - 3F'(y_n)\right)$$

and

$$B_n = \left(F'(x_n) + aF'(y_n)\right)^{-1}\left(bF'(x_n) + cF'(y_n)\right).$$

Theorem 2.1 *Let $F : \mathbb{D} \subset \mathbb{X} \to \mathbb{Y}$ be a Fréchet-differentiable operator. Suppose that there exist $x^* \in \mathbb{D}$, $L_0 > 0$, $L > 0$, $M_0 > 0$, $M \geq 1$, $\alpha > 0$, $a \in S \setminus \{-1\}$,b, c, ξ, λ, $\mu \in S$ such that for all $x, y \in D$ the following conditions hold:*

$$M|1 - \xi| < 1,$$

$$\frac{p(0) M_0}{4}(1 + |\mu| M) < 1,$$

$$F(x^*) = 0, \quad F'(x^*)^{-1} \in L(\mathbb{Y}, \mathbb{X}), \quad \|F'(x^*)\| \leq \alpha, \tag{4}$$

$$\|F'(x^*)^{-1}(F'(x) - F'(x^*))\| \leq L_0 \|x - x^*\|, \tag{5}$$

$$\|F'(x^*)^{-1}(F'(x) - F'(y))\| \leq L \|x - y\|, \tag{6}$$

$$\|F'(x)\| \leq M_0, \tag{7}$$

$$\|F'(x^*)^{-1} F'(x)\| \leq M, \tag{8}$$

and

$$\bar{U}(x^*, r) \subseteq \mathbb{D}, \tag{9}$$

where functions p *and* φ *are defined as:*

$$p(t) = 4M(2 + M_0|\lambda|), \text{ if } \lambda \neq 0$$

or

$$p(t) = 0, \text{ if } \lambda = 0,$$

and

$$\varphi(t) = \frac{|1+b|L_0(1+g_1(t))t}{|1+a|\left(1-\left(\frac{L_0}{|1+a|}(1+|a|g_1(t))t\right)t\right)}, \quad \text{if } a+b+c+1=0$$

or

$$\varphi(t) = \frac{(|1+b|+|a+c|)M}{|1+a|\left(1-\left(\frac{L_0}{|1+a|}(1+|a|g_1(t))t\right)t\right)}, \quad \text{if } (|1+b|+|a+c|)M < |1+a|$$

where,

$$g_1(t) = \frac{1}{2(1-L_0t)}[Lt + 2M|1-\xi|]$$

Moreover, we define

$$\gamma(t) = \frac{3L_0(1+g_1(t))t}{4\left(1-\frac{L_0}{4}(1+3g_1(t))t\right)},$$

and r *is the minimum of*

- $\dfrac{2(1-M|1-\xi|)}{2L_0+L}$.
- *The smallest positive root of:*

$$3L_0(1+g_1(t))t + L_0(1+3g_1(t))t - 4 = 0$$

- *The smallest positive root of:*

$$\frac{1}{2\left(1-L_0t\right)}\left[Lt+\frac{M_0p\left(t\right)}{2\left(1-\left(\left(\frac{L_0}{4}\left(1+3g_1\left(t\right)\right)\right)t\right)\left(1-\gamma\left(t\right)\right)\right)}\right]-1=0$$

- *The smallest positive root of:*

$$\frac{L_0}{|1+a|}\left(1+|a|g_1\left(t\right)\right)t-1=0$$

- *The smallest positive root of:*

$$|1+b|L_0\left(1+g_1\left(t\right)\right)t+|1+a|\left(\frac{L_0}{|1+a|}\left(1+|a|g_1\left(t\right)\right)t\right)t-|1+a|=0.$$

- *The smallest positive root of:*

$$\left(\frac{1}{2\left(1-L_0t\right)}\left[Lt+\frac{M_0p\left(t\right)}{2\left(1-\left(\frac{L_0}{4}\left(1+3g_1\left(t\right)\right)\right)t\right)\left(1-\gamma\left(t\right)\right)}\right]\right)$$
$$\times\left[1+\frac{|\mu|M}{\left(1-L_0t\right)\left(1-\varphi\left(t\right)\right)}\right]-1=0.$$

Then, sequence $\{x_n\}$ *generated for* $x_0\in U\left(x^*,r\right)\setminus\{x^*\}$ *by method* (3) *is well defined, remains in* $\bar{U}\left(x^*,r\right)$ *for each* $n=0,1,2,\ldots$ *and converges to* x^*.

As family (3) is a parametric family, we can study the dynamics by menas of using the techniques presented in [1, 2, 3, 4, 5].

References

[1] I. K. Argyros, Á. A. Magreñán, On the convergence of an optimal fourth-order family of methods and its dynamics, Appl. Math. Comput. 252(1), 336–346 (2015).

[2] F. Chicharro, A. Cordero, J.R. Torregrosa, Drawing dynamical and parameters planes of iterative families and methods, The Scientific World Journal Volume 2013 Article ID 780153.

[3] A. Cordero, J. García-Maimó, J.R. Torregrosa, M.P. Vassileva, P. Vindel, Chaos in King's iterative family, Applied Mathematics Letters 26 (2013) 842–848.

[4] Á. A. Magreñán, Different anomalies in a Jarratt family of iterative root-finding methods, Applied Mathematics and Computation 233 (2014), 29–38.

[5] Á. A. Magreñán, A new tool to study real dynamics: The convergence plane Applied Mathematics and Computation, 248 (2014), 215–224.

[6] R. Sharma, Iterative methods for the solution of nonlinear equations (PhD Thesis), Punjab University, 2011.

A Two-level ILU preconditioner for electromagnetism applications

J. Cerdán*, J. Marín and J. Mas

Instituto de Matemática Multidisciplinar

Universitat Politècnica de València, Camí de Vera s/n 46022 Valencia Spain

1 Introduction

The numerical solution of the Maxwell's equations [7] plays a crucial role in numerous large scale industrial and scientific applications related with electromagnetism phenomena. To name a few, the computation of the antenna radiation pattern, electromagnetic interference and compatibility studies of an electrical device with their environment, and scattering problems as the computation of the radar cross-section of 3D body are important for aerospace industry.

The most common techniques for obtaining a numerical solution of Maxwell's equations can be classified either into methods that solve the differential equations or methods that consider their integral formulation. Integral equations methods (IEMs) have emerged as an attractive alternative for CEM applications. These methods solve the problem by reformulating the Maxwell's equations as a set of integral equations with equivalent sources [2]. The integral equations relate the electric and magnetic fields to the equivalent electric and magnetic currents on the surface of the object. This leads to a reduction on the dimensionality of the problem by one, and therefore allows significant reduction on the number of unknowns of the associated linear systems. Because the boundary conditions are incorporated into the surface integral equations, IEMs can handle general geometries in open domains without formulating any artificial boundary.

*e-mail: jcerdan@mat.upv.es

The integral equations are usually discretized by means of the boundary element method (BEM) or the Method of Moments (MoM) [8]. The matrices arising from IEMs are dense and expensive to solve. Since in large-scale industrial applications the size of the matrices can be very large the application of direct gaussian elimination methods is out of context, leaving the use of Krylov-type iterative methods as the only practical alternative. As it is well known, the success of an iterative method for ill-conditioned problems depends on applying a suitable preconditioning technique to the system matrix. In [1] the author show that most algebraic factorized preconditioners fail to produce good converge rates or even fail to converge. The best results on medium size problems were obtained with sparse approximate inverse preconditioners based on Frobenious norm minimization [5].

Our aim in this work is to present a technique for the compuation of ILU-type preconditioners for ill-conditioned CEM applications. The method is based on graph partitioning techniques applied to the near field matrix of the linear systems. The paper is organized as follows. In Sections 2 and 3 we review the main ideas of graph partitioning and the algorithm for computing a two-level ILU for CEM applications. Then, the numerical results are presented in Section 4.

2 Graph partitioning

The idea behind graph partitioning is the computation of a $p-$way partitioning of the adjacency graph of a matrix keeping the size of the p subgraphs balanced while minimizing to some extent the number of edges that are cut. Let A be a sparse structurally symmetric matrix. The associated undirected adjacency graph $G = (V, E)$ consists of a set nodes $V = \{1, \ldots, n\}$, one node for each row or column of the matrix, and the edge set E. There is an edge $\langle i, j\rangle$ for any matrix entry $a_{ij} \neq 0$. Note that there is not distinction between $\langle i, j\rangle$ and $\langle j, i\rangle$. For nonsymmetric sparse patterns the adjacency graph of $A + A^T$ is considered instead. We define the separator set as the group of nodes which are connected by edges that are cut in the graph partition. We also define the group of interior nodes as the nodes which are connected with the separator set. Thus, there are p groups of interior nodes, one for each subgraph of the partition. By numbering first the interior nodes and taking the separator set last, the matrix is permuted into the following block

angular form:

$$P^T AP = \begin{pmatrix} A_1 & & & & B_1 \\ & A_2 & & & B_2 \\ & & \ddots & & \vdots \\ & & & A_p & B_p \\ C_1 & C_2 & \dots & C_p & A_S \end{pmatrix} \tag{1}$$

where P is a permutation matrix. The diagonal blocks $A_1, ..., A_p$ correspond to subgraphs induced by the interior nodes in the graph decomposition, the off-diagonal blocks B_i y C_i represent the connections between interior nodes and the separator set, and A_s correspond to the subgraph induced by the separator set.

3 Two-level ILU preconditioner

Consider a linear system of n equations with n unknowns given by

$$Ax = b \tag{2}$$

obtained after the discretization of the integral form of the Maxwell's equations using the Method of Moments. The matrix A is called the impedance matrix and it is dense, non-hermitian and with complex elements. Moreover, the impedance matrix is often characterized by a large condition number which results in a slow convergence of iterative methods. A preconditioning technique consists in finding a matrix M for wich an approximate solution of the equivalent linear systems

$$M^{-1}Ax = M^{-1}b\,, \quad \text{or} \quad AM^{-1}y = b\,, \quad x = M^{-1}y \tag{3}$$

is obtained more efficiently. The matrix M is called the preconditioner. If the preconditioned matrix $M^{-1}A$ (or AM^{-1}) has a better condition number than A or its eigenvalues have a favourable distribution, one can expect an improvement of the convergence rate of the iterative method [9].

The impedance matrix relates the induced currents with the incident fields on the surface of the 3D body. Each equation represents the interaction between an edge of the mesh and its neighbourhood. In general, the magnitudes of its entries associated with the electric and magnetic field operators decrease with the distance between edges of the mesh. If we decompose

the impedance matrix into its near-field and far-field entries, equation (2) can be rewritten as

$$(A^{near} + A^{far})x = b \ , \tag{4}$$

where the near-field matrix A^{near} contains only those entries representing the interactions between source and test basis functions lying within some threshold distance. A good preconditioner should approximate the inverse of A or, at least, its near-field entries. Therefore, the preconditioner will be formulated using the matrix A^{near}. After computing a $p-$way partitioning for the adjaceny graph of A^{near} and permuting the matrix as described in the previous section, one obtains the block angular form (1) which has a block LU factorization given by

$$P^T A^{near} P = \begin{pmatrix} L_1 & & & & \\ & L_2 & & & \\ & & \ddots & & \vdots \\ & & & L_p & \\ F_1 & F_2 & \dots & F_p & L_S \end{pmatrix} \begin{pmatrix} U_1 & & & & E_1 \\ & U_2 & & & E_2 \\ & & \ddots & & \vdots \\ & & & U_p & E_p \\ & & \dots & & U_S \end{pmatrix} , \tag{5}$$

where $A_i = L_i U_i$, $E_i = L_i^{-1} B_i$, $F_i = C_i U_i^{-1}$. The matrices L_S and U_S are the triangular factors of the Schur complement matrix

$$S = A_S - \sum_{i=1}^{p} C_i A_i^{-1} B_i. \tag{6}$$

From (5) an incomplete factorization $P^T A^{near} P$ is obtained by computing a ILU decomposition for each diagonal block, i.e., $A_i \approx \hat{L}_i \hat{U}_i$. Moreover, since $A_i^{-1} \approx U_i^{-1} L_i^{-1}$, it also follows from equation (6) that an approximation of the Schur complement matrix S is computed and factorized as $\hat{S} \approx \hat{L}_S \hat{U}_S$. This double factorization and approximation characterizes the two-level nature of the algorithm. The ILU factorization described is used as a preconditioner for the iterative solution of the permuted linear system

$$(P^T A P)y = P^T b, y = P^T x,$$

where A is the full matrix, i.e., $A = A^{near} + A^{far}$.

4 Numerical experiments

In this section we show the results of the numerical experiments obtained for a set of model problems which are listed in Table 1. All matrices were kindly provided to us by the EADS-CASA company. A threshold distance of $0,04$ meters was used to filter the impedance matrix A. All codes developed for the tests were written in FORTRAN 95 in double precision complex arithmetic, and have been compiled with Intel Fortran Composer XE 2013 and linked with the Intel Math Kernel Library. For the experiments we used one AMD Opteron CPU of a Sun Fire X2200 M2 Server. The GMRES(m) method [10] with right preconditioning was used to solve the linear systems. The iterative method was stopped when the initial residual was reduced by at least six orders of magnitude. The METIS software package [6] was used to obtain a $p-$way partitioning of the near-field adjacency graph. The ILUT algorithm was used to compute the incomplete factorizations [9]. In our tests only the threshold option was used to reduce the fill-in and it is indicated in the caption of the tables.

The numerical experiments were conducted to show the effect of the number of partitions on the quality of the ILU preconditioner, the effect of applying reorderings to the diagonal blocks of the permuted near-field matrix before computing the incomplete factorization, and finally a comparison of the proposed ILU preconditioner with SPAI which is widely used in CEM applications.

Table 1: Tested matrices

Matrix	n	$\rho(A^{near})$
CETAF3	3097	0.02
CETAF5	5021	0.01
CETAF10	10022	0.01
CN2	2038	0.05
CN3	3020	0.02
CN5	5005	0.01

It is well known that incomplete factorizations for nonsymmetric matrices can benefit from symmetric reorderings applied to the coefficient matrix,

such as fill-in reducing orderings and level set reorderings. Therefore, we tested the effect of the multiple minimum degree, quotient minimum degree, reverse Cuthill-McKee and nested dissection [3, 4]. In table 2 we refer to them as *mmd*, *qmd*, *rcm* and *nd*, respectively, whereas *no* indicates natural ordering. This table shows the effect of the symmetric reorderings for the matrix CETAF3, and the effect of the number of partitions of the adjacency graph of the near field matrix. For the rest of matrices a similar behaviour was observed. The parameter p indicates the number of partitions, *Rmethod* is the symmetric reordering method applied to the diagonal blocks and Schur complement matrix, *Tr* is the time spent to compute these reorderings, ρ is the density of the preconditioner with respect the number of nonzeros of A^{near}, *iter* is the number of iterations, *Tp* and *Tsol* are the preconditioner computation time and iterative solution time, respectively. All the timings correspond to the CPU time in seconds. Note that the value $p = 1$ indicates that the ILUT preconditioner was computed for the full matrix, i.e., without the two-level approach proposed.

First, with respect the effect of symmetric reorderings we observe that applying the minimum degree or reverse Cuthill-Mckee produces sparser preconditioners and, at the same time, decreases the number of iterations needed to converge. In general we found that multiple minimum degre performed the best, although no big differences compared with quotient minimum degree or reverse Cuthill-Mckee were found. Moreover, the time needed to compute and apply this preprocessing to the matrix is very small compared with the overall time solution. Second, we note that the two-level ILUT preconditioner increases the convergence rate and also helps to obtain sparser preconditioners. The total time was reduced considerably for almost all the tested matrices. Moreover, a nice feature is that the number of iterations almost remain constant with the number of partitions. Finally, it is worth to mention that a SPAI preconditioner with the same nonzero density spent 9.9 seconds and 169 iterations to converge. The computation time was 40 seconds for CETAF3. We found for all the matrices tested that the two-level ILUT preconditioner allways converged faster in time and number of iterations. Taking into account that computing and incomplete LU is also considerably cheaper, we think that the two-level ILU preconditioner proposed is a competitive alternative for CEM applications.

Table 2: Effect of symmetric reorderings for the matrix CETAF3, GMRES(200), ILUT(0.04).

p	Rmethod	Tr	ρ	Tp	Tsol	iter
1	no	-	4.7	4.4	81.2	796
	mmd	0.4	3.2	1.9	53.8	570
2	no	-	2.7	3.7	8.1	112
	mmd	0.4	2.3	3.4	10.4	107
	qmd	0.4	2.5	3.2	9.1	115
	rcm	0.4	2.1	2.7	8.0	111
	nd	0.4	2.6	2.5	10.1	117
3	no	-	2.5	4.3	10.9	110
	mmd	0.4	2.1	3.0	9.0	104
	qmd	0.5	2.2	3.5	9.9	104
	rcm	0.4	2.0	2.9	9.7	103
	nd	0.5	2.3	3.8	11.7	115
4	no	-	2.5	2.2	10.5	110
	mmd	0.5	2.0	2.1	8.0	107
	qmd	0.5	2.3	3.3	10.9	110
	rcm	0.5	2.0	2.9	9.7	102
	nd	0.5	2.4	2.1	11.7	112
5	no	-	2.4	4.4	10.4	106
	mmd	0.4	2.1	3.2	7.8	103
	qmd	0.4	2.2	3.0	7.8	104
	rcm	0.4	2.2	2.3	7.5	103
	nd	0.4	2.3	3.4	11.2	112

References

[1] B. Carpentieri. *Sparse preconditioners for dense linear systems, from electromagnetic applications*. PhD thesis, PhD Thesis, l'Institut National Polytechnique de Toulouse, CERFACS, 2002.

[2] W. C. Chew, J.-M. Jin, E. Michielssen, and J. Song. *Fast and Efficient Algorithms in Computational Electromagnetics and Elastic Waves*. Morgan & Claypool Publishers, 2009.

[3] I. S. Duff, A. M. Erisman, and J. K. Reid. *Direct Methods for Sparse Matrices*. Oxford University Press, London, 1986.

[4] A. George and J. W. H. Liu. *Computer Solution of Large Sparse Positive Definite Systems*. Prentice-Hall, Englewood Cliffs, NJ., 1981.

[5] M. J. Grote and T. Huckle. Parallel preconditioning with sparse approximate inverses. *SIAM J. on Scientific Computing*, 18(3):838–853, 1997.

[6] G. Karypis and V. Kumar. METIS: A software package for partitioning unstructured graphs, partitioning meshes, and computing fill-reducing orderings of sparse matrices (version 3.0). Technical report, University of Minnesota, Department of Computer Science and Army HPC Research Center, October 1997.

[7] J. C. Maxwell.. A dynamical theory of the electromagnetic field. 155:459–512, 1865.

[8] S. M. Rao, D. R. Wilton, and A. W. Glisson. Electromagnetic scattering by surfaces of arbitrary shape. *IEEE Trans. Antennas Propagat.*, AP-30:409–418, 1982.

[9] Y. Saad. *Iterative Methods for Sparse Linear Systems*. PWS Publishing Co., Boston, 1996.

[10] Y. Saad and M. H. Schulz. GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM Journal on Scientific and Statistical Computing*, 7:856–869, 1986.

Numerical Simulation of Needle Movement Nozzle Flow Coupled with Spray for a Diesel Injector Using an Eulerian Spray Atomization Model

R. Payri[♭], J. Gimeno[♭], P. Martí-Aldaraví[♭]*, M. Alarcón[♭]

(♭) CMT - Motores Térmicos, Universitat Politècnica de València,

Edificio 6D, Camino de Vera s/n, 46022, Valencia, Spain.

November 30, 2015

1 Introduction

The study of flow characteristics inside Diesel nozzles is tricky due to their small characteristic length and hard testing conditions (very high pressure and velocity). That is the reason why Computational Fluid Dynamics (CFD) is widely used to study such kind of flows [1, 2]. These techniques model the real domain generally by finite volume discretization. Therefore, the geometry of the nozzle must be very well known in advance.

In order to obtain geometrical details of Diesel nozzle, silicone mold technique [3] has been widely used over the years. Nevertheless, new powerful techniques such as X-ray tomography [4] have been also successfully employed. In fact, this last technique not only allows obtaining the sizing of the nozzle with a resolution of microns, but also the time resolved position of the needle and measurements in the dense region of the spray [5]. It is known that needle movement affect spray behavior (and so combustion efficiency) specially during opening and closing transients [6, 7].

*e-mail: pedmar15@mot.upv.es

From another point of view, the in-nozzle flow pattern defines the spray behavior and characteristics [8]. Though nozzle flow and spray are clearly linked. Under this framework, the objective of this work is to employ an already existing Eulerian model, able to simulate nozzle flow and spray at the same time [10], to computationally address the effects of needle movement on the spray behavior. To do so, Spray A conditions of Engine Combustion Network (ECN) [11] are considered. Nozzle of Spray A is very well characterized in terms of geometry and flow characteristics, including needle lift measurements. However, the time resolution used for measuring needle velocity may not be high enough, then it is possible that some post-processing (i.e. moving average) of the experimental signal is required.

2 Methodology

The Eulerian Spray Atomization (ESA) is a homogeneous model built to simulate non-evaporative conditions so far [10], and it has been deeply described in previous works [10, 12, 13]. In this type of models, the air-fuel mixture is considered as a single phase and the dispersion is calculated through a balance equation [9].

Simulations (which are peformed in OpenFOAM 2.3.0 ®) set-up is similar to the one configured in previous works [10, 12], being the main difference the needle position and movement. The needle lift was measured by means of X-ray tomography [4]. The velocity of the needle has a peaky shape, which may lead to unstable or unrealistic simulations, so the signal has been filtered using a moving average filter. By comparing simulations perfomed with these two needle velocity laws (original and filtered) it is possible to check if this has an impact on the spray behavior. However, the combination of needle movement and time varying pressure at inlet boundary leads to mass flow rate curves quite different from experimental ones. Therefore, an additional simulation is performed in which the pressure is fixed and the non-filtered needle velocity is used.

OpenFOAM 2.3.0 ® supports mesh morphing six degree of freedom. However, for this work only the axial movement is needed. Mesh motion is based on solving the cell-center Laplacian for the given component of the motion velocity, and cells are deformed accordingly in all directions to keep cell quality (skewness, aspect ratio) under acceptable levels.

3 Results and discussion

First parameters to be analyzed are the mas flow rate and momentum flux, together with the non-dimensional coefficients that define the nature of the flow: C_v, C_a and C_d, following Desantes et al. [14] procedure. All simulations predict almost the same value for the discharge coefficient C_d, 2% lower than the experimental. This is due to the under-prediction of the velocity coefficient, which means that simulations give a velocity slightly lower than experiments. This difference may be related to uncertainties on reference density value. Nevertheless, differences are small.

Comparing mass flow rate and momentum flux signals, it can be observed that time varying pressure inlet boundary by itself is not enough to obtain the oscillations in mass flow and momentum experimentally observed. In fact, when that condition is combined with needle movement, oscillations grow but do not follow experimental data. This means that pressure signal on the rail should not be used as boundary condition in internal nozzle flow simulations.

Spray parameters need to be also analyzed. Simulation with fixed needle slightly over-estimates the near-field spray penetration, whilst simulations with moving needle give quite similar curves than experiments. Both simulation curves overlaps, one more evidence that filtering the needle velocity may not be needed.

4 Conclusions

Spray A conditions have been simulated by using the ESA model. Needle movement has been taken into account under different approaches: fixed needle, the raw experimental data for needle lift, and filtered signal to remove possible instabilities. The following conclusions can be drawn from the current work:

- No clear difference has been observed between raw lift data and the filtered signal. There is no need of post-processing the needle lift data. Although the acquisition frequency seemed to be low, is enough for performing simulations.

- Needle movement is needed to capture oscillations experimentally observed in mass flow rate and momentum flux. However, it must be

combined with the right inlet boundary condition and the proper initialization.

- Errors in spray penetration are minimized if needle movement is considered.
- Spray structure is generally well predicted, except that ESA computational model over-predicts fuel concentration for distances longer than 5 mm.

Therefore, when the needle lift is available, it is recommended to consider it in the simulations to improve predictions in both nozzle and spray parameters. Special attention must be paid to inlet boundary condition and initialization of simulations. Nonetheless, there is still work to do to improve Diesel spray models.

References

[1] F. J. Salvador, J.-V. Romero, M.-D. Roselló and D. Jaramillo, *Numerical simulation of primary atomization in diesel spray at low injection pressure*, Journal of Computational and Applied Mathematics, In Press, Available on-line, (2015).

[2] G. Strotos, P. Koukpuvinis and A. Theodorakakos, *Transient heating effects in high pressure Diesel injector nozzles*, International Journal of Heat and Fluid Flow, 51 (2015), pp. 257–267.

[3] V. Macian, V. Bermúdez, R. Payri and J. Gimeno, *New technique for determination of internal geometry of a Diesel nozzle with the use of silicone methodology*, Experimental Techniques, 39 (2003), pp. 39–43.

[4] A. L. Kastengren, F. Z. Tiloco, C. F. Powell, J. Manin, L. M. Pickett, R. Payri and T. Bazyn, *Engine Combustion Network (ECN): Measurements of Nozzle Geometry and Hydraulic Behavior*, Atomization and Sprays, 22 (2012), pp. 1011–1052.

[5] A. L. Kastengren, F. Z. Tiloco, D. J. Duke, C. F. Powell, X. Zhang and S. Moon, *Time-resolved X-Ray Radiography of Sprays from Engine Combustion Network Spray A Diesel Injectors*, Atomization and Sprays, 24 (2014), pp. 251–272.

[6] W. G. Lee and R. D. Reitz, *A Numerical Investigation of Transient Flow and Cavitation within Minisac and VCO Diesel Injector Nozzles*, Proceedings of the 2009 Spring Technical Conference of the ASME Internal Combustion Engine Division (2009), pp. 643–653.

[7] R. Payri, J. Gimeno, J. P. Viera and A. H. Plazas, *Needle lift profile influence on the vapor phase penetration for a prototype diesel direct acting piezoelectric injector*, Fuel, 113 (2013), pp. 257–265.

[8] V. Bermdez, R. Payri, F. J. Salvador and A. H. Plazas, *Study of the influence of nozzle seat type on injection rate and spray behavior*, IMechE. Part D: Journal of Automobile Engineering, 219 (2005), pp. 677–689.

[9] A. Vallet, A. A. Burluka and R. Borghi, *Development of a Eulerian Model for the "Atomization" of a Liquid Jet*, Atomization and Sprays, 11 (2001), pp. 619–642.

[10] F. J. Salvador, J. Gimeno, J. M. Pastor and P. Martí-Aldaraví, *Effect of turbulence model and inlet boundary condition on the Diesel spray behavior by an Eulerian Spray Atomization (ESA) model*, International Journal of Multiphase Flow, 65 (2015), pp. 108-116.

[11] M. Bardi, R. Payri, L. M. Malbec, G. Bruneaux, L. M. Pickett, J. Manin, T. Bazyn and C. Genzale, *Engine Combustion Network: Comparison of Spray Development, Vaporization, and Combustion in Different Combustion Vessels*, Atomization and Sprays, 22 (2012), pp. 807-842.

[12] J. M. Desantes, R. Payri, J. Gimeno and P. Martí-Aldaraví, *Simulation of the First Millimeters of the Diesel Spray by an Eulerian Spray Atomization Model Applied on ECN Spray A Injector*, SAE Technical Paper 2014-01-1418 (2014).

[13] R. Payri, S. Ruiz, J. Gimeno and P. Martí-Aldaraví, *Verification of a new CFD compressible segregated and multi-phase solver with different flux update-equations sequences*, Applied Mathematical Modelling, 39 (2015), pp. 851–861.

[14] J. M. Desantes, R. Payri, J. M. Pastor and J. Gimeno, *Experimental characterization of internal nozzle flow and diesel spray behavior. Part 1: Non-evaporative conditions*. Atomization and Sprays, 17 (2005), pp. 315–345.

Study of the influence of momentum interpolation methods in the SIMPLE algorithm for incompressible flows by an open-source CFD code

J. Martínez[♭] *, F. Piscaglia[♭], A. Montorfano[♭], A. Onorati[♭], and S. M. Aithal[†]

(♭) Politecnico di Milano,

Dip. Energia, Via Lambruschini 4, 20156, Milano-Italy,

(†) Argonne National Laboratory,

Lemont, IL 60439, United States.

November 30, 2015

1 Introduction

Finite-volume methods have been extensively applied in the recent years for the solution of incompressible Navier-Stokes equations. From the point of view of variables storage two options may be considered, either a staggered-grid or a co-located-grid arrangement. While the former has proved to improve stability and robustness of the solver, the latter significantly diminishes the memory requirements, causing important reduction on computational time. As a result, co-located grids have gained popularity in both general-purpose and commercial flow solvers.

The most significant problem derived from the use of the co-located grid arrangement is the so called checker-board pressure fields. Application of central-difference spatial discretization to the pressure gradient term in the velocity equations produces a decoupling of pressure and velocity cell values, leading to saw-tooth oscillations. To avoid this problem, Rhie and Chow [1] proposed in 1983 a momentum-based interpolation of mass fluxes on cell faces, imitating the staggered-grid discretization (Original Momentum Interpolation Method). This technique removes the checker-boarding pressure in the most part, which is the reason of its wide acceptance.

However Rhie-Chow OMIM presents some additional problems. In particular, Majumdar [2] and Miller et al. [3] independently reported that solutions obtained with the original Rhie-Chow interpolation method are underrelaxation factor-dependent. Besides, the use of very small under-relaxation factors could make the checker-board pressure reappear.

*e-mail: jmartrubio@gmail.com

In parallel, for the solution of incompressible steady-state flow problems, several algorithms have been proposed during the years, to deal with the pressure-velocity coupling of Navier-Stokes equations. In particular SIMPLE (Semi-Implicit Method for Pressure Linked Equations), SIMPLER (SIMPLE-Revised), SIMPLEC (SIMPLE-Consistent) and PISO (Pressure Implicit with Splitting of Operators) are some the most well known algorithms.

The aim of the present paper is to apply the method for mass fluxes interpolation proposed by Majumdar in combination with different pressure-velocity coupling algorithms, to study the overall solver performance in terms of accuracy of the solution, dependency on the under-relaxation factors and convergence rate. Two simple 2D cases have been selected with this purpose: a laminar lid-driven cavity and a NACA-0012 aifoil profile.

2 Numerical methodology

The open-source finite-volume CFD code OpenFOAM®, extended with the implementation of Majumdar correction and alternative pressure-velocity coupling algorithms (SIMPLE-R, SIMPLE-C), has been used for this study. The two cases selected for the tests tried to evaluate the performance of the correction in a theoretical simple laminar case (cavity), where an orthogonal quadrilateral mesh or a triangle mesh with low non-orthogonality was used, and a more complex case (airfoil) with a non-orthogonal mesh where turbulence was present and lack of convergence of results might have limited the influence of the correction. For each of the two cases the numerical procedure followed is described below.

2.1 Lid-driven cavity

A laminar lid-driven cavity case was simulated using either a uniform orthogonal mesh, a non-uniform orthogonal mesh or a triangle mesh as shown in Fig. 1.

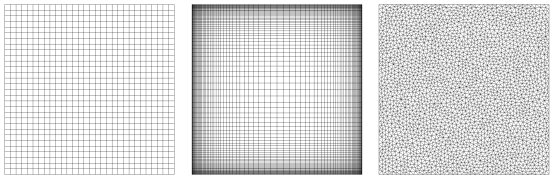


Figure 1: 2D meshes chosen for cavity case

Two Reynolds numbers, namely 1000 and 5000, were used for the simulations. Linear interpolation was used for convective and diffusive terms. Convergence was assumed to be achieved when normalized residuals of both components of velocity and pressure equations was below $1.0 \cdot 10^{-8}$. Benchmark solutions from Ghia et al. [4] have been used to compare vertical velocity profiles along the horizontal center-line.

2.2 NACA 0012 airfoil profile

The turbulent flow ($Re \approx 6 \cdot 10^6$ based on the chord length) around a NACA 0012 airfoil profile was studied under incompressible conditions ($M = 0.15$). A structured grid of 14625 cells (225×65 with 129 points on airfoil surface) was employed, giving an approximate average $y^+ \approx 1$ over the airfoil profile (see Fig. 2). Numerical results have been compared to experimental references, including pressure distribution [5], lift and drag [6] coefficients on upper wall, for different angles of attack.

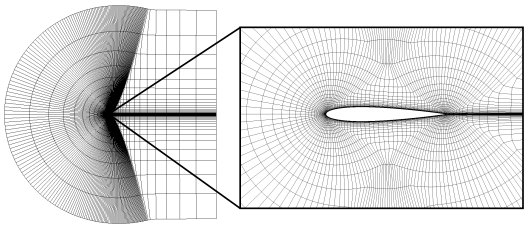


Figure 2: 2D mesh of NACA 0012 profile and zoom to airfoil region

3 Results

As expected, the dependency of the solution of OMIM with respect to the velocity under-relaxation is quickly reduced as mesh refinement increases. Majumdar correction completely eliminates the dependency, at the cost of reducing the accuracy of the solution. It can be proved that Majumdar solution tends to that obtained by OMIM when velocity under-relaxation factor approaches 0. In fact, if the error in the determination of vertical velocity in one point is plotted with respect to the number of cells of the simulation (see Fig. 3), we can see how Majumdar solution is completely independent of under-relaxation factors but generates always a bigger error than standard OMIM for any value of velocity equations under-relaxation factor.

When comparing the performance of different pressure-velocity coupling algorithms, the use of SIMPLE-C significantly reduces the number of iterations required to arrive to a given threshold in the equations residuals, when compared to other classical algorithms (SIMPLE, SIMPLE-R, PISO) .

Given the slight dependency of results for fine meshes, NACA 0012 was used to test how important these small variations in velocity profiles affect the prediction of pressure (C_p), drag (C_D) and lift coefficient (C_L). As shown in Fig. 4, for those parameters mainly affected by pressure distribution (C_p, C_L) no significant differences are found when OMIM is used for different under-relaxation factors. However, for drag coefficients, where velocity derivatives become of importance, some differences might appear. It was checked that Majumdar correction eliminates most part of these dependencies for the level or convergence achieved in the simulations.

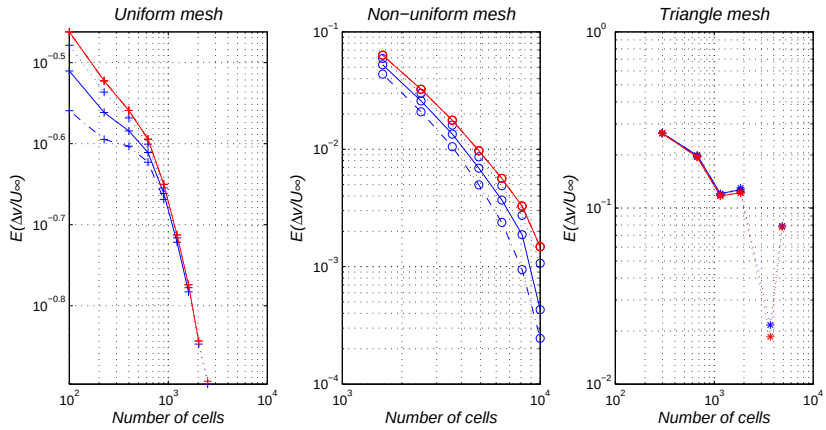


Figure 3: Error in the determination of v-component of velocity at $x = 0.07765$, $y = 0.5$ for different meshes: uniform mesh (+); Non-uniform mesh ($\circ$); Triangle mesh ($*$); and for different under-relaxation factors: $\alpha_u = 0.8$ ($-\cdot-$); $\alpha_u = 0.5$ (—); $\alpha_u = 0.2$ ($\cdots$). Results are shown for standard OMIM (blue) and Majumdar correction (red).

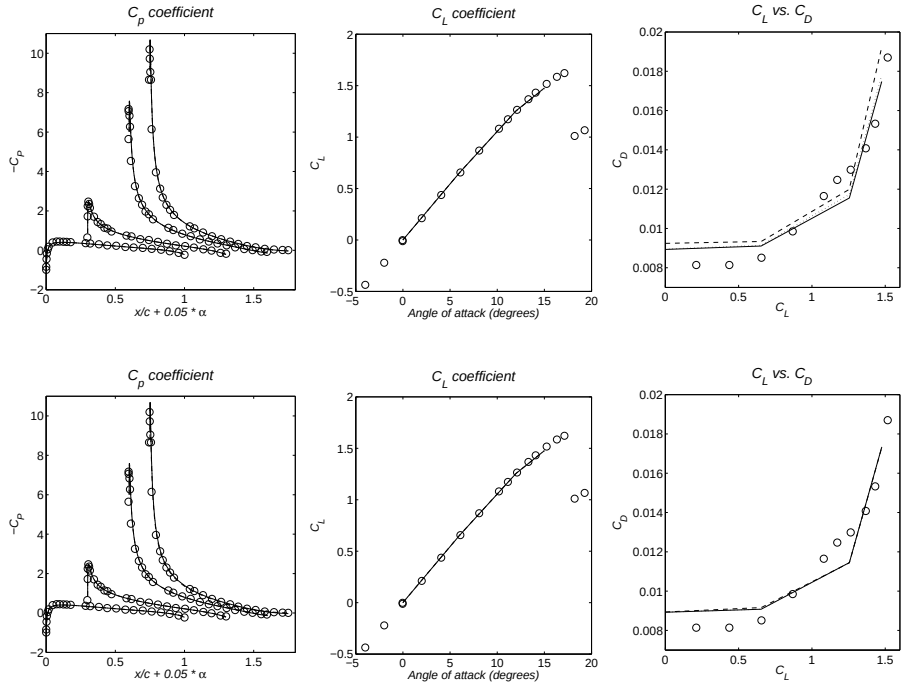


Figure 4: C_p, C_L and C_D coefficients for OMIM (top) and Majumdar correction (bottom). Results are shown for: $\alpha_u = 0.1$ (- - -); $\alpha_u = 0.5$ ($\cdots$); $\alpha_u = 0.8$ (—)

4 Conclusions

In the present paper Majumdar correction and two alternative pressure-velocity coupling algorithms (SIMPLE-C and SIMPLE-R) have been implemented in OpenFOAM®. Majumdar correction has proved to eliminate the dependency of the solution on velocity under-relaxation factors. While the dependency of OMIM is reduced when a fine mesh is used, its influence in coefficients depending on velocity derivatives should still be considered. For the cases studied, SIMPLE-C algorithm has proved to improve the convergence speed with respect to other classical algorithms (SIMPLE, SIMPLE-R or PISO).

References

[1] C.M. Rhie and W. L. Chow. A numerical study of the turbulent flow past an isolated airfoil with trailing edge separation *AIAA J.*, vol. 21, pp. 1525-1532, 1983

[2] S. Majumdar Role of underrelaxation in momentum interpolation for calculation of flow with nonstaggered grids *Numerical Heat Transfer*, vol. 13, pp. 125-132, 1988

[3] T.F. Miller and F.W. Schmidt Use of a Pressure-Weighted Interpolation Method for the Solution of Incompressible Navier-Stokes Equations on a Non-Staggered Grid System *Numerical Heat Transfer*, vol. 14, pp. 213-233, 1988

[4] U.Ghia, K. N. Ghia and C. T. Shin High-Re Solutions for Incompressible Flow using the Navier-Stokes Equation and a Multigrid Method *Journal of Computational Physics*, vol. 48, pp. 387-411, 1982

[5] N. Gregory and C. L. O'Reilly Low-Speed Aerodynamic Characteristics of NACA 0012 Aerofoil Sections, including the Effects of Upper-Surface Roughness Simulation Hoar Frost *NASA R&M* 3726, 1970

[6] C. L. Ladson, A. S. Hill and W. G. Johnson Jr. Pressure Distributions from high Reynolds Number Transonic Tests of a NACA 0012 Airfoil in the Langley 0.3 Meter Transonic Cryogenic Tunnel *NASA TM* 100826, 1987

Improved railway wheelset-track interaction model in the high–frequency domain

José Martínez-Casas[1], Juan Giner-Navarro[1], F. D. Denia[1], Paloma Vila[1] and L. Baeza[1]

[1]Centro de Investigación en Ingeniería Mecánica, Universitat Politècnica de València, Camino de Vera s/n, 46022 Valencia, Spain.

e-mail address of corresponding author: jomarc12@mcm.upv.es

ABSTRACT

As it is well known, there are various phenomena related to railway train-track interaction, some of them caused by the high–frequency dynamics of the system, such as rolling noise, squeal noise and short-pitch rail corrugation. Due to these phenomena and some others unsolved so far, a large effort has been made over the last 40 years in order to define suitable models for studying train-track interaction. In most of the early works the vehicle was modelled as a rigid body system. Later, the introduction of the wheelset flexibility in the model was required to have a more realistic representation of the wheel–rail interaction effects at higher frequencies. Recently, more advanced train-track interaction models have been developed for vehicle running on straight track and negotiating a curved track, where the rails are modelled by means of Timoshenko beam elements with cyclic boundary conditions. Thus, the rail model is valid up to 1.5 kHz for lateral vibration and up to 2 kHz for vertical vibration. Therefore, this confines the frequency range of validity for the complete train-track model to 1.5 kHz.

The aim of this paper is to propose an improvement of the train-track interaction models so as to extend their range of validity above 1.5 kHz. With this purpose, a track model based on the 3D Moving Element Method (MEM) is developed to replace the Timoshenko beam considered in earlier studies, adopting cyclic boundary conditions and Eulerian coordinates. The MEM approach considers a mobile finite element (FE) mesh which moves with the vehicle speed, so the mass of the rail 'flows' with the velocity of the vehicle but in the opposite direction through the mesh. The MEM permits to fix the contact area in the middle of the finitely long track and, hence, to refine longitudinally the mesh only around the contact area. This Eulerian-modal approach reduces substantially the number of degrees of freedom of the problem whereby, together with a refinement just around the fixed contact area, the computational cost of the proposed model is much lower than FEM models widely used. Simulation results for the proposed modelling approach are presented and discussed for different excitation sources including random rail roughness and singularities such as wheel flats. All the simulation cases are made for a Timoshenko beam and a 3D MEM rail models, in order to point out the differences in the contact forces above the range of validity of the Timoshenko beam.

1. INTRODUCTION

The complexity of the train-track interaction comes from the vibration coupling between the railway vehicle and a flexible track, in which wheel–rail contact forces couple both sub-systems and their surface imperfections, such as rail roughness and wheel out-of-roundness, excite the global system. Unwanted phenomena such as damage of the rolling surfaces in the form of high levels of noise and vibration [1], corrugation [2], rolling contact fatigue [3] and dynamic stresses may appear in some cases due to large level of vibration and large dynamic fluctuations of the contact forces, which need to be carefully considered in order to avoid failures due to metal fatigue.

Accordingly, many suitable train-track interaction models have been developed over the last 40 years. Finite Element (FE) models have strongly entered in railways research to extend the frequency range above 1 kHz to address the previous phenomena, but only very recently, further works have considered the flexibility and inertial effects due to wheelset rotation [4]. In order to widen the range of valid frequencies, this wheelset model is used in this paper, which takes into account the gyroscopic and inertial effects associated with the rotation by using Eulerian modal coordinates and, hence, reducing the computational cost.

In recently published train-track interaction models, the rails are modelled by means of Timoshenko beam elements with cyclic boundary conditions [5], thus the rail model being valid up to 1.5 kHz for lateral vibration and up to 2 kHz for vertical vibration according to [1]; this confines the frequency range of validity for the complete train-track model to 1.5 kHz. The aim of this paper is, hence, to propose an improvement of the train-track interaction model [4] so as to extend their range of validity above 1.5 kHz. With this purpose, a track model based on the Moving Element Method (MEM) technique is developed to replace the Timoshenko beam considered in earlier studies, adopting cyclic boundary conditions and Eulerian coordinates.

The proposed model is based on the Moving Element Method (MEM) presented by Koh *et. al.* [6, 7] for 1D and 2D beam tracks. In the present work, this Eulerian approach has been extended to a straight 3D real track extruded from the UIC-60 profile (3D MEM model). The Eulerian coordinate system is attached to the moving vehicle instead of a fixed coordinate system. Hence, the mesh is moving with this mobile frame and consequently the material of the rail 'flows' into this mesh. This relative motion requires considering the material derivative for the formulation of the rail dynamics. In this model, the contact area is fixed on the same elements, allowing refine the mesh just around the fixed contact area. Therefore, the 3D MEM model permits to decrease the computational cost compared to the FEM models commonly employed.

Results for the proposed modelling approach are presented and discussed for a selected vehicle type and a straight track for different excitation sources including random rail roughness and singularities such as wheel flats. All the simulation cases are made for a Timoshenko beam and a 3D MEM models representing the track. The vertical contact force for both model are analysed and compared with special attention to discrepancies above the range of validity of the Timoshenko beam (from 1.5 to 8.5 kHz).

2. THE VEHICLE–TRACK INTERACTION MODEL

For the vehicle-track interaction model, a sub-structuring technique [5] is followed in this paper, permitting to divide the whole system into three substructures: the vehicle, the rails and the discrete rail supports, all coupled by the wheel-rail contact forces and the forces generated at the rail pads. The vehicle is confined to one wheelset with primary suspension, and two static forces on the primary suspensions are included to represent the weight of the bogie and the car body. The wheelset has been modelled as a flexible and rotating wheelset [4], in which Eulerian-modal coordinates are employed due to the axial symmetry of the body. This modal approach leads to a significant reduction of the computation cost of the simulation.

In the present paper, the track is included by two different flexible models: Timoshenko beam model [5], with a limited frequency range of validity (up to 1.5 kHz for lateral vibration and up to 2 kHz for vertical vibration), and the 3D Moving Element Method (MEM) proposed in this work with the purpose of extending the range to the high-frequency domain. A cyclic approach is considered in both cases since it provides some benefit with respect to classical track modelling or the finite one.

For the wheel-rail contact forces in the contact patch, Hertzian model is adopted here for the normal contact and FASTSIM is the software used for the tangential contact that depends on the normal contact force and of the creepages. The contact model considers the real geometry of the contacting profiles and includes the non-linearities between creepages and creep forces The discrete rail supports are introduces in the form of lumped parameter systems, where the rail pads are modelled as lumped viscoelastic elements, sleepers as lumped masses, and the ballasts as equivalent stiffness and damping below the lumped masses.

3. 3D MOVING ELEMENT METHOD (MEM)

The Moving Element Method (MEM) was presented first by Koh *et. al.* [6, 7] as an efficient Eulerian approach to treat moving loads applied on finite beams. This section summarise the effort of extending this technique to a 3D finite track with cyclic boundary conditions at the edges aiming to widen the frequency range of validity of the Timoshenko beam model used in previous works. Taking into account the material derivative and convective terms, and using the virtual work principle and modal superposition, the equation of motion is obtained below:

$$\tilde{\mathbf{M}}\ddot{\mathbf{q}} - 2V\tilde{\mathbf{C}}\dot{\mathbf{q}} + \left(\tilde{\mathbf{K}} - V^2\tilde{\mathbf{A}}\right)\mathbf{q} = \mathbf{Q}, \tag{1}$$

where V is the velocity of the vehicle, $\mathbf{q}$ is the Eulerian-modal coordinate vector, $\mathbf{Q}$ is the vector of generalised forces from contact and rail pad forces, and the matrices are the following:

$$\tilde{\mathbf{M}} = \int_{Vol} \rho\, \mathbf{\Phi}^{\mathrm{T}}\mathbf{\Phi}\, d\upsilon, \qquad \tilde{\mathbf{C}} = \int_{Vol} \rho\, \mathbf{\Phi}^{\mathrm{T}} \frac{\partial \mathbf{\Phi}}{\partial x}\, d\upsilon, \qquad \tilde{\mathbf{A}} = \int_{Vol} \rho\, \frac{\partial \mathbf{\Phi}}{\partial x}^{\mathrm{T}} \frac{\partial \mathbf{\Phi}}{\partial x}\, d\upsilon, \tag{2}$$

being $\mathbf{\Phi}$ the mode shape functions matrix of the cyclic rail, $\tilde{\mathbf{M}}$ the modal mass matrix, $\tilde{\mathbf{K}}$ the modal stiffness matrix, $\tilde{\mathbf{C}}$ introduces the inertial force due to the convective velocity, and $\tilde{\mathbf{A}}$ contains the inertial force due to the convective acceleration.

The modal differential equation of motion of the rail in Eq. (1) is linear and then the matrices are only calculated one time at the beginning of the simulation. Therefore, the proposed 3D MEM model is suitable in the high–frequency domain and has a low computational cost.

4. RESULTS

In this section, results of wheelset-track interaction simulations for the proposed modelling approach are presented considering different sources of excitation: random rail roughness and wheelflat. All the simulations consider the vehicle running at 300 km/h through a tangent track; each case is run separately for a Timoshenko beam and a 3D MEM modelling the track. Both models are compared by the vertical contact force obtained from the simulations above the range of validity of the Timoshenko beam (from 1.5 to 8.5 kHz) in order to study the contributions of the 3D MEM in the high–frequency range.

Fig. 1 shows the time history of the vertical contact force for an excitation caused by randomly corrugated rails, assuming a corrugation spectrum corresponding to the ISO 3095 limit, which establishes a third-octave band spectrum of the rail roughness. As expected, the dynamic fluctuations of the vertical contact force show a complex waveform, arising from the dynamic response of the wheelset-track system to broadband random excitation. The results for both track models present a similar trend and mean values, although a higher frequency content is observed for the MEM.

An alternative plot for the wheel-rail contact force can be made in the frequency domain. Fig. 2 represents the third-octave band spectrum of the vertical contact forces (in dB). There appear peaks in the antiresonance of the track frequency (between the P2 and pinned-pinned frequencies, in the 100-300 Hz band), whereas the smaller responses are at the P2 (below 100 Hz) and pinned-pinned (below 1 kHz) frequencies. It can be observed a similar behaviour in low-medium frequency range for both track model up to 1.5 - 2 kKz (range of validity for the Timoshenko beam), but MEM generally shows a higher frequency content for higher frequencies. These higher harmonics seem to be crucial to describe the high frequency noise, as rolling noise or squeal noise.

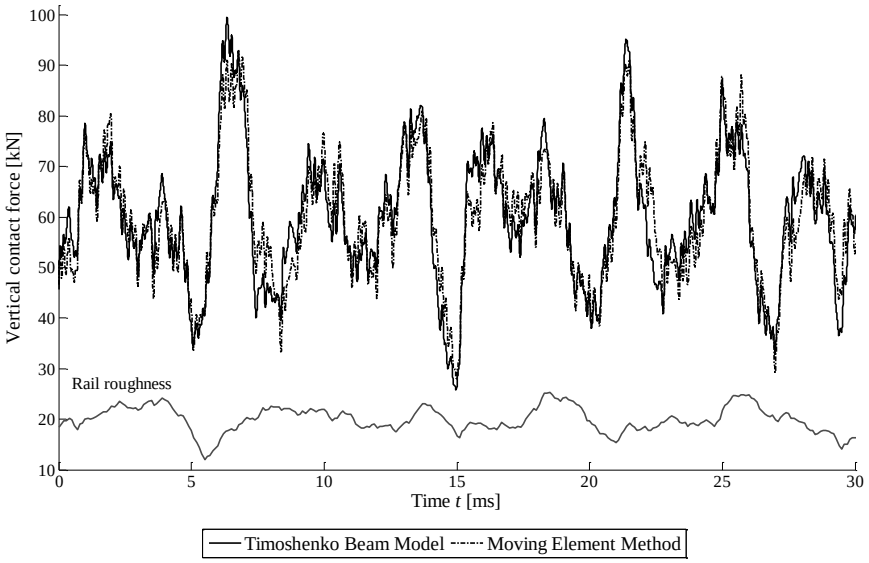


Fig. 1. Vertical wheel-rail contact forces when the vehicle circulates at 300 km/h speed on a randomly corrugated tangent track. Amplitudes corresponding to the ISO 3095 limit.

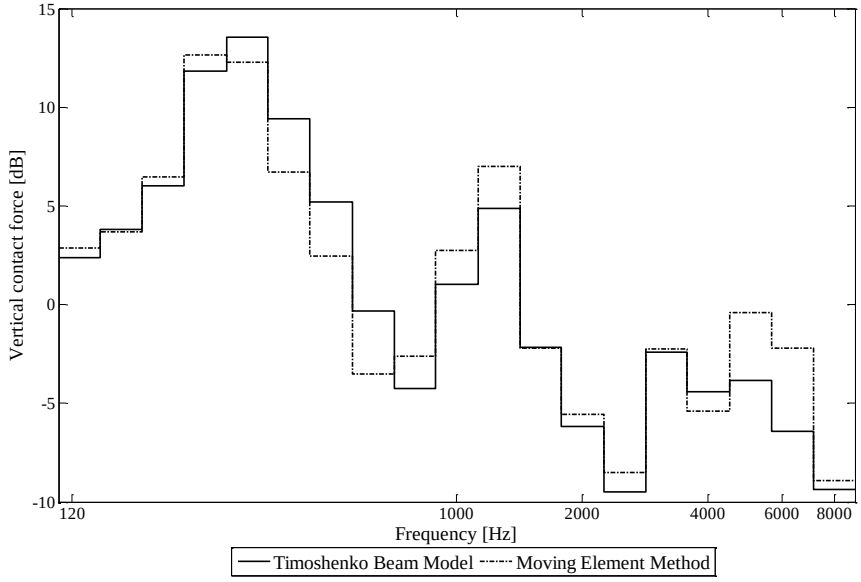


Fig. 2. Frequency domain plot of the vertical wheel-rail contact forces referred to 1 kN when the vehicle circulates at 300 km/h speed on a randomly corrugated tangent track. Amplitudes corresponding to the ISO 3095 limit.

Fig. 3 shows the time history of the vertical contact forces with both rail models caused by a wheelflat when the wheelset runs over a perfectly even tangent track. In the simulations, a rounded geometry of the wheel flat with size 50 mm is adopted. Intense dynamic effects are

observed, initially leading to the occurrence of full loss of contact in the wheels, then followed by a severe impact causing peaks, and finally by a transient vibration that generates further dynamic fluctuations in all the force components. The results for both track models present again a similar trend and mean values, but the MEM shows a larger load peak and a higher frequency content.

Fig. 4 represents the harmonics from Discrete Fourier Transform of the vertical contact forces (in kN). The results displayed are consistent with the previous figure, showing much larger harmonics for the MEM, whose frequency content is markedly higher than the Timoshenko beam in the high frequency domain.

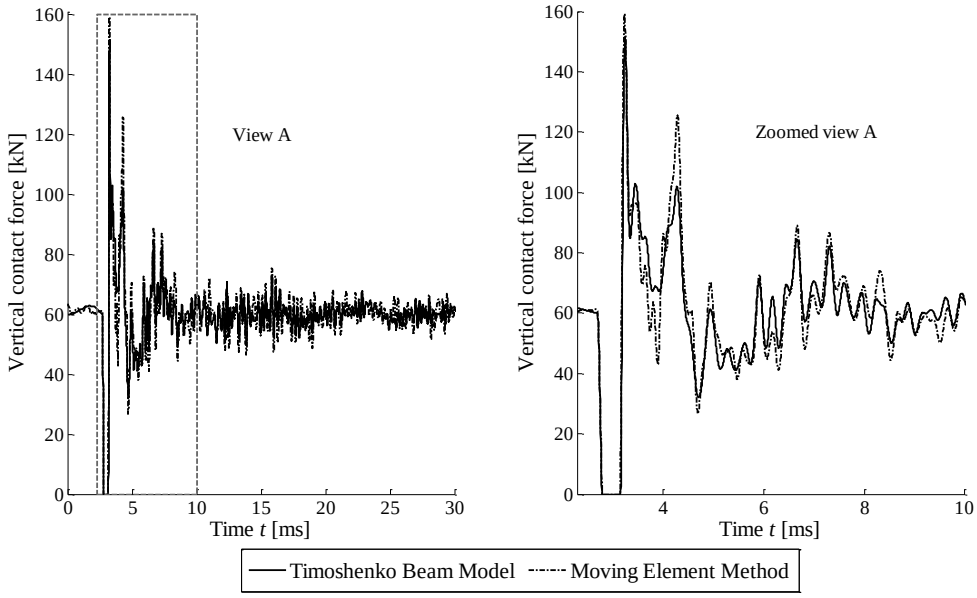


Fig. 3. Vertical wheel-rail contact forces when the vehicle circulates at 300 km/h speed on a perfectly even tangent track in presence of a 50 mm wheelflat.

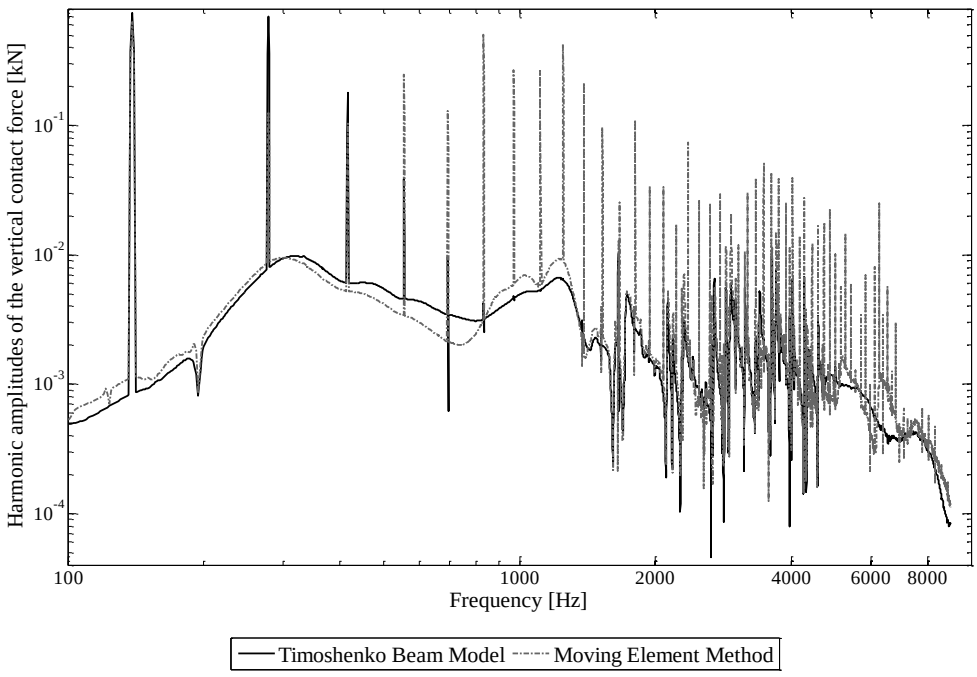


Fig. 4. Frequency domain plot of the vertical wheel-rail contact forces when the vehicle circulates at 300 km/h speed on a perfectly even tangent track in presence of a 50 mm wheelflat.

5. CONCLUSIONS

This paper has presented a new 3D model for a finite railway track which has been formulated through the Moving Element Method (MEM) technique and developed as a Finite Element (FE) model in order to improve the modelling of high frequency dynamics. The model considers an Eulerian coordinate system attached to the moving vehicle instead of a fixed coordinate system and adopts cyclic boundary conditions. This approach permits to decrease the computational cost compared to the FEM models commonly used.

Results for the vertical contact force are presented for two types of excitation: randomly corrugated tangent track and excitation arising from a wheelflat when the wheelset running at 300 km/h along. The MEM and Timoshenko beam models have been compared in all cases. Both models show a similar behaviour in the low-medium frequency domain (up to 2 kHz) for two excitation cases, where a similar trend and mean values can be observed. Significant discrepancies appear, however, in the high frequency range, above 2-3 kHz, where the MEM model has higher frequency content and larger peaks than Timoshenko beam. Hence, the proposed MEM model seems suitable to describe the high frequency dynamics associated to different railway phenomena such as short pitch rail corrugation, wheelflat excitation, axle fatigue, rolling noise, squeal noise and braking noise.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the financial support of Ministerio de Economía y Competitividad and the European Regional Development Fund (project TRA2013-45596-C2-1-R), as well as Generalitat Valenciana (project Prometeo/2012/023) and Ministerio de Educación, Cultura y Deporte (project SP20140659) as part of Programa Campus de Excelencia Internacional.

REFERENCES

[1] D. J. Thompson, *Railway Noise and Vibration: Mechanisms, Modelling and Means of Control*, Elsevier, 2009.

[2] J. C. O. Nielsen, R. Lundén, A. Johansson and T. Vernersson, *Train-track Interaction and Mechanisms of Irregular Wear on Wheel and Rail Surfaces*, Vehicle System Dynamics 40 (1-3) (2003) 3-54.

[3] A. Ekberg, *Fatigue of railway wheels*, The Wheel/Rail Interface Handbook Woodhead Publishing in Mechanical Engineering (2009) 211-244.

[4] J. Martínez-Casas, L. Mazzola, L. Baeza and S. Bruni, *Numerical estimation of stresses in railway axles using a train-track interaction model*, International Journal of Fatigue 47 (2013) 18-30.

[5] L. Baeza and H. Ouyang, *A railway track dynamics model based on modal substructuring and a cyclic boundary condition*, Journal of Sound and Vibration 330 (2011) 75–86.

[6] C. G. Koh, J. S. Y. Ong, D. K. H. Chua and J. Feng, *Moving Element Method for train-track dynamics*, International Journal for Numerical Methods in Engineering 56 (2003) 1549–1567.

[7] C. G. Koh, G. H. Chiew, C. C. Lim, *A numerical method for moving load on continuum*, Journal of Sound and Vibration 300 (2007) 126–138.

A new mathematical model for automated production lines based on probabilistic subcycle times

E.Garcia[♭] and N.Montés[†] *

(♭) Ford España,

Polígono industrial Norte Almussafes S/N, Valencia (Spain),

(†) Universidad CEU Cardenal Herrera,

C/ San bartolomé 55, Alfara del Patriarca, Valencia (Spain),

November 30, 2015

1 Introduction

A production line is an arrangement of machines or a set of sequential operations established in a factory whereby a product moves along while it is being built or produced. Each machine or worker performs a particular job that must be finished before the product moves to the next position in the line. The design of such lines is of considerable importance, [1]. There are a large number of crucial decisions to be made in flow line design as, product design, process selection, line layout configuration, line balancing, machine selection, available technology, etc. Usually, these problems are considered one at a time because of their complexity, [1]. The last and crucial step in the process design is the line balancing, [1]. Here tasks are assigned to the workstations and resources that will be employed on the line (this is a complex combinatorial problem and the solution mostly determines the efficiency of the line designed). Due to the relevance of this task, a large number of researchers have been working on this topic, cite1.

*nicolas.montes@uch.ceu.es

But, what happens when the line is designed and installed in the factory?. The designers defined the maximum production rate, mainly in jobs per hour (JPH), knowing as *Engineering Running Capacity (ERC)*. The goal of the factory employees is to achieve this maximum production rate, defining the throughput of the line. Reality shows that the ERC is extremely difficult or impossible to achieve so the factory employees define a new based-experience maximum production rate and is known as the *Engineering Running Rate (ERR)*. The literature does not offer a reason about this throughput reduction, being the objective of the present paper.

Therefore, in the present paper a novel mathematical model that will allow us to demonstrate that the ERR is due to the machine variability parts is developed . The proposed model is tested in a real welding line in Ford S.A. located at the Almussafes factory. This line has 35 welding units distributed in 8 workstations. The paper is organized as follows. Section 2 presents a mathematical model. Section 3 presents the mini-terms characterization. Section 4 presents a a real welding line and the simulation results. Section 5 concludes the paper with an emphasis on future research challenges.

2 Mathematical model. From micro-term to long term

The literature classifies the data used in the analysis into *long-term* and *short-term*. *Long-term* is mainly used for process planning, while *short-term* focuses primarily on process control. Therefore, following the definition in [2], the *short-term* is referred to an operational period not large enough for a machines failure period to be described by a statistic distribution. The machine's cycle time is considered *short-term*. The present study redefines short-term into two new terms, *mini-term* and *micro-term*. A *mini-term* could be defined as a machine part, in a preventive maintenance policy or in a breakdown, where it could be replaced in an easier and faster manner than another machine part subdivision. Also a *mini-term* could be defined as a subdivision that allows us to understand and study the machine behavior. In the same way, a *micro-term* is defined as each *mini-term* part that could be divided itself. In general, Factory Plan time can be represented as a tensor;

$$F = \sum_{1 \leqslant l \leqslant p} \sum_{1 \leqslant s \leqslant k} \sum_{1 \leqslant M \leqslant j} \sum_{1 \leqslant m \leqslant i} = \tau_{l,s,M,m} \tag{1}$$

where $\tau_{l,s,M,m}$ for $1 \leqslant l \leqslant p$, $1 \leqslant s \leqslant k$, $1 \leqslant M \leqslant j$, $1 \leqslant m \leqslant i$ are independent random times with median and variance

$$\mathbb{E}[\tau_{l,s,M,m}] = \mu_{l,s,M,m} \qquad Var[\tau_{l,s,M,m}] = \sigma_{l,s,M,m} \tag{2}$$

and where m corresponds to the number of *micro-terms*, M corresponds to the number of *mini-terms*, s corresponds to the number of *short-terms* and l corresponds to the number of *long-terms*. In the same way a workstation with k *short-terms* (machines) working in serial can be defined as;

$$W = \sum_{1 \leqslant s \leqslant k} \sum_{1 \leqslant M \leqslant j} \sum_{1 \leqslant m \leqslant j} = \tau_{s,M,m} \tag{3}$$

where the cycle time for each *short-term* can be defined as

$$T^s_{TC} = \sum_{1 \leqslant M \leqslant j} \sum_{1 \leqslant m \leqslant i} = \tau_{s,M,m} \tag{4}$$

for $1 \leqslant s \leqslant k$, and

$$T^M_{TM,s} = \sum_{1 \leqslant m \leqslant i} = \tau_{s,M,m} \tag{5}$$

for $1 \leqslant s \leqslant k$, and $1 \leqslant M \leqslant i$. We can write then

$$W = \sum_{1 \leqslant s \leqslant k} = T^s_{TC} = \sum_{1 \leqslant s \leqslant k} \sum_{1 \leqslant M \leqslant j} = \tau^m_{TM,s} \tag{6}$$

If the workstation has k *short-terms* (machines) working in parallel, the model can be rewritten as;

$$W = MAX[T^s_{TC}] \forall s \in [1, k] \tag{7}$$

In both cases, as the times $\tau_{s,M,m}$ are pseudorandom and independent, we have;

$$\mu^s_{TC} = \mathbb{E}[T^s_{TC}] = \sum_{1 \leqslant M \leqslant j} \sum_{1 \leqslant m \leqslant i} \mu_{s,M,m} \tag{8}$$

and

$$(\sigma^s_{TC})^2 = Var[T^s_{TC}] = \sum_{1 \leqslant M \leqslant j} \sum_{1 \leqslant m \leqslant i} \mu_{s,M,m} \tag{9}$$

for $1 \leqslant s \leqslant k$, and

$$\mu^s_{TM,s} = \mathbb{E}[T^s_{TM,s}] = \sum_{1 \leqslant m \leqslant i} = \mu_{s,M,m} \tag{10}$$

and

$$(\sigma^s_{TM,s})^2 = Var[T^s_{TM,s}] = \sum_{1 \leqslant m \leqslant i} \sigma^2_{s,M,m} \tag{11}$$

Then, we have that

$$\mu^s_{TC} = \mathbb{E}[T^s_{TC}] = \sum_{1 \leqslant m \leqslant i} \mu^s_{TM,s} \tag{12}$$

and

$$(\sigma^s_{TC,})^2 = Var[T^s_{TC}] = \sum_{1 \leqslant m \leqslant i} (\sigma^s_{TM,s})^2 \tag{13}$$

for $1 \leqslant s \leqslant k$. the last step to develop a line model is to simulate the workstation joined with the other ones. The common way is to use a simplified machine state, see [4] and the technique used in the present paper

3 A real welding line. Modelling and simulation.

The goal of the present study is to analyze the effect of some *mini-terms* on the throughput of the line. For this propose, a car welding station is taken as an example. The welding station is one of the most relevant stations because there are 4.500 welding points in a car. In the present study, a real welding line in Ford S.A. located at the Almussafes factory is selected.The welding line was installed in 1980. The staff group that designed the line defined the maximum running capacity, ECR (engineering running capacity), 60 JPH. However, the plant engineers have another maximum running capacity, that is the ERR (engineering running rate), in this case defined in 51 JPH.

The welding line is composed by welding workstations that in itself has welding stations working in parallel, see [6].The behavior of the welding station is simple. First, the robot arm moves the welding clamp to the point to weld. Then, a pneumatic cylinder moves the welding clamp in two phases:

One to approximate the clamp and the second one to weld. The pressure applied by the clamp is controlled by a control system. Each of these devices need a certain time to develop their task and within each of these devices, there are also components that also need a certain time to develop their own tasks. In order to analyze it, the welding unit is divided in three *mini-terms*, the robot arm, the welding clamp motion and the welding task. This *mini-temrs* are measured in an isolated welding station. The experimental methodology is designed as follows, in order to obtain enough precision. The clamping task welds the same point 6 times in order to obtain enough time precision. The robot arm trajectory is the same in all the movements. Then, the clamping task is repeated 40 times in order to obtain a sufficient number of samples to measure the mean value and the standard deviation for each *mini-term*. The next table shows the experimental result measurements for each mini-term. (μ, σ).

	Robot arm	Motion Clamp	Welding clamp
Measurements	(35.55, 0.67)	(2.49, 1.15)	(8.62, 2.05)
Single cycle times	(1,0.11)	(0.42, 0.47)	(1.44, 0.84)

Table 1: Experimental test measurements for each mini-term and the total cycle time. 6 repetitions

By means of this experimental measurements and using the mathematical models proposed in the present study, the simulation is executed where the mean and variance of the jobs produced in each hour is computed. The simulation is running during 50 hours and the the result is, (51,1.05) JPH. The mean value is equal than the ERR, as we wanted to prove, demonstrating that the reason for the loss of jobs is the machine variability, and ERC cannot be achieved due to the mini-term and micro-term time deviation of each machine.

4 Conclusions

This paper presents a new mathematical model that allows us to determine how the variability of a machine part can affect the production rate in a line. This model could be a deep impact in the production industry for some reason. It could help to analyze deterioration pathologies and their effect in

the throughput. In this sense, a deep characterization of the machine part deterioration is required. Although, manufactures have a threshold for the lifespan of the parts, however, for maintenance workers the evolution of the deterioration during the lifespan time could be crucial in the throughput of the line, as well as for the detection of pathologies with great variability.

5 Acknowledgments

Authors wish to thank Ford España S.A and in particular Almussafes Factory for their support in the present research.

References

[1] O.Battaïa and A.Dolghi. A taxonomy of line balancing problems and their solution approaches. *International Journal of production Economics*, Volume(142):259–277, 2013.

[2] PhD Q. Chang. Supervisory factory control based on real-time production feedback, 2011.

[3] L.Li, Q.Chang, J.Ni. Real time production improvement through bottleneck control. *International Journal of production research*, Volume(47:21):6145–6158, 2009.

[4] F.Leal, R.F.S.Costa, J.A.B.Montevechi, D.A.Almeida, F.A.S.Marins A practical guide for operational validation of discrete simulation models *pesquisa Operational*, Volume(31(1)):57–77, 2011.

[5] PhD C.E.López. Unbalanced workload allocation in large assembly lines, 2014.

[6] PhD E. Garcia. Analisis de los sub-tiempos de ciclo técnico en la mejora del rendimiento de las líneas de fabricación, Summer 2016.

Mathematical Model and Implementation of Rational Processing

H. Mora[♭] [*], J. Mora Pascual[♭], J.M. García Chamizo[♭]
and M.T. Signes Pont[♭]

(♭) Specialized Processors Architecture Laboratory

Department of Computer Technology, University of Alicante

San Vicent del Raspeig, Alicante (E-03080). Spain

November 30, 2015

1 Introduction

High precision computing is a very active research area due to the number of interesting applications that need it.

This work aims to propose a mathematical model to represent and operate rational numbers without error. This model constitutes the formal framework of an arithmetic architecture where computational techniques are defined to build the operators with rational numbers.

The key idea of this research is based on representing explicitly the non-zero periodic part of the rational number expressed by the positional number system. The main challenge of this notation is to develop arithmetic methods to compute the numbers, especially if they are also represented in floating point.

[*]e-mail: hmora@ua.es

2 Related Work

There are several alternatives to represent binary and decimal numbers in a computer [2,3] and even some proposals allow to represent rational numbers without error [4]. The standard representation formats [1] are not among them.

The main drawbacks of the more accurate ways of representing numbers are the complexity of their arithmetic methods and the lack of precision when the numbers are transformed into positional number system expressions.

The most recent proposals based on the interval arithmetic and on-line [5] methods offer alternatives of interest which improve the precision of the results. However, they are unable to provide an exact value but only an approximation for irrational or periodic rational numbers.

3 Formal Framework for High Precision Computing

Let f be a general mathematical function. Any computable function whose result approximates to f according to a particular implementation is defined implementation function of f. In this way:

$$codomain(\Gamma_f(\vec{x})) \subseteq codomain(f(\vec{x})) \tag{1}$$

and then,

$$\forall \vec{x} \in domain(\Gamma_f(\vec{x})), |\Gamma_f(\vec{x}) - f(\vec{x})| \leq \varepsilon \tag{2}$$

where,
$\vec{x}$: function arguments
ε : Approximation of f by Γ.

An architecture Λ is characterised both by the set of functions that it provides and by the way in which they are implemented. Let the following set of functions be:

$$\Phi = \{f_1, f_2, ..., f_n\} \tag{3}$$

An architecture Λ_Φ that provides these functions will be made up of:

$$\Lambda_\Phi = \{\Gamma_f 1, \Gamma_f 2, ..., \Gamma_f n\} \tag{4}$$

The framework for high precision computing works into a rational domain. The set of functions provided is the following:

$$\Lambda_\Phi = \{identity, addition, multiplication\} \tag{5}$$

Along with this method that provides exact representation of the numbers, the arithmetic functions of Λ define a commutative ring whose nonzero elements form an abelian group under multiplication and addition operations:

(, +) and (,) are abelian groups.
(, +,) is a commutative ring.

4 Arithmetic Unit Architecture

In general terms, our proposal consists of developing an arithmetic architecture which contains a set of operators which achieve the exact result for rational operands. The arithmetic unit architecture provides the implementation function of each of the functions of Λ as it is defined in expression (6).

$$\Lambda_\Phi = \{\Gamma_i dentity, \Gamma_a ddition, \Gamma_m ultiplication\} \tag{6}$$

These operators produce the exact results of the functions. Figure 1 shows a diagram summarising the unit design. This work develops the representation format and the addition operator. The implementation of the multiplication can be made based on the same principles.

5 Empirical evaluation

The experiments demonstrate that the proposed method for high precision computing of rational numbers developed in this work, produces exact results regardless of the number of chained operations executed, whereas the standard binary IEEE 754 format representation causes an error which increases with the amount of operations and causes significant deviations in the results.

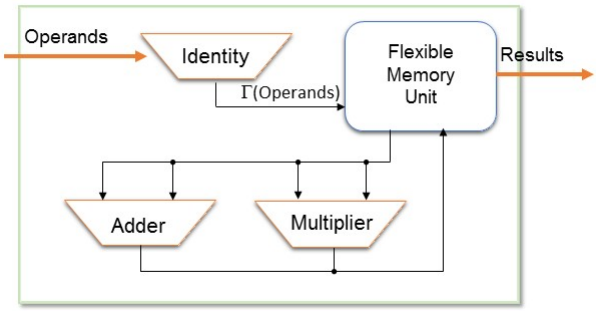


Figure 1: General overview of the Rational Arithmetic Architecture

6 Conclusions

The research presented in this paper describes a mathematical model for representing and processing rational numbers. This proposal is an alternative calculation model when the accuracy in storage and processing of rational numbers is a key aspect to be considered.

References

[1] ANSI/IEEE. IEEE Standard for Binary Floating-Point Arithmetic. ANSI/IEEE Standard 754-2008 *IEEE*, 2008.

[2] JL. Sánchez et al. An iterative method for improving decimal calculations on computers *Mathematical and Computer Modelling*, Volume(50):869–878, 2009.

[3] A. Vázquez, E. Antelo A High-Performance Significand BCD Adder with IEEE 754-2008 Decimal Rounding *IEEE Symposium on Computer Arithmetic*, 135–144, 2009.

[4] D. Piso and J.D. Bruguera. Obtaining Accurate Error Expressions and Bounds for Floating-Point Multiplicative Algorithms *The Computer Journal*, Volume(2):319–331, 2014.

[5] C. Garca, S. Gonzalez, J. Villalba, E.L. Zapata On-line Decimal Adder with RBCD Representation *IEEE International Conference on Application-Specific Systems, Architectures and Processors*, 53–60, 2012.

Formalization of a multi-agent system using Z notation: Application to a system for oil spill location

Francisco J. Mora *, Ramon Rizo, Mar Pujol,
Fidel Aznar, and Mireia Sempere

Department of Computer Science and Articial Intelligence

University of Alicante, San Vicent del Raspeig, Alicante (E-03080). Spain

November 30, 2015

1 Introduction

In this paper, we present the formal specification of multiagent systems using the Z notation to model a system able to detect the presence of oil spills at sea. We will describe the modeling using Z-notation of a multiagent system capable of monitoring, cover and mark the perimeter of a resource. This model has been designed to be implemented with a low cost swarm of drones. To this end, we have designed a microscopic model for a swarm, bearing in account the properties of oil spills. This system allows monitoring of spots produced by a marine oil spill. Subsequently, we have designed a macroscopic model to analyze the functioning of the swarm globally. The proposed system is based on GNOME model developed by National Oceanic and Atmospheric Administration (NOAA) of the United States that allows modeling the oil spill [1]. In this model, you can use real weather maps to simulate ocean currents and winds. This will allow us a realistic simulation of an oil spill on the Spanish coasts. Finally we present the analytical and experimental results of the system operation.

*e-mail: mora@ua.es

2 Formalization with Z language

The language Z is accessible to researchers from variety of different backgrounds and allows us to provide a consistent unified formal account of an abstract agent system. Precisely and unambiguously provide meanings for the common concepts. Allows formally define system modeling and system behavior. Well-structured to provide a foundation for subsequent development of increasingly more refined concepts. ZMichael Luck and Mark d'Inverno wrote Structuring a Z Specification to Provide a Formal Framework for Autonomous Agent Systems in 1995 [2].

The main components of the model are environment, objects, agents and autonomous agents. An agent is an instantiation of an object together with an associated goal or set of goals.For objects, agents and autonomous agents we define their state when situated in an environment in ObjectState, AgentState and AutonomousAgentState. We define how they act in an environment in the schemes ObjectAction, AgentAction and AutonomousAgentAction respectively. For agents and autonomous agents, we detail how they perceive in a given environment in AgentPerception and AutonomousAgentPerception.

We will modeled a system to mark and cover a oil spills, defining the different schemes.

3 Microscopic System Behavior

We propose a homogeneous behavior that run in all the agents of the swarm. This behavior has three states: Wander, Resource and InResource. In broad terms, the agents initially sought by the environment trying to find some trace of the spill. Once the spill is detected, the agent is directed towards it. Finally it will try to keep inside (to cover) or on its perimeter (to mark) depending on the desired behavior. When agents are inside the spill and therefore do not perceive their limit, they do a random motion until it meets the sea.

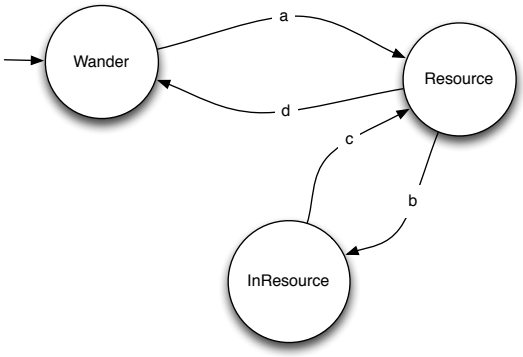


Figure 1: Finite state machine which governs the operation of each agent of the swarm. (a) Detects resource (b) No detects resource. (c) <80%. (d) <80%.

In ObjectAction we define the perpection system (vision system, radar system) and the Locomotor system (actuators). In Agent we define two goals: cover the resource or mark the perimeter, and three actions: Vw, Vr and Vs for each different state (Wander, Resource, InResource). The action Vr depends on Vo and Vc. Vo is a function that defines a reactive behavior for avoidance of other robots. Vc calculates the resource zone with greater intensify.

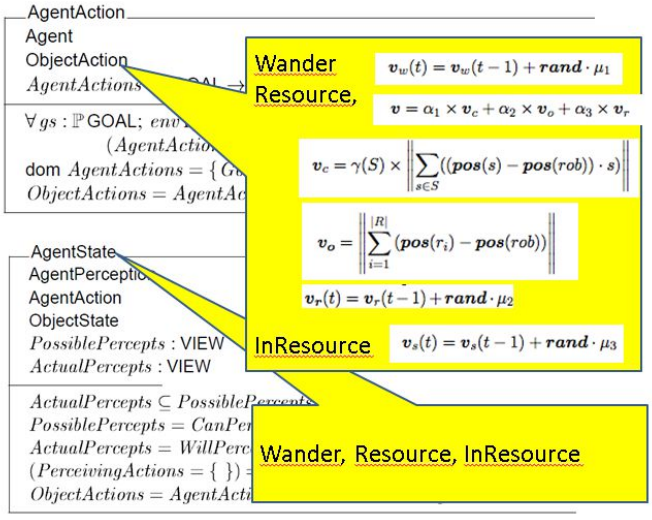


Figure 2: Definition of states and actions in Z language

4 Macroscopic System Behavior

Once you set the microscopic behavior of agents, it is interesting to analyze the overall behavior of the swarm. There are several techniques to analyze such behavior, such as using recurrence equations generated from microscopic behavior defined by a finite state machine or the definition of differential equations. However, most of these methods only allow globally analyze the evolution of transitions between states.

The platform proposed in [3] allows to obtain the probability distribution of the position of the swarm for any time t. This will allow us to predict in detail, the overall system performance. Once defined the microscopic behavior, it is possible to calculate the overall performance of the system using the Fokker-Planck equation:

$$\frac{\partial \rho(\boldsymbol{r},t)}{\partial t} = -\nabla\left(\boldsymbol{A}(\boldsymbol{r},t)\rho(\boldsymbol{r},t)\right) + \frac{1}{2}Q\nabla^2\left(B^2(\boldsymbol{r},t)\rho(\boldsymbol{r},t)\right) \tag{1}$$

Where Q is the displacement caused by a collision. $\rho\left(\boldsymbol{r},t\right)dr_x dr_y$ is the probability of finding a robot in the position r, in the rectangle defined by dr_x and dr_y, at the moment *position t*. This equation provides a statistical method for modeling a swarm of robots, based on multi-particle from the field of quantum physics systems. From a Langevin equation which represents the behavior of a single particle, the Fokker-Planck equation for the entire set is derived.

As described in [4], the Fokker-Planck equation implements the necessary abstractions about microscopic details. The equation continues to be correct if this noise is generated by a Gaussian process. I has as result the evolution of the probability density of the positions of the agents the time. In Figure 3, we can see the functions that define the macroscopic behavior.

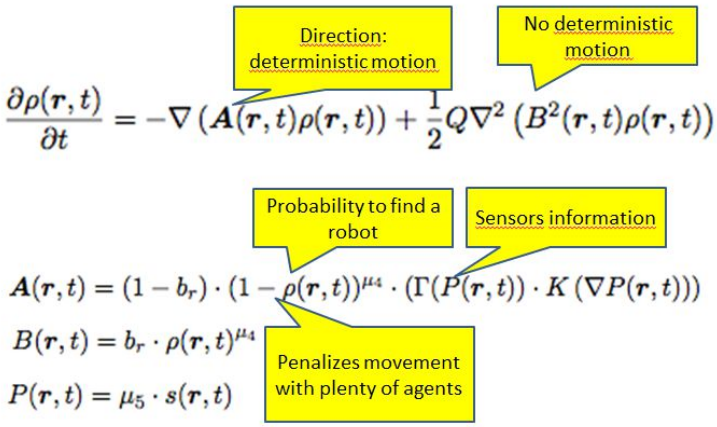


Figure 3: Functions that define the macroscopic behavior

5 Experimentation

To perform the experimentation, we have been performed a simulated oil spill in the coastal area of "San Antonio Abad ", on the island of Ibiza, using an area of approximately $350km^2$. We have chosen this area due to the various ocean currents that affect the area and its proximity to the mainland. they have been used actual meteorological data for both streams of water to the air. For the simulation we used a swarm of 200 agents initially randomly distributed in the environment. These agents circulate evenly 60km/h.

On the left side of the figure 4 we can see a map of the location of the spill and the initial distribution of the agents. On the right side, the image shows the position of the agents in the instant t =15.000s and 30.000s. In the figure 4 you can see how at the time 30,000s swarm completely marks the oil spill.

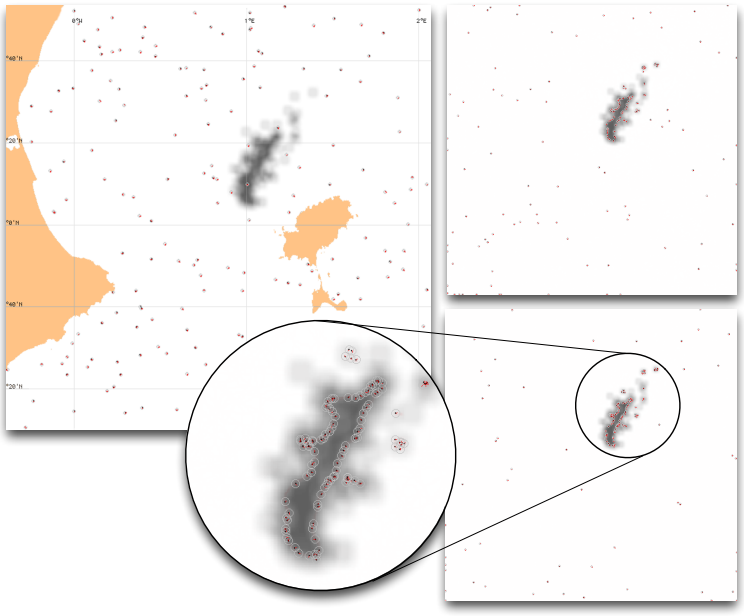


Figure 4: Marking a perimeter of an oil spill

Figure 5 shows the percentage of agents on an oil spill over time for 5 separate executions. The number of agents that are located on the stain gradually increases over time.

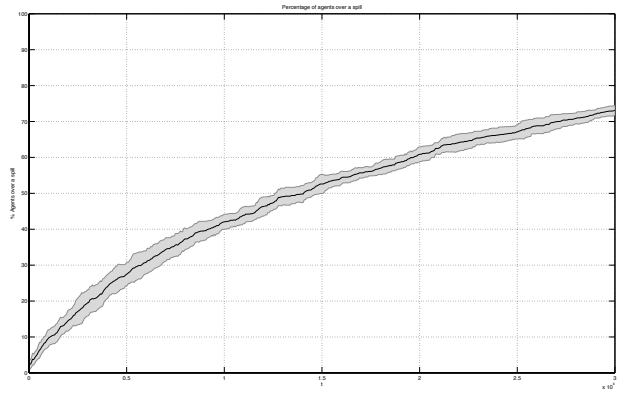


Figure 5: Percentage of agents on an oil spill over time (in seconds)

In addition to the above tests, using a microscopic model it is possible to set the map areas most likely to contain a robot, regardless of the number of agents to be used. This allows viewing behavior more accurately for large swarms are not limited by the number of agents to simulate.

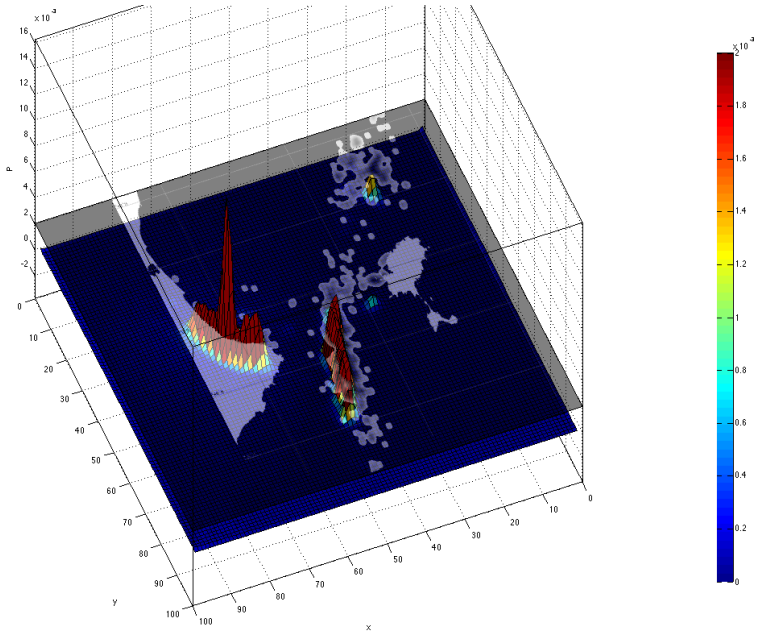


Figure 6: 3D representation of probability distribution of an agent to be in a certain position the environment at time t=168h

We can see that the macroscopic model correctly predict the functioning observed at microscopically level in the swarm.

6 Conclusions

We have formally defined a Swarm System based on states, actions, and goals using the Z notation. The microscopic behavior defines individual behaviors and is simple to understand locally. The macroscopic model establishes the swarm behavior. The macroscopic model is complex (it use differential equations), but the execution times are insignificant compared with microscopic simulation model. We have demonstrated the viability of the model building a system to mark the perimeter and cover oil spills in the sea.

References

[1] United States. National Ocean Service. Office of Response and Restoration. *General noaa operational modelling environment (gnome technical documentation. U.S. Dept. of commerce, National Oceanic and Atmospheric Administration, National Ocean Service, Office of Response and Restoration, 2012*

[2] Justin Clark and Rafael Fierro. Mobile robotic sensors for perimeter detection and tracking. *ISA Transactions, 46(1):3-13, 2007*

[3] Hamann, Heiko and Wörn, Heinz. A framework of space–time continuous models for algorithm design in swarm robotics. *Swarm Intelligence. Spinger. 2008 pag.209-239*

[4] Hamann, Heiko. Space-Time Continuous Models of Swarm Robotic Systems: Supporting Global-to-Local Programming. *Springer.2010*

Acknowledgement

This work has been supported by the Spanish Ministerio de Economía y Competitividad project TIN2013-40982-R. Project co-financed with FEDER funds.

An algorithm for trajectory semantic similarity

Francisco Moreno[♭] [*], Santiago Román[†], and Vania Bogorny[‡]

(♭ †) Universidad Nacional de Colombia, Sede Medellín.,

Carrera 80 No 65-223 - Núcleo Robledo Medellín, Colombia,

(‡) Universidad Federal de Santa Catarina,

CEP 88040-900 - Campus Universitario Cx.P.476 Florianpolis S.C.

November 30, 2015

1 Introduction

Trajectory data are collected as raw trajectories, represented as a sequence of space time points, that correspond to the position (x, y) of an object in a space at an instant (t). The discovery of similar movement behavior from trajectory data is interesting for several domains, such as trajectory clustering and nearest neighbor queries. In this extended abstract, we propose a new similarity function for semantic trajectories, which supports both the semantics of the places visited by the trajectories and the activities performed at each place, what to the best of our knowledge has not been addressed before. While previous works [1, 2] do only consider the exact matching of the semantic dimension, we propose to use a taxonomy of places and activities to consider partial matching of places and activities performed at a place. The main contribution of this work is a new similarity function for measuring different semantic information of trajectory data.

[*]e-mail: fjmoreno@unal.edu.co

2 Trajectories similarity

Similarly to [3] we consider a Category (concept) Tree for the Classification of the Sites (CTCS). For simplicity, each site is associated with a unique category corresponding to a leaf node of the tree. Similarly, we consider a Category Tree for the Classification of the Activities (CTCA).
Let S be a set of m sites $S = \{s_1, s_2, \ldots, s_m\}$ where $s_i = (s\_id, s\_name, s\_cat)$, where s_id is the site identifier, s_name its name and s_cat represents the CTCS category (*leaf node*) which is associated with the site. Thus, one site is (*directly*) associated with one leaf node of the CTCS and (*indirectly*) with all the ancestor nodes of that leaf node in the CTCS.
Example. Let $S = \{s_1, s_2, s_3, s_4, s_5, s_6, s_7\}$ where s_1 = (1, Cinema Central, Cinema), s_2 = (2, Bocagrande, Beach), s_3 = (3, Universidad de Cartagena, University), s_4 = (4, El Rosario, Beach), s_5 = (5, Golden Disco, Discotheque), s_6 = (6, Universidad de Bolvar, University), and s_7 = (7, Botanical Garden, Park).
Similarly, we define a set of p activities $A = \{a_1, a_2, \ldots, a_p\}$ where $a_i = (a\_id, a\_name, a\_cat)$, where a_id is the identifier of the activity, a_name its name and a_cat represents the CTCA category (leaf *node*) which is associated with the activity.
Example. Let $A = \{a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8\}$ where a_1 = (1, Studying math, studying), a_2 = (2, Bicycling, Playing sports), a_3 = (3, Reading science fiction, Reading), a_4 = (4, Dancing electronic, Dancing), a_5 = (5, Studying Spanish, Studying), a_6 = (6, Swimming, Playing sports), a_7 = (7, Singing rock, Karaoke) and a_8 = (8, Watching adventure movies, watching movies).
On the other hand, a trajectory T is a set of n episodes $T = \{e_1, e_2, \ldots, e_n\}$, where $e_i = (s_i, a_i, t_i)$:

$s_i \in S$ represents the site where the episode occurred,
$a_i \in A$ represents the activity accomplished in the site s_i,
$t_i = (t_{ini}, t_{fin})$ represents the time of start (t_{ini}) and end (t_{fin}) ot the episode, $t_{ini} < t_{fin}$.

Example. Let $T_1 = \{e_1, e_2, e_3, e_4\}$ where $e_1 = (s_6, a_5, t_1)$, $e_2 = (s_4, a_6, t_2)$, $e_3 = (s_1, a_8, t_3)$, and $e_4 = (s_7, a_3, t_4)$; and where $t_1 = (8am, 12m)$, $t_2 = (3pm, 4pm)$, $t_3 = (4pm, 5:30pm)$, and $t_4 = (8pm, 9pm)$, all times correspond to 18/Feb/2015. .
To calculate the similarity we extend the proposal of [4]. They propose a formula to determine whether two trajectories are *spatial similarity complete* based on the set of points of interest (POI) of each one and a threshold θ.

Let POI_{ns,T_i} the set of all associated sites (either directly or indirectly) with one node $ns \in$ CTCS included in the episodes of trajectory T_i. The similarity between two trajectories T_i and T_j with regard to ns, C_{ns,T_i,T_j}, is calculated by Equation (1).

$$C_{ns,T_i,T_j} = \frac{\left|POI_{ns,T_i} \cap POI_{ns,T_j}\right|}{\left|POI_{ns,T_i} \cup POI_{ns,T_j}\right|} \quad (1)$$

That is, C_{ns,T_i,T_j} is the relationship between the total number of sites *common* to the two trajectories associated with the node ns and the *total* number of sites of the two trajectories associated with that node. $C_{ns,T_i,T_j} = INDEF$ (undefined) if $POI_{ns,T_i} \cup POI_{ns,T_j} \neq \varnothing$, i.e., when none of the two trajectories have sites associated with the node ns.
Consider two trajectories T_i and T_j. In our method, we compute the similarity of each node $ns \in$ CTCS (or $ns \in$ CTCA if the similarity is considered with regard to activities) by Equation (1), i.e., $SIM_{ns} = C_{ns,T_i,T_j}$. In this way, the user can analyze the trajectories similarity with regard to each CTCS node. For instance, if ns is the root of CTCS, then C_{ns,T_i,T_j} indicates the similarity of the trajectories from a general point of view (node "Site"). The user can then analyze the similarity from a more specific point of view as he descends through the levels of CTCS or a CTCA (a "drill-down"). Note that in our method, to calculate the similarity of a *non-leaf node, calculating the* similarity of its child nodes, is not required.

References

[1] Josh Jia-Ching Ying, Eric Hsueh-Chan Lu, Wang-Chien Lee, Tz-Chiao Weng, and Vincent S Tseng. Mining user similarity from semantic trajectories. In *Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Location Based Social Networks*, pages 19–26. ACM, 2010.

[2] Vania Bogorny, Chiara Renso, Artur Ribeiro Aquino, Fernando Lucca Siqueira, and Luis Otavio Alvares. Constant–a conceptual data model for semantic trajectories of moving objects. *Transactions in GIS*, 18(1):66–88, 2014.

[3] Min-Joong Lee and Chin-Wan Chung. A user similarity calculation based on the location for social network services. In *Database Systems for Advanced Applications*, pages 38–52. Springer, 2011.

[4] Hongbin Zhao, Qilong Han, Haiwei Pan, and Guisheng Yin. Spatio-temporal similarity measure for trajectories on road networks. In *Internet Computing for Science and Engineering (ICICSE), 2009 Fourth International Conference on*, pages 189–193. IEEE, 2009.

Capturing the Subjacent Risk of Death from a Population: the Wavelet Approximation

I. Baeza Sampere * and F. Morillas Jurado †

Universitat de Valencia,

Dpt. Economia Aplicada, Avgda. Tarongers s/n, 46122, Valencia (Spain).

November 30, 2015

1 Introduction

In the study of the biometric behavior of a population it is usual to analyze some biometrics characteristics: life expectancy, risk of death (or survival) at age x, and others. This work focuses in the study of the risk of death at each age x. It is relevant to remark that the value of the risk of death it is unknown (generally).

In this work a method based in the wavelet decomposition with multiresolution schema, combining approximation by Thresholding and Piecewise Polynomial Harmonic (PPH) techniques is articulated to estimate a subjacent risk of death of a population.

It is important to emphasize (for the demographic and actuarial fields) the importance that has to obtain good estimates of the true probabilities of death. The estimation of the subjacent risk can be used to estimate the premiums of several types of insurance policies; or the mathematics reserves (obligatory) that an insurance company must set aside; or to estimate the Sustainability Factor to determine the age of the retirement, or the quantity of the pensions (in the spanish case) (view [14]). In this context, the instrument widely used to summarize the experience of mortality observed in a region or period is the Life Table ([2]). This construction it is organized by age, and the information that we can find corresponds to the called biometric functions. Among these functions, the most commons are: the number of exposed at risk of dying at age x,l_x; the number of death at age x, d_x; the crude rates of death at age x, q_x; the life expectancy at several ages, e_x^{ω}; and others.

In particular, the study of the age of death is easier if we consider that the variable is continuous and we assume that there is a real rate but it is altered by random fluctuations. From this, this work considers that this phenomenon has a random nature and that it is impossible to replicate. This fact makes that the values of biometric functions are always unknown and that only estimations can be obtained. Then, the efforts of this study it is in making estimates of the *underlying* values of $\{q_x\}$, $x = 0, \ldots, \omega$ (ω being the highest age considered in the study) . We perceive the sum of both (real rates and random fluctuations) and we assume that true rates have a structural behavior. This hypothesis justifies the extensive development of the graduation techniques.

*e-mail:Ismael.Baeza@uv.es

†e-mail:Francisco.Morillas@uv.es

Graduation can be defined (view [9]) as "the principles and methods by which a set of observed (or gross) probabilities are adjusted to provide a smooth base that will allow us to make inferences and also practical calculations of bonuses, reserves, etc". Graduation is necessary and has an eminently statistical estimation nature. London [12] explains the reason why we have to change and therefore graduate our initial estimates sequence obtained. This is because for each specific period, given the corresponding data, we can obtain the sequence of initial estimates ages that has sometimes abrupt changes, but it is a particular realization of the evolution of mortality. It should be satisfied that the difference between the probabilities of death of two consecutive age is not excessively high, requiring the setting of a function that meets that condition.

In literature we can find different types of graduation techniques. Parametric graduation tries to find the parameters of a function that adjusts the rate: De Moivre in 1724 and tries to model. Gompertz ([8]) represents an exponential growth for mortality; Makeham (1860) adds a constant component A > 0 to this exponential growth; Weibull (1939) suggests that the mortality force grows as a power of t rather than exponentially. These laws have different problems: for example, some of them are applied only to adult ages; or many fail to represent the hump accidents in adulthood. Heligman and Pollard ([10]) obtained encouraging results for the *full width of the interval of life*. Other techniques are the semiparametric, for example, in [6] the authors make a brief introduction to splines graduation. The last group of techniques is non-parametric models. They no functional form for the behavior of the data is assumed. Mortality rates are obtained by applying smoothing methods that combine adjacent death rates. The works ([4],[5] and [7]) are examples of kernel graduation. In (Baeza and Morillas (2011)) it is proposed a wavelet graduation as a non-parametric method. This study focuses on this latter type of techniques and its objective is to generalize the results obtained [3], which is valid for all ages above 30 (when the social hump is not strong). The work is structured as, in section two introduces some basics concepts about wavelet and its use as a graduation technique. Section 3 enumerates some problems to be applied a simple wavelet graduation in all range of ages; also, the PPH interpolation is briefly describes. Section 4 introduces the Wavelet-PPH graduation and shows the comparison results between Wavelet-PPH graduation and kernel graduation, using a synthetic process (via Heligman and Pollard law). An application to real-observed data is also reported here. Finally, in section 5 there are some final remarks.

2 Basics: Wavelet graduation

At this point, graduation and smoothing can be considered similar and one of the fields where wavelets have been used most successfully is in signal reconstruction/cleaning or cleaned. Noise is removed in order to reproduce the characteristics and patterns that original (or subjacent) signal have. In the actuarial field is not usual to know the probabilities of death, we have only the called mortality experiences.

In Figure 1 we can see, on a logarithmic scale, an example of mortality experience with actual data provided by the Spanish National Statistics Institute (INE). We can split the data into three components: *adaptation to environment*, *social hump* and *natural mortality*. The first represents infant mortality, the second one represents mortality in adult ages which includes death by accidents or maternity. The third component reflects the increased risk of death due to natural causes.

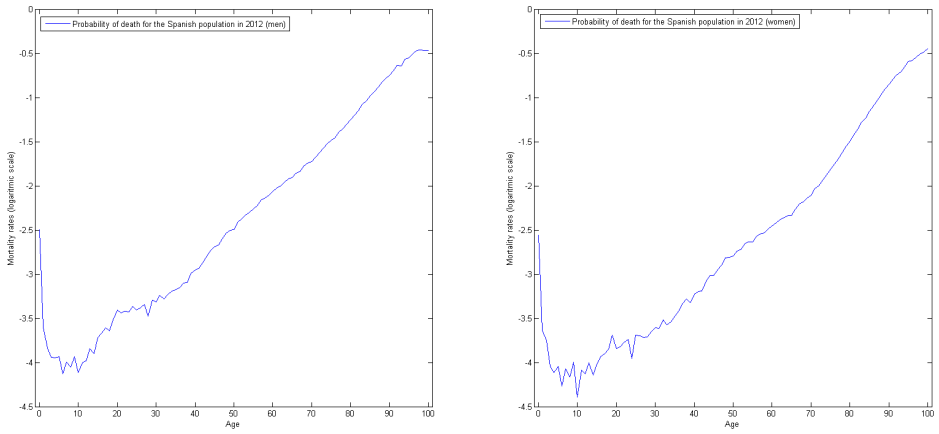


Figure 1: Probability of death for the Spanish population in 2012. Men (left) and women (right). Source: Authors using data published by INE.

In this work, it's assumed that every experience is composed of two terms additives: the true values (unknown) of the series, and a random fluctuation. The wavelet graduation, trying random fluctuation as noise, treats to recovery the true values of the biometric function considered. Let us briefly what is a *wavelet*.

A wavelet is a family of functions is characterized from a generator element: *mother wavelet* function $\psi(t)$. This is a function of real variable t which should range in time and well localized (decay to zero when the variable $t \to \infty$). From the mother wavelet other family elements are defined, these are generated by simultaneous changes of scaling and translations of $\psi(t)$. This family of functions is denoted as: $\{\psi_{a,b}(t), a > 0,\ b \in \mathbb{R}\}$:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}}\psi\left(\frac{t-b}{a}\right), \tag{1}$$

where a is the scaling parameter associated to the stretching or compression of the mother function; b is called the translation parameter and it "locates" temporary the distribution of energy. From these functions $\psi_{a,b}(t)$ we define the Continuous Wavelet Transform of function through the following expressions:

$$W_f(a,b) := \langle f(t), \psi_{a,b}(t)\rangle = \int_{-\infty}^{+\infty} f(t)\,\overline{\psi}_{a,b}(t)\,dt. \tag{2}$$

In this way, a Inverse Wavelet Transform ($\widehat{W_f}$) it is defined and verify that $\widehat{W_f}W_f = f$. Also, *Discrete Wavelet Transform* (and its inverse) is defined similarly to the Continuous Wavelet. In this case the parameters a and b can only take discrete values. The transform is obtained similarly to Mallat in [13] using the discrete scalar product between the function to decompose and the wavelet functions, $\psi_{a,b}(t)$.

The result of applying the Wavelet Transform (continuous or discrete) is formed by two functional elements (or data series): a first element is called *scaling part*, and the second one named *wavelet part*. Scaling part give us a first approach that includes the trend of the initial series, f. This part lost the details of the original series which are contained in the wavelet part. Often, these 'details', the differences between the original series and the obtained by scaling, are considered disturbances. Appropriately combining the two parts we can reconstruct accurately the initial values of the function (or data set). This process can be

applied iteratively and application leads: on another scale, a new part scaling and wavelet and so on. If the process only is applied in the scaling part, it is named Multiresolution Analysis; if, also, it is applied in the wavelet part, it is named Wavelets Packets. Often, the wavelet part and the scaling part obtained are orthogonal, so is, the information contained in a part is not contained in the other, i.e., the information it is complementary, not redundant. It is usual, for example in signal processing, to assume that when a series of data has random fluctuations (noise or not), these are in the wavelet part. So the elimination or reduction of noise is linked to the treatment of wavelet part (as previous to apply inverse transform to recovery the signal). The aim of the wavelet graduation is to reduce or even eliminate random fluctuations using *thresholding techniques*: truncating the values of wavelet part, assuming certain *threshold* as an element that determines whether a value is or not considered as random fluctuation.

3 Limitations of the *simple* wavelet graduation: The PPH interpolation

The wavelet graduation may have more or less significant drawbacks according to the available information or the functional relationship of the data. In the case of life tables, in [3], this technique only can be applied to ranges above 30 years of age. For younger age mortality curve has a non-linear relationship that complicates the analysis by the lack of information. We consider important to highlight some aspects:

- When we apply the wavelet technique, the incorporation of symmetric information at the ends of the series introduces *noise discontinuity*.
- The problem of discontinuity reappears if we use a wavelet family with a big support or if we use *Multiresolution Analysis*.
- Some effect, similar to the Gibbs phenomenon, has also been detected by smoothing the central area, the accident or social hump.

To avoid these limitations we consider the PPH interpolation as a previous step to applied the wavelet graduation, increasing synthetically the available data. The PPH technique is a nonlinear interpolation scheme of forth order and depending of the data. It was introduced in [1] and based on a piecewise polynomial harmonic operator. The PPH interpolation has several desirable features:

- Each polynomial piece is constructed with four centered point,
- On smooth region it is as accurate as its linear equivalent,
- No introduce oscillations and
- Preserve the concavity / convexity of the function.

The PPH polynomial value at is given by the expression:

$$\widetilde{P}_j\left(x_{j+\frac{1}{2}}\right) = \frac{f_j + f_{j+1}}{2} - \frac{1}{4}\widetilde{D}h^2 \tag{3}$$

where

$$\widetilde{D} = \begin{cases} \frac{2D_jD_{j+1}}{D_j+D_{j+1}} \text{ if } D_jD_{j+1} > 0 \\ 0 \text{ otherwise.} \end{cases} \tag{4}$$

and $D_j = f\left[x_{j-1}, x_j, x_{j+1}\right]$ are the divided differences associates to the interpolatory stencil.

4 Combining wavelet and PPH techniques to graduate:validation

As we have seen, graduation wavelet has problems when we apply it to the entire range of age of the biometric function. To solve this we will introduce some additional information for inter annual data. These new data will be given by the PPH interpolation since it allows incorporating additional data without introducing oscillations and preserving the concavity (or convexity) of the function.

The calibration and validation of the Wavelet-PPH technique that we proposed, it is do measuring the capacity of the recovery the true probabilities of death. In the practice this probabilities are unknown, for this reason we articulate a synthetic procedure, which give us an arbitrary number of *synthetic death experiences* ($N = 10^4$), all of them equally plausible from the same -and known- true probabilities of death(via Helligman and Pollard law,HPL). The numeric procedure it is based on the HPL to obtain a (*true*) probabilities of death, and estimates the synthetics values of the next variables: number of deaths for each age,d_x; and the exposed at risk of death l_x. The relation between them and the numerical procedure are, for each age $x = 0, 1, \ldots, \omega$: $d_x \sim Bi(l_x, q_x)$, $d_x = l_x - l_{x+1}$, with initial value $l_0 = 10^5$ (arbitrary). Briefly, to obtain a first synthetic scenario: (i) the initial value of l_0 is fixed in all experiment, 10^4.(ii) It is obtained d_0^1 as a random number from $Bi(l_0, q_0)$. (iii) The number of exposed at risk to age $x = 1$ is estimated as survival at age 0 via $l_1^1 = l_0 - d_0^1$. (iv) The synthetic observed rates is $q_0^1 = \frac{d_0^1}{l_0}$. (v) Nextly, we iterate the process for next age, so is: $d_1^1 \sim Bi(l_1^1, q_1)$; $l_2^1 = l_1^1 - d_1^1$; and $q_1^1 = \frac{d_1^1}{l_1^1}$. The procedure continue for $x = 2, \ldots, \omega$; obtaining the first sequence of synthetic observed rates $q_0^1, q_1^1, \ldots, q_\omega^1$. To obtain the others synthetic observed rates, the procedure it is initialized obtaining alternative values for each variables d_x^k, l_x^k and q_x^k, for $k = 1, 2, \ldots, N$. The procedure described ability us to measue The simulation procedure allows us to apply the graduation method to each series of synthetic values,$q_0^k, q_1^k, \ldots, q_\omega^k, k = 1, \ldots, N$, and enables us to measure the capacity of the recover the true values (prefixed via HPL). Also, these synthetic series are used to compare the Wavelet-PPH technique with Gaussian kernel graduation. We consider two types of indicators: goodness of fit and smoothness:

The Mean relative indicator (IRM): $IRM\,(q) = \frac{1}{\omega+1}\sum_{x=0}^{\omega}\frac{|q_x - \widehat{q}_x|}{q_x}$.

The Mean squared relative indicator (IRCM): $IRCM\,(q) = \frac{1}{\omega+1}\sum_{x=0}^{\omega}\frac{|q_x - \widehat{q}_x|^2}{q_x}$.

The Whittaker-Henderson smoothness indicator [15]: $S = |S\,(\widehat{q}_x) - S\,(q_x)|$ where $S\,(q_x) = \sum_{x=0}^{\omega-2}\left(\Delta^2 q_x\right)^2$

In the above definitions q_x denotes the theoretical probability of death (HPL), and $\widehat{q}_x$ indicates the graduated probability, the value obtained by applying the graduation to each generated realizations. The defined indicators suggest that the lower value is the best estimate of the theoretical probability is obtained, suggesting that another technical improvement in this regard.

Table 1 shows the results of the comparison between the Wavelet-PPH graduation and the Gaussian kernel graduation. In columns 3 and 4 we can see the mean value of the indicators for the 10^4 synthetic death experiences. The last column presents the percentage of times that Wavelet-PPH obtains better results than Gaussian kernel graduation.

Table 1. Indicators. Comparison Wavelet-PPH vs. Kernel

N	Indicator	Wavelet-PPH	Kernel	Better W-PPH(%)
	IRM	0,0338977	0,03903211	86,74
100	IRCM	0,00370256	0,00885898	99,84
	S	0,00058177	0,00945352	100,00
	IRM	0,034489398	0,0392806	85,37
200	IRCM	0,003280213	0,009398363	100,00
	S	0,000700086	0,009761079	100,00
	IRM	0,034557725	0,039082103	82,74
400	IRCM	0,003582433	0,009322405	99,96
	S	0,000870859	0,00998631	100,00
	IRM	0,035888371	0,038810222	74,07
800	IRCM	0,00369508	0,009175132	99,97
	S	0,002292422	0,010123821	100,00

The table 2 shows the parameters for the Wavelet-PPH graduation. We opted for the biorthogonal wavelet family. Wavelet is selected by means of a criterion based on energy retention instead an exhaustive strategy like in [3]. The measure is given by $H = \frac{\|\widehat{q}_x\|^2}{\|q_x\|^2}$. We work with a criterion for thresholding based on [13].

Table 2. Parameters used in Wavelet-PPH technique

N	Wavelet	Scales	Thresholding
100	Biorthogonal 3.3	2	0,15
200	Biorthogonal 3.3	3	0,2
400	Biorthogonal 3.3	4	0,25
800	Biorthogonal 3.3	5	0,3

In Figure 2 (left), are presented for the entire range of ages: the Heligman and Pollard theorical model; a random realization; and the two approximations by graduation, the kernel graduation and the Wavelet-PPH graduation (with N=400 and the parameters given by table 2). In Figure 2(right) we observe details of these functions more closely.

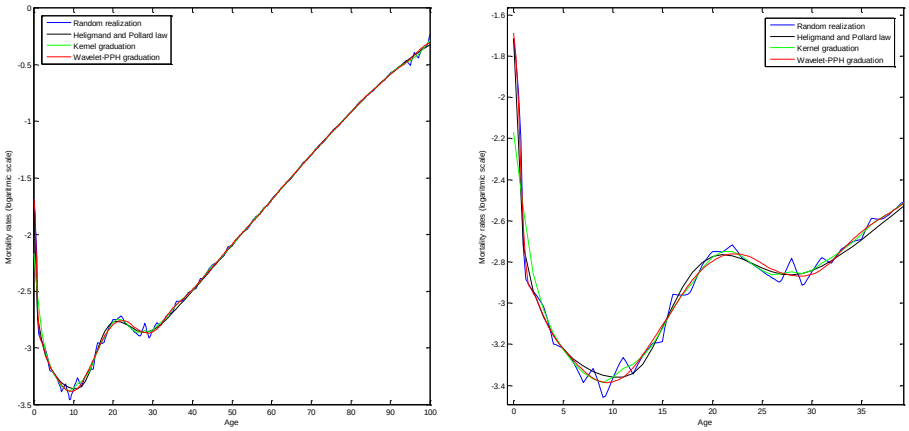


Figure 2: Left: Comparison. All range of ages. Source: Authors. Right: Details.

5 An Application to Observed (real) Data

We apply the Wavelet-PPH technical data to actual mortality rate of Spain to check the results of this work. As the actual rates are unknown, we cannot compare it but, using different values of Table 2 and considering the results of the previous section we believe that the approach of Figure 3 is a good reconstruction of the mortality rate. This can be used by the different agents of actuarial science in their fields.

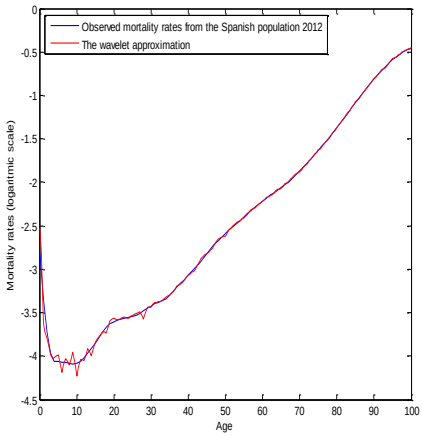


Figure 3: Spanish population 2012 mortality rates and Wavelet-PPH graduation. Source: Authors.

6 Final Remarks

This paper presents a process in two stages to graduate mortality rates. The process combine wavelets and PPH interpolation trying to generalize the results obtained in [3]. In the first stage the PPH interpolation

allows to incorporate additional data without introducing oscillations and preserving the concavity (or convexity) of the function. The second stage use wavelets eliminate the noise (or random fluctuations) in order to reconstruct the *true* values of the biometric function considered. That technique can be applied to all range of ages and, in the sense of the indicators used, give us better results than kernel graduation. Also, the technique presented is more robust in the sense that, when the indicator considered is better (minus value) for the kernel graduation than Wavelet-PPH technique, the relative difference is higher than if we consider the reverse relation.

Acknowledgements

This research was partially supported by Ministerio de Economía y Competitividad under grant MTM2012-31698.

References

[1] Amat S., Busquier S. and Candela V. A polynomial approach to the piecewise hyperbolic method. *I.J. Comput. Fluid Dynam.*,3(17): 205–217, 2003.

[2] Ayuso M., Corrales H., Guillen M., Pérez-Marín A.M. and Rojo J.L. *Estadistica Actuarial Vida*. Barcelona: UBe, 2007.

[3] Baeza I. and Morillas F.G. Using wavelet to non-parametric graduation of mortalily rates *Anales del Instituto de Actuarios Españoles*, 17: 135–164,2011.

[4] Copas J. and Haberman S. Non parametric graduation using kernel methods *Journal of the Institute of Actuaries*, (110): 135–156, 1983.

[5] Felipe A., Guillen M., Nielsen J. Longevity studies based on kernel hazard estimation. *Insurance: Mathematics and Economics*, (28):191–204,2001.

[6] Forfar D.,McCutcheon J. and Wilkie A. On graduation by mathematical formulae *Journal of the Institute of Actuaries*, (115): 693–694,1988.

[7] Gavin J.,Haberman S. and Verrall R. Moving weighted average graduation using kernel estimation *Mathematics and Economics*, 12 (2):113–126, 1993.

[8] Gompertz B. On the nature of the function of the law of human mortality and on a new mode of determining the value of life contingencies *Transactions of The Royal Society*, (115): 513–585,1825.

[9] Haberman S. and Renshaw A. Generalized linear models and actuarial science *The Statistician*, 4(45):113–126,1996.

[10] Heligman L. and Pollard J. The age pattern of mortality *Journal of the Institute of Actuaries*,(107): 49–80,1980.

[11] Instituto Nacional de Estadistica. *Tablas de mortalidad de la población de España por año, sexo, edad y funciones*[Online].*Madrid: INEbase*. http://www.ine.es [visited:02/09/2015].

[12] D. London, Graduation: The Revision of Estimates. Coonecticut, ACTEX Publications, 1985.

[13] Mallat S.G. A theory for multiresolution signal decomposition: The wavelet representation *IEEE Translation*, 11(7):84–95,1980.

[14] Meneu Gaya R., Devesa Carpio J.E., Devesa Carpio M. and Nagore García A. El Factor de Sostenibilidad: Diseños alternativos y valoración financiero-actuarial de sus efectos sobre los parámetros del sistema *Economía Española y Protección Social*, V:63–96,2013.

[15] Whittaker E.T. On a new method of graduation *Proc. Edinburgh Math. Soc.*, (41): 63–75,1923.

Modeling a fishery problem using random differential equations: The randomized Bertalanffy model

M.-C. Casabán[♭], J.-C. Cortés[♭] [*], A. Navarro-Quiles[♭], J.V. Romero[♭], M.-D.Roselló[♭] and R.-J. Villanueva[♭].

(♭) Instituto Universitario de Matemática Multidisciplinar,

Universitat Politècnica de València, Spain.

November 30, 2015

1 Introduction

Starting with the contributions of Newton, Leibniz, Jacob and Johan Bernoulli, the usefulness of the differential equations have been demonstrated. Usually, inputs parameters are based in sample data, which lead to measurement errors. Then, there are some situations where we must to consider this parameters like randomness magnitudes.

In this paper we solve the following initial value problem (i.v.p.) based on a random homogeneous Bernoulli-type differential equation

$$\left.\begin{array}{rcl} \dot{X}(t) & = & CX(t) + D(X(t))^A, \quad t \geq t_0\,, \\ X(t_0) & = & X_0\,, \end{array}\right\} \tag{1}$$

where all the input parameters X_0, D, C and A are assumed to be absolutely continuous random variables (r.v.'s) defined on a common probability space, $(\Omega, \mathfrak{F}, \mathbb{P})$.

Solving random differential equations involves not only the computation of the solution, as an ordinary differential equations but the determination of

[*]e-mail: jccortes@imm.upv.es

the main statistical properties associated with the solution stochastic process (s.p.), such as the mean and variance functions. If it is possible, the determination of the first probability density function (p.d.f.) is also recommendable since from it one can get a full statistical description of the solution in every time instant t. In this work, we deal with the computation of the 1-p.d.f. of the solution $X(t)$ of the i.v.p. (1).

Notice that this problem is compound by 15 cases (obviously, excluding the deterministic case where all the inputs are constants). In this paper, we will focus on the general scenario, where all inputs parameters are r.v.'s, and we will assume that

$$\mathbb{P}[\{\omega \in \Omega \,:\, A(\omega) \neq 1\}] = 1. \tag{2}$$

In order to deal with the general scenario all the cases listed in Table 1 must be undertaken, however in this piece only Case I will be developed since the rest of the casuistries can be dealt following an analogous reasoning to the one showed below.

Case I	$\mathbb{P}[\{\omega \in \Omega \,:\, X_0(\omega) \neq 0\}] = 1 = \mathbb{P}[\{\omega \in \Omega \,:\, C(\omega) \neq 0\}]$
Case II	$\mathbb{P}[\{\omega \in \Omega \,:\, X_0(\omega) = 0\}] = 1 = \mathbb{P}[\{\omega \in \Omega \,:\, C(\omega) \neq 0\}]$
Case III	$\mathbb{P}[\{\omega \in \Omega \,:\, X_0(\omega) \neq 0\}] = 1 = \mathbb{P}[\{\omega \in \Omega \,:\, C(\omega) = 0\}]$
Case IV	$\mathbb{P}[\{\omega \in \Omega \,:\, X_0(\omega) = 0\}] = 1 = \mathbb{P}[\{\omega \in \Omega \,:\, C(\omega) = 0\}]$

Table 1: List of different cases considered to compute the 1-p.d.f. of the solution s.p. to the Bernoulli random i.v.p. (1) under assumption (2).

With this aim, we apply the Random Variable Transformation (R.V.T.) method, [1]. Further, the following result will be also useful to conduct our study:

In the case that $\mathbf{U} = (U_1, U_2)$ be an absolutely continuous real random vector defined on a probability space $(\Omega, \mathfrak{F}, \mathbb{P})$, with joint p.d.f. $f_{\mathbf{U}}(u_1, u_2)$. Assume that $U_1(\omega) \neq 0$ and $U_2(\omega) \neq 1$ for all $\omega \in \Omega$. Then, the p.d.f. $f_V(v)$ of the transformation $V = (U_1)^{\frac{1}{1-U_2}}$ is given by

$$f_V(v) = \int_{\mathcal{D}(U_2)} f_{\mathbf{U}}(v^{1-u_2}, u_2) \left|(1-u_2)v^{-u_2}\right| \, \mathrm{d}u_2, \tag{3}$$

where $\mathcal{D}(U_2)$ denotes the domain of U_2.

2 Solving the Case I

Let us assume that (X_0, D, C, A) is a random vector with joint p.d.f. $f_{X_0,D,C,A}(x_0, d, c, a)$.

Our objective is to compute the 1-p.d.f. of the solution s.p. of the i.v.p. (1) in the Case I, thus, where the parameters X_0 and C are nonzero with probability one.

With this aim, making the change of variable

$$X(t) = (Z(t))^{\frac{1}{1-A}},$$

the nonlinear problem (1) can be transformed into the linear i.v.p.

$$\begin{aligned} \dot{Z}(t) &= (1-A)CZ(t) + (1-A)D, \quad t \geq t_0\,, \\ Z(t_0) &= (X_0)^{1-A}, \end{aligned}$$

whose exact closed form solution s.p. is given by

$$Z(t) = (X_0)^{1-A}\,\mathrm{e}^{(1-A)C(t-t_0)} + \frac{D}{C}\,\mathrm{e}^{(1-A)C(t-t_0)} - \frac{D}{C}.$$

Now, fixing $t \geq t_0$, denote $Z = Z(t)$ and applying R.V.T. method in its general form, we get the joint p.d.f. of the random vector (Z, D, C, A). Then, marginalizing, the p.d.f. of the random vector (Z, A) is given by

$$\begin{aligned} &f_{Z,A}(z,a) \\ &= \int_{\mathcal{D}(D)}\int_{\mathcal{D}(C)} f_{X_0,D,C,A}\left(\left(\frac{-d+\mathrm{e}^{-c\,(t-t_0)(1-a)}(zc+d)}{c}\right)^{\frac{1}{1-a}}, d, c, a\right) \\ &\times \left|\frac{\mathrm{e}^{-c(t-t_0)(1-a)}}{1-a}\left(\frac{-d+\mathrm{e}^{-c(t-t_0)(1-a)}(zc+d)}{c}\right)^{\frac{a}{1-a}}\right| \mathrm{d}c\;\mathrm{d}d. \end{aligned}$$

Considering $X(t) = (Z(t))^{\frac{1}{1-A}}$, applying (3) to $U_1 = Z$, $U_2 = A$ and $V = X$, and considering $t \geq t_0$ arbitrary, the 1-p.d.f. of the solution s.p. $X(t)$ of the i.v.p. (1) is given by

$$\begin{aligned} &f_1(x,t) \\ &= \int_{\mathcal{D}(A)}\int_{\mathcal{D}(D)}\int_{\mathcal{D}(C)} f_{X_0,D,C,A}\left(\left(\frac{-d+\mathrm{e}^{-c\,(t-t_0)(1-a)}(x^{1-a}\,c+d)}{c}\right)^{\frac{1}{1-a}}, d, c, a\right) \\ &\times \left|\frac{\mathrm{e}^{-c(t-t_0)(1-a)}}{x^a}\left(\frac{-d+\mathrm{e}^{-c(t-t_0)(1-a)}(x^{1-a}c+d)}{c}\right)^{\frac{a}{1-a}}\right| \mathrm{d}c\,\mathrm{d}d\,\mathrm{d}a\,. \end{aligned} \tag{4}$$

3 An application to modelling: Bertalanffy model

The Bertalanffy model [2, 3] is useful to describe the fish weight growth over the time. It is given by

$$\left.\begin{array}{rcl} \dot{W}(t) & = & -\lambda W(t) + \eta (W(t))^{2/3}, \quad t \geq t_0\,, \\ W(t_0) & = & W_0\,, \end{array}\right\} \tag{5}$$

where $W(t)$ is the fish growth at time instant t, η is the intrinsic growth rate and λ is the linear coefficient. We observe that the i.v.p. (5) corresponds a particular problem of the random i.v.p. (1). Notice that this case does not fit to any of the four cases listed in Table 1, since now the exponent A in the Bernoulli i.v.p. (1) is the constant $2/3$. We follow four steps to show a full study of this application.

Step 1: To randomize the classical model. Then, taking into account the i.v.p. (5) we consider that the initial condition, W_0, and we assume that the coefficients, η and λ are absolutely continuous r.v.'s defined on a common probability space, $(\Omega, \mathfrak{F}, \mathbb{P})$, being $f_{W_0,\eta,\lambda}(w_0, \eta, \lambda)$ their joint p.d.f. Moreover, let us suppose that

$$\mathbb{P}\left[\{\omega \in \Omega : W_0(\omega) \neq 0\}\right] = 1, \quad \mathbb{P}\left[\{\omega \in \Omega : \lambda(\omega) \neq 0\}\right] = 1.$$

Step 2: We will determine the 1-p.d.f. of the solution s.p. $W(t)$. This is done by following the same development exhibited in Section 2, this yields

$$\begin{array}{l} f_1(w,t) \\ = \dfrac{1}{3} f_Z\left(w^{1/3}\right) |w|^{-2/3} \\ = \displaystyle\int_{\mathcal{D}(\eta)} \int_{\mathcal{D}(\lambda)} f_{W_0,\eta,\lambda}\left(\left(\frac{\mathrm{e}^{(1/3)\lambda(t-t_0)}\,\lambda w^{1/3} + \eta - \mathrm{e}^{(1/3)\lambda(t-t_0)}\,\eta}{\lambda}\right)^3, \eta, \lambda\right) \\ \times \left(\dfrac{\mathrm{e}^{(1/3)\lambda(t-t_0)}\,\lambda w^{1/3} + \eta - \mathrm{e}^{(1/3)\lambda(t-t_0)}\,\eta}{\lambda}\right)^2 \mathrm{e}^{(1/3)\lambda(t-t_0)}\,|w|^{-2/3} \mathrm{d}\lambda\,\mathrm{d}\eta. \end{array} \tag{6}$$

Step 3: We use real data in order to assign a probabilistic distribution to input random vector, $\mathbf{Q} = (W_0, \eta, \lambda)$. The data are the fish weights w_i for wallaye species in lbs every year t_i, $1 \leq i \leq 33 = N$. They are collected in Table 2.

t_i (years)	1	2	3	4	5	6	7	8	9	10	11
w_i (lbs)	0.2	0.4	0.6	0.9	1	1.3	1.6	1.8	2.3	2.6	2.9
t_i (years)	12	13	14	15	16	17	18	19	20	21	22
w_i (lbs)	3.1	3.4	3.7	4.5	5.2	5.7	6.2	6.5	6.7	6.8	7.2
t_i (years)	23	24	25	26	27	28	29	30	31	32	33
w_i (lbs)	8.2	9	9.5	10	10.5	11	11.5	12	12.5	13	14

Table 2: Fish weights w_i for *walleye* species in lbs every year t_i, $1 \leq i \leq 33 = N$.

Step 4: To obtain the distribution of the random vector $\mathbf{Q} = (W_0, \eta, \lambda)$ by using an inverse frequentist technique. This leads to the following distribution for the input random vector

$$\mathbf{Q} = (W_0, \eta, \lambda) \sim \mathrm{N}_3(\mu_{\mathbf{Q}}; \Sigma_{\mathbf{Q}}),$$

where $\mu_{\mathbf{Q}} = (0.365934, 0.305461, 0.0880184)$ and

$$\Sigma_{\mathbf{Q}} = \begin{bmatrix} 0.0029288 & -0.000812275 & -0.000400288 \\ -0.00081227 & 0.000268075 & 0.000136915 \\ -0.000400288 & 0.000136915 & 0.0000705259 \end{bmatrix}.$$

In the Figure 1, the 1-p.d.f. of $W(t)$ at different time intervals is shown.

Finally, we can construct both punctual and probabilistic predictions based on confidence intervals. In the Figure 2 one observes that the proposed randomized Bertalanffy model captures satisfactorily data uncertainty.

Acknowledgements

This work has been partially supported by the Ministerio de Economía y Competitividad grants MTM2013-41765-P. Ana Navarro Quiles acknowledges the doctorate scholarship granted by Programa de Ayudas de Investigación y Desarrollo (PAID), Universitat Politècnica de València.

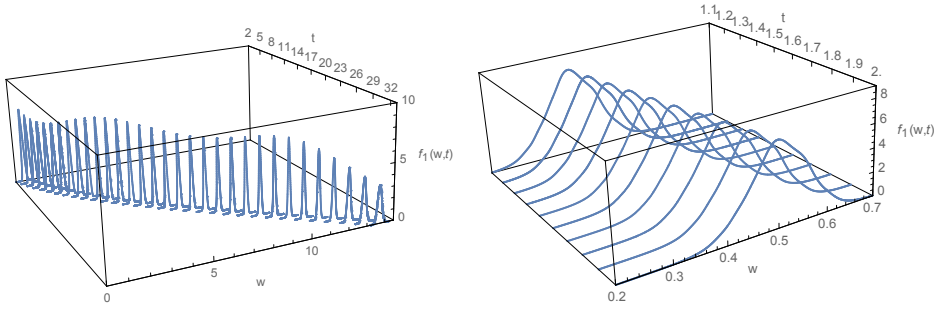


Figure 1: Left: 1-p.d.f. of $W(t)$ given by (6) for all the times of the sample, $t \in \{2, \ldots, 33 = N\}$. Right: Detailed representation of the 1-p.d.f. for the times $t \in \{1.1, 1.2, \ldots, 2\}$.

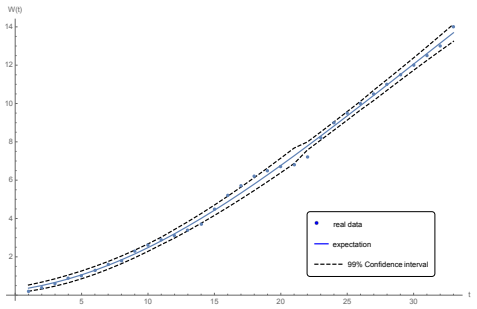


Figure 2: Expectation (solid line) and 99%–confidence intervals (dotted lines). Points represent fish weigh.

References

[1] T.T. Soong, Random Differential Equations in Science and Engineering. New York, Academic Press, 1973.

[2] Bertalanffy, L. von. A quantitative theory of organic growth (Inquiries on Growth Laws II) *Human Biology*, Volume(10): 181–213, 1938.

[3] Bertalanffy, L. von. Quantitative laws in metabolism and growth *The Quarterly Review of Biology*, Volume(32): 217–231, 1957.

[4] M. C. Casabán, J. C. Cortés, J. V. Romero, M. D. Rosell ó. Determining the first probability density function of linear random initial value problems by the Random Variable Transformation (RVT) technique: A comprehensive study *Abstract and Applied Analysis*, Volume(2014) : 1–25, 2014.

[5] M. C. Casabán, J. C. Cortés, J. V. Romero, M. D. Roselló. Probabilistic solution of random SI-type epidemiological models using the random variable transformation technique *Communications in Nonlinear Science and Numerical Simulation* Volume(24): 1–3, 2015.

[6] M. C. Casabán, J. C. Cortés, A. Navarro-Quiles, J. V. Romero, M. D. Roselló, R. J. Villanueva. Probabilistic solution of the homogeneous Riccati differential equation: A case-study by using linearization and transformation techniques *Journal of Computational of Applied Mathematics* Volume(291): 20–35, 2016.

[7] L. v. Qiming, J. W. Pitchford. Stochastic von Bertalanffy models with applications to fish recruitment *Journal of Theoretical Biology* Volume(244): 640–655, 2007.

A front-fixing numerical method for a free boundary nonlinear diffusion logistic population model

M. A. Piqueras[†] [*], R. Company[†] , and L. Jódar[†]

([†]) Instituto de Matemática Multidisciplinar, Universitat Politècnica de València,

Camino de Vera s/n, 46022 Valencia, Spain.

November 30, 2015

1 Introduction

The spatial-temporal spreading of a new invasive species in a habitat has interest in ecology and is modeled by a moving boundary diffusion logistic partial differential problem, where the moving boundary represents the unknown expanding front of the species. In this paper a front-fixing approach is applied in order to transform the original moving boundary problem into a fixed boundary one. A finite difference method preserving qualitative properties of the theoretical solution is proposed. Results are illustrated with numerical experiments.

The diffusive logistic model of [2] for the density of population of the invasive species $U(t,x)$ depending on time t and spatial variable x states as follows

$$\frac{\partial U}{\partial t} - D\frac{\partial^2 U}{\partial x^2} = U(a-bU), \;\; t>0, \; 0<x<H(t), \tag{1}$$

together with the boundary conditions

$$\frac{\partial U}{\partial x}(t,0) = 0, \quad U(t,H(t)) = 0, \quad t>0, \tag{2}$$

[*]e-mail: mipigar@cam.upv.es

the Stefan condition

$$H'(t) = -\mu \frac{\partial U}{\partial x}(t, H(t)), \quad t > 0, \tag{3}$$

and the initial conditions

$$H(0) = H_0, \quad U(0, x) = U_0(x), \quad 0 \leq x \leq H_0. \tag{4}$$

The initial function $U_0(x)$ satisfies the following properties

$$U_0(x) \in C^2([0, H_0]), \quad U_0'(0) = U_0(H_0) = 0, \quad U_0(x) > 0, \ \ 0 \leq x < H_0. \tag{5}$$

Here $H(t)$ is the unknown moving boundary such that the population is distributed in the interval $[0, H(t)]$, $D > 0$ is the dispersal rate and the positive parameters a and b are the intrinsic growth rate and the intraspecific competition, respectively.

2 Transformation and discretization of the continuous problem

Let us consider the Landau transformation, [1, 3],

$$z(t, x) = \frac{x}{H(t)}, \quad W(t, z) = U(t, x). \tag{6}$$

Under substitution (6) problem (1)-(4) takes the form

$$G(t)\frac{\partial W}{\partial t} - G'(t)\frac{z}{2}\frac{\partial W}{\partial z} - D\frac{\partial^2 W}{\partial z^2} = G(t)W(a - bW), \ \ t > 0, \ 0 < z < 1, \tag{7}$$

where

$$G(t) = H^2(t), \ \ t \geq 0. \tag{8}$$

Boundary conditions (2) and Stefan condition (3) take the form

$$\frac{\partial W}{\partial z}(t, 0) = 0, \quad W(t, 1) = 0, \quad t > 0, \tag{9}$$

and

$$G'(t) = -2\mu \frac{\partial W}{\partial z}(t, 1), \quad t > 0, \tag{10}$$

respectively, while the initial conditions (4) become

$$G(0) = H_0^2, \quad W(0,z) = W_0(z) = U_0(zH_0), \quad 0 \le z \le 1. \tag{11}$$

Conditions (5) for the initial function $U_0(x)$ are translated to $W_0(z)$ as follows

$$W_0(z) \in C^2([0,1]), \quad W_0'(0) = W_0(1) = 0, \quad W_0(z) > 0, \;\; 0 \le z < 1. \tag{12}$$

Let us consider the step size discretization $k = \Delta t$, $h = \Delta z = 1/M$, and the mesh points (t^n, z_j), with $t^n = kn$, $n \ge 0$, $z_j = jh$, $0 \le j \le M$ and M positive integer. Let us consider the forward approximation of the time derivatives

$$\frac{w_j^{n+1} - w_j^n}{k} \approx \frac{\partial W}{\partial t}(t^n, z_j), \quad \frac{g^{n+1} - g^n}{k} \approx G'(t^n), \tag{13}$$

and the central approximation of the spatial derivatives

$$\frac{w_{j+1}^n - w_{j-1}^n}{2h} \approx \frac{\partial W}{\partial z}(t^n, z_j), \quad \frac{w_{j-1}^n - 2w_j^n + w_{j+1}^n}{h^2} \approx \frac{\partial^2 W}{\partial z^2}(t^n, z_j). \tag{14}$$

Then, we have the following explicit scheme

$$g^{n+1} = g^n + \frac{k\mu}{h}(4w_{M-1}^n - w_{M-2}^n), \quad n \ge 0, \tag{15}$$

$$w_j^{n+1} = a_j^n w_{j-1}^n + b_j^n w_j^n + c_j^n w_{j+1}^n, \quad n \ge 0, \quad 0 \le j \le M-1, \tag{16}$$

where the coefficients are given by

$$\begin{aligned} a_j^n &= \frac{k}{h^2}\left(\frac{D}{g^n} - \frac{z_j \mu (4w_{M-1}^n - w_{M-2}^n)}{4g^n}\right), \\ b_j^n &= 1 + k(a - bw_j^n) - \frac{k}{h^2}\frac{2D}{g^n}, \\ c_j^n &= \frac{k}{h^2}\left(\frac{D}{g^n} + \frac{z_j \mu (4w_{M-1}^n - w_{M-2}^n)}{4g^n}\right). \end{aligned} \tag{17}$$

3 Consistency

Let us consider the problem (7)-(11), denoted in vector form as $\mathcal{L}(W,G) = (\mathcal{L}_1(W,G), \mathcal{L}_2(W,G), \mathcal{L}_3(W,G))$ where equations (7),(9), (11) are written in the form

$$\mathcal{L}_1(W,G) = \frac{\partial W}{\partial t} - \frac{G'(t)}{G(t)}\frac{z}{2}\frac{\partial W}{\partial z} - \frac{D}{G(t)}\frac{\partial^2 W}{\partial z^2} - W(a - bW) = 0, \quad t > 0,\ 0 < z < 1, \tag{18}$$

$$\mathcal{L}_2(W,G) = \frac{\partial W}{\partial z}(t,0) = 0, \quad t > 0, \tag{19}$$

$$\mathcal{L}_3(W,G) = G'(t) + 2\mu\frac{\partial W}{\partial z}(t,1) = 0, \quad t > 0, \tag{20}$$

and the finite difference scheme, written together as $L(w,g)$ $= (L_1(w,g), L_2(w,g), L_3(w,g))$ where

$$L_1(w,g) = \frac{w_j^{n+1} - w_j^n}{k} - \frac{z_j}{2}\frac{w_{j+1}^n - w_{j-1}^n}{2h}\left(\frac{g^{n+1} - g^n}{g^n k}\right)$$
$$- \frac{D}{g^n}\frac{w_{j-1}^n - 2w_j^n + w_{j+1}^n}{h^2} - w_j^n(a - bw_j^n) = 0, \quad n \geq 0, \quad 0 \leq j \leq M-1, \tag{21}$$

$$L_2(w,g) = \frac{w_1^n - w_{-1}^n}{2h} = 0, \quad n \geq 0, \tag{22}$$

$$L_3(w,g) = \frac{g^{n+1} - g^n}{k} - \frac{\mu}{h}(4w_{M-1}^n - w_{M-2}^n) = 0, \quad n \geq 0. \tag{23}$$

It holds that the scheme $L(w,g)$ is consistent with the problem $\mathcal{L}(W,G)$ and the local truncation error behaves

$$T_j^n(W,G) = \mathcal{O}(k) + \mathcal{O}(h^2). \tag{24}$$

4 Positivity and Stability

The next numerical analysis result establishes that the proposed numerical scheme provides conditionally positive and stable solutions.

Theorem 1. *With previous notation, let k_0 be*

$$k_0 = \min\left\{k_1 = \frac{h^2}{\frac{\mu C}{g^0} + h^2(bC - a)}, k_2 = \frac{h^2}{\frac{2D}{g^0} + h^2(be^{aT}B(0) - a)}\right\}. \quad (25)$$

Under condition $k < k_0$ for small enough values of h the solution $\{w_j^n, g^n\}$ of the numerical scheme (15)-(17) verifies that g^n is positive monotone increasing and

$$0 \leq w_j^n \leq B(0)e^{aT};\ \ 0 \leq j \leq M,\ \ 0 \leq n \leq N,\ \ Nk = T, \quad (26)$$

where B(0) denotes the maximum value of the initial condition $U_0(x)$, $0 \leq x \leq H_0$. Consequently, the scheme (15)-(17) is $\|\cdot\|_\infty$-stable.

5 Numerical dichotomy: spreading versus vanishing

The following example is devoted to spreading case showing that the numerical solution of problem (7)-(11) computed by the proposed scheme (15)-(17) converges to the habitat carrying capacity a/b confirming that the numerical spreading occurs under spreading condition $H_0 \geq L = \frac{\pi}{2}\sqrt{\frac{D}{a}}$ given in [2].

Example 1. In the logistic diffusion model (1)-(4) with parameters values $(D, \mu, a, b, H_0) = (1, 1, 2, 1, 4)$ and $U_0 = cos(\pi x/8)$, Figure 1 shows the spreading behavior under condition $H_0 = 4.00 > L = 1.11$.

The next example illustrates the vanishing behavior of the numerical solution according to the theoretical results of [2].

Example 2. In this example we take $(D, \mu, a, b, H_0) = (0.1, 0.2, 0.04, 0.04, 1)$, with $U_0 = cos(\pi x/2)$. There is vanishing behavior with $H_0 = 1.00 < L = 2.48$ and $\mu = 0.20$. Figure 2 shows that numerical population density tends to zero and the free boundary is always upper bounded by L.

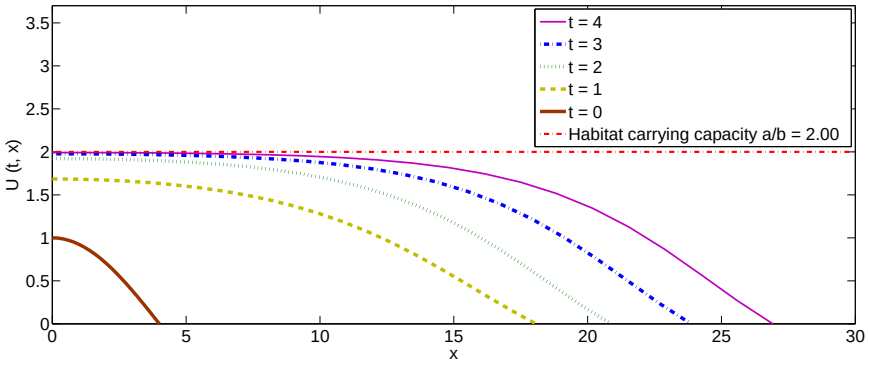


Fig. 1. Numerical solution for several values of time in a spreading case.

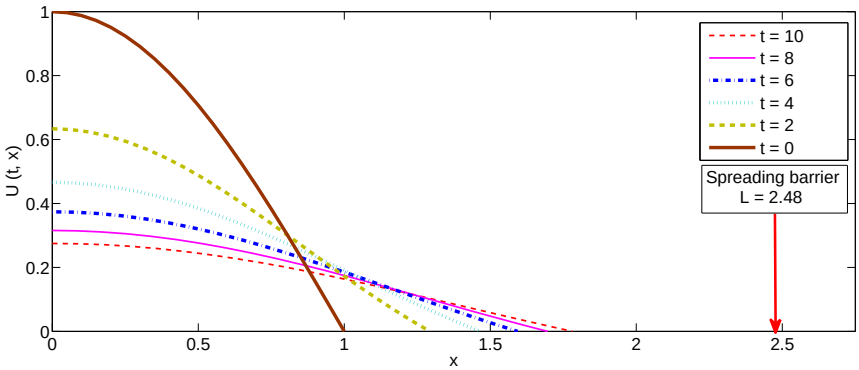


Fig. 2. Numerical solution for several values of time in a vanishing case.

References

[1] J. Crank, Free and Moving Boundary Problems, Oxford Science Publications, 1984.

[2] Y. Du, Z. Lin, Spreading-vanishing dichotomy in the diffusive logistic model with a free boundary, *SIAM J. Math. Anal.*, 42(1):377–405, 2010.

[3] H. G. Landau, Heat condution in a melting solid, *Quarterly of Applied Mathematics*, 8:81–95, 1950.

A computational study on the influence of convergent-divergent orifices on the inner flow and cavitation development in diesel injector nozzles

F.J. Salvador[♭]*, D. Jaramillo-Císcar[♭], J.-V. Romero[†] and M.-D. Roselló[†]

(♭) CMT-Motores Térmicos, Universitat Politècnica de València,

Camino de Vera s/n, Edificio 6D, 46022 Valencia, Spain,

(†) Instituto de Matemática Multidisciplinar, Universitat Politècnica de València,

Camino de Vera s/n, Edificio 8G, 2°, 46022 Valencia, Spain

November 30, 2015

1 Introduction

Nowadays modern diesel engines are facing to the improvement in the fuel consumption and the reduction of pollutant emissions. The first step before to study the combustion process or the air-fuel mixing process is to study the flow behaviour inside the diesel injector that will be the responsible for the subsequent processes.

In order to accomplish with the required improvement, non-conventional orifices have been studied in the recent years [1–5]. Those studies involve the use of elliptical nozzles [1–3] and divergent nozzles [4, 5]. The aim of this study is to increase the available information on non-conventional orifices in diesel injectors. For that reason 3 convergent-divergent nozzles have been

*Corresponding author, Dr. F. Javier Salvador, e-mail: fsalvado@mot.upv.es, Telephone: 34-963879659, Fax: 34-963877659

numerically studied. Due to the divergent component of the orifices, the cavitation phenomenon plays an important role in the flow behaviour. This kind of geometry is common in other areas as acoustic [6] and refrigeration systems [7].

The present extended abstract has been divided into 5 sections. First of all, a brief description of the computational fluid dynamics (CFD) approach used to study the cavitation phenomenon is performed in section 2. The geometry characteristics for the three convergent-divergent nozzles that have been studied are explained in section 3. The results of the study are presented in section 4 and finally, the main conclusions are drawn in section 5.

2 Description of the CFD approach

The study has been performed with a code previously validated [8,9] and able to simulate cavitation phenomenon using a homogeneous equilibrium model (HEM). The HEM is based in the assumption that the liquid and vapour phases are in each cell completely mixed and in equilibrium. The model use a barotropic equation of state to relate the density and the pressure. The compressibility on each cell depends on the liquid and the vapour phases.

The turbulence is modelled using a RANS approach. This method solves the Reynolds-averaged Navier Stokes (RANS) equations which models turbulent quantities, decomposing the fluid properties into an averaged and a fluctuating component. From previous studies the RNG k-ε model has been shown to be the best option to model the turbulence.

3 Geometry and boundary conditions

A multi-hole nozzle with 7 orifices has been used as a basis. Due to the symmetry of the geometry the full nozzle is reduced to only 1 of the 7 orifices (51.4º). The orifices of the 3 convergent-divergent nozzles that have been studied are composed as shown in the Figure 1 of an initial convergent part until the middle of the orifice and a divergent part from the middle until the outlet. The convergent and the divergent part are selected to be symmetrical, so the inlet and the outlet diameters, denoted by D_{i} and D_0 respectively, are the same. The geometrical characteristics regarding the orifices can be found at Table 1, where D_{mid} is the middle diameter and k-factor the conicity factor.

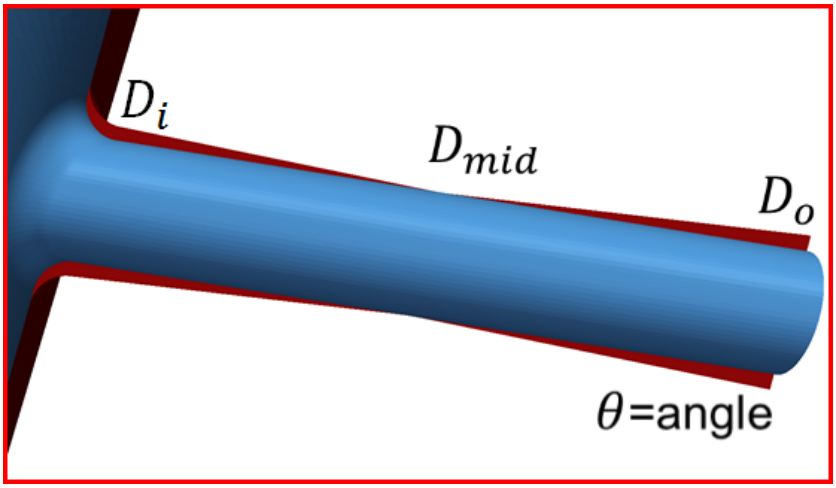


Figure 1: Orifice's geometry.

Nozzle	θ [°]	D_{mid} [µm]	D_0 [µm]	k-factor$_{\text{con}}$ [–]
CD-2.5	2.5	124.6	156.3	3.2
CD-3.5	3.5	124.6	169.0	4.4
CD-4.5	4.5	124.6	181.7	5.7

Table 1: Geometrical characteristics of the orifices

For each part of the orifice the k-factor is defined as follows:

$$k\text{-factor}_{\text{con}} = \frac{D_{\text{i}} - D_{\text{mid}}}{10\mu\text{m}} \tag{1}$$

$$k\text{-factor}_{\text{div}} = \frac{D_{\text{mid}} - D_0}{10\mu\text{m}} \tag{2}$$

As the inlet and the outlet diameters are the same both k-factors have the same absolute value but different sign. For that reason any reference to the k-factor will be in absolute value.

$$k\text{-factor}_{\text{div}} = -k\text{-factor}_{\text{con}} \tag{3}$$

For the simulations a fixed pressure condition has been used at the entry while for the outlet a mean pressure condition has been used. The mean

pressure condition allows the existence of very low pressure regions as required for the presence of vapour inside the flow, while the mean pressure is kept at the desirable value. For the walls a non-slip condition has been used for the velocity, while a zero-normal gradient for the pressure.

The study has been performed at a low injection pressure of 40 MPa and over a backpressure swept from 1 MPa to 25 MPa as shows at the Table 2.

Injection pressure [MPa]	Backpressure [MPa]
40	1,3,5,7,9,11,13,15,17,19,21,23,25

Table 2: Test matrix.

4 Results

4.1 Flow parameters

For the comparison in terms of flow parameters has been using the following parameters at the outlet, mass flow, $\dot{m}_\mathrm{f}$ (Eq. 4), momentum flux, $\dot{M}_\mathrm{f}$ (Eq. 5) and effective velocity, u_eff (Eq. 6):

$$\dot{m}_\mathrm{f} = \int \rho \left(\vec{u} \cdot \vec{n}\right) \mathrm{d}S \tag{4}$$

$$\dot{M}_\mathrm{f} = \int \rho \left(\vec{u} \cdot \vec{n}\right)^2 \mathrm{d}S \tag{5}$$

$$u_\mathrm{eff} = \frac{\dot{M}_\mathrm{f}}{m_\mathrm{f}} \tag{6}$$

Where ρ is the density, $\vec{u}$ the velocity and $\vec{n}$ the normal vector to the surface S (in this case the outlet section).

These parameters are represented at Figure 2. As can be observed the k-factor affects significally the flow behaviour. Regarding the mass flow can be observed that the nozzle CD-2.5 increases linearly with square root of the differences of pressure until a point where it stabilizes. At this point, a flow choking occurs, the backpressure condition needed to reach this situation is called critical cavitation condition (CCC). As the k-factor increases can be observed that the CCC increases, i.e. the square root of the differences of pressure decrease.

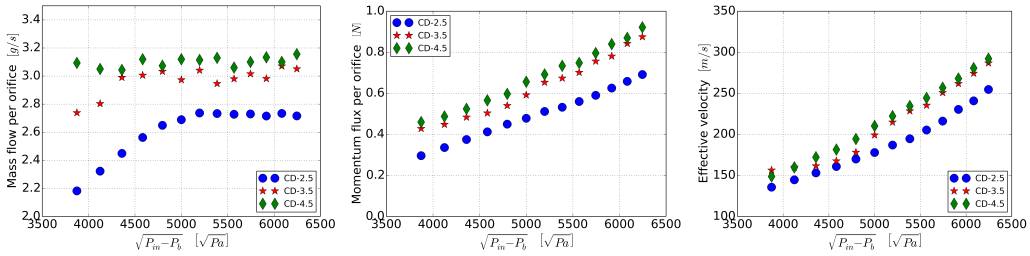


Figure 2: Flow parameters.

For the nozzle CD-3.5 almost all the studied conditions have been found in cavitating conditions. Furthermore for the nozzle CD-4.5, which is the one with the highest k-factor, all the studied conditions are at cavitating conditions. Ordering the nozzles from more to less prone to cavitate the result is CD-4.5 > CD-3.5 > CD-2.5. For the momentum flux and the effective velocity the results are similar, they increase when the backpressure decrease and the increment of the k-factor also increases the value of the parameters.

4.2 Cavitation morphology

Regarding at the cavitation pattern at the Figure 3. For the CD-2.5, cavitation phenomenon occurs as well in the rounding radius but doesn't spread. As it is known in convergent orifices cavitation phenomenon decreases when the k-factor increases. That is exactly what can be observed for the convergent part, for the CD-3.5 and CD-4.5 at the rounding radius there is no cavitation.

However in the divergent part, it can be observed that appears cavitation and wrap all the wall circularly while it spreads to the outlet. As was observed for the mass flow, when the k-factor increases the CCC increases and for the CD-4.5 there is a vapour ring even at a high backpressure of 25 MPa.

4.3 Flow coefficients

With regard to the dimensionless coefficients, three coefficients are used for the comparison. The discharge coefficient, C_d (Eq. 7), is representative of the global losses in the nozzle. The theoretical velocity used at Eq. 7 is defined at Eq. 8, where P_in is the injection pressure and P_b is the discharge

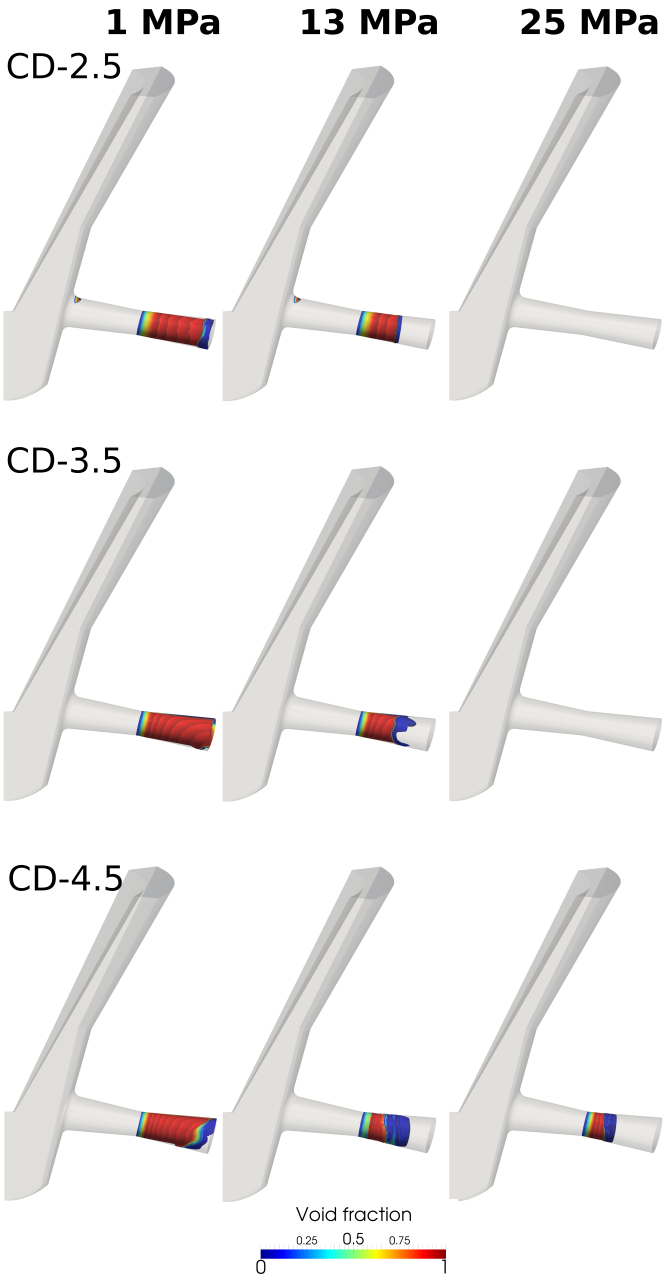


Figure 3: Cavitation pattern.

back pressure. The velocity coefficient C_{v} (Eq. 9) and the area coefficient C_{a} (Eq. 10), those 3 coefficients are related (Eq. 11). As the outlet area change with the k-factor the denominator for the discharge and the area coefficients will be different even at the same conditions of pressure.

$$C_{\mathrm{d}} = \frac{\dot{m}_{\mathrm{f}}}{\rho A_0 u_{\mathrm{th}}} \tag{7}$$

$$u_{\mathrm{th}} = \sqrt{\frac{2(P_{\mathrm{in}} - P_{\mathrm{b}})}{\rho}} \tag{8}$$

$$C_{\mathrm{v}} = \frac{u_{\mathrm{eff}}}{u_{\mathrm{th}}} \tag{9}$$

$$C_{\mathrm{a}} = \frac{A_{\mathrm{eff}}}{A_0} = \frac{\dot{m}_{\mathrm{f}}{}^2}{\rho \dot{M}_{\mathrm{f}} A_0} \tag{10}$$

$$C_{\mathrm{d}} = C_{\mathrm{v}} C_{\mathrm{a}} \tag{11}$$

As can be observed at Figure 4, at cavitating conditions the discharge and the area coefficient are ordered in the following way, CD-2.5 > CD-3.5 > CD-4.5. That is the inverse order as when they are ordered respect to the proneness to cavitate, due to that the apparition of cavitation reduce the effective area. As reminder that CD-4.5 is the only nozzle that is always cavitating the area coefficient for CD-4.5 is always lower than 1.

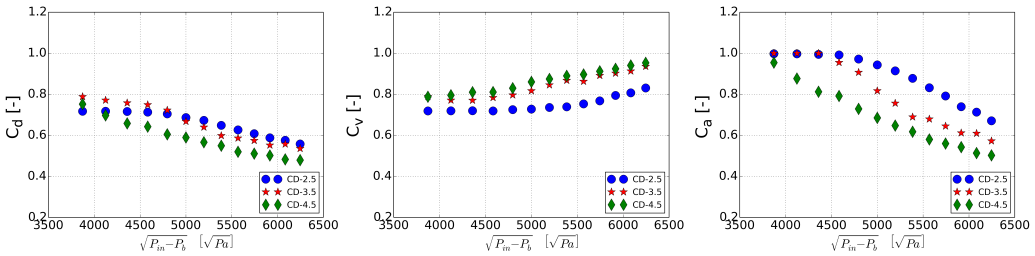


Figure 4: Dimensionless coefficients.

Regarding the velocity coefficient, it increases when the backpressure decrease and in accordance with the previous results, more cavitating geometries show higher values in the velocity coefficient due to the effect of the vapour.

5 Conclusions

The main conclusions of this study are summarized in the following points:

- Due to the divergent part of the orifice, the Convergent-Divergent nozzles are more prone to cavitate (even at high backpressure conditions). The k-factor influences the appearance of cavitation (CD-4.5 > CD-3.5 > CD-2.5).
- The effective velocity is increased with the k-factor in this types of nozzles. Together with the increase of cavitation, a better quality in the air-fuel mixing process is presupposed for higher k-factor.
- As it is known, high k-factor in convergent nozzles prevents cavitation to appear in the rounding radius. And that has been obtained. For CD-2.5, a small amount of vapour appears at the rounding radius aside the divergent part, while for CD-3.5 and CD-4.5 cavitation just appears in the divergent part but not in the rounding radius.

Acknowledgments

The authors would like to express gratitude for the computer resources, technical expertise and assistance provided by the Universidad de Valencia relating to the use of the supercomputer "Tirant".

References

[1] Chih Ming Ho and Ephraim Gutmark. Vortex induction and mass entrainment in a small-aspect-ratio elliptic jet. *Journal of Fluid Mechanics*, 179:383–405, 1987.

[2] Fazle Hussain and Hyder S. Husain. Elliptic jets. Part 1. Characteristics of unexcited and excited jets. *Journal of Fluid Mechanics*, 208:257–320, 1989.

[3] S. Molina, F.J. Salvador, M. Carreres, and D. Jaramillo. A computational investigation on the influence of the use of elliptical orifices on the inner nozzle flow and cavitation development in diesel injector nozzles. *Energy Conversion and Management*, 79:114–127, mar 2014.

[4] J. Benajes, S. Molina, C. González, and R. Donde. The role of nozzle convergence in diesel combustion. *Fuel*, 87(10-11):1849–1858, aug 2008.

[5] Ozgur Oguz Taskiran and Metin Ergeneman. Effect of nozzle dimensions and fuel type on flame lift-off length. *Fuel*, 115:833–840, jan 2014.

[6] Chih-Chung Hu, Win-Ti Lin, Chun-Min Su, and Wen-Jay Liu. Discharge characteristics of small sonic nozzles in the shape of pyramidal convergent and conical divergent. *Flow Measurement and Instrumentation*, 25:26–31, jun 2012.

[7] Jinghui Liu, Jiangping Chen, and Zhijiu Chen. Critical flashing flow in convergentdivergent nozzles with initially subcooled liquid. *International Journal of Thermal Sciences*, 47(8):1069–1076, aug 2008.

[8] F.J. Salvador, J.-V. Romero, M.-D. Roselló, and J. Martínez-López. Validation of a code for modeling cavitation phenomena in Diesel injector nozzles. *Mathematical and Computer Modelling*, 52(7-8):1123–1132, oct 2010.

[9] F. Payri, R. Payri, F.J. Salvador, and J. Martínez-López. A contribution to the understanding of cavitation effects in Diesel injector nozzles through a combined experimental and computational investigation. *Computers & Fluids*, 58:88–101, apr 2012.

Mixed Truck Delivery Systems with both Hub-and-Spoke and Direct Shipment

B. Royo[♭] [*], D. Escuin[†], A. Fraile[♭], JA. Sicilia[‡], and E. Larrodé[♭]

(♭) Universidad de Zaragoza,

Calle María de Luna 3, 50018, Zaragoza,

(†) Instituto Tecnológico de Aragón,

Calle María de Luna 7, 50018, Zaragoza,

(‡) Universidad Internacional de La Rioja (UNIR),

Gran Vía Rey Juan Carlos I, 41. 26002 Logroño, La Rioja.

November 30, 2015

1 Introduction

Progress of the supply chain management and the increase of e-commerce places the freight motor carriers to consider strategies and tactics to minimize costs while maintaining a high service quality. Therefore, [1] defines different distribution strategies networks: *many to many, hub&spoke and peddling&collecting*, which are usually studied separately.

Several studies show positive effects of mixing strategies. The following papers: [2], [3], [4], [5], [6], [7], etc., developed metaheuristics to solve different combined scenarios.

In this case, a linear programming model of a mixed delivery system solved in [6], is proposed. Previous researches have developed mathematical formulation as in [8], [9] and [10], however, to the best of our knowledge this is the first study to model a mixed delivery system with several hubs, stopovers and time constraints.

[*]e-mail:broyoa@unizar.es

2 Problem description

This paper deals with the load planning problem of lessthantruckload motor carriers, the determination of how freight should be routed through the network to minimize costs while maintaining service.
The main features of the problem are: an order consists on a pickup and a delivery node with small freight and time windows; to be managed it is allowed to go through at least one *"hub"*, to stop among direct shipments, to stop between *"hub"* node or node *"hub"* connections and the *"hubs"* are linked but stopovers are not allowed.

3 Mathematical Formulation

The following lists the assumption, the required parameters and the decision variables to formulate the problem.

Assumptions:
Each order $o \in O$ consists on a pickup $i \in N$ and a delivery node $j \in N$ with c_{ij} as the shortest distance among them. The service time in a node is taken as t_i.

Input Data:
H_o, H_d the set of origin and destination *"hubs"* respectively and R and E are the set of pickup and delivery nodes where $N = H_o \cup H_d \cup R \cup E$.
Let $d_i \in N$ the delivery node of the pickup node $i \in N$ and $a_h \in N$ the *"hubs"* pickup node of the delivery node $h \in N$.
V consists on the set of vehicles with a capacity Q and they are not allowed to arrived before e_i and after l_i and m_i as the amount of freight in node $i \in N$.

Zero one decision variables:

$$x_{ij}^{v} = \begin{cases} 1 & \text{If } v \in V \text{goes form node } i \in N \text{to node } j \in N. \\ 0 & otherwise. \end{cases} \tag{1}$$

$$y_{ih} = \begin{cases} 1 & \text{If } i \in R \cup E \text{ starts or ends in } \textit{"hubs"}\ h \in H_o. \\ 0 & otherwise. \end{cases} \tag{2}$$

$$u_{ij} = \begin{cases} 1 & \text{If a line trip goes from } \textit{"hubs"} i, \text{to } \textit{"hubs"} j \in H_o. \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

Other variables are $s_i^v, i \in N, v \in V$ is the arrival time at node i by the vehicle v and $q_i^v, i \in N, v \in V$ is the amount of available capacity of the vehicle v in the node i.

Objective function
The minimization of the total distance including the distance among *"hubs"*:

$$min \sum_{k \in V} \sum_{i \in N} \sum_{j \in N} c_{ij} x_{ij}^k + \sum_{i \in H_o} \sum_{j \in H_o} c_{ij} u_{ij} \tag{4}$$

Subject to constraints:

$$\sum_{h \in H_o} \sum_{i \in H_d \cup E} x_{hi}^v = 1 \qquad \forall v \in V \tag{5}$$

$$\sum_{i \in H_o \cup R} \sum_{h \in H_d} x_{ih}^v = 1 \qquad \forall v \in V \tag{6}$$

$$x_{dp}^v = 0 \qquad \forall v \in V, \forall p \in R, \forall d \in E \tag{7}$$

$$\sum_{v \in V} \sum_{j \in N} x_{ij}^v = 1 \qquad \forall i \in N \tag{8}$$

$$x_{ii}^v = 0 \qquad \forall v \in V, \forall i \in N \tag{9}$$

$$\sum_{i \in H_o \cup R \cup E, i != j} x_{ij}^v - \sum_{i \in H_d \cup R \cup E, i != j} x_{ji}^v = 0 \qquad \forall v \in V, \forall j \in R \cup E \tag{10}$$

5, 6 ensure no pickup node after arriving at H_o and no delivery node before leaving H_o, 7 to prevent any pickup node precedes any delivery node, 8,9 guarantee every node is visited only once and 10 is the flow conservation equation that ensures the continuity of each vehicle route.

$$1 - x^v_{d_{o_2} d_{o_1}} - \sum_{h \in H_d} \sum_{o \in R} x^v_{oh} \leq (1 - x^v_{o_1 o_2})M \qquad \forall v \in V, \forall o_1, o_2 \in R \tag{11}$$

$$1 - \sum_{o_1 \in R} \sum_{h_2 \in H_d} x^v_{o_1 h_2} - x^v_{h_2 d_{o_1}} \leq (1 - x^v_{oh})M \qquad \forall v \in V, \forall o \in R, \forall h \in H_o \tag{12}$$

$$x^v_{oe} = 0 \qquad \forall v \in V, \forall o \in R, \forall e \in E, d_o! = e \tag{13}$$

11, 12 y 13 ensure all pickup nodes precede delivery nodes following last-in-first-out policy in direct shipments.

$$1 - y_{d_o h} \leq (2 - \sum_{j \in R \cup H_o} x^v_{jh} - \sum_{j \in H_o \cup R} x^v_{oj})M \qquad \forall v \in V, \forall o \in R, \forall h \in H_d \cup H_o \tag{14}$$

$$1 - y_{dh} \leq (2 - \sum_{j \in E \cup H_d} x^v_{hj} - \sum_{j \in E} x^v_{dj})M \qquad \forall v \in V, \forall d \in E, \forall h \in H_o \tag{15}$$

$$1 - u_{h_1 h_2} \leq (2 - y_{o h_1} - y_{d_o h_2})M \qquad \forall o \in R, \forall h_1 \in H_o, h_2 \in H_d \tag{16}$$

14 and 15 identify pickup or delivery routes from a node to a H_o and from a H_d to a node respectively. 16 refers routes among *"hubs"*.

$$s^v_h = 0 \qquad \forall v \in V, \forall h \in H_d \tag{17}$$

$$s^v_i + t_i - s^v_j + c_{ij} \leq (1 - x^v_{ij})M \qquad \forall v \in V, \forall i, j \in N \tag{18}$$

$$e_i \leq s^v_i \qquad \forall v \in V, \forall i \in N \tag{19}$$

$$s_i^v \leq l_i \qquad \forall v \in V, \forall i \in N \tag{20}$$

17and 18 show how to obtain values related to time. 19 y 20 refers time windows boundaries.

$$q_i^v \leq Q \qquad \forall v \in V, \forall i \in N \tag{21}$$

$$q_h^v \leq (1 - x_{hj}^v)M \qquad \forall v \in V, \forall j \in R, \forall h \in H_o \tag{22}$$

$$q_i^v + m_j - q_j^v \leq (1 - x_{ij}^v)M \qquad \forall v \in V, \forall i \in H_o \cup R, \forall j \in R \cup H_d \tag{23}$$

$$q_i^v \leq (1 - x_{ij}^v)M \qquad \forall v \in V, \forall i \in E, \forall j \in H_d \tag{24}$$

$$q_j^v + m_j - q_i^v \leq (1 - x_{ij}^v)M \qquad \forall v \in V, \forall i \in E \cup H_o \tag{25}$$

$$q_i^v + m_j - q_j^v \leq (1 - x_{ij}^v)M \qquad \forall v \in V, \forall i \in R, \forall j \in E \tag{26}$$

Vehicles cannot exceed capacity as is ensured in 21,22,23,24,25 and 26.

4 Conclusions

To the best of our knowledge, our is the first study to develop a mathematical model representing a transport scenery where features of different problems converge as *'less than truckload, hub allocation, long distances, time windows, route planing, direct shipments allowing stopovers and hub shipments allowing stopovers'*.
This research could be extended by considering more constraints as limits on the hubs' capacity or on hubs' operational time.

References

[1] Estrada, M. Efficient distribution strategies for less-than-truckload carriers. PhD. thesis. *School of Civil Engineering, Technical University of Catalonia, Spain.*, 2007

[2] Jiyin Liu, Chung-Lun Li and Chun-Yan Chan. Mixed truck delivery systems with both hub-and-spoke and direct shipment *Transportation Research Part E: Logistics and Transportation Review*, Volume(39):325-339, 2003

[3] Kuby, M. J. and Gray, R. G. The hub network design problem with stopovers and feeders: The case of federal express. *Transportation Research Part A: Policy and Practice*, Volume(27):1-12, 1993

[4] Kuo, Y. and Wang, C.-C. A variable neighborhood search for the multidepot vehicle routing problem with loading cost. *Expert Systems with Applications*, Volume(39):6949-6954, 2012.

[5] Crevier, B., Cordeau, J.-F. and Laporte, G. The multi-depot vehicle routing problem with inter-depot routes. *European Journal of Operational Research*, Volume(176):756-773, 2007.

[6] Royo, B. Fraile, A. Larrode, E. and Muerza, V. Route planning for a mixed delivery system in long distance transportation and comparison with pure delivery systems. *Journal of Computational and Applied Mathematics*, Volume(291):488-496, 2015.

[7] Caramia, M. and Guerriero, F. A heuristic approach to long-haul freight transportation with multiple objective functions. *Omega*, Volume(37):600-614, 2009.

[8] Aykin, T. The hub location and routing problem. *European Journal of Operational Research*, Volume(83):200-219, 1995.

[9] Cheng-Chang Lin and Sheu-Hua Chen. An integral constrained generalized hub-and-spoke network design problem. *Transportation Research Part E: Logistics and Transportation Review*, Volume(44):986-1003, 2008.

[10] Hande Yaman, Bahar Y. Kara and Barbaros Ç. Tansel. The latest arrival hub location problem for cargo delivery systems with stopovers. *Transportation Research Part B: Methodological*, Volume(41):906-919, 2007.

A proposal to classify the epidemiological behavior of a network model of meningococcal C using ROC method

L. Acedo[♭] [*], R.-M. Shoucri[†], R.-J. Villanueva[‡]

(♭) Instituto Universitario de Matemática Multidiciplinar,

Universitat Politècnica de Valéncia, Valencia, Spain,

(†) Department of Mathematics & Computer Science,

Royal Military College of Canada,

(‡) Instituto Universitario de Matemática Multidiciplinar,

Universitat Politècnica de Valéncia, Valencia, Spain.

November 30, 2015

1 Introduction

Meningococcal disease is caused by the bacterium Neisseria meningitidis, also called meningococcus. About 10% of people have this type of bacteria in the back of their nose and throat with no signs or symptoms of disease, being called *carrier*. But sometimes Neisseria meningitidis bacteria can invade the body causing certain illnesses, which are known as meningococcal diseases [1].

Neisseria meningitidis bacteria are spread through the exchange of respiratory and throat secretions like spit (e.g., living in close quarters, kissing, sharing drinks). Fortunately, these bacteria are not as contagious as what causes the common cold or the flu. Besides, the bacteria are not spread by

[*]e-mail: luiacrod@imm.upv.es

casual contact or by simply breathing the air where a person with meningococcal disease has been. Sometimes Neisseria meningitidis bacteria spread to people who have had close or lengthy contact with a patient with meningococcal disease. People in the same household, roommates, or anyone with direct contact with a patient's oral secretions, meaning saliva or spit, (such as a boyfriend or girlfriend) would be considered at increased risk of getting the infection [1]

Meningitis is an infection of the brain and spinal cord and can even infect the blood. Nowadays the main cause of Meningitis is the bacterium Neisseria meningitidis. This bacterium is transmitted exclusively among humans, mainly during adolescence. An individual may get infected by contact with a carrier individual, that is, healthy carriers transmit the bacteria. It is treated with specific antibiotic, however, even properly treated, there is up to 10% of mortality and 10% of survivors have sequelae [2, 3].

Recently, network models under a computational (simulations) point of view are being used to study the short-medium term transmission dynamics of infectious diseases and to find sets of feasible model parameters by fitting model with data. Also, the network modeling allows to include randomness in a natural way in the model simulations [4, 5].

One of the main problems in the study of the transmission dynamics of Men C is that there is a lack of reliable data, because the time of carrier is very short (around 3 months), much less than the duration of a typical clinical trial. Moreover, there are facts well know, as Men C may have cyclical behavior: every 7 (more or less) years there is an outbreak of Men C in Africa.

However in a network under a computational point of view, we need tools that allow us to classify the behavior of the model outputs. With these tools:

- we could determine the area of the parameters space responsible of a determined behavior,
- the knowledge of these parameters could help us to understand better its transmission dynamics,
- the knowledge of these parameters could help us to determine if the cyclic behavior is also characteristic of other places (for instance, Europe).

2 Network model

Now, we are going to describe a network model to study the transmission dynamics of Men C. An individual can be

- Susceptible (S), healthy individual, or
- Carrier (C), infected individual who can transmit the bacterium.

We are going to consider the transmissions among adolescents (β_A) and transmission among the rest of people (β_O), where $\beta_A \geq \beta_O$. The recovering time, that is, when an individual losses the carrier state and becomes susceptible, is around 3 months.

We build a network with a million nodes with $e = k \times 10^6/2$ edges (assigned randomly) where k is the average number of contacts in the network. We also consider an underlying demographic model to assign ages to the nodes.

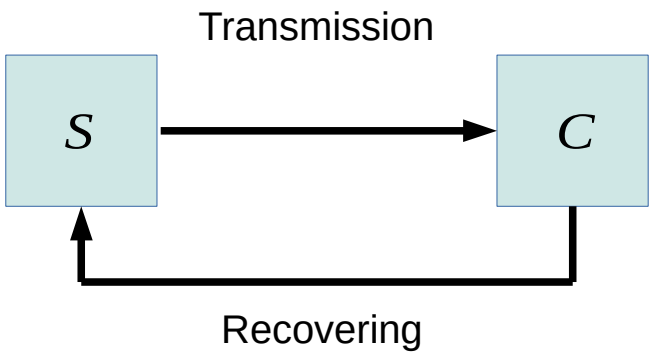


Figure 1: Flowchart of the transmission dynamics of the Men C.

The model dynamics follow the next procedure:

- Initialize the network with $1\% - 5\%$ of carriers. The time step is 1 month.
- For every time instant (month) t

  - For every node i

    – If node i is susceptible, we search the carrier nodes connected to node i

        * For every carrier node j connected to node i, we generate a random number r between 0 and 1
            - If nodes i and j are ADOLESCENTS and $r < \beta_A$, node i becomes carrier and the number of months being carrier is initialized as 0.
            - If node i or node j are NOT ADOLESCENTS, and $r < \beta_O$, node i becomes carrier and the number of months being carrier is initialized as 0.
    - If node i is carrier and the number of months being carrier is equal to 3, this node becomes susceptible. Otherwise, the number of months being carrier increases in 1 unit.

- We calculate the total number of carriers

Following the above procedure, for different model parameter values, we perform some simulations with different behavior. Some of the are illustrated in Figure 2.

3 A proposal to classify the model outputs: ROC curves

In this section we present a first approach to an automatic procedure to classify model outputs. To do that, we are going to use ROC (Receiver Operating Characteristic), a technique used in signal detection and biomedical informatics to classify the model outputs.

A ROC curve, is a graphical plot that illustrates the performance of a binary classifier system as its discrimination threshold is varied. The curve is created by plotting the true positive rate against the false positive rate at various threshold settings.

Thus, the higher the ROC curve, the better the segregation and consequently, the better classification.

We performed 46 simulations and the percentage of carriers have been retrieved from the model output. Then, they have been classified depending on the behavior shown in the Figure 2:

- Group 1 or disease-free: Curves 1, 15, 27, 28, 29, 38, 39, 42.

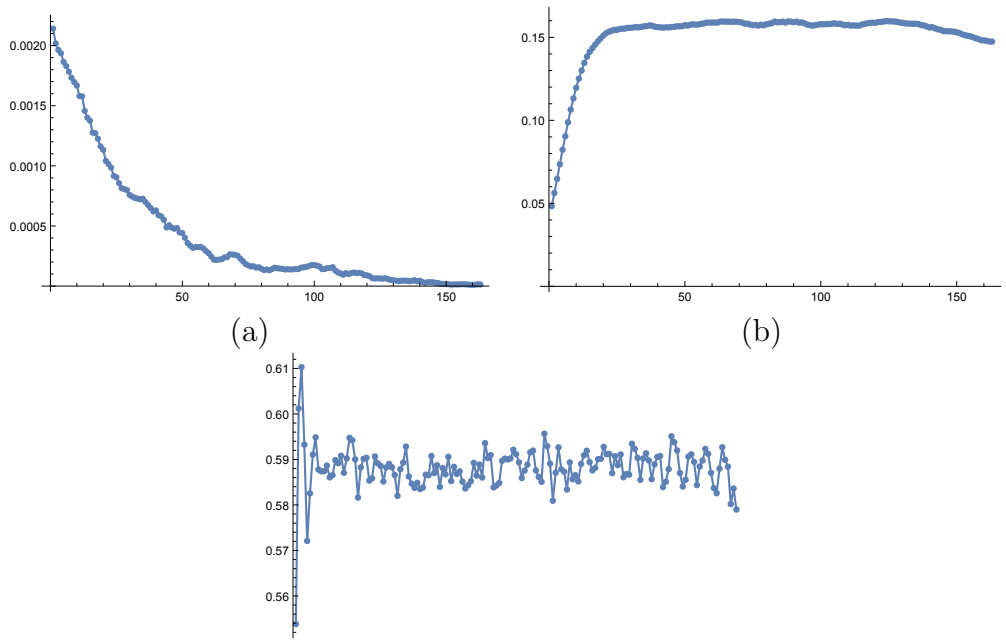


Figure 2: Three different simulations of the men C network model. (a) Disease-free. (b) Endemic. (c) Endemic with oscillations.

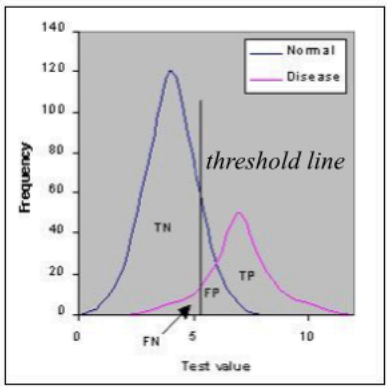
- Group 2 or endemic: Curves 2, 5, 11, 21, 23, 26, 31, 32, 34, 35, 40, 41, 43, 44, 45, 46.
- Group 3 or endemic with oscillations: Curves 3, 4, 6, 7, 8, 9, 10, 12, 13, 14, 16, 17, 18, 19, 20, 22, 24, 25, 30, 33, 36.

In order to design a classification procedure, we consider the values of a curve

$$x_{k,1}, x_{k,2}, x_{k,3}, x_{k,4}, \ldots, x_{k,n}$$

where k is the curve number and i is the time in months. Taking curve k^* and time i^*, we calculate $dx = x_{k^*,i^*} - x_{k^*,i^*-1}$. Now, we form the differences $X_d = x_{k,i^*} - x_{k,i^*-1}$ for all the curves, and we apply ROC curve to place dx in the correct group by comparing with X_d.

Bootstrap analysis and AUC (Area Under the Curve) are used for the classification in one of the three groups. The correct group corresponds to the



	p' (Predicted)	n' (Predicted)
p (Actual)	True positive	False negative
n (Actual)	False positive	True negative

Figure 3: Procedure to generate a ROC curve. The ROC curve is created by plotting the true positive rate against the false positive rate at various threshold settings curve.

ROC curve with minimum AUC (maximum AUC corresponds to maximum segregation).

3.1 Examples

Here, we present 3 examples, one for every group. In the first one, we take $k^* = 27$ (Group 1) and time $i^* = 45$. The ROC curves can be seen in the Figure 4. Minimum area correspond to correct classification, Group 1.

In the second example, we take $k^* = 11$ (Group 2) and time $i^* = 90$. The three ROC curves can be seen in the Figure 5. Minimum area correspond to correct classification, Group 2.

Finally, the third example. Here we take $k^* = 12$ (Group 3) and time $i^* = 50$. The three ROC curves can be seen in the Figure 6. Minimum AUC correspond to correct classification, Group 3.

4 Conclusion

There are epidemics where to have reliable data is very difficult. However, it is interesting to study all the possible behavior.

In network models under a computational point of view, we need tools that allow us to classify the behavior of the model outputs.

Here we have presented a first approach with the goal of classifying the behavior of the output of a network model of Men C. We have selected a difference on one curve, and then tried to classify it in one of the three groups; generally it works, we get the correct group; but there is still work to be done to improve the classification algorithm, study accuracy, etc.

References

[1] http://www.cdc.gov/meningococcal/index.html

[2] Cartwright K.,Meningococcal carriage and disease. *In: Cartwright K., editor. Meningococcal disease*. Chichester, UK: John Wiley & Sons; p.71-114, 1995.

[3] De Walls P. Immunization strategies for the control of serogroup C meningococcal disease in developed countries. *Expert Rev Vaccines* 5: 269-75, 2006.

[4] L. Acedo, J.-A. Moraño, R.-J. Villanueva, J. Villanueva-Oller, J. Diez-Domingo, Using random networks to study the dynamics of respiratory syncytial virus (RSV) in the Spanish region of Valencia, *Mathematical and Computer Modelling*, Volume 54 (7?8): 1650-1654, 2011.

[5] Gilberto González-Parra, Rafael-J. Villanueva, Javier Ruiz-Baragaño, Jose-A. Moraño, Modelling influenza A(H1N1) 2009 epidemics using a random network in a distributed computing environment, *Acta Tropica*, Volume 143: 29-35, 2015.

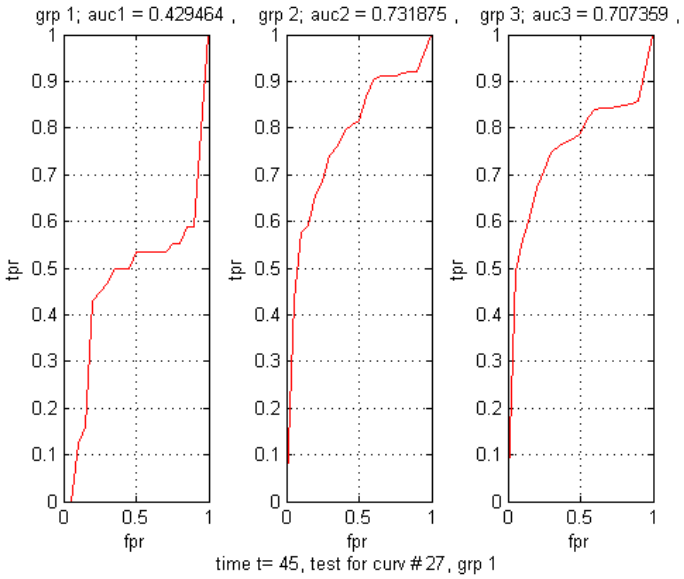


Figure 4: Example 1. The minimum AUC correspond to correct classification, Group 1.

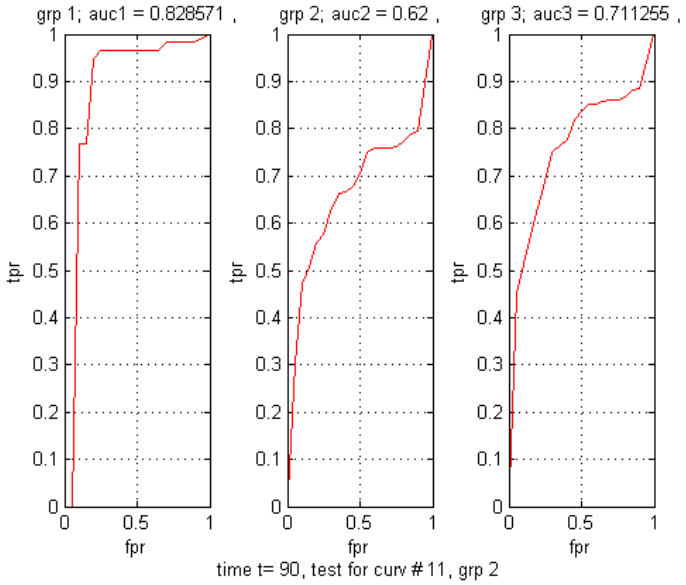


Figure 5: Example 2. The minimum AUC correspond to correct classification, Group 2.

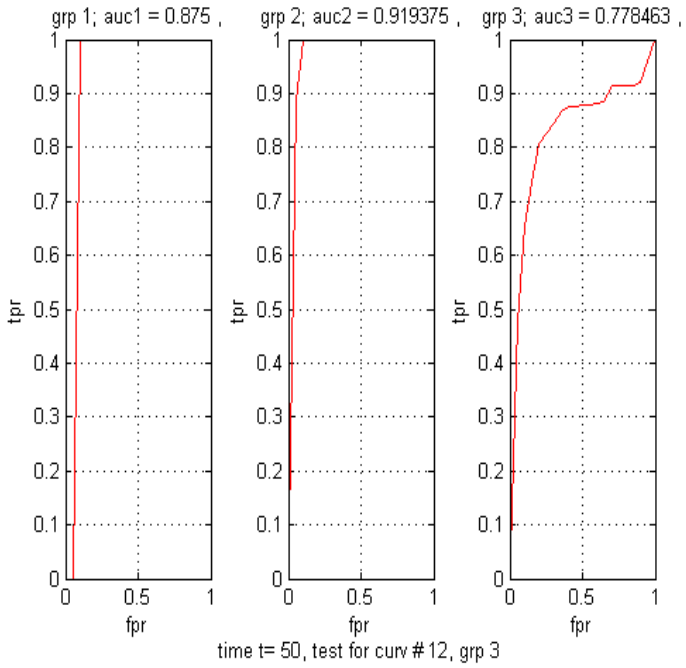


Figure 6: Example 3. The minimum AUC correspond to correct classification, Group 3.

An approach for extrapolating turbocharger compression ratio maps for engine simulations

J. Galindo[♭], R. Navarro[♭], L.M. García-Cuevas[♭], and D. Tarí[♭] *

(♭) CMT - Motores Térmicos

Universitat Politècnica de València

Camino de Vera, 46022 Valencia

November 30, 2015

1 Introduction

As pollutant regulations of internal combustion engines are becoming more restrictive year by year, manufacturers are forced to design, certify and construct new engines as fast as possible. This rush combined with the high amount of independent variables and parameters available in the engine and its subsystems lead the characterization process to be simulated using 0D-1D codes, since the modern computational units are powerful enough to handle large number of simulations at the same time without risk of damaging actual engines. Thus, the improvement of the accuracy and convergence of the codes is critical.When a turbocharged ICE needs to be simulated, the most used compressor submodel in the 0D-1D codes relies on a compressor map where the operating condition is interpolated. This work is addressed to extend a common compressor map range to assure the accuracy and convergence of the simulation.

Compressors behave quite differently in each of their regions, implying a difficulty for a single extrapolation method to predict all of them. So that, for the extrapolation approach proposed in this work, each zone is going to

*e-mail: datade@mot.upv.es

be treated separately, taking advantage of the benefits of the most suitable model for each region.

2 Assessment of available models

2.1 Jensen based models

Jensen [1] proposed a relation between the flow ϕ and load ψ dimensionless coefficients. This equation (1), is controlled with six adjustable parameters that are tuned with a least squares method using all the available points in the measured compressor map. Reasonable results are obtained generally with this method at low speeds, but due to the definition of the equation itself, it is impossible to capture the shape of the iso-speed lines within the choke region.

Martin et al. [2] used a simplified non-iso-speed-dependant Jensen equation for adjusting the compressor map (Eq.2). In this case, the coefficients are defined in a linear piece-wise between each iso-speed line.

$$\psi = \frac{k_1 + k_2 M_{inl} + k_3 \phi + k_4 M_{inl} \phi}{k_5 + k_6 M_{inl} - \phi} \tag{1}$$

$$\psi = \frac{k_{1_i} + k_{2_i} \phi}{k_{3_i} - \phi} \tag{2}$$

2.2 Leufvén model

Oskar Leufvén [3] used a generalized ellipse definition to fit the compressor map. Equation (3) reflects the ellipse definition aided with eleven adjustable parameters, which are tuned with the whole measured compressor map. This method allows low and high speed extrapolation and is able to conform fully choked flow behaviors.

$$\Pi = \Pi_{zsl} \left(1 - \left(\frac{\dot{m}^* - \dot{m}^*_{zsl}}{\dot{m}^*_{max} - \dot{m}^*_{zsl}} \right)^{C_1} \right)^{\frac{1}{C_2}} \tag{3}$$

3 Proposed approach

This section describes the developed extrapolation approach, which employs a combination of the models described in Section 2 together with enhancements for improved extrapolation.

3.1 Preprocessing

Before the extrapolation itself, some data processing must be carried out to prepare the measured points to be used. This step consists in removing redundant points if there are any and removing the positive slope region part of the line since the models cannot take it into account.

3.2 Low pressure ratio extrapolation

Of the above explained methods found in the literature, the Martin [2] model and the ellipse model developed by Leufvén [3] are going to be compared extrapolating in this zone for a certain automotive turbocharger compressor.

Figure 1 show the comparison between Martin and Leufvén models used for extrapolating measured iso-speeds at low pressure ratio. The most critical point of this extrapolation, the mass flow at $\Pi_c = 1$, is also compared with the measured set. In this case, the Leufvén model produced an error between 1 and 4 % while Martin was less accurate with an error between 4 and 8 %. These errors grow as the compressor speed increases since the choke phenomena is hard to predict due to its non-linear behavior. Therefore, it is decided to select the Leufvén model for extrapolating this region in the developed approach. In addition, blending coefficients are used to guarantee a smooth transition between measured and extrapolated points at these iso-speeds.

3.3 Low speed extrapolation

Leufvén method [3] includes the possibility of extrapolating at different compressor speeds in a direct way. Regarding the Martin model [2], since each line is adjusted individually, additional assumptions must be carried out for extrapolating different compressor speeds. In this case, the information extracted from the previous step is used. In addition to the lowest measured speed line, the extrapolated points corresponding to this actual line at low

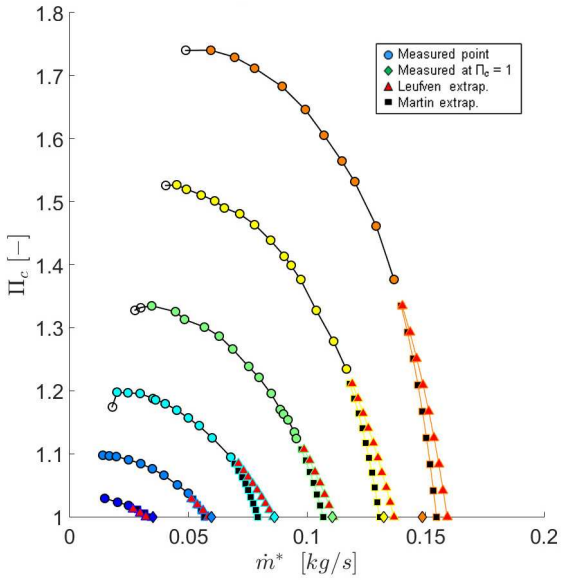


Figure 1: Comparison of low pressure ratio extrapolating methods

pressure ratio are used to tune the Jensen equation. Then, the adjusted coefficients are kept constant for the extrapolation of lower speeds. With this explained modification, the lower speed lines are consistent with the lowest low pressure ratio extrapolated line.

The extrapolation of the low speed lines is shown in Fig.2. For calculating the error of each model, the compression ratio is taken in the first place. It is calculated the mean error of all possible points in terms of compression ratio. Then, the error of the corrected mass flow at $\Pi_c = 1$ is calculated as well. Regarding the results of these errors, Martin model is more accurate. The errors in terms of compression ratio are around 7% while with Leufvén increase up to 20%. Concerning the error of corrected mass flow at $\Pi_c = 1$, the Martin model achieve values below 2% while Leufvén model is less stable. As Martin model has demonstrated stability and accuracy at extrapolating at low speeds, it is going to be selected for the final approach of this work.

3.4 High speed extrapolation

For this final region of extrapolation, only the Leufvén model is going to be compared with the measured data since the Martin model does not provide a

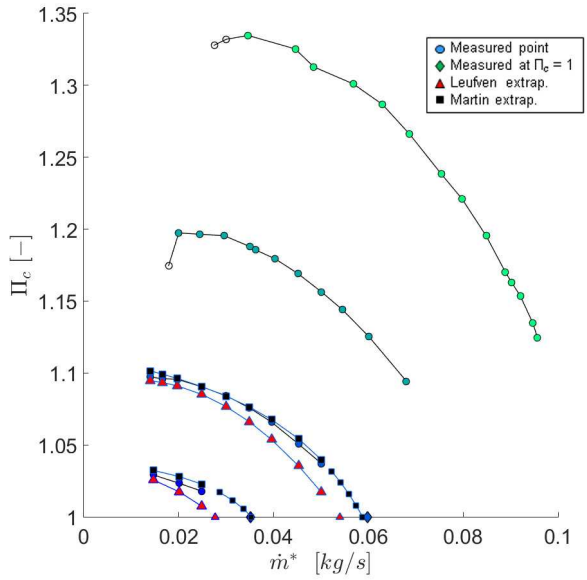


Figure 2: Comparison of low speed extrapolating methods

satisfactory high speed extrapolation approach. In addition, the non-linearity behavior of the compressor performance at this region (very high compressor speeds and choke appearance) implies difficulties when extrapolating at even higher speeds. The extrapolation using Leufvén method is direct, nevertheless special attention must be paid over the negative slope region of the line. This method can not predict negative slopes, so that this part has to be extrapolated independently. The error of the Leufvén model is calculated following the same definition used in the previous steps, using the corrected mass flow at $\Pi_c = 1$. Figure 3 shows the high speed extrapolation, where the error of the Leufvén model is found to be around 7%.

4 Concluding remarks

A complete compressor extrapolation model has been developed based on a zonal approach differentiating between low pressure ratio, high and low speeds. Several models have been considered from the literature review, implemented and compared. The best resulting model of each zone has been selected to form the final model. Additional considerations and modifications have been introduced to allow the integration of different models in

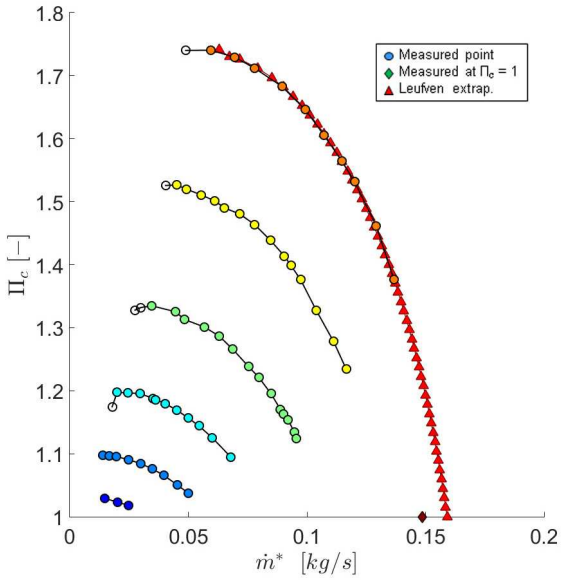


Figure 3: Comparison of high speed extrapolating methods

the same approach, avoiding inconsistencies and discontinuities, generating smooth maps which will ease the convergence of the 0D-1D simulation codes that will use them. Over ten additional maps have been tested, resulting in satisfactory extrapolations.

References

[1] Jensen, J-P., A. F. Kristensen, S. C. Sorenson, N. Houbak, et al., "Mean value modeling of a small turbocharged diesel engine," SAE Technical Paper 910070, 1991, doi:10.4271/910070.

[2] Martin, G., Talon, V., Higelin, P., Charlet, A. et al., "Implementing Turbomachinery Physics into Data Map-Based Turbocharger Models," SAE Int. J. Engines 2(1):211-229, 2009, doi:10.4271/2009-01-0310.

[3] Leufvén, O., "Modeling for control of centrifugal compressors," PhD thesis, Linköping University, 2013.

The inverse problem associated to $\{K, s+1\}$-potent matrices

L. Lebtahi[♭] [*], O. Romero[†] [†], and N. Thome[‡] [‡]

(♭) Universidad Internacional de La Rioja, E–26002 Logroño, Spain,

(†) Departamento de Comunicaciones,

Universitat Politècnica de València, E–46022 Valencia, Spain,

(‡) Instituto Universitario de Matemática Multidisciplinar,

Universitat Politècnica de València, E–46022 Valencia, Spain.

November 30, 2015

1 Introduction

In [3], the authors introduced and characterized a new kind of matrices called $\{K, s+1\}$-potent where K is involutory. We recall that a matrix $A \in \mathbb{C}^{n\times n}$ is called $\{K, s+1\}$-potent if $KA^{s+1}K = A$ and for some $s \in \{1, 2, \dots\}$. These matrices generalize all the following classes of matrices: $\{s+1\}$-potent matrices, periodic matrices, idempotent matrices, involutory matrices, centrosymmetric matrices, mirrorsymmetric matrices, circulant matrices, etc. The class of $\{K, s+1\}$-potent matrices was linked to other kind of matrices such as $\{s+1\}$-generalized projectors, $\{K\}$-Hermitian matrices, normal matrices, Hamiltonian matrices, etc. [4]. Moreover, some related results were given in [1] from an algebraic point of view. Furthermore, in [5] the authors developed an algorithm to perform this class of matrices. This problem is called the direct problem.

[*]e-mail: leila.lebtabi@unir.net

[†]e-mail: oromero@dcom.upv.es

[‡]e-mail: njthome@mat.upv.es

The aim of this paper is to design an algorithm to solve the inverse problem, that is, to find all the involutory matrices K for which a given matrix A is $\{K, s+1\}$-potent.

2 Obtaining the involutory matrices K

It is well known that the Kronecker product is an important tool to solve some matrix problems, as for example the Sylvester and Lyapunov equations. The Kronecker sum, obtained as a sum of two Kronecker products, is applied, for example, to solve the two-dimensional heat equation, to rewrite the Jacobi iteration matrix, etc. [6]. The notations $\otimes$ and $\oplus$ used in this paper refer to the Kronecker product and Kronecker sum of two matrices, respectively; and X^T denotes the transpose of the matrix X [2]. For any matrix $X = [x_{ij}] \in \mathbb{C}^{n\times n}$, let $v(X) = [v_k] \in \mathbb{C}^{n^2\times 1}$ be the vector formed by stacking the columns of X into a single column vector. The expression $[v(X)]_{\{(j-1)n+1,\ldots,(j-1)n+n\}}$, for $j = 1, \ldots, n$, denotes the j^{th} column of X.

In what follows, we will need the following property: if $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{C}^{n\times n}$ then

$$\operatorname{Ker}(A) \cap \operatorname{Ker}(B) = \operatorname{Ker}\left(\begin{bmatrix} A \\ B \end{bmatrix}\right), \tag{1}$$

which is also valid for a finite number of matrices of suitable sizes, where Ker(.) denotes the null space of the matrix (.).

We recall that the principal idempotents associated with the eigenvalues $\lambda_1, \ldots, \lambda_l$ are given by

$$P_t = \frac{p_t(A)}{p_t(\lambda_t)} \qquad \text{where} \qquad p_t(\eta) = \prod_{\substack{i=1 \\ i\neq t}}^{l} (\eta - \lambda_i). \tag{2}$$

By using the function φ and the projectors introduced in [3], it is possible to consider the matrix

$$M = \begin{bmatrix} P_0^T \oplus -P_{\varphi(0)} \\ P_1^T \oplus -P_{\varphi(1)} \\ \vdots \\ P_{(s+1)^2-2}^T \oplus -P_{\varphi((s+1)^2-2)} \\ P_{(s+1)^2-1}^T \oplus -P_{(s+1)^2-1} \end{bmatrix}. \tag{3}$$

The algorithm presented below solves the aforementioned inverse problem.

ALGORITHM

Inputs: A matrix $A \in \mathbb{C}^{n\times n}$ and an integer $s \geq 1$.

Outputs: All the involutory matrices $K \in \mathbb{C}^{n\times n}$ such that A is a $\{K, s+1\}$-potent matrix.

Step 1 Compute $\sigma(A) = \{\sigma_1, \sigma_2, \ldots, \sigma_l\}$.

Step 2 Define $\Lambda = \{0\} \cup \Omega_{(s+1)^2-1} = \{\lambda_0, \lambda_1, \ldots, \lambda_{(s+1)^2-1}\}$ in the following order $0, \omega^1_{(s+1)^2-1}, \ldots, \omega^{(s+1)^2-2}_{(s+1)^2-1}, 1$.

Step 3 If $\sigma_i \notin \Lambda$ for some $i \in \{1, 2, \ldots, l\}$ or $A^{(s+1)^2} \neq A$ then go to Step 13.

Step 4 Rearrange the eigenvalues σ_i's as in Λ. Rename them by $\{\beta_1, \beta_2, \ldots, \beta_l\}$.

Step 5 For every $i = 1, 2, \ldots, l$, identify β_i with the corresponding λ_j, for all $j = 0, 1, \ldots, (s+1)^2 - 1$. The set of these subscripts j's will be called $\{j_1, j_2, \ldots, j_l\}$.

Step 6 For every $t \in \{1, 2, \ldots, l\}$, compute the principal idempotents associated with $\{\lambda_{j_1}, \lambda_{j_2}, \ldots, \lambda_{j_l}\}$ by using (2).

Step 7 Compute $\varphi(j_1), \varphi(j_2), \ldots, \varphi(j_l)$.

Step 8 Define the matrix M as in (3) containing only the row blocks corresponding to the eigenvalues $\{\lambda_{j_1}, \lambda_{j_2}, \ldots, \lambda_{j_l}\}$.

Step 9 Compute the null space of the matrix M obtaining the solution $v(K)$.

Step 10 If $v(K) = 0$ then go to Step 13.

Step 11 Reshape the entries of $v(K)$ in the square matrix K obtaining

$$K = \left[\ [v(K)]_{\{1,\ldots,n\}} \quad [v(K)]_{\{n+1,\ldots,2n\}} \quad \ldots \quad [v(K)]_{\{(n^2-n+1,\ldots,n^2\}}\ \right]$$

Step 12 The outputs are those matrices K's that satisfy $K^2 = I_n$. Go to End. Otherwise, go to Step 13.

Step 13 'There is no matrix $K \in \mathbb{C}^{n\times n}$ such that A is $\{K, s+1\}$-potent'.

End

Example 1 *For* $s = 4$ *and*

$$A = \begin{bmatrix} i & 0 & 0 \\ 0 & 5 & -2 \\ 0 & 15 & -6 \end{bmatrix}$$

our algorithm gives all the following solutions K*:*

$$K = \begin{bmatrix} z_9 & 0 & 0 \\ 0 & -5z_1 + 6z_5 & 2z_1 - 2z_5 \\ 0 & -15z_1 + 15z_5 & 6z_1 - 5z_5 \end{bmatrix},$$

where $z_1, z_5, z_9 \in \{-1, 1\}$.

References

[1] G. Bourgeois Similar powers of a matrix *Linear and Multilinear Algebra*, (61): 699–709, 2013.

[2] P. Lancaster, The Theory of Matrices, 2nd Edition. London, Academic Press, 1985.

[3] L. Lebtahi, O. Romero, N. Thome, Characterizations of $\{K, s + 1\}$-Potent Matrices and Applications *Linear Algebra and its Applications*, (436): 293-306, 2012.

[4] L. Lebtahi, O. Romero, N. Thome, Relations between $\{K, s+1\}$-Potent Matrices and Different Classes of Complex Matrices *Linear Algebra and its Applications*, (438): 1517–1531, 2013.

[5] L. Lebtahi, O. Romero, N. Thome, Algorithms for $\{K, s+1\}$-potent matrix constructions *Journal of Computational and Applied Mathematics*, (249): 157–162, 2013.

[6] D.S. Watkins, Fundamentals of Matrix Computations, 3rd Edition. New Jersey, Wiley, 2010.

High-order iterative methods for solving nonlinear models *

A. Cordero[♭], A. Franques[♭], and Juan R. Torregrosa[♭] †

(♭) Instituto de Matemáticas Multidisciplinar,

Universitàt Politecnica de València, Camino de Vera, s/n, 46022-Valencia, Spain

November 30, 2015

1 Introduction

The *polar decomposition* is a generalization to complex matrices of the trigonometric representation of a complex number. Specifically, let A be a complex matrix of size $m \times n$, $m \geq n$ (in other case, we work with the transpose matrix). Then there exist a matrix $U \in \mathbb{C}^{m\times n}$, with orthonormal columns and a Hermitian positive semi-definite $H \in \mathbb{C}^{n\times n}$ such that

$$A = UH, \qquad U^*U = I_n, \tag{1}$$

where U^* denotes the conjugate transpose of U and I_n is the identity matrix of size $n \times n$. The Hermitian factor H is always unique and can be expressed as $H = (A^*A)^{1/2}$. If matrix A has full rank, then H is positive definite and the unitary factor U is uniquely determined. The polar decomposition is well known and can be found in many textbooks, see, for example, [4].

Let us observe that, once the unitary factor U is calculated, the other factor is obtained in a simple way, $H = U^*A$. So, our goal in this work is to obtain factor U.

*This research was partially supported by Ministerio de Economía y Competitividad MTM2014-52016-C2-2-P.

†e-mail: jrtorre@mat.upv.es

In this paper, we are interested in computing the polar decomposition by means of an iterative method of the fixed-point form $U_{k+1} = G(U_k)$, provided that the initial guess matrix U_0 is given. Let us remember that from the unitary factor U the other factor of the polar decomposition is obtained easily. In [3] Higham proposed a fixed-point algorithm based on Newton's method to compute the square root of a number, for obtaining the unitary factor U of a nonsingular $n \times n$ matrix A. Starting with $U_0 = A$, the sequence U_k is computed by

$$U_{k+1} = \frac{1}{2}(U_k + U_k^{-*}), \tag{2}$$

where U_k^{-*} denotes $(U_k^{-1})^*$. The quadratic convergence of sequence $\{U_k\}_{k\geq 0}$ was proved.

Fifteen years later, Du in [1] generalized Higham's algorithm for rectangular matrices by means of the iterative expression

$$U_{k+1} = \frac{1}{2}(U_k + U_k^{\dagger *}), \tag{3}$$

where $U_k^{\dagger}$ denotes the Moore-Penrose pseudoinverse of U_k. This scheme keeps the order of convergence of the previous one.

In a similar way as Hihgam with Newton's method, Gander in [2] used Halley's scheme for scalar equations for designing the following algorithm that converges to the unitary factor with order of convergence three for nonsingular matrices.

$$U_{k+1} = \left[U_k(3I + U_k^* U_k)\right]\left[I + 3U_k^* U_k\right]^{-1}. \tag{4}$$

Recently, a fourth-order iterative method for computing the polar decomposition was developed by Khaksar and Soleymani in [5] from a fourth-order method for solving nonlinear equations. The iterative expression of this scheme is

$$U_{k+1} = \left[U_k(7I + Y_k)(I + 3Y_k)\right]\left[I + 18Y_k + 13Z_k\right]^{-1}, \tag{5}$$

where $Y_k = U_k^* U_k$, $Z_k = Y_k Y_k$ and $U_0 = A$.

The rest of the paper is organized as follows: in Section 2, a new root-finding scheme for scalar equations is designed and, from it, an iterative scheme for computing the polar decomposition of any rectangular matrix is

derived. The sixth-order of convergence, for as proper initial matrix, of the designed method is proving. Some numerical test are presented in Section 3 to confirm the theoretical results and for comparing our scheme with other known ones.

2 The proposed iterative scheme and its convergence

Many multi-point iterative schemes for solving nonlinear scalar equations $f(x) = 0$ have been designed in the last years. The main interest of these fixed-point methods is the possibility to reach any order of convergence. The connection between the matrix iterations for computing polar decomposition, or in general for approximating the solution of a nonlinear matrix equation, and iterative methods for nonlinear scalar equations were described by Iannazzo in [6].

In fact, the matrix form of fixed-point type methods for polar decomposition is the generalization of applying the nonlinear equation solvers to the matrix equation

$$F(U) := U^*U - I = 0,$$

where I is the identity matrix of the appropriate size. This reveals the relation between the polar decomposition and matrix sign function.

Let us consider the following iterative expression for finding the simple zeros of a nonlinear equation $f(x) = 0$

$$\begin{aligned} y_k &= x_k - \frac{10 - 4L_f(x_k)}{10 - 9L_f(x_k)} \frac{f(x_k)}{f'(x_k)}, \\ x_{k+1} &= y_k - \frac{f(y_k)}{f'(y_k)}, \end{aligned} \tag{6}$$

where $L_f(x_k) = \dfrac{f(x_k)f''(x_k)}{f'(x_k)^2}$ is the degree of logarithmic convexity. This scheme is the composition of the scheme proposed in [7] for finding matrix sign functions and Newton's method. By using Taylor expansion of the different elements of the iterative expression (6), we can prove the following result.

Theorem 1 *Let $\alpha \in I$ be a simple zero of a sufficiently differentiable function $f : I \subseteq \mathbb{R} \to \mathbb{R}$ for an open interval I, and $x_0 \in I$ an initial guess*

close enough to α. Then, iterative expression (6) converge to α with order of convergence six, being its error equation

$$e_{k+1} = \frac{1}{25}c_2(c_2^2 - 5c_3)^2 e_k^6 + O(e_k^7),$$

where $c_j = \dfrac{1}{j!}\dfrac{f^{(j)}(\alpha)}{f'(\alpha)}$, $j = 2, 3, \ldots$, *and* $e_k = x_k - \alpha$.

Solving the equation $u^2 - 1 = 0$ by (6), we have the iterative expression in the reciprocal form

$$u_{k+1} = \frac{36u_k + 314u_k^3 + 384u_k^5 + 66u_k^7}{4 + 141u_k^2 + 435u_k^4 + 211u_k^6 + 9u_k^8}, \quad k = 0, 1, \ldots \tag{7}$$

The global convergence of this iterative formula on quadratic polynomials allows us to extend it in the matrix form

$$U_{k+1} = \left[U_k(36I + 314Y_k + 384Z_k + 66X_k)\right]\left[4 + 141Y_k + 435Z_k + 211X_k + 9W_k\right]^{-1}, \tag{8}$$

where $Y_k = U_k^* U_k$, $Z_k = Y_k Y_k$, $X_k = Y_k Z_k$ and $W_k = Z_k Z_k$.

Expression (8) is a new iterative fixed-point scheme for finding the polar decomposition via calculating the unitary matrix U. In the following result we analyze the convergence of sequence $\{U_k\}_{k\geq 0}$ generated by the iterative scheme (8).

Theorem 2 *Let A be an arbitrary $m \times n$ complex matrix of rank r. Then, the sequence of matrix iterates $\{U_k\}_{k\geq 0}$ obtained from (8) converges to the unitary factor U, with order of convergence six, for $U_0 = A$.*

3 Numerical results

In this section we are going to present the numerical results. The numerical tests have been made in Matlab in double precision arithmetics, with 50 digits of mantissa. The computer specifications are Intel(R) Core(TM), i5-2500, CPU 3.30 GHz, with 16 GB of RAM. We compare our scheme, denoted as CTM, with several known iterative methods such as (3) denoted by NM, (4) denoted by HM and (5) denoted by KSM, of orders of convergence two,

three and four, respectively, which also require one inverse per iteration. We use two different stopping criterium: the difference between the last iterates

$$\|U_{k+1} - U_k\|_M < tol$$

and this other one

$$\left\|U_{k+1}^* U_{k+1} - I\right\|_M < tol,$$

where *tol* is the tolerance and $\|\cdot\|_M$ is a proper matrix norm.

Example 1 In this experiment, we analyze the behavior of the different methods on a random rectangular matrix of size 510×500 generated in Matlab by $A = rand(510, 500)$, working in double precision arithmetics.

In Table 1 we show the results obtained by applying the different methods on matrix A, with tolerance $tol = 10^{-10}$ and initial estimation $U_0 = A$. Specifically, we present the values of $\|U_{k+1} - U_k\|_2$ and $\|U_{k+1}^T U_{k+1} - I_{500}\|_2$ in the last iteration, where I_{500} is the identity matrix of order 500, the number of iterations and the elapsed time, in seconds, being the mean execution time for 50 performances of each method (the command cputime of Matlab has been used).

	Newton	Halley	KSM	CTM
iterations	13	9	7	6
$\|U_{k+1} - U_k\|_2$	1.70e-14	1.12e-15	1.64e-14	1.31e-15
$\|U_{k+1}^T U_{k+1} - I_{500}\|_2$	3.34e-15	1.50e-15	1.51e-15	1.31e-15
e-time	5.67	1.93	1.84	1.78

Table 1: Numerical results for a random matrix of size 510×500 in double precision aritmethics

Example 2 In this example, we analyze the behavior of the different methods on Hilbert matrix of size 75×75 generated in Matlab by $B = hilb(75)$, working in double precision arithmetics and under the same conditions as Example 1. Let us remember that it is an example of ill-conditioned matrix.

In this case, the number of iterations increases significantly, in fact Newton's method is not convergent. However, the elapsed time decreases because matrix B has a small size.

	Newton	Halley	KSM	CTM
iterations	> 1000	46	26	22
$\|U_{k+1} - U_k\|_2$	-	8.88e-16	2.49e-11	4.26e-15
$\|U_{k+1}^T U_{k+1} - I_{75}\|_2$	-	1.11e-15	1.30e-15	8.41e-16
e-time	-	0.16	0.11	0.10

Table 2: Numerical results for a Hilbert's matrix of size 75×75

References

[1] Du K. The iterative methods for computing the polar decomposition of rank-deficient matrix. *Appl. Math. Comput.*, 162: 95–102, 2005.

[2] Gander W. Algorithms for the polar decomposition. *SIAM J. Sci. Statist. Comput.* 11(6): 1102–1115, 1990.

[3] Higham N.J. Computing the polar decomposition - With applications. *SIAM J. Sci. Statist. Comput.*, 7: 1160–1174, 1998.

[4] R.A. Horn and C.A. Johnson, Matrix Analysis, Cambridge University Press, London, 1985.

[5] Khaksar F. and Soleymani F. On a fourth-order matrix method for computing polar decomposition. *Comp. Appl. Math.*, 34: 389–399, 2015.

[6] Iannazzo B. A family of rational iterations and its application to the computation of the matrix Pth root. *SIAM J. Matrix Analysis and Applications*, 30: 1445–1462, 2008.

[7] M. Sharifi M., Karimi S., Khaksar F., Arab M., Shateyi S. On a cubically convergent iterative method for matrix sign. *The Scientific World Journal*, Volume 2015 (2015), Article ID 964257, 6 pages.

Implementation/Adaptation of a Total Variation Diminishing (TVD) scheme to a non-linear 1D finite volume method for engine gas-exchange modelling

A.J. Torregrosa*, A. Broatch, F.J. Arnau, and M. Hernández

CMT - Motores Térmicos. Universitat Politècnica de València.

Camino de Vera s/n. 46022–Valencia, Spain.

November 30, 2015

1 Introduction

A suitable compromise between the quality of the solution and the computational cost when predicting wave dynamics in intake and exhaust systems of internal combustion engines is provided by quasi-3D models, in which the momentum equation is solved in a simplified fashion on a staggered mesh [1]. However, those methods are affected by overshoots at discontinuities when used in their basic form. It was shown in previous work by the authors [2] that a Flux Corrected Transport (FCT) methodology provided satisfactory results through the combination of dissipation via damping together with the phoenical form of the anti-diffusion term. However, FCT methods can distort the solution and produce significant errors in mass conservation under some circumstances, and thus flux limiters based on the total variation diminishing (TVD) criterion [3] are usually preferred.

In this paper, the benefits of rendering the quasi-3D model equations in TVD form were explored, both in terms of the removal of flow oscillations and of the quality of the frequency domain behaviour of the solution.

*e-mail: atorreg@mot.upv.es

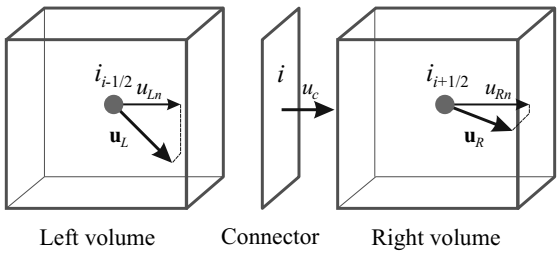


Figure 1: Basic mesh elements, definition of velocity projections and notation of volumes and connectors.

2 Model formulation

Full details of the basic model can be found in [2]. Two basic elements are considered: volumes and connector. Volumes contain information about scalar magnitudes, and connectors contain information on vector quantities. Therefore, the continuity and energy equations are solved in the volumes, the resulting discretized form being:

$$\rho^{n+1} = \rho^n + \frac{\Delta t}{V} \sum_{c=1}^{N_c} W_c^n$$

$$(\rho e_0)^{n+1} = (\rho e_0)^n + \frac{\Delta t}{V} \sum_{c=1}^{N_c} W_c^n e_0^n + \frac{\Delta t}{V} \sum_{c=1}^{N_c} p_c^n u_c^n A_c$$

where ρ is the density, p is the pressure and e_0 is the specific internal energy (superscript n indicates the time step), Δt represents the time interval, V the cell volume , N_c the number of connectors and $W_c^n = \rho_c^n u_c^n A_c$, where u is the flow velocity and A the equivalent section (subscript c refers to variables taken at the connectors, and otherwise the variable is taken at the volumes).

Momentum is computed at the connectors by projecting the flow velocity in the connected volumes onto the direction orthogonal to the connector surface (see Figure 1). As a result, a one-dimensional momentum equation is solved whose form after discretization is:

$$W_c^{n+1} = W_c^n + (\Delta t/\Delta L) \left[(\rho u_n^2 + p)_L + (\rho u_n^2 + p)_R \right] A_c$$

where u_{Ln} and u_{Rn} are the projections of the velocity at the volumes (see Figure 1).

Considering the specific structure of the staggered grid discretized equations, where the solution of the momentum equation in the connector is used to compute the fluxes for the mass and energy equations in the volumes, it appears that avoiding the numerical oscillations in the momentum equation is sufficient to stabilize the method. Therefore, the flux limiter should be added to me momentum equation, using the needed variables from the neighbour connectors. The momentum expression was thus modified so that the overall scheme is rendered in TVD form, for which the particular method proposed by Davis [4] for 1D finite difference methods in collocated grids was adapted to the present formulation. This choice is justified in view that, for a quality of results comparable to that given by the original scheme proposed by Harten [3], the increase in the computational cost with respect to the basic formulation is considerably smaller (around an order of magnitude), which may be critical for the practical viability of the method. The resulting modified momentum equation is:

$$\bar{W}^{n+1} = W^{n+1} + \left[\bar{G}^{+}(r_i^{+}) + \bar{G}^{-}(r_{i+1}^{-})\right] \Delta W^{n}_{i+\frac{1}{2}} - \left[\bar{G}^{+}(r_{i-1}^{+}) + \bar{G}^{-}(r_i^{-})\right] \Delta W^{n}_{i-\frac{1}{2}}$$

where $\bar{G}^{\pm}(r_i^{\pm}) = \left[1 - \min(r_i^{\pm}, 1)\right]/8$, and

$$r_i^{\pm} = \left[\Delta W^{n}_{i-\frac{1}{2}}, \Delta W^{n}_{i+\frac{1}{2}}\right] / \left[\Delta W^{n}_{i\pm\frac{1}{2}}, \Delta W^{n}_{i\pm\frac{1}{2}}\right]$$

Here $[\cdot,\cdot]$ is a scalar product, and $\Delta W^{n}_{i+\frac{1}{2}} = W^{n}_{i+1} - W^{n}_{i}$. Finally, the momentum vector of each volume can be written as a weighted vector average of those corresponding to the adjacent connectors:

$$(\rho_c \mathbf{u} V)_v^{n+1} = \frac{1}{2} \sum_{c=1}^{N_c} (\rho \mathbf{u}_c A_c \Delta L)_c^{n+1}$$

3 Application to the shock-tube problem

The method was applied to this simple case [5], in which two gases with different thermo- and fluid-dynamic states are put into contact, so that the contact discontinuity propagates with the flow velocity, whereas a shock wave propagates in the same direction with the speed of sound plus the flow velocity, and a rarefaction wave propagates in the opposite direction at the speed of sound minus the flow velocity. The fact that all kinds of possible propagating perturbations (contact discontinuity, and shock and rarefaction

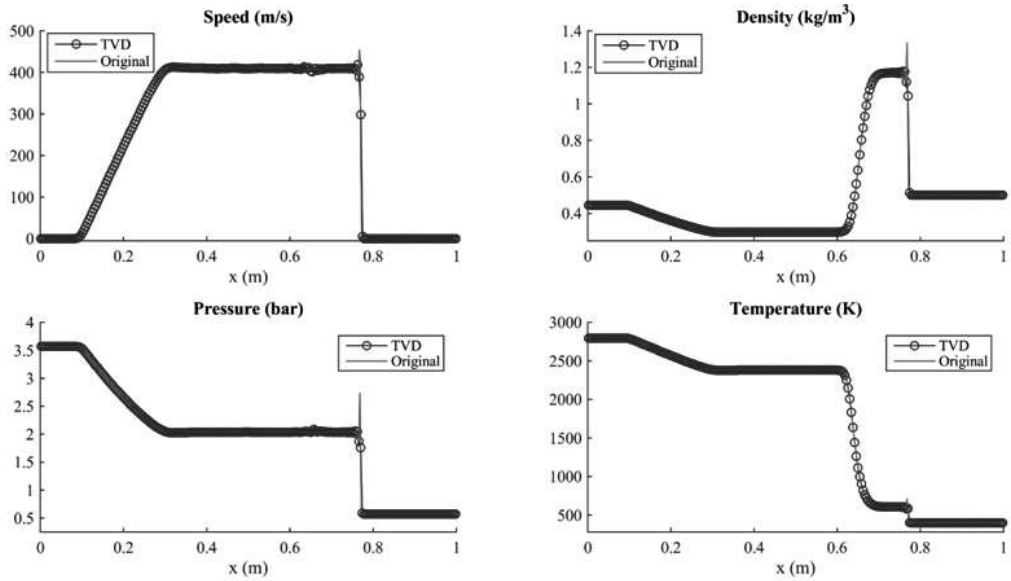


Figure 2: Shock tube results: comparison of initial solution and proposed TVD method.

waves) are present and easy to identify in the solution, has made that this problem is widespread used in all the literature.

The results obtained are shown in Fig. 2, in which comparison is given between the results obtained from the basic method and those obtained with the proposed TVD scheme. A noticeable overshooting associated with the propagation of the shock wave can be clearly seen in the basic solution, whereas it can be observed that the overshoots have been successfully removed by the inclusion of the additional TVD terms, even if the wavefronts are smoothed to a certain extent. It is also apparent that the contact discontinuity spreads in space, and thus the method is introducing some diffusion at this point due to the procedure used for the solution of the mass and energy equations. This effect, however, should not be considered as a serious shortcoming of the method from the point of view of its practical application, as this is likely to provide a more realistic representation of actual discontinuities occurring in real flow situations.

4 Conclusion

A staggered-mesh time domain model which makes use of a non-linear second order time and space discretization based on finite volumes has been complemented with suitable terms so that a Total Variation Diminishing formulation has been achieved. This has allowed suppressing overshoots at discontinuities. However, the behaviour of the solution in the vicinity of the discontinuities present in the shock-tube problem indicates that additional work is needed so that the dissipative character of the method may be reduced.

References

[1] G. Montenegro, A. Onorati and A. Della Torre. The prediction of silencer acoustical performances by 1D, 1D-3D and quasi-3D non-linear approaches. *Computers & Fluids*, 71: 208–223, 2013.

[2] A.J. Torregrosa, A. Broatch, F.J. Arnau and M. Hernández. A non-linear quasi-3D model with Flux-Corrected-Transport for engine gas-exchange modelling. *Journal of Computational and Applied Mathematics*, 291: 103-111, 2016.

[3] A. Harten. High resolution schemes for hyperbolic conservation laws. *Journal of Computational Physics*, 49: 357–393, 1983.

[4] S.F. Davis. A simplified TVD finite difference scheme via artificial viscosity. *SIAM Journal on Scientific and Statistical Computing*, 8: 1–18, 1987.

[5] G.A. Sod. Survey of several finite-difference methods for systems of non-linear hyperbolic conservation laws. *Journal of Computational Physics*, 135: 172–186, 1997.

Vibrations induced on the railway structure by the vehicle passage on a turnout. Study of mitigation measures

J. L. Velarte[1*], A. E. Blanco[1], S. Morales[1], J. Real[1]

[1]Institute for Multidisciplinary Mathematics, Polytechnic University of Valencia, 46022, Valencia, Spain

November 30, 2015

1. Introduction

A turnout is a simple mechanism that allows the intersection of two different tracks at the same level. Nevertheless, the presence of singular points such as crossing noses or switch blades hinders its dynamic behavior. The main reason lies in the rail geometry variations, which induces irregularities in the wheel-rail contact. As a consequence, great dynamic forces take place [1] which damage both vehicle and track elements and force to the railway operator to adopt expensive corrective measures.

In order to delve into this topic, this investigation studies the vertical dynamic response of a railway turnout once some corrective measures are introduced. For this purpose, a numerical feed-back interaction between a Finite Elements model of the track and a multi-body model of the vehicle is developed. Then, the results of both models are coupled and the validity of the whole process is assessed with real data obtained during a data gathering campaign.

2. Methodology

In order to assess the influence of different corrective measures in the vibration generation phenomenon, a numerical feed-back interaction between a Finite Elements model of a track developed by means of ANSYS software and a multi-body model of a rail vehicle developed by means of VAMPIRE software is proposed (Figure 1).

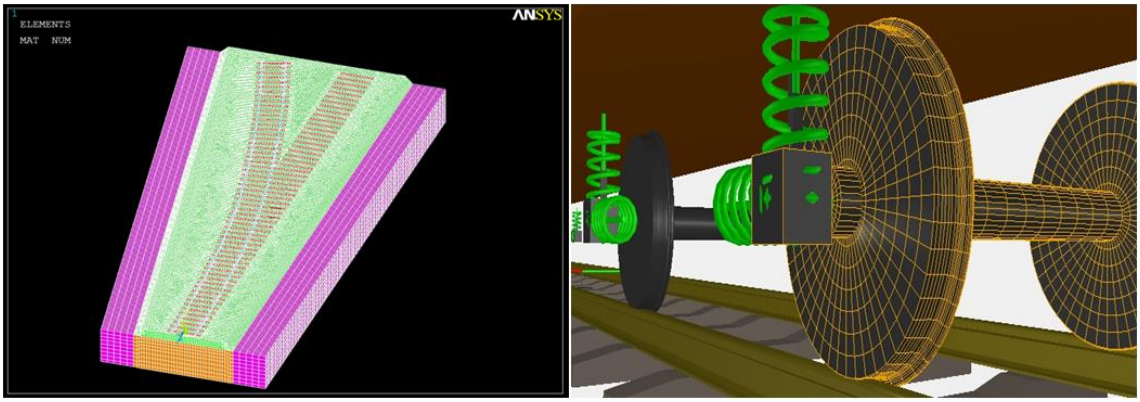


Figure 1: Track model (left) and vehicle model (right)

Nevertheless, the process is not trivial due to the lack of knowledge of some mechanical track parameters. Thus, in order to overcome this setback, the following process is proposed:

In a first step, once the FE geometry is performed, a calibration process of the track materials is carried out with the accelerations registered during the data gathering campaign in a point located far enough from any track singularity. Then, the equivalent track stiffness is obtained as the quotient between the load applied (Qa) and the track deformation (δ).

In a second step, once the multi-body model is developed, the aforementioned stiffness is incorporated together with other track parameters and a non-Linear transient analysis is performed in order to obtain the wheel-rail contact forces.

In a third step, the aforementioned wheel-rail contact forces are incorporated to the FE model of the track, beginning a validation process with the accelerations registered during the data gathering campaign in the crossing nose rail web.

Finally, once the whole process is validated, both FE and multi-body models are modified in order to perform the following corrective measures: first, a replacement from wooden to concrete sleepers; second, an elastomeric mat introduction between the ballast layer and the ground layer and finally, a rail pads replacement under the rail.

3. Case of study

As previously explained, in order to assess the validity of the whole process, the results obtained by the numerical models should be compared to a real case. For this purpose, the vibrations generated by the passage of a train were registered in a railway turnout located in Sant Cugat del Vallés (Barcelona).

During the data gathering campaign, two accelerometers were located (Figure 2). The first one was set in the crossing nose rail web, while the second one was set in a point located far enough from any track singularity. This point was defined as checkpoint and

was located 11 meters far from the crossing nose and 31 meters far from the switch blades. The registered vehicle was a CAF 112 and circulated through the main direction.

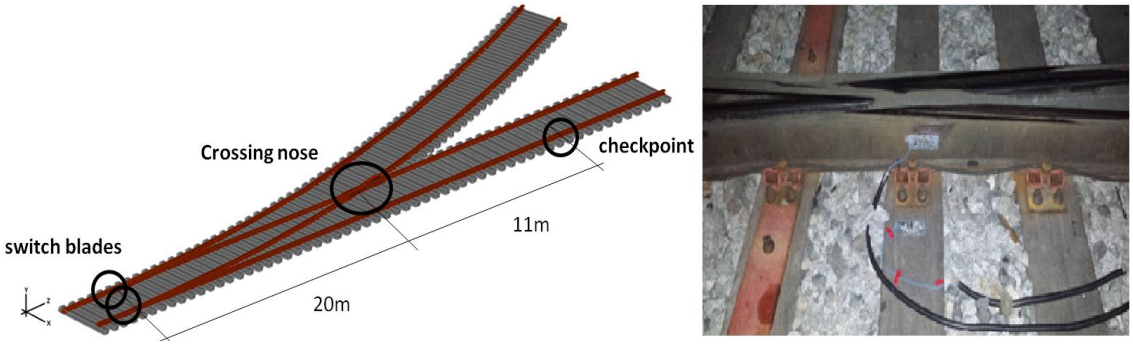


Figure 2: Studied stretch (left) and crossing nose accelerometer location (right)

4. Model's development

The computational considerations of both models are summarized below.

- FEM model:

The track geometry is simplified according to [2]. Thus, hexahedral SOLID45 elements are set in rail pads, sleepers, ballast layer, ground layer and reinforced concrete layer. Nonetheless, in order to simplify the rails geometry and in order to consider the rails inertia variations, linear BEAM4 elements are considered for rails. In this sense, the area moment of inertia of each section is obtained by means of AutoCAD software by defining the 2-Dimensional geometry.

For the simulation, a succession of punctual forces is performed as if the vehicle circulation were. Thus, the dynamic Eq. (1) is solved in every node of the model. In this equation, [M], [C] and [K] are the mass, damping and stiffness matrixes; u, $\dot{u}$ and $\ddot{u}$ are the displacements, velocities and accelerations vectors and {$F(t)$} is the external forces vector.

$$[M]\{\ddot{u}\}+[C]\{\dot{u}\}+[K]\{u\}=\{F(t)\} \quad (1)$$

To calculate the damping matrix, it has been considered the Rayleigh's damping theory by assuming that [C] depends on two coefficients α and β according to Eq. (2).

$$[\mathrm{C}] = \alpha^*[\mathrm{M}] + \beta^*[\mathrm{K}] \quad (2)$$

The diagonalization of this matrix allows obtaining the i-th term in the diagonal from Eq. (3), where [Ø] is the normalized eigenvector of the mass matrix and w_i is the natural frequency of the system.

$$[Ø]^T[\mathrm{C}][Ø] = \alpha[Ø]^{\mathrm{T}}[\mathrm{M}][Ø] + \beta[Ø]^{\mathrm{T}}[\mathrm{K}][Ø] = [\mathrm{C}] = \begin{bmatrix} \alpha+\beta w_1^2 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \alpha+\beta w_n^2 \end{bmatrix} \quad (3)$$

The elements of the resulting matrix, from analogy to single-degree of freedom systems, can be expressed as Eq. (4) where ξ_i is the modal damping coefficient of the i-th term.

$$c_i = 2 * \xi_i * w_i \tag{4}$$

- Multi-body model:

For the multi-body model development, a mathematical simplification of a two-masses vehicle composed by sprung and unsprung masses is adopted. The vehicle suspensions are performed as a series of parallel springs and dampers provided by stiffness and damping coefficients. Furthermore, the equivalent track properties and geometrical track parameters are also included.

For the simulation, a Non-Linear Transient Analysis is adopted and the wheel-rail contact forces are obtained. The time response calculation is carried out with the displacements (q), velocities ($\dot{q}$) and accelerations ($\ddot{q}$) of all degrees of freedom of the vehicle integrated every time step *n* according to Eqs. (5-6).

$$\dot{q}_{(n+1)} = \dot{q}_{(n)} + \ddot{q}_{(n)} \cdot dt \tag{5}$$

$$q_{(n+1)} = q_{(n)} + \dot{q}_{(n+1)} \cdot dt \tag{6}$$

- Multi-body – FE model interaction:

In order to validate the results of the calibration process, the wheel-rail contact forces calculated by the multi-body model are introduced in the FE model of the track together with the quasi-static load performing a new simulation. Once the simulation ends, the resulting accelerograms are compared in the time domain (Figure 3).

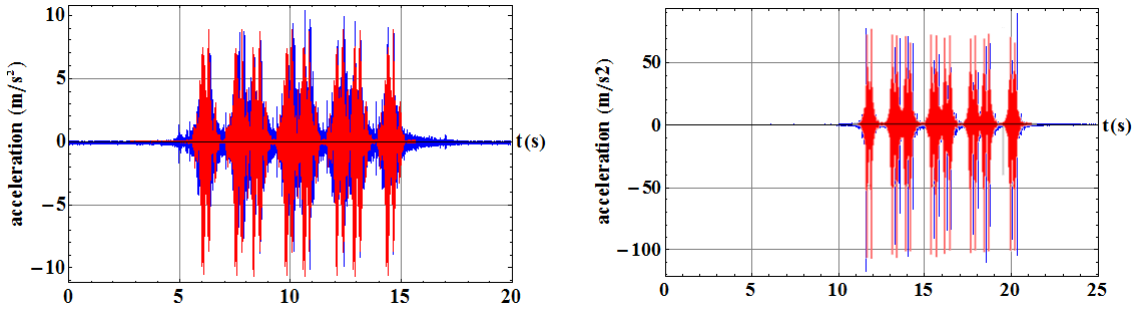


Figure 3: Model's calibration (left) and validation (right). FEM results (red) and data gathering campaign (blue).

The results show an acceptable matching between the data gathering campaign and the FE results. Hence, the model is validated.

5. Simulations and results

Once the Finite Elements model is validated, the different corrective measures can be studied. For this purpose, new VAMPIRE and ANSYS simulations are performed. With

regard to the VAMPIRE results, Table 1 shows the wheel-rail contact forces variations in the singular points of the turnout in the three studied solutions.

		CASE 1 /REALITY	CASE 2 /REALITY	CASE 3 /REALITY
Crossing nose	Left wheel (kN)	-17%	-7%	-15%
	Right wheel (kN)	-10%	-7%	-7%
Switch blades	Left wheel (kN)	-33%	-25%	-25%
	Right wheel (kN)	-35%	-27%	-30%

Table 1: Wheel-rail contact forces variations after the introduction of the corrective measures

It is noticed that the major dynamic overloads reductions are obtained in case 1. Nonetheless, in order to assess the validity of the proposed corrective measures in the vibration generation phenomenon, the aforementioned overloads are introduced in the validated Finite Elements model and the vibratory response is obtained (Figures 4-6).

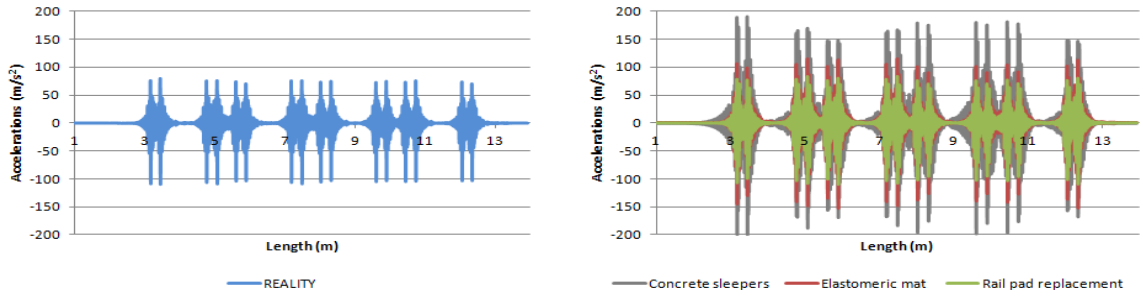


Figure 4: Accelerations calculated in the crossing nose rail web in the real case (left) and after the implementation of the different corrective measures (right).

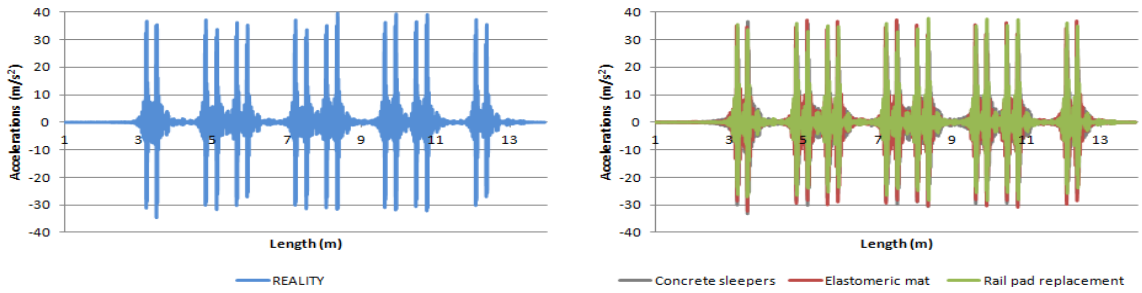


Figure 5: Accelerations calculated in the crossing nose sleeper surface in the real case (left) and after the implementation of the different corrective measures (right).

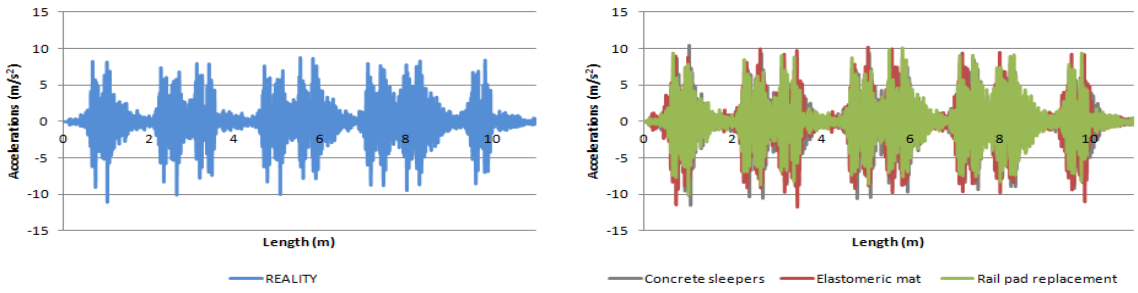


Figure 6: Accelerations calculated in the crossing nose sleeper surface in the real case (left) and after the implementation of the different corrective measures (right).

From the vibratory point of view, with regard to the crossing nose rail web accelerations (Figure 4), it is noticed that contrary to the expectations, all the proposed measures involve a growth in the vibratory response of the crossing nose rail web.

With regard to the accelerations calculated in the crossing nose sleeper surface (Fig. 5), only a slight reduction of the maximum accelerations is obtained in all the studied cases.

Finally, with regard to the accelerations obtained in the switch blades sleeper surface (Figure 6), it can be appreciated that all the proposed measures compel a growth in the accelerations, exception made from the case 3 in which the minimum accelerations are reduced a 10%.

6. Conclusions

The present investigation develops a FE – multi-body interaction in the time domain to study the influence of several corrective measures in the reduction of the vibration generation phenomenon. Thus, according to the above results, the following conclusions can be obtained.

First, regarding the wheel-rail contact forces generation, a noticeable drop of the dynamic loads has been obtained in all the studied measures. In this sense, solution 1 (sleepers replacement) should be the chosen one since it achieves the greatest reductions.

Second, regarding the vibration generation phenomenon, solutions 1 and 2 compel a noticeable growth in the accelerations calculated in the crossing nose rail web. Hence, solution 3 (rail pads replacement) should be the chosen one since it achieves the greatest reductions.

Third, regarding the relation between the wheel-rail contact forces and the vibration generation phenomenon, it must be highlighted that lower overloads does not necessary compel lower vibrations. Thus, solution 1 (sleepers replacement) is the optimum solution for the wheel-rail contact forces reduction, while solution 3 (rail pads replacement) is the optimum solution for the vibrations reduction. Hence, from a railway operator point of view, the optimum solution should be a combination from both solutions 1 and 3.

References

[1] Alfi, S., & Bruni, S. 2009. *Mathematical modelling of train-turnout interaction*. Vehicle System Dynamics, 47(5): 551-574.

[2] Real, J. I., Zamorano, C., Comendador, R., & Real, T. (2014). Computational considerations of 3-D finite element method models of railway vibration prediction in ballasted tracks. *Journal of Vibroengineering*, *16*(4), 1709-1722.

Assessment of train derailment risk in ballasted an slab tracks with failed fasteners

S. Morales[1], J.I. Real[2], L. Montalbán [3], T. Real[4]
[1, 2,3,4] University Institute for Multidisciplinary Mathematics, Polytechnic University of Valencia, 14 Camino de Vera, 46022 Valencia, Spain
[1]Corresponding author. (Tel. +34 963877000; Fax. +34 963877371

E-mail: *[1]silmoiv@upv.es, [2]jureaher@tra.upv.es; [3]laumondo@upv.es, [4]tereaher@upv.es*

1. Introduction

It is well known that curved layouts have a great influence on the dynamic behavior of the vehicle-track system. Thus, correct designs as well as an effective maintenance plan are key aspects to prevent and correct the deterioration of the vehicle and the track.

Vehicle-track dynamics in curves are closely related to derailments [1]. For this reason, in [2] track and vehicle conditions are related with the proclivity to derailment in curves of small radii. This study revealed that the derailment coefficient decreases as the curve radius increases. Another similar research was conducted in [3] using the simulation software SIMPACK. A deeper analysis about the influence of the curve parameters and vehicle speed on derailment was presented in [4], employing the commercial software ADAMS/Rail. Derailment in a high-radius curve (7000 m) including the effect of the fastener failure is analyzed in [5]; this study demonstrated that in this high-radius concrete curve, the stability of a high speed train is not compromised if the number of failed fasteners is lower than 15.

The relationship between flawed fasteners and derailment risk has been also studied in [6] in a tangent track using a coupled vehicle/track model. The results show the strong influence of the disabled fasteners on the derailment coefficient and on the track widening, which also increases the derailment risk. The research [7], conducted in slab tracks, concludes that failed fasteners cause increased displacements on the rail and the slab which become greater with increasing vehicle speed. In [8], the derailment risk in a curved track is assessed considering failures in the track supports, demonstrating their great influence on the running safety; the vehicle and the track infrastructure are modeled as a multi-body system while the rails are modeled as Timoshenko beams resting on discrete supports. The importance of the lateral dynamics on vehicle stability has prompted the development of innovative elements, such as frictional sleepers [9] to increase the lateral resistance of railway tracks, and of new monitoring techniques to maintain the fasteners in perfect condition [10].

In this context, the present study aims to evaluate the dynamic behavior of the vehicle in curved tracks where fasteners have failed. However, two different vehicles have been considered (passengers and freight), with their respective speeds, and two curves of different radii. Furthermore, all these situations have been evaluated and then compared in two different scenarios: a ballasted track and a concrete slab track.

2. Model of the studied track

To assess the risk of derailment, two models have been developed: a track FEM model and a track-vehicle model.

The FEM model is used to calculate the values of track vertical and lateral stiffness in different scenarios in which fasteners fail. For this reason, it has been developed to represent in detail the upper part of the track superstructure, including the rails, the fasteners and the sleepers which rest on elastic supports that simulate the granular layers.

Track geometry has been implemented as a mesh of hexahedral elements with a maximum side of 0.04m. The fastener clip has been represented in detail and the screw effect has been represented by compressive forces acting on each clip. The lateral and plan view of the fastener is imported in a CAD file to construct the model in ANSYS.

Regarding the boundary conditions, the length of the model is 9m, long enough to avoid the influence of the extreme boundaries on the static results, which are always calculated in the center of the model. Rotations and longitudinal movements of the rail extremes are constrained in the model boundaries, while the sleeper displacements are conditioned by spring elements that account for the elastic support.

Once the track model is developed, its stiffness is calibrated with data recorded in a gathering campaign carried out in a real stretch.

Meanwhile, track-vehicle interaction has been analyzed by means of a multi-body model performed in the commercial software VAMPIRE Pro. In this model two vehicles (a passengers and a freight train) have been represented as dynamic systems of three masses (unsprung, semi-sprung and sprung) connected by spring-damper elements that simulate the contact with the rail and the vehicle suspensions.

The procedure followed to calculate the risk of derailment in each case is as follows: First of all, in the track model, fasteners are steadily removed to calculate the variations produced on the vertical stiffness k_v and on the lateral stiffness k_h when the fasteners fail.

Subsequently, vertical and lateral track stiffness values are introduced into the track-vehicle model, the passage of a vehicle is simulated and the wheel-rail contact forces calculated. Once these forces are known, derailment coefficient can be finally obtained as the relationship between lateral and vertical forces in the contact.

3. Simulation of different scenarios

In the present study different scenarios have been presented and their influence on the risk of derailment has been assessed. Among the studied cases, in first term, the influence of the number of failed fasteners has been evaluated; then, the importance of

the uncompensated acceleration of the vehicle while negotiating the curve is assessed; and finally, the behavior of a ballasted and a slab track have been compared.

To study the influence of the number of failed fasteners, they have been progressively removed from the track model and the results for track stiffness have been plotted in Fig.1.

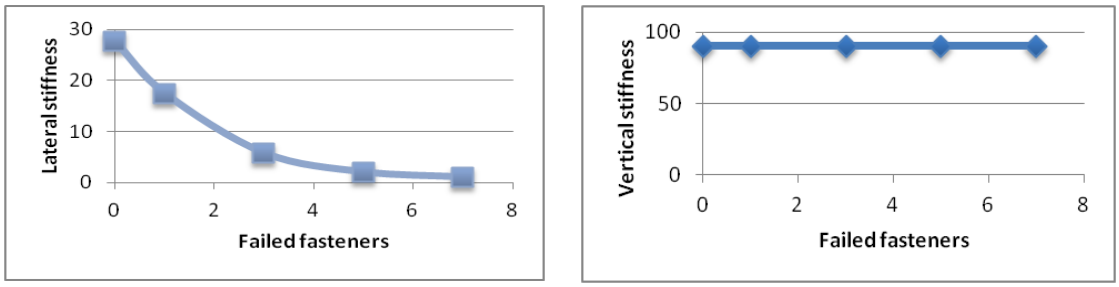


Fig.1. Variations in lateral (left) and vertical (right) stiffness

From Fig. 1, it can be seen that as the number of consecutive failed fasteners increases, the track lateral stiffness decreases and the vertical remains constant. Furthermore, the lateral stiffness approaches asymptotically the 0 value, so it is assumed that for this value track does not oppose any resistance to lateral forces.

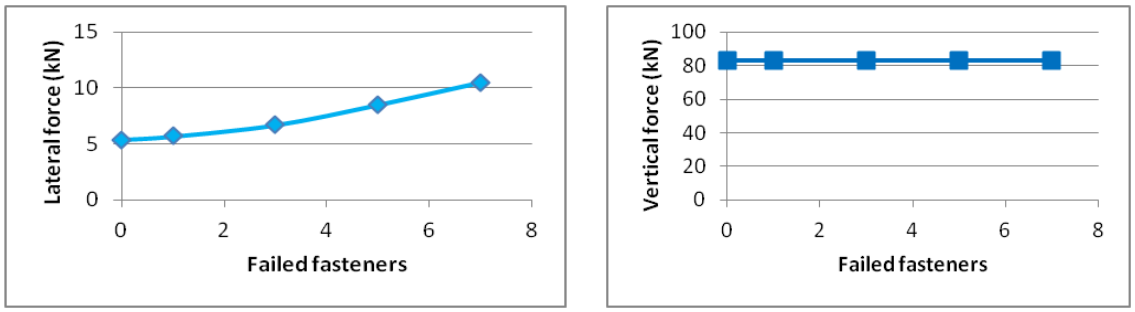


Fig.2. Variations in lateral (left) and vertical (right) forces

Being the stiffness known for each case, its value has been introduced in the vehicle-track model so as to obtain the forces in the wheel rail contact. From Fig.2, it may be concluded that the higher the number of failed fasteners, the higher the lateral forces. Meanwhile, regarding vertical loads, they are not affected by the number of failed fasteners. Relating lateral and vertical forces, derailment coefficient is obtained, concluding that the higher the number of failed fasteners, the lower the vertical stiffness. This situation leads to an increase in lateral forces and derailment coefficient.

In the second scenario, the influence of the parameters of two studied curves, which can be found in the real line, as well as the running speeds of the passenger and freight trains have been studied. Speeds have been specifically defined to induce that, the passenger train negotiates the curves with cant deficit (exerting a greater force on the high rail as a consequence of the centrifuge acceleration) while the freight train circulates in the curve with cant excess. The publication [4] revealed, implicitly, that both the cant excess and deficiency cause high forces in the low and high rails, respectively, modifying the derailment coefficient.

In this case, as expected, no changes in stiffness values are obtained. On the other hand, according to Fig.3, lateral forces are highly affected by failed fasteners, while vertical forces almost remain constant.

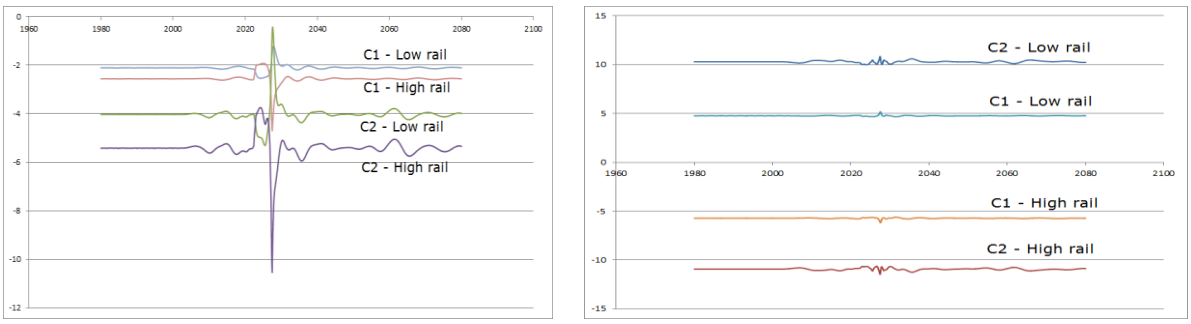


Fig.3. Lateral (left) and vertical (right) forces with a failed fastener for a passengers train running through two curves (C1 and C2) with different radius

Furthermore, the influence of uncompensated acceleration may be clearly seen in Fig 3, where the higher the uncompensated acceleration, the higher the lateral forces. The passengers train presents a higher uncompensated acceleration in curve 2 (C2) since it presents a smaller radius and, again as expected, the higher lateral forces are reached in the high rail since the train negotiates the curve with cant deficit.

Analogously, for the freight trains the higher lateral forces are obtained in the low rail and in the curve with the biggest radius.

Since vertical forces remain constant in the section where the fastener has failed, derailment coefficient will follow the same trend as the lateral forces. Thus, the higher the uncompensated acceleration, the higher the risk of derailment.

Finally, in the third scenario, the influence of the type of track has been assessed. With this aim, firstly, the influence on track lateral stiffness is evaluated. In Fig.4 the variations of stiffness along the track are shown when one and seven consecutive fasteners fail both in a ballasted and in a slab track.

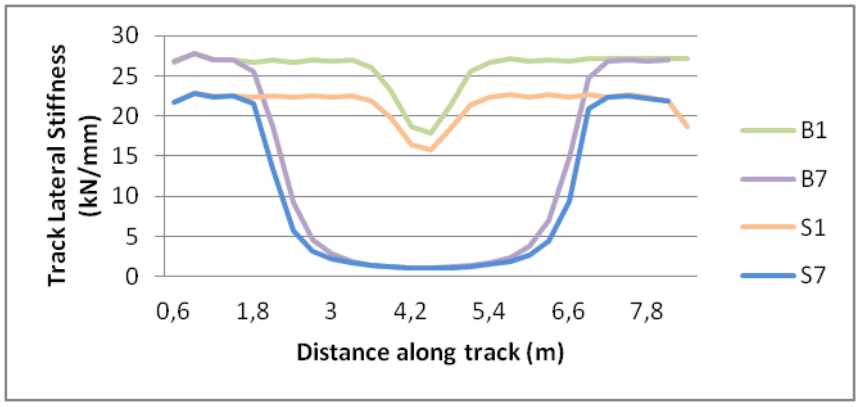


Fig 4. Variations of track lateral stiffness along the track

From this figure, two main considerations may be highlighted. On the one hand, far enough from the center of the model or, in other words, far enough from the failed fasteners, noticeable difference between the stiffness of the slab track and ballasted track are shown. In contrast, in those points where the fasteners have failed, these differences have been diminished. On the other hand, although both tracks have different stiffness values, they present a similar lost of stiffness.

Regarding derailment coefficient, from Fig.5 it can be seen that the higher the number of failed fasteners, the higher the derailment coefficient and the higher the differences between the behavior of the slab and the ballasted track.

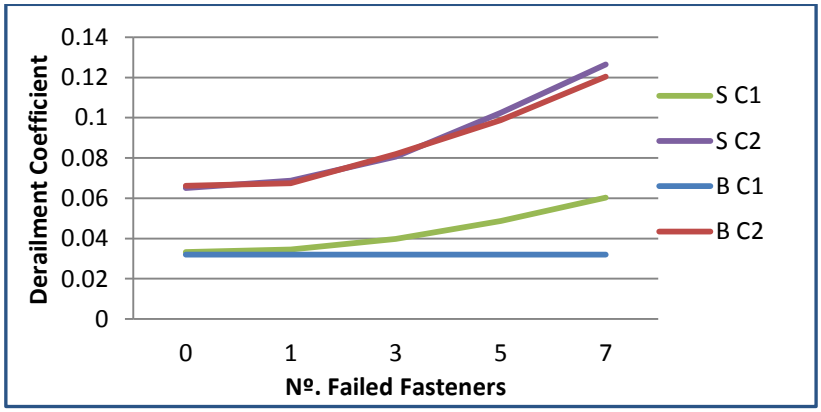


Fig.5. Derailment coefficient in a slab and in a ballasted track with flawed fasteners

Furthermore, it can be seen that for both tracks the same pattern has been obtained when analyzing the influence of uncompensated accelerations. In C1 (the curve with the biggest radius) derailment coefficient is lower than in C2 for a passengers train.

4. Conclusions

In this study, track vertical stiffness has been estimated from experimental results of track settlements using a FEM model. Then the vehicle-track system dynamics in curves and the derailment risk have been studied. From the results obtained, it can be concluded that:

- The higher the number of failed fasteners, the lower the lateral stiffness of the track. This leads to increases in lateral loads in the contact as well as the risk of derailment.
- The influence of failed fasteners on lateral stiffness is significantly higher than on vertical stiffness.
- The lateral forces caused by the fasteners' failure are proportional to the uncompensated acceleration regardless of whether the vehicle negotiates the curve with cant excess or deficiency. Vertical forces are not significantly influenced by the fasteners' failure.
- Slab and ballasted track present different stiffness but a similar lost of stiffness in presence of flawed fasteners.

- Differences in derailment coefficient are higher as the number of failed fasteners grows.

References

1. Brabie D (2007) On derailment-worthiness in railway vehicle design. Analysis of vehicle features influencing derailment processes and consequences, PhD Thesis, Royal Institute of Technology, Stockholm

2. Matsumoto A, Sato Y, Ohno H, Shimizu M, Kurihara J, Saitou T, Michitsuji Y, Matsui R, Tanimoto M, Mizuno M (2014) Actual states of wheel/rail contact forces and friction on sharp curves – Continuous monitoring from in-service trains and numerical simulations. Wear 314:189-197

3. Wang L, Huang A, Liu G (2013) Analysis on curve negotiation ability of the rail vehicle based on SIMPACK. Advanced Materials Research 721:551-555

4. Eom BG, Lee H (2010) Assessment of running safety of railway vehicles using multibody dynamics. International Journal of Precision Engineering and Manufaturing 11(2): 315-320

5. Zhou L, Shen Z (2013) Dynamic analysis of a high-speed train operating on a curved track with failed fasteners. Journal of Zhejiang University- SCIENCE A 14(6): 447-458

6. Xiao X, Jin X, Wen Z (2007) Effect of disabled fastening systems and ballast on vehicle derailment. Journal of Vibration and Acoustics 129: 217-229

7. Shi W, Cai C (2011) Influence of slab track fastener failure on track dynamic performance. Advances in Environmental Vibration-Proceedings of the 5th International Symposium on Environmental Vibration, Chengdu, pp 686-692

8. Xiao X, Jin X, Deng Y, Zhou Z (2008) Effect of curved track support failure on vehicle derailment. Vehicle System Dynamics: International Journal of Vehicle Mechanics and Mobility 46(11):1029-1059

9. Zakeri J, Fakhari M, Mirfattahi B (2012) Lateral resistance of railway track with frictional sleepers. Proceedings of the Institution of Civil Engineers. Transport 165(2):151-155

10. Gibert X, Patel V, Chellappa R, (2015) Robust fastener detection for autonomous visual railway track inspection. IEEE Winter Conference on Applications of Computer Vision, Waikoloa Beach Hawai, pp 694-701

Finite Element analysis of transitions between ballasted tracks and slab tracks in a tram line

T. Real[1], J. Alcañiz[2], C. Zamorano [3], J.I. Real[4]
[1, 4] University Institute for Multidisciplinary Mathematics, Polytechnic University of Valencia, 14 Camino de Vera, 46022 Valencia, Spain
[2] Universidad Católica San Antonio de Murcia. Avenida de los Jerónimos, Nº 135 30107. Murcia. Spain
[3]Foundation for the Research and Engineering in Railways, 160 Serrano, 28002, Madrid, Spain
[1]Corresponding author. (Tel. +34 963877000; Fax. +34 963877371

E-mail: *[1]tereaher@upv.es, [2]jalcaniz@ucam.edu, [3]claraz@fundacioncdh.com, [4]jureaher@tra.upv.es*

1. Introduction

Sudden changes of stiffness in railway tracks can produce dynamic effects affecting track durability as well as passengers' comfort. As a consequence of the recent improvements in railway networks design, many resources have been invested in the study of transitions, putting forward solutions to smooth the stiffness variations in order to minimize their negative effects.

In the current study, two different transitions between three different track typologies (ballasted track, concrete slab track and asphalt slab track) are analyzed by means of Finite Element models calibrated with real data.

1.1 Sections and transitions

Since a 3D finite element model is necessary to study the transitions between the different track typologies, it is important to know the detailed geometry of the real infrastructure from which the model is created. In Fig.1, the cross section of the studied track stretches is shown:

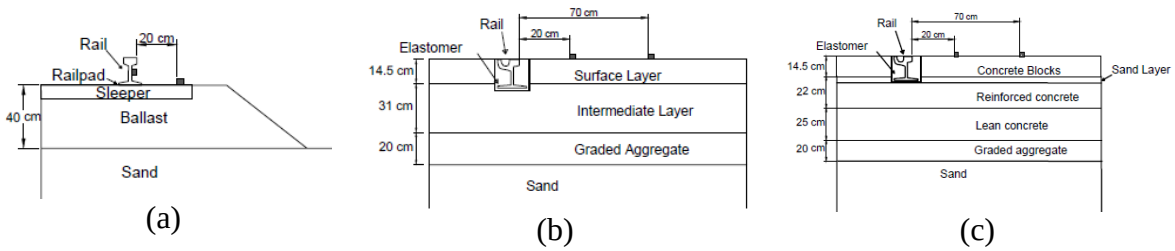
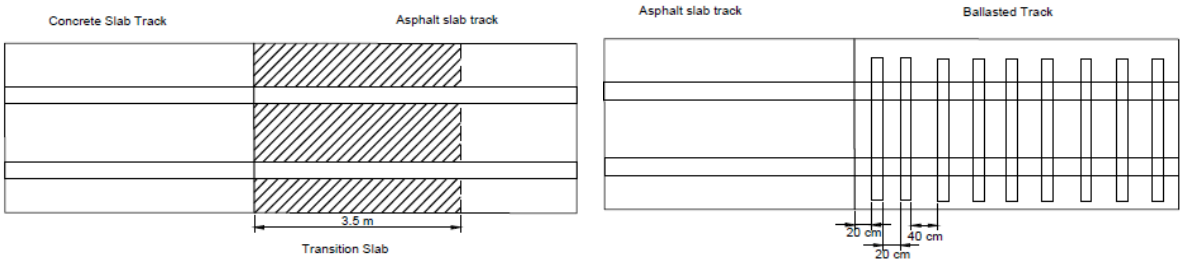


Fig.1 Cross section of the different track typologies

In the ballasted track, the ballast is placed directly on the natural sandy ground. The concrete monoblock sleepers lie on the ballast layer and give support to the UIC-54 rails, with a railpad placed in between as depicted in Fig. 2 (a). This track type is linked to the asphalt slab track represented in Fig. 2 (b) in which a change to a Ph37N grooved rail embedded in elastomer is observed. The surface and intermediate layers are made of a SMA16 asphalt mixture with a 0.5% of recycled plastics and rest on a 20 cm thick layer of graded aggregate placed on the natural soil. The concrete slab track represented in Fig. 2 (c) shares the graded aggregate layer with the adjacent asphalt slab track but a lean concrete bed is placed on this instead to support the reinforced concrete layer. Finally, a thin sand layer is placed under the surface concrete blocks and the reinforced concrete.

It is known that concrete slab track has the highest vertical stiffness, followed by the asphalt slab track and the ballasted track. For this reason, different measures addressed to lessen the abrupt stiffness change were taken during the construction of the infrastructure.

The stiffness difference existing in this case is not as marked as in transition 2. Even so, a lean concrete transition was executed between the graded aggregate and the intermediate layer of the asphalt slab track. This transition slab shown in Fig. 2a, is 15 cm thick and 3.5 m long in the direction of the track.



(a) (b)

Fig.2. Plan view of transitions

With the objective of increasing the stiffness of the ballasted track in the transition vicinity, the distance between the two sleepers closer to the asphalt slab track was reduced as shown in Fig. 2b. These sleepers were placed at a distance of 20 cm, instead of the 40 cm existing in the rest of the line to allow the tamping operations.

Decreasing gradually the separation between the sleepers near the transition to a stiffer track is an effective manner to avoid the sudden stiffness change and its negative effects on the track and the vehicle.

2. Finite elements model of the studied track

The three-dimensional finite elements model of the studied track in Fig. 5 has been constructed following the research carried out by Real et al. (2014). In the model, the geometry of the different sections described in Fig.1 is represented and the transitions

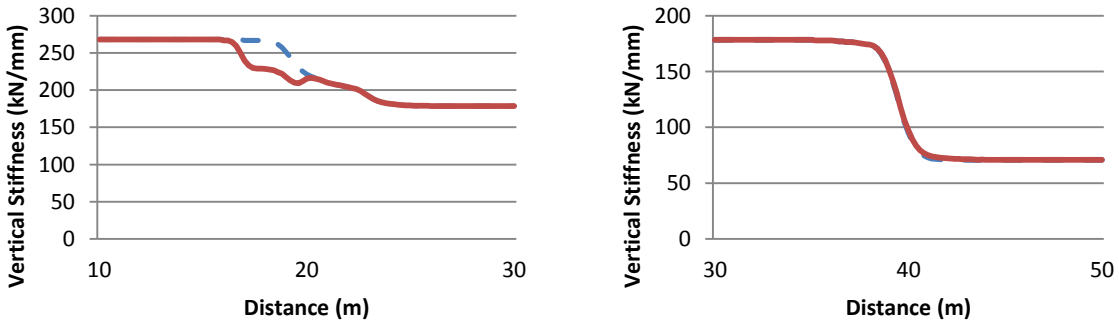
shown in Fig.2 have been implemented as well. The total length of the model is 60m, divided in 3 stretches of the same length correspondent to each track typology. Two different analyses are run on this model: one dynamic to obtain the accelerations induced by the train loads passage and another static to calculate the deflections caused by these loads and, subsequently, the vertical track stiffness.

The considered load corresponds to that transmitted by the considered vehicle, carrying 4 passengers per square meter in each of the three carriages. The final load result (90 kN/axle) has been incremented to include the dynamic overloads according to the German criterion explained in Fastenrath (1981). The final load is 63.95 kN/wheel considering medium track conditions and high statistical precision (95.5%). Vehicle speed has been set in 35 km/h since all the real registers were measured when the trains run at speeds of 33-36 km/h.

3. Evaluation of improvements for the transition design

Since the most important property that the transition must fulfill is to provide a progressive variation of the track vertical stiffness, three different methods existing in the literature to achieve this aim are compared. To do so, from the FEM model, both the static (stiffness) and dynamic (acceleration) responses are calculated on the track in the different scenarios considered.

In the first case, a change of the stiffness of the elastomer is proposed, since authors as Lei and Hang (2011) stated that elastomers have a noticeable influence on the vertical stiffness of the track. Thus, a more elastic elastomer is chosen for the concrete slab track while the stiffness elastomers placed over the sleepers on the ballasted track is progressively increased, being it higher in those sections adjacent to the asphalt slab track. The result of these changes is plotted in Fig. 3 (stiffness) and Fig. 4 (accelerations).

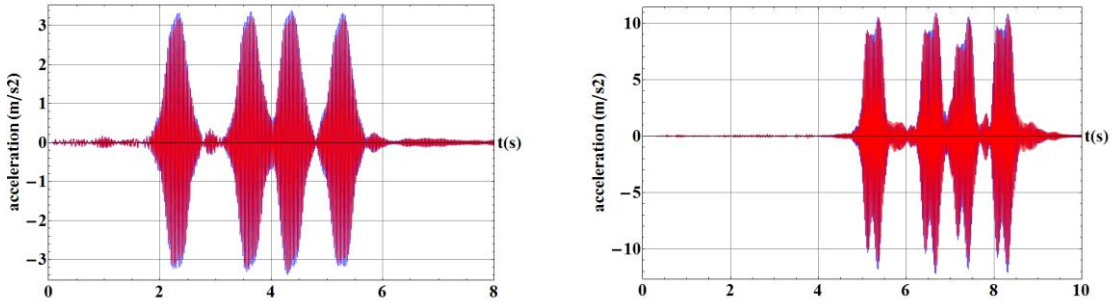


(a) (b)

Fig.3. Evolution of the track vertical stiffness at transition 1 (a) and at transition 2 (b) in the original case (red) and the new scenario (blue)

From the results in Fig.4a, it can be deduced that the transition has been improved between the concrete and the asphalt slab tracks after the substitution of the elastomer in which the rail is embedded by other more elastic. On the other hand, the change of the railpads in transition 2 does not yield the expected results and the stiffness variation is as the original, being the graphs totally overlapped.

The dynamic behavior of the track after the changes is shown in Fig. 4b. According to the previously said, the vibration responses in transition 2 are almost identical while in transition 1, the acceleration slightly increases when a more elastic elastomer is considered.



(a) (b)

Fig.5. Calculated accelerations at 0.2 m from the rail side on the concrete slab track (a) and on the transition sleeper (b) before (red) and after (blue) the change of the stiffness of the elastomers

Other simple strategy to increase the stiffness was studied by Kerr and Moroney (1995) and consists of two additional parallel rails fixed to the sleepers between the rails on which the train moves. This measure results very useful since it can be performed in an operating track without affecting the infrastructure. However, it is only suitable for ballasted tracks.

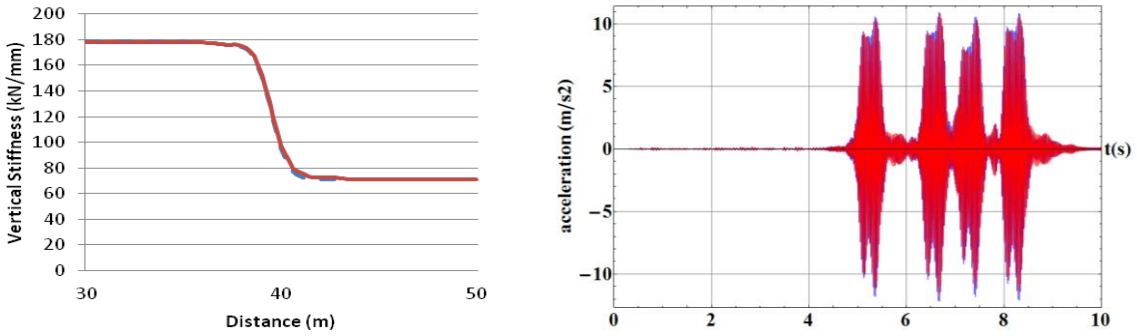


Fig.5. Effect of the additional rails in transition 2 on vertical stiffness (left) and accelerations (right) in the original case (blue) and the new scenario (red)

From Fig. 5, there is no evidence of improvement after the implementation of the additional rails what may suggest that the previous transition, in which the distance between sleepers was reduced, was properly designed. However, track stiffness along the transition is higher

after the extra rails placement (the continuous line is always over the dashed one at the transition). Regarding the dynamic response, it is observed how in a stiffer transition the acceleration peaks on the sleeper decrease slightly for the same loading condition. This acceleration diminution brings with a lower deterioration rate of the ballast layer in the transition vicinity, thus expanding the life span of the infrastructure.

Finally, the use of resilient mats was studied by Xin and Gao (2011) in floating slab tracks on bridges and Costa et al., (2012) in ballasted tracks. In the studied case, two resilient mats of different stiffness have been placed under the reinforced concrete layer in the transition 1 and under the ballast in the transition 2. The objective is to diminish the vertical stiffness of the concrete slab track and increase it in the ballasted track at the transition zones respectively. The results of these changes are seen in Fig. 6a (static) and Fig. 6b (dynamic).

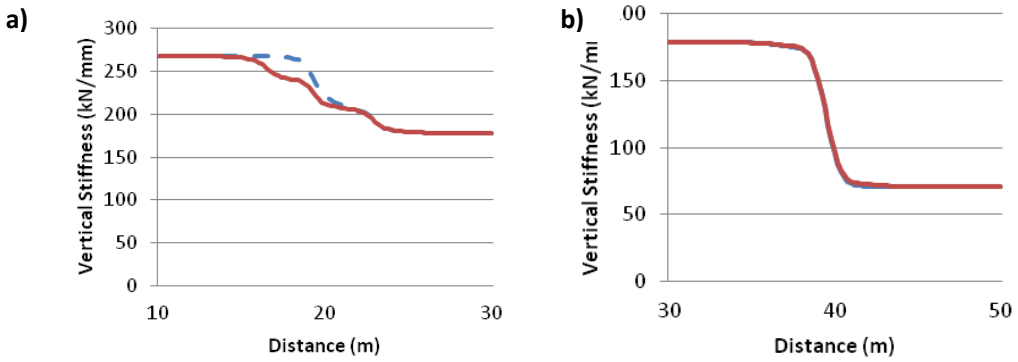


Fig.6. Evolution of the track vertical stiffness at transition 1 (a) and at transition 2 (b) after the placement of the resilient mats. Blue line represents the original results

From this analysis it is seen that transition 2 is not affected by the changes performed. On the other hand, the high-elasticity mat placed in the concrete slab track produces a stiffness reduction in this zone achieving a smooth transition 1. Analogously, the vibration response of the sleepers remains unchanged while on the surface of the concrete slab the acceleration increases as a consequence of the low stiffness of the mat (see Fig.7).

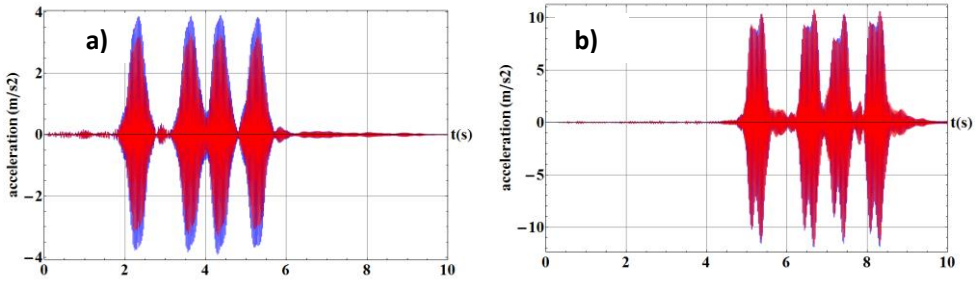


Fig.7. Calculated accelerations at 0.2 m from the rail side on the concrete slab track (a) and on the transition sleeper (b) before (red) and after (blue) the implementation of the resilient mats

CONCLUSIONS

Two different transitions have been analyzed in this study. These transitions were located between three different track typologies: ballasted track; asphalt slab track and concrete slab track. To perform the analysis, a three-dimensional finite elements model has been developed and validated with real acceleration measurements. From this model, different scenarios have been assessed. The results of the study revealed:

- The stiffness slope in the transition zone between a concrete and an asphalt slab track can be improved either reducing the elastomer stiffness in the concrete slab track or using resilient mats under the concrete slab track.
- Additional rails on the ballasted track do not contribute achieve a smoother transition. The sole reduction of the sleeper spacing in the sleepers closer to the transition is recommended instead.
- Neither the elastomer stiffness nor the resilient mats have improved the transition behavior between asphalt and concrete slab tracks. In the cases in which the stiffness slope is too acute, it would be convenient to increase the length of the transition slab located beneath the asphalt slab track.

The dynamic behavior of the infrastructure is slightly modified when an elastic resilient mat is placed under the concrete slab track, being the accelerations increased. Otherwise, no significant changes have been reached.

REFERENCES

Costa, P. A., Calçada, R., & Cardoso, A. (2012). Ballast mats for the reduction of railway traffic vibrations. Numerical study. *Soil Dynamics and Earthquake Engineering , 42*, 137-150.

Fastenrath, F. (1981). *Railroad Track: theory and practice: material properties, cross sections, welding and treatment* (1 ed.). New York: F. Ungar Pub. Co.

Kerr, A., & Moroney, B. (1995). Track transition problems and remedies. *American Railway Engineering Association , 742*, 267-297.

Lei, X., & Zhang, B. (2011). Analysis of dynamic behavior for slab track of high-speed railway based on vehicle and track elements. *Journal of Transportation Engineering , 137*, 227-240.

Real, J., Zamorano, C., Hernández, C., Comendador, R., & Real, T. (2014). Computational considerations of 3-D finite element method models of railway vibration prediction in ballasted tracks. *Journal of Vibroengineering , 16* (4), 1709-1722.

Xin T., G. L. (2011). Reducing slab track vibration into bridge using elastic materials in high speed railway. *Journal of Sound and Vibration , 330*, 2237-2248.

Comparison between analytical and numerical models to predict vibrations in railway tunnels

Fran Ribes[1]*, César Hernández[1], Teresa Real[1], Julia Real[1]

[1]Institute for Multidisciplinary Mathematics, Polytechnic University of Valencia, 46022,

Valencia, Spain

*Corresponding author. E-mail: frarilla@cam.upv.es

1. INTRODUCTION

The increase of underground railway lines in urban areas has prompted the number of people affected by railway externalities, such as noise and vibrations. Furthermore, railway induced vibrations in tunnels may also compromise the stability of the tunnel or the nearby buildings; for this reason, many mathematical models have been developed in the last decades to model railway vibrations in the tunnel vicinity.

These mathematical models can be divided in two main groups: two-dimensional (2-D) and three-dimensional (3-D). As 2-D models only consider wave propagation in the plane which is perpendicular to the tunnel axis and underestimate soil damping, Gupta et al. (2007) assert that this approach is not suitable to predict train induced vibrations. This opinion is also supported by Andersen and Jones (2006), who compared the results of 2-D and 3-D models, stating that 2-D models can be used to assess vibrations qualitatively as a consequence of their low computational cost, but that do not yield precise results.

Simultaneously, vibration prediction models can also be classified in two groups: analytical and numerical models. Analytical models compute rapidly the exact solution solving complex differential equations which, in order to be solved, require many simplifications. The best-extended model available in the current literature is the *Pipe-in-Pipe* model, developed by Forrest and Hunt (2006) and subsequently completed by Hussein and Hunt (2007), Jones and Hunt (2011) and Kuo et al. (2011).

On the contrary, numerical vibration models in tunnels apply the Finite Element Method (FEM), the Boundary Element Method (BEM) or both together to solve the dynamic problem. Although numerical models require higher computational effort than analytical methods, they allow introducing discontinuities and the different elements of the system in detail.

With the objective of comparing two of the most-extended modeling methodologies in the literature, the current study computes the vibrations induced by the passage of a train following two different strategies: the analytical method *Pipe-in-Pipe* and a numerical model constructed exclusively using a FEM software.

2. STUDIED SECCTION

The tunnel used in this study has a circular section with an inner diameter of 8.43m. The tunnel structure is made of reinforced concrete segments, with a thickness of 0.32 cm, and a 0.18 cm concrete grout. There are two parallel tracks of Iberian gauge (1.668m), formed by UIC-60 rails embedded in the concrete slab using the system EDILON Corkelast®.

The soil in which the tunnel is embedded is formed by alternating layers of silty sand and clayey gravel. The water table is located at a depth equivalent to the tunnel center, for this reason two different soils have been considered: saturated and non-saturated.

3. NUMERICAL MODEL

The FEM model is solved using the software ANSYS LS-DYNA. The dynamic problem is expressed by the forced vibration equation where [*M*] is the global mass matrix, [*C*] the damping matrix, [*K*] de stiffness matrix. In the other side of the equation, the time dependent forces vector appears.

$$[M]\{\ddot{\boldsymbol{u}}\} + [C]\{\dot{\boldsymbol{u}}\} + [K]\{\boldsymbol{u}\} = \{\boldsymbol{F^a}(\boldsymbol{t})\} \quad (1)$$

The damping matrix must be determined in order to solve the dynamical problem. the damping matrix is determined as a linear combination of the mass matrix [M] and the stiffness matrix [K].

$$[C] = \alpha[M] + \beta[K] \quad (2)$$

Regarding to the structure, the modeling must accurately represent the details and sizes of the different elements (Fig.1). However, in order to reduce the computational calculation time and meet the elements dimensions criteria imposed by the software, some simplifications have been assumed. These simplifications are carefully chosen in order to avoid any loss of accuracy in the final results.

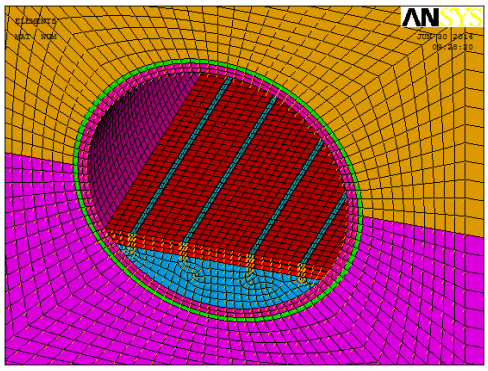


Fig.1. FEM model of the track-tunnel-soil system

The material behavior is assumed to be linear elastic. This hypothesis is assumed because it has been previously checked that the moving train does not induce large strains in the soil in this case.

The vehicle is simulated as a three-mass system where the wheel–rail interaction is modeled as a node-to-beam contact allowing for sliding and loss of contact, using the Penalty algorithm. The contact elements provide an elastic support between the rail and the wheel simulating the hertzian contact.

The full Newton–Raphson method is used to solve the non-linear equations of the problem. The Newmark implicit time integration method is used, to solve the transient dynamic equilibrium equations.

In numerical model, the elasticity modulus of the soil is subjected to a calibration process to represent in the modeled homogeneous soil the same dynamic properties of the real stratified soil. The different accelerograms obtained from the different values chosen are compared with real accelerations registers measured in a point located at the tunnel support.

Numerical model has been calibrated in time domain, but in order to compare it with the numerical models results the frequency domain results have to be also in accordance with measured results.

4. ANALYTICAL MODEL

The main characteristic of the *Pipe-in-Pipe* model (Forrest and Hunt (2006)) is that considers the soil as a hollow cylinder of infinite length into which the tunnel is embedded. Into the tunnel, there is a concrete slab which supports the rails. The concrete slab and the rails are modeled as Euler-Bernouilli beams subjected to bending forces. Between the rails and the slab there are railpads which provide an elastic support for the rail while between the slab and the tunnel invert there is an elastic mat. The displacements among the different elements in the model are coupled in the frequency and wavenumber domain, in which the dynamic problem is also solved.

The wave propagation phenomenon through the soil is studied via the wave equation, considering a solid, homogeneous, isotropic and linear-elastic medium for the tunnel structure and the soil. The wave equation is displayed in Eq.(3), in which u is the displacement vector; λ and μ are Lamé constants of the material; ρ is the density of the material; f the volumetric forces and t is the independent variable time. Although it is a 3-D model, the dynamic behavior in the tunnel direction is assumed to remain invariable in order to simplify the calculations.

$$(\lambda + \mu)\nabla\nabla \cdot u + \mu\nabla^2 u + \rho f = \rho \frac{\partial^2 u}{\partial t^2} \tag{3}$$

Eq. (3) is also solved in the frequency and wavenumber domain, obtaining the field of displacements u produced by a unit load acting in the tunnel invert in the vertical and horizontal direction (See Hussein and Hunt (2007)). The coupling between the concrete structure of the tunnel and the natural soil is solved taking into account that the values of the field of displacements in the tunnel outer surface must be equal to the

displacements in the soil adjacent to the tunnel and that there must exist equilibrium of external forces in that tunnel-soil interface.

The unit load is introduced as a harmonic excitation along the tunnel axis, taking into account only the unsprung mass of the vehicle. Moreover, this load is decomposed using Fourier series along the circular cross section as a summation of sinusoidal terms in the three space directions (Gupta et al. (2007)). The resulting expression of the unit load, acting at time t in the position x, in the j-th special direction in wavenumber-frequency domain is represented in Eq.4, in which $\tilde{Q}_{jjn}$ is the value of the unit load in frequency-wavenumber domain; n accounts for the term of the Fourier series decomposition; θ is the angle; ϕ depends on the load direction and ω and ξ are, respectively, the frequency and the wavenumber.

$$q_j(x,t) = \tilde{Q}_{jn} sin\,(n\theta + \phi) e^{j(\omega t + \xi x)} \tag{4}$$

After solving the wave equation (3) for the unit load condition, the resulting displacements for each space direction have an analogous expression, where the term $\tilde{U}_{jjn}$ represents the displacement calculated in the frequency-wavenumber domain.

$$u_j(x,t) = \tilde{U}_{jn} sin\,(n\theta + \phi) e^{j(\omega t + \xi x)} \tag{5}$$

Hussein and Hunt (2009) implemented the complex model *Pipe-in-Pipe* in a software programmed in *Matlab*. Introducing the vehicle, track, tunnel and soil inputs the system's response is directly computed in a specified frequency range.

The aim of this paper is to compare the calculated results from both models in order to check the strengths of each methodology. For that reason, in analytical model, the same characteristics as in the numerical model have been considered. However, there are some differences between the input parameters each model requires. Despite the majority could to be converted as an equivalent, the damping coefficients used in numerical models are not available in the "Pipe-in-Pipe" model. Hence loss factor has been calibrated in order to match the calculated results with the measured registers in frequency domain.

5. RESULTS

In this section, the results obtained from both models are compared. To do this, the results from three different points will be obtained. These points are located in the surrounding soil. Must be noticed that Analytical model do not provide results in the track elements.

Observing the results (Fig.2) is noticed that the frequency band from 20 to 80 Hz are relatively close each other. However in the low frequency range there are noticeable differences.

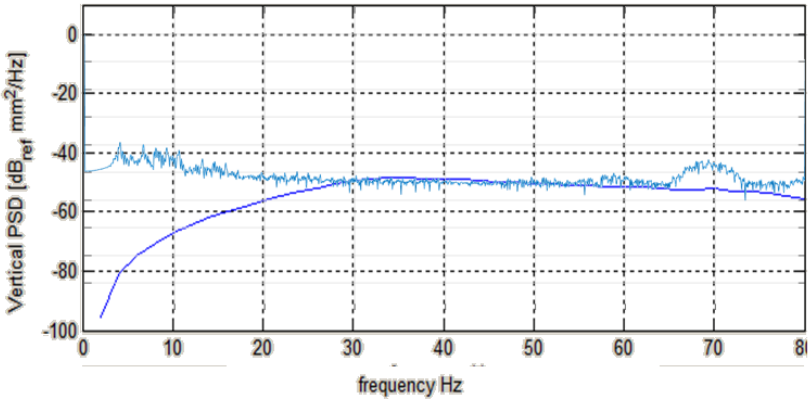


Fig.2. Frequency domain calculated displacements at tunnel invert

The results in the other points are compared and the same trend is observed. The frequency brand from 2 to 25 presents noticeable differences.

These differences could be due to some reasons. Should to be considered that despite the analytical model allows to reproduce the pass of the train though the tunnel, it do not reproduce the wave motion in the longitudinal direction. These facts underestimate the wave propagation phenomena, which is a pure 3D mechanism. The vehicle simulation is other of the differences within models that could produce the differences between results. Meanwhile in the numerical model the train is simulated as multibody approach coupled with the rail, the analytical model is simulated as an infinite train loads. Despite these differences, the results of both models match perfectly for both models from 30 hz.

6. CONCLUSIONS

In this paper numerical 3D finite elements model to predict railway vibrations in tunnel have been compared with an analytical model "PiP". For this purpose, 3D finite elements model has been developed, calibrated with real field measurements. Analytical model Pipe-in-Pipe (Forrest and hunt (2006)) is calibrated in frequency domain with the same field measurements. Comparing the results from the different models developed, the next conclusions have been obtained:

- Results do not match as much as expected, but certainly there is certain similitude in results in medium range frequencies (25-80 Hz).
- Analytical model, despite is considered as a 3D model, do not allow waves to propagate in the longitudinal direction. This fact introduce some inaccuracy in the results since some phenomena such a wave propagation has a clearly spatial behavior.
- Analytical model only provide support to reproduce circular tunnels with just one track due to the complexity of the equations to be solved. In this way, numerical FEM models allow to reproduce any section without restrictions, allowing to use as a prediction tools in geometrically complex tunnels.
- Numerical model allow reproducing the vehicle as a multibody system considering all the masses and wheel rail interaction. This provides a more

accurate reproduction of the train circulating that the infinite train of loads moving along the tunnel considered in the analytical model

- On the other hand, computing time is one of the most important factors when defining the modeling strategy. In this regard, the simulation of the unbounded domain as absorbing boundaries and the great amount of elements needed stand as the major disadvantage of the 3D FEM models. In the studied case, computing time is around of 20 hours for the FEM model while this time can be reduced up to 2 minutes in the case of Pipe-in-Pipe model. Therefore, even though 3D fem model provides a better reproduction of the dynamic response, the lower calculation time make analytical models interesting tools when there exist time or computing limitations.

References

Andersen, L, y C.J.C Jones. «Coupled boundary and finite element analysis of vibration from railway tunnels - a comparison of two- and three-dimensional models.» *Journal of Sound and Vibration, 293 (3-5)*, 2006: 611-625.

Forrest, J.A, y H.E.M Hunt. «A three-dimensional tunnel model for calculation of train-induced ground vibration.» *Journal of Sound and Vibration* 294 (2006): 678-705.

Gupta, S, M.F.M Hussein, G Degrande, H.E.M Hunt, y D. Clouteau. «A comparison of two numerical models for the prediction of vibrations from underground railway traffic.» *Soil Dynamics and Earthquake Engineering* 27 (2007): 608-624.

Hussein, M.F.M, y H.E.M Hunt. «A computationally efficient software application for calculating vibration from underground railways.» *Jounal of Physics: Conference Series* 181 (2009): 1-6.

Hussein, M.F.M, y H.E.M Hunt. «A numerical model for calculating vibration from a railway tunnel embedded in a full-space.» *Journal of Sound and Vibration* 305 (2007): 401-431.

Kuo, K. A, H.E.M Hunt, y M.F.M Hussein. «The effect of a twin tunnel on the propagation of ground-borne vibration from an underground railway.» *Journal of Sound and Vibration* 330 (2011): 6203-6222.

Influence of rail corrugation on the vibrations induced by traffic loads

Julia Real[1]*, José Luis Velarte[1], Clara Zamorano[2], Teresa Real[1]

[1]Institute for Multidisciplinary Mathematics, Polytechnic University of Valencia, 46022, Valencia, Spain

[2]Foundation for the Research and Engineering in Railways, 160 Serrano, 28002, Madrid, Spain

September 9, 2015

1. Introduction

The imperfections in the wheel-rail contact are one of the main sources of generation of vibrations and dynamic forces. In this sense, rail corrugation has become one of the major problems in the field of railway engineering, since its presence affects to the vehicle dynamics and represents one of the main transport externalities.

This pathology is characterized by the presence of quasi-sinusoidal waves on the rail surface performing crests and valleys [1], and is defined by its amplitude and wavelength. Thus, depending on the values of these parameters, the dynamic response of both vehicle and track varies. For this reason, the present investigation deals with the consequences derived from the presence of rail corrugation phenomenon, analyzing the influence that the wavelength, amplitude and vehicle speed have on the wheel-rail contact forces and the vibration generation phenomena.

2. Methodology

In order to study in depth this pathology, this investigation proposes a numerical feedback process between a Finite Elements (FE) model of the track developed by means of ANSYS software and a multi-body model of the vehicle developed by means of VAMPIRE software. The whole process is divided into three steps.

In a first step, the FE model (Figure 1) is performed according to [2] and the unknown mechanical properties of the track are obtained thanks to a calibration process

performed with the accelerations registered during the passage of a tram vehicle in a straight stretch with no rail corrugation (Figure 2).

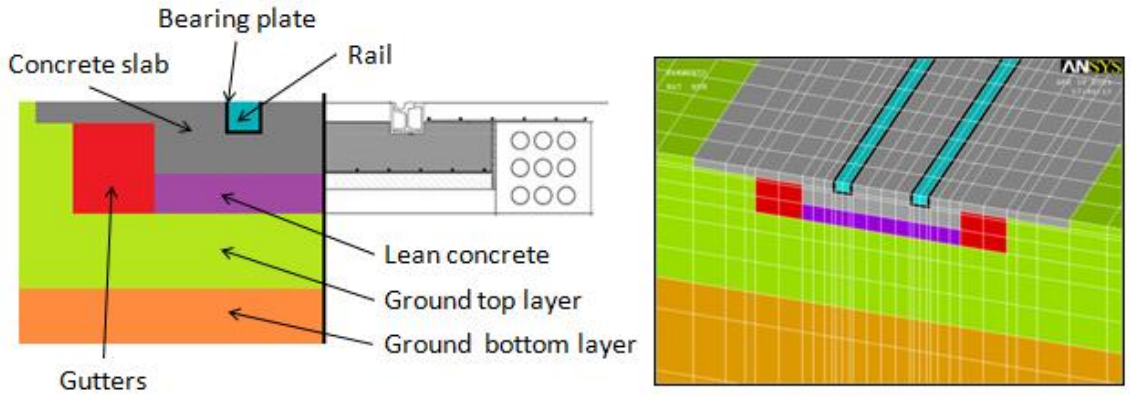


Figure 1: Geometrical simplifications of the FE model

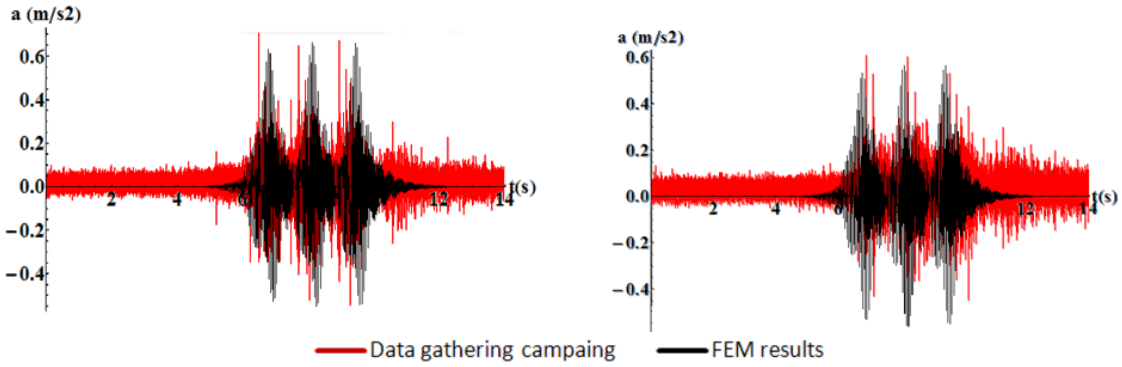


Figure 2: Comparison between the FE results and the data gathering campaign. Accelerograms obtained at 0,3 meters far from the rail web for calibration (left) and at 1,2 meters far from the rail web for validation (right).

In a second step, the multi-body model (Figure 3 right) is developed as a two masses vehicle (Figure 3 left) according to [3]. A total of 60 simulations are performed in order to study the influence of rail corrugation parameters when the vehicle circulates through a damaged track to different speeds. Thus, the variations of the wheel-rail contact forces resulting from the passage of the vehicle through different rail corrugations are obtained.

Figure 3: Comparison between the mathematical simplification of the vehicle (left) and VAMPIRE simulation (right).

In a third step, the aforementioned wheel-rail contact forces are introduced as punctual forces in the calibrated FE model and the vibratory response of the track is analyzed in each case.

The result of the whole process allows obtaining several conclusions about the influence of rail corrugation parameters on both dynamic forces generation and the vibrations generation phenomena.

3. Simulations and results

In order to study the influence of rail corrugation on vehicle-track dynamics, a sensitivity analysis of rail corrugation parameters has been performed according to the typical values collected in [4]. Thus, a total of 6 combinations have been studied (Table 1) for a speed range from 10 to 100 km/h in steps of 10 km/h (60 simulations).

	λ=0.25m	λ=0.5m	λ=1m
A=0.15mm	YES	YES	YES
A=0.3mm	NO	YES	YES
A=0.5mm	NO	NO	YES

Table 1: Combinations of the rail corrugation parameters

The influence of rail corrugation parameters on the vehicle-track dynamics has been studied from two different points of view. The first one is focused on the wheel-rail contact forces generation phenomenon, while the second one is focused on the vibration generation phenomenon.

With regard to the first one, all the simulations have been performed in VAMPIRE software. As an example of the results, the following Figure shows a comparison between the wheel-rail contact forces obtained when a vehicle circulates at two different speeds (30km/h and 70 km/h) through a track with the presence of a rail corrugation of A=0.3 mm amplitude and λ=0.5 meters wavelength.

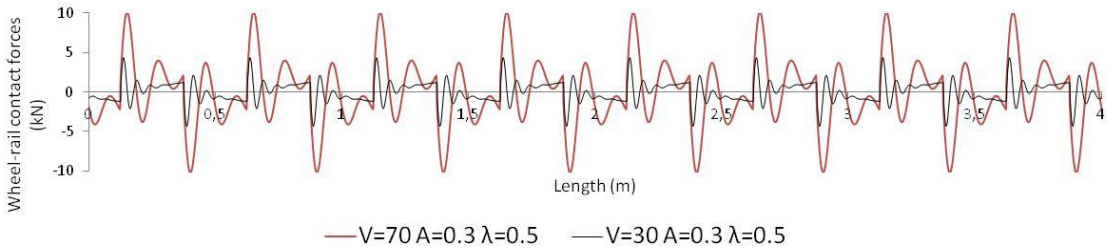


Figure 4: Wheel-rail contact forces generated in a track with the presence of a rail corrugation of 0.3 mm amplitude and 0.5 meters wavelength when a vehicle circulates at two different speeds: 30km/h and 70 km/h

It can be appreciated that, in both cases, a heterogeneous pattern takes place. Nevertheless, the distance between the maximum and minimum peaks matches with the rail corrugation wavelength (0.5 meters). The main reason of this heterogeneous pattern is that the wheel-rail contact patch varies as a consequence of the combination between the rail corrugation geometry and vehicle speed.

For the sake of simplicity, the following Figure summarizes the influence of rail corrugation parameters on the wheel-rail contact forces generation phenomenon.

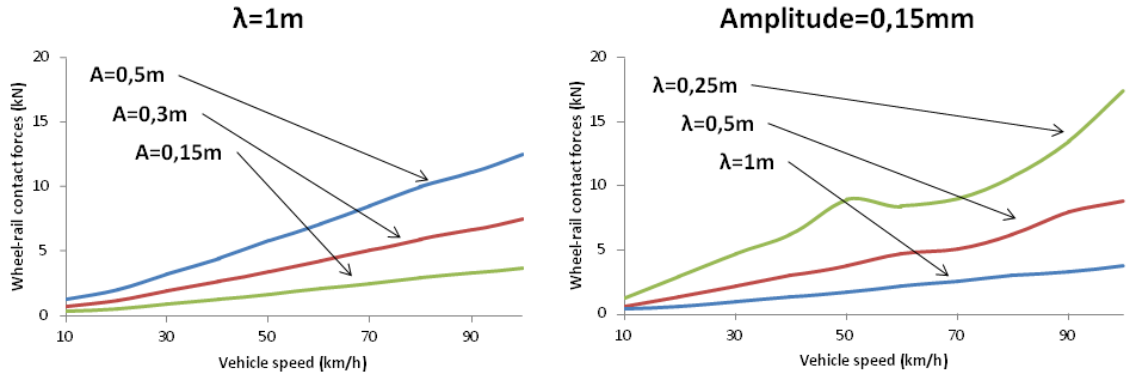


Figure 5: Influence of rail corrugation amplitude (left) and wavelength (right) on the wheel-rail contact forces generation

According to Figure 5 left, on the one hand, it can be stated that the deeper the amplitude (A) is, the greater the dynamic overloads are. Meanwhile, according to Figure 5 right, it can be stated that the shorter the wavelength (λ) is, the greater the dynamic overloads are. On the other hand, according to both Figure 5 left and Figure 5 right, the greater the circulation speed is, the grater the dynamic overloads are.

With regard to the vibration generation phenomenon, the combinations shown in Table 1 have been performed in ANSYS software at two circulation speeds: 30 km/h and 70 km/h. Thus, a total of 12 cases have been studied. To do so, no rail corrugation has been performed in the Finite Elements model. By contrast, the maximum and minimum wheel-rail contact forces obtained by means of VAMPIRE simulations are set as punctual forces together with the quasi-static load through the track.

At the end of each simulation, the vibrations generated at three different locations have been analyzed: rail web (in order to study the effects on rails and bearing plates), at 0.4 meters far from the rail web (in order to study the effects on the slab track) and at 1.8 meters far from the rail web (in order to study the effects on track surroundings).

Next, Figures 6-8 show the results of the numerical simulations when a rail corrugation of 0.15mm amplitude and a variable wavelength (λ) is performed.

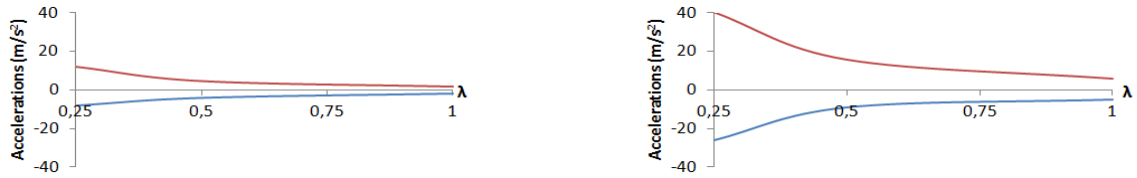


Figure 6: Maximum accelerations calculated on the rail web. Circulation speeds of 30 km/h (left) and 70 km/h (right).

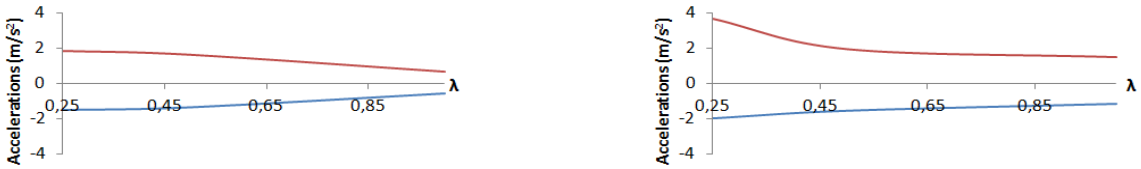


Figure 7: Maximum accelerations calculated at 0.4 meters far from the rail web. Circulation speeds of 30 km/h (left) and 70 km/h (right).

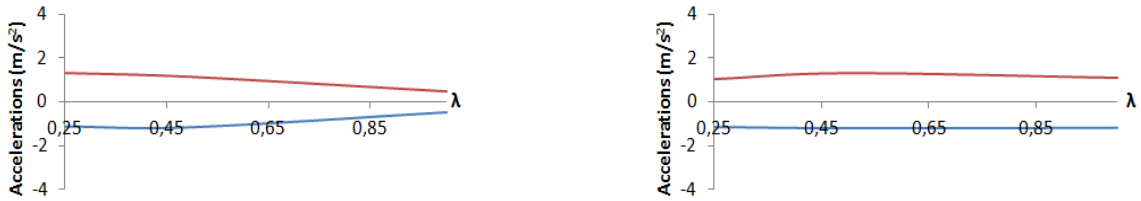


Figure 8: Maximum accelerations calculated at 1.8 meters far from the rail web. Circulation speeds of 30 km/h (left) and 70 km/h (right).

According to Figures 6-8, on the one hand, the shorter the wavelength is, the greater the accelerations are. On the other hand, the greater the circulation speed is, the greater the accelerations are. Nevertheless, the effect of circulation speed decreases with the distance to the track. Thus, the vibrations generated at the rail web (Figure 6) are 16 m/s^2 when the vehicle circulates at 30 km/h and 40 m/s^2 when the vehicle circulates at 30 km/h (230% higher). Nevertheless, this difference is reduced to 103% in Figure 7 and 14% in Figure 8.

Next, Figures 9-11 show the results of the numerical simulations when a rail corrugation of λ=1m wavelength and a variable amplitude (A) is performed.

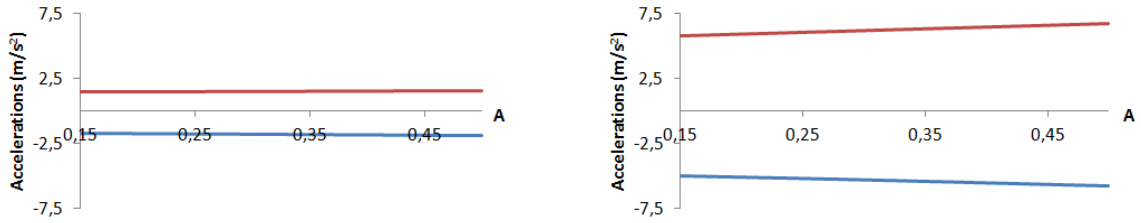


Figure 9: Maximum accelerations calculated on the rail web. Circulation speeds of 30 km/h (left) and 70 km/h (right).

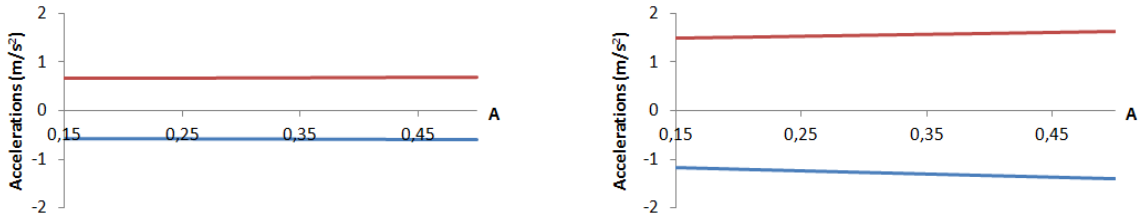


Figure 10: Maximum accelerations calculated at 0.4 meters far from the rail web. Circulation speeds of 30 km/h (left) and 70 km/h (right).

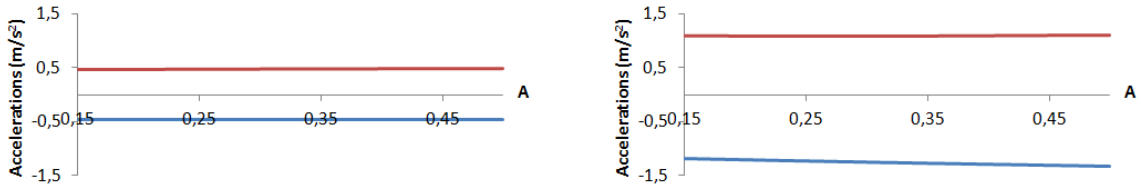


Figure 11: Maximum accelerations calculated at 1.8 meters far from the rail web. Circulation speeds of 30 km/h (left) and 70 km/h (right).

According to Figures 9-11, on the one hand, the deeper the amplitude is, the greater the dynamic overloads are. On the other hand, as occurred in Figures 6-8, the greater the circulation speed is, the greater the accelerations are. Meanwhile, the effect of circulation speed decreases with the distance to the rail web.

4. Conclusions

In this work, a method consisting in a numerical feedback of a FE model of a track and a multi-body model of a vehicle has been presented in order to study the influence of rail corrugation parameters and vehicle circulation speed on the dynamic response of a tram track. Thus, according to the numerical results, it can be stated that the shorter and the deeper the rail corrugation is, the greater the wheel-rail contact forces and vibrations are.

In this sense, the vehicle speed is also a key factor. Thus, the greater the circulation speed is, the greater the dynamic overloads and accelerations are. Nevertheless, its influence decreases with the distance to the rails.

REFERENCES

[1] Grassie, S. L. (2009). Rail corrugation: characteristics, causes, and treatments. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, *223*(6), 581-596.

[2] Real, J. I., Zamorano, C., Comendador, R., & Real, T. (2014). Computational considerations of 3-D finite element method models of railway vibration prediction in ballasted tracks. *Journal of Vibroengineering*, *16*(4), 1709-1722.

[3] Real, J. I., Zamorano, C., Velarte, J. L., & Blanco, A. E. (2015). Development of a vehicle–track interaction model to predict the vibratory benefits of rail grinding in the time domain. *Journal of Modern Transportation*, *23*(3), 189-201.

[4] Spanish regulations N.R.V. 7-5-2.1. (1993). Conservación de la vía. Amolado de las superficies activas del carril, en vía. RENFE, Dirección de Mantenimiento de Infraestructura.

A Schwarz preconditioner for the neutron diffusion equation

A. Vidal-Ferràndiz[♭] , S. González-Pintor[†], D. Ginestar[‡] [*],
G. Verdú[♭] and C. Demazière[†]

(♭) Instituto de Seguridad Industrial: Radiofísica y Medioambiental,

Universitat Politècnica de València, Camino de Vera s/n, 46022, València, Spain,

(†) Department of Applied Physics

Chalmers University of Technology, Maskingränd 2, 412 58 Göteborg, Sweden

(‡)Instituto de Matemática Multidisciplinar.

Universitat Politècnica de València, Camino de Vera s/n, 46022, València, Spain.

November 30, 2015

1 Introduction

Simulation codes of a nuclear power reactor traditionally use a multigroup approximation of the neutron diffusion equation to describe the neutron distribution inside a reactor core. The steady state of of the reactor in the approximation of two groups of energy is described by a differential generalized eigenvalue problem of the form,

$$\begin{pmatrix} -\vec{\nabla}(D_1\vec{\nabla}) + \Sigma_{a1} + \Sigma_{12} & 0 \\ -\Sigma_{12} & -\vec{\nabla}(D_2\vec{\nabla}) + \Sigma_{a2} \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} = \frac{1}{k_{\text{eff}}} \begin{pmatrix} \nu\Sigma_{f1} & \nu\Sigma_{f2} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}, \tag{1}$$

where D_g, $g = 1, 2$ are the diffusion coefficients, Σ_{ag}, Σ_{fg} and Σ_{12} are the macroscopic cross sections of absorption, fission and scattering, respectively. ϕ_1 and ϕ_2 are the fast and thermal neutron fluxes, respectively.

[*]e-mail: dginesta@mat.upv.es

Once a spatial discretization has been selected, the problem is reduced to its corresponding algebraic set of equations. To solve the algebraic eigenvalue problem it is necessary to solve linear systems associated with t block matrices, corresponding tot the energy groups. These systems are solved using a preconditioned Krylov method. For the spatial discretization of equation (1) a high order Finite Element Method (FEM) is used [1], which makes use of Lagrange polynomials whose nodes form a Gauss-Lobbato quadrature set. These polynomials provide a partition of the shape functions into vertices, edges, faces and interior functions. Using this natural partition the linear systems of equations associated with each energy group can be solved using a Schur Complement method that algebraically decouples the interior degrees of freedom from the other ones. This method is advantageous when a high polynomial degree, p, is used in the FEM discretization. To precondition the resulting Schur complement system two different strategies are described. First a substructuring block Jacobi preconditioner is studied where the coupling between the different elements is neglected. Also, a domain decomposition algorithm with overlapping between subdomains, like the additive Schwarz method, is considered. The performance of the different approaches is studied numerically using a three-dimensional model of a reactor core.

2 Schur complement method

Consider the solution of a generic linear system,

$$A\,x = f \ , \tag{2}$$

which can be structured as,

$$\begin{pmatrix} A_{bb} & A_{bi} \\ A_{bi}^T & A_{ii} \end{pmatrix} \begin{pmatrix} x_b \\ x_i \end{pmatrix} = \begin{pmatrix} f_b \\ f_i \end{pmatrix} \ , \tag{3}$$

where the degrees of freedom related to the interior of the finite elements are decoupled algebraically using a technique known as the Schur complement method. This manipulation decomposes the system into the contributions related to the subdomain boundaries, b, union of the vertices, edges and faces shape functions; and the interior shape functions, i. The resulting system is of the same size as of the boundary degrees of freedom and is written as,

$$S x_b = f_S \ , \tag{4}$$

$$S = A_{bb} - A_{bi}A_{ii}^{-1}A_{bi}^{T} \ , \quad f_S = f_b - A_{bi}A_{ii}^{-1}f_i \ , \quad \text{and} \quad x_i = A_{ii}^{-1}(f_i - A_{bi}^{T}x_b) \ .$$

2.1 Subtructuring preconditioner

In the same way as the original matrix, the Schur complement matrix has an structure that can be algebraically separated into vertices, edges and faces degrees of freedom, as follows

$$\begin{pmatrix} S_{vv} & S_{ve} & S_{vf} \\ S_{ve}^T & S_{ee} & S_{ef} \\ S_{vf}^T & S_{ef}^T & S_{ff} \end{pmatrix} \begin{pmatrix} x_{b,v} \\ x_{b,e} \\ x_{b,f} \end{pmatrix} = \begin{pmatrix} f_{S,v} \\ f_{S,e} \\ f_{S,f} \end{pmatrix} \ . \tag{5}$$

A substructuring preconditioner for this system can be defined as,

$$P_{vef} = \begin{pmatrix} S_{vv} & 0 & 0 \\ 0 & \text{block-diag}(S_{ee}) & 0 \\ 0 & 0 & \text{block-diag}(S_{ff}) \end{pmatrix}^{-1} \ . \tag{6}$$

The submatrices S_{ee} and S_{ff} represent the whole set of edges and faces degrees of freedom, respectively. However, in the proposed preconditioner, each edge and each face is considered independently neglecting the interactions between different edges or faces forcing a block diagonal structure for the submatrices.

For comparative reasons a simple diagonal Jacobi preconditioner, P_d, and a classical incomplete LU decomposition, P_{ILU}, are also implemented.

2.2 Restricted Additive Schwarz

Another possibility of preconditioning consists of introducing and overlapping between the subdomains by including the degrees of related to the vertices and edges in the different blocks. The last preconditioner is referred as a Restricted Additive Schwarz (RAS),

$$P_{RAS} = \sum_{k}^{K} R_k^T X_k \left(R_k S R_k\right)^{-1} R_k, \tag{7}$$

where the subdomain k is the complete union of two finite element cells, R_k denote the restriction operator from the global domain to the subdomain k, R_k^T in the corresponding extension operator and X_k is a partition of unity matrix that scales the contribution of each degrees of freedom depending the number subdomain of in which it is present.

3 Numerical Results

Table 1 shows the performance of the different preconditioners in terms of the average number of iterations for the required linear systems resolutions, memory used by the matrix elements and CPU time with 5th degree polynomials in the FEM. A decrease of the number of iterations can be seen as the preconditioners become more complete. However, this improvement in the number of iterations does not always represent a time improvement because an extra overhead to compute the preconditioner exists. The results show that the fastest preconditioner for Schur complement system is the RAS.

Table 1: Comparative of preconditioners for IAEA reactor [2] with $p = 5$.

Preconditioner	its $L_{g=1}$	its $L_{g=2}$	Memory (MB)	Time (s)
None	48.7	47.6	2500	1500
P_d	35.1	25.1	2500	1200
P_{vef}	29.0	22.1	2500	890
P_{RAS}	23.9	17.9	2800	800
P_{ILU}	11.0	8.9	4600	1280

Acknowledgements

The work has been partially supported by the spanish Ministerio de Economía y Competitividad under projects ENE 2014-59442-P and MTM2014-58159-P, the Generalitat Valenciana under the project PROMETEO II/2014/008 and the Universitat Politècnica de València under the project FPI-2013.

References

[1] A. Vidal-Ferrandiz, R. Fayez, D. Ginestar, and G. Verdú. Solution of the lambda modes problem of a nuclear power reactor using an h-p finite element method. *Annals of Nuclear Energy*, 72:338–349, 2014.

[2] Computational Benchmark Problems Committee. *Benchmark Problem Book*. American Nuclear Society, Argonne National Laboratory, Jun 1977.

Randomizing the Bessel differential equation: Solution and probability properties

J.-Cortés[♭] , L. Jódar[♭], and L. Villafuerte[†*]

(♭) Instituto Universitario de Matemática Multidisciplinar,

Universitat Politècnica de València,

Camino de Vera s/n, 46022, Valencia, Spain

(‡) Facultad de Ciencias en Física y Matemáticas,

Univesridad Atónoma de Chiapas,

Ciudad Universitaria, 29050, Tuxtla Gutiérrez, México

December 4, 2015

1 Introduction

Much of mathematical physics is dominated by Laplace's equation and when it is translated into cylindrical coordinates, Bessel's equation inevitably appears [1]. This paper deals with the construction of a mean square solution for the Bessel random differential equation (r.d.e.)

$$t^2\ddot{X}(t) + t\dot{X}(t) + (t^2 - A^2)X(t) = 0, \qquad t > 0, \tag{1}$$

with initial conditions

$$X(t_0) = \eta_1, \qquad \dot{X}(t_0) = \eta_2. \tag{2}$$

Input data A, η_1, η_2 are assumed to random variables defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Using the so-called L_p-random calculus and assuming moment conditions on the random variables in the equation,

*e-mail:laura.villafuerte@unach.mx

a mean square convergent generalized power series solution is constructed. As a result of this convergence, the sequences of the mean and standard deviation obtained from the truncated power series solution are convergent as well. The results obtained in the random framework extend their deterministic counterpart. The theory is illustrated in two examples in which several distributions on the random inputs are assumed. Finally, we show through examples that the proposed method is computationally faster than Monte Carlo method.

The consideration of randomness in the A parameter that appears in the Bessel differential equation (1) can be motivated from physical considerations. The wave propagation generated by a electric field and its variations in the medium can be considered as being randomly varying due to unhomogeneous physical properties of the medium. As it is shown in [2], the governing equation for the electric field in a specific direction is given by a Bessel equation of the form (1), where A coefficient depends upon random medium parameters. Another contributions solving random differential equations in the mean square sense include [3, 4, 5].

2 L_p–random calculus and main contribution

In order to clarify the meaning of the derivatives in the Bessel equation (1), we will summarize the main definitions and results of the L_p–random calculus. Further details about them can be found in [4, 5]. Let $p \geq 1$ be a real number. A real random variable X defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called of order p (in short, p-r.v.), if $\mathrm{E}[|X|^p] < \infty$, where $\mathrm{E}[\,]$ denotes the expectation operator. The set $\mathrm{L}_p(\Omega)$ of all the p-r.v.'s endowed with the norm

$$\|X\|_p = (\mathrm{E}[|X|^p])^{1/p} \,,$$

is a Banach space. Let $\{X_n : n \geq 0\}$ be a sequence in $\mathrm{L}_p(\Omega)$. We say that it is convergent in the p-th mean to $X \in \mathrm{L}_p(\Omega)$, if $\lim_{n\to\infty} \|X_n - X\|_p = 0$. For $p = 2$, this 2-th mean convergence is usually referred to as mean square convergence. If there exists a stochastic process $\frac{dX(t)}{dt} \in \mathrm{L}_p(\Omega)$ such that

$$\left\| \frac{X(t+h) - X(t)}{h} - \frac{dX(t)}{dt} \right\|_p \xrightarrow[h\to 0]{} 0 \,, \quad t, t+h \in \mathcal{T}, \tag{3}$$

then we say that the stochastic process $X(t)$ is p-th mean differentiable at $t \in \mathcal{T}$ and its p-th mean derivative at t is given by $\frac{dX(t)}{dt}$. The notation $\dot{X}(t)$

is also used for the p-th mean derivative of the stochastic process $X(t)$ at the point t. Establishing new results related to the so-called L_p-random calculus, it is shown that under moment conditions on the r.v.'s in the Bessel equation, there exists a stochastic process with first and second 2-th mean derivatives, which is a mean square solution of the Bessel equation, precisely:

Theorem 1 *Let A be a random variable satisfying conditions*

i) $\mathbb{P}\left[\{\omega \in \Omega :\ A(\omega) \geq 0\}\right] = 1.$

ii) $A \in L_{16}(\Omega)$.

iii) There exist positive numbers $r_1 > 16$ and $\delta_1 > 0$ such that

$$\sup_{s\in[-\delta_1,\delta_1]} \mathrm{E}\left[e^{Ar_1 \ln(t+s)}\right] < \infty.$$

iv) There exist positive numbers $r_2 > 16$ and $\delta_2 > 0$ such that

$$\sup_{s\in[-\delta_2,\delta_2]} \mathrm{E}\left[e^{-Ar_2 \ln(t+s)}\right] < \infty.$$

v) $A(\omega) \in \cup_{n=0}^{\infty}[a_n, b_n]$ *w.p. 1, where* $n < a_n < b_n < n+1$.

Assume that the r.v.'s $\eta_i \in L_4(\Omega)$, $i = 1, 2$, and the r.v. A is independent of random variables η_i, $i = 1, 2$. If $d_n := \min\{n - b_{n-1}, |n - a_n|\}$ and there exists $\epsilon > 0$ such that $0 < \epsilon < d_n$ for all $n = 1, 2, 3, ...$ then the initial value problem

$$t^2\ddot{X}(t) + t\dot{X}(t) + (t^2 - A^2)X(t) = 0, \qquad X(t_0) = \eta_1, \quad \dot{X}(t_0) = \eta_2, \tag{4}$$

$t_0, t \in [r_1, r_2]$, $0 < r_1 < r_2 < \infty$, *has a solution stochastic process $X(t) \in L_2(\Omega)$ given by*

$$X(t) = \alpha_1 X_1(t) + \alpha_2 X_2(t), \tag{5}$$

where

$$X_1(t) = t^A \left[1 + \sum_{n=1}^{\infty} \frac{(-1)^n}{4^n n! \prod_{i=1}^{n}(A+i)} t^{2n}\right]$$

$$X_2(t) = t^{-A} \left[1 + \sum_{n=1}^{\infty} \frac{(-1)^n}{4^n n! \prod_{i=1}^{n}(-A+i)} t^{2n}\right]$$

with

$$\alpha_1 = \frac{\eta_1 \dot{X}_2(t_0) - \eta_2 X_2(t_0)}{W(X_1, X_2)(t_0)}, \qquad \alpha_2 = \frac{\eta_2 X_1(t_0) - \eta_1 \dot{X}_1(t_0)}{W(X_1, X_2)(t_0)}, \tag{6}$$

being $W(X_1, X_2)(t_0) = X_1(t_0)\dot{X}_2(t_0) - \dot{X}_1(t_0)X_2(t_0)$, *for each* $t \in [r_1, r_2]$.

Note that the stochastic solution given by (5) is a convergent series in $L_2(\Omega)$. Thus, by truncating this series, we obtain convergent sequences of the mean and the standard deviation of such solution. In the following, several distributions on the r.v.'s A, η_1, η_2 are taken in order to illustrate the theory. It is not hard to show that such r.v.'s with their respective distributions satisfy the hypotheses of Theorem 1.

Example 1 *Assume that A has uniform distribution on the interval* $[\frac{1}{10}, \frac{9}{10}]$, *and the r.v.'s* η_i, $i = 1, 2$, *have Beta distributions,* $\eta_i \sim Be(a_i; b_i)$, $i = 1, 2$, *where* $a_1 = 1$, $b_1 = 3$, $a_2 = 2$ *and* $b_2 = 5$. *The obtained results for the mean and the standard deviation are shown in Tables 1-2, respectively. Approximations using Monte Carlo sampling with m simulations for the mean,* $\tilde{\mu}_X^m(t)$, *and the standard deviation,* $\tilde{\sigma}_X^m(t)$, *are also collected in these tables. From these data we observe that both methods agree.*

Table 1: Approximations of the mean by the proposed truncated series method ($\mathrm{E}[X_N(t)]$) and Monte Carlo sampling ($\tilde{\mu}_X^m(t)$) using different orders of truncation N and number m of simulations, respectively, at some selected time points t in the context of Example 1.

t	$\mathrm{E}[X_N(t)]$; $N = 10$	$\mathrm{E}[X_N(t)]$; $N = 20$	$\tilde{\mu}_X^m(t)$; $m = 50000$	$\tilde{\mu}_X^m(t)$; $m = 100000$
1.0	0.250000	0.250000	0.247345	0.249427
2.0	0.343619	0.343619	0.341780	0.343579
2.5	0.276392	0.276392	0.275382	0.276553
3.5	0.031725	0.031725	0.032386	0.032022
4.0	-0.088935	-0.088935	-0.087747	-0.088672

Table 2: Approximations of the standard deviation by the proposed truncated series method ($\sigma_N(t)$) and Monte Carlo sampling ($\tilde{\sigma}_X^m(t)$) using different orders of truncation N and number m of simulations, respectively, at some selected time points t in the context of Example 1.

t	$\sigma_N(t)$; $N = 10$	$\sigma_N(t)$; $N = 20$	$\tilde{\sigma}_X^m(t)$; $m = 50000$	$\tilde{\sigma}_X^m(t)$; $m = 100000$
1.0	0.193646	0.193649	0.192932	0.193486
2.0	0.165496	0.165132	0.193486	0.165337
2.5	0.128705	0.128705	0.128607	0.128744
3.5	0.078538	0.078538	0.078478	0.078603
4.0	0.090635	0.090635	0.090362	0.090572

Example 2 *Assume that A has a truncated beta distribution on [d,1-d] ,* $d = 1 \times 10^{-7}$*, with parameters* $\alpha = 1$ *and* $\beta = 3$*;* η_1 *has a standard Gaussian distribution,* $\eta_1 \sim N(0;1)$ *and* η_2 *has with uniform distribution on* $[0,1]$*,* $\eta_2 \sim U([0,1])$*. From Tables 3-4 we observe a high agreement between both approximations.*

Table 3: Approximations of the mean by the proposed truncated series method ($\mathrm{E}[X_N(t)]$) and Monte Carlo sampling ($\tilde{\mu}_X^m(t)$) using different orders of truncation N and number m of simulations, respectively, at some selected time points t in the context of Example 2.

t	$\mathrm{E}[X_N(t)];\ N=10$	$\mathrm{E}[X_N(t)];\ N=20$	$\tilde{\mu}_X^m(t);\ m=50000$	$\tilde{\mu}_X^m(t);\ m=100000$
1.0	0	0	0.005396	-0.005113
2.0	0.293729	0.293729	0.297004	0.289315
2.5	0.307795	0.307795	0.309329	0.304980
3.5	0.147978	0.147978	0.146352	0.148902
4.0	0.024559	0.024559	0.022073	0.026854

Table 4: Approximations of the standard deviation by the proposed truncated series method ($\sigma_N(t)$) and Monte Carlo sampling ($\tilde{\sigma}_X^m(t)$) using different orders of truncation N and number m of simulations, respectively, at some selected time points t in the context of Example 2.

t	$\sigma_N(t);\ N=10$	$\sigma_N(t);\ N=20$	$\tilde{\sigma}_X^m(t);\ m=50000$	$\tilde{\sigma}_X^m(t);\ m=100000$
1.0	1.000000	1.000000	1.004220	1.002310
2.0	0.670807	0.670807	0.674305	0.672227
2.5	0.364257	0.364257	0.366328	0.364783
3.5	0.322779	0.322779	0.323623	0.323559
4.0	0.484010	0.484010	0.485911	0.485238

A comparison of the CPU time used in Mathematica® 7.0 to compute some numerical results presented in Tables 1-4 is shown in Table 5. These data show that the proposed truncated series method is faster than the Monte Carlo Method.

Table 5: Execution time for computing the mean and variance for Examples 1 and 2 implemented on Intel® Core™ 2 Duo, 4GB, 2.4GHz.

Methods	Monte Carlo 10×10^4 simulations CPU(seconds)	Truncated series method truncation order N=20 CPU(seconds)	% Increase
Example 1	94.30	31.46	300.2
Example 2	94.32	3.49	2702.6

3 Conclusions

In this paper mean square convergent generalized power series solution of the random Bessel differential equation (4) have been constructed taking advantage of L_p-random calculus. The results obtained extend their deterministic counterpart under mild conditions. In addition, general expressions to approximate both the mean and the variance of the solution have been determined. An important feature of our analysis is that these approximations are guaranteed to converge to their respective exact values. To illustrate the reliability of the results, two examples have been provided. Finally, we want to point out that our approach can be very useful to continue studying, from a probabilistic standpoint, other kind of Bessel differential equations.

Acknowledgements. This work has been partially supported by the Spanish Ministerio de Economía y Competitividad grant MTM2013-41765-P and by the European Union in the FP7-PEOPLE-2012-ITN Program under Grant Agreement no. 304617 (FP7 Marie Curie Action, Project Multi-ITN STRIKE-Novel Methods in Computational Finance) and Mexican Conacyt.

References

[1] F. E. Relton, Applied Bessel Functions, London, Blackie and Son Limited, 1946.

[2] T. T. Soong, S. N. Chuang, , Solutions of a class of random differential equations, SIAM Journal on Applied Mathematics 24 (4): 449–459, 1973.

[3] M. El-Sohaly, Mean square convergent three and five points finite difference scheme for stochastic parabolic partial differential equations, Electronic Journal of Mathematical Analysis and Applications 2: 66–84, 2014.

[4] T. T. Soong, Random Differential Equations in Science and Engineering, New York, Academic Press, 1973.

[5] L. Villafuerte, C. A. Braumann, J. C. Cortés, L. Jódar, Random differential operational calculus: Theory and applications, Computers and Mathematics with Applications 59 (1): 115–125, 2010.

A stochastic capacitated lot sizing problem under vendor managed inventory for the paper industry

Lorena Polo[♭] [*], David Escuín[♭], and David Ciprés[♭]

(♭) Aragon Institute of Technology,

María Luna 7-8, 50018 Zaragoza, Spain

November 30, 2015

1 Introduction

Production lot-sizing has a great impact on inventory, particularly under seasonal fluctuations of demand and constrained production capacity. Many companies adopt the MTO (Make To Stock) policy in which products are not built until a confirmed order for products is received by the manufacturer. Other companies maintain high levels of inventory (stock) to face periods of uncertain demand. However, a production schedule which does not adjust accurately the real demand may lead to overstocks for some products and stock-outs for other. Inventory sizing by product is especially important under uncertainty, when the inventory is necessary to guarantee a service level in a stochastic environment. One of the integration practices that can contribute to reduce inventory in the supply chain is Vendor Managed Inventory (VMI). VMI programs allow for consumer demand information to be disseminated up the supply chain, thus mitigating upstream demand fluctuations due to the bullwhip effect [4] and [6] . Due to this demand anticipation, VMI may allow to reduce logistic and manufacturing costs, reduce overall lead-times, improve service level and reduce transportation costs.

[*]e-mail: lpolo@itainnova.es

The aim of this paper is to develop a mathematical model to seek the most effective inventory mix to face stochastic demand at minimum cost in a two-level supply chain. We focus on a multi-product dynamic lot-sizing problem under stochastic demand subject to capacity and service level constraints. Unlike previous studies, this model is executed to compare a MTO strategy to a VMI partnership between the supplier and their customers [2]. Both policies are developed within the model, and their results are compared within the numerical application.

In the problem, a schedule of production orders is determined over the planning horizon in order to minimize the inventory holding costs of the supply chain, taking into consideration that the supplier is also responsible of initiating the replenishment orders and deliveries of their customers according to the VMI partnership. The model also considers features such as service level required, the production capacity at machine level, set up time or product-machine allocation. The integration of stochastic demand in the production/inventory model is performed through the statistical distribution of the forecast accuracy. Historical data are analyzed to select the most suitable forecasting model for each reference (also called SKU). The selected model is triggered to forecast the demand during the rolling horizon.

The applicability of the proposed model is illustrated empirically using a real case study: a paper manufacturing company that pursues to improve customer service level and supply chain inventory costs through a proper production planning of their paper machines and a suitable VMI order replenishment schedule. A cost analysis of the supply chain inventory under different service levels and under different adoption rates of VMI demonstrates the potential of this model to improve performance in the supply chain.

2 Methodology for the the paper industry problem

The case study addressed in this paper is related with the supply chain for the pulp and paper industry . The sector comprises companies that use wood as raw material and produce pulp, paper, board and other cellulose-based products. This work is focused in the two-echelon operations from the manufacturers (facilities that processes the raw material) and the customers

that receives the paper and produces strong, lightweight cardboard boxes. The manufacturers must meet the customer's requirements in terms of paper quality, wide and other paper properties. Several approaches are launched to schedule the production into the paper machines. Board must be dispatched to the customers by different transport modes as in [3].

The study compares the performance of VMI and MTO strategies in the two-echelon serial supply chain with finite production capacity. The performance is investigated using difference equations forming a simulation model. The purpose of using simulation is to provide a simplified environment into which a number of situations and ideas can be tested. From the simulation outputs, the effect of the strategies on manufacturing on-costs, inventory holding costs and transport costs will be quantified and discussed.

As said before, with an MTO strategy, the manufacturer produces in response to a customer order, whereas the manufacturer produces in anticipation of orders with an VMI strategy (see [1] and [5]). The paper industry considered in the use case triggers orders in a calendar agreed with the manufacturer (weekly or every two weeks). To calculate the size of the order, the customers estimate the demand during the lead time and dispatched as [7] [8]. It cause large orders to cover the demand under the uncertainty. This is the key aspect that VMI seeks to lower.

2.1 Overall architecture.

Firstly, at initial stages, it is necessary to analytically calculate the customers safety stock (how much extra stock that is maintained to mitigate risk of stockouts due to uncertainties in supply and demand), the manufacturer inventory, and fill rate for VMI and MTO strategies. They are based on the lead time, the delay between the time the reorder point is reached and renewed availability. This will be the starting point prior to adjust a iterative approach aimed at performing forecasting and optimization tasks iteratively following similar optimization as in [10]. The overall vision of the paper is reflected in the Figure 1 and is also generalized by [9]. Manufacturers and customers provide qualitative and quantitative information to adopt the forecasting and decision models. Experts work in a collaborative environment and demand forecasts and error estimates (an estimate of how far actual demand may be from forecasted demand) feed the decision model which optimizes the lot sizing (quantity of paper to be produced) of the manufacturers. Historical data stored at the manufactures and warehouses

is used as an input to update models and fit forecasts.

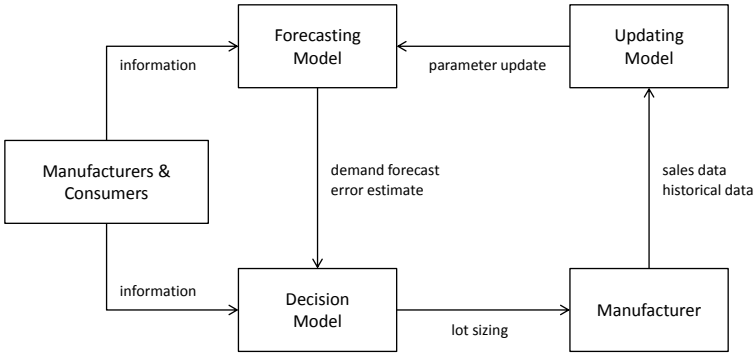


Figure 1: Iterative approach general flow chart

2.2 Objective function

The problem can be described as follows. Let K be the set of customers and M the number of manufacturers. Let S be the set of SKUs and D be the number of days of the simulation. Let $s_{k,s,d}$ be the stock for the customers. The stock VMI is represented by $svmi_{m,s,d}$ and stock for MTO $smto_{m,s,d}$. Other variables are the quantity of product sent by the manufacturers ($e_{m,s,d}$) and $f_{m,s,d}$, that is the quantity of product manufactured. Let us define $\alpha^1_{s,m}$ as the cost of manufacturing a unit of item s in manufacturer m. Let $\beta^1_{s,m}, m \in M, \beta^2_{s,k}, k \in K$ be the holding cost of unit of item s for the manufacturers and customers respectively. The transportation costs from the manufacturers to the customers are represented by $\zeta^1_{m,k}$. The objective function 1 minimizes the total cost that aggregates production, inventory and transportation costs is:

$$min(\alpha + \beta + \zeta) \tag{1}$$

where

$$\alpha = \sum_{d \in D} \sum_{s \in S} \sum_{m \in M} \alpha^1_{s,m} f_{m,s,d} \tag{2}$$

$$\beta = \sum_{s \in S} (\sum_{m \in M} \beta^1_{s,m} (svmi_{m,s,d} + smto_{m,s,d}) + \sum_{k \in K} \beta^2_{s,k} s_{k,s,d}) \tag{3}$$

$$\zeta = \sum_{m \in M} \sum_{k \in K} \zeta^{1}_{m,k} e_{m,k,s,d} \tag{4}$$

3 Conclusions

The work addresses the comparison between two replenishment policies. The results demonstrate the effectiveness of the model and the algorithms under VMI. The safety stock is reduced and, in turn, the inventory levels. By timing production and shipments, the manufacturer is able to manage the inventory level at the customers. VMI has proved to get a better performance than MTO replenishment strategy. The paper simulates the VMI and traditional supply chain MTO strategy in a scenario which previously has been configured with the forecasted demand of the customer. During the rolling horizon, the designed network of supply chain makes replenishment orders to cover the forecasted demand with the objective of minimizing the total costs. It would be interesting to extend the model to address some aspects more deeply. How the shipments are scheduled based on the availability of the transport modes (rail of road transport) and the influence of the machine cycles derive on a great impact on the supply chain.

Acknowledgment

The dissemination of this work has been partly financed by the FSE Operative Programme for Aragon (2007-2013).

References

[1] Disney, S.M. and Towill, D.R., 2003 The effect of VMI dynamics on the bullwhip effect in supply chains, International Journal of Production Economics, Vol. 85, No. 2, pp199-215. DOI: 10.1016/S0925-5273(03)00110-5.

[2] Fraile, A. Larrode, E., Magrean, A. Sicilia, J.A. Decision model for siting transport and logistic facilities in urban environments: A methodological approach. *Journal of Computational and Applied Mathematics* Volumen 290: 100-120, 2015.

[3] Gallego, J., Larrode, E., Sicilia, J.A., Royo, B. y Fraile, A. Optimization and sizing methodology of necessary infraestructure for the incorporation of hydrogen to the transportation sector. *DYNA - Ingeniería e Industria.*, Volume 89: 405–412, 2014.

[4] Govindan, K. The optimal replenishment policy for time-varying stochastic demand under vendor managed inventory. *European Journal of Operational Research.* Volume 242, Issue 2, 402-423, 2015.

[5] Johnson M., Eric, M., D, Tom. Vendor-managed inventory in the retail supply chain Waller, Journal of Business Logistics; 1999; 20, 1; ProQuest Business Collection pg. 183

[6] Kwangyeol, R., Ilkyeong M., Seungjin O., Mooyoung J. A fractal echelon approach for inventory management in supply chain networks. *International Journal of Production Economics* Volume 143, Issue 2, 316-326. 2013.

[7] Royo, B., Fraile, A., Larrode, E. y Muerza, M.V. Route planning for a mixed delivery system in long distance transportation and comparison with pure delivery systems. *Journal of Computational and Applied Mathematics.* Volumen 291: 488-496. 2016.

[8] Royo, B. Sicilia, J.A., Oliveros, M.J., Larrode, E. Solving a Long-Distance Routing Problem using Ant Colony Optimization. *Applied Mathematics & Information Sciences* Volumen 9. 415-421. 2015.

[9] Peter A. Salzarulo. Vendor managed inventory programs and their effect on supply chain performance. Submitted to the faculty of the University Graduate School in partial fulfillment of the requirements for the degree Doctor of Philosophy in the Kelley School of Business. Indiana University 2006.

[10] Sicilia, J.A., Royo, B., Quemada, C. Oliveros, M.J, Larrode, E. An decision support system to long haul freight transportation by means of ant colony optimization. *DYNA - Ingeniería e Industria.* Volume 90: 105-113. 2015.

A Tensor Bézier Shape Deformation for cluttered and uncertain spaces

L.Hilario[♭] [*], A. Falcó[♭], N. Montés[†], F. Chinesta[‡] and M.C.Mora[○]

(♭) Cardenal Herrera University,

Departamento de Ciencias, Físicas, Matemáticas y de la Computación

San Bartolomé 55, 46115 Alfara del Patriarca (Valencia), Spain,

(†) Cardenal Herrera University,

Departamento de Ingeniería de la Edificación y Producción Industrial

San Bartolomé 55 46115 Alfara del Patriarca (Valencia), Spain,

(‡) Ecole Centrale Nantes,

Institut Universitaire de France

Nantes (France),

(○) Universitat Jaume I ,

Departamento de Ingeniería Mecánica y Construcción

Avd. Vicent Sos Baynat s/n 12071 Castellón, Spain

November 30, 2015

1 Abstract

In this paper we propose an improvement of the Tensor Bézier Shape Deformation algorithm developed in our previous work, [Hilario(2015]. In this case the algorithm is applied in cluttered and uncertain spaces. The most common example of these spaces is the traffic congestion (cars, trucks and pedestrians) where an automated system wants to guide you from a initial position to a final position.

[*] luciah@uch.ceu.es

For instance how to cross a crowded square. In [Hilario(2015] a trajectory free of collisions of a mobile robot was computed obtaining the deformation of the initial trajectory through a field of known vectors. The modified trajectory to avoid the obstacles passes through these vectors. The improvement of the algorithm developed in cluttered spaces looks for the more probable trajectory to cross the cluttered space in a safely manner. Some simulations are shown at the end of the paper to illustrate the algorithm.

2 Introduction

The aim of this study is how to guide a robot in a dynamic, cluttered and uncertain space (DCUS). The problem in this situation is how the robot must work interacting with many robots or obstacles. If the space is uncertain or cluttered is difficult to predict the moviment of all of those. Our objective is how to obtain the optimal trajectory of the mobile robot in a DCUS. An example of a DCUS is a square full of cars and motorbikes, pedestrians! and the aim of this algorithm is to find the best trajectory to cross this intersection avoiding all the obstacles.

Our previous work defined an algorithm based on the deformation of a Bézier trajectory through a field of vectors (forces). The consequence is how to guide a mobile robot with a smooth trajectory free of collisions. DCUS represents better the real situation and our previous algorithm does not work fine in this realistics environments. In fact, our new algorithm will look for a preferred trajectory because it will be computed when possible. There are few works studying this situation, for example we can find [Toit(2012), Toit(2010)]. This publication develops a procedure to approximately solve the stochastic dynamic problem. Classical motion planning algorithms, [LaValle(2006)] mostly ignore uncertainty when planning in static or dynamic environments. When the future locations of moving agents are known, the two common approaches are to add a time-dimension to the configuration space, or velocity tuning, [LaValle(2006)]. When the future locations are unknown, the planning problems is either solved locally, [Burgard(1997), Fiorini(1998), Choset(2007)], or global planners guides the robot towards the goal and a local planner reacts to the dynamic component of the environment, [LaValle(2006), Choset(2007), Clodic(2005)]. One of the attempt to extend the local planner to uncertain environments is the Probabilistic Velocity Obstacle approach, [Fulgenzi(2008)].

[Toit(2010)] represents the first formal effort to incorporate the effect of anticipated future measurements in the motion planning process.

Taking into account some different techniques, our article proposes a new algorithm designing an optimal-preffered trajectory using the deformation of a Bézier curve in a DCUS of a mobile robot. This algorithm is a redefinition of an algorithm that we proposed in a previous work.

3 Previous Work

In our previous work, an algorithm to compute an optimal deformation of a Bézier trajectory curve was defined. It was called Bézier Shape Deformation (**BSD**). The idea is to compute the deformation of a Bézier curve through a field of vectors. These field of vectors are joining a Start Point, S_i, with a Target Point, T_i. These vectors are related to repulsive forces of each obstacle. To compute the new Bézier curve as a modification of the original one it is necessary to obtain the matrix of the perturbations, X_n of the control points.

Definition 1. *The definition of the **Bézier Trajectory curve** is written in equivalent matrix form,*

$$\boldsymbol{\alpha}_t^n(u) = P_n(t)\,\mathbf{B}_n(u); u \in [0,1] \tag{1}$$

where

$$P_n(t) = \left[\begin{array}{ccc} \mathbf{P}_n^0(t) & \cdots & \mathbf{P}_n^n(t) \end{array}\right] \in \mathbb{R}^{2\times(n+1)} \tag{2}$$

$$\mathbf{B}_n(u) = \left[\begin{array}{ccc} B_{0,n}(u) & \cdots & B_{n,n}(u) \end{array}\right]^T \in \mathbb{R}^{(n+1)\times 1}. \tag{3}$$

where, n is the order of the Bézier curve, $B_{i,n}(t)$ are the Bernstein Basis, $u \in [0,1]$ is the intrinsic parameter and $(n+1)$ Control Points, $\boldsymbol{P}_n^i$ such that $i = 0,1,\cdots,n$.

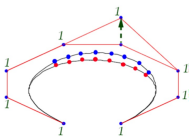
Definition 2. *The matrix expression of the **modified Trajectory Bézier curve** $\boldsymbol{\alpha}_{t+\Delta t}^n$ is given by,*

$$\boldsymbol{\alpha}_{t+\Delta t}^n(u) = (P_n(t) + X_n)\,\mathbf{B}_n(u); \quad u \in [0,1]. \tag{4}$$

*where X_n is the matrix of **the perturbations of the control points**,*

$$X_n = \left[\begin{array}{ccc} \mathbf{X}_n^0 & \cdots & \mathbf{X}_n^n \end{array}\right] \in \mathbb{R}^{2\times(n+1)}. \tag{5}$$

To deform a given Bézier curve, the control points must be changed and the perturbation, $\mathbf{X_n^i}$, of every control point must be computed.



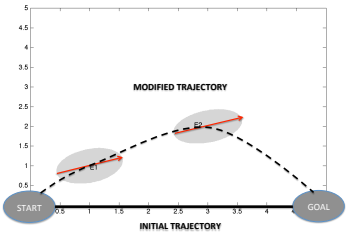
This problem is solved defining a constrained optimization problem. It is solved with the Lagrange Multipliers Theorem. The **optimization function minimizes the distance** between the orginal Bézier curve, $\alpha_t^n(u)$, and the modified Bézier curve, $\alpha_{t+\Delta t}^n(u)$.

Definition 3. *The **optimization function** using a finite set of **concatenate Bézier curves** $\alpha_t^{n_1},\ldots,\alpha_t^{n_k}$ of degrees $n_1,\ldots,n_k$, respectively is defined as,*

$$\Phi_n(X_{n_1},\cdots,X_{n_k}) := \sum_{i=1}^{k}\int_0^1 \|\alpha_t^{n_i}(u)-\alpha_{t+\Delta t}^{n_i}(u)\|_2^2\,du \tag{6}$$

4 Reassesment of the algorithm: Cluttered-Tensor Bézier Trajectory Deformation (C-T-BTD)

As a consequence of the problematic of the real situation, it is necessary the redefinition of our previous algorithm because we would like to guide the robot in a cluttered and uncertain spaces. For that reason the T-BSD has been rethought and now the new algorithm is called **Cluttered-Tensor Bézier Shape Deformation** (C-T-BSD). Considering a set of probabilities represented by ellipses, the eigenvector of each ellipse represents the right direction to follow by the Bézier modified curve. We would like to compute the preferred trajectory joining the Initial Point and the Goal Point. The idea is how to obtain the best tangent behaviour between the Bézier modified curve and the set of eigenvectors given by the ellipses.



Considering the eigenvector of an ellipse, $\mathbf{v}\in\mathbb{R}^{2\times 1}$ and he derivative of the modified Bézier curve. The new optimization function is defined as,

$$f(X_n) =< \mathbf{v}, \frac{d\alpha_{t+\Delta t}^n(u)}{du} > \tag{7}$$

considering $< .,. >$ the scalar euclidean product.

If we want to obtain the best tangent between the Bézier curve and the eigenvector of the ellipse, we should maximize this function.

5 Conclusions

The critical situations for the autonomous systems are the cluttered and uncertain spaces. It is difficult to predict the movements of the humans and the movements of obstacles when the path planning is solved in a congested environment. We have defined an initial framework to look for the preferred trajectory between a Initial Point and a Goal Point guiding the autonomous system trying to be free of collisions. A preferred Bézier trajectory is defined computing the best tangent between the ellipses and the modified Bézier Trajectory.

References

[Toit(2012)] N.E.Du Toit, and J.W. Burdick: *Robot Motion Planning in Dynamic Uncertain Environments*. IEEE Transactions on Robotics, Vol. 28, NO. 1, February 2012.

[Toit(2010)] N.E.Du Toit and J.W. Burdick: *Robot Motion Planning in Dynamic, Cluttered, Uncertain Environments*. IEEE International Conference on Robotics and Automation, May 2010.

[LaValle(2006)] S.M. LaValle: *Planning Algorithms*. Cambridge University Press, 2006.

[Burgard(1997)] D. Fox, W. Burgard and S. Thrun: *The dynamic window approach to collision avoidance*. IEEE Robotics and Automation Magazine, Vol.4, no.1, pp. 23-33, 1997.

[Fiorini(1998)] P. Fiorini and Z. Shiller: *Motion planning in dynamic environments using velocity obstacles*. International Journal of Robotics Research, Vol. 17, no. 7, pp. 760-722, July 1998.

[Choset(2007)] H. Choset, K.M.Lynch, S. Hutchinson, G. Kantor, W. Burgard, L.E.Kavraki and S. Thrun: *Principles of Robot Motion*. MIT Press, 2007.

[Clodic(2005)] A. Clodic, V. Montreuil, R. Alami and R. Chatila: *A decisional framework for autonomous robots interacting with humans*. IEEE International Workshop on Robot and Human Interactive Communication, August 2005, pp. 543-548.

[Fulgenzi(2008)] C. Fulgenzi, A. Spalanzani and C. Laugier, *Dynamic obstacle avoidance in uncertain environment combining pvos and occupancy grid*. IEEE International Conference on Robotics and Automation, April 2007.

[Hilario(2015] Hilario L., Falcó A., Montés N., Mora M.C. *A tensor optimization algorithm for Bézier Shape Deformation*. Journal of Computational and Applied Mathematics, doi: 10.1016/j.cam.2015.02.035.

Closed-form formulae vs. PDE based numerical solution for FRAP data processing: Theoretical and practical comparison

Š. Papáček[♭] *, B. Macdonald[†], and C. Matonoha[‡]

(♭) University of South Bohemia in Č. Budějovice, Institute of Complex Systems, FFPW, Zámek 136, 373 33 Nové Hrady, Czech Republic,

(†) University of Glasgow, School of Mathematics and Statistics, Glasgow, G12 8QW, Scotland,

(‡) Institute of Computer Science, Academy of Sciences of the Czech Republic, Pod Vodárenskou věží 2, 182 07 Prague 8, Czech Republic.

November 30, 2015

1 Introduction

Fluorescence recovery after photobleaching (FRAP) is a widely used method to analyse (usually using fluorescence microscopy) the mobility of either fluorescently tagged or autofluorescent proteins (e.g., photosynthetic) in living cells. The FRAP method works by imaging the recovery of fluorescence intensity over time in a region of interest previously bleached by a high-intensity laser pulse. Whilst the basic principles of FRAP are simple and the experimental setup is usually fixed, quantitative FRAP data analysis is not well developed. Several models and numerical procedures are used for the underlying model parameter estimation, i.e., the data processing step is not currently well established. Hence, a natural inquiry: Are all FRAP data processing methods well suited?

*e-mail: spapacek@frov.jcu.cz

In this paper we rigorously formulate the inverse problem of model parameter estimation (including the sensitivity analysis) making possible the comparison of different FRAP data processing methods. Then, in a case study with simulated datasets, we focus on how different methods for data (pre)processing influence the error in parameter estimation. We explain why our new method based on a general initial-boundary value problem formulation for the Fick diffusion PDE should exhibit less bias and narrower confidence intervals of the estimated model parameters, than other state-of-the-art methods.

2 Model parameter estimation based on spatio-temporal data

We aim to determine both the model parameter value $p \in \mathbb{R}^q,\ q \in \mathbb{N}$, and the corresponding confidence interval, subject to the noise in data and a quantity related to the sensitivity, see (4). The data are represented by a (measured) signal on a Cartesian product of the space-points $(x_i)_{i=1}^n$ and time-points $(t_j)_{j=1}^m$. Let $N_{\mathrm{Data}} := m \times n$ be the total number of spatio-temporal data points. We define the operator $S : \mathbb{R}^q \to \mathbb{R}^{N_{\mathrm{Data}}}$ that maps parameter values $p_1, \ldots, p_q$ to the solution of the underlying initial-boundary value problem, e.g. (5), evaluated at points (x_i, t_j): $S(p) = \{y(x_i, t_j, p) \in \mathbb{R}, \quad 1 \le i \le n, \quad 1 \le j \le m\}$.

It is common in the FRAP community to perform a preprocessing, e.g., space averaging [4], or not to employ all the N_{Data} measured values at points $\{(x_i, t_j),\ i = 1, \ldots, n,\ j = 1, \ldots, m\}$ [6]. Hence we define the observation operator $G : \mathbb{R}^{N_{\mathrm{Data}}} \to \mathbb{R}^{N_{\mathrm{data}}}$ that evaluates the set of values $S(p)$ on a certain subset of the full data space ($N_{\mathrm{data}} \le N_{\mathrm{Data}}$): $G(S(p)) = (z(x_l, t_l, p))_{l=1}^{N_{\mathrm{data}}}$ The forward map $F = G \circ S$, representing the parameter-to-output mapping, is defined as the composition of the PDE solution operator S and the observation operator G; $F : p \to z(x_l, t_l, p)_{l=1}^{N_{\mathrm{data}}}$. Our regression model is now

$$F(p) = \mathrm{data}. \tag{1}$$

The data are modelled and contaminated with additive Gaussian noise.

Given some data, the aim of the parameter estimation problem is to find p_T, such that (1) is satisfied in some appropriate sense. Since (1) usually consists of an overdetermined system (there are more data points than un-

knowns), it cannot be expected that (1) holds with equality, but instead an appropriate notion is that of a least-squares solution $\hat{p}$

$$\| \text{ data} - F(\hat{p}) \|^2 = \min_p \| \text{ data} - F(p) \|^2 . \tag{2}$$

Fisher information matrix and confidence intervals

For the sensitivity analysis we require the Fréchet-derivative $F'[p_1, \ldots, p_q] \in \mathbb{R}^{N_{\text{data}} \times q}$ of the forward map F, that is

$$F'[p_1, \ldots, p_q] = \left(\frac{\partial}{\partial p_1} F(p_1, \ldots, p_q), \quad \ldots \quad , \frac{\partial}{\partial p_q} F(p_1, \ldots, p_q) \right).$$

A corresponding quantity is the Fisher information matrix (FIM)

$$M[p_1, \ldots, p_q] = F'[p_1, \ldots, p_q]^T F'[p_1, \ldots, p_q] \in \mathbb{R}^{q \times q}. \tag{3}$$

Using [1], we can estimate confidence intervals. Suppose we have computed $\hat{p}$ as the least-squares solution in (2), we can then define the residual as $res^2(\hat{p}) = \|F(\hat{p}) - \text{data}\|^2 = \sum_{i=1}^{N_{\text{data}}} \left[\text{data}_i - z(x_i, t_i, \hat{p})\right]^2$. Then according to [1], it is possible to quantify the error between the computed parameters $\hat{p}$ and the true parameters p_T.

With the further comparison of certain methods of FRAP data processing in mind, we restrict ourselves on one single scalar parameter p estimation. Then, the Fisher information matrix M in (3) reduces into the scalar quantity $\sum_{i=1}^{N_{\text{data}}} \left[p \frac{\partial}{\partial p} z(x_i, t_i, p) \mid_{p=\hat{p}} \right]^2$ (corresponding to the semi-relative sensitivity and the estimate of the relative parameter error variance). The $1 - \alpha$ confidence interval for the observations, where α is the significance level, is then described as follows

$$(\frac{\hat{p} - p_T}{\hat{p}})^2 \sum_{i=1}^{N_{\text{data}}} \left[p \frac{\partial}{\partial p} z(x_i, t_i, p) \mid_{p=\hat{p}} \right]^2 \leq \frac{res^2(\hat{p})}{N_{\text{data}} - 1} f_{1, N_{\text{data}}-1}(\alpha), \tag{4}$$

where $f_{1, N_{\text{data}}-1}(\alpha)$ corresponds to the upper α quantile of the Fisher distribution with 1 and $N_{\text{data}} - 1$ degrees of freedom. Notice the term $\frac{res^2(\hat{p})}{N_{\text{data}}-1}$ is an estimator of the error variance, such that the approximation $\frac{res^2(\hat{p})}{N_{\text{data}}-1} \sim \sigma^2$ holds if N_{data} is large [1].

3 Comparison of several FRAP data processing methods

The fluorescence intensity measured in a region of interest Ω is proportional to the concentration of non-bleached particle's c_{nb}. It holds $\frac{c_{nb}}{c}+\frac{c_b}{c}=1$, where c_b is the bleached particles concentration, and c is the overall concentration. The Fick diffusion equation holds for both variables, c_{nb} and c_b respectively. Hence, for the sake of simplicity, we further study the spatio-temporal dynamics of the fraction $\frac{c_b}{c} := y$, as a function of one spatial coordinate x, time t and diffusion coefficient p

$$\frac{\partial y}{\partial t} - p\frac{\partial^2 y}{\partial x^2} = 0 \ , \tag{5}$$

in $(t_0, T) \times \Omega$, with suitable boundary conditions on $(t_0, T) \times \partial\Omega$ and initial conditions in Ω, where $\Omega \subset \mathbb{R}$, being the bleached particles distribution just after the bleaching.

The one-point Mullineaux method

C. W. Mullineaux *et al.* [6] measured one-dimensional bleaching profiles along the specimen long axis. They took the initial bleaching profile (of bleached particles) as a Gaussian with half-width r_0 at height (or depth) $y_{0,0}e^{-2}$, i.e., the initial condition in the form of

$$y(x, t_0) = y_{0,0} \exp \frac{-2x^2}{{r_0}^2}, \tag{6}$$

where t_0 corresponds to the time instant just after bleaching and can be set to zero. For the infinite domain, i.e., $x \in \mathbb{R}$, the solution $y(x, t, p)$ of (5) is

$$y(x, t, p) = \frac{y_{0,0}\ r_0}{\sqrt{{r_0}^2 + 8pt}} \exp \frac{-2x^2}{{r_0}^2 + 8pt}, \quad x \in \mathbb{R}, \quad t \in [0, T]. \tag{7}$$

In order to determine the diffusion coefficient, Mullineaux *et al.* [6] took only the time evolution of the normalized fluorescence intensity in the center of the bleached area ($x = 0$), i.e., $z(t) := y(0, t, p)/y_{0,0}$, where the maximum bleached level is reached. Continuing with our terminology developed in Section 2, we have $z_M(t, p) = \frac{r_0}{\sqrt{{r_0}^2+8pt}}$, and the corresponding FIM, based on the

semi-relative sensitivities and normalized output (divided by the maximum value $y_{0,0}$), collapses to the scalar quantity

$$M_M = \sum_{j=1}^{m} \left[p \frac{\partial z_M(t_j, p)}{\partial p} \right]^2 = \sum_{j=1}^{m} \frac{(4 r_0 p t_j)^2}{({r_0}^2 + 8 p t_j)^3} = \frac{1}{4} \sum_{j=1}^{m} \frac{(8 s_j)^2}{(1 + 8 s_j)^3}, \qquad (8)$$

where we introduced the dimensionless variable $s_j := \frac{p t_j}{{r_0}^2}$.

Initial boundary value problem for (5) and the FIM calculation

As the closed-form formulae approaches have several limitations or inaccuracies, we propose to model the diffusion process by the Fick diffusion equation with realistic initial and boundary conditions instead, e.g., as an initial boundary value problem (5-6). The parameter estimation is then formulated as an ordinary least squares problem (2) resulting in the estimate $\hat{p}_{PDE}$, see e.g., [7, 4].

The uncertainty assessment of the estimate $\hat{p}_{PDE}$ is based on the numerical evaluation of FIM. For each time instant t_j we denote $p_j = \hat{p}_{PDE}$, our software CA-FRAP solves (5) and using the simulated output $y(x_i, t_j, p_j)$, $i = 1, \ldots, n.$, according to (3), evaluates the FIM (diagonal in this case).

Our examples on simulated FRAP datasets show an agreement with the theoretical prediction for the low data noise levels (for signal-to-noise ratio until 30) only. FRAP data analysis is a challenging problem and a recent research direction, and scaling up to larger noise scenarios is work in progress.

Conclusion: We have developed a unified modelling framework for processing of FRAP data sets in order to compare two basic approaches of the fluorescent compounds mobility determination from the spatio-temporal FRAP measurement. The first and simplest approach is based on the curve fitting to a closed-form formula and obviously needs some unrealistic or hard-to-accomplish conditions. The second approach is based on a numerical approximation of the Fick diffusion PDE (5). Both approaches are implemented in our software CA-FRAP, which advantageously uses the UFO (Universal Functional Optimization) system [5]. CA-FRAP simultaneously provides the parameter estimate and the corresponding standard error (with is thereby used to calculate the confidence interval).

Acknowledgements

This work was supported by the MEYS of the Czech Republic - projects CENAKVA (No. CZ.1.05/2.1.00/01.0024), CENAKVA II (No. LO1205 under the NPU I program), and by the long-term strategic development financing of the Institute of Computer Science (RVO:67985807).

References

[1] Bates D. M., and Watts D. G. Nonlinear regression analysis: Its applications. New York, John Wiley & Sons, 1988.

[2] Cintrón-Arias A., Banks H. T., Capaldi A., and Lloyd A. L. A sensitivity matrix based methodology for inverse problem formulation. *J. Inv. Ill-Posed Problems*, Volume (17):545–564, 2009.

[3] Ellenberg J., *et al.* Nuclear membrane dynamics and reassembly in living cells: targeting of an inner nuclear membrane protein in interphase and mitosis. *The Journal of Cell Biology,* Volume (138):1193-1206, 1997.

[4] Kindermann, S., and Papáček, Š. On Data Space Selection and Data Processing for Parameter Identification in a Reaction-Diffusion Model Based on FRAP Experiments. *Abstract and Applied Analysis*, Volume (2015):17 pp., Article ID 859849, 2015.

[5] L. Lukšan, M. Tůma, C. Matonoha, J. Vlček, N. Ramešová, M. Šiška, and J. Hartman, *UFO 2014 - Interactive system for universal functional optimization*. Technical Report V-1191, Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague 2014 (http://www.cs.cas.cz/luksan/ufo.html).

[6] Mullineaux C. W., Tobin M. J., and Jones G. R. Mobility of photosynthetic complexes in thylakoid membranes. *Nature*, Volume (390):421-424, 1997.

[7] Papáček Š., Kaňa R., and Matonoha C. Estimation of diffusivity of phycobilisomes on thylakoid membrane based on spatio-temporal FRAP images. *Mathematical and Computer Modelling*, Volume (57):1907–1912, 2013.

Custering based on rules and post-conceptualization for learning profiles on Borderline Personality Disorder

Karina Gibert[♭] *and Marc Ferrer[†]
and Óscar Andión[†] and Luis Salvador-Carulla[‡]

(♭)Knowledge Engineering and Machine Learning group and
Dep. Statistics and Operations Research, Universitat Politècnica de Catalunya, Barcelona, Spain,

(†) Psychiatry Department, Hospital Universitari Vall d'Hebron, Barcelona, CIBERSAM, Spain

(‡) Mental Health Policy Unit, Brain and Mind Research Institute,University of Sydney, Australia

November 30, 2015

1 Introduction

Borderline Personality Disorder (BPD) is recognized as a complex syndrome characterized by a pervasive pattern of instability in regulation of emotion, interpersonal relationships, self-image, and impulse control [1]. The overall prevalence is estimated to be between 0.5% and 5.9%. BPD is considered the most common personality disorder in clinical populations representing 10% of psychiatric outpatients and 15-25% of inpatients.

BPD is associated with high risk of suicide behavior, complex pattern of comorbidity, severe functional impairment, overuse of multiple health care resources, and high costs to society.

In the last decades, considerable efforts have been made to reach a better definition of the disorder, trying further shape the current conceptualization of this prevalent and disabling condition. Indeed, it is still no clear whether

*e-mail: karina.gibert@upc.edu

BPD represents either a unitary construct or different subgroups of patients who simultaneously share certain clinical features [2]. Therefore, any contribution to conceptualization of BPD will be of benefit to clarify BPD's etiology, diagnosis, treatment and prognosis.

Knowledge Discovery (KDD) [3] arised as a new powerful approach to provide useful and understandable models from very complex phenomena, overcoming the performance of classical data analysis techniques in complex domains, as BPD is.

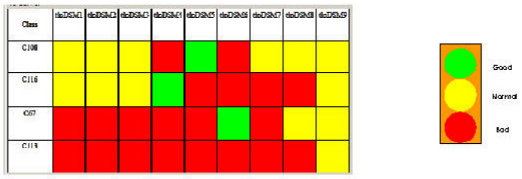
2 Data and methodological approach

In this paper, a real BDP database has been analyzed, containing 488 patients followed at the BPD Programm in Valle Hebron hospital (Barcelona, Spain). The Valle Hebrón Hospital is a University Hospital in Barcelona, managing ..... patients of all specialities, and has the single outpatient BPD Programm in Spain, started in 2005. The 488 patients analyzed in this work are decribed by 130 variables including patient's anagraphic information (Sex, age, academic level, etc), assessment scales on BDP (SCDI-II and DIB-r) and comorbidities on Mental disorders (Axis-I), Personality disorders (Axis II) and other non-mental diseases.

The main goal was to find distinct patterns of BPD patients, by using Clustering based on rules (ClBR)[4], a hybrid KDD method which combines statistical hierarchical clustering with inductive learning techniques, coming from Artificial Intelligence. Following the more integrative KDD approach, ClBR allows the formal analysis of both numerical and categorical data and can take also into account the prior expert knowledge available in the target domain, thus providing better results than classical clustering methods, where no prior knowledge is considered to biass the class formation, nor heterogeneous variables are treated together. It has proven to provide successful results in several real applications [4][5][6].

Once the classes are obtained, innovative post-processing techniques are used to understand the classes' structure. Post-processing techniques are used to bridge the gap [7] between the raw data mining results and the effective decision support layers, and are a clue to guarantee that the data mining really impacts into the clustering domain. In this work, annotated Traffic Lights Panels are used, transforming the raw data mining results into knew actionable knowledge that effectively contributes to a better understanding

Figure 1: Traffic lights pannel of the resulting 4-classes partition



of the BPD phenomenon.

3 Results

After a preprocessing step where a selection of relevant variables were guided by the experts and missing data was inputed, a priori knowledge was acquired from the experts to be taken into account for the ClBR process. Two simple rules were provided by the experts, describing extreme values in the BPD range: one regarding to patients without even impulsivity, associated to most light BPD cases, and another regarding to the most severe profiles where most of the sympthoms appear together. A ClBR with Ward method and euclidean distance was used, and Calinski-Harabaz criterion optimized over the resulting dendrogram, which suggested a partition in 4 classes. The resulting TLP is in figure 1:

From this TLP, and the additional information of patients available, the following interpretation of profiles is provided:

- Lower severity (C108):lack of suicidal behaviours and self-harming, some do not show unstable relationships; only severe Impulsivity and affective instability; high comorbid eating disorder (bulimia and atracón).
- Affective Dysregulated (C116): high affective instability and self-harming behaviours, lack of impulsivity, some do not show unstable relationships
- Behavioural Dysregulated (C67): high impulsivity and suicidal behavior, high ansiety disorder (in between impulsivity and compulsivity), patients who do not show affective instability, older and more damage than previous group
- Severe (C113): have most of the sympthoms, only paranoid ideation appears partially, as in all groups, recent self-harming behaviours

4 Conclusions

The use of ClBR and TLP as a post-processing tool provides four understandable profiles of patients with BDP. The group C108 is likely to be a non-BDP group, whereas for the other three, it is the absence of a single psychopathologycal component that is discriminating among the different profiles. Characterization by absence is uncommon in clinics and this is a promising result that might be extremely useful to better deal with the complex BPD construct.

References

[1] Sanislow, C., McGlashan, T.: Treatment outcome of personality disorders. The Canadian Journal of Psychiatry **43** (1998) 237–250

[2] Oldham, J.: Borderline personality disorder and suicidality. The American Journal of Psychiatry **163** (2006) 20 26

[3] Fayyad, U., *et alt.*: From Data Mining to Knowledge Discovery: An overview. In: Advances in Knowledge Discovery and Data Mining. AAAI/MIT Press. (1996)

[4] Gibert, K., Sonicki, Z.: Classification based on rules and thyroids dysfunctions. Applied Stochastic Models in Business and Industry **15** (1999) 319–324

[5] Gibert, K., Rodríguez-Silva, G., Annicchiarico, R.: Post-processing: bridging the gap between modelling and effective decision-support. the profile assessment grid in human behaviour. Mathematical and Computer Modelling **57(7-8)** (2013) 1633–1639 10.1016/j.mcm.2011.10.046.

[6] Gibert, K., Valls, A., Batet, M.: Introducing semantic variables in mixed distance measures. impact on hierarchical clustering. Knowledge and Information Systems **40(3** (2013) 559–593

[7] Gibert, K., Conti, D.: On the understanding of profiles by means of post-processing techniques: An application to financial assets. Environmental Engineering and Management Journal **93** (2014) in press