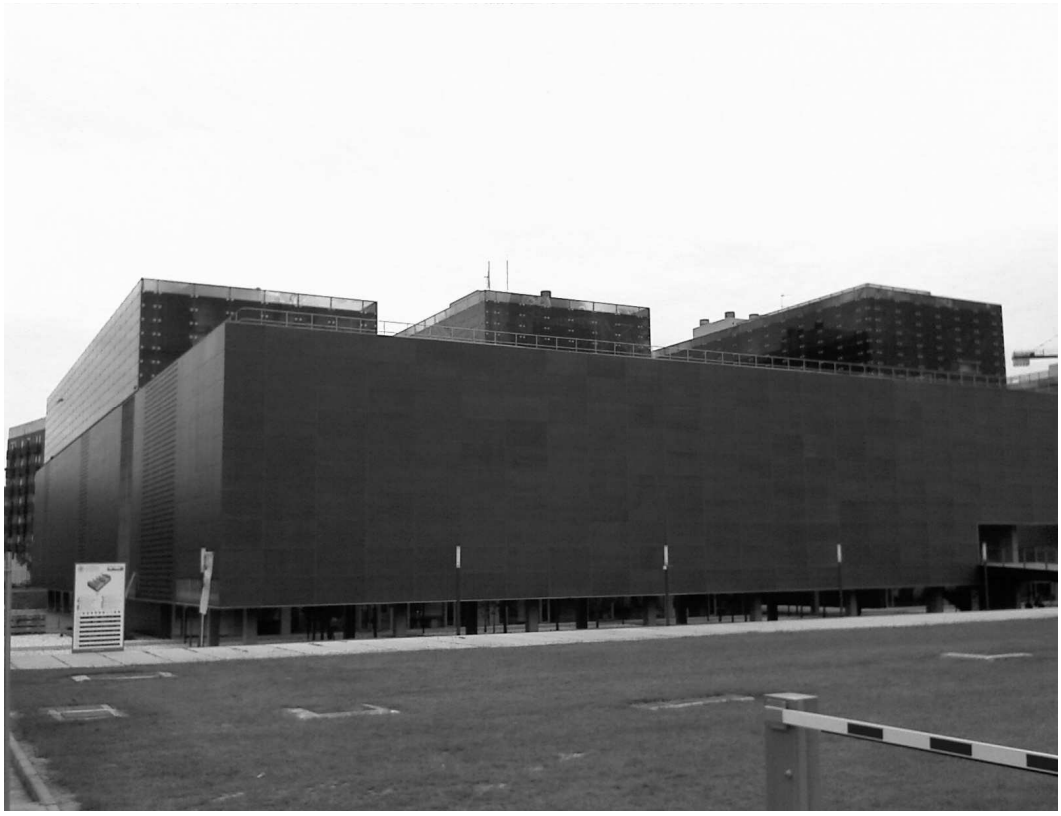
MODELLING FOR MEDICINE,BUSINESS AND ENGINEERING 2009

Instituto de Matemática Multidisciplinar



Edited by: L. Jódar, Instituto de Matemática Multidisciplinar

im² Instituto de Matemática Multidisciplinar

UNIVERSIDAD POLITECNICA DE VALENCIA

GENERALITAT VALENCIANA

MODELLING FOR MEDICINE, BUSINESS AND ENGINEERING 2009

Instituto de Matemática Multidisciplinar
Universidad Politécnica de Valencia
Valencia 46022, SPAIN

Edited by
Lucas Jódar,
Instituto de Matemática Multidisciplinar, Director
I.S.B.N.: 978-84-692-9972-2

CONTENTS

Exact and numerical solutions of the Michaelis-Menten equation.

L. Acedo* *, G. González-Parra†, and A. J. Arenas‡

(*) Instituto de Matemática Multidisciplinar,

Universidad Politécnica de Valencia,

Edificio 8G, Piso 2, 46022 Valencia, España.

(†) Departamento de Cálculo,

Universidad de los Andes,

Mérida, Venezuela.

(‡) Departamento de Matemáticas y Estadística,

Universidad de Córdoba,

Montería, Colombia.

December 10, 2009

1 Introduction

Enzymes are proteins that increase the rate of biochemical reactions. These catalyzers played a fundamental role on the origin of life on Earth and control all the reactions in living organisms.

The Michaelis-Menten equation has proven to be a simple yet powerful approach to describe enzyme processes. Its power resides in the time-independent hyperbolic relation of the initial velocity with initial substrate concentration that leads to a linear double reciprocal, from which the reaction parameters, namely the rate constant K_m and maximum velocity V_m can be determined.

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

Evoking the pseudo-steady state approximation, the Michaelis-Menten equation reduces to a single first-order nonlinear ordinary differential equation which describes the rate of depletion of the substrate of interest [20]. While this equation can be readily integrated, the resulting expression is implicit in the substrate concentration. Thus, it is common to compute the substrate concentration is by root-solving techniques such as the bisection and Newton-Raphson methods [12]. Alternatively, substrate concentration can be estimated by numerically integrating the differential form of the Michaelis-Menten equation as shown in several studies [13, 14]. In [11] a comparison of the estimation of enzyme kinetic parameters by nonlinear fitting reaction curve to different integrated MichaelisMenten rate equations was developed.

Here we are interested in improving the aforementioned methods to compute the substrate concentration in the Michaelis-Menten equation. Algebraic solutions provides a simpler alternative to numerical approaches such as differential equation evaluation and root-solving techniques that are currently used [20]. An explicit closed-form solution to the Michaelis-Menten equation has been proposed only recently [8]. This solution is based on the Lambert W function [10]. Having an accurate closed-form solution to the Michaelis-Menten equation opens up the possibility of examining the utility of other non-conventional solution techniques to solve the Michaelis-Menten equation.

In order to improve the computation of the substrate concentration in the Michaelis-Menten equation we rely on DTM and a global modal series. The DTM is a semi-analytical numerical technique depending on Taylor series that has been applied in various fields of mathematics. This technique derives from the differential equation system with initial conditions a system of recurrence equations that finally leads to a system of algebraic equations whose solutions are the coefficients of a power series solution. In order to obtain a global accurate solution we divide the time domain into several equally spaced subintervals. In this way, the differential equation can then be solved in each subdomain to obtain a piecewise finite series solution [5, 2, 1].

Here the DTM is compared numerically with multistage Adomian decomposition, Runge-Kutta methods and the global modal series. The classical Adomian method has been used in the past to obtain solutions to the Michaelis-Menten ordinary differential equation [15, 16, 17]. In [20] the Adomian decomposition with domain split has been applied successfully to the Michaelis-Menten ordinary differential equation.

It is important to remark that numerical comparisons are made regarding

two issues: accuracy and computation time. Computation time is important since in several cases the kinetic parameters are estimated by minimizing the residual sum of squares error between experimental and calculated substrate concentration data [11, 12, 20, 19]. Thus a faster computation time will improve overall estimating procedure time [20]. It is important to remark that an analytic forms are obtained with the DTM, Adomian and modal series methods. However with Runge-Kutta methods and nonlinear solvers only discrete solutions are obtained.

2 The DTM method applied to the Michaelis-Menten equation.

In this section, the differential transformation technique is applied to solve the Michaelis-Menten ordinary differential equation. The Michaelis-Menten equation (1) considered here describes the rate of depletion of the substrate of interest.

$$\dot{s}(t) = -\frac{V_m s(t)}{K_m + s(t)}, \tag{1}$$

where $s(t)$ is the substrate concentration, and V_m and K_m are the limiting rate and Michaelis constant, respectively In the next Section numerical simulations are performed with $V_m = 1$, $K_m = 1$ and initial condition $s_0 = 10$. However, in some simulations these values are modified but are clearly mentioned. From the properties of DTM ([1]), the corresponding spectrum for the system (1) can be determined as one recurrence system given by,

$$\mathbf{S}(k+1) = -\frac{V_m H_i \mathbf{Z}(k)}{k+1}, \tag{2}$$

where $Z(k)$ is the spectrum of $z(t) = -\frac{s(t)}{K_m + s(t)}$, given by

$$\mathbf{Z}(k) = \frac{\mathbf{S}(k) + \sum_{l=0}^{k-1} \mathbf{Z}(l)\Big(K_m \delta(k-l) + \mathbf{S}(k-l)\Big)}{K_m \delta(0) + \mathbf{S}(0)},\ k \geq 1\ , \mathbf{Z}(0) = \frac{\mathbf{S}(0)}{K_m \delta(0) + \mathbf{S}(0)}, \tag{3}$$

where the initial condition is $\mathbf{S}(0) = s(0)$. Thus, from a process of inverse differential transformation, it can be obtained the solutions on each subdomain taking $n+1$ terms for the power series like equation, i.e.,

$$s_i(t) = \sum_{k=0}^{n} \left(\frac{t}{H_i}\right)^k \mathbf{S}(k),\ 0 \leq t \leq H_i, \tag{4}$$

provided that the solution holds with:

$$s(t) = \sum_{i=0}^{n} s_i(t). \tag{5}$$

3 Solution of Michaelis-Menten equation by modal series

Here we obtain a global analytical solution by the modal series which avoid the split of time domain in subintervals and produces accurate solutions. This method is based on the following modal expansion series:

$$s(t) = \sum_{k=0}^{\infty} \mathcal{A}_k e^{-k\omega t}, \omega > 0, \tag{6}$$

where the frequencies and the coefficients $\mathcal{A}_k, k = 1, 2...$ are determined by a recurrence relation in terms of $\mathcal{A}_1$ and the parameters of the model, to see more details of this method we refer to [3]. Indeed, the equation can be rewritten in the following form

$$K_m \dot{s}(t) + s(t)\dot{s}(t) = -V_m s(t). \tag{7}$$

Now, from (6) it follows that

$$\dot{s}(t) = -\sum_{k=1}^{\infty} k\omega \mathcal{A}_k e^{-k\omega t} = -\sum_{k=0}^{\infty} k\omega \mathcal{A}_k e^{-k\omega t}, \tag{8}$$

and using the Cauchy product we obtain

$$s(t)\dot{s}(t) = -\sum_{k=0}^{\infty} e^{-k\omega t}\Big(\sum_{j=0}^{k} \omega j \mathcal{A}_j \mathcal{A}_{n-j}\Big). \tag{9}$$

Inserting (6), (8) and (9) into (7) one gets that

$$\sum_{k=0}^{\infty}\Big(-K_m k\omega \mathcal{A}_k - \sum_{j=0}^{k} j\omega \mathcal{A}_j \mathcal{A}_{n-j} + V_m \mathcal{A}_k\Big)e^{-k\omega t} = 0. \tag{10}$$

Thus, one gets that

$$-K_m k\omega \mathcal{A}_k - \sum_{j=0}^{k} j\omega \mathcal{A}_j \mathcal{A}_{k-j} + V_m \mathcal{A}_k = 0, \text{ for } k = 0, 1, ..., \tag{11}$$

where we have taken into account that $e^{-k\omega t}$ $(k = 0, 1, ...)$ are a linearly independent base of exponential functions. It is clear that $\mathcal{A}_0 = 0$, and if we assume that $\mathcal{A}_1 \neq 0$, then we have that

$$-K_m\omega\mathcal{A}_1 - \omega\mathcal{A}_1\mathcal{A}_0 + V_m\mathcal{A}_1 = 0, \tag{12}$$

i.e., $\omega = V_m/K_m$. Therefore, we can establish a recurrence formula for $\mathcal{A}_k$ as follows

$$\mathcal{A}_k = \frac{w\sum_{j=1}^{k-1} j\mathcal{A}_j\mathcal{A}_{k-j}}{V_m - k\omega K_m}, \text{ for } k \geq 2, \tag{13}$$

where $\mathcal{A}_1 \neq 0$ is choose suitably. The initial condition is related with the $\mathcal{A}_k s$ by

$$s(0) = \sum_{k=1}^{\infty} \mathcal{A}_k. \tag{14}$$

Theorem 3.1 *If $\mathcal{B}_k = |\mathcal{A}_k|$ is a monotone decreasing sequence, then the series $\sum_{k=1}^{\infty}(-1)^k\mathcal{B}_k$ converges absolutely, and furthermore $\sum_{k=1}^{\infty}\mathcal{B}_k < K_m$.*

Proof. If k is even, we have that

$$\mathcal{B}_{k+1} = \frac{k+1}{k}\frac{1}{K_m}\Big\{\mathcal{B}_1\mathcal{B}_k + \mathcal{B}_2\mathcal{B}_{k-1} + \cdots + \mathcal{B}_{\frac{k-1}{2}}\mathcal{B}_{\frac{k+3}{2}} + \mathcal{B}^2_{\frac{k+1}{2}}\Big\}. \tag{15}$$

Since $\mathcal{B}_k < \mathcal{B}_{k-1} < \mathcal{B}_{k-2} < \cdots$, then $\mathcal{B}_{k+1}$ is lower bounded as follows

$$\mathcal{B}_{k+1} > \frac{k+1}{k}\frac{1}{K_m}\sum_{j=1}^{\frac{k+1}{2}}\mathcal{B}_j. \tag{16}$$

Suppose that $\sum_{k=1}^{\infty}\mathcal{B}_k > K_m$, then there exists N such that $\sum_{k=1}^{\frac{N+1}{2}}\mathcal{B}_k > K_m + \epsilon$, where $\epsilon > 0$. Therefore,

$$\frac{\mathcal{B}_{N+1}}{\mathcal{B}_N} > \frac{N+1}{N}\frac{K_m+\epsilon}{K_m} > 1. \tag{17}$$

This contradicts the hypothesis of the statement. Thus, is $\sum_{k=1}^{\infty}\mathcal{B}_k < K_m$

4 Concluding remarks

In this talk we discuss a piecewise finite series approximate solutions of the Michaelis-Menten ordinary differential equation using the differential transformation method DTM, and Series modal method has been obtained. The Michaelis-Menten equation considered here describes the rate of depletion of the substrate of interest. The time domain has been splitted in subintervals and the approximating solutions are obtained in a sequence of time intervals in order to obtain accurate solutions. However, the accuracy can be increased easily by means of additional new terms. The DTM develops from the differential equation with initial condition a recurrence equation that finally leads to the solution of an algebraic equation as coefficients of a power series solution. Here we compare the effectiveness of DTM with multistage Adomian and Runge-Kutta methods. The DTM solutions shown an excellent agreement with those obtained by the Runge-Kutta methods and with the analytical solution.

In addition, we obtain a global analytical solution by a modal series expansion which avoid the split of time domain in subintervals and produces accurate solutions. However, this method is restricted to certain initial conditions. Future works include the investigation of the feasibility of the extension of the modal series expansion method to any initial condition.

The numerical results show that the DTM is accurate, easy to apply and the obtained approximate solutions preserves the positivity property of the Michaelis-Menten ordinary differential equation. Furthermore, high accuracy can be obtained without using large computation. Moreover, the DTM does not compute the derivatives or integrals symbolically and this give advantages over other methods such Taylor, power series or Adomian method. Finally, the analytic form of the DTM solution and its relatively high accuracy make this an competitive approach to solve the Michaelis-Menten equation and can be used for estimating the parameters V_m and K_m through minimizing the residual sum of squares error between experimental and calculated substrate concentration data.

References

[1] Abraham J. Arenas, Gilberto González-Parra, and Benito M. Chen-Charpentier. Dynamical analysis of the transmission of seasonal dis-

eases using the differential transformation method. *Mathematical and Computer Modelling*, 50(5-6):765–776, 2009.

[2] Hsin-Ping Chu and Chieh-Li Chen. Hybrid differential transform and finite difference method to solve the nonlinear heat conduction problem. *Communications in Nonlinear Science and Numerical Simulation*, 13(8):1605 – 1614, 2008.

[3] L. Acedo, Gilberto González-Parra and Abraham J. Arenas, An exact global solution for the classical SIRS epidemic model, Nonlinear Analysis: Real World Applications, DOI: 10.1016/j.nonrwa.2009.04.007.

[4] Hugues Berry, Monte Carlo Simulations of Enzyme Reactions in Two Dimensions: Fractal Kinetics and Spatial Segregation, Biophysical Journal, 83(4) (2002) 1891-1901.

[5] C.L. Chen, S.H. Lin, C.K. Chen, Application of Taylor transformation to nonlinear predictive control problem, Appl. Math. Model. 20 (1996) 699710.

[6] U. Kettling, A. Koltermann, P. Schwille and M. Eigen, Real-time enzyme kinetics monitored by dual-color fluorescence cross-correlation spectroscopy, Proceedings of the National Academy of Sciences of the United States of America, 95(4) (1998) 1416-1420.

[7] J.K. Zhou, Differential Transformation and its Applications for Electrical Circuits, Huarjung University Press, Wuuhahn, China, 1986 (in Chinese).

[8] S. Schnell and C. Mendoza, Closed form solution for time-dependent enzyme kinetics, Journal of Theoretical Biology, 187 (1997), 207212.

[9] L. Michaelis and M. L. Menten, Die kinetik der invertinwirkung, Biochemistry Z, 49 (1913), 333369.

[10] R. M. Corless, G. H. Gonnet, D. E. Hare, D. J. Jeffrey and D. E. Knuth, On the Lambert W function, Advances in Computational Mathematics, 5 (1996), 329359.

[11] Fei Liao, Xiao-Yun Zhu, Yong-Mei Wang and Yu-Ping Zuo, The comparison of the estimation of enzyme kinetic parameters by fitting reaction

curve to the integrated MichaelisMenten rate equations of different predictor variables, Journal of Biochemical and Biophysical Methods, 62(1) (2005), 1324.

[12] R. G. Duggleby, Analysis of enzyme progress curves by non-linear regression, Methods in Enzymology, 249 (1995), 6190.

[13] R. G. Duggleby, Quantitative analysis of the time course of enzyme catalyzed reactions, Methods, 24 (2001), 168174.

[14] C. T. Zimmerle and C. Frieden, Analysis of progress curves by simulations generated by numerical integration, Biochemical Journal, 258 (1989), 381387.

[15] B. T. Bullman and P. W. Kuchel, Comparison of biochemical simulations using integrators derived from Adomian decomposition with traditional methods, Biomedica Biochimica Acta, 8/9 (1990), 661670.

[16] A. Sen, An application of the Adomian decomposition method to the transient behavior of a model biochemical reaction, Journal of Mathematical Analysis and Applications, 131 (1988), 232245.

[17] A. Sen, An approximate solution for the transient behavior of enzyme-catalyzed reactions: The irreversible Michaelis-Menten kinetics, Mathematical Biosciences, 85 (1987), 141151.

[18] A. Sen, On the time course of the reversible Michaelis-Menten reaction, Journal of Theoretical Biology, 135 (1988), 483493.

[19] C. T. Goudar and J. R. Sonnad and R. G. Duggleby, Parameter estimation using a direct solution of the integrated Michaelis-Menten equation, Biochimica et Biophysica Acta, 1429 (1999), 377383.

[20] J.R. Sonnad and C.T. Goudar, Solution of the Michaelis-Menten equation using the decomposition method, Mathematical Biosciences and Engineering, 6 (2009), 173-188.

[21] G. Adomian, Solving frontier problems of physics: the decomposition method. Boston, Kluwer Academic Publishers, 1994.

[22] A. Repaci, Nonlinear dynamical systems: On the accuracy of Adomian's decomposition method. Appl. Math. Lett., 3(4):35-39, 1990.

Age-structured network model for the transmission of human papillomavirus.

L. Acedo*, J.-A. Moraño†,
R.-J. Villanueva‡and J.Díez-Domingo§

∗, †, ‡ Instituto de Matemática Multidisciplinar,

Universidad Politécnica de Valencia,

Edificio 8G, Piso 2, 46022 Valencia, España.

§ Centro Superior de Investigación en Salud Pública,

CSISP, Valencia, España.

December 10, 2009

1 Introduction

Human papillomaviruses (HPV) encompass more than 100 types of viruses that infects cutaneous, genital and respiratory epithelia of humans worldwide. They are the direct cause of clinically important diseases such as cervical carcinoma associated with the oncoviruses HPV16 and 18. In the last years about 500000 new cases were diagnosed anually and 200000 womend died as a consequence of the progression of cervical dysplasias to invasive carcinomas [1].

Other genotypes of HPV are also clinically relevant. In particular, HPV6 and 11 are related with genital warts. Although benign, these lesions are the most common sexually transmitted disease in the world. As much as 1 million new cases are reported each year and the burden of treatment is

*e-mail: acedo@imm.upv.es
†e-mail: jomofer@imm.upv.es
‡e-mail:rjvillan@imm.upv.es

increased by the tendency of these warts to recur after initial clearance The economic burden of the treatment of these genital warts was estimated to exceed $ 3.8 billion only in the U.S. for 1997 [2].

Taken into account the magnitude of this pandemic a lot of research effort on the development of vaccines for the most dangerous of HPV genotypes was carried out since the 80's of past century. This research was crowned with success in 1993 when a research group at the US National Cancer Institute developed non-infectious virus-like particles with the same structure that VPH16 [3]. These particles contains the capsid protein L1 of the virus but not the viral DNA. Consequently, they trigger an antibody response that protects the individual against future infections by the real virus. As it is well-known, these advances in the laboratory finally have led to the development of a quadrivalent vaccine for types VPH6, 11, 16 and 18 named Gardasil or Silgard (Merck & Co.). This vaccine is a prophylactic measure to prevent initial VPH infections by the aforementioned genotypes, which are the most important from the clinical point of view worldwide.

In order to fulfill its preventive role the vaccine must be administered to adolescent girls before becoming sexually active. On the other hand, if a person is already infected by one of the virus types VPH6, 11, 16 or 18 the vaccine could still be effective against future infections by the rest of types. In the Autonomous Community of Valencia, Spain the vaccine is already been administered to girls aged < 15 years. Similar vaccination strategies are adopted in many other countries. This way Health Services predict an important reduction in the number of cases of cervical cancer, and its precursor, cervical dysplasia in the next few years. Moreover, fewer cases also means less resources devoted to cervical biopsies, to treat children infected by their mothers, and also to treat cases of genital warts and precancerous lesions both in men and women because of the herd immunity effect.

Vaccination strategies of this kind has been studied by Elbasha et al. [6, 7, 8] by means of a compartmental model with 17 age groups for each sex. This model focus mainly on the development of cervical intraepithelial neoplasia (CIN) and its progression from CIN1 to CIN3. According to these authors vaccination must be implemented for adolescent girls aged between 12 and 14 years. Elbasha et al. also found some evidence that the vaccination of boys could also be cost-effective [6]. By vaccinating girls alone a reduction of 83% in the incidence of genital warts is expected but this reduction is increased to 97 % if boys are also vaccinated.

In a recent work by Fairley et al. [9] a decrease on the number of infected

persons and the number of persons with genital warts is already reported for Australia after one year of vaccinations of young girls. The most interesting fact unveiled by this statistical study is a clear diminishment also in the number of infections and symptoms of the disease also in boys (not vaccinated) as a consequence of herd immunity. These results are even more optimistic than the predictions of continuous models.

In this paper we propose an age-structured SIR model for the transmission of VPH 6 and VPH 11 viruses in the Autonomous Community of Valencis (Spain). We will fit these model by using data for the prevalence of VPH in Spain as given by Langeron et al. [10]. Once this model is fitted, we will simulate two different vaccination strategies: (i) only 14 year old girls are vaccinated (before the first relation) as recommended by Elbasha et al. [6]. This is the policy that is currently applied in Valencia, and (ii) both girls and boys are vaccinated at the age of 14 years. We conclude that the first policy is more suitable from the point of view of cost-effectiveness because the public health benefit obtained by vaccinating boys is marginal.

2 Mathematical model

The Spanish region of Valencia [4] is located in eastern Mediterranean Spain, with an extension of $23,255\ km^2$ and a population of $5,029,601$ million inhabitants (2008), composed by three provinces, Castellón (north), Alicante (south) and Valencia (middle). Vaccination with Gardasil was authorised in the European Union since September 2006. In 2009 a total of 18606 girls received the first dose in the Autonomous Community of Valencia, two additional booster doses are also administered in a period of six months.

In [10], it is presented Table 2 where incidence of genital warts in women per age group is shown. Let us assume that the total population is divided into the age groups determined by Table 2.

We assume that this data are also representative of the situation in Valencia prior to the vaccination. To the best of our knowledge, there is no prevalence study specific to the region we are considering. Our first objective is to propose a model that could be fitted to the data in Table 1.

Following [14, Section 6.1, p. 634 - 635], we can introduce an age-structured SIR model with ten age groups: $i = 1$ corresponds to population from 15 to 19 years old, $i = 2$ to population from 20 to 24 years old, $i = 3$ to population from 25 to 29 years old, $i = 4$ to population from 30 to 34 years

Age group	**Incidence genital warts / 100,000**
15-19	58
20-24	277
25-29	215
30-34	208
35-39	96
40-44	70
45-49	70
50-54	36
55-74	36
75+	36

Table 1: Incidence of genital warts in women per age group [10] .

old, $i = 5$ to population from 35 to 39 years old, $i = 6$ to population from 40 to 44 years old, $i = 7$ to population from 45 to 49 years old, $i = 8$ to population from 50 to 54 years old, $i = 9$ to population from 55 to 74 years old and $i = 10$ to population older than 75. Moreover, due to the papillomavirus types 6 and 11 are sexual transmitted diseases (STD) and taking into account that we assume a society mainly heterosexual, each one of the age groups should be also divided into two groups: one for women and other for men. Furthermore, for each age group we have also three subpopulations according to the state of the individuals with respect to the disease:

- Susceptible women, $SW_i(t)$, $1 \leq i \leq 10$ the fraction of those women at risk of contracting the disease per age group,
- Susceptible men, $SM_i(t)$, $1 \leq i \leq 10$ the fraction of those men at risk of contracting the disease per age group,
- Infective women, $IW_i(t)$, $1 \leq i \leq 10$ the fraction of those women infected and capable to transmit the papillomavirus types 6 or 11 per age group,
- Infective men, $IM_i(t)$, $1 \leq i \leq 10$ the fraction of those men infected and capable to transmit the papillomavirus types 6 or 11 per age group,
- Recovered women, $RW_i(t)$, $1 \leq i \leq 10$ the fraction of those women who are recovered of the disease per age group, and

- Recovered, $RM_i(t)$, $1 \leq i \leq 10$ the fraction of those men who are recovered of the disease, per age group.

Then the two-sex age-structured model is defined by the following system of differential equations

We assume that before 15, boy and girls do not have sexual contacts, i.e., they enter in the system as susceptibles. Let us assume homogeneous population mixing, i.e., each man or woman can contact sexually with any other individual of the other sex [13]. We also assume that the population of women and men are equal, i.e., the same number of individuals per age group.

Under the above assumptions, a transmission dynamics of HPV 6/11 model for Valencian population is given by the following nonlinear system of 60 ordinary differential equations:

$$
\begin{aligned}
SW_i'(t) &= \mu_i/2 + c_{i-1}SW_{i-1}(t) - (c_i + d_i)SW_i(t) - \\
&\quad \sum_{k=1}^{10} \beta_{ik} SW_i(t) IM_k(t), && (1)\\
IW_i'(t) &= c_{i-1}IW_{i-1}(t) - (c_i + d_i)IW_i(t) + \\
&\quad \sum_{k=1}^{10} \beta_{ik} SW_i(t) IM_k(t) - \gamma_i IW_i(t), && (2)\\
RW_i'(t) &= c_{i-1}RW_{i-1}(t) - (c_i + d_i)RW_i(t) + \gamma_i IW_i(t), && (3)\\
&\quad SW_i(0) = SW_i^0, IW_i(0) = IW_i^0, RW_i(0) = RW_i^0, && (4)\\
SM_i'(t) &= \mu_i/2 + c_{i-1}SM_{i-1}(t) - (c_i + d_i)SM_i(t) - \\
&\quad \sum_{k=1}^{10} \beta_{ki} SM_i(t) IW_k(t), && (5)\\
IM_i'(t) &= c_{i-1}IM_{i-1}(t) - (c_i + d_i)IM_i(t) + \\
&\quad \sum_{k=1}^{10} \beta_{ki} SM_i(t) IW_k(t) - \gamma_i IM_i(t), && (6)\\
RM_i'(t) &= c_{i-1}RM_{i-1}(t) - (c_i + d_i)RM_i(t) + \gamma_i IM_i(t), && (7)\\
&\quad SM_i(0) = SM_i^0, IM_i(0) = IM_i^0, RM_i(0) = RM_i^0, && (8)\\
&\quad i = 1, 2, 3, \cdots, 10. && (9)
\end{aligned}
$$

where

- $\mu_i = 0$ for $i = 2, \ldots, 10$ and $\mu_1 = \mu = 0.00985894$ is the birth rate in the Spanish region of Valencia [15]. We assume that the birth rate is the same for women and men,

- d_i for $i = 1, \ldots, 10$, are the death rate per each age group in the Spanish region of Valencia, that is, $d_1 = 0.000317077$, $d_2 = 0.000427157$, $d_3 = 0.000478085$, $d_4 = 0.00056897$, $d_5 = 0.000890828$, $d_6 = 0.00142877$, $d_7 = 0.00220999$, $d_8 = 0.0033346$, $d_9 = 0.0106009$, $d_{10} = 0.0802736$ [15],

- c_i for $i = 1, \ldots, 10$, are the growth rates, with $c_0 = c_{10} = 0$ (nobody is in the system before 15 and nobody grows after his/her death),

- β_{ki} for $i, j = 1, \ldots, 10$, are the transmission rate of the disease due to sexual contacts between a susceptible woman (man) of age group k and an infected man (woman) of age group i. We are assuming that the transmission rate for a sexual contact β_{ki} is the same with independence of which of both is the infected, woman or man.

- γ_i for $i = 1, \ldots, 10$, are the recovering rate per age group and they are supposed to be independent on the sex.

In each age group type-epidemiological model underlies a demographic model that allows to determine, under some assumptions, the unknown growth rates. Assuming that we are in the Spanish region of Valenica in a situation of constant population where each of the ten age groups has constant population, following [14] we have the demographic model

$$N_i'(t) \;=\; \mu_i + c_{i-1}N_{i-1}(t) + (-c_i - d_i)N_i(t),\ 1 \leq i \leq 10. \qquad (10)$$

where $N_i(t)$, $i = 1, \ldots, 10$, is the fraction of the total population corresponding to age group i for both together, women and men and μ_i, d_i, c_i have been defined in the above section.

Thus, as we assume that $N_i'(t) = 0$ for all $1 \leq i \leq 10$, then the growth rates can be computed from (10) and we obtain that $c_1 = 0.159269$, $c_2 = 0.132305$, $c_3 = 0.0991838$, $c_4 = 0.0877746$, $c_5 = 0.0914135$, $c_6 = 0.0970967$, $c_7 = 0.105703$, $c_8 = 0.119228$, $c_9 = 0.0336361$. Moreover the percentage of population per age group are $N_1(t) = 0.061778$, $N_2(t) = 0.0741296$, $N_3(t) = 0.0984097$, $N_4(t) = 0.110485$, $N_5(t) = 0.105063$, $N_6(t) = 0.0974792$, $N_7(t) =$

0.0877088, $N_8(t) = 0.0756438$, $N_9(t) = 0.203875$, $N_{10}(t) = 0.0854274$. This data is a necessary requisite to set up the initial conditions of the model.

In Figure 1 and Figure 2 we present two simulations: the vaccination of the 80% of 15-years-old girls and the vaccination of the 80% of 15-years-old girls and boys.

3 Concluding remarks

In this paper we have studied the possible benefits that could be obtained by vaccinating also boys with the standard VPH vaccine. This is an important Public Health issue because oncoviruses VPH 16 and 18 are the direct cause of half a million cases of cervical cancer all around the world. In the case of men, VPH 6 and 11 also cause genital warts that require treatment to recede and imply also an important cost to Health Services. For this reason, the U.S. Food and Drug Administration approved very recently the indication of Gardasil to prevent genital warts in men and boys.

References

[1] Kari Syrjänen and Stina Syrjänen, Papillomavirus Infections in Human Pathology, John Wiley & Sons, Chichester, West Sussex, UK, 2000.

[2] John St. Clair Roberts, Vaccine for Genital Warts, in Vaccines for Human Papillomavirus Infections and Anogenital disease, Medical Intellegence Unit 14, Robert W. Tindle, Ed., R. G. Landes Company, Austin, Texas, 1999.

[3] Caroline McNeil, Who invented the VLP Cervical Cancer Vaccines ?, Journal of the National Cancer Institute, 98(7), 2006, 433.

[4] Valencia (autonomous community) [on-line]. Available from: http://en.wikipedia.org/wiki/Valencian_Community [Accessed November 19, 2009]

[5] Bosch, F. Xavier, Ann N. Burchell, Mark Schiffman, Anna R. Giuliano, Silvia de Sanjose, Laia Bruni, Guillermo Tortolero-Luna, Susanne

Kruger Kjaer, Nubia Muoz. Epidemiology and Natural History of Human Papillomavirus Infections and Type-Specific Implications in Cervical Neoplasia. Vaccine 26 (August 2008): K1-K16.

[6] Elbasha EH, Dasbach EJ, Insinga RP. Model for assessing human papillomavirus vaccination strategies. Emerg Infect Dis 13(1): 2841. 2007 Jan. Available from http://www.cdc.gov/ncidod/EID/13/1/28.htm

[7] Elbasha EH, Galvani AP Vaccination against multiple HPV types. Math Biosci 2005;197:88117 doi: 10.1016/j.mbs.2005.05.004.

[8] Elamin H. Elbasha, Merck Research, Erik J. Dasbach, and Ralph P. Insinga, Assessing the public health and economic impact of HPV vaccination strategies, http://dimacs.rutgers.edu/Workshops/WSEconEpi/abstracts.html

[9] Fairley G, Hocking J, Chen M, Donovan, Bradshaw C. Rapid decline in warts after national quadrivalent VPH vaccine program. The 25th International Papillomavirus Conference; 2009 May 8-14; Malmo, Sweden.

[10] Langeron N, Remy V, Oyee J, San-Martín M, Cortés J, Olmos, l.: Análisis de coste-efectividad de la vacunación frente al virus del papiloma humano tipos 6, 11, 16 y 18 en España. Vacunas 2008; 9: 3 - 1129.

[11] W. O. Kermack, A. G. McKendrick, Contributions to the mathematical theory of epidemics, Part I, Proc. R. Soc. A 115 (1927) 700.

[12] L. Edelstein-Keshet, Mathematical Models in Biology, Random House, New York, 1988.

[13] J. D. Murray, Mathematical Biology, Springer-Verlag, Heidelberg, 1993.

[14] H. W. Hethcote, The mathematics of infectious diseases, SIAM Review 42-4 (2000) 599.

[15] Instituto Valenciano de Estadística, [on-line]. Available from http://www.ive.es.

[16] F. Brauer and C. Castillo-Chavez, Mathematical Models in Population Biology and Epidemiology, Springer Verlag, 2001.

[17] W.H. Press, B.P. Flannery, S.A. Teukolsky, et al., Numerical Recipes: The Art of Scientic Computing, Cambridge Univ. Press, 1986.

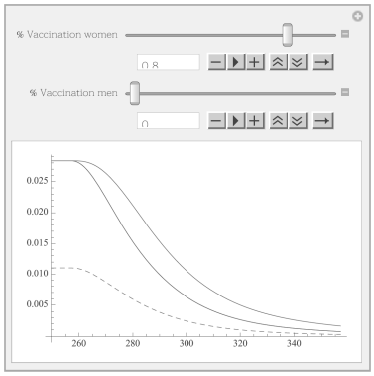


Figure 1: Simulations of vaccination strategies. On the left, we suppose the vaccination of the 80% of 15-years-old girls. The upper line corresponds to the total percentage of infected men. The middle line, to the total percentage of infected women. The lower-dashed line, to the total percentage of infected women with warts. It can be seen that the reduction of infected men is due to the herd immunity produced by the girls vaccination.

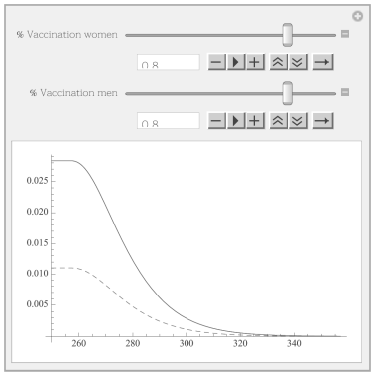


Figure 2: The same as figure 1 but for the vaccination the 80% of 15-years-old girls and boys. In this case the reduction of infected men is going some time faster.

A Mathematical Model for Measuring Hospital Performance

Caballer Tarazona, María (macata@upvnet.upv.es), Moya Clemente, Ismael (imoya@esp.upv.es), Vivas Consuelo, David (dvivas@upvnet.upv.es) and Barrachina Martínez, Isabel (ibarach@ade.upv.es).

CIEGS. Research Center for Health Economics & Management.
Universidad Politécnica de Valencia. Spain.
Edificio 7I. 3piso.
Camino de Vera s/n. c.p:46022
Valencia. Spain.
(December 10th, 2009)

1.- Introduction:

One of the main objetives of the policy in most countries is to improve their health system both in terms of quality services and efficiency, and the extent to which its resources are put to good use.
The fundamental reason to promote research into the efficiency of publicly financed hospitals in the Valencian Community is the necessity to establish the bases for the best distribution and use of healthcare resources (optimum planning), and to detect the set of problems of various kinds which affect their efficiency and capacity to offer top-quality services to the population .
The fact that several healthcare managing models exist in the same AC calls for conducting comparative studies regarding operative efficiency within hospitals. The idea is to optimise the available resources and, at the same time, to guarantee a quality homogeneous health and welfare system for all its citizens. In other words, stimulating and diffusing comparisons and benchmarking based on the works done with already existing data are initiatives that must be taken into account for the system to work well and with the willingness to promote new information transparency.

2.- Data sources and Methodology:

In order to fulfil the main objective of this article, a comparative study was conducted at 22 hospitals in the Valencian Community. In particular, the study into hospital efficiency was conducted in three main hospital units: general surgery, ophthalmology and traumatology-orthopaedic surgery. These three hospital units were selected in terms of longest waiting lists.
It is noteworthy that the hospitals studied included both public hospitals and those run by the Administrative Allowance System. All the data has been provided by the Regional Ministry of Health for 2005.

2.1. Definition of Inputs and Outputs

The Output variables considered were as follows:
- *Income*: (income x case-mix) Number of admissions weighted by the case-mix to consider the complexity of the cases.
- *Consultations*: (first consultations).
- *Successive consultations*.
- *Surgical interventions*: number of surgical interventions.

The Input variables used were the following:
- *Number of doctors*
-*Number of beds.*

2.2. Data Envelopment Analysis

The DEA Model (Data Envelopment Analysis) was the first methodology used to analyse the efficiency of the services considered.
The main objective of the DEA Model is to find a limit for efficiency formed by those combinations of resources which optimise the amount of products made by minimising production costs. Then with this limit, it assesses the relative efficiency of the combinations of resources that do not belong to it.
The model evaluates the efficiency solving this problem:

$$\max \; e_1 = \frac{\sum u_s * y_{s0}}{\sum v_m * x_{m0}}$$

$$\text{s. a :} \quad \frac{\sum u_s * y_{si}}{\sum v_m * x_{mi}} \leq 1 \qquad i = 1,\ldots\ldots\ldots, I$$

$$u_s \; v_m \geq 0 \qquad m = 1,\ldots,M \qquad s = 1,\ldots,S$$

Where:
y_{s0} = quantity of output s per DMU.
u_s = weight regarding output s.
x_{s0} = quantity of input m per DMU.
v_m = weight regarding input m.

3. Efficiency indicators and discriminate analysis

As an alternative methodology to DEA, a simpler and more operative measure, by means of indicators, is proposed which may prove more useful for the hospital management domain.
These indicators offer two basic advantages for an efficiency analysis done by the DEA Model:

- Firstly, it's simple methodology converts them into an efficiency measurement instrument that any hospital director may use.

- Secondly, fewer variables are needed to calculate the indicators than to estimate the DEA Model. So, when faced with a situation where data are scarce, it may also prove more advantageous to estimate the efficiency of the service with indicators.

The indicators proposed are:

Indicator I_1: Incomes /doctors.
Indicator I_2: Interventions /doctors.

To verify whether these indicators correctly classify services as efficient or inefficient, their efficacy has been checked by a discriminate analysis. A discriminate analysis is a statistical technique which allows an activity to be assigned to a group defined a priori (a dependent variable) in terms of a series of the group's characteristics.
Where (1) is the expression of a discriminate D_s function.

$$D_s = B_{s1} X_1 + \ldots\ldots + B_{sp} X_p + B_{s0} \qquad (1)$$

The percentage of correctly classified cases will be an effectiveness index of the discriminate functions. If these functions were effective with regards to the sample observed, it is expected that they would also be effective when classifying individuals for whom the group they belong to is unknown.
Therefore with this particular study, the belonging group will be the score obtained with the DEA Model, that is, 0 if the hospital unit is inefficient and 1 if it is efficient. Otherwise, the classifying variables will be the previously calculated indicators.

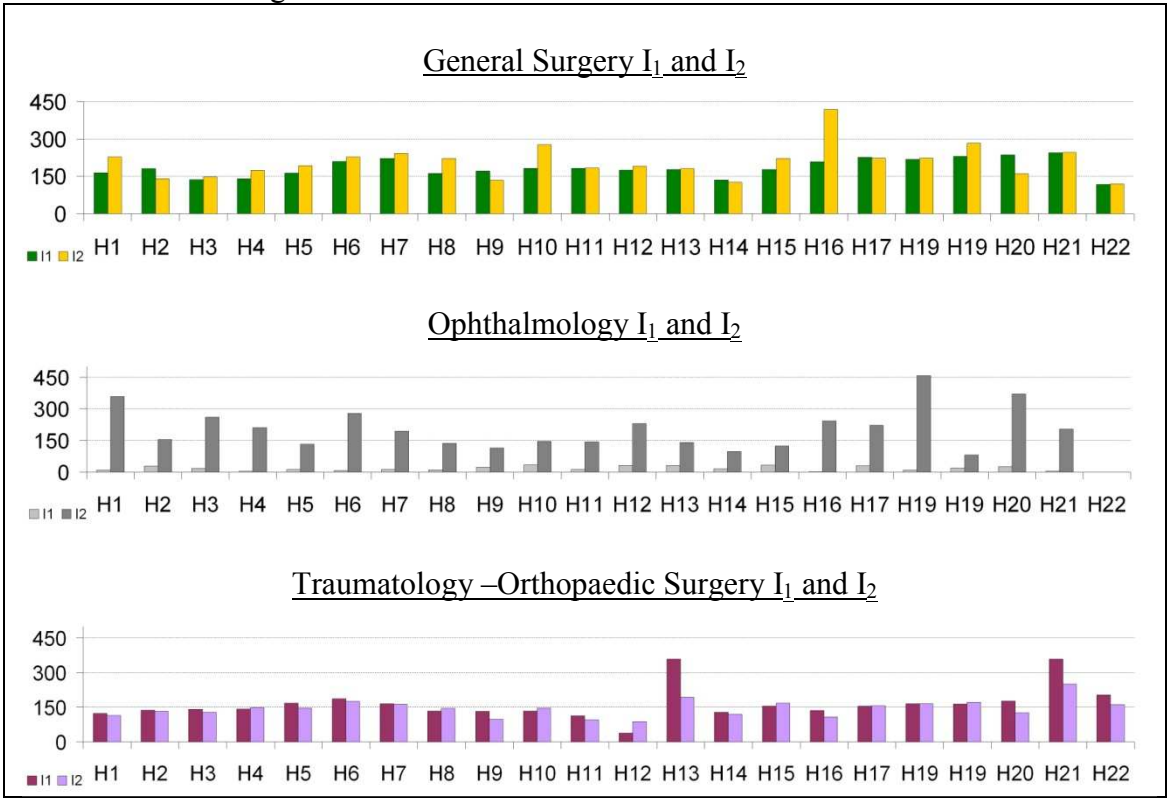
4. Results:

Applying the DEA Model to the data we obtained the following results:
Eight efficient units and 14 inefficient units were identified for the general surgery hospital unit, while 9 efficient and 12 inefficient units were found for ophthalmology. Finally, only 6 efficient but 16 inefficient units were identified for traumatology-orthopaedic surgery.

What clearly comes across from these results is that if a hospital runs one of the studied services efficiently, this does not necessarily mean that the rest of its services are also run efficiently.
Table 1 presents the results of the calculation of the indicators proposed which involves approach a simple to measure efficiency, where:

Indicator I_1: Incomes/doctors
Indicator I_2: Interventions/doctors

Table 1: Calculating the indicators.



We can see in Table 1 how the services with the highest scores for the indicators are classified as efficient as they will be the services which are capable of producing more outputs with less resources or inputs.
Then if we take the score obtained in the DEA analysis as the dependent variable, and indicators I_1 and I_2 as the classifying variables, the discriminate analysis can be done to identify the exact percentage that these indicators are able to correctly classify the different hospital units as either efficient or inefficient.
This analysis will verify whether the considered indicators are capable of classifying the different hospital units into the corresponding group, and which of these two indicators contributes to this classification to a greater extent.
Therefore, 3 discriminating functions have been obtained (2, 3 and 4), one for each hospital service under study.

General surgery service:

The results of the discriminant analysis are:

$$\mathbf{Y = -5.919 + 0.024\ I_1 + 0.007\ I_2} \qquad \mathbf{(2)}$$

Contraste de las funciones	Lambda de Wilks	Chi-cuadrado	gl	Sig.
1	,686	7,158	2	,028

		I_1	I_2
I_1	Cor. Pearson	1	,534
	Sig. (bilateral)		,010
	N	22	22
I_2	Cor. Pearson	,534	1
	Sig. (bilateral)	,010	
	N	22	22
score	Cor. Pearson	,516	,625
	Sig. (bilateral)	,014	,002
	N	22	22

For General Surgery, a discriminate function was obtained which is capable of classifying 86.4% of the hospital units correctly by using the previously calculated indicators as classifying variables.

Ophthalmology service:

The results of the discriminant analysis are:

$$\mathbf{Y = -3.657 + 0.090\ I_1 + 0.010\ I_2} \qquad \mathbf{(3)}$$

Contraste de las funciones	Lambda de Wilks	Chi-cuadrado	gl	Sig.
1	,707	6,230	2	,044

		score	I_1	I_2
score	Cor. Pearson	1	,123	,558
	Sig. (bilateral)		,594	,009
	N	21	21	21
I_1	Cor. Pearson	,123	1	-,286
	Sig. (bilateral)	,594		,208
	N	21	21	21
I_2	Cor. Pearson	,558	-,286	1
	Sig. (bilateral)	,009	,208	
	N	21	21	21

With the Ophthalmology Service, the discriminate function was capable of classifying 76.2% of the different hospital units correctly as efficient and inefficient.

Traumatology-Orthopaedic Surgery service:

The results of the discriminant analysis:

$$\mathbf{Y = -3.129 + 0.015\ I_1 + 0.005\ I_2} \qquad \mathbf{(4)}$$

Contraste de las funciones	Lambda de Wilks	Chi-cuadrado	gl	Sig.
1	,666	7,734	2	,021

		score	I_1	I_2
score	Cor. Pearson	1	,557	,701
	Sig. (bilateral)		,007	,000
	N	22	22	22
I_1	Cor. Pearson	,557	1	,838
	Sig. (bilateral)	,007		,000
	N	22	22	22
I_2	Cor. Pearson	,701	,838	1
	Sig. (bilateral)	,000	,000	
	N	22	22	22

Finally, the discriminate function of the traumatology-orthopaedic surgery service correctly classified 81.8% of the different hospital units as efficient or inefficient.
The correction percentages, with which the discriminate functions classify the hospital units, are considered high enough to accept the indicators as suitable classifying variables.

5. Conclusions:

The fundamental objective of this research work is to offer simple tools to measure efficiency in hospitals in the Valencian Community. This type of efficiency analysis becomes particularly relevant if we apply it in the context of the Valencian Community because both the DEA Model and the construction of efficiency indicators are still highly novel and relevant operative performance analysis methods for hospitals as efficiency is still a unsolved problem in Valencian hospitals and one that has been poorly dealt with.
Several conclusions have been obtained from the different analyses done to approach the measure of efficiency.
Firstly, the efficiency analysis using the DEA Model is considered more useful when studying the efficiency of each service separately instead of studying the overall efficiency of a given hospital. However, the DEA Model may present hospital directors

with practical difficulties. For this reason, the design of two easily constructed, user-friendly indicators has been proposed as an operative tool to measure efficiency in various hospital services; their effectiveness was verified by means of a discriminant analysis. This analysis offers a discriminate function for each service under study whereby high percentages of correct classifications of hospital units as efficient and inefficient were obtained by simply using the aforementioned indicators as classifying variables.

Therefore, we may conclude that the indicators here proposed are indeed an alternative measure of efficiency to the DEA Model.

With these results, healthcare administrations are recommended to provide hospitals with the mean and standard deviation of the efficiency indicators to serve as self-assessment guidelines for each hospital's activity.

Additionally, hospitals may also be provided with the previously calculated discriminate functions so that each hospital service could accurately calculate if it is indeed efficient or inefficient.

6. References

MAGMUSSEN, J et al. Measuring efficiency in clinical departments. Health Policy 87. 2008; p:1-7.

JACOBS, R. SMITH, P. STREET, A. Measuring Efficiency in Health Care. Analytic Techniques and Health Policy. Cambridge University Pess. 2006. Chapter 5.

CHARNES, A. COOPER, W. LEWIN, A. SEIFORD, L. Data Envolment Analysis. Theory, Methodology and Applications. Kluwer Academic Publishers. 1994. Chapters 1, 2 y 3.

COELLI, T. RAO, D. BATTESE, G. An Introduction to Efficiency and Productivity Analysis. Kluwer Academic Publishers. London. 1999.

GONZÁLEZ, B. Análisis Multivariante. Aplicación al Ámbito Sanitario. SG Editores, S.A. Barcelona. 1991.

BAREA, J. GÓMEZ, A. "El problema de la eficiencia del sector público en España. Especial consideración de la sanidad." Publicación del Instituto de Estudios Económicos. ISBN: 84-88533-09-8. Madrid. 1994.

BAREA TEJEIRO. J. "Organización hospitalaria y eficiencia". Rev. Gestión y Evaluación de Costes sanitarios; Vol. (2) Num. 1, Marzo.2001.

PUIG-JUNOY, J. Eficiencia en la atención primaria de salud: una revisión crítica en las medidas de frontera. Revista Española de salud pública; 2000. Vol. (74) pag. 483-495.

ADAY, L.A, BEGLEY, C.E, LARISON, D.R , SLATER. "Evaluating the medical care system: Effectiveness, efficiency, and equity". Health Administration Press, Ann Arbor: MI 1993. p.30.

KATHARAKI, M. “Approaching the management of hospital units with an operation research technique: The case of 32 Greek obstetric and gynaecology public units”. Health Policy. 2007.

KONTODIMOPOULOS, N, NIAKAS, D. “Efficiency measurement of haemodialysis units in Greece with data envelopment analysis” Health Policy 71. 2005. p. 195-204.

Mean square power series solutions of random Hermite differential equations *

G. Calbo, J.-C. Cortés†, L. Jódar
Instituto Universitario de Matemática Multidisciplinar
Universidad Politécnica de Valencia
Edificio 8G, 2ª, P.O. Box 22012, Valencia, Spain
{gcalbo, jccortes, ljodar}@imm.upv.es

December 10, 2009

Complexity, uncertainty and ignorance are present in most of real problems, not only due to natural phenomena, such as earthquakes, but also to human behavior such as greed or panic behavior of investors in financial markets. These facts motivate an increasing interest in the consideration of randomness in the mathematical models. Individual behavior may be erratic, but aggregate behavior is often predictable. Differential equations are powerful tools for representing reality up to certain point. The quantification of uncertainty requires a model specifying the mechanism by which randomness is generated. Random differential equations have been used in the last few decades to deal with errors and uncertainty. For example, see [1] for the general randomness case and [2, 3] for the case of white noise uncertainty. Theoretical approaches of random differential equations probably started with [1, 4]. However, thinking of applications, it is crucial not only to guarantee existence and uniqueness of the solution but also to be enable to compute approximations of the probability density function or the main statistical properties of the solution process such as mean and variance functions. Recent papers in this address are [5, 6].

In this talk, we develop a power series method to construct mean square convergent series solutions of random Hermite differential equation

$$\ddot{X}(t) - 2t\dot{X}(t) + AX(t) = 0, \quad -\infty < t < +\infty, \tag{0.1}$$

where the coefficient A is a random variable satisfying certain conditions to be specified later. An important difficulty to be overcame is the lack of sub-

*This work has been partially supported by the Spanish M.C.Y.T. grant MTM2009-08587 and Universidad Politécnica de Valencia grant PAID06-09-2588

†Corresponding author

multiplicativity of the mean square norm and the necessity of bounding products of random variables. Particularly interesting is the case where A is a discrete random variable taking even integer values with certain probabilities. In such case, it appears mean square series solutions of problem (0.1), called random Hermite polynomial.

The main concepts, definitions and results that will be referred throughout this talk may be found in [7, chap.4], [8, part IV], [9, chap.1-3]. Let $(\Omega, \mathcal{F}, P)$ be a probability space, we will consider the set L_2 which elements are second order real random variables (2-r.v.'s), i.e., $X : \Omega \to \mathbb{R}$ such that $\mathrm{E}[X^2] < +\infty$, where $\mathrm{E}[\cdot]$ denotes the expectation operator. The set L_2 endowed with the so-called 2-norm

$$\|X\| = \left(\mathrm{E}\left[X^2\right]\right)^{1/2}, \tag{0.2}$$

has a Banach space structure. We say that a sequence of 2-r.v.'s $\{X_n : n \geq 0\}$ is mean square (m.s.) convergent to $X \in L_2$ if

$$\lim_{n\to\infty} \|X_n - X\| = \lim_{n\to\infty} \left(\mathrm{E}\left[(X_n - X)^2\right]\right)^{1/2} = 0.$$

For the coefficient A appearing into the random differential equation (0.1) we will assume that there exists a nonnegative number p such that

$$\left\|A^{n+1}\right\| = O\left(n^p\right) \|A^n\| \,. \tag{0.3}$$

By considering $O(\cdot)$ definition (see, [10, p.335]), condition (0.3) entails that there is a constant $M > 0$ and a positive integer n_0 such that

$$\|A^n\| \leq C\, M^{n-1} \left((n-1)!\right)^p, \quad p \geq 0\,, \quad \forall n > n_0\,, \tag{0.4}$$

where C and M are positive constants. Note that by considering the 2-norm expression given by (0.2), this condition can be written as follows

$$\mathrm{E}\left[A^{2n}\right] \leq D\, M^{2(n-1)} \left((n-1)!\right)^{2p}, \quad p \geq 0\,, \quad \forall n > n_0\,, \tag{0.5}$$

where $D > 0$. The set of this type of r.v.'s is not empty. In fact, several relevant families of r.v.'s used extensively in probability, such that, uniform, exponential, beta, gamma or gaussian, belong to this class. For instance,

Example 0.1 *Let A be a beta r.v. of parameters $\alpha > 0$ and $\beta > 0$, i.e., $A \sim Be(\alpha; \beta)$. It is well-known that its statistical moments with respect to the origin are given by*

$$E[A^m] = \frac{\alpha(\alpha+1)\cdots(\alpha+m-1)}{(\alpha+\beta)(\alpha+\beta+1)\cdots(\alpha+\beta+m-1)} \quad , \quad m = 1, 2, \ldots,$$

then

$$\frac{\|A^{m+1}\|}{\|A^m\|} = \left(\frac{E[A^{2m+2}]}{E[A^{2m}]}\right)^{1/2} = \sqrt{\frac{(\alpha+2m)(\alpha+2m+1)}{(\alpha+\beta+2m)(\alpha+\beta+2m+1)}} = O(1),$$

that is, r.v. $A \sim Be(\alpha;\beta)$ *satisfies condition (0.3) for* $p = 0$. *Since uniform r.v.'s on interval* $[0,1]$ *are beta r.v.'s with parameters* $\alpha = 1 = \beta$, *then these basic but important type of r.v.'s are also included in this case. As a consequence, a uniform r.v. defined on an arbitrary interval, say* $[a,b]$, *belongs to this class too.*

Example 0.2 *Let* A *be a gamma r.v. of parameters* $r > 0$ *and* $a > 0$, *i.e.,* $A \sim Ga(r;a)$. *In this case its statistical moments with respect to the origin are given by*

$$E[A^m] = \frac{r(r+1)\cdots(r+m-1)}{a^m} \quad , \quad m = 1,2,\ldots,$$

then

$$\frac{\|A^{m+1}\|}{\|A^m\|} = \sqrt{\frac{(r+2m)(r+2m+1)}{a^2}} = O(m) \ ,$$

that is, $A \sim Ga(r;a)$ *satisfies condition (0.3) for* $p = 1$. *Note that exponential r.v.'s of parameter* $\lambda > 0$ *are also included in this case because they can be considered as gamma r.v.'s of parameters* $r = 1$ *and* $a = \lambda$.

Next we extent from the deterministic framework to the random one the concept of fundamental set of solutions. Let A_1 and A_2 be r.v.'s, and let $X_1(t)$ and $X_2(t)$ be two solutions of the second-order random differential equation

$$\ddot{X}(t) + A_1\dot{X}(t) + A_2X(t) = 0, \quad -\infty < t < +\infty. \tag{0.6}$$

We say that $\{X_1(t), X_2(t)\}$ is a fundamental set of solution processes of (0.6) in $-\infty < t < +\infty$, if any solution $X(t)$ of (0.6) admits a unique representation of the form

$$X(t) = C_1X_1(t) + C_2X_2(t), \quad t \in (-\infty,+\infty), \tag{0.7}$$

where C_1 and C_2 are r.v.'s uniquely determined by $X(t)$. Sufficient conditions for a pair of solutions $\mathcal{S} = \{X_1(t), X_2(t)\}$ of (0.6) defines a fundamental set in $-\infty < t < +\infty$ in terms of the so-called wronskian determinant process of $\mathcal{S}$, $W_{\mathcal{S}}(t) = X_1(t)\dot{X}_2(t) - X_2(t)\dot{X}_1(t)$ are given by

$$\text{there exists } t_0 \in (-\infty,+\infty) \text{ such that } W_{\mathcal{S}}(t_0) \neq 0.$$

In this talk, we are interested in determining solutions of differential equation (0.1) by means random power series which under certain conditions become random polynomials. Given a collection of r.v.'s $\{X_k : k \geq 0\}$ and $t \in \mathcal{T}$, $\sum_{k\geq 0} X_k t^k$ is called a random power series. If $P[X_k = 0] = 1$ for all $k > m$, that is,

$P\left[\{\omega \in \Omega :\ X_k(\omega) = 0,\ \forall k > m\}\right] = 1$, then $\sum_{k=0}^{m} X_k t^k$ is said to be a random polynomial in t of degree m.

An important fact is that the 2-norm $\|\cdot\|$ given by (0.2) does not provide a Banach algebra structure to L_2, i.e., it is not submultiplicative because the property $\|XY\| \leq \|X\|\,\|Y\|$ does not hold. In fact, let Z be a non-constant positive 2-r.v. and let us take $X = Y = Z^{1/2}$, then

$$\|XY\|^2 - \|X\|^2\,\|Y\|^2 = \mathrm{E}[Z^2] - (\mathrm{E}[Z])^2 = \mathrm{Var}[Z] > 0.$$

Therefore $\|XY\| > \|X\|\|Y\|$. This situation difficult the next target. Indeed, once a random power series solution of (0.1) has been constructed the justification of its mean square convergence is required. The general term of such a series is given by means the product of certain r.v.'s. In order to establish the absolute m.s. convergence of thaat series, we will require to bound the 2-norm of a product of r.v.'s in terms of the 2-norm of each factor. Next result provide some results to overcome this difficulty and it constitutes a generalization of Schwarz inequality.

Proposition 0.3 *Let $\{X_i\}_{i=1}^n$, $n \geq 2$ be r.v.'s such that $E\left[(X_i)^{2^n}\right] < +\infty$, $i = 1, 2, \ldots, n$, then*

$$E\left[\left|\prod_{i=1}^{n} X_i\right|\right] \leq \left(\prod_{i=1}^{n} E\left[(X_i)^{2^{n-1}}\right]\right)^{1/2^{n-1}}, \quad n \geq 2. \tag{0.8}$$

Inequality (0.8) can be expressed in terms of the 2-norm as follows by assuming that $\mathrm{E}\left[(Y_i)^{2^{n+1}}\right] < +\infty$, $i = 1, 2, \ldots, n$:

$$\left\|\prod_{i=1}^{n} Y_i\right\| \leq \prod_{i=1}^{n} \left\|(Y_i)^{2^{n-1}}\right\|^{1/2^{n-1}}, \quad \mathrm{E}\left[(Y_i)^{2^{n+1}}\right] < +\infty,\ i = 1, \ldots, n. \tag{0.9}$$

In advance, let us seek a formal solution process of problem (0.1) of the form

$$X(t) = \sum_{n\geq 0} X_n t^n = \sum_{n\geq 1} X_n t^n, \tag{0.10}$$

where coefficients X_n are 2-r.v.'s. By applying a Frobenius method one gets

$$X(t) = X_0 X_1(t) + X_1 X_2(t), \tag{0.11}$$

where

$$\begin{aligned} X_1(t) &= \left(1 + \sum_{k\geq 0} \frac{t^{2k+2}}{(2k+2)!} \prod_{j=0}^{k} (4j - A)\right), \\ X_2(t) &= \left(t + \sum_{k\geq 0} \frac{t^{2k+3}}{(2k+3)!} \prod_{j=0}^{k} (4j + 2 - A)\right). \end{aligned} \tag{0.12}$$

Next work is addressed to establish the m.s. convergence of the two random series given by (0.12), and since $(L_2, \|\cdot\|)$ is a Banach space, that is equivalent to prove that both series are absolutely convergent in the 2-norm. Thus, for each $t \in (-\infty, +\infty)$, we consider the numerical series associated to the first series in (0.12) given by

$$\sum_{k\geq 0} \frac{|t|^{2k+2}}{(2k+2)!} \left\| \prod_{j=0}^{k} (4j - A) \right\|, \tag{0.13}$$

and note that by hypothesis $\mathrm{E}\left[|A|^n\right] < +\infty$, $\forall n \geq 0$, then applying (0.9) one gets

$$\left\| \prod_{j=0}^{k} (4j - A) \right\| \leq \prod_{j=0}^{k} \left\| (4j - A)^{2^k} \right\|^{\frac{1}{2^k}}. \tag{0.14}$$

We can bound each factor of the above right-hand side for each $j = 0, 1, \ldots, k$ as follows

$$\begin{aligned} \left\| (4j - A)^{2^k} \right\|^{\frac{1}{2^k}} &= \left(\mathrm{E}\left[|4j - A|^{2^{k+1}} \right] \right)^{\frac{1}{2^{k+1}}} \\ &\leq \left(2^{2^{k+1}-1} \left((4j)^{2^{k+1}} + \mathrm{E}\left[|A|^{2^{k+1}} \right] \right) \right)^{\frac{1}{2^{k+1}}}. \end{aligned} \tag{0.15}$$

On the other hand, since we are assuming that r.v. A satisfies condition (0.4) (and therefore (0.5)), one gets

$$\mathrm{E}\left[|A|^{2^{k+1}} \right] = \mathrm{E}\left[|A|^{2\times 2^k} \right] \leq D\, M^{2(2^k-1)} \left((2^k - 1)! \right)^{2p}, \quad p \geq 0, \quad \forall k \geq k_0,$$

with D and M are positive constants, that is

$$\mathrm{E}\left[|A|^{2^{k+1}} \right] = \mathrm{E}\left[|A|^{2\times 2^k} \right] \leq L\, M^{2^{k+1}} \left((2^k - 1)! \right)^{2p}, \quad p \geq 0, \quad \forall k \geq k_0, \tag{0.16}$$

where $L > 0$. As we are interested in establishing the convergence of series (0.13), in the following we can assume without loss of generality that $k_0 = 0$. Thus considering (0.16) into (0.15) one gets for each $j = 0, 1, \ldots, k$

$$\begin{aligned} \left\| (4j - A)^{2^k} \right\|^{\frac{1}{2^k}} &\leq \left(2^{2^{k+1}-1} \left((4j)^{2^{k+1}} + L\, M^{2^{k+1}} \left((2^k - 1)! \right)^{2p} \right) \right)^{\frac{1}{2^{k+1}}} \\ &\leq \left(2^{2^{k+1}-1} \left((4k)^{2^{k+1}} + L\, M^{2^{k+1}} \left((2^k - 1)! \right)^{2p} \right) \right)^{\frac{1}{2^{k+1}}}. \end{aligned} \tag{0.17}$$

Now we bound each factor appearing in the right-hand side of (0.14) by means (0.17). In this way the positive numerical series (0.13) is majorized by the series

$$\sum_{k\geq 0} \frac{\left(2^{2^{k+1}-1} \left((4k)^{2^{k+1}} + L\, M^{2^{k+1}} \left((2^k - 1)! \right)^{2p} \right) \right)^{\frac{k+1}{2^{k+1}}}}{(2k+2)!} |t|^{2k+2}. \tag{0.18}$$

By using standard techniques it can be established that numerical series given by (0.13) is convergent, thus the first random series of (0.12) is m.s. convergent. Following an analogous procedure, it is easy to establish the m.s. convergence of the second series in (0.12). Then both solution series $X_1(t)$ and $X_2(t)$ given by (0.12) are m.s. uniformly convergent, therefore the formal differentiation considered in the application of Frobenius method is justified. On the other hand, taking $t_0 = 0$ and considering that $X_1(0) = 1$, $\dot{X}_1(0) = 0$, $X_2(0) = 0$ and $\dot{X}_2(0) = 1$, one gets that $W_{\mathcal{S}}(0) = 1 \neq 0$, then the solution of random differential equation (0.1) with random initial conditions $X(0) = Y_0$ and $\dot{X}(0) = Y_1$ is given by

$$X(t) = Y_0 X_1(t) + Y_1 X_2(t), \quad t \in (-\infty, +\infty)\,, \tag{0.19}$$

where $X_1(t)$ and $X_2(t)$ are defined in (0.12). Summarizing the following result has been established

Theorem 0.4 *The random differential equation (0.1) with initial conditions $X(0) = Y_0$ and $\dot{X}(0) = Y_1$, where A is a continuous r.v. satisfying condition (0.3), admits as random power series solution of the form (0.19) where $X_1(t)$ and $X_2(t)$ are given by (0.12). Moreover the solution is m.s. convergent for each $t \in (-\infty, +\infty)$.*

The case where A is a discrete r.v. deserves a special treatment because, in general, condition (0.5) is difficult to check due to the lack of explicit expressions for the moments of relevant discrete r.v.'s. In order to illustrate how we can handle this situation, let us assume that r.v. A follows a binomial distribution, i.e., $A \sim \mathrm{Bi}(n; p)$, then A takes its values on the finite set $A = 0, 1, 2, \ldots, n$, for a fixed positive integer n and $p \in (0, 1)$. Then,

$$\mathrm{E}\left[|A|^{2^{k+1}}\right] = \sum_{a=0}^{n} a^{2^{k+1}} p_A(a) \leq (n+1) n^{2^{k+1}} \sum_{a=0}^{n} p_A(a) = (n+1) n^{2^{k+1}}.$$

Now we proceed in the same way as in the previous exposition: we bound each factor appearing in the right-hand side of (0.14) and one obtains that the positive numerical series (0.13) is majorized by

$$\sum_{k \geq 0} \frac{\left(2^{2^{k+1}-1}\left((4k)^{2^{k+1}} + (n+1)n^{2^{k+1}}\right)\right)^{\frac{k+1}{2^{k+1}}}}{(2k+2)!}\, |t|^{2k+2}\,, \tag{0.20}$$

which is convergent. An analogous treatment can be made for the case that A is a Poisson r.v., and the following result can be established:

Corollary 0.5 *The theorem 0.4 holds true if A is a discrete r.v. belonging to some of the following classes of r.v.'s having finite moments:*

- *A is a discrete r.v. taking a finite number of values.*

- *A is a Poisson r.v.*

From (0.12), one deduces that if there exists $n \geq 0$ such that $P[A = 2n] = 1$, then (0.1) has a random polynomial solution: if there exists $k \geq 0$ such that $P[A = 4k] = 1$ (or $P[A = 4k + 2] = 1$), then $X_1(t)$ (or $X_2(t)$) given by (0.12) generates a (random) polynomial solution of degree $2k$ (or $2k+1$). These random (degenerate) solutions can be interpreted as corresponding Hermite polynomials that one presents in the deterministic framework. However in the random scenario there are richer situations that deserves to be considered. Indeed, in the case that A is a continuous r.v., like beta, gamma or gaussian r.v.'s, since $P[A = 2n] = 0$ for every integer $n \geq 0$, then with probability 1 one can conclude that there are not random polynomial solutions of (0.1). Whereas if A is a discrete r.v. that only takes different even values (not concentrated in just one even value), then there will exist with probability 1, random polynomial solutions. This case generalizes the concept of Hermite polynomial solution from the deterministic framework.

For the case that A is a discrete r.v. whose values lie in a set containing even numbers (like binomial or Poisson r.v.'s), the random differential equation does not have random polynomial solutions but it admits some sample representations which are (deterministic) polynomials. Then, considering the solution stochastic process as a family of trajectories, we can assign the probability p_{pol} of the random power series given by (0.19) and (0.12) has random polynomial sample solutions. Note that this series becomes a random polynomial if and only if $p_{pol} = 1$. For instance, if A is a Poisson r.v., i.e., $A \sim \text{Po}(\lambda)$, then the probability of having random polynomial solutions is given by

$$p_{pol} = P[A = 2n] = \exp(-\lambda) \sum_{n \geq 0} \frac{\lambda^{2n}}{(2n)!} = \exp(-\lambda) \cosh(\lambda).$$

Table 1 shows these probabilities for different values of the parameter $\lambda > 0$. One observes that these values decrease from 1 to 0.5 as λ increases from 0 to $+\infty$. The above considerations motivate the following result:

λ	0.01	0.1	0.5	1	2	5	10000
p_{pol}	0.990099	0.909365	0.683940	0.567668	0.509158	0.500023	0.5

Table 1: Probabilities of generating random polynomial (sample) solutions when A is a Poisson r.v. of parameter $\lambda > 0$

Corollary 0.6 *Let us consider the random differential equation (0.1) with initial conditions $X(0) = Y_0$ and $\dot{X}(0) = Y_1$, where the (discrete) r.v. A takes only a finite number of even integer values, that is, $P[A = 2m_j] = p_j > 0$, $1 \leq j \leq n$ with $\sum_{j=1}^{n} p_j = 1$. Then this i.v.p. has a random polynomial solution $H_{m_n}(t)$ of the degree m_n.*

Since r.v. A can only take a finite number of values, then for obtaining random polynomial solutions, we do not not require the assumption $\mathrm{E}\left[|A|^n\right] < +\infty$ for each $n = 1, 2, \ldots$ because it is satisfied.

Definition 0.7 *Let m_n be a positive integer. The m_n-th random Hermite polynomial or the random Hermite polynomial of degree m_n, is the random m.s. solution of problem (0.1) with initial conditions $X(0) = Y_0$ and $\dot{X}(0) = Y_1$, being A the discrete r.v. taking the finite number of even integer values $2m_j$, with probabilities $P\left[A = 2m_j\right] = p_j > 0$, $1 \leq j \leq n$ with $\displaystyle\sum_{j=1}^{n} p_j = 1$.*

Under conditions of corollary 0.6 a random Hermite polynomial solution can be interpreted as a collection of deterministic Hermite polynomials, which, for each $j : 1 \leq j \leq n$, has a probability p_j of sampling. The degree m_n of the random Hermite polynomial $H_{m_n(t)}$ is the greatest of all degrees corresponding to each (deterministic) Hermite polynomials, but it is not necessary that its sample associated probability p_n be also greater than p_j for all $j : 1 \leq j \leq n$. The situation where A is the discrete r.v. taking all the even integer values with $P\left[A = 2j\right] = 2^{-(j+1)}$, $j = 0, 1, 2, \ldots$, then for each j one obtains a sample polynomial solution, but since j lies in the positive integer numbers, the degree of the random polynomial solution cannot be defined according with definition 0.7. We finish this talk by computing of the main statistical functions of the m.s. solution of (0.1) given by (0.11)-(0.12) such that mean and variance in terms of the data. For this goal, we will consider a truncation of order N of the solution:

$$\begin{aligned} X_N(t) &= X_0 \left(1 + \sum_{k=0}^{N} \frac{t^{2k+2}}{(2k+2)!} \prod_{j=0}^{k} (4j - A)\right) \\ &+ X_1 \left(t + \sum_{k=0}^{N} \frac{t^{2k+3}}{(2k+3)!} \prod_{j=0}^{k} (4j + 2 - A)\right). \end{aligned} \tag{0.21}$$

In the sequel, we will assume that r.v. A is independent of $X(0) = Y_0$ and $\dot{X}(0) = Y_1$. Therefore one gets

$$\begin{aligned} \mu_{X_N}(t) &= \mathrm{E}\left[Y_0\right] \left(1 + \sum_{k=0}^{N} \frac{t^{2k+2}}{(2k+2)!} \mathrm{E}\left[\prod_{j=0}^{k} (4j - A)\right]\right) \\ &+ \mathrm{E}\left[Y_1\right] \left(t + \sum_{k=0}^{N} \frac{t^{2k+3}}{(2k+3)!} \mathrm{E}\left[\prod_{j=0}^{k} (4j + 2 - A)\right]\right). \end{aligned} \tag{0.22}$$

On the other hand, it is easy to check that

$$\mathrm{E}\left[\left(X_N(t)\right)^2\right] = \sum_{n=0}^{N} \mathrm{E}\left[\left(X_n\right)^2\right] t^{2n} + 2 \sum_{n=1}^{N} \sum_{m=0}^{n-1} \mathrm{E}\left[X_n\, X_m\right] t^{n+m}.$$

where with the usual convection $\prod_{i=u}^{v} f(i) = 1$ if $v < u$ the terms involved in the two above sums can be computed as follows:

$$\mathrm{E}\left[X_n X_m\right] = \frac{1}{n!\,m!} \begin{cases} \mathrm{E}\left[(Y_0)^2\right] \mathrm{E}\left[P_1\left(\dfrac{n-2}{2}\right) P_1\left(\dfrac{m-2}{2}\right)\right] & \text{if} \quad \begin{array}{l} n=0,2,4,\ldots, \\ m=0,2,4,\ldots, \end{array} \\ \mathrm{E}\left[Y_0 Y_1\right] \mathrm{E}\left[P_1\left(\dfrac{n-2}{2}\right) P_2\left(\dfrac{m-3}{2}\right)\right] & \text{if} \quad \begin{array}{l} n=0,2,4,\ldots, \\ m=1,3,5,\ldots, \end{array} \\ \mathrm{E}\left[Y_0 Y_1\right] \mathrm{E}\left[P_2\left(\dfrac{m-3}{2}\right) P_1\left(\dfrac{n-2}{2}\right)\right] & \text{if} \quad \begin{array}{l} n=1,3,5,\ldots, \\ m=0,2,4,\ldots, \end{array} \\ \mathrm{E}\left[(Y_1)^2\right] \mathrm{E}\left[P_2\left(\dfrac{n-3}{2}\right) P_2\left(\dfrac{m-3}{2}\right)\right] & \text{if} \quad \begin{array}{l} n=1,3,5,\ldots, \\ m=1,3,5,\ldots, \end{array} \end{cases} \tag{0.23}$$

where we have denoted

$$P_1(k) = \prod_{j=0}^{k} (4j - A), \quad P_2(k) = \prod_{j=0}^{k} (4j + 2 - A),$$

An important feature is that because of the m.s. convergence of random series appearing into (0.21) when $N \to \infty$ then the mean and the variance of the truncated solution (0.21) to the corresponding exact values are warranted.

References

[1] J.L. Strand, Random ordinary differential equations, J. Diff. Equat. **7** (1973) 538–553.

[2] L. Arnold, Stochastic Differential Equations Theory and Applications, John Wiley, New York, 1974.

[3] B. Øksendal, Stochastic Differential Equations. An Introduction with Applications (5th Ed.), Springer-Verlag, Berlin Heildelberg 1998.

[4] J.L. Strand, Stochastic Ordinary Differential Equations, Ph.D. Thesis, Univ. of California, Berkeley, California, 1968.

[5] M.A. El-Tawil, The approximate solutions of some stochastic differential equations using transformations, Appl. Math. Comput., **164(1)** (2005) 167–178.

[6] J.C. Cortés, L. Jódar, L. Villafuerte, Random linear-quadratic mathematical models: computing explicit solutions and applications, *Math. Comput. Simulat.* **79 (7)** 2076–2090 (2009).

[7] T.T. Soong, *Random Differential Equations in Science and Engineering*, Academic Press, New York (1973).

[8] M. Loève, *Probability Theory*, Van Nostrand, Princeton, New Jersey (1963).

[9] E. Wong, B. Hajek, *Stochastic Processes in Engineering System*, Springer Verlag, New York (1985).

[10] S. Elaydi, *An Introduction to Difference Equations*, Third Ed., Springer, New York (2005).

[11] I.S. Gradshteyn, I.M. Ryzhik, *Table of Integrals, Series and Products*, Academic Press, England (1994).

[12] N. N. Lebedev, *Special Functions and their Applications*, Dover, New York (1972).

[13] G.R. Grimmett, D.R. Stirzaker, *Probability and Random Processes*, Claredon Press, Oxford (2000).

[14] J.C. Cortés, P. Sevilla-Peris, L. Jódar, Analytic-numerical approximating processes of diffusion equation with data uncertainty, *Comput. Math. Appl.* **49** 1255–1266 (2005).

[15] J.B. Keller, Stochastic equations and wave propagation in random media, Procc. Symp. Appl. Math. Am. Math. Soc. Providence Rhode Island 1963 **16** (1963) 145–170.

[16] D. Henderson, P. Plaschko, *Stochastic Differential Equations in Science and Engineering*, World Scientific, Singapore (2006).

[17] M. El-Tawil, W. El-Tahan, A. Hussein, A proposed technique of SFEM on solving ordinary random differential equation, *Appl. Math. Comput.* **161** 35–47 (2005).

[18] G. Calbo, J.C. Cortés, L. Jódar, Random analytic solution of coupled differential models with uncertain initial condition and source term, *Comput. Math. Appl.* **56** 785–798 (2008).

Identifiability of a structured parametric economic model. *

B. Cantó[†], C. Coll [‡] and E. Sánchez[§]

Instituto de Matemática Multidisciplinar,

Universidad Politécnica de Valencia,

Edificio 8G, Piso 2, 46022 Valencia, España.

December 10, 2009

1 Introduction

Process models are applied for design different control problems. In particular, modelling the tourism sector is useful since governments are interested, for example, in the level tourist expenditures at a country, hotels in tourism areas with higher demand and airlines in routes in greatest demand. The obtained information of the model could be used to analyze the level of economic activity, rental cars unused, etc., and the interconnection of the various services involved in the sector. Very often the equation of the model involves unknown parameters that must be estimated from experimental data. This problem is known as the identifiability problem. This consists on investigate whether unknown parameters in a given model structure can be uniquely recovered from experimental data. The identifiability framework helps us test the unique relationship between parameters sets and model response. It is possible to distinguish between the following types of identifiability: global identifiability (a model is global identifiable if and only if there is an unique

*Supported in part by Grant MMT2007-64477.

[†]bcanto@imm.upv.es

[‡]mccoll@imm.upv.es

[§]esanchezj@imm.upv.es

input-output behavior for every parameter set) and local identifiability (a model is local identifiable if the identifiability analysis problem yields multiple solutions). A formal definition is given in [1].

This paper addresses the structural parameter identifiability problem for singular dynamic system. We analyze this problem in the global identifiability sense. To facility the study of the model of a real dynamical process on the impact produced by changes of tourist's income on the demand for tourism services we need a Leontief model (see for instance [2] and [3]). The dynamic model is described by n tourism services interconnected among themselves such that the $i-$th tourism service depends on itself and the next $(i+1)-$th service. In this case, the technological coefficient matrix has a bidiagonal structure. This matrix is denoted by $P(\mathbf{p})$. By the nature of this matrix $0 \leq p_{ij}^p \leq 1$, and $\sum_{i=1}^{n} p_{ij}^p \leq 1$. Must be take into account that the demand only affects on the n-th tourism service. Then, the demand coefficient matrix $D(\mathbf{p})$ is such that has one column multiple of the n-th canonical vector. The (i,j)-entry of the capital matrix $C(\mathbf{p})$ represents the amounts of capital of the $i-$th good necessary for the production of one unit of the $j-$th tourism service. The singularity of this matrix arises because no output from one sector is used in the production of some products. Note that these matrices are nonnegative matrices and the Capital matrix can be or not be singular.

Let the dynamic model be represented by this generalized system:

$$C(\mathbf{p})x\,(k+1) = (I - P(\mathbf{p}) + C(\mathbf{p}))\,x\,(k) + D(\mathbf{p})u\,(k) \tag{1}$$

where $x(k)$ is the production level vector of the tourism services and $u(k)$ is the demand level vector. The coefficient matrices represent the different factors involved in the economic process. These matrices have a fixed structure with the parameter vector $\mathbf{p}$ belongs to a subset $\mathcal{P} \subseteq \mathbb{R}^r$. These parameters suggest the relation among variables of the economic process. If the capital matrix $C(\mathbf{p})$ is invertible, then we have a standard system $S(\mathbf{p}) = (A(\mathbf{p}), B(\mathbf{p}))$ where $A(\mathbf{p}) = I + C^{-1}(\mathbf{p})(I - P(\mathbf{p}))$ and $B(\mathbf{p}) = C^{-1}(\mathbf{p})D(\mathbf{p})$. This system belongs to a kind of structured systems. The identifiability of the parameters of this system is concerned with the determination of them from the external behavior of the system. This input-output behavior of $S(\mathbf{p})$ is given by the Markov parameters $V(j,\mathbf{p}) = A^j(\mathbf{p})B(\mathbf{p}),\ j \geq 0$.

A number of methods have been proposed in the literature to test identifiability of linear parametric systems. Although some methods for structural

identifiability analysis seems appealing, some of them are limited by the ability to solve an algebraic equation set with solutions is in general no feasible. In this work the Markov parameters are used because they are useful to construct an algorithm to test if the parameters can be determinated. The aim of this work is analyze the identifiability problem of the structured system (1), which it models the dynamic process that represents the tourism services.

2 Structured standard system

Consider a structured standard system $S(\mathbf{p}) = (A(\mathbf{p}), B(\mathbf{p}))$ given by

$$A(\mathbf{p}) = \begin{cases} p_{ij}^a & i = 1, \dots n-1,\ j = i, i+1 \quad \text{and} \quad i = n,\ j = n \\ 0 & i = 1, \dots n-1,\ j \neq i, i+1 \quad \text{and} \quad i = n,\ j \neq n \end{cases}$$

and $B(\mathbf{p})$ is such that has almost one monomial column of kind $b(\mathbf{p}) = (0\ 0\ \cdots\ 0\ p_n^b)^T$ with all parameters not equal to zero. We denote this set of parameter vectors $\mathbf{p}$ by $\mathcal{P}$. In this case, by the structure of these matrices and by a technical process we can prove that the Markov parameters, $v(k) = \left(v_i^{(k)}(\mathbf{p})\right)_{i=1,\dots,n}$, $k = 0, \dots, n$, under the condition $v_{n+1}^{(k-1)}(\mathbf{p}) = 0$, satisfy the following relations,

$$\begin{aligned} v_n^{(0)}(\mathbf{p}) &= p_n^b \qquad & v_i^{(k)}(\mathbf{p}) &= 0,\ i = 1, \dots, n-k-1, \\ v_i^{(k)}(\mathbf{p}) &= \sum_{j=0}^{1} p_{i,i+j}^a v_{i+j}^{(k-1)}(\mathbf{p}), \qquad & i &= n-k, \dots, n. \end{aligned}$$

Hence, the structured system $S(\mathbf{p})$ is globally identifiable.

3 Testing the global identifiability of a structured economic model

Since the Leontief economic model described in the first section is a generalized system, solve the identification problem in this case is not easily determined in general. For that, now consider the economic model (1) showed in the introduction with a particular structure. Suppose that the $i-$th good produce the $i-$th tourism service. Then the Capital matrix is a non-singular diagonal matrix, that is $C = \text{diag}(c_1, c_2, \dots, c_n)$. The Technological matrix

represents the amount of material required per unit of service. In this case, we consider that the Technological matrix $P(\mathbf{p}) = (p^p_{ij})$ satisfies that for each $i = 1, \ldots, n$ the entries p^p_{ii} and p^p_{ii+1} are positive and the entries p^p_{ij} are zero for $j \neq i, i+1$, that is $P(\mathbf{p})$ is a bidiagonal matrix. Finally a Demand matrix $D(\mathbf{p})$, that represents the demand of end consumers (customers), is a monomial column. Note that, the parameter vector $\mathbf{p}$, comprising the unknown parameters, is given by $\mathbf{p} = p^b_n,\ p^a_{i,i+j},\ j = 0, 1,\ i = n-k, \ldots, n$.

In this section parameter estimation is performed using the standard system associated to the Leontief model. By the structure of this standard model $S(\mathbf{p})$, we can check that this system is globally identifiable. It is seen that there is good correspondence between the matrices of the system $S(\mathbf{p})$ and the initial parameters of the vector $\mathbf{p}$. The analysis remains valid for the parameter vector and the initial Leontief model is structurally globally identifiable. Since it is known the response of the economic process, that is, it is known the collection of matrices $\{V(k), k = 0, \ldots, n\}$, the results obtained permit to build up an algorithm for identification problem. That is, a condition for testing the identifiability for the discrete-time generalized system is formulated as follows.

Algorithm

In this subsection we give an algorithm to solve the identifiability problem. The algorithm is composed in two phases where each phase encompasses several steps. The first phase concerns the construction the Markov parameters of the associated standard system. The second phase investigates the parameter identification of the economical model. Step 1 of the algorithm concerns model formulation of the system. Step 2, 3 and 4 concerns determination of Markov parameters using the relationship analyzed above. In the step 5 the algorithm obtain the parameters of the structured standard system. Finally, steps 6 and 7 give the formulation of the structured economic system and the values of the initial parameters. The steps are detailed in the following.

Step 1. Introduce the size of the state vector: n and the capital matrix C. Introduce the matrices $\{V(k), k = 0, \ldots, n\}$ that determine the known external behavior of the process. And introduce the position of the monomial column of matrix $V(0)$: j .

For $k = 0, \ldots, n$

Step 2. Choose and denote the j-th column of $V(k)$ as $v^{(k)}$. Put $v^{(k-1)}_{n+1} = 0$.

Step 3. Construct the following system: $v_n^{(0)} = p_n^b$, $v_i^{(k)} = \sum_{j=0}^{1} p_{i,i+j}^a v_{i+j}^{(k-1)}$, $i = n-k, \ldots, n$.

Step 4. Solving the above system, we obtain the parameters p_n^b and $p_{i,i+j}^a$, $j = 0, 1$, and $i = n-k, \ldots, n$.

Step 5. Using parameters obtained in *Step 4* we construct the matrices of the associated standard Leontief model $S(\mathbf{p}) = (A(\mathbf{p}), B(\mathbf{p}))$ and from them we obtain p_n^d and $p_{i,i+j}^p$, $j = 0, 1$, and $i = 1, \ldots, n$.

With this process we have determinate all parameters of the system.

Step 6. Finally, construct the matrices of the system (1).

4 Conclusions

This paper considers the identifiability of Leontief dynamic model. Convenient representation of the associated system is sought in order to determine identifiability of the original model. The problem has been addressed by proving that the model is globally identifiable. The procedure that has been introduced relates to a more constructive means for the calculations of identifiable parameter. The first result shows that a standard structured system associated to the Leontief economic model is globally identifiable. In this case, it would appear that the Markov parameters plays an important role in the identification problem. The second result shows that, the origin economic model is also globally identifiable. Finally, an algorithm is performed to test the identifiability problem and for obtaining the unknown parameters in a economic process.

References

[1] A. Ben-Zvi, P.J. McLellan and K.B. McAuley, Identifiability of linear time-invariant differential-algebraic systems. I. The generalized Markov parameter approach. *Ind. Eng. Chem. Res.* **42** 6607-6618 (2003).

[2] B. Cantó, C. Coll and E. Sánchez, Positive N-periodic descriptor control systems. *Systems and Control Letters.* **53** 407-414 (2004).

[3] M.S. Silva and T. Lima, Looking for nonnegative solutions of a Leontief dynamic model. *Linear Algebra and its Applications* **364** 281-316 (2003).

An Algorithm to obtain a minimal realization of a polynomial transfer matrix. *

R. Cantó[†], B. Ricarte[‡], A. M. Urbano[§]

Instituto de Matemática Multidisciplinar,

Universidad Politécnica de Valencia,

Edificio 8G, Piso 2, 46022 Valencia, España.

December 10, 2009

1 Introduction

In this paper we study the realization problem in polynomial transfer matrices, and consider some applications on electrical circuits. More properly, on the state variable models of the corresponding regular singular systems for circuits [1], so that, by applying Kirchoff's laws, systems with matrices of big size and quite sparce are obtained. But of course some of the corresponding equations of this system will be redundant. The problem is that the determination of a minimal set of equations, and thus a minimum number of variables, is not computationally trivial. We introduce an efficient and stable computational algorithm to obtain this minimal realization.

*Supported by the Spanish DGI grant DGI MTM2007-64477 and the UPV under its research program.

[†]rcanto@mat.upv.es

[‡]bearibe@mat.upv.es

[§]amurbano@mat.upv.es

2 Definitions and Preliminary results

Consider the continuous time-invariant singular system

$$\begin{cases} \bar{E}\dot{\bar{x}}(t) &= \bar{A}\bar{x}(t) + \bar{B}\bar{u}(t) \\ \bar{y}(t) &= \bar{C}\bar{x}(t) \end{cases} \tag{1}$$

where $\dot{\bar{x}}(t) = \frac{d\bar{x}(t)}{dt}$, t is the time, $\bar{x}(t) \in \mathbb{R}^n$ is the vector of internal variables, $\bar{u}(t) \in \mathbb{R}^p$ is the vector of control inputs and $\bar{y}(t) \in \mathbb{R}^m$ is the vector of outputs. If $\bar{E}$ is a singular matrix the system is called *singular*.

Given a rational matrix $G(s) \in \mathbb{R}^{m\times p}[s]$, matrices $\bar{E}$, $\bar{A} \in \mathbb{R}^{n\times n}$, $\bar{B} \in \mathbb{R}^{n\times p}$ and $\bar{C} \in \mathbb{R}^{m\times n}$ such that $G(s) = \bar{C}\left[s\bar{E} - \bar{A}\right]^{-1}\bar{B}$ are called a *realization* of $G(s)$. It is denoted by $(\bar{E}, \bar{A}, \bar{B}, \bar{C})$. The size of $\bar{A}$ is called the *dimension* of the realization. The realization is *minimal* if it has minimum dimension.

If this system satisfies the regularity condition, i.e. there exists an scalar $\lambda \in \mathbb{C}$ such that $\det[\lambda\bar{E} - \bar{A}] \neq 0$ then it is equivalent to the canonical forward-backward form (see [4]) where $E = \mathrm{diag}(I_{n_1}, N)$, $A = \mathrm{diag}(A_1, I_{n_2})$, $n_1 + n_2 = n$, with n_1 the degree of polynomial $\det[sE - A]$, $A_1 \in \mathbb{R}^{n_1\times n_1}$ and $N \in \mathbb{R}^{n_2\times n_2}$ a nilpotent matrix of index t, $B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$ with $B_1 \in \mathbb{R}^{n_1\times p}$ and $B_2 \in \mathbb{R}^{n_2\times p}$, and $C = [C_1 \;\; C_2]$ with $C_1 \in \mathbb{R}^{m\times n_1}$ and $C_2 \in \mathbb{R}^{m\times n_2}$.

Lemma 1 *[5, Lemma 4.2] The transfer matrix of a canonical forward-backward system is $G(s) = G_1(s) + G_2(s)$ where $G_1(s) = C_1[sI_{n_1} - A_1]^{-1}B_1$ is the strictly proper transfer matrix of a standard (forward) subsystem and $G_2(s) = C_2(sN - I_{n_2})^{-1}B_2$ is the polynomial transfer matrix of a complete singular (backward) subsystem.*

Consequently, we deduce that the realization problem in regular singular systems can be dealt as two realization subproblems, a realization problem in a standard system and a realization problem of a complete singular system. The computation of a minimal realization (N, I, B_2, C_2) of $G_2(s)$ is studied in this paper.

3 Algorithm to obtain a minimal realization

Consider a polynomial transfer matrix

$$\begin{aligned} G(s) &= C[sN - I]^{-1}B = -CB - sCNB - \ldots - s^{t-1}CN^{t-1}B = \\ &= W_0 + sW_1 + \ldots + s^{t-1}W_{t-1} \end{aligned}$$

where t is the nilpotence index of N and $W_i \in \mathbb{R}^{m\times p}$, $i = 0, 1, \ldots, t-1$. This algorithm is an interesting improvement of the realization algorithm presented by Silverman and Ho (see [4, pp. 63]).

We can suppose, without loss of generality, that $m \geq p$. In other case, we work with the polynomial transfer matrix $G^T(s)$ so that if (N, I, B, C) is a minimal realization of $G(s)$, then (N^T, I, C^T, B^T) is a minimal realization of $G^T(s)$.

From now on, for a matrix $A \in \mathbb{R}^{n\times m}$, we denote by $A(i_1 : i_2, j_1 : j_2)$ the submatrix of A with rows $\{i_1, i_1 + 1, \ldots, i_2\}$ and columns $\{j_1, j_1 + 1, \ldots, j_2\}$. If the submatrix has all rows (resp. columns) of A, then it is denoted by $A(:, j_1 : j_2)$ (resp. $A(i_1 : i_2, :)$).

Definition 1 *A matrix* $U \in \mathbb{R}^{n\times m}$ *with* $\operatorname{rank}(U) = n$ *is* a basic upper matrix *if it is an upper reduced echelon matrix and the leading* 1*'s are in the* n *first columns.*

Definition 2 *A block matrix* $U = [U_1\ U_2\ \ldots\ U_t]$ *is* a basic upper block matrix *if it is an upper reduced echelon matrix and the leading* 1*'s in each block are in the first columns. That is,* U *has the following structure*

$$U = \left[\begin{array}{cc|cc|c|cc} I_{r_1} & U^{(1,1)} & O & U^{(2,1)} & \cdots & O & U^{(t,1)} \\ O & O & I_{r_2} & U^{(2,2)} & \cdots & O & U^{(t,2)} \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ O & O & O & O & \cdots & I_{r_t} & U^{(t,t)} \end{array}\right].$$

Algorithm 1
Step 1. *Define the matrix*

$$W = \begin{bmatrix} -W_{t-1} & -W_{t-2} & \cdots & -W_1 & -W_0 \\ O & -W_{t-1} & \cdots & -W_2 & -W_1 \\ \vdots & \vdots & & \vdots & \vdots \\ O & O & \cdots & -W_{t-1} & -W_{t-2} \\ O & O & \cdots & O & -W_{t-1} \end{bmatrix}$$

Step 2. *Obtain, applying the quasi-Gauss method [3], the following full rank factorization of matrix W*

$$W = G * U = \begin{bmatrix} G_1 \\ G_2 \\ \vdots \\ G_{t-1} \\ G_t \end{bmatrix} [U_1\, U_2\, \cdots\, U_{t-1}\, U_t]$$

where $G_i \in \mathbb{R}^{m\times n}$ and $U_i \in \mathbb{R}^{n\times p}$ for $i = 1, 2, \ldots, t$, with $n = \operatorname{rank}(W)$ and U is a basic upper block matrix.

Step 3. *Define*

$$\begin{aligned} B &= U(:, (t-1)p+1 : tp), \quad C = G(1:m, :) \\ q_{t-1} &= \operatorname{rank}(U(:, 1:p)) \\ q_{t-i} &= \operatorname{rank}(U(:, 1:ip)) - \operatorname{rank}(U(:, 1:(i-1)p)) \quad \text{for } i = 2, 3, \ldots, t. \\ N &= [O_{n\times q_{t-1}}\ U(:, 1:q_{t-2})\ U(:, p+1:p+q_{t-3})\ \ldots \\ &\ldots\ U(:, (t-2)p+1:(t-2)p+q_0)] \end{aligned}$$

Proposition 1 *(N, I, B, C) given by Algorithm 1 is a minimal realization of the polynomial transfer matrix $G(s) = W_0 + sW_1 + \ldots + s^{t-1}W_{t-1}$ where $W_i \in \mathbb{R}^{m\times p}$, $i = 0, 1, \ldots, t-1$.*

Remark 1

1. *Note that q_{t-i}, $i = 1, 2, \ldots, t$, are directly obtained from the number of leading 1's of the corresponding blocks of U. Hence, G is given from the first n linearly independent columns of W. Observe that we only need to save the column indices of U with leading 1's.*

2. *By construction N is nilpotent, with nilpotent index equal to t.*

3. *The computational cost of the Algorithm 1 is based on the cost of the quasi-Gauss method because B, C and N are specific rows and columns of G and U.*

4. *The Silverman and Ho algorithm [4, pp. 63] starts with a full rank factorization L_1L_2 with $L_1 \in \mathbb{R}^{tm\times n}$ and $L_2 \in \mathbb{R}^{n\times tp}$, and then computes $N = (L_1^TL_1)^{-1}L_1^TM_1L_2^T(L_2L_2^T)^{-1}$ with $M_1 \in \mathbb{R}^{tm\times tp}$. The computational cost of N is more expensive because transpose, inverse and products of matrices are included.*

Remark 2 *If U is not a basic upper block matrix, in order to apply Algorithm 1, it is necessary to permute some rows and columns of U because the linearly independent columns in each block are not in the first columns.*

4 Example

In [1, section 3] some principal applications in terms of electrical circuits, or more properly, in terms of state variable models for circuits are introduced. Concretely, the continuous time-invariant singular system of a linear RLC circuit is given by (1) where $\bar{E} = (\bar{e}_{ij}) \in \mathbb{R}^{11\times 11}$ is the zero matrix except for $\bar{e}_{22} = \bar{e}_{99} = 1$; $\bar{B} = (\bar{b}_{ij}) \in \mathbb{R}^{11\times 2}$ is the zero matrix except for $\bar{b}_{11} = \bar{b}_{10,2} = 1$; $\bar{C} = I_{11}$ is the identity matrix and

$$\bar{A} = \left[\begin{array}{rrrrrrrrrrr} -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ \hline 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -3 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ \hline 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & -3 & -1 & -1 & -1 & 0 \end{array}\right] \in \mathbb{R}^{11\times 11}.$$

The corresponding transfer matrix is:

$$\begin{aligned} G(s) &= \bar{C}\left[s\bar{E} - \bar{A}\right]^{-1}\bar{B} = G_1(s) + G_2(s) = \\ &= \frac{1}{s}\left[\begin{array}{cccccccccccc} 0 & 0 & 0 & 0 & 0 & 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]^T + \\ &+ \left[\begin{array}{cccccccccccc} 1 & 3 & 1 & 1 & 1 & 6s+1 & -3s-1 & 3s+2 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & 1 & 0 & 1 & 0 \end{array}\right]^T. \end{aligned}$$

By [2] a minimal realization (A_1, B_1, C_1) of $G_1(s)$ is

$$A_1 = [0] \quad B_1 = [1\ 0] \quad C_1 = \left[\begin{array}{cccccccccccc} 0 & 0 & 0 & 0 & 0 & 1 & -1 & 1 & 1 & 0 & 0 \end{array}\right]^T.$$

By Algorithm 1, we compute a minimal realization (N, I, B_2, C_2) of $G_2(s)$. Let

$$\begin{aligned} G_2(s) &= W_1 s + W_0 = \\ &= s \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 6 & -3 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^T \\ &+ \begin{bmatrix} 1 & 3 & 1 & 1 & 1 & 1 & -1 & 2 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & 1 & 0 & 1 & 0 \end{bmatrix}^T . \end{aligned}$$

The matrix $W = \begin{bmatrix} -W_1 & -W_0 \\ O & -W_1 \end{bmatrix}$, has the following full rank factorization

$$W = G * U = [W(:,1)\ W(:,3:4)] \left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right]$$

where U is a basic upper block matrix. Therefore,

$$B_2 = U(:,3:4) = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$C_2 = G(1:11,:) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & -6 & 3 & -3 & 0 & 0 & 0 \\ -1 & -3 & -1 & -1 & -1 & -1 & 1 & -2 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 & -1 & 0 & -1 & 0 \end{bmatrix}^T$$

$$q_{t-1} = q_1 = 1,\ q_{t-2} = q_0 = 2,\ \text{then } N = [O_{3\times 1}\ U(:,1:2)] = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

From (A_1, B_1, C_1) of $G_1(s)$ and (N, I, B_2, C_2) of $G_2(s)$, a minimal realization (E, A, B, C) of $G(s)$ is obtained. Note that we have begun with a space-state representation of order 11 and we have obtained a minimal realization of order 4. Therefore, the number of variables of this electrical circuit has been reduced significantly. Moreover, the computational cost has been of 2567 flops whereas with the Silverman-Ho algorithm has been of 4341 flops. Flops is a MatLab command that returns the cumulative number of floating point operations.

References

[1] S.L. Campbell, *Singular systems of differential equations II*, Pitman Publishing, London, 1982.

[2] R. Cantó, B. Ricarte and Ana M. Urbano, *Positive Realizations of Transfer Matrices with real poles*, IEEE Trans. Circuits Syst. II, Expr. Briefs, vol. 54, No. 6, (2007), 517-521.

[3] R. Cantó, B. Ricarte and Ana M. Urbano, *Full rank factorization and the Flanders Theorem*, Electronic Journal of Linear Algebra, vol. 18, (2009), 352-363.

[4] L. Dai, *Singular Control Sistems*, Lecture Notes in Control and Information Sciences, Springer-Verlag, vol. 118, 1989.

[5] T. Kaczorek, *Positive 1D and 2D Systems*, Springer, London, 2002.

An optimal control for a diabetes model*

C. Coll†, A. Herrero ‡,
E. Sánchez§, N. Thome¶

Instituto de Matemática Multidisciplinar,

Universidad Politécnica de Valencia,

46022 Valencia, Spain

December 10, 2009

1 Introduction

People with diabetes are at greater risk of dying from heart and kidney diseases than adults without diabetes. The number of people with diabetes mellitus type 2 has increased considerably worldwide. Concretely, in Spain, according to different studies [8], the prevalence of diabetes has increased to 10-15% of the population. This increase is due to several factors such as changes in the diagnostic criteria (the fasting blood glucose value has been changed from 140mg/dl to 126mg/dl,), the age of the population or actually new cases of diabetes.

Given the significant negative impact of diabetes in very serious diseases, the control of diabetes prevalence should be considered as part of a comprehensive management strategy on the population of a country. It is very interesting to reconsider the model for undiagnosed diabetes and determine which measures should be taken to prevent progression of diabetes in different groups. However, little is known about the impact of optimal control on the

*This paper has been partially supported by DGI grant MTM2007-64477.

†mccoll@mat.upv.es

‡aherrero@mat.upv.es

§esanchezj@mat.upv.es

¶njthome@mat.upv.es

prevalence of diabetes. It is likely that a model which includes some control over the factors in the prevalence of diabetes represents an important role in the design of strategies to mitigate disease progression. This is because the use of optimal control techniques allows the incorporation of functional constraints and desired requirements as a starting point for the design of the process. Recall that, in the optimization problem, the performance of the system can be improved by adjusting the weighting factors in the functional according to the country needs.

Recently, several works, including information on the prevalence of diabetes have been published [1, 3, 5, 7, 11]. In this paper a dynamic model for type 2 (commonly undiagnosed) diabetes prevalence is considered taking into account the number of people diagnosed with diabetes, those not diagnosed and those who have dead in each year. General risk factors like age, sex and race that may have influenced in prevalence diabetes are also considered in the model, but it was not included women who had diabetes only during pregnancy.

The problem of minimization of a quadratic cost functional along the trajectories of the controlled system is introduced. Such an optimization problem is usually known as the *linear quadratic control* (LQC) problem. The solution of the LQC problem is closely related to the solution of certain Riccati matrix equation.

The dynamic model involves a difference-algebraic equation whose coefficients are rectangular matrices, so the Moore-Penrose inverse plays an important role in its solution. Recall that the matrix $A^\dagger$ denotes the Moore-Penrose inverse of the matrix $A \in \mathbb{R}^{m\times n}$. When A has full row rank, it is well-known that the expression of its Moore-Penrose inverse is given by $A^\dagger = A^T(AA^T)^{-1}$. A square matrix A is positive definite (semi-definite) if A is symmetric and $x^TAx > 0$ ($x^TAx \geq 0$) for all $x \neq 0$.

2 Mathematical description and preliminaries

Several authors have modelled the diabetes prevalence by using a singular control system [2, 3, 4, 6]. In this work we use a similar model to the considered in [3] but with controls. Then, the control process is described by the

system (E, A, B) formed by the equation

$$Ex(k+1) = Ax(k) + Bu(k) \tag{1}$$

where the coefficient matrices are given by

$$E = \begin{bmatrix} I & -I & -I \end{bmatrix} \in \mathbb{R}^{n\times 3n}, \quad A = \begin{bmatrix} M^T & O & O \end{bmatrix} \in \mathbb{R}^{n\times 3n}, \quad B \in \mathbb{R}^{n\times l},$$

the state vector is $x(k) = \begin{bmatrix} p(k) & b(k) & pm(k) \end{bmatrix} \in \mathbb{R}^{3n}$, with $n = 2(m+1)$, and the control vector $u(k) \in \mathbb{R}^l$. Here

$$M = \begin{bmatrix} O & \operatorname{diag}(M_i)_{i=0}^{m-1} \\ O & M_m \end{bmatrix}, \quad \text{with} \quad M_i = \begin{pmatrix} \alpha_i & \beta_i \\ 0 & \gamma_i \end{pmatrix}, \quad \text{and}$$

- The population vector $p(k) \in \mathbb{R}^{2(m+1)}$ contains the number of people which are non-diabetic $p_i^{nd}(k)$ or diabetic $p_i^d(k)$ at age i in the year k, that is $p(k) = \begin{bmatrix} p_i^{nd}(k) & p_i^d(k) \end{bmatrix}_{i=0}^{m}$. The age is considered from 0 to m, where m is sufficiently large for neglecting people of age greater than m.

- The population migration vector $pm(k) \in \mathbb{R}^{2(m+1)}$ is a row containing the net migrations for year k with the same structure of vector $p(k)$.

- The birth vector is denoted by $b(k) = \begin{bmatrix} b_1(k) & 0 & \dots & 0 \end{bmatrix} \in \mathbb{R}^{2(m+1)}$ with $b_1(k) \in \mathbb{R}^2$ representing the number of non-diabetic births and diabetic births for year k.

- α_i, β_i, and γ_i are the probabilities of being non-diabetic at age $i+1$ given non-diabetic at age i, of being diabetic at age $i+1$ given non-diabetic at age i, of being diabetic at age $i+1$ given diabetic at age i, respectively.

By the structure of the coefficient matrices, $\operatorname{rank}(E-A) = \operatorname{rank}(E) = n$, and then the Moore-Penrose inverse of $E - A$ is given by $(E-A)^\dagger = (E-A)^T((E-A)(E-A)^T)^{-1}$.

From this, we can define the $n \times n$ matrices $\hat{E} = E(E-A)^\dagger$ and $\hat{A} = A(E-A)^\dagger$, which satisfy $\hat{A} = \hat{E} - I$ and consequently they commute [3]. Moreover, the projector $P = (E-A)^\dagger(E-A)$ has index 1 and so $\mathbb{R}^{3n}$ can be decomposed into direct sum of $\operatorname{Im}(P)$ and $\operatorname{Ker}(P) = \operatorname{Im}(I-P)$. From this, system (1) can be written as

$$EPx(k+1) = APx(k) + Bu(k) - E(I-P)x(k+1) + A(I-P)x(k).$$

Using the matrices $\hat{E}$, $\hat{A}$, and P we obtain the associated standard system

$$\hat{x}(k+1) = \hat{E}^{-1}\hat{A}\hat{x}(k) + \hat{E}^{-1}v(k),$$

given that the matrix $\hat{E}$ is regular and taking $\hat{x}(k) = (E - A)x(k)$ and $v(k) = Bu(k) - E(I - P)x(k+1) + A(I - P)x(k)$. This system has solution for any initial vector $\hat{x}(0) = \widehat{x}_0$ and it is given by

$$\hat{x}(k) = (\hat{E}^{-1}\hat{A})^k\widehat{x}_0 + \sum_{i=0}^{k-1}(\hat{E}^{-1}\hat{A})^{k-i-1}\hat{E}^{-1}\big(Bu(i) + g(i)\big), \tag{2}$$

where $g(i) = -E(I - P)h(i+1) + A(I - P)h(i)$, being $h(i)$ an arbitrary vector for each $i = \{0, \ldots, k-1\}$. Then, since $x(k) \in \mathrm{Ker}(I - P)$, we get $x(k) = (E - A)^{\dagger}\hat{x}(k)$.

3 Linear Quadratic Control Problem

Optimal control allows the incorporation of functional constraints and requirements as a starting point for the design process. In this section, the following linear quadratic control problem, which models a diabetes prevalence problem, is presented for the control system (1),

LQC problem 1. *Determining a control law $u(k) = -Fx(k)$ to minimize the cost functional given as follows:*

$$J(u) = \sum_{k=0}^{\infty}(x^T(k)Qx(k) + u^T(k)Ru(k))$$

where

$$Q = \begin{bmatrix} Q_1 & -Q_1 & -Q_1 \\ -Q_1 & Q_2 & Q_1 \\ -Q_1 & Q_1 & Q_3 \end{bmatrix}$$

is a positive semi-definite matrix and R is a positive definite matrix.

To solve this LQC problem we transform the original system into a control system equivalent by using the change of basis transformation $x(k) = Nz(k)$, where the matrix N is

$$N = \begin{bmatrix} I & I & I \\ O & I & O \\ O & O & I \end{bmatrix}.$$

So, we have the following system

$$\tilde{E}z(k+1) = \tilde{A}z(k) + Bu(k)$$

with $\tilde{E} = \begin{bmatrix} I & O & O \end{bmatrix}$ and $\tilde{A} = M^T \begin{bmatrix} I & I & I \end{bmatrix}$. Denoting $z^T(k) = \begin{bmatrix} z_1^T(k) & z_2^T(k) & z_3^T(k) \end{bmatrix}$ and $\tilde{u}^T(k) = \begin{bmatrix} z_2^T(k) & z_3^T(k) & u^T(k) \end{bmatrix}$, we have

$$z_1(k+1) = M^T z_1(k) + \tilde{B}\tilde{u}(k) \tag{3}$$

where $\tilde{B} = \begin{bmatrix} M^T & M^T & B \end{bmatrix}$. Note that solving the LQC problem 1 is equivalent to solve a new LQC problem for system (3) given by

LQC problem 2. *Determining a control law $\tilde{u}(k) = -Lz_1(k)$ to minimize the cost functional given by*

$$\tilde{J}(\tilde{u}) = \sum_{k=0}^{\infty} (z_1^T(k)\tilde{Q}z_1(k) + \tilde{u}^T(k)\tilde{R}\tilde{u}(k))$$

where $\tilde{Q}$ is a positive semi-definite matrix and $\tilde{R}$ is a block-diagonal positive definite matrix. Note that, $\tilde{Q} = Q_1$ and $\tilde{R} = diag(Q_2 - Q_1, Q_3 - Q_1, R)$.

To solve this LQC problem we need that the system (3) be completely reachable. Remember that a system $x(k+1) = Ax(k) + Bu(k)$ is completely reachable if and only if matrix $[B\ AB \cdots A^{n-1}B]$ has full rank. In this case we show the following result

Proposition 1 *System (3) is completely reachable if and only if the matrix $B = \begin{bmatrix} B_1^T & B_2^T \end{bmatrix}^T$ with $B_1 \in \mathbb{R}^{2\times l}$, $B_2 \in \mathbb{R}^{(n-2)\times l}$, and* $\text{rank}(B_1) = 2$.

If we assume that system (3) is completely reachable then, the minimizing solution of $\tilde{J}$ with respect to control input $\tilde{u}$ is given by $L = (\tilde{R} + \tilde{B}^T P \tilde{B})^{-1}\tilde{B}^T P M^T$ where P is the positive definite matrix satisfying the Riccati equation

$$P = Q_1 + MP(I + \tilde{B}\tilde{R}^{-1}\tilde{B}^T P)^{-1}M^T.$$

Next, we have to return to the LQC problem 1. For that we split $L^T = \begin{bmatrix} L_1^T & L_2^T & L_3^T \end{bmatrix}$ and from $\tilde{u}(k) = -Lz_1(k)$, we have

$$z_2(k) = -L_1 z_1(k), \qquad z_3(k) = -L_2 z_1(k), \qquad u(k) = -L_3 z_1(k).$$

Thus, by the transformation $z(k) = N^{-1}x(k)$ we have that the solution of the LQC problem 1 is $u(k) = -Fx(k)$ with

$$F = L_3 \left[\begin{array}{ccc} -I & I & I \end{array} \right],$$

and this feedback is well-defined if the following condition holds:

$$x(k) \in \mathrm{Ker} \left(\left[\begin{array}{ccc} L_1 & I - L_1 & -L_1 \end{array} \right] \right) \cap \mathrm{Ker} \left(\left[\begin{array}{ccc} L_2 & -L_2 & I - L_2 \end{array} \right] \right) \qquad (4)$$

Summarizing, we have the following result.

Theorem 1 *Let the system (1) be satisfying condition of Proposition 1 and consider the LQC problem 1. If the matrices Q_1, Q_2 and Q_3 satisfy that $Q_2 - Q_1$ and $Q_3 - Q_1$ are positive definite matrices and the condition (4) holds then the LQC problem 1 has solution.*

4 Conclusions

A problem of optimal control for a singular system, which represents the diabetes prevalence in a hypothetical country, has been considered. A solution of this system has been given by using the Moore-Penrose inverse. Finally, the solution of the LQC problem has been obtained by means of an auxiliary system which satisfies the reachability condition. This solution is closely related to a Riccati matrix equation.

References

[1] A. F. Amos, D. J. McCarty, P. Zimmet, The rising global burden of diabetes and its complications: estimates and projections to the year 2010, Diabet Med 14 (Suppl. 5), S1–S85 (1997).

[2] S. L. Campbell, *Singular systems of differential equations,* Pitman Advanced Publishing Program, (San Francisco, 1980).

[3] C. Coll, A. Herrero, E. Sánchez, N. Thome, A dynamic model for a study of diabetes. Mathematical and Computer Modelling, 50, 713-716 (2009).

[4] L. Dai, *Singular Control Systems*, Lecture notes in Control and Inform. Sci. 118, Springer-Verlag, (Berlín, 1989).

[5] A. Honeycutt, J. P. Boyle, K. R. Broglio, T. J. Thompson, T. J. Hoerger, L. S. Geiss, K. M. Venkat, A dynamic Markov model for forecasting diabetes prevalence in the United State through 2050, *Helath Care Management Science* 6, 155-164 (2003).

[6] T. Kaczorek, *Linear Control Systems,* Wiley and Sons, (New York, 1992).

[7] H. King, R. E. Aubert, W. H. Herman, Global burden of diabetes 1995-2025: Prevalence, numerical estimates, and projections. *Diabetes Care* 21, 1414-1431 (1998).

[8] S. Valdés, G. Rojo-Martínez, F. Soriguer, Evolution of prevalence of type 2 diabetes in adult Spanish population, *Medicina Clinica*, 129, 352-355 (2007).

[9] S. Valdés, P. Botas, E. Delgado, F. Alvares, F. Diaz, Population-Based Incidence of Type 2 Diabetes in Northern Spain, *Diabetes Care* 30 (9), 2258-2263 (2007).

[10] J. Tuomilehto, J. Lindstrom, J. Eriksson, T.Valle, H. Hamalainen , P. Ilanne-Parikka , et al. Prevention of type 2 diabetes mellitus by changes inlifestyle among subjects with impaired glucose tolerance, *N Engl J Med.*, 344,1343-1350 (2001).

[11] S. Wild, G. Rogilc, A. Green, R. Sicree, H. King, Global prevalence of diabetes, *Diabetes Care*, 27 (5), 1047-1053 (2004).

Numerical solution for a nonlinear problem modeling option pricing in illiquid markets. *

R. Company†, L. Jódar‡and J.–R. Pintos§

Instituto de Matemática Multidisciplinar,

Universidad Politécnica de Valencia,

Camino de Vera s/n, 46022 Valencia, España

December 10, 2009

1 Introduction

This paper deals with the numerical analysis and computing of nonlinear models of option pricing that appear when illiquid markets effects are taken into account. It is well-known that Black-Scholes (B-S) model is acceptable in idealized financial markets. One of the major assumptions of Black-Scholes model is that the market in the underlying asset is perfectly elastic so that large trades do not affect prices in equilibrium. This occurs in perfectly liquid markets, but the case is clearly unrealistic.

The presence of price impact of investors' trading has been widely documented and extensively analyzed in the literature, see, for instance, [1, 2, 3]. In [4], the way in which price impact in the underlying asset market affects the replication of a European contingent claim is examined. They obtain a generalized Black-Scholes pricing PDE that for the case where interest rate and the reference volatility are constant, takes the form

*This paper has been supported by the Spanish Department of Science and Education grant TRA2007-68006-C02-02 and the Generalitat Valenciana grant GVPRE/2008/092.

†e-mail: rcompany@imm.upv.es

‡e-mail: ljodar@imm.upv.es

§e-mail: jrpt60@gmail.com

$$\frac{\partial v}{\partial t}(S,t)+\frac{\sigma^2 S^2}{2\left(1-\lambda(S,t)S\dfrac{\partial^2 v}{\partial S^2}(S,t)\right)^2}\frac{\partial^2 v}{\partial S^2}(S,t)+rS\frac{\partial v}{\partial S}(S,t)-rv(S,t)=0,$$

$$(S,t)\in]0,+\infty[\times]0,T] \tag{1}$$

$$v(S,T)=f(S),\quad 0<S<+\infty. \tag{2}$$

[4] establish the existence and uniqueness of a classical solution to this PDE, for the case where the payoff function $f(S)$ of the European contingent claim is a continuous piecewise linear function. In (1), $\lambda(S,t)$ is the price impact factor

$$\lambda(S,\ t)=\begin{cases}\dfrac{\gamma}{S}\left(1-e^{-\beta(T-t)}\right) & if\ \underline{S}\leq S\leq \overline{S}\\ \\ 0 \qquad\qquad\qquad , & otherwise,\end{cases} \tag{3}$$

where the constant price impact coefficient $\gamma>0$ measures the price impact per traded share, and $\underline{S}$ and $\overline{S}$ represent respectively, the lower and upper limit of the stock price within which there is a price impact.

If $x=[x_1,\ldots,x_p]^t$ is a vector in $\mathbb{R}^p$, its 1-norm is denoted by $\|x\|=\sum_{i=1}^{p}|x_i|$ and its supremum norm is denoted by $\|x\|_\infty=\max\limits_{1\leq i\leq p}|x_i|$.

2 Transformation of the problem.

For the sake of convenience the PDE (1)-(2) is going to be transformed into a nonlinear diffusion equation.

Let us consider the substitution defined by

$$\begin{gathered}X=e^{r(T-t)}S;\quad \tau=\tfrac{\sigma^2}{2}(T-t)\,;\quad u=e^{r(T-t)}v\\ \nu(X,\tau)=\lambda(S,t)\ S\ v_{SS}\,;\quad \rho=\frac{r}{\sigma^2}.\end{gathered} \tag{4}$$

From (1)-(2) and (4) one gets

$$L(u)=\frac{\partial u}{\partial \tau}-\frac{1}{\left[1-\nu(X,\tau)\right]^2}X^2\frac{\partial^2 u}{\partial X^2}=0,\quad (X,\tau)\in\mathbb{R}^+\times]0,\frac{\sigma^2 T}{2}], \tag{5}$$

together with the initial condition

$$u(X,0) = f(X)\,, \qquad X > 0. \tag{6}$$

Note that from (3) and (4), $\nu(X,\tau)$ takes the form

$$\nu(X,\tau) = \begin{cases} \gamma e^{2\rho\tau}\left(1 - e^{-\frac{2\beta\tau}{\sigma^2}}\right) u_{XX}, & \underline{X} \le X \le \overline{X} \\ \\ 0 \qquad\qquad\qquad\qquad , & otherwise. \end{cases} \tag{7}$$

3 Numerical scheme construction

The numerical domain is $(X,\tau) \in [0,b] \times [0, \frac{\sigma^2 T}{2}]$ and the nodes $X_j = jh$, $\tau^n = nk$, with $0 \le j \le N$, $0 \le n \le \ell$, $Nh = b$ and $\ell k = \frac{\sigma^2 T}{2}$. The numerical approximation of the exact theoretical solution $u(X_j, \tau^n)$, is denoted by U_j^n. Let us introduce the approximations of the partial derivatives and the operator $\triangle_j^n$:

$$\left.\begin{aligned} &\frac{\partial u}{\partial \tau}(S_j,\tau^n) = \frac{U_j^{n+1} - U_j^n}{k} + O(k), \\ &\frac{\partial^2 u}{\partial X^2}(S_j,\tau^n) = \frac{U_{j-1}^n - 2U_j^n + U_{j+1}^n}{h^2} + O(h^2) = \triangle_j^n(U) + O(h^2). \end{aligned}\right\} \tag{8}$$

The approximations U_{-1}^n and U_{N+1}^n are obtained by using linear extrapolation.

By replacing the partial derivatives of equation (5) by the approximations given in (8) one gets the numerical scheme at the internal mesh points:

$$\begin{gathered} U_j^{n+1} = \left(1 - \frac{2k}{h^2}\beta_j^n\right) U_j^n + \frac{k}{h^2}\beta_j^n \left(U_{j-1}^n + U_{j+1}^n\right) \\ 1 \le j \le N-1, \qquad 0 \le n \le \ell - 1, \end{gathered} \tag{9}$$

where

$$\beta_j^n = \frac{X_j^2}{[1-\nu_j^n]^2}, \quad 0 \le j \le N, \quad 0 \le n \le \ell, \tag{10}$$

$$\nu_j^n = \begin{cases} \gamma e^{2\rho nk}\left(1 - e^{-\frac{2\beta nk}{\sigma^2}}\right) \triangle_j^n, & \underline{X} \le X \le \overline{X} \\ \\ 0 \qquad\qquad\qquad\qquad , & otherwise. \end{cases} \tag{11}$$

From (9) evaluated at $j = 0$ and $j = N$, at the boundaries we obtain

$$U_0^{n+1} = U_0^n = ... = U_0^0 = f(0), \tag{12}$$

$$U_N^{n+1} = U_N^n = \ldots = U_N^0 = f(b). \tag{13}$$

For the sake of clarity in the study of the properties of the numerical solution of (9)-(13) it is convenient the study of behaviour of $\triangle_j^n$ appearing in the coefficients of (10)-(11).

Taking into account the definition of $\triangle_j^n$ in (8) and the evolution of the numerical solution from n to $n+1$ given by (9)-(13), one gets

$$\triangle_j^{n+1} = \left(1 - \frac{2k}{h^2}\beta_j^n\right)\triangle_j^n + \frac{k}{h^2}\beta_{j-1}^n\triangle_{j-1}^n + \frac{k}{h^2}\beta_{j+1}^n\triangle_{j+1}^n, \;\; 1 \le j \le N-1, \tag{14}$$

Let us denote the vector $\triangle^n$ in $\mathbb{R}^{N+1}$ defined by

$$\triangle^n = [\triangle_0^n \;\; \triangle_1^n \; ... \; \triangle_N^n]^t, \quad 0 \le n \le \ell. \tag{15}$$

The following result shows that under certain conditions, the operator $\triangle^n$ is decreasing in $1 - norm$ as n increases.

Lemma 1 *Let r be the riskless interest rate, σ the volatility and γ the impact factor parameter of the illiquid market. Let us assume the conditions*

$$\gamma e^{2\rho\tau}\left(1 - e^{-\frac{2\beta\tau}{\sigma^2}}\right)\|\triangle^0\| = \delta < 1 \tag{16}$$

$$\frac{k}{h^2} \le L(h), \quad \textit{and} \quad L(h) = \frac{(1-\delta)^2}{2b^2} \tag{17}$$

then

$$\|\triangle^{n+1}\| \le \|\triangle^n\|, \qquad 0 \le n \le \ell.$$

4 Positivity, monoticity, and stability

4.1 Positivity

A suitable property of the numerical solution of an equation pricing a contract is the positivity. Under hypothesis (16), (17) of lemma 1 inequality

$$1 - \frac{2k}{h^2}\beta_j^n \ge 0, \quad 0 \le j \le N, \;\; 0 \le n \le \ell - 1. \tag{18}$$

holds true and taking into account (9) and (10) one gets that for a nonnegative payoff $\{U_j^0\}$ the numerical solution $\{U_j^n\}$ is nonnegative.

4.2 Monoticity

We introduce the following definition.

Definition 1 *Consider the scheme* $F(U_j^n) = 0\,,\ j \in J,\ n \in L$ *where* J *and* L *are sets of nonnegative integers. We say that the scheme is monoticity-preserving, if, assuming that* $U_{j+1}^n \geq U_j^n\,, \quad j \in J\,, \quad j+1 \in J$, *then, it occurs that* $U_{j+1}^{n+1} \geq U_j^{n+1}\,, \quad j \in J\,, \quad j+1 \in J$.

Theorem 1 *Under hypotheses (16)-(17) of Lemma 1, then the numerical scheme (9), (12), (13) is monoticity-preserving , with* $0 \leq j \leq N$, $0 \leq n \leq \ell$.

Corollary 1 *Under hypotheses (16)-(17) and notation of theorem 1, and assuming that the payoff function* $f(X)$ *is non decreasing and nonnegative with* $f(0) = 0$, *then the scheme (9), (12), (13) is nonnegative and nondecreasing in variable* j *for each time stage* n.

4.3 Stability

Let us denote the numerical solution vector $u^n = [u_0^n\ u_1^n \cdots u_N^n]^t$.
We introduce the following definition.

Definition 2 *The numerical scheme (9)-(13) for the initial value problem (5) is said to be* $\|\ \|_\infty$*-stable in the fixed station sense in the domain* $[0, b] \times [0, \frac{\sigma^2 T}{2}]$, *if given* τ *with* $0 < \tau \leq \frac{\sigma^2 T}{2}$, *for every partition with* $k = \triangle\tau$, $h = \triangle X$, *with* $\tau = \ell k$, *and every* N *with* $Nh = b$ *one gets* $\|\, U^n \,\|_\infty \leq C$, *where* $C > 0$ *is independent of* $h,\ k$ *and* N.

From (13), theorem 1 and corollary 1, $\|U^n\|_\infty = f(b)$ and the following result holds true.

Theorem 2 *Under conditions (16)-(17), the numerical scheme (9)-(13) for solving initial value problem (5) with a nondecreasing payoff function is* $\|\ \cdot\ \|_\infty$*-stable.*

5 Consistency

Consistency of a numerical scheme with respect to a partial differential equation means that the exact theoretical solution of the PDE approximates well

the exact theoretical solution of the difference scheme as the stepsize discretization tends to zero, [5]. Let us write the scheme (9) in the form

$$F(U_j^n) = \frac{U_j^{n+1} - U_j^n}{k} - \beta_j^n(U)\triangle_j^n(U) = 0, \tag{19}$$

where discrete operator Δ_j^n is given by (8) and coefficients $\beta_j^n(U)$ are given by (10)-(11). In accordance with [5, pag.100], the scheme (19) is said to be consistent with problem (5) if the local truncation error

$$T_j^n(u) \;=\; F(u_j^n) \;-\; L(u_j^n), \tag{20}$$

satisfies

$$T_j^n(u) \;\to\; 0\,, \quad as \quad h = \Delta X \;\to\; 0\,, \quad k = \Delta\tau \;\to\; 0\,, \tag{21}$$

where u_j^n denotes the value of the theoretical solution of $L(u) = 0$ at the mesh point (X_j, τ^n), $X_j = jh$, $\tau^n = nk$.

Using Taylor's expansions about (X_j, τ^n) the following result can be established:

Theorem 3 *The numerical scheme (9)-(11) is consistent with equation (5), and the local truncation error T_j^n satisfies*

$$T_j^n(u) = O(h^2) + O(k).$$

6 Examples

In this section we check the properties of the proposed numerical scheme (9)-(13) for the model (5). Furthermore, simulations are performed for different values of the price impact coefficient γ.

Example 1 *Consider the vanilla call option for an illiquid market with Strike $E = 50$, $r = 0.06$, $\sigma = 0.4$, $T = 1$, $\underline{S} = 20$, $\overline{S} = 80$, $\beta = 100$, $\gamma = 1$ and $h = 1.5928$.*

Figure 1 shows the option pricing value. For $k = 5.3333 \cdot 10^{-6}$ the sufficient stability conditions are satisfied, dot line. For $k = 5.6537 \cdot 10^{-5}$ the stability conditions are broken appearing spurious oscillations in the numerical solution, continuous line.

Example 2 *Consider the problem treated in example 1 under stability step size requirements. Figure 2 shows the variation of the option price with the parameter γ, showing that price grows with the value of γ, mainly in the proximities of the strike price.*

7 Conclusion

A consistent monotone finite difference scheme is proposed and a relationship between the discretization stepsize is obtained ensuring nonnegative and stable numerical solutions and avoiding spurious oscillations.

References

[1] W. F. Sharpe, G. J. Alexander, J. V. Bailey, Investments, Prentice Hall, New Jersey, 1999.

[2] P. Jorion, Value at Risk, McGraw-Hill, New York, 2000.

[3] R. Frey, P. Patie, Risk management for derivatives in illiquid markets: A simulation study, Advances in Finance and Stochastics (2002) 137–159.

[4] H. Liu, J. Yong, Option pricing with an illiquid underlying asset market, J.Economic Dynamics and Control 29 (2005) 2125–2156.

[5] G. D. Smith, Numerical solution of partial differential equations: finite difference methods, 3rd Edition, Clarendon Press, Oxford, 1985.

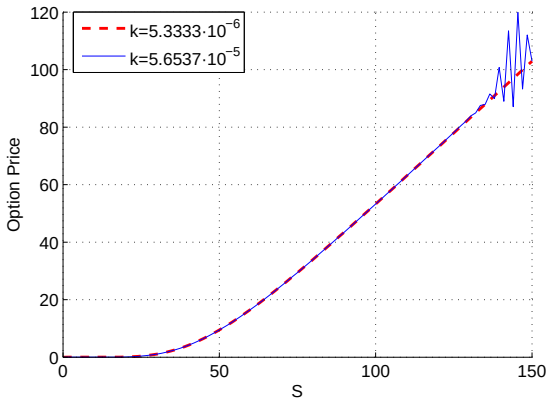


Figure 1: Numerical solutions for several temporal step size discretization values.

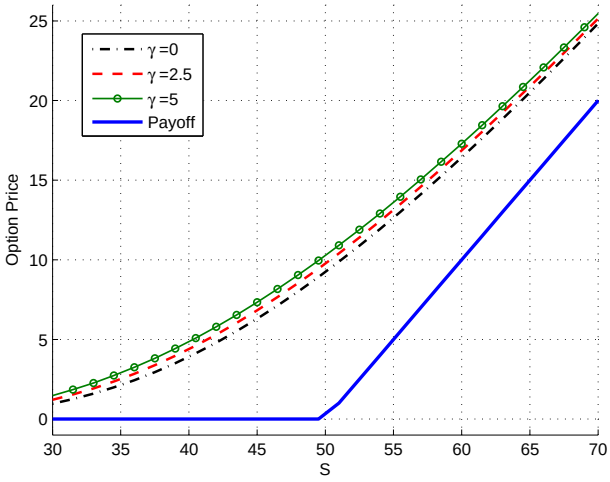


Figure 2: Variation of the option price with the parameter γ.

Multi-point iterative methods for nonlinear equations with multiple roots. *

A. Cordero†, J.-L. Hueso‡, E. Martínez§, J.-R. Torregrosa¶

Instituto de Matemática Multidisciplinar,

Universidad Politécnica de Valencia,

Edificio 8G, Piso 2, 46022 Valencia, España.

December 10, 2009

1 Introduction

Iterative methods play a main role in the mathematical modeling of engineering problems: frequently they derive to nonlinear equations of type $f(x) = 0$. Usually, iterative methods for solving nonlinear equations require the root to be simple. Nevertheless, the problem about multiple roots is less studied. The aim of this work is to find efficient methods in the solution of this kind of nonlinear equations.

A family of iterative methods for approximating a root α of a nonlinear equation $f(x) = 0$ has been presented recently in [1]. The general method has the expression

$$x_{k+1} = x_k - \frac{1}{f'(x_k)} \sum_{j=1}^{m} A_j f(\eta_j(x_k)), \tag{1}$$

*This research was supported by Ministerio de Ciencia y Tecnología MTM2007-64477

†acordero@mat.upv.es

‡jlhueso@mat.upv.es

§eumarti@mat.upv.es

¶jrtorre@mat.upv.es

with $\eta_j(x_k) = x_k - \tau_j \dfrac{f(x_k)}{f'(x_k)}$, where τ_j and A_j are parameters to be chosen in $[0,1]$ and $\mathbb{R}$, respectively. In the case of simple roots it was proved that the value of these parameters plays an important role in the order of convergence of the proposed methods, obtaining at least order of convergence three.

In Section 2 we prove that the iterative methods (1) have order of convergence one if α is a multiple root of multiplicity $p > 1$. If this multiplicity is known we may modify the methods to obtain order of convergence two, by using different types of correction. When p is unknown, we prove in Section 3 that three of the most efficient methods of family (1) converge faster than the classical Newton's method. These methods, denoted by $M1$, $M2$ and $M3$, are the following:

M1. (Potra and Ptak [3])
By choosing $A_1 = A_2 = 1$, $\tau_1 = 0$ and $\tau_2 = 1$, we obtain $x_{k+1} = x_k - \dfrac{f(x_k) + f(\eta_2(x_k))}{f'(x_k)}$, with $\eta_2(x_k) = x_k - \dfrac{f(x_k)}{f'(x_k)}$.

M2. (see [1] and [4])
With $A_1 = \dfrac{3+\sqrt{5}}{2}$ and $\tau_1 = \dfrac{\sqrt{5}-1}{2}$, we have $x_{k+1} = x_k - \dfrac{3+\sqrt{5}}{2}\dfrac{f(\eta_1(x_k))}{f'(x_k)}$, where $\eta_1(x_k) = x_k - \dfrac{\sqrt{5}-1}{2}\dfrac{f(x_k)}{f'(x_k)}$.

M3. (see [1])
By taking $A_1 = A_3 = 4$, $A_2 = -6$, $\tau_1 = 1/4$, $\tau_2 = 1/2$ and $\tau_3 = 3/4$, we have $x_{k+1} = x_k - \dfrac{4f(\eta_1(x_k)) - 6f(\eta_2(x_k)) + 4f(\eta_3(x_k))}{f'(x_k)}$, where $\eta_i(x_k) = x_k - \tau_i \dfrac{f(x_k)}{f'(x_k)}$.

Finally, last section is dedicated to numerical results obtained by applying methods $M1$, $M2$, $M3$ and Newton's method to several nonlinear equations.

2 Analysis of convergence: the correction factor.

Let $f : I \subseteq \mathbb{R} \longrightarrow \mathbb{R}$, be a sufficiently differentiable function and $\alpha \in I$ a multiple root of multiplicity $p > 1$, of the nonlinear equation $f(x) = 0$. Let

g be the fixed point function that allows us to describe (1)

$$g(x) = x - \frac{1}{f'(x)} \sum_{j=1}^{m} A_j f(\eta_j(x)), \tag{2}$$

with $\eta_j(x) = x - \tau_j \dfrac{f(x)}{f'(x)}$.

If $f(x)$ is a sufficiently differentiable function and α is a multiple root of $f(x)$, with multiplicity $p > 1$, it can be proven that the iterative fixed point methods, described by (1), have only linear convergence independently of values of parameters A_j and τ_j.

Theorem 1 *If α is a root of $f(x)$ with multiplicity $p > 1$, then*

$$g'(\alpha) = 1 - \frac{1}{p} \sum_{j=1}^{m} A_j (1 - \frac{\tau_j}{p})^p \neq 0.$$

If we denote $C = \dfrac{p}{\sum_{j=1}^{m} A_j \left(1 - \dfrac{\tau_j}{p}\right)^p}$, in virtue of Theorem 1 it is possible to restore the quadratic convergence to the iterative methods (1), when α is a root of multiplicity $p > 1$, by correcting (1) in the form

$$x_{k+1} = x_k - C \sum_{j=1}^{m} \frac{A_j f(\eta_j(x_k))}{f'(x_k)}. \tag{3}$$

It is interesting to note that C depends on the multiplicity p and the parameters A_j and τ_j. Unfortunately, the methods (3) have no longer order of convergence three. By using the correction factor C in the methods described by $M1$, $M2$ and $M3$ we obtain the corrected methods, which we denote by $CM1$, $CM2$ and $CM3$, respectively.

On the other hand, it is well known (see [2]) that the Newton's method converges only linearly if α is a multiple root with multiplicity p, but if we define as $x_{k+1} = x_k - p\dfrac{f(x_k)}{f'(x_k)}$ the iterate of the corrected Newton's method (CNM), we still have quadratic convergence. In the case of multiple roots we can hope to improve the order of convergence of our iterative methods (1) by using, instead of $\eta_j(x)$, the function

$$\eta_j^*(x) = x - p\tau_j \frac{f(x)}{f'(x)}. \tag{4}$$

Theorem 2 *Let α be a root of $f(x)$ with multiplicity $p > 1$. If we use the iteration given by (1), with $\eta_j^*(x)$ instead of $\eta_j(x)$, then*

$$g'(\alpha) = 1 - \frac{1}{p}\sum_{j=1}^{m} A_j(1-\tau_j)^p \neq 0.$$

So, the new correction factor is $D = \dfrac{p}{\sum_{j=1}^{m} A_j(1-\tau_j)^p}$. By using Theorem 2 it is possible again to restore the quadratic convergence to the methods (1), when α is a multiple root of multiplicity $p > 1$, by correcting the mentioned methods in the form

$$x_{k+1} = x_k - D\sum_{j=1}^{m} \frac{A_j f(\eta_j^*(x_k))}{f'(x_k)}. \tag{5}$$

If we use this new correction factor D in the methods described by $M1$, $M2$ and $M3$ we obtain new corrected methods, which we denote by $DM1$, $DM2$ and $DM3$, respectively.

3 The correction factor when p is unknown

Theoretically, we have proven that corrected iterative methods have quadratic convergence, but we need to know the multiplicity p of root α. If we do not know p, we cannot use the corrected methods but, in this case, we can prove that methods $M1$, $M2$ and $M3$ converge faster than Newton's method, since the mentioned methods have $g'(\alpha)$ smaller than $g'_{Newton}(\alpha) = 1 - 1/p$.

Theorem 3 *If α is a multiple root of $f(x)$, with multiplicity $p > 1$, method $M2$ converges faster than method $M1$, $M1$ converges faster than $M3$ and this one faster than Newton's method.*

In Table 1,we give the convergence factors of the methods $M1$, $M2$, $M3$ and Newton's method, for $p = 2, 3, \ldots, 10$.

p	M2	M1	M3	NM
2	0.3750	0.3750	0.3750	0.5000
3	0.5632	0.5679	0.5679	0.6667
4	0.6655	0.6709	0.6710	0.7500
5	0.7293	0.7345	0.7346	0.8000
6	0.7727	0.7775	0.7777	0.8333
7	0.8042	0.8086	0.8087	0.8571
8	0.8280	0.8320	0.8322	0.8750
9	0.8467	0.8504	0.8505	0.8889
10	0.8617	0.8651	0.8653	0.9000

Table 1: Convergence factors

4 Numerical results

In this section we will check the effectiveness of the different numerical corrected methods introduced in this work (including the corrected Newton's method), in order to estimate the multiple roots of the following nonlinear functions:

(a) $f(x) = (sin(x) - x/2)^2$; $\alpha = 0$ with $p = 2$.

(b) $f(x) = exp(x) - x - 1$; $\alpha = 0$ with $p = 2$.

(c) $f(x) = (x-1)(e^{x-1} - 1)$; $\alpha = 1$ with $p = 2$.

(d) $f(x) = x^5 - 8x^4 + 24x^3 - 34x^2 + 23x - 6$; $\alpha = 1$ with $p = 3$.

(e) $f(x) = (x-2)^4 e^x$; $\alpha = 2$ with $p = 4$.

Numerical computation have been carried out in double precision in MATLAB 7.1. The stopping criterion used is $|x_{k+1} - x_k| + |f(x_k)| < 10^{-9}$. For every method, we analyze the number of iterations needed to converge to the solution and the computational order of convergence ρ, approximated by (see [6]):

$$\rho \approx \frac{\ln(|x_{k+1} - \alpha| \,/\, |x_k - \alpha|)}{\ln(|x_k - \alpha| \,/\, |x_{k-1} - \alpha|)}. \tag{6}$$

The value of ρ that appears in Table 2 is the quotient (6) when the variation from one iteration to another is small. In some cases, the computational order of convergence is not stable and it is not shown in the table.

In Table 2, we estimate the zeros of functions from (a) to (e). For every function, the following items are specified: the initial estimation x_0, the solution, and, for each method, the number of iterations and the estimation of the computational order of convergence ρ.

$f(x)$	x_0	**Solution**	**Iterations**				ρ			
			CNM CNM	CM1 DM1	CM2 DM2	CM3 DM3	CNM CNM	CM1 DM1	CM2 DM2	CM3 DM3
(a)	0.5	0	4	4	4	4	3.0	3.0	3.0	3.0
	0.5	0	4	4	5	5	3.0	3.0	3.0	3.0
(b)	-1	1	6	6	6	6	2.0	2.0	2.0	2.0
	-1	1	6	6	11	6	2.0	2.0	2.0	2.0
(c)	0	1	-	5	5	5	-	2.0	2.0	2.0
	0	1	-	6	5	6	-	2.0	2.0	1.8
(d)	0.5	0	6	4	4	4	2.0	2.0	2.0	2.0
	0.5	0	6	4	4	4	2.0	2.0	2.0	2.0
(e)	1.5	2	5	5	5	5	2.0	2.0	2.0	2.0
	1.5	2	5	5	6	5	2.0	2.0	2.0	2.0

Table 2: Numerical results for different iterative methods

As we can see, the numerical tests confirm the theoretical results presented in this work. The order of convergence is quite similar for all methods, but we observe that the corrected methods $CM1$, $CM2$ and $CM3$ are slower than the corrected one $DM1$, $DM2$ and $DM3$.

5 Conclusions

We have shown that the family of multi-point iterative methods described by (1) for solving nonlinear equations with simple roots, can be corrected in order to restore the quadratic convergence when the equation has a multiple root. If we do not know the multiplicity of the root, three of the most efficient methods in the family (1) have order of convergence one, as the classical Newton's method, but have a smaller convergence factor.

References

[1] A. Cordero, Juan R. Torregrosa, A class of multi-point iterative methods for nonlinear equations, *Applied Mathematics and Computation*, 197 (2008), 337-344.

[2] J.F. Traub, *Iterative methods for the solution of equations*, Chelsea Publishing Company, New York, 1982.

[3] F.A. Potra, V. Ptack, Nondiscrete induction and iterative processes, Research notes in Mathematics, vol. 103, Pitman, Boston, 1984.

[4] J. Kou, Y. Li, Modified Chebyshev's method free from second derivative for nonlinear equations, *Applied Mathematics and Computation*, 187 (2007) 1027-1032.

[5] A.M. Ostrowski, *Solutions of equations and systems of equations*, Academic Press, New York-London, 1966.

[6] S. Weerakoon, T.G.I. Fernando, A variant of Newton's method with accelerated third-order convergence, *Applied Mathematics Letters*, 13 (8) (2000), 87-93.

A mathematical model for the electronic commerce in Spain*

J.-C. Cortés†, I.-C. Lombana, R.-J. Villanueva
Instituto Universitario de Matemática Multidisciplinar
Universidad Politécnica de Valencia
Edificio 8G, 2ª, P.O. Box 22012, Valencia, Spain
{jccortes, rjvillan}@imm.upv.es, ivanclombana@gmail.com

1 Introduction

The diffusion of electronic commerce, in short, e-commerce, has dramatically increased its business volume in recent years due to the possibility of saving purchase time, accessing to a greater variety and quality of articles, etc.

The formulation of a reliable mathematical technological diffusion model must consider the particular features of the technology as well as its users. In spite of several authors have developed mathematical diffusion models for study some technologies [1, 2, 3], they do not consider important aspects like the different consumption among people depending on their age and a major impact of innovations on certain age groups. Models of diffusion are a powerful tool to explain and predict the process of adoption of a new good or innovation over time. In our case, taking into account official data from INE, (Spanish Statistical Institute) [4], the population is divided into 6 groups between 16 and 74 years old in Spain to estimate the number of consumers who use internet to buy any good over time. The adopters of the technology can be classified in two groups: the innovators who dare to use e-commerce due to the advertising impact through the media or another external factors regardless of the decision of others, and the imitators, consumers who will only begin to use the e-commerce once they have looked that others use it and as a result of the interaction and influence with innovators.

A compartmental model is proposed to study the diffusion of e-commerce in Spain based on a nonlinear differential system. The proposed model provides a

*This work has been partially supported by the Spanish M.C.Y.T. grant MTM2009-08587 and Universidad Politécnica de Valencia grant PAID06-09-2588

†Corresponding author

tool for forecasting the short-term trends of this technology in Spain and therefore to enable strategic decision-making marketing.

2 Demographic model

2.1 Available data

We have considered in this study available official data from the INE about the percentage of Spanish people who have bought using internet from 2006 to 2008 [4]. These data, the only available at the moment, are classified by age groups is shown in Table 1.

Table 1: Percentage of Spanish people who have bought using internet from 2006 to 2008.

Age groups	Percentage in 2006	Percentage in 2007	Percentage in 2008
16 to 24 years old	28.7	28.7	32.2
25 to 34 years old	31.6	33.4	37.9
35 to 44 years old	23.2	24.3	26.5
45 to 54 years old	14.6	18	20
55 to 64 years old	7.2	7.8	8.6
65 to 74 years old	1.3	2	2

We propose a model of population with the age groups determined by Table 1 with two positions toward innovation: those who adopt the technology (e-commerce) and those who do not adopt it.

2.2 Age-structured demographic model

The age-structured demographic model considers the age interval in accordance with Table 1 and the corresponding growth rates and death rates, from official data from INE [4]. Therefore, we define the age groups:

- Group 1 (G_1) : Population of people aged between 16 and 24 years old,
- Group 2 (G_2) : Population of people aged between 25 and 34 years old,
- Group 3 (G_3) : Population of people aged between 35 and 44 years old,
- Group 4 (G_4) : Population of people aged between 45 and 54 years old,
- Group 5 (G_5) : Population of people aged between 55 and 64 years old,
- Group 6 (G_6) : Population of people aged between 65 and 74 years old.

The age-demographic model that we consider is given by the following system of ordinary differential equations [5]:

$$\begin{aligned} G_1'(t) =& \mu - c_1 G_1(t) - d_1 G_1(t), \\ G_j'(t) =& c_{j-1} G_{j-1}(t) - c_j G_j(t) - d_j G_j(t), \quad 2 \leq j \leq 6. \end{aligned} \tag{2.1}$$

From INE's data, the average population between 16-74 is 33,955,126, and death rates per age group are $d_1 = 0.00053$, $d_2 = 0.00062$, $d_3 = 0.00126$, $d_4 = 0.00277$, $d_5 = 0.00613$ and $d_6 = 0.10230$. The coefficient d_6 has been calculated taking into account that the last age group, apart from the death of the population, there is a proportion of people that reach 75 years old and leaves the system. The birth rate corresponds to young people that enter into the system, i.e, 16-years-old people, and then $\mu = 0.01313$. We will assume that the whole population as well as the corresponding to each age group remain constant over time [5], i.e.,

$$G_1'(t) = G_2'(t) = G_3'(t) = G_4'(t) = G_5'(t) = G_6'(t) = 0. \tag{2.2}$$

Once coefficients μ and d_i have been computed, we address to estimate the parameters c_k, $k = 1, \ldots, 5$ (we consider that the growth-rate c_6=0, because group 6 is the oldest considered in the target population). For this task, we solve in cascade the system (2.1) imposing (2.2), then $c_1 = \dfrac{\mu}{G_1} - d_1$, $c_j = c_{j-1}\dfrac{G_{j-1}}{G_j} - d_j, 2 \leq j \leq 5$, and therefore, their numerical values are given by $c_1 = 0.09621$, $c_2 = 0.05781$, $c_3 = 0.05889$, $c_4 = 0.06923$ and $c_5 = 0.08084$.

3 Age-structured diffusion mathematical model

We introduce two subpopulations where we have divided the individuals aged between 16-74 years old into 6 age groups:

- $N_i(t)$, $i = 1, \ldots, 6$, denotes the percentage of population of the $i-th$ group that does not adopt the technology, at time t.
- $Y_i(t)$, $i = 1, \ldots, 6$, denotes the percentage of population of $i-th$ group that adopts the technology, at time t.

The diffusion of technology will be represented by the transition of an individual in the population N_i(t) to Y_i(t) through the coefficients of innovation or imitation described by:

- P_i, $i = 1, \ldots, 6$, is the coefficient of innovation for the $i-th$ age group.
- C_i, $i = 1, \ldots, 6$, is the e-commerce influence coefficient of the $i-th$ age group on the other age groups.

- S_j, $j = 1, \ldots, 6$, is the susceptibility coefficient to the $j-th$ age group to be influenced by the use of e-commerce by individuals in other age groups.
- $Q_{ij} = C_i S_j$, $i, j = 1, \ldots, 6$, represents the transmission rate of the use of e-commerce by encounters between individuals of the $i-th$ group that have already adopted the technology and the individuals of the $j-th$ who have not adopted it yet.

Furthermore, we will consider the following assumptions:

- Let us assume homogeneous population mixing, i.e., each individual can contact with any other individual.
- Let us consider that people of 16 years old enter into the system to $Y_1(t)$ with a rate given by $\mu \dfrac{Y_1(t)}{N_1(t)+Y_1(t)}$ or $N_1(t)$ with a rate $\mu \dfrac{N_1(t)}{N_1(t)+Y_1(t)}$ if they have already adopted the use of e-commerce or not, respectively.
- There are two ways for that an individual belonging to $N_i(t)$ transits to $Y_i(t)$: the first possibility is by means of the innovation coefficient. It is given by $P_i N_i(t)$, $i = 1, \ldots, 6$. The second way is by means of the imitation coefficient. It is given by a nonlinear term modeled by $N_j(t) \sum_{i=1}^{6} Q_{ji} Y_i(t)$, $j = 1, \ldots, 6$.

Under the above assumptions, our age-structured mathematical diffusion model is based on the nonlinear system of ordinary differential equations given by (3.1)-(3.4). The Figure 1 shows a compartmental representation of this system.

$$N_1'(t) = \mu \frac{N_1(t)}{N_1(t)+Y_1(t)} + (-c_1 - d_1 - P_1) N_1(t) - N_1(t) \sum_{j=1}^{6} Q_{1j} Y_j, \tag{3.1}$$

$$Y_1'(t) = \mu \frac{Y_1(t)}{N_1(t)+Y_1(t)} + P_1 N_1(t) + (-c_1 - d_1) Y_1(t) + N_1(t) \sum_{j=1}^{6} Q_{1j} Y_j, \tag{3.2}$$

$$N_i'(t) = c_{i-1} N_{i-1}(t) + (-c_i - d_i - P_i) N_i(t) - N_i(t) \sum_{j=1}^{6} Q_{ij} Y_j, \quad 2 \le i \le 6, \tag{3.3}$$

$$Y_i'(t) = c_{i-1} Y_{i-1}(t) + P_i N_i(t) + (-c_i - d_i) Y_i(t) + N_i(t) \sum_{j=1}^{6} Q_{ij} Y_j, \; 2 \le i \le 6. \tag{3.4}$$

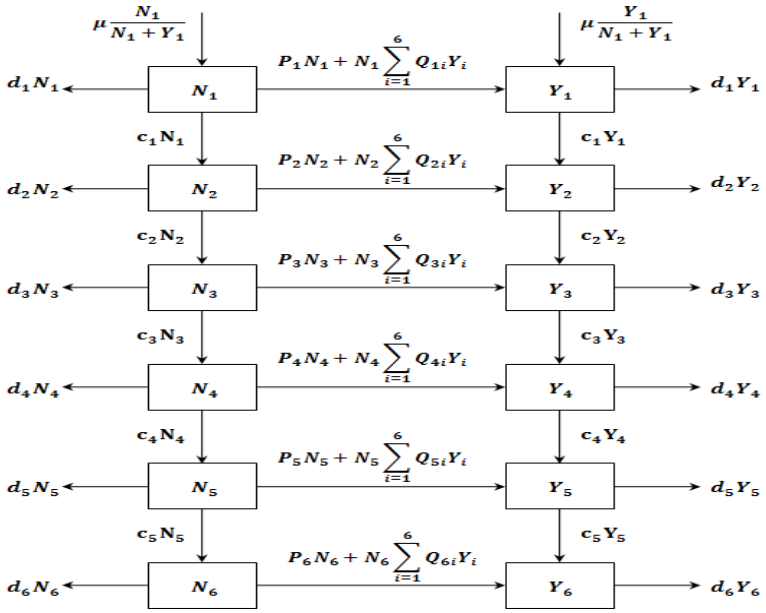


Figure 1: Diagram of the compartmental age-structured diffusion model for the e-commerce in Spain, (3.1)-(3.4).

4 Fitting and forecasting

4.1 Model fitting

Taking data in Table 1, let us fit data with the model (3.1)-(3.4). As initial conditions of the model (3.1)-(3.4) we have taken the year 2000 (corresponding to $t = 0$) as the starting of the e-commerce in Spain [8]. We have assigned the values $Y_2(2000) = Y_3(2000) = 0.01$ and zero otherwise because in general people belong to groups G_2 and G_3 were more active with purchases on internet.

In order to compute the best fitting, we carried out computations with *Mathematica*®[6] and we implemented the function:

$$E : \mathbb{R}^{18} \longrightarrow \mathbb{R}^{+} \cup \{0\},$$

which variables are P_i, C_i and S_j with $i, j = 1, \ldots, 6$ such that:

i. Solve numerically (*NDSolve[]*) the system of differential equations (3.1)-(3.4) previously described.

ii. Taking data in Table 1 evaluate the computed numerical solution for each subpopulation N_i, $i = 1, \ldots, 6$ and Y_i, $i = 1, \ldots, 6$.

iii. Compute the least square error between the values obtained in step 2 and percentage of Spanish people who have bought using internet from 2006 to 2008, (see Table 1).

Function E takes values in $\mathbb{R}^{18}$ and returns a positive real number. Hence, we can try to minimize this function using the Nelder-Mead algorithm [7], that does not need the computation of any derivative or gradient, which is impossible to know in this case.

Next, in order to find a global minimum we chose randomly $10,000$ 18-tuples $(P_1,\ldots,P_6,C_1,\ldots,C_6,S_1,\ldots,S_6)$ with $P_i,C_i,S_j \in [0,1]$, $i,j = 1,\ldots,6$, and for each one, the Nelder-Mead algorithm is applied. We stored all the minima obtained and, among them, the non-negative values of P_i, C_i and S_j with $i,j = 1,\ldots,6$, that minimize the function E are given in Table 2. The value of the function in the global minimum, that is, the least square error, is 0.0035.

A greater intensity in Figure 2 indicates greater value of imitation coefficient. From Table 2 and Figure 2, we can obtain the following conclusions: the obtained results for groups G_5 and G_6 can be explained because they are not familiarized with this technology and, for G_1, because in general they do not have a significative purchasing power that can influence other groups. People belonging to groups G_2 and G_3 are not influenced by other groups since they perform their opinion about the e-commerce mainly from advertising and purchasing campaigns (see P_2 and P_3 in Table 2). On the other hand groups G_2-G_4 exert an important influence group G_1 becuase they are older people that have more experience with the use of this technology and even in many cases they have a familiar relationship with people of group G_1. Finally, groups G_2 and G_4 also influence over groups G_3 and G_4 because people aged between 45-64 are predisposed to be convinced by people of its own group and people aged 25-34 who have purchasing power as well as they can have a good technological skills.

Table 2: Values of the coefficients P_i, C_i, $i = 1,\ldots,6$ and S_j, $j = 1,\ldots,6$ obtained from fitting the model (3.1)-(3.4) with data of Table 1.

Parameter	Estimation	Parameter	Estimation
P_1	0.04651	P_4	0
P_2	0.02635	P_5	0.00489
P_3	0.02568	P_6	0.00045
C_1	0.03877	C_4	0.49476
C_2	0.67826	C_5	0.05316
C_3	0.20927	C_6	0.00218
S_1	1.21112	S_4	0.91891
S_2	0.14628	S_5	0.9732
S_3	0.15495	S_6	0.6346

4.2 Short-term e-commerce forecasting

We have determined the values of the parameters that provide the best fitting of the nonlinear system (3.1)-(3.4). This solution is represented in Figure 3. Note that despite having few data the obtained fit is acceptable. Moreover, our approach permits to forecast the future trends of this technology in the next few

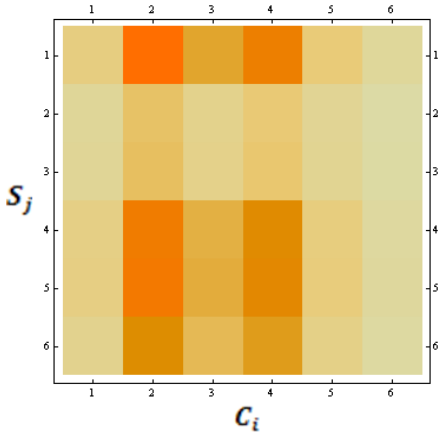


Figure 2: Representation for the imitation coefficients $Q_{ij} = C_i S_j$.

years. From Figure 3, the rise in the use of e-commerce in all age groups is expected.

5 Sensitivity analysis (SA)

5.1 Latin Hypercube Sampling (LHS)

Sensitivity analysis are used to determine the degree of uncertainty in model outcomes that is due to uncertainty in the input parameters. We use uniform distribution centered at deterministic parameters estimators by the absence of data to inform on the distribution for a given parameter [9], [10]. Then, the model can be simulated by sampling a single value from each parameter distribution. Latin Hypercube Sampling, a type of stratified Monte Carlo sampling, is a sophisticated and efficient method for achieving equitable sampling of all inputs parameters simultaneously [9], [10], [11].

We use standard LHS in SA. The procedure is the following [12]:

i. Let N denote the number of samples and K the number of random inputs (initial conditions and/or transmission parameters).

ii. P is an $N \times K$ matrix where each of the K columns is a random permutation of $1, \ldots, N$ and R is an $N \times K$ matrix of independent random numbers from the uniform $(0, 1)$ distribution.

iii. Compute S as $S = (P - R)/N$.

iv. Each element of $S = (s_{ij})$ is mapped according to its marginal probability distribution as $x_{ij} = F_j^{-1}(s_{ij})$, where F^{-1} represents the inverse of the

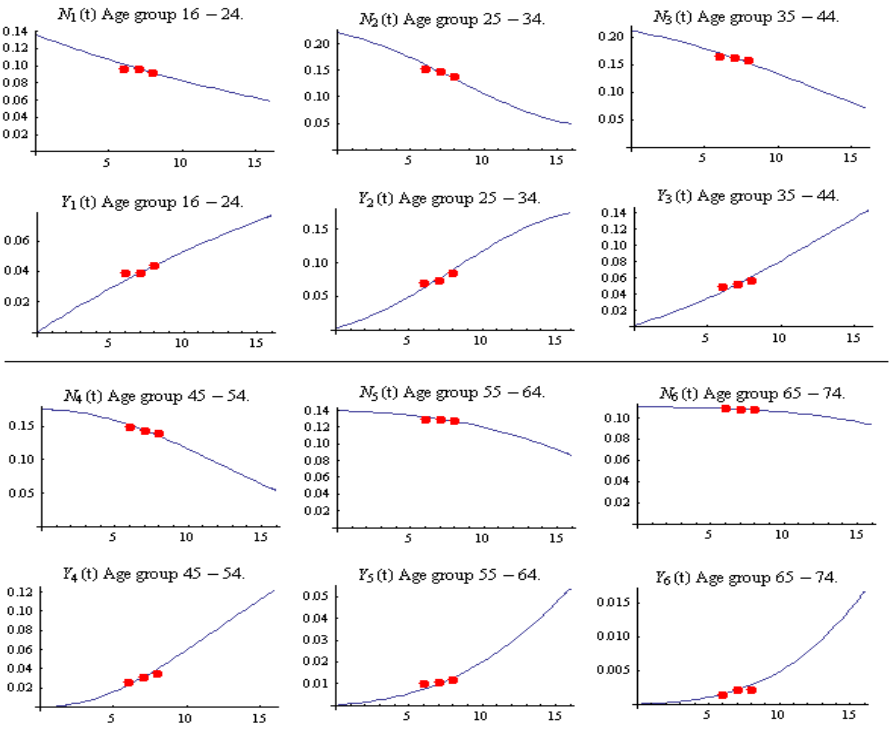


Figure 3: E-commerce fitting and forecasting in Spain from 2009 to 2015, $t = 0$ is year 2000 and $t = 15$ is year 2015.

cumulative probability distribution function for input j. A vector $x_i = (x_{i1}, x_{i2}, \ldots, x_{ik})$ contains input values (initials conditions and transmission parameters) for one deterministic run time.

Significativity tests can be performed to assess if a correlation is significative, i.e., if a Partial Rank Correlation Coefficients (PRCC) is significantly different from zero. Confidence level used is 95%. Thus, by this method we can quantify which parameters are the most important to the variability in the outcome of the model [9], [10], [11].

It is convenient to take a high enough number of samples N [9], in our case, we take $N = 2500$ realizations, apply the above mentioned procedures, compute the matrix outputs and perform sensitivity analysis by PRCC to study the sensitivity of the model. See Table 3 for the corresponding results at time instant $t = 2012$. The percentages of times the correlations is significative and the mean of correlations are very low, therefore, any parameter has a noticeable influence in the output.

Table 3: Results of the sensitivity analysis at $t = 2012$. In column "Mean of correlations" appears the mean of the 2500 computed correlations and in the column "percentage", the rate of times the correlation is significative.

Parameter	Mean of correlat.	Percentage	Parameter	Mean of correlat.	Percentage
P_1	-0.00004	0.2	P_4	0.00046	4.8
P_2	-0.00118	3.3	P_5	0.00017	4.6
P_3	0.00103	6.2	P_6	0.0003	5.2
C_1	-0.00026	5.6	C_4	0.00042	5.3
C_2	0.00088	5.1	C_5	0.00054	3.6
C_3	0.00012	6.3	C_6	0.00175	5.6
S_1	0.00008	5.6	S_4	0.00014	4.0
S_2	-0.00004	3.5	S_5	-0.00034	5.4
S_3	0.00037	5.2	S_6	-0.00064	3.5

6 Conclusions

We build an age-structured mathematical model for the diffusion of e-commerce in Spain that allows us to predict future trends of e-commerce in the next few years. Also, we performed a sensitivity analysis where any of the model parameters has a noticeable influence on the output model. The obtained results of the model may be a useful tool for helping e-commerce companies in order to adopt better investment strategies and show that investment must be global in our target population. In spite of some limitations in the current use of electronic commerce in Spain, the proposed diffusion model of this technology shows an increasing trend in the next few years.

References

[1] R.T. Frambach, An integrated model of organizational adoption and diffusion of innovations, European Journal of Marketing 27 (1993) 22-41.

[2] V. Mahajan, E. Muller, F.M. Bass, New product diffusion models in marketing: A review and directions for research, The Journal of Marketing 54 (1990) 1-26.

[3] D. Zhang, A. Ntoko. Mathematical model of technology diffusion in developing countries, Computational Methods in Decision-Making, Economics and Finance (2002) 526-539.

[4] *http://www.ine.es.*

[5] H.W. Hethcote, The mathematics of infectious diseases, Society For Industrial and Applied Mathematics 42 (2000) 599-653.

[6] *http://www.wolfram.com/products/mathematica.*

[7] J.A. Nelder, R. Mead, A simplex method for function minimization. The Computer Journal 7 (1964) 308-313.

[8] F. García Mas, Comercio y Firma Electrónicos: Análisis Jurídico de los Servicios de la Sociedad de la Información, (E-commerce and Electronic Signature: Law Analysis of Service of Information Society), 2^a ed., Editorial Lex Nova, 2004, pp. 103-110.

[9] S. Marino, I.B. Hogue, C.J. Ray, D.E. Kirschner, A methodology for performing global uncertainty and sensitivity analysis in systems biology, Journal of Theorical Biology 254 (2008) 178-196.

[10] A. Hoare, D.G. Regan, D.P Wilson, Sampling and sensitivity analysis tools (SaSAT) for computational modeling. Theoretical Biology and Medical Modeling, 5:4 (2008).

[11] S.M Blower, H. Dowlatabadi, Sensitivity and uncertainty analysis of complex models of disease transmission: and HIV model, as an example. International Statistical Review, 62(2) (1994) 229-243.

[12] A. Olsson, G. Sandberg, O. Dahlblom, On Latin hypercube sampling for structural reliability analysis. Structural Safety 25 (2003) 47-68.

[13] A.B. Owen, Controlling correlations in Latin Hypercube samples. Journal of American Statistical Association, 89(428) (1994) 1517-1522.

Numerical solutions of matrix differential models in engineering using higher-order matrix splines. *

E. Defez⋆, M.M. Tung⋆, J.J. Ibáñez♮, J. Sastre†

⋆ Instituto de Matemática Multidisciplinar

♮ Instituto de Aplicaciones de las Tecnologías de la Información

y de las Comunicaciones Avanzadas

† Instituto de Telecomunicaciones y Aplicaciones Multimedia

Universidad Politécnica de Valencia, Spain

{edefez, mtung}@imm.upv.es, jjibanez@dsic.upv.es, jorsasma@iteam.upv.es

December 10, 2009

1 Introduction

In this paper we propose a novel algorithm to tackle matrix differential equations of the first order. Matrix differential models are relevant for the description of many phenomena in physics and engineering, ranging from such diverse applications as control theory to game theory [1]. In particular, we will develop in this work a method for the numerical integration of first-order matrix differential equations with initial conditions. For different examples of this class of problems, we also refer to Ref. [2].

In their seminal work, Loscalzo and Talbot introduce spline function approximations for solutions of scalar differential equations [3]. These spline solutions $S(x)$ are of degree $m = 2, 3$ and continuity class $\mathcal{C}^{m-1}$. Recently, this

*This work has been supported by grant PAID-06-07/3283 from the Universidad Politécnica de Valencia, Spain.

method has been used in the resolution of other scalar problems as discussed in Ref. [4]. The corresponding generalizations to the matrix framework have been carried out in Refs. [5, 6].
Unfortunately, as detected by Loscalzo and Talbot, their scalar procedure is divergent when higher-order spline functions are used [3, p. 444–445]. They have explicitly shown by numerical computations that the equation $y' = y, y(0) = 1$ contains noticeable divergences for splines of order $m > 3$. However, our new method avoids these problems with divergences for splines $S(x)$ of order m but only require them to be of differentiability class $\mathcal{C}^1$.

This paper is organized as follows. In Section 2, we give a description of the proposed method and give details of the corresponding procedure. Section 3 concludes the discussion with some numerical examples for the scalar, vector and matrix cases, respectively.

2 Description of the method

As usual, let us consider the following first-order matrix problem

$$\left.\begin{array}{rcl} Y'(x) & = & f(x, Y(x)) \\ Y(a) & = & Y_a \end{array}\right\} \quad , \quad a \leq x \leq b, \tag{2.1}$$

where the unknown matrix is $Y(x) \in \mathbb{R}^{r\times q}$ with initial condition $Y_a \in \mathbb{R}^{r\times q}$. The matrix-valued function $f : [a, b] \times \mathbb{R}^{r\times q} \to \mathbb{R}^{r\times q}$ is of differentiability class $f \in \mathcal{C}^s(T)$, $s \geq 1$, with

$$T = \left\{(x, Y);\ a \leq x \leq b,\ Y \in \mathbb{R}^{r\times q}\right\}, \tag{2.2}$$

and f fulfills the global Lipschitz's condition

$$\|f(x, Y_1) - f(x, Y_2)\| \leq L \|Y_1 - Y_2\|,\ a \leq x \leq b, Y_1, Y_2 \in \mathbb{R}^{r\times q} \tag{2.3}$$

to guarantee the existence and uniqueness of the continuously differentiable solution $Y(x)$ of problem (2.1), see Ref. [7, p.99].

The partition of the interval $[a, b]$ shall be given by

$$\Delta_{[a,b]} = \{a = x_0 < x_1 < \ldots < x_n = b\},\ x_k = a + kh,\ k = 0, 1, \ldots, n, \tag{2.4}$$

where n is a positive integer with the corresponding step size $h = (b-a)/n$. We will construct in each subinterval $[a+kh, a+(k+1)h]$ a matrix spline $S(x)$ of order $m \in \mathbb{N}$ with $1 \leq m \leq s$, where s is the order of the differentiability class of f. This will approximate the solution of problem (2.1) so that $S(x) \in C^1([a,b])$.

In the first interval $[a, a+h]$, we define the matrix spline as

$$\begin{aligned} S_{\big|_{[a,a+h]}}(x) &= Y(a) + Y'(a)(x-a) + \frac{1}{2!}Y''(a)(x-a)^2 + \frac{1}{3!}Y^{(3)}(a)(x-a)^3 \\ &+ \cdots + \frac{1}{(m-1)!}Y^{(m-1)}(a)(x-a)^{m-1} + \frac{1}{m!}A_0(x-a)^m, \quad (2.5) \end{aligned}$$

where $A_0 \in \mathbb{R}^{r\times q}$ is a matrix parameter to be determined. It is straightforward to check

$$S_{\big|_{[a,a+h]}}(a) = Y(a), \quad S'_{\big|_{[a,a+h]}}(a) = Y'(a) = f(a, Y(a)),$$

and therefore the spline satisfies the differential equation Eq. (2.1) at $x = a$. We must obtain the values $Y''(a), Y^{(3)}(a), \ldots, Y^{(m-1)}(a)$, and A_0 in order to determine the matrix spline (2.5). To compute the second-order derivative $Y''(x)$, we follow the procedure and nomenclature as described in Ref. [6] to obtain

$$\begin{aligned} Y''(x) &= \frac{\partial f(x,Y(x))}{\partial x} + \Big[[\text{vec } f(x,Y(x))]^T \otimes I_r\Big] \frac{\partial f(x,Y(x))}{\partial \,\text{vec } Y(x)} \\ &= g_1(x, Y(x)), \quad (2.6) \end{aligned}$$

where $g_1 \in \mathcal{C}^{s-1}(T)$ and vec is the column-vector operator defined in Ref. [6, p. 658]. We are now in the position to evaluate $Y''(a) = g_1(a, Y(a))$ using (2.6).

Similarly, we can assume that $f \in \mathcal{C}^s(T)$ for $s \geq 2$. Then, the second partial derivatives of f exist and are continuous. This yields the third derivative:

$$\begin{aligned} Y^{(3)}(x) = \frac{\partial^2 f(x,Y(x))}{\partial x^2} &+ \Big([\text{vec } f(x,Y(x))]^T \otimes I_r\Big) \frac{\partial}{\partial x}\left(\frac{\partial f(x,Y(x))}{\partial \,\text{vec } Y(x)}\right) \\ &+ \left(\frac{\partial\,[\text{vec } f(x,Y(x))]^T}{\partial x} \otimes I_r\right) \frac{\partial f(x,Y(x))}{\partial \,\text{vec } Y(x)} \end{aligned}$$

$$
\begin{aligned}
&+\left([\text{vec } f(x,Y(x))]^T \otimes I_r\right) \frac{\partial}{\partial \text{vec } Y(x)} \left(\frac{\partial f(x,Y(x))}{\partial x}\right) \\
&+\left([\text{vec } f(x,Y(x))]^T \otimes I_r\right) \left(\frac{\partial [\text{vec } f(x,Y(x))]^T}{\partial \text{vec } Y(x)} \otimes I_r\right) \frac{\partial f(x,Y(x))}{\partial \text{vec } Y(x)} \\
&+\left([\text{vec } f(x,Y(x))]^T \otimes I_r\right) \left([\text{vec } f(x,Y(x))]^T \otimes I_{r^2q}\right) \frac{\partial^2 f(x,Y(x))}{(\partial \text{vec } Y(x))^2} \\
&= g_2(x,Y(x)) \in \mathcal{C}^{s-2}(T). \qquad (2.7)
\end{aligned}
$$

Now we can evaluate $Y^{(3)}(a) = g_2(a, Y(a))$ using (2.7). For all higher-order derivatives $Y^{(4)}(x), \ldots, Y^{(m-1)}(x)$ we proceed in like manner and calculate

$$
\left.
\begin{aligned}
Y^{(4)}(x) &= g_3(x,Y(x)) \in \mathcal{C}^{s-3}(T) \\
&\vdots \\
Y^{(m-1)}(x) &= g_{m-2}(x,Y(x)) \in \mathcal{C}^{s-(m-2)}(T)
\end{aligned}
\right\}. \qquad (2.8)
$$

A list of all these derivatives can be easily established by employing standard computer algebra systems. Substituting $x = a$ in (2.8), one gets $Y^{(4)}(a)$, ..., $Y^{(m-1)}(a)$. In summary, all matrix parameters of the spline which were to be determined are known, except for A_0. To determine A_0, we suppose that (2.5) is a solution of problem (2.1) at $x = a + h$, which gives

$$
S'_{\big|_{[a,a+h]}}(a+h) = f\left(a+h, S_{\big|_{[a,a+h]}}(a+h)\right). \qquad (2.9)
$$

Next, we obtain from (2.9) the matrix equation with only one unknown A_0:

$$
\begin{aligned}
A_0 &= \frac{(m-1)!}{h^{m-1}} \left[f\left(a+h, Y(a) + Y'(a)h + \cdots + \frac{h^{m-1}}{(m-1)!} Y^{(m-1)}(a) + \frac{h^m}{m!} A_0\right)\right. \\
&- \left. Y'(a) - Y''(a)h - \frac{1}{2} Y^{(3)}(a)h^2 + \cdots + \frac{1}{(m-2)!} Y^{(m-1)}(a) h^{m-2}\right]. \qquad (2.10)
\end{aligned}
$$

Assuming that the implicit matrix equation (2.10) has only one solution A_0, the matrix spline (2.5) is totally determined in the interval $[a, a+h]$.

In the following interval $[a+h, a+2h]$, the matrix spline takes the form

$$
\begin{aligned}
&S_{\big|_{[a+h,a+2h]}}(x) = S_{\big|_{[a,a+h]}}(a+h) + \overline{Y'(a+h)}(x-(a+h)) + \\
&\tfrac{1}{2!}\overline{Y''(a+h)}(x-(a+h))^2 + \cdots + \tfrac{1}{(m-1)!}\overline{Y^{(m-1)}(a+h)}(x-(a+h))^{m-1}
\end{aligned}
$$

$$+\tfrac{1}{m!}A_1(x-(a+h))^m, \tag{2.11}$$

where

$$\overline{Y'(a+h)} = f\left(a+h, S_{\big|_{[a,a+h]}}(a+h)\right), \tag{2.12}$$

and $\overline{Y''(a+h)}, \ldots, \overline{Y^{(m-1)}(a+h)}$ are the similar results obtained after evaluating the respective derivatives of $Y(x)$ using $S_{\big|_{[a,a+h]}}(a+h)$ in (2.6)–(2.8). In more compact form, we may write

$$\begin{aligned} \overline{Y''(a+h)} &= g_1\left(a+h, S_{\big|_{[a,a+h]}}(a+h)\right), \\ &\vdots \\ \overline{Y^{(m-1)}(a+h)} &= g_{m-2}\left(a+h, S_{\big|_{[a,a+h]}}(a+h)\right). \end{aligned} \tag{2.13}$$

Note that matrix spline $S(x)$ defined by (2.5) and (2.11) is of differentiability class $\mathcal{C}^1\left([a,a+h]\cup[a+h,a+2h]\right)$, contrary to the splines introduced by Loscalzo and Talbot [3], which were of class $\mathcal{C}^{m-1}\left([a,a+h]\cup[a+h,a+2h]\right)$. By construction, spline (2.11) satisfies the differential equation (2.1) at $x = a+h$. and all of its coefficients are determined with the exception of $A_1 \in \mathbb{R}^{r\times q}$.

The value of A_1 can be found by taking the spline (2.11) as a solution of (2.1) at point $x = a+2h$:

$$S'_{\big|_{[a+h,a+2h]}}(a+2h) = f\left(a+2h, S_{\big|_{[a+h,a+2h]}}(a+2h)\right).$$

An expansion yields the matrix equation with the only unknown A_1:

$$\begin{aligned} A_1 &= \frac{(m-1)!}{h^{m-1}}\left[f\left(a+2h, S_{\big|_{[a,a+h]}}(a+h) + \overline{Y'(a+h)}h + \frac{h^2}{2!}\overline{Y''(a+h)} + \right.\right. \\ &+ \cdots + \frac{h^{m-1}}{(m-1)!}\overline{Y^{(m-1)}(a+h)} + \left.\left.\frac{h^m}{m!}A_1\right) - \overline{Y'(a+h)} - \overline{Y''(a+h)}h \right. \\ &- \left. \cdots - \frac{1}{(m-2)!}\overline{Y^{(m-1)}(a+h)}h^{m-2}\right]. \end{aligned} \tag{2.14}$$

Let us assume that the matrix equation (2.14) has only one solution A_1. This way the spline is totally determined in the interval $[a+h, a+2h]$.

Iterating this process, we can construct the matrix spline taking $[a+(k-1)h, a+kh]$ as the last subinterval. For the next subinterval $[a+kh, a+(k+1)h]$, we define the corresponding matrix spline as

$$\begin{aligned} S_{\big|_{[a+kh,a+(k+1)h]}}(x) &= S_{\big|_{[a+(k-1)h,a+kh]}}(a+kh) + \overline{Y'(a+kh)}(x-(a+kh)) \\ &+\tfrac{1}{2!}\overline{Y''(a+kh)}(x-(a+kh))^2 + \cdots + \tfrac{1}{(m-1)!}\overline{Y^{(m-1)}(a+kh)}(x-(a+kh))^{m-1} \\ &+\tfrac{1}{m!}A_k(x-(a+kh))^m, \end{aligned} \tag{2.15}$$

where

$$\overline{Y'(a+kh)} = f\left(a+kh, S_{\big|_{[a+(k-1)h,a+kh]}}(a+kh)\right), \tag{2.16}$$

and in a similar manner one abbreviates

$$\begin{aligned} \overline{Y''(a+kh)} &= g_1\left(a+kh, S_{\big|_{[a+(k-1)h,a+kh]}}(a+kh)\right), \\ &\vdots \\ \overline{Y^{(m-1)}(a+kh)} &= g_{m-2}\left(a+kh, S_{\big|_{[a+(k-1)h,a+kh]}}(a+kh)\right). \end{aligned} \tag{2.17}$$

With this definition, the matrix spline $S(x) \in \mathcal{C}^1\left(\bigcup_{j=0}^{k}[a+jh, a+(j+1)h]\right)$ fulfills the differential equation (2.1) at point $x = a+kh$. As an additional requirement, we assume that $S_{\big|_{[a+kh,a+(k+1)h]}}(x)$ satisfies (2.1) at point $x = a+(k+1)h$:

$$S'_{\big|_{[a+kh,a+(k+1)h]}}(a+(k+1)h) = f\left(a+(k+1)h, S_{\big|_{[a+kh,a+(k+1)h]}}(a+(k+1)h)\right),$$

and expanding this expression gives

$$\begin{aligned} A_k &= \tfrac{(m-1)!}{h^{m-1}}\Big[f\left(a+(k+1)h, S_{\big|_{[a+kh,a+(k+1)h]}}(a+(k+1)h) + \overline{Y'(a+kh)}h\right. \\ &\quad \left. + \cdots + \tfrac{h^{m-1}}{(m-1)!}\overline{Y^{(m-1)}(a+kh)} + \tfrac{h^m}{m!}A_1\right) - \overline{Y'(a+kh)} - \overline{Y''(a+kh)}h \\ &\quad - \cdots - \tfrac{h^{m-2}}{(m-2)!}\overline{Y^{(m-1)}(a+kh)}\Big]. \end{aligned} \tag{2.18}$$

Observe that the final result (2.18) relates directly to equations (2.10) and (2.14), when setting $k = 0$ and $k = 1$. We will demonstrate that these equations have a unique solution using a fixed-point argument.

For a fixed h and k, we consider the matrix function $g : \mathbb{R}^{r\times q} \to \mathbb{R}^{r\times q}$ defined by

$$\begin{aligned} g(T) &= \tfrac{(m-1)!}{h^{m-1}}\Big[f\Big(a+(k+1)h, S_{\big|_{[a+kh,a+(k+1)h]}}(a+(k+1)h) + \overline{Y'(a+kh)}h \\ &\quad + \cdots + \tfrac{h^{m-1}}{(m-1)!}\overline{Y^{(m-1)}(a+kh)} + \tfrac{h^m}{m!}T\Big) - \overline{Y'(a+kh)} - \overline{Y''(a+kh)}h \\ &\quad - \cdots - \tfrac{h^{m-2}}{(m-2)!}\overline{Y^{(m-1)}(a+kh)}\ \Big]. \end{aligned} \tag{2.19}$$

Relation (2.18) holds if and only if $A_k = g(A_k)$, that is, if A_k is a fixed point for function $g(T)$. By using the definition (2.19) of g and applying the global Lipschitz's condition (2.3) for f, it immediately follows that

$$\|g(T_1) - g(T_2)\| \le \frac{Lh}{m}\|T_1 - T_2\|.$$

Taking $h < m/L$, the matrix function g is contractive. Therefore equation (2.18) has unique solutions A_k for $k = 0, 1, \ldots, n-1$, and the matrix spline is completely determined. In summary, we have proved the following theorem:

Theorem 2.1 *For the first-order matrix differential equation (2.1), let L be the corresponding Lipschitz constant defined by (2.3). We also consider the partition (2.4) with step size $h < m/L$. Then, the matrix spline $S(x)$ of order $m \in \mathbb{N}$ exists in each subinterval $[a+kh, a+(k+1)h]$, $k = 0, 1, \ldots, n-1$, as defined in the previous construction and is of class $\mathcal{C}^1[a,b]$.*

Observe that the so constructed splines have a global error of $O(h^{m-1})$, which follows from an analysis similar to Loscalzo and Talbot's work [3].

The approximate solution of (2.1) can be computed by means of matrix splines of order m in the interval $[a,b]$ with an error of the order $O(h^{m-1})$ under the conditions of Theorem 2.1. The procedure is as follows:

- Compute the functions $g_1(x, Y(x)), \ldots, g_{m-2}(x, Y(x))$ given by (2.6)–(2.8) to determine the constant $Y''(a), \ldots, Y^{(m-1)}(a)$. Choose $n > \frac{L(b-a)}{m}$ so that $h = \frac{(b-a)}{n}$ with the partition $\Delta_{[a,b]}$ defined by Eq. (2.4).
- Solve equation (2.10) to find A_0, and determine $S_{\big|_{[a,a+h]}}(x)$ of Eq. (2.5).
- Solve equations (2.18) iteratively for $k = 1, \ldots, n-1$ to find all A_k, and then compute the splines $S_{\big|_{[a+kh,a+(k+1)h]}}(x)$ according to Eq. (2.15).

In order to find A_k for $k = 0, 1, \ldots, n-1$, one may solve equations (2.10) and (2.18) either explicitly [8], or by employing an iterative method [9]. For example, we can consider the recursion relation $T^s_{l+1} = g(T^s_l)$. Here, T^s_0 is an arbitrary matrix in $\mathbb{R}^{r\times q}$ for $s = 0, 1, \ldots, n-1$, and $g(T)$ is given by (2.19).

3 Numerical Examples

3.1 A scalar test problem

This simple test problem is motivated by Loscalzo and Talbot's seminal work on scalar spline function approximation for ordinary differential equations [3]. Unfortunately, their otherwise very efficient method had the drawback to be divergent for higher degree spline functions ($m > 3$). Here, we will compare our procedure with their test case for the spline solution of $y' = y$ with initial condition $y(0) = 1$.

Figure 1 depicts the error of fourth-order spline solutions for the Loscalzo-Talbot problem which were constructed by our proposed method. Observe that for $h = 0.01$ the results already reach the accuracy of 10^{-14}, compared to the serious error of the conventional Loscalzo-Talbot method [3]. It also becomes clear that a further reduction in step size h does not necessarily improve the approximation.

It may be interesting to study the increasing quality of the approximation with higher-order splines. Figure 2 shows how the solutions improve by taking $m = 4, 5, 6$, respectively, with a constant step size $h = 0.1$.

3.2 A non-linear vector system

As a second example of our method, we choose the following vector differential system for the interval $x \in [0, 1]$, which is clearly non-linear:

$$\left.\begin{aligned} y_1'(x) &= -1 + e^x - \sin x + \sin\left(y_2(x)\right) \\ y_2'(x) &= \frac{1}{4 + y_1^2(x)} - \frac{1}{5 + e^{2\,x} + 2\,e^x\,\cos x - \sin^2 x} \end{aligned}\right\} \tag{3.1}$$

with the initial values

$$\left.\begin{array}{rcl} y_1(0) & = & 2 \\ y_2(0) & = & \frac{\pi}{2} \end{array}\right\} .$$

We can then rewrite the problem using vector notation $Y(x) = \begin{pmatrix} y_1(x) \\ y_2(x) \end{pmatrix}$ with $Y(0) = \begin{pmatrix} 2 \\ \frac{\pi}{2} \end{pmatrix}$ to obtain the nonlinear vector problem $Y'(x) = f(x, Y(x))$, where

$$f\left(x, Y(x)\right) = \begin{pmatrix} -1 + e^x - \sin x + \sin\left(y_2(x)\right) \\ \dfrac{1}{4 + y_1^2(x)} - \dfrac{1}{5 + e^{2\,x} + 2\,e^x\,\cos x - \sin^2 x} \end{pmatrix}. \quad (3.2)$$

According to Ref. [6] this problem has the exact solution $y_1(x) = e^x + \cos x$ and $y_2(x) = \pi/2$, and hence for this test case we will be able to assess the exact error of our numerical estimates. Our proposed method serves to construct the splines of fifth order for the problem given in Eq. (3.1). For this we require to calculate $Y''(x)$, $Y^{(3)}(x)$ and $Y^{(4)}(x)$, which in general is straightforward. We may derive $Y''(x) = \begin{pmatrix} y_1''(x) \\ y_2''(x) \end{pmatrix}$ using a computer algebra system such as *Mathematica*, which readily produces:

$$\left.\begin{array}{rcl} y_1''(x) & = & e^x - \cos\left(x\right) + \cos\left(y_2(x)\right) y_2'(x) \\ y_2''(x) & = & \dfrac{2e^{2x} + 2e^x\cos\left(x\right) - 2e^x\sin\left(x\right) - 2\cos\left(x\right)\sin\left(x\right)}{\left(5 + e^{2x} + 2e^x\cos\left(x\right) - \sin\left(x\right)^2\right)^2} - \dfrac{2y_1(x)y_1'(x)}{\left(4 + y_1(x)^2\right)^2} \end{array}\right\} . \quad (3.3)$$

Taking into account that $y_1(0) = 2, y_1'(0) = 1, y_2(0) = \frac{\pi}{2}$, and $y_2'(0) = 0$, it follows by Eq. (3.3) that $Y''(0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. We similarly calculate the third-

order derivative $Y^{(3)}(x) = \begin{pmatrix} y_1^{(3)}(x) \\ y_2^{(3)}(x) \end{pmatrix}$ with components:

$$\left.\begin{aligned} y_1^{(3)}(x) &= e^x + \sin(x) - \sin(y_2(x))\left(y_2'(x)\right)^2 + \cos(y_2(x))y_2''(x) \\ y_2^{(3)}(x) &= -\frac{8\left(\cos(2x) - 2e^x\left(e^x - \sin(x)\right)\right)}{\left(9 + 2e^{2x} + 4e^x\cos(x) + \cos(2x)\right)^2} - \frac{64\left(e^x + \cos(x)\right)^2\left(e^x - \sin(x)\right)^2}{\left(9 + 2e^{2x} + 4e^x\cos(x) + \cos(2x)\right)^3} \\ &+ \frac{8\left(y_1(x)\right)^2\left(y_1'(x)\right)^2}{\left(4 + \left(y_1(x)\right)^2\right)^3} - \frac{2\left(y_1'(x)\right)^2}{\left(4 + \left(y_1(x)\right)^2\right)^2} - \frac{2y_1(x)y_1''(x)}{\left(4 + \left(y_1(x)\right)^2\right)^2} \end{aligned}\right\} \tag{3.4}$$

In like manner as before, we consider $y_1(0) = 2, y_1'(0) = 1, y_1''(0) = 0, y_2(0) = \pi/2, y_2'(0) = 0$, and $y_2''(0) = 0$ with (3.4) to deduce $Y^{(3)}(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. We may then easily derive $Y^{(4)}(x) = \begin{pmatrix} y_1^{(4)}(x) \\ y_2^{(4)}(x) \end{pmatrix}$ with components:

$$\left.\begin{aligned} y_1^{(4)}(x) &= e^x + \cos(x) - \cos(y_2(x))\left(y_2'(x)\right)^3 - 3\sin(y_2(x))y_2'(x)y_2''(x) + \cos(y_2(x))y_2^{(3)}(x) \\ y_2^{(4)}(x) &= \frac{6\left(2e^{2x} + 2e^x\cos(x) - 2e^x\sin(x) - 2\cos(x)\sin(x)\right)^3}{\left(5 + e^{2x} + 2e^x\cos(x) - \sin(x)^2\right)^4} - \frac{48y_1(x)^3y_1'(x)^3}{\left(4 + y_1(x)^2\right)^4} \\ &+ \frac{8e^{2x} - 4e^x\cos(x) - 4e^x\sin(x) + 8\cos(x)\sin(x)}{\left(5 + e^{2x} + 2e^x\cos(x) - \sin(x)^2\right)^2} + \frac{24\left(y_1(x)y_1'(x)^3 + y_1(x)^2y_1'(x)y_1''(x)\right)}{\left(4 + y_1(x)^2\right)^3} \\ &- \frac{6\left(2e^{2x} + 2e^x\cos(x) - 2e^x\sin(x) - 2\cos(x)\sin(x)\right)\left(4e^{2x} - 2\cos(x)^2 - 4e^x\sin(x) + 2\sin(x)^2\right)}{\left(5 + e^{2x} + 2e^x\cos(x) - \sin(x)^2\right)^3} \\ &- \frac{6y_1'(x)y_1''(x) - 2y_1(x)y_1^{(3)}(x)}{\left(4 + y_1(x)^2\right)^2} \end{aligned}\right\}$$

In the final step, it remains to substitute the known values $y_1(0) = 2, y_1'(0) = 1, y_1''(0) = 0, y_1'''(0) = 1, y_2(0) = \frac{\pi}{2}, y_2'(0) = 0, y_2''(0) = 0, y_2'''(0) = 0$, into the last expression to obtain $Y^{(4)} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$.

Also, it is not difficult to see that f, defined by (3.2), fulfills the global Lipschitz's condition

$$\|f(x,Y) - f(x,Z)\| \leq \|Y - Z\|, \quad 0 \leq x \leq 1, \quad Y, Z \in \mathbb{R}^2. \tag{3.5}$$

Comparing with the general form (2.3), we note that $L = 1$. Therefore, by Theorem 2.1 we need to take $h < 5$. In the following, for example we choose $h = 0.1$ and summarize the numerical results in Table 2. In each interval, we evaluated the difference between the estimates of our numerical approach and the exact solution, and then take the Fröbenius norm of this difference, following the procedure explained in Ref. [6]. Table 1 lists the maximum of these errors for each subinterval.

For the solution of the vector differential system (3.1), Figure 3 illustrates the approximation behavior of various splines of the fourth order ($m = 4$) with the different step sizes $h = 0.1, 0.01$, and $h = 0.001$. All vector splines lie well in the predicted range of Theorem 2.1 and provide excellent approximations for the problem at hand with the benefit of very low computational cost. Observe that at step size $h = 0.001$ the limit of machine precision is practically reached and explains the random fluctuations around 10^{-15}. Hence, it obviously is of lesser interest to obtain more accurate approximations for $m = 4$ and $h = 0.001$.

3.3 Sylvester matrix differential equation

In many areas of science and engineering linear matrix differential equations appear of the type

$$\left.\begin{array}{rcl} Y'(x) & = & A(x)Y(x) + Y(x)B(x) + C(x) \\ Y(a) & = & Y_a \end{array}\right\} \quad a \leq x \leq b, \tag{3.6}$$

where $Y(x), A(x), B(x), C(x) \in \mathbb{R}^{r\times r}$. The case of constant coefficients has been studied by several authors [10], whereas the variable-coefficient case has so far received little numerical treatment in the literature.

Following Ref. [6], we choose the following Sylvester problem (3.6) as a final

example:

$$\begin{aligned} A(x) &= \begin{pmatrix} 0 & xe^{-x} \\ x & 0 \end{pmatrix}, \quad B(x) = \begin{pmatrix} 0 & x \\ 0 & 0 \end{pmatrix}, \\ C(x) &= \begin{pmatrix} -e^{-x}(1+x^2) & -2e^{-x}x \\ 1-e^{-x}x & -x^2 \end{pmatrix} \\ Y(0) &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad Y(x) \in \mathbb{R}^{2\times 2},\ 0 \le x \le 1. \end{aligned} \tag{3.7}$$

According to [6] we know that this problem has the exact solution

$$Y(x) = \begin{pmatrix} e^{-x} & 0 \\ x & 1 \end{pmatrix}$$

with the Lipschitz constant $L = 2$. The higher-order derivatives of $Y(x)$ are required for the construction of the spline approximation and can be readily obtained.

For splines of the fifth order ($m = 5$), we take $n = 10$ partitions and $h = 0.1$. The results are summarized in Table 3, where the numerical estimates have been rounded to the sixth relevant digit. In Table 4, we evaluated the difference between the estimates of our numerical approach and the exact solution, and then take the Fröbenius norm of this difference. The maximum of these errors are indicated for each subinterval.

For the solution of the Sylvester matrix problem (3.6), Figure 4 depicts the approximation behavior of various splines of the fifth order ($m = 5$) with the different step sizes $h = 0.1, 0.01$, and $h = 0.001$. As before, all matrix splines lie well in the predicted range of Theorem 2.1. It becomes evident that the splines for step sizes $h = 0.01$ and $h = 0.001$ are almost indistinguishable and reach the same precision of almost 10^{-14}.

We also carried out the computations for the sixth order matrix splines ($m = 6$) with the step sizes $h = 0.1, 0.01$, and $h = 0.001$, and as expected, we could observe that $h = 0.01$ yields an accuracy close to machine precision. Interestingly, higher step sizes do not improve these approximations—the quality of approximation indeed deteriorates due to the accumulation of rounding errors.

4 Conclusions

This work focuses on the presentation of a new method for the numerical integration of first-order matrix differential equations of the type $Y'(x) = f(x, Y(x))$ in the interval $[a, b]$ using higher-order matrix splines ($m > 3$). Contrary to existing spline methods in the literature, this new algorithm only requires first-order derivatives for the construction of the splines to provide a continuous approximation of order $O(h^{m-1})$. Additionally, our method is well-suited for implementation on numerical and/or symbolical computer systems.

For an explicit demonstration of our proposed method and its advantages over existing conventional methods, we discussed three numerical test cases with excellent results. It is hoped that this new approach to approximating matrix differential models will motivate and open up alternative avenues to tackle different related problems in science and engineering.

References

[1] G. Freiling and A. Hochhaus, On a class of rational matrix differential equations arising in stochastic control, Linear Algebra Appl. 379 (2004) 43–68.

[2] U.M. Ascher, R.M.M. Mattheij, R.D. Russell, Numerical solutions of boundary value problems for ordinary differential equations, Prentice Hall, New Jersey, 1988.

[3] F.R. Loscalzo, T.D. Talbot, Spline function approximations for solutions of ordinary differential equations, SIAM J. Numer. Anal. 4(3) (1967) 433–445.

[4] E.A. Al-Said, M.A. Noor, Cubic splines method for a system of third-order boundary value problems, Appl. Math. Comput. 142 (2003) 195–204.

[5] E. Defez, L. Soler, A. Hervás, C. Santamaría, Numerical solutions of matrix differential models using cubic matrix splines, Comput. Math. Appl. 50 (2005) 693–699.

[6] E. Defez, A. Hervás, L. Soler, M.M. Tung, Numerical solutions of matrix differential models using cubic matrix splines II, Math. Comp. Modelling 46 (2007) 657–669.

[7] T.M. Flett, Differential Analysis, Cambridge University Press, 1980.

[8] P. Lancaster, Explicit solutions of linear matrix equations, SIAM Review 12 (1970) 544–566.

[9] J.M. Ortega, W.C. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, 1972.

[10] A.Y Barraud, Noveaux développements sur la résolution numérique de $X' = AX + XB + C; X(0) = C$, R.A.I.R.O. 16(4) (1982) 341–356.

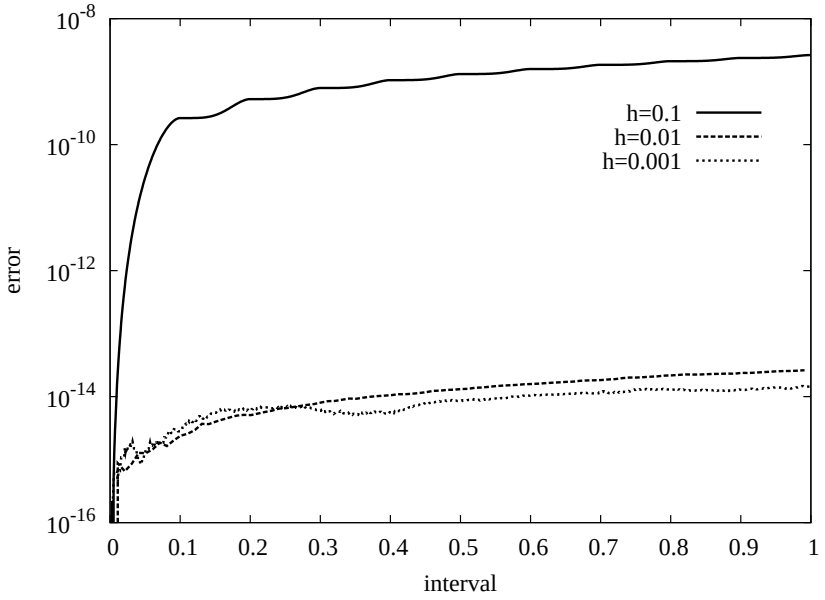


Figure 1: Error for the Loscalzo-Talbot problem with splines of fourth order ($m = 4$) using our proposed method for various step sizes.

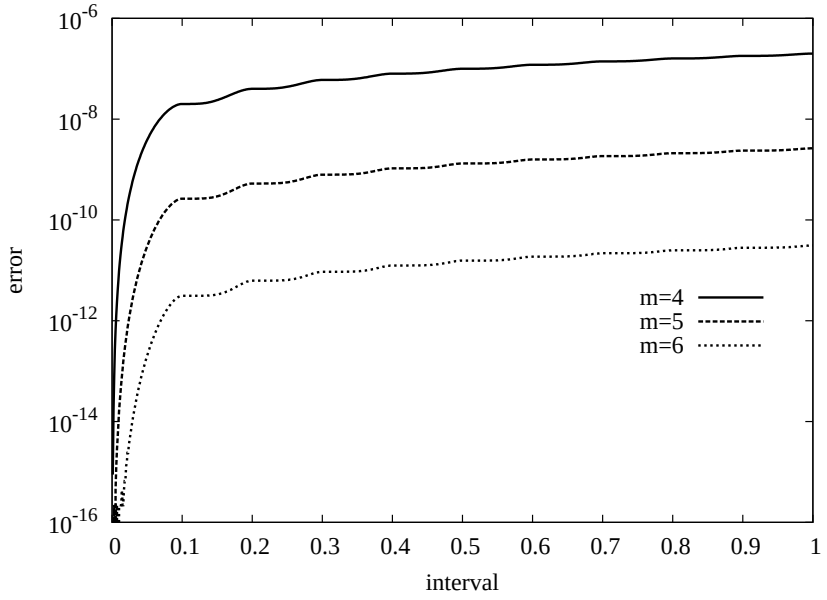


Figure 2: Errors for increasing spline orders ($m = 4, 5, 6$) solving the Loscalzo-Talbot problem. The step size is constant ($h = 0.1$).

Interval	$[0, 0.1]$	$[0.1, 0.2]$	$[0.2, 0.3]$	$[0.3, 0.4]$	$[0.4, 0.5]$
Max. error	8.2362×10^{-12}	4.8717×10^{-11}	1.27357×10^{-10}	2.50353×10^{-10}	4.24194×10^{-10}
Interval	$[0.5, 0.6]$	$[0.6, 0.7]$	$[0.7, 0.8]$	$[0.8, 0.9]$	$[0.9, 1.0]$
Max. error	6.55672×10^{-10}	9.51896×10^{-10}	1.32033×10^{-9}	1.7688×10^{-9}	2.30555×10^{-9}

Table 1: Approximation error for vector problem (3.1).

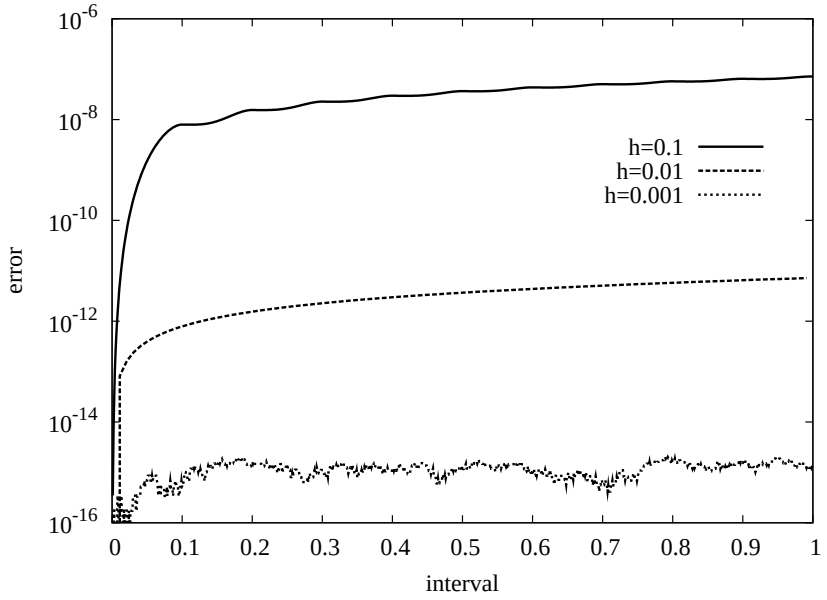


Figure 3: Representing the 2-norm error for the vector differential system (3.1) using splines of fourth order ($m = 4$).

Interval	Approximation
$[0, 0.1]$	$\begin{pmatrix} 2. + x + 0.166667x^3 + 0.0833333x^4 + 0.00833619x^5 \\ 1.5708 \end{pmatrix}$
$[0.1, 0.2]$	$\begin{pmatrix} 2. + 1.x - 3.98676 \times 10^{-7}x^2 + 0.166671x^3 + 0.0833075x^4 + 0.0083996x^5 \\ 1.5708 + 1.02341 \times 10^{-9}x^2 - 9.53254 \times 10^{-9}x^3 + 4.27272 \times 10^{-8}x^4 - 7.27159 \times 10^{-8}x^5 \end{pmatrix}$
$[0.2, 0.3]$	$\begin{pmatrix} 2. + 1.x - 9.78808 \times 10^{-6}x^2 + 0.166723x^3 + 0.0831609x^4 + 0.00856703x^5 \\ 1.5708 - 2.80891 \times 10^{-9}x + 2.68447 \times 10^{-8}x^2 - 1.2696 \times 10^{-7}x^3 + 2.96285 \times 10^{-7}x^4 - 2.72203 \times 10^{-7}x^5 \end{pmatrix}$
$[0.3, 0.4]$	$\begin{pmatrix} 2. + 1.00001x - 0.000070073x^2 + 0.166941x^3 + 0.0827649x^4 + 0.00885657x^5 \\ 1.5708 - 2.3641 \times 10^{-8}x + 1.51773 \times 10^{-7}x^2 - 4.84484 \times 10^{-7}x^3 + 7.68203 \times 10^{-7}x^4 - 4.83576 \times 10^{-7}x^5 \end{pmatrix}$
$[0.4, 0.5]$	$\begin{pmatrix} 2. + 1.00005x - 0.000295117x^2 + 0.167541x^3 + 0.0819626x^4 + 0.00928717x^5 \\ 1.5708 - 1.04291 \times 10^{-7}x + 5.05234 \times 10^{-7}x^2 - 1.21958 \times 10^{-6}x^3 + 1.46618 \times 10^{-6}x^4 - 7.01984 \times 10^{-7}x^5 \end{pmatrix}$
$[0.5, 0.6]$	$\begin{pmatrix} 1.99998 + 1.0002x - 0.000921692x^2 + 0.168862x^3 + 0.080566x^4 + 0.00987867x^5 \\ 1.5708 - 3.25859 \times 10^{-7}x + 1.26869 \times 10^{-6}x^2 - 2.46395 \times 10^{-6}x^3 + 2.3864 \times 10^{-6}x^4 - 9.21882 \times 10^{-7}x^5 \end{pmatrix}$
$[0.6, 0.7]$	$\begin{pmatrix} 1.99993 + 1.00062x - 0.00237386x^2 + 0.171395x^3 + 0.0783551x^4 + 0.0106518x^5 \\ 1.5708 - 8.18293 \times 10^{-7}x + 2.66421 \times 10^{-6}x^2 - 4.32971 \times 10^{-6}x^3 + 3.51164 \times 10^{-6}x^4 - 1.13697 \times 10^{-6}x^5 \end{pmatrix}$
$[0.7, 0.8]$	$\begin{pmatrix} 1.9998 + 1.00162x - 0.00534205x^2 + 0.175805x^3 + 0.0750753x^4 + 0.0116284x^5 \\ 1.5708 - 1.76297 \times 10^{-6}x + 4.93332 \times 10^{-6}x^2 - 6.89355 \times 10^{-6}x^3 + 4.80962 \times 10^{-6}x^4 - 1.34027 \times 10^{-6}x^5 \end{pmatrix}$
$[0.8, 0.9]$	$\begin{pmatrix} 1.99947 + 1.00376x - 0.0108784x^2 + 0.18297x^3 + 0.0704351x^4 + 0.0128313x^5 \\ 1.5708 - 3.38486 \times 10^{-6}x + 8.30591 \times 10^{-6}x^2 - 0.0000101804x^3 + 6.23218 \times 10^{-6}x^4 - 1.52432 \times 10^{-6}x^5 \end{pmatrix}$
$[0.9, 1.0]$	$\begin{pmatrix} 1.99873 + 1.00796x - 0.0205098x^2 + 0.19401x^3 + 0.0641039x^4 + 0.0142844x^5 \\ 1.5708 - 5.93162 \times 10^{-6}x + 0.000012961x^2 - 0.0000141487x^3 + 7.71598 \times 10^{-6}x^4 - 1.68162 \times 10^{-6}x^5 \end{pmatrix}$

Table 2: Vector approximation for system (3.1) in the interval $[0, 1]$.

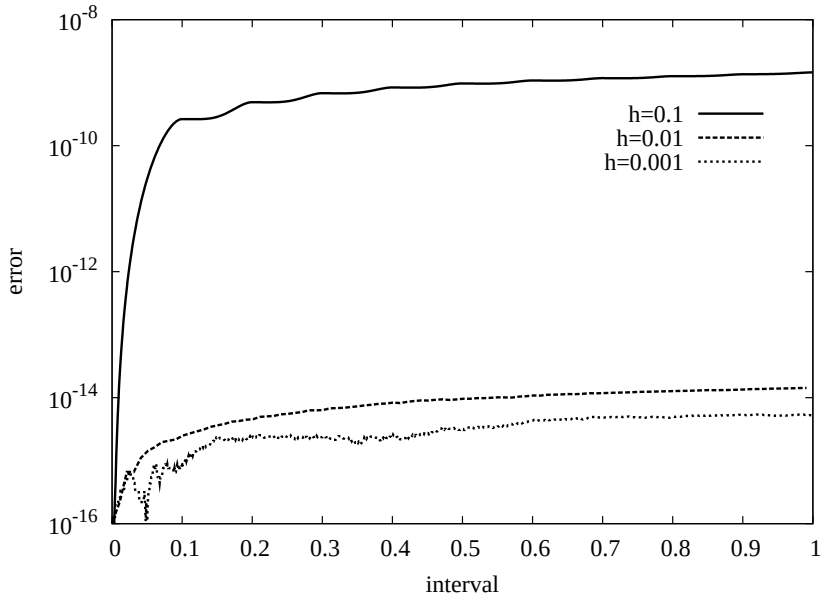


Figure 4: Representing the 2-norm error for the Sylvester matrix differential equation (3.6) using splines of fourth order ($m = 4$).

Interval	Approximation
$[0, 0.1]$	$\begin{pmatrix} 1. - 1.x + 0.5x^2 - 0.166667x^3 + 0.0416667x^4 - 0.00816941x^5 & 0. \\ x. & 1. \end{pmatrix}$
$[0.1, 0.2]$	$\begin{pmatrix} 1. - 1.x + 0.499997x^2 - 0.166626x^3 + 0.0413976x^4 - 0.00739198x^5 & 0. \\ 1.x & 1. \end{pmatrix}$
$[0.2, 0.3]$	$\begin{pmatrix} 1. - 0.999997x + 0.499961x^2 - 0.166422x^3 + 0.0408023x^4 - 0.00668854x^5 & 0. \\ 1.x & 1. \end{pmatrix}$
$[0.3, 0.4]$	$\begin{pmatrix} 0.999999 - 0.999979x + 0.499834x^2 - 0.165957x^3 + 0.0399455x^4 - 0.00605204x^5 & 0. \\ 1.x & 1. \end{pmatrix}$
$[0.4, 0.5]$	$\begin{pmatrix} 0.999995 - 0.999925x + 0.499542x^2 - 0.16517x^3 + 0.0388822x^4 - 0.00547612x^5 & 0. \\ 1.x & 1. \end{pmatrix}$
$[0.5, 0.6]$	$\begin{pmatrix} 0.999983 - 0.999797x + 0.499x^2 - 0.16402x^3 + 0.0376596x^4 - 0.00495499x^5 & 0 \\ 1.x & 1. \end{pmatrix}$
$[0.6, 0.7]$	$\begin{pmatrix} 0.999954 - 0.999547x + 0.498127x^2 - 0.16249x^3 + 0.0363175x^4 - 0.00448346x^5 & 0 \\ 1.x & 1. \end{pmatrix}$
$[0.7, 0.8]$	$\begin{pmatrix} 0.999896 - 0.999117x + 0.496844x^2 - 0.160578x^3 + 0.0348899x^4 - 0.00405681x^5 & 0 \\ 1.x & 1. \end{pmatrix}$
$[0.8, 0.9]$	$\begin{pmatrix} 0.999792 - 0.998438x + 0.495083x^2 - 0.158291x^3 + 0.033405x^4 - 0.00367075x^5 & 0 \\ 1.x & 1. \end{pmatrix}$
$[0.9, 1.0]$	$\begin{pmatrix} 0.999617 - 0.997437x + 0.492785x^2 - 0.155651x^3 + 0.0318868x^4 - 0.00332143x^5 & 0 \\ 1.x & 1. \end{pmatrix}$

Table 3: Approximation for the Sylvester matrix problem (3.6).

Interval	$[0, 0.1]$	$[0.1, 0.2]$	$[0.2, 0.3]$	$[0.3, 0.4]$	$[0.4, 0.5]$
Max. error	2.6999×10^{-10}	5.1438×10^{-10}	7.36134×10^{-10}	9.38797×10^{-10}	1.1268×10^{-9}
Interval	$[0.5, 0.6]$	$[0.6, 0.7]$	$[0.7, 0.8]$	$[0.8, 0.9]$	$[0.9, 1.0]$
Max. error	1.30572×10^{-9}	1.48252×10^{-9}	1.66579×10^{-9}	1.86603×10^{-9}	2.09601×10^{-9}

Table 4: Approximation error for the Sylvester matrix problem (3.6).

Analytic Hierarchy Process for Assessing Externalities in Water Leakage Management

X. Delgado-Galván[*], R. Pérez-García[*], J. Izquierdo[*], J. Mora-Rodríguez[†]

[*]Instituto Matemático Multidisciplinar
[†]Departamento de Ingeniería Hidráulica y Medio Ambiente
Universidad Politécnica de Valencia
Camino de Vera s/n, 46022, Valencia, Spain

1 Introduction

One of the main challenges for water supply managers is the minimization of water losses caused by leakage. In pursuance of this aim, substantial sums of money are invested every year in leak detection and repairs.

By considering exclusively an economic point of view, this investment is usually balanced by the benefits derived from the use of the recovered water. Nevertheless, this scheme does not reflect the whole dimension of the profit-earning capacity of repairing leaks. The associated benefits may include more aspects than just the economic value of the recovered water. In this paper, we examine a new economic assessment approach that includes all the costs incurred by the existence of leaks and the benefits derived from their control. These are mainly social and environmental costs and benefits, which are called externalities from an economic point of view. Their inclusion renders the assessment of leaks more realistic, but raises important problems. The main problem derives from the fact that comparisons with regard to property will only work for properties with well-defined scales of measurement. Nevertheless, direct comparisons are necessary to establish measurements for intangible properties that have no scales of measurement.

The analytic hierarchy process (AHP) provides a useful way to establish relative scales. The various factors are arranged in a hierarchy or a network structure, and measured according to the criteria represented within these structures. The leading eigenvalue and the principal (Perron) eigenvector of the matrix of criteria [1,2] provide, taking into account the properties this matrix exhibits, the necessary information to deal with complex decisions in the context of benefits, opportunities, costs, and risks [3].

In this paper, a comparison between active leakage control (ALC) and passive leakage control (PLC) is considered. The obtained results show that the inclusion of social and environmental costs clearly points in the direction of ALC as the best alternative in leakage control. As a conclusion, it can be stated that water supply managers and authorities should, accordingly, shift direction from purely economic policies towards new social and environmental policies.

2 Analytic hierarchy process; application to leakage control

The method of AHP initially breaks a problem into a certain hierarchical form. As a consequence it is easy to identify levels or hierarchies that discriminate objectives, criteria, and alternatives. Then, the people involved in the process compare the criteria and alternatives in pairs, make judgments, and create a relative scale of those judgements.

Finally, the main target is accomplished – a synthesis of priorities to determine the best decision to make.

The main objective is the minimization of water loss by means of suitable leakage control. The criteria used to decide on the eventual alternative(s) are manifold. We consider here the following: (a) planning development cost and its implementation; (b) damage to properties and other service networks; (c) effects of supply disruptions; and (d) inconveniences caused by closed or restricted streets. Regarding alternatives for leakage management, a comparison between active leakage control (ALC) and passive leakage control (PLC) is considered. ALC is the process of undertaking actions in a distribution system or an individual hydrometric district area, to locate and repair detectable leaks that have not been reported. PLC refers to repairing only reported or evident leaks.

In this case, the assessment is made from the point of view of the company managers; and aims to provide a base of support for decision-making. Referring to leakage management, the company decides if it wishes to carry out ALC or PLC.

2.1 The matrix of criteria

The hierarchical tree used to evaluate the alternatives, can be observed in Figure 1.

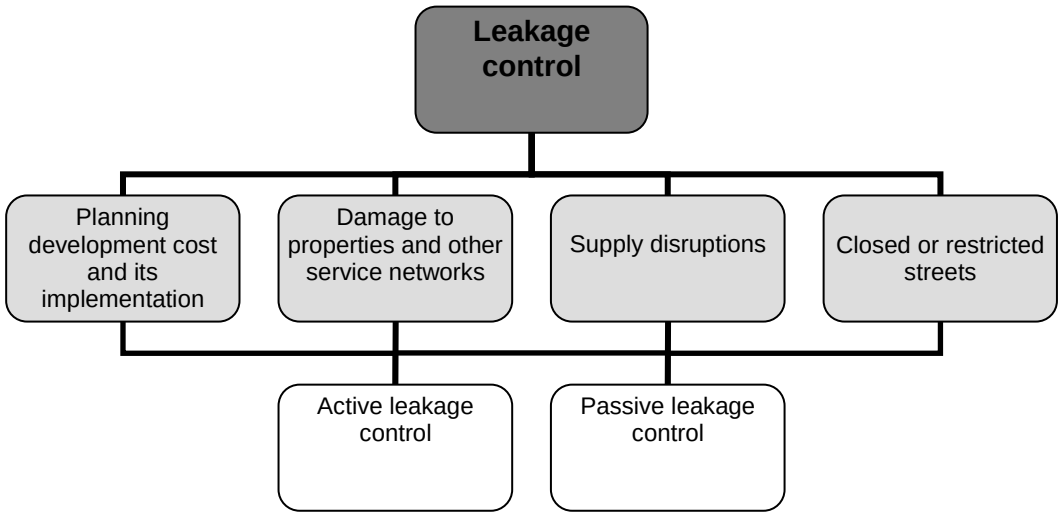


Figure 1. Hierarchical structure for leakage management alternatives

When using judgment to estimate dominance in making comparisons between two alternatives, a single number that is drawn from a fundamental scale of absolute numbers is assigned. Judgment must be based on knowledge, that is to say, on data. One method to collect data is by directly interviewing managers; and this data can be supplemented with different methods.

Table 1 summarizes an option to give values to opinions or verbal observations which can be expressed about the elements into which the problem has been divided. The scale developed by Saaty [4] enables comparisons to be made between pairs of components.

Table 1. Saaty numerical scale for pairwise comparisons in AHP

Judgment term	Saaty scale (a_{ij})
Absolute preference (element *i* over element *j*)	9
Very strong preference (*i* over *j*)	7
Strong preference (*i* over *j*)	5
Weak preference (*i* over *j*)	3
Indifference of *i* and *j*	1
Weak preference (*j* over *i*)	1/3
Strong preference (*j* over *i*)	1/5
Very strong preference (*j* over *i*)	1/7
Absolute preference (*j* over *i*)	1/9

The AHP starts by constructing a square matrix $A_{nxn} = (a_{ij})$ with positive entries, according to the numerical scale in Table 1, where a_{ij} represents the comparison between element *i* and element *j* from the values of the fundamental scale.

Let us take the point of view of a supply company manager who seeks the best option among the two alternatives for leakage management – while considering the four abovementioned criteria. Let us suppose that, upon evaluation, the plausible matrix in Figure 2 is obtained.

j / ***i***	Planning development cost and its implementation	Damage to properties and other service networks	Supply disruptions	Closed or restricted streets
Planning development cost and its implementation	1	1/3	5	3
Damage to properties and other service networks	3	1	7	5
Supply disruptions	1/5	1/7	1	1/2
Closed or restricted streets	1/3	1/5	2	1

Figure 2. Comparison matrix of criteria to evaluate leakage management alternatives

2.2 Properties of the matrix of criteria

Matrices such as the one in Figure 2 obtained by using the scale in Table 1 are positive matrices (matrices with only positive entries) and exhibit two clear properties:

(i) Homogeneity: if elements *i* and *j* are considered equally important, then $a_{ij} = a_{ji} = 1$.

(ii) Reciprocity: $a_{ji} = 1/a_{ij}$, for all i, j.

Homogeneity is crucial to compare elements of the same class, or with similar characteristics. In the case of comparisons between pairs, homogeneity is expressed by values

of 1 when comparing two elements with no clear importance of one over the other. In particular, all the elements in the diagonal are equal to 1.

Reciprocity emerges from the mutual dependency of one element on another. It refers to a comparison of an element i with an element j, for which the reciprocal comparison takes the reciprocal value.

A third property, namely consistency, should theoretically be desirable for matrix A.

(iii) A positive $n \times n$ matrix is consistent if $a_{ij}a_{jk} = a_{ik}$, for $i, j, k = 1, ..., n$.

Consistency expresses the coherence that should (perhaps) exist between judgments about the elements of a set. Since preferences are expressed in a subjective manner it is reasonable (and arguably even desirable) for some kind of incoherence to exist. One source of inconsistency may arise from ordinal intransitivity (x is preferred to y and y to z, but z is preferred to x). This kind of inconsistency usually originates in errors. Another example derives from inaccuracy in estimations: if x is indifferent with respect to y – the corresponding a_{ij} equals 1 –, and y is twice preferred to z –the corresponding $a_{jk} = 2$ –, x should be then be twice preferred to z as well. This forces the corresponding a_{ik} to equal 2; nevertheless, the original judgment could give three, or another amount, as the measure for the preference of x over z.

For a consistent matrix the following properties hold.

a) A has rank 1. In fact, each column A_k is a multiple of any other A_i, as $A_k = a_{ik}A_i$, since $a_{jk} = a_{ji}a_{ik}$, $j = 1,..., n$, using the consistency definition.

b) As a consequence of (a), all the eigenvalues of A are 0, except for one.

c) Since the sum of the eigenvalues equals the trace of a matrix, the non-zero eigenvalue of A is n, which, consequently is the so-called principal or Perron eigenvalue of A.

d) Each column, A_i, of A is an eigenvector corresponding to the eigenvalue n, since $AA_i = a_{1i}A_1 + ... + a_{ni}A_n = a_{1i}a_{i1}A_i + ... + a_{ni}a_{in}A_i = nA_i$, because of reciprocity.

e) The corresponding unique (within a multiplicative constant) eigenvector may then be obtained by dividing any column by the sum of its entries.

As a consequence of (d), any column of A, after normalization using the sum of its entries, gives an eigenvector whose entries sum up to 1, thus representing a coherent scale of priorities used to relatively assess the considered alternatives, and called the *priority vector*.

In the general case, A is not consistent, because only estimates of the comparison values are known through numerical judgment. For most problems we can consider that estimates of these values by an expert are assumed to be small perturbations of the 'right' values. This implies small perturbation of the eigenvalues (see, for instance, [5]). Now, the problem to solve is the eigenvalue problem $Aw = \lambda_{max}w$, where λ_{max} is, according to the Perron-Frobenius theorem, the unique largest eigenvalue of A, which also gives the so-called Perron eigenvector, that is an estimate of Z, the priority vector.

By using the power method to calculate the largest eigenvalue and the Perron eigenvector of our matrix A, $\lambda_{max} = 4.0685$ is obtained, the priority vector being

$Z = (0.2643, 0.5693, 0.0609, 0.1055)^t$.

Higher values for Z's components correspond to more weighted criteria in the evaluation process; while lower values refer to criteria regarded as less important. In this case, the greater value corresponds to costs related to *damage to personal properties and other*

service networks, corresponding to floods, building deterioration, damage to electricity networks, gas, telecommunications, etc. On the other hand, the lowest value is attributed to *supply disruptions* that are related with the cost of alternative supply.

4. The decision making process

As stated in [6] the additive approach is crucial in performing composition using the limiting powers of a priority – rather than a judgment – matrix when dependence and feedback are considered in decision-making. In this way, once we have obtained the priority vector for the four considered criteria, the next step is to obtain vectors of priorities for our two alternatives, namely ALC and PLC, for each criterion. These vectors will reflect the weight or relative importance of each alternative for each criterion [3]. Calculation of these priority vectors is straightforward since the four matrices are 2×2. In fact, it can be proven that positive, reciprocal 2×2 matrices are always consistent. As a result, any column of one such matrices is a principal eigenvector (corresponding to $\lambda_{max} = 2$). Consequently, normalization of any of these columns directly gives the sought priority vector. The four priority vectors can be observed in Figure 3.

	ALC	PLC	
ALC	1	7	0.8750
PLC	1/7	1	0.1250

Damage to properties and other service networks

	ALC	PLC	
ALC	1	1/5	0.1666
PLC	5	1	0.8333

Planning development cost and its implementation

	ALC	PLC	
ALC	1	3	0.7500
PLC	1/3	1	0.2500

Closed or restricted streets

	ALC	PLC	
ALC	1	2	0.6666
PLC	1/2	1	0.3333

Supply disruption

Figure 3. Matrices of alternative comparisons according to established criteria and their corresponding priority vectors

Finally, a score is computed for an alternative by multiplying its priority value times the priority of any criterion and summing all the criteria. This score is shown in the last column of Figure 4. The highest value in this column will be associated with 'the best alternative' and the lowest with 'the worst alternative' [2].

	Weight of leakage management alternatives for each valuation criterion				Weight of criteria	Weight of alternatives based on criteria
ALC	0.1666	0.8750	0.6666	0.7500	0.2643 0.5693	0.6619
PLC	0.8333	0.1250	0.3333	0.2500	0.0609 0.1055	0.3381

Figure 4. Weighting of leakage management alternatives

Since the priority vector of alternatives is

$W = \{0.6619, 0.3381\}$,

an ALC policy should clearly be preferred over PLC. According to the considerations already made, the social costs that a water utility could incur for ALC, as much as PLC, play a leading role in the decision. The second factor influencing this decision must be attributed to planning development costs and their implementation.

5. Conclusions

The objective of considering other costs, not only those of an economic nature, associated with the assessment of the actual level of leakage in a water supply system, and the development of management alternatives has created interest in the search for alternative methods of valuation. In this paper, we have shown that AHP can be considered suitable for the inclusion of social and environmental evaluation costs. The fact that AHP can include social and environmental externalities in an evaluation of actual level of leakage and management alternatives seems to indicate that AHP is superior to traditional methods. Additionally, AHP can be used to consider not only the manager's point of view, but also the opinion of other involved actors, such as users, authorities, associations, etc.

AHP, which is based on the properties of the matrix of criteria, enables the evaluation of complex multi-criteria problems through a hierarchical visualization of the problem, including objectives, alternatives, and criteria. For the studied problem, we have shown that (only including the point of view of managers) the alternative of active control leakage clearly outperforms the classical passive control leakage. The main factor influencing this fact comes from the consideration of social costs due to damage to properties and other service networks.

Acknowledgements

This work has been performed under the support of the project IDAWAS, **DPI2009-11591** of the Dirección General de Investigación del Ministerio de Educación y Ciencia (Spain). The use of English in this paper was revised by John Rawlins; and the revision was funded by the Universidad Politécnica de Valencia, Spain.

References

[1] T.L. Saaty, Relative Measurement and Its Generalization in Decision Making. Why Pairwise Comparisons are Central in Mathematics for the Measurement of Intangible Factors. The Analytic Hierarchy/Network Process, Rev. R. Acad. Cien. Serie A. Mat. 102(2) (2008) 251–318.

[2] B. Srdjevic, Linking analytic hierarchy process and social choice methods to support group decision-making in water management, Decision Support Systems 42 (2007) 2261-2273.

[3] J. Aznar, F. Guijarro, Nuevos métodos de valoración. Modelos Multicriterio. Universidad Politécnica de Valencia. España, 2008.

[4] P.K. Dey, An integrated assessment model for cross-country pipelines, Environmental Impact Assessment Review 22 (2002) 703–721.

[5] G. Stewart, J.G. Sun, Matrix Perturbation Theory, Academic Press, New York, 1990.

[6] T.L. Saaty, G. Hu, Ranking by the eigenvector versus other methods in the analytic hierarchy process, Applied Mathematical Letters 11 (4) (1998) 121–125.

A quasi-steady model for gas-dynamic prediction of centrifugal compressor behaviour*

J. Galindo[†], F. J. Arnau [‡], A. Tiseira[§] and P. Piqueras[¶]

CMT-Motores Térmicos,

Universidad Politécnica de Valencia,

Edificio 6G, 46022 Valencia, España.

December 10, 2009

1 Introduction

Models for the simulation of internal combustion engines are important tools as much in design as in optimization [1]. Unlike CFD codes, which provide great accurate simulations but with high time consuming that only allow simulations of isolated engine elements, one-dimensional gas-dynamics codes provide a good balance between accuracy and computational cost [2, 3]. These codes are widely used to reproduce the whole engine behaviour. They consider the different thermodynamics processes (i.e. heat transfer, friction) that the fluid undergoes through the different elements of the engine, and despite of the possibility of calculating ducts with non constant section, only the axial dimension is considered. Besides the flow movement through the elements of the engine, 1-D codes must be able to reproduce how different techniques like turbocharging, exhaust gas recirculation or exhaust gas after-treatment can affect the engine behaviour.

*The authors wish to thank the economical support of this work to Spanish Project TRA2007-65433 from Ministerio de Ciencia e Innovacin

[†]email galindo@mot.upv.es

[‡]email farnau@mot.upv.es

[§]email anti1@mot.upv.es

[¶]email pedpicab@mot.upv.es

In recent years, turbocharged diesel engines, which is a technique largely used to improve the power output of such engines, has spread out to engines of small displacement capacity as those used in automotive applications. This technique has become a major actor in downsizing. The main objective of downsizing is to decrease engine displacement while maintaining or even increasing power output and reducing fuel consumption and pollutant emissions. Recently, this technique is also extending to the new gasoline direct injection engines.

Due to the extension of turbocharging technique, gas-dynamic codes must provide to engineers the capability to simulate the real behaviour of the turbocompressors coupled to the alternative engine not only working in normal conditions but also detecting extreme behavior like surge [4].

This paper presents the mathematical base of a one-dimensional compressor model validated experimentally in [5]. The model uses the compressor map information and acts as a quasi-steady boundary condition connecting two ducts. The objective of the model is to simulate both stable working conditions and surge phenomena. The boundary uses the Method of Characteristics to determine the flow conditions at compressor inlet and outlet. Unlike the compressor model presented by Katrasnik [6], this model is based on the perfect gas hypothesis due to the fact that the temperature drop between compressor inlet and outlet does not cause significant variation in the gas properties and then, specific heat ratio can be considered constant. This assumption simplifies the boundary solution and reduces time consumption.

2 Compressor model

The compressor boundary condition is based on the use of the compressor map, which relates the compressor ratio and the efficiency to the corrected air mass flow for every compressor speed. Usually, the compressor map is provided by manufacturers within the operation zone in a flow range from surge to choke. The model presented in this work needs to extend the range flow of the iso-speed curves to negative values. This is very important to reach the capability of surge prediction.

The most accurate way to obtain an extended map is by means of experiments. To measure in the surge flow range or in the negative flow range a valve is placed very near the compressor outlet in order to reduce the volume between the compressor and the valve. The small volume reduces consider-

ably the generation of oscillations due to the surge phenomenon [7] and the compressor can be measured in a pseudo-stable working point.

However, the measurements of the compressor in these conditions can be very destructive [8]. The compressor suffers severe working conditions and it can break. Therefore, in many cases it is preferable to estimate the shape of the iso-speed curve in this zone of the compressor map. To help this estimation the compressor ratio for air mass flow equal to zero can be theoretically obtained by means of the radial equilibrium at compressor blades with geometrical information of the compressor.

$$\pi_{\dot{m}=0} = \left[1 + \frac{\gamma - 1}{2\gamma R T_{01}} \omega^2 \left(r_2^2 - r_1^2\right)\right]^{\frac{\gamma}{\gamma-1}} \tag{1}$$

In equation 1 r_2 represents the outlet tip radius of the compressor wheel and $r_1 = \sqrt{(r_{1tip}^2 + r_{1hub}^2)/2}$ is a mean radius of the inducer section.

In a gas-dynamic code, the compressor boundary conditions presented behaves as a connection between ducts. Every time step, the boundary is solved supposing the quasi-steady hypothesis. The physical equations describing the behaviour of the compressor relate the flow conditions at the compressor inlet and at the compressor outlet.

The first equation refers to mass conservation. Due to the fact that the compressor is considered by the model as a point in the flow path, instantaneously, the mass flow at the compressor inlet must be the same as at the compressor outlet. This hypothesis is represented by equation 2.

$$\rho_1 U_1 S_1 = \rho_2 U_2 S_2 \tag{2}$$

The pressure increment generated by the compressor is represented by the momentum equation. This equation provides the stagnation pressure at the compressor outlet as a function of the stagnation pressure at the compressor inlet times the compressor ratio.

$$p_{02} = p_{01}\pi \tag{3}$$

In equation 3, π represents the compressor ratio. The compressor model presented includes an inertia term avoiding sudden changes in this parameter. Thus the compressor ratio at every time step is calculated using the equation 4

$$\frac{\mathrm{d}\pi}{\mathrm{d}t} = \frac{\pi_{map} - \pi}{\tau} \tag{4}$$

where π_{map} represents the compressor ratio interpolated in the compressor map, and $\tau = L_c/u$ is the time that a gas particle spend to cross a compressor with a characteristic length L_c and with a gas velocity u.

Finally, the flow that pass through the compressor undergoes a heating process due to the compression of the flow and the efficiency of the compressor. This phenomenon can be described by equation 5.

$$T_{02} = T_{01}(1 + k) \tag{5}$$

where $k = (\pi^{\frac{\gamma-1}{\gamma}} - 1)/\eta$

These three equations can be written as a function of known parameters at the boundary, i.e. the Riemann variables or the entropy level at both sides of the boundary. Firstly they can be written as a function of the compressor inlet and outlet gas velocity, speed of sound and entropy level.

Rearanging the terms of the equations it is possible to obtain a non-linear two-equation system where the unknown parameters are the match number at compressor inlet and at compressor outlet. This equation system can be solved using a root finding method like Newton-Rapson Method for non-linear system of equations [9].

Figure 1 shows a flow diagram explaining how the compressor model is integrated in the 1D gas dynamic code.

3 Conclusiones

A compressor model has been presented which is used as a quasi-steady boundary condition in one-dimensional gas dynamic codes and it connects two ducts. The mathematical aspects of the model have been analysed in depth. The boundary solution is based on The Method of Characteristics and it needs an iterative process. The Newton-Rapshon Method for non-linear system of equations is used.

The compressor map is essential information for the model. Working conditions must be interpolated in the map, and as it has been shown in this work, the transformation of the compressor map data in polar coordinates increases the accuracy of the model.

Another important aspect of the model is the inertia term introduced in the boundary equations, which allows surge prediction. The aim of this inertia term is avoid sudden changes in compressor ration simulating the inertia of the flow.

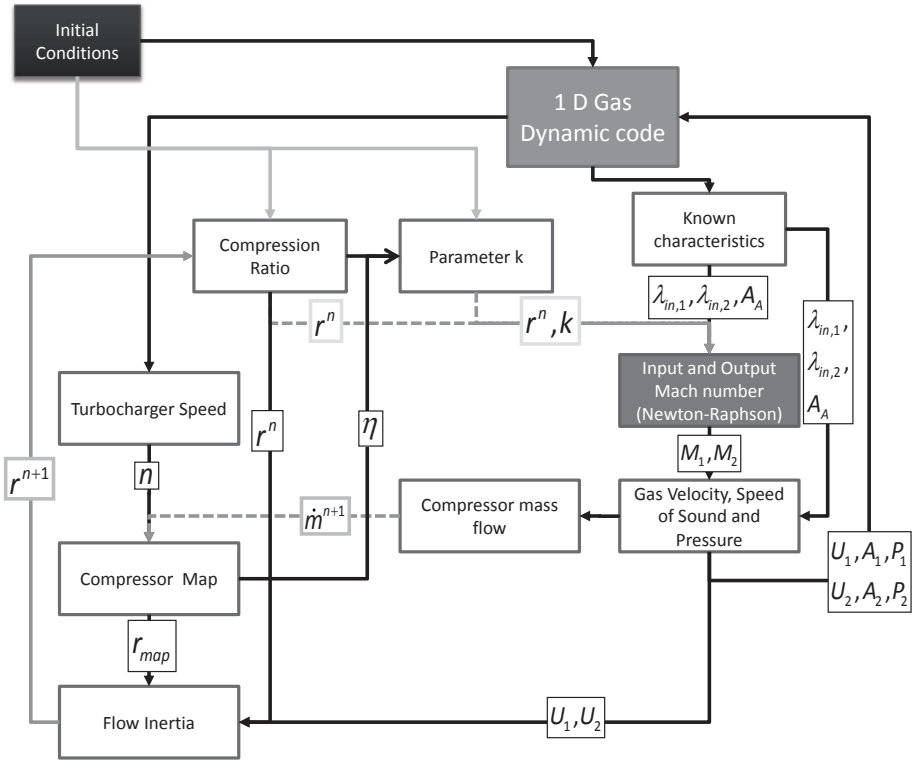


Figure 1: Flow diagram of the compressor model

Finally, the model has been compared with experimental data obtained in a turbocharger test bench. The good agreement with measured data of the results shown confirms the capabilities of the compressor model to work in normal working conditions and to predict surge.

References

[1] D. E. Winterbone, R. J. Pearson, Theory of engine manifold design: wave action methods for IC engines, Professional Engineering Publishing, 2000.

[2] J. Galindo, J. Serrano, F. Arnau, P. Piqueras, Description of a semi-independent time discretization methodology for a one-dimensional gas dynamics model, Journal of Engineering for Gas Turbines and Power 131 (034504) (2009) 5.

[3] J. Galindoa, J. Serrano, F. Arnau, P. Piqueras, High-frequency response of a calculation methodology for gas dynamics based on independent time discretisation, Mathematical and Computer Modelling 50 (5-6) (2009) 812–822.

[4] E. Greitzer, Surge and rotating stall in axial flow compressors, Journal of Engineering for Power Transactions of the ASME 98 (1976) 190 – 198.

[5] J. Galindo, J. Serrano, H. Climent, A. Tiseira, Experiments and modelling of surge in small centrifugal compressor for automotive engines, Experimental Thermal and Fluid Science 32 (2008) 818–826.

[6] T. Katrasnik, Improved model to determine turbine and compressor boundary conditions with the method of characteristics, International Journal of Mechanical Science 48 (2006) 504–516.

[7] J. Galindo, J. Serrano, C. Guardiola, C. Cervelló, Surge limit definition in a specific test bench for the characterization of automotive turbochargers., Experimental Thermal and Fluid Science 30 (5) (2006) 449 – 462.

[8] H. Bloch, A Practical Guide to Compressor Technology., McGraw-Hill, 1996, iSBN 0-07-005937-3.

[9] J. Ortega, W. Rheinboldt, Iterative solution of nonlinear equations in several variables, New York: Academic Press, 1970.

Modeling Bladder Cancer Using a Markov Process With Multiple Absorbing States.

B. García-Mora [*], C. Santamaría[†], E. Navarro [‡], and G. Rubio [§]

Instituto Universitario de Matemática Multidisciplinar,

Universidad Politécnica de Valencia,

Edificio 8G, piso 2, 46022, Valencia, Spain.

December 10, 2009

1 Introduction and Aims

Bladder carcinoma is an aggressive neoplasm where most superficial tumors (60% to 70%) have a propensity to *recurrence* after transurethral tumor resection (TUR). Some (15% to 25%) *progress* to muscle invasion [1] leading even the *bladder extirpation*. We calculate the associated survival functions with both *progression* and *extirpation* to establish predictions after TUR.

Several extensions of the Cox model have been proposed to analyze the reappearance of *multiple recurrences* [2]. An application of all these in bladder carcinoma is in [3], but the analysis becomes more complicated as the number of recurrences increases. *Markov models* have proven to be useful in the study of chronic diseases with relapse times, since subjects can spend times in the different stages of the disease (transient and absorbing ones). References in cancer are [4, 5] and [6] applied to bladder carcinoma where only one absorbing state is considered. The literature is not extensive [7] when this number is higher.

[*]magarmo5@imm.upv.es

[†]crisanna@imm.upv.es

[‡]entorres@imm.upv.es

[§]grubio@imm.upv.es

We consider the sum of two random continuous independent variables representing two absorption times in bladder carcinoma. These variables are associated with two homogeneous time–continuous Markovian processes, each of them with multiple absorbing states. The distribution function of the variable sum leads to an increment of the dimension of the problem. In order to avoid this drawback we make use of the Fréchet derivative and the Kronecker matrix representation to compute the exponential function of a block upper triangular matrix in terms of its respective blocks.

2 Survival function of the sum of two independent absorption times in Markov processes with multiple absorbing states.

Let $\{M(x) : x \geq 0\}$ be a homogeneous Markov process with state space $\{1, 2, \ldots, m, m+1, \ldots, m+p\}$, where $\{1, 2, \ldots, m\}$ are the transient states and $\{m+1, \ldots, m+p\}$ are the absorbing ones. $(\alpha, \alpha_{m+1} \ldots, \alpha_{m+p})$ is an initial probability vector where α is a row probability m–vector with $\alpha_i = P(M(0) = i),\ i = 1, \ldots m$ and $\alpha_{m+j} = P(M(0) = m+j),\ j = 1, 2, \ldots, p$.

The infinitesimal generator Q associated with the process $M(x)$ is

$$Q = \left[\begin{array}{c|c} T & T^0 \\ \hline 0 & 0 \end{array}\right] \tag{1}$$

$T \in \mathbb{R}^{m\times m}$ is a non–singular matrix. The entries t_{ij} represent the rates among the transient states and $|t_{ii}|^{-1}$ is the mean holding time in the ith transient state and exponentially distributed, so the structure of T is

$$t_{ij} = \begin{cases} -v_i & \text{if } i = j \\ v_i\, p_{ij} & \text{if } i \neq j \end{cases} \tag{2}$$

where $v_i > 0$ is the parameter of the exponential distribution in the ith state and p_{ij} is the transition probability from the state i to the state j when the process leaves the state i. It follows that $|t_{ii}| \geq \sum_{j\neq i} t_{ij}$, $t_{ij} > 0$ and $t_{ii} < 0$.

$T^0 \in \mathbb{R}^{m\times p}$ denotes a matrix of order $m \times p$ where the entries represent the absorbing rates from ith-transient state to jth absorbing state.

The matrices T and T^0 satisfy the relation $T\mathbf{e_m} + T^0\mathbf{e_p} = 0$ where $\mathbf{e_k} = (1, 1, \ldots, 1)' \in \mathbb{R}^{k\times 1}$, since each row in Q adds up to zero, given

that Q is conservative. Furthermore $\alpha \mathbf{e_m} + \sum_{j=1}^{p} \alpha_{m+j} e_j = 1$ given that $(\alpha,\, \alpha_{m+1} \ldots, \alpha_{m+p})$ is a row probability m–vector.

$P(x) = \{p_{ij}(x);\, i,j = 1, \ldots, m+p\}$ is the matrix of transition probabilities where $p_{ij}(x)$ is the probability of the process being in state j at time x after being in state i initially $(p_{ij}(x) = P(M(x) = j | M(0) = i))$. $P(x)$ satisfies $\left\{\frac{dP(x)}{dx} = P(x)Q,\ P(0) = I\right\}$, with solution

$$P(x) = \exp(Qx) = \left[\begin{array}{c|c} \exp(Tx) & (\exp(Tx) - \mathrm{I})\, T^{-1}T^0 \\ \hline 0 & \mathrm{I} \end{array}\right]$$

If we denote X_{m+h} as the time until absorption by the state $m+h$ ($X_{m+h} = \inf\{x :\, M(x) = m+h\}$), then the distribution function of X_{m+h} is

$$\begin{aligned} F_h(x) &= P(X_{m+h} \le x) = (\alpha,\, \alpha_{m+1} \ldots, \alpha_{m+p}) \left[\begin{array}{c} p_{1,m+h}(x) \\ p_{2,m+h}(x) \\ \vdots \\ p_{m+p,m+h}(x) \end{array}\right] \\ &= \left[\begin{array}{cc} \alpha \exp(Tx) & \alpha\,(\exp(Tx) - \mathrm{I})\, T^{-1}T^0 + (\alpha_{m+1} \ldots, \alpha_{m+p}) \end{array}\right]_{m+h} \\ &= \alpha\,(\exp(Tx) - \mathrm{I})\, T^{-1}T_h^0 + \alpha_{m+h} \end{aligned} \tag{3}$$

where T_k^0 denotes the kth column of the matrix $T^0 \in \mathbb{R}^{m \times p}$. The associated Survivor function will be

$$S_h(x) = 1 - F_h(x) = P(X_{m+h} > x), \tag{4}$$

and the density function and its Laplace transform are respectively

$$f_h(x) = \alpha \exp(Tx) T_h^0 \quad \text{and} \quad \mathcal{L}[f_h(x)] = \alpha\,(sI - T)^{-1}\, T_h^0 \tag{5}$$

In order to compute the distribution function for the sum of two independent absorption times in Markov processes with multiple absorbing states, let $\{M_1(x) : x \ge 0\}$ and $\{M_2(x) : x \ge 0\}$ be two homogeneous processes with state space $\{1,\, 2,\, \ldots, m,\, m+1, \ldots, m+p\}$ and $\{1,\, 2,\, \ldots, n,\, n+1, \ldots, n+q\}$ respectively where, $\{1,\, 2,\, \ldots, m\}$ and $\{1,\, 2,\, \ldots, n\}$ are the transient states and $\{m+1,\, \ldots,\, m+p\}$ and $\{n+1,\, \ldots, n+q\}$ the absorbing ones.

The initial probability vectors are $(\alpha,\, \alpha_{m+1} \ldots, \alpha_{m+p})$ and $(\beta,\, \beta_{n+1} \ldots, \beta_{n+q})$ for both Markov processes and the infinitesimal generators are respectively

$$Q_1 = \left[\begin{array}{c|c} T & T^0 \\ \hline 0 & 0 \end{array}\right] \text{and } Q_2 = \left[\begin{array}{c|c} S & S^0 \\ \hline 0 & 0 \end{array}\right] \tag{6}$$

where the matrices $S \in \mathbb{R}^{n\times n}$ and $S^0 \in \mathbb{R}^{n\times q}$ are similar to T and T^0 satisfying the relation $S\mathbf{e_n} + S^0\mathbf{e_q} = 0$.

If X_{1h} denotes the absorption time by the absorbing state $m+h$ in the first process and X_{2k} the absorption time by the absorbing state $n+k$ in the second process, then $Y_{hk} = X_{1h} + X_{2k}$ will denote the absorption time by the absorbing state $n+k$ of the second process after the first process has previously arrived at its absorbing state $m+h$. We consider a new Markov process $\{M(x) : x \geq 0\}$ associated with Y_{hk} with state space $\{1, 2, \ldots, m, m+1, m+2, \ldots, m+n, m+n+1, \ldots, m+n+p-1, m+n+p, \ldots, m+n+p+q-1\}$, where $\{1, 2, \ldots, m, m+1, \ldots, m+n\}$ are the transient states and $\{m+n+1, \ldots, m+n+p-1, m+n+p, \ldots, m+n+p+q-1\}$ the absorbing ones. In $M(x)$ the transient states are the set of transient states of the first process and the transient ones of the second after the first process has arrived at its absorbing state $m+h$. The absorbing states are the set of absorbing states of the first process except for the state $m+h$ and all absorbing ones from the second process after the first one arrives at $m+h$. $M(x)$ maintains the initial and transition probabilities of $M_1(x)$ and $M_2(x)$. So the initial probability vector and the infinitesimal generator will be respectively

$$\gamma = (\alpha, \alpha_{m+h}\beta, \alpha_{m+1}, \ldots, \alpha_{m+h-1}, \alpha_{m+h+1}, \ldots, \alpha_{m+p}, \alpha_{m+h}\beta_{n+1}, \ldots, \alpha_{m+h}\beta_{n+q}),$$

and

$$Q = \left[\begin{array}{c|c} L & L^0\beta \\ \hline 0 & 0 \end{array}\right] \text{ where } L = \left[\begin{array}{c|c} T & T_h^0\beta \\ \hline 0 & S \end{array}\right],$$

with T and S given in (6), and

$$L^0 = \left[\begin{array}{c|c} T_1^0, T_2^0, \ldots, T_{h-1}^0, T_{h+1}^0, \ldots T_p^0 & T_h^0(\beta_{n+1}, \ldots, \beta_{n+q}) \\ \hline 0 & S^0 \end{array}\right]$$

The distribution function for Y_{hk} will be defined by (3) and given by

$$\begin{aligned} F_{hk}(x) &= P(X_{hk} \leq x) = (\alpha\ \alpha_{m+h}\beta)L^{-1}\big(\exp(Lx) - I_{m+n}\big)L^0_{p-1+k} + \alpha_{m+h}\beta_{n+k} \qquad (7) \\ &= (\alpha\ \alpha_{m+h}\beta)\left(\exp(Lx) - I_{m+n}\right)\left(\begin{array}{c} T^{-1}T_h^0\beta_{n+k} - T^{-1}T_h^0\beta S^{-1}S_k^0 \\ S^{-1}S_k^0 \end{array}\right) + \alpha_{m+h}\beta_{n+k} \end{aligned}$$

As the distribution functions $F_h(\cdot)$, $F_k(\cdot)$ and $F_{hk}(\cdot)$ of X_h, X_k and Y_{hk} are continuous and derived with derivative functions the density functions $f_h(\cdot)$, $f_k(\cdot)$ and $f_{hk}(\cdot)$ also continuous, then $f_{hk} = f_h * f_k$, or equivalently

$$\mathcal{L}[f_{hk}] = \mathcal{L}[f_h]\mathcal{L}[f_k] \qquad (8)$$

must be true. As

$$\begin{aligned}\mathcal{L}[f_{hk}] &= (\alpha\ \alpha_{m+h}\beta)\big(sI-L\big)^{-1}\begin{pmatrix} T_h^0\beta_{n+k} \\ S_k^0\end{pmatrix}\\ &= \alpha(sI-T)^{-1}T_h^0\beta_{n+k}+\alpha(sI-T)^{-1}T_h^0\beta(sI-S)^{-1}S_k^0+\alpha_{m+h}\beta(sI-S)^{-1}S_k^0\\ &= \alpha(sI-T)^{-1}T_h^0\beta_{n+k}+\mathcal{L}[f_h]\mathcal{L}[f_k]+\alpha_{m+h}\beta(sI-S)^{-1}S_k^0.\end{aligned}$$

Then (8) is true if $\alpha_{m+h}=\beta_{n+k}=0$ and so any process can start in the absorbing states α_{m+h} and β_{n+k}. So the distribution function (7) is

$$F_{hk}(x) = (\alpha\ 0)\exp(Lx)\begin{pmatrix} -T^{-1}T_h^0\beta S^{-1}S_k^0 \\ S^{-1}S_k^0\end{pmatrix}+\alpha T^{-1}T_h^0\beta S^{-1}S_k^0 \qquad (9)$$

The problem increases in size as the matrix L is greater than T and S. We use the following Lemma to avoid this and to compute $\exp(Lx)$ in terms of its respective blocks T and S.

Lemma 2.1 (Schur–Fréchet) *Let H be an analytic function in a open set containing the spectrum of A with $A=D+Z$ a block upper triangular matrix where D and Z have the same block structure*

$$D=\begin{bmatrix} A_{11} & 0\\ 0 & A_{22}\end{bmatrix};\quad Z=\begin{bmatrix} 0 & A_{12}\\ 0 & 0\end{bmatrix};$$

where $A_{11}\in\mathbb{R}^{m\times m}$, $A_{22}\in\mathbb{R}^{n\times n}$, $A_{12}\in\mathbb{R}^{m\times n}$. Then, it follows that

$$H(A)=H(D)+L_H(Z,D) \qquad (10)$$

where $L_H(Z,D)$ denotes the Fréchet derivative of H at D in the matrix direction Z.

Property 2.2 (Fréchet derivative of the exponential map). *If $H(X)=\exp(X)$ then it follows that (see [8] for more details):*

$$L_H(Z,\,D)=\int_0^1\exp((1-s)D)Z\exp(sD)\,ds, \qquad (11)$$

We apply the Lemma 2.1 to the term $\exp(Lx)$ with the property (11) and we substitute in the distribution function (9)

$$\begin{aligned}F_{hk}(x) &= (\ \alpha\ \ 0\ )\left\{\begin{pmatrix}\exp Tx & 0\\ 0 & \exp Sx\end{pmatrix}+\right. \qquad (12)\\ &\left.\int_0^1\begin{pmatrix}\exp(1-s)Tx & 0\\ 0 & \exp(1-s)Sx\end{pmatrix}\begin{bmatrix}0 & T_h^0\beta x\\ 0 & 0\end{bmatrix}\begin{pmatrix}\exp sTx & 0\\ 0 & \exp sSx\end{pmatrix}ds\right\}\\ &\begin{bmatrix}-T^{-1}T_h^0\beta S^{-1}S_k^0\\ S^{-1}S_k^0\end{bmatrix}+\alpha T^{-1}T_h^0\beta S^{-1}S_k^0\end{aligned}$$

and operating we arrive at the following expression

$$F_{hk}(x) = \alpha\left(\mathrm{I}-\exp[Tx]\right)T^{-1}T_h^0\beta S^{-1}S_k^0 + \int_0^1 \alpha\exp\left[(1-s)Tx\right]T_h^0\beta x\exp sSxS^{-1}S_k^0\,ds \tag{13}$$

Let Y denote the first addend of (13), we apply to the integral twice the property of the *Kronecker matrix form* and we can obtain

$$F_{hk}(x) = Y + \left((S^{-1}S_k^0)' \otimes \alpha\right)\left[S'x \oplus (-Tx)\right]^{-1}\left(\exp(S'x)\otimes \mathrm{I_m} - \mathrm{I_n}\otimes\exp(Tx)\right)\mathrm{vec}\left(T_h^0\beta\right)x,$$

since the derivative $\frac{d}{ds}\left[sS'x \oplus (1-s)Tx\right] = S'x \oplus (-Tx)$ is constant with respect to s, commutes with $sS'x \oplus (1-s)Tx$ and is non–singular if and only if $\Gamma(S)\cap\Gamma(T) = \emptyset$. So the total expression for the Survivor Function is $S_{hk}(x) = 1 - F_{hk}(x)$ with $F_{hk}(x)$ given in last expression.

3 Bladder carcinoma application

Bladder carcinoma is characterized by multiple recurrences (reappearance of superficial tumors) and progression (reappearance of a more highly aggressive tumor and muscle invasion leading even in some cases to bladder extirpation. So, two different stages in the evolution of the bladder cancer are considered, each one of them modeled with a Markov process.

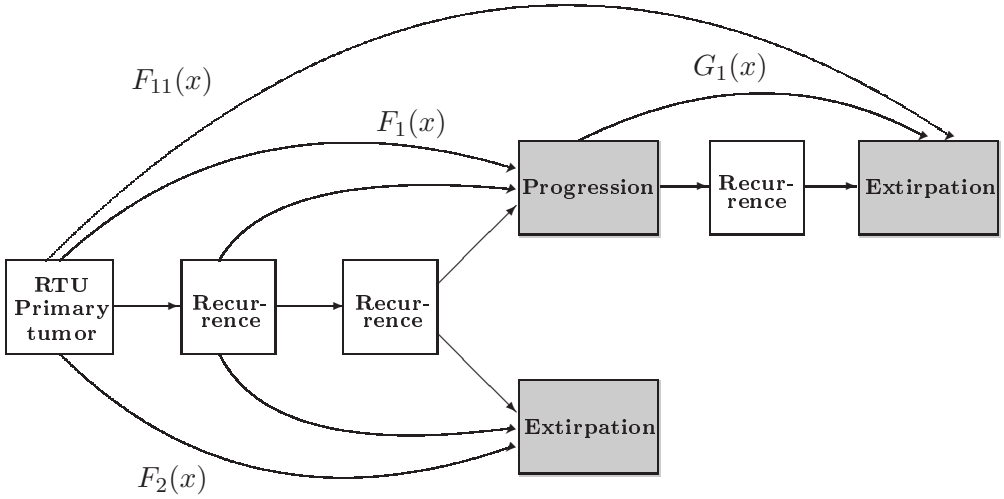


Diagram: The Markov process $M(x)$.

The first stage begins from the appearance of a primary superficial tumor until *progression* or bladder *extirpation*, whichever occurs first, both are considered absorbing states. During that period of time we have considered up to two recurrences per patient as transient states. So the first stage is modeled with a Markov process with three transient states (the primary superficial tumor, the first and second recurrence) and two absorbing states (*progression* and *extirpation*), (see Diagram). The second stage begins from the first progression until bladder extirpation. In this period only one new recurrence of muscle invasive tumor until extirpation. So, the second stage is modeled by means of a Markov process with two transient states (first and second muscle invasive tumors) and only one absorbing state (*extirpation*).

Let $\{M_1(x) : x \geq 0\}$ and $\{M_2(x) : x \geq 0\}$ be the above described independent Markov processes with two and only one absorbing states respectively. The distribution functions $F_1(\cdot)$ and $F_2(\cdot)$ associated with the absorption times *progression* and *extirpation* respectively in the first process $M_1(x)$ are defined by (3):

$$F_1(x) = \alpha T^{-1} \left(\exp(Tx) - \mathrm{I}\right) T^0 \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{14}$$

$$F_2(x) = \alpha T^{-1} \left(\exp(Tx) - \mathrm{I}\right) T^0 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{15}$$

where $T \in \mathbb{R}^{3\times 3}$ and $T^0 \in \mathbb{R}^{3\times 2}$. Nevertheless the absorption time in the second process $M_2(x)$ associated with the time until extirpation is Phase type distributed due to it only having one absorbing state [[9], Chap.2]. The distribution function in this case is given by

$$G_1(x) = 1 - \beta \exp(Sx)\mathbf{e} \tag{16}$$

where $S \in \mathbb{R}^{2\times 2}$ and $S^0 \in \mathbb{R}^{2\times 1}$. In this case the matrices S and S^0 satisfy that $S^0 = -S\mathbf{e}$, and $\beta\mathbf{e} + \beta_3 = 1$ with $\mathbf{e} = (1, 1)'$.

The rates of transition in the matrices T and S are estimated using the maximum–likelihood method as in [6]. For this two independent data sets were used. The first data set contains information on superficial bladder tumors and was collected from 540 patients with a mean follow-up period of 48 months. The second data set was made up of 60 patients, all of whom presented a muscle invasive tumor with a mean follow-up period of 38.9 months. The transition intensity matrices T and S were estimated using the R program and the `msm` package [10].

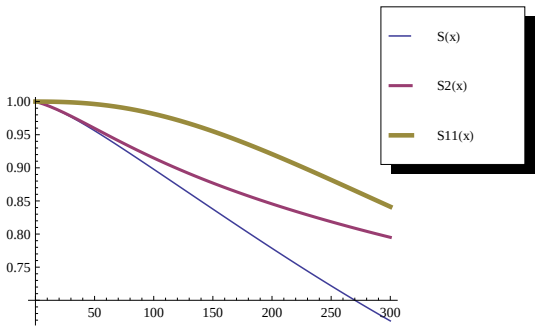


Figure 1: Survival Curves

$$T = \begin{bmatrix} -0.01935 & 0.01846 & 0 \\ 0 & -0.02729 & 0.02386 \\ 0 & 0 & -0.002215 \end{bmatrix}; S = \begin{bmatrix} -0.01033 & 0.006572 \\ 0 & -0.01738 \end{bmatrix}$$

where the entries t_{12} and t_{23} represent the transition rates between the primary superficial tumor and the first recurrence, and between the first and second recurrence in the first protocol. In a similar way the entrie s_{12} is the transition rate between the two possible transient muscle invasive tumors in the second protocol.

In order to compute the Survivor Function to extirpation let a new Markov process $M(x)$ consider constructed from $M_1(x)$ and $M_2(x)$, depicted in Diagram. Let X_{11} denote the absorption time by *extirpation* in the second process after the first process has arrived at its first absorption state, the *progression*. So, using result (14), the Survivor Function for X_{11} is:

$$S_{11}(x) = \alpha \left(\exp(Tx) - \mathrm{I}_3\right) T^{-1} T^0 \begin{bmatrix} 1 \\ 0 \end{bmatrix} - \tag{17}$$

$$(\mathbf{e}' \otimes \alpha) \left[S'x \oplus (-Tx)\right]^{-1} \left[\exp(S'x) \otimes \mathrm{I}_3 - \mathrm{I}_2 \otimes \exp(Tx)\right] \mathrm{vec}\left(T^0 \begin{bmatrix} 1 \\ 0 \end{bmatrix} \beta x\right)$$

$S_{11}(x)$ gives the probability that a patient does not undergo bladder extirpation when the patien begun with a primary superficial tumor suffering this before progression. $S_2(x) = 1 - F_2(x)$ gives the probability that a patient does not undergo bladder extirpation when the patient begun with a primary superficial tumor without undergoing any progression before bladder extirpation. So $S(x) = S_{11}(x)S_2(x)$ is exactly the probability that a

patient does not undergo bladder extirpation. As for both processes $M_1(x)$ and $M_2(x)$ the patients are initially in the primary superficial tumor and the first muscle invasive tumor states respectively, the initial probabilities are $\alpha = (1, 0, 0)$ and $\beta = (1, 0)$. As the initial probability of an absorbing state is zero, then $\alpha_3 = \alpha_4 = \beta_3 = 0$. Finally, the resulting survival curves are depicted in Figure 1.

References

[1] C. L. Amling, Diagnosis and management of superficial bladder cancer, Curr. Probl. Cancer. 25(4) (2001) 219–278.

[2] T. M. Therneau, P. M. Grambsch, Modelling Survival Data: Extending the Cox Model, Springer, Nueva York, 2000.

[3] B. García-Mora, C. Santamaría, G. Rubio, J. L. Pontones, Modeling the recurrence–progression process in bladder carcinoma, Comput. Math. Appl. 56 (2008) 619–630.

[4] O. O. Aalen, On phase type distributions in survival analysis, Scand. J. Stat. 22 (1995) 447–463.

[5] C. H. Jackson, L. D. Sharples, Hidden markov models for the onset and progression of bronchiolitis obliterans syndrome in lung transplant recipients, Stat. Med. 21 (2002) 113–128.

[6] C. Santamaría, B. García-Mora, G. Rubio, E. Navarro, A markov model for analyzing the evolution of bladder carcinoma, Math. Comput. Model. in press, doi: 10.1016/j.mcm.2008.12.019.

[7] R. Pérez-Ocón, J. Ruiz-Castro, A multiple absorbent markov process in survival studies: Application to breast, Biometrical J. 45(7) (2003) 783–797.

[8] C. S. Kenny, A. J. Laub, Condition estimates for matrix functions, SIAM J. Matrix Anal. Appl. 10 (3) (1998) 191–209.

[9] M. F. Neuts, Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach., John Hopkins University Press, 1981.

[10] C. Jackson, msm: Multi-state Markov and hidden Markov models in continuous time, r package version 0.8.1 (2008).

Simulation of a cubic-like Chua's oscillator with variable characteristic

D. Ginestar *, E. Parrilla, J.L. Hueso, J. Riera,

Instituto de Matemática Multidisciplinar.

Universidad Politécnica de Valencia.

Camino de Vera, 14. 46022 Valencia (Spain).

1 Introduction

Chua's circuit is a widely studied nonlinear electronic device [1], [2], [3]. This circuit is shown in Figure 1 and can be simulated using the following system of differential equations, [4],

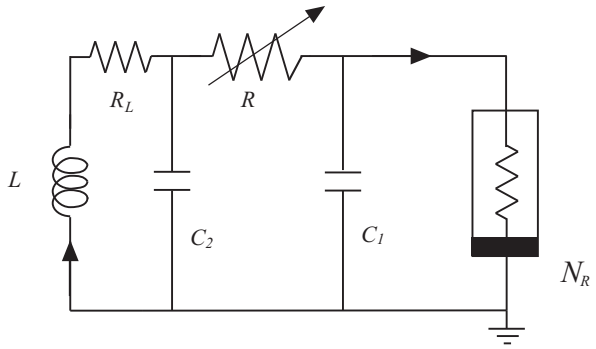


Figure 1: Chua's circuit.

$$\begin{aligned} \frac{dV_{C_1}}{dt} &= \frac{1}{RC_1}\left(V_{C_2} - V_{C_1}\right) - \frac{1}{C_1} g\left(V_{C_1}\right) \ , \\ \frac{dV_{C_2}}{dt} &= \frac{1}{RC_2}\left(V_{C_1} - V_{C_2}\right) + \frac{1}{C_2} i_L \ , \end{aligned}$$

*Corresponding author. e-mail: dginesta@mat.upv.es, Tel.:+34-963877665; Fax: +34-963877669.

$$\frac{di_L}{dt} = -\frac{1}{L}V_{C_2} - \frac{i_L}{L}R_L \ , \tag{1}$$

where $g(V)$ is the characteristic function of the nonlinear resistive element, which depends on the considered Chua's circuit configuration. The standard Chua's circuit has a nonlinear resistor with a piece-wise linear characteristic function. This function can be realized using two operational amplifiers and several linear resistors [4]. But the characteristic function of a nonlinear resistor realized in a real circuit is always a smooth function and not all the features of a physical circuit are captured correctly by a piecewise-linear characteristic. For this reason, some authors consider a different Chua's oscillator with a characteristic function given by a cubic polynomial, but the realization of this circuit requires a significant amount of circuitry and a simpler realization of a cubic-like nonlinear resistor has been proposed in [5]. The nonlinear resistor in this case is realized using four transistors, which can be visualised as a pair of cross-coupled CMOS inverters (see Figure 2).

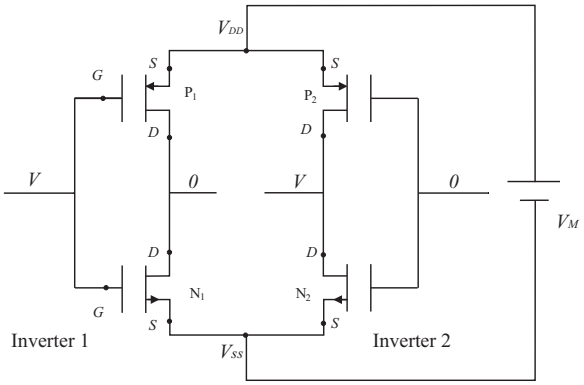


Figure 2: Standard cubic-like Chua's nonlinear resistor.

The shape of the characteristic function only depends on the transistors implemented in the chip selected to build the circuit and it can be difficult to find the characteristic attractors of the Chua's circuit in a reasonable range of the resistance, R, and typical values of the capacitances, C_1, and C_2 and the autoinduction L. Thus, we propose a modification in the realization of the cubic-like nonlinear resistor that allows to change easily the shape of the characteristic function of the circuit adding only two linear resistors.

2 Cubic-like variable characteristic function

2.1 Standard cubic-like characteristic function

In Figure 2, the realization of the cubic-like nonlinear resistor is shown. This realization uses a pair of cross-coupled CMOS inverters composed of four transistors [6].

A n-type transistor can be treated as a three-terminal device, the gate (G), the source (S) and the drain (D), which can work in the states cutoff, triode and saturation. The charcteristic function of the n-type transistor can be described by the function

$$i_N = \begin{cases} 0 \ , & \text{if} \quad V_{GSN} \leq V_{tN} \ , \\ k'_N \frac{W}{L} \left(\left(V_{GSN} - V_{tN} \right) V_{DSN} - \frac{1}{2} V_{DSN}^2 \right) \ , & \text{if} \quad V_{GSN} \geq V_{tN}; \\ & \qquad V_{DSN} < V_{GSN} - V_{tN} \ , \\ \frac{1}{2} k'_N \frac{W}{L} \left(V_{GSN} - V_{tN} \right)^2 \ , & \text{if} \quad V_{GSN} \geq V_{tN}; \\ & \qquad V_{DSN} \geq V_{GSN} - V_{tN} \ . \end{cases} \tag{2}$$

(For a detailed description of the n-type transistors and the notation used see [6]).

For the p-type transistor there is a similar characteristic function that depends on the p-type transistor parameters K'_P, V_{tP}.

Nowadays, most of the integrated circuits do not use individual MOSFET transistors but complementary MOS circuits (CMOS). The basic CMOS inverter utilizes two matched enhancement-type MOSFETs: one, with an n-channel and the other, with a p-channel constituting an inverter circuit (see Figure 2).

For the cubic-like circuit realized with two inverters the following assumptions are made, [5],

$$V_{tN} = -V_{tP} = V_t \ , \quad \frac{k'_N W}{L} = \frac{k'_P W}{L} = \mathcal{K} \ . \tag{3}$$

To obtain the characteristic function of this nonlinear resistor, we take into account that it is symmetric and we consider only the right half side. Three regions can be distinguished where transistors $N1$ and $P1$ of inverter 1 and transistors $N2$ and $P2$ of inverter 2 work in the states shown in Table 1.

The characteristic function for the standard cubic-like nonlinear resistor

Table 1: Regions for the standard nonlinear resistor.

Region 1	$N1$: sat.	$N2$: sat.
	$P1$: sat.	$P2$: sat.
Region 2	$N1$: triode	$N2$: sat.
	$P1$: sat.	$P2$: triode
Region 3	$N1$: triode	$N2$: cutoff
	$P1$: cutoff	$P2$: triode

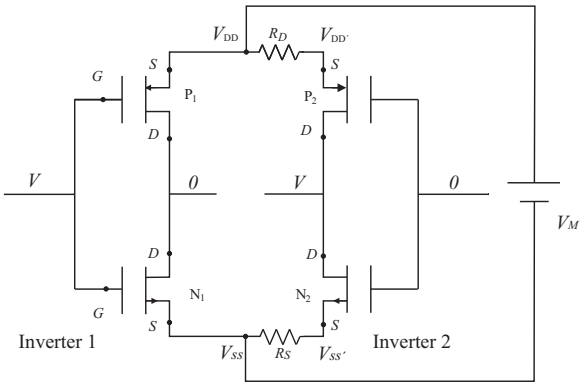


Figure 3: Variable cubic-like Chua's nonlinear resistor.

can be expressed as [5]

$$i = \begin{cases} \frac{\mathcal{K}}{2} V \left(V_M - 2V_t\right) \ , & \text{if} \quad V \in [0, V_t) \ , \\ \frac{\mathcal{K}}{2} \left(V^2 - V V_M + V_t^2\right) \ , & \text{if} \quad V \in [V_t, V_M - 2V_t) \ , \\ \frac{\mathcal{K}}{2} \left(\frac{3}{4} \left(V - \frac{1}{3} \left(V_M + 2V_t\right) \right)^2 - \frac{1}{3} \left(V_M - V_t\right)^2 \right) \ , & \text{if} \quad V \in [V_M - 2V_t, V_M] \ . \end{cases} \tag{4}$$

2.2 Variable cubic-like characteristic function

In order to be able to change the shape of the characteristic function of the nonlinear resistor, two new linear resistors, R_D and R_S, are added to the inverter 2, connected as it is shown in Figure 3. Since the values of R_D and R_S have not to be equal, the symmetry of the characteristic is not guaranteed. Now, seven regions can be distinguished in the operation of the circuit, where the transistors $N1$, $P1$, $N2$ and $P2$ work in the states shown in Table 2.

To obtain a model for the variable characteristic function of the nonlinear resistor we proceed in a similar way as it has been shown above for the

Table 2: Regions for the nonlinear resistor with resistors.

Region 1	$N1$: sat.	$N2$: sat.			
	$P1$: sat.	$P2$: sat.			
Region 2a	$N1$: triode	$N2$: sat.	Region 2a$'$	$N1$: sat.	$N2$: triode
	$P1$: sat.	$P2$: sat.		$P1$: triode	$P2$: sat.
Region 2b	$N1$: sat.	$N2$: cutoff	Region 2b$'$	$N1$: sat.	$N2$: sat.
	$P1$: sat.	$P2$: triode		$P1$: sat.	$P2$: cutoff
Region 3	$N1$: triode	$N2$: cutoff	Region 3$'$	$N1$: sat.	$N2$: triode
	$P1$: sat.	$P2$: sat.		$P1$: triode	$P2$: cutoff
Region 4	$N1$: triode	$N2$: cutoff	Region 4$'$	$N1$: cutoff	$N2$: triode
	$P1$: cutoff	$P2$: triode		$P1$: triode	$P2$: cutoff

standard cubic-like nonlinear resistor.

Given a value of V, in order to compute the intensity through the circuit, on has first to determine in which region is working. Thus, for each one of the regions, using that $V_{DD} = V_{SS} + V_M$ and the relation $i_{N1} - i_{P1} = i_{N2} - i_{P2}$, we obtain a set of nonlinear equations. For each region, we choose the only solution of these equations that satisfies $V_{DD'} > 0$ and $V_{SS'} < 0$.

Since, for a given V, only one state of each transistor is possible, the corresponding working region of Table 2 is determined, and the current through the circuit is computed, obtaining in this way the charcteristic function of the circuit.

3 Simulation and experimental results

3.1 Characteristic function simulation

To check the performance of the model presented above, we have used a polarization battery with voltage $V_M = 9$V and we have selected two different commercial chips to build the Chua's nonlinear resistor. Particularly, the chip CD4007UBE of Texas Instruments and the chip HEF4007UBP of Phillips Semiconductors have been analysed.

To obtain the parameters V_t and $\mathcal{K}$ of the transistors, we have set the values of the resistors $R_D = R_S = 0$ and we have used the experimental values for an extremal value of the intensity and the value of the voltage for which the extremal value is achieved. Assuming that the segment of the characteristic function (4) corresponding to the region 3 describes the

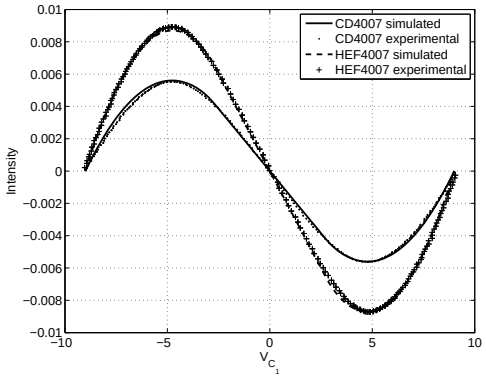


Figure 4: Experimental and simulated characteristic function for the chips CD4007 and HEF4007.

behaviour of the characteristic function near the extremal value we obtain the values presented in Table 3. The experimental and simulated values for the characteristic functions corresponding to the chips CD4007 and HEF4007 are presented in Figure 4.

Table 3: V_t and $\mathcal{K}$ for chips CD4007 and HEF4007.

	V_t	$\mathcal{K}$
CD4007	2.66 V	0.938 $10^{-3}\,(\Omega V)^{-1}$
HEF4007	2.63 V	1.288 $10^{-3}\,(\Omega V)^{-1}$

3.2 Chua's circuit simulation

A large variety of attractors are found for the Chua's circuit. Some of these attractors are: a limit cycle (1T), shown in Figure 5(a). A doubling-period transition to chaotic behaviour is observed. In Figure 5(b) an orbit with two periods (2T) is presented. An orbit with four periods is presented in Figure 5(c). A chaotic attractor, called Rössler attractor is presented in Figure 5(d). Another characteristic chaotic attractor of this circuit is the Double Scroll (DS), which is presented in Figure 5(e). Finally, an outer limit cycle (CL) is observed. This orbit is presented in Figure 5(f).

Using as a guide the α-β bifurcation diagram presented in [3] and [5] where

$$R = \frac{-1.143}{G_a}\ , \quad C_2 = \alpha C_1\ , \quad L = \frac{C_2 R^2}{\beta}\ , \tag{5}$$

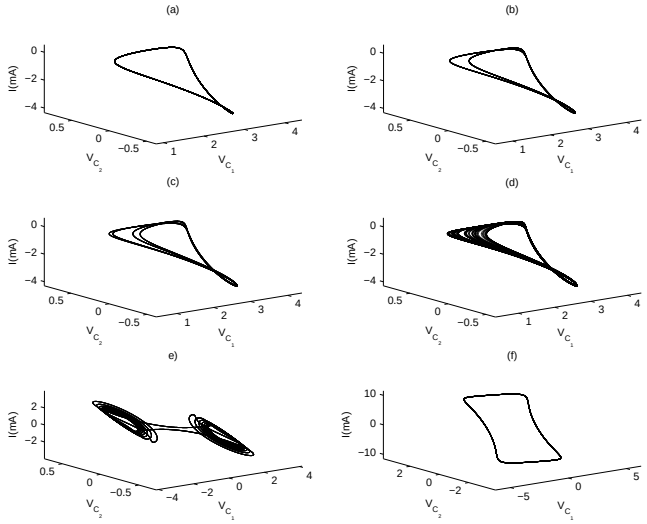


Figure 5: Typical attractors of Chua's circuit.

being G_a is the slope at the origin of the characteristic function $g(V)$ of the nonlinear resistor, we have selected the chip CD4007 and fixed the values $C_2 = 150\text{nF}$, $L = 20\text{mH}$, and $R_L = 39\Omega$. Varying the values of R_D and R_s we have obtained characteristic functions with different values of G_a, which using relations (5) implies different values of β. For each value of β, the different attractors of Chua's circuit are obtained for different values of C_1. In Figure 6 we show the diagram α-β obtained for chip CD4007. We observe that, for large values of β, that is, for low absolute values of G_a the interval of values of α where the different attractors appear is wider than configurations of Chua's circuit with lower values of β. This makes this configuration easier to be observed in an experimental device.

4 Conclusions

An implementation of a cubic-like Chua's oscillator has been presented in [5], which is based on a nonlinear resistor realized with four transistors that constitute two CMOS inverters. In this paper a new realization of the nonlinear resistor is proposed introducing only two new linear resistors. This new realization allows to change the shape of the characteristic function easily. A mathematical model is proposed to simulate the characteristic of the new nonlinear resistor and its validity is tested using experimental data obtained from real circuits.

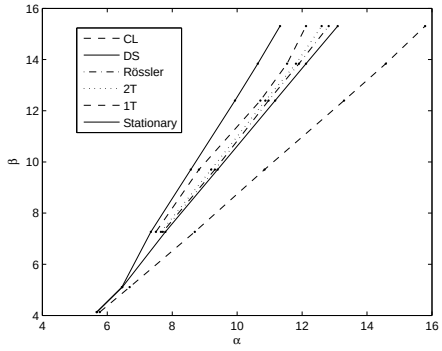


Figure 6: Diagram α-β for chip CD4007.

References

[1] T. Matsumoto, *A Chaotic Attractor from Chua's Circuit*. IEEE Transactions on Circuits and Systems, Vol. Cas-31, No 12, 1055-1058, (1984).

[2] G. Zhong, F. Ayrom, *Periodicity and Chaos in Chua's Circuit*. IEEE Transactions on Circuits and Systems, Vol. Cas-32, No 5, 501-503, (1985).

[3] L.O. Chua, M. Komuro, T. Matsumoto, *The Double Scroll Family*. IEEE Transactions on Circuits and Systems, Vol. Cas-33, No 11, 1073-1118, (1986).

[4] M.P. Kennedy, *Robust OP AMP Realization of Chua's Circuit*. Frequenz, Vol 46, No 3-4, 66-80, (1992).

[5] K. O'Donoghue, M.P. Kennedy, P. Forbes, M. Qu, S. Jones, *A fast and simple implementation of Chua's oscillator with cubic-like nonlinearity*. Interantional Journal of Bifurcation and Chaos, Vol. 15, No 9, 2950-2971, (2005).

[6] A.S. Sedra, K.C. Smith, *Microelectronic Circuits*. Oxford University Press, (2004).

Time Integration of the Neutron Diffusion Equation on Hexagonal Geometries

S. González-Pintor[a], D. Ginestar[b], G. Verdú[a]

[a] Departamento de Ingeniería Química y Nuclear.
Universidad Politécnica de Valencia.
Camino de Vera 14, E-46022 Valencia. Spain.

[b] Instituto de Matemática Multidisciplinar.
Universidad Politécnica de Valencia.
Camino de Vera 14, E-46022 Valencia. Spain.

December 10, 2009

1 Introduction

New nuclear power plants are mainly being built in east countries of Europe and Asia and most of the new reactors under construction are of VVER type. These reactors differ from the PWR and BWR, of western design, mainly in the geometry of fuel elements. In the BWR and PWR the fuel elements are rectangular prisms and in the VVER reactors the fuel elements are hexagonal prisms, in this way, it is interesting to develop efficient nuclear reactors simulators based on hexagonal meshes.

Under general assumptions, the neutronic population inside a nuclear power reactor can be modelled by the time dependent neutron diffusion equation in the approximation of two energy groups. This model is of the form [1]

$$\begin{aligned} \left[v^{-1}\right] \frac{\partial \Phi}{\partial t} + \mathcal{L}\Phi &= (1-\beta)\mathcal{M}\Phi + \sum_{k=1}^{K} \lambda_k \chi \mathcal{C}_k \ , \\ \frac{\partial \mathcal{C}_k}{\partial t} &= \beta_k [\nu\Sigma_{f1} \ \nu\Sigma_{f2}]\Phi - \lambda_k \mathcal{C}_k \ , \quad k = 1, \ldots, K \ , \end{aligned} \tag{1}$$

where, K is the number of delayed neutron precursors groups considered and standard notation is used.

Associated with this problem there is the generalized eigenvalue problem

$$\mathcal{L}\Phi_i = \frac{1}{k_i}\mathcal{M}\Phi_i \ , \tag{2}$$

known as the Lambda Modes problem. The fundamental eigenvalue (the largest one) is called the k-effective of the reactor core, and this eigenvalue and its corresponding eigenfunction describe the steady state neutron distribution in the core.

To solve both problems (1) and (2), a spatial discretization of the equations has to be selected. Once this discretization has been selected, the semidiscrete version of the time dependent neutron diffusion equation is solved. Since the ordinary differential equations resulting of the discretization of diffusion equations are, in general, stiff, implicit methods are necessary for the time discretization. We have used first a finite differences method, that needs to solve a large system of linear equations for each time step. With the aim of reducing the computational cost of this method, we have also studied a modal method based on expanding the neutron flux in terms of the dominant Lambda modes of the reactor core.

2 Spatial discretization

In a nuclear reactor core with hexagonal geometry, the spatial mesh is naturally defined by the different compositions of the materials that compose the core.

To carry out the spatial discretization of the equation each hexagon is divided into six equilateral triangles and the neutron flux is expanded into a set of polynomial functions [4] over each element of the mesh, such as it is shown in [2], A high order finite element method has been derived from a variational formulation where the neutron diffusion equation can be expressed as a stationary condition for a suitable functional [3]. Then, a semi-discrete system of equations is obtained

$$\begin{aligned} \left[v^{-1}\right]\dot{\psi} + L\psi &= (1-\beta)M\psi + \sum_{k=1}^{K}\lambda_k X C_k \ , \\ X\dot{C}_k &= \beta_k M\psi - \lambda_k X C_k \ , \quad k = 1, \ldots, K \ . \end{aligned} \tag{3}$$

Albedo boundary conditions have been considered. The discrete Lambda modes problem is of the form

$$L\psi_l = \frac{1}{k_l} M\psi_l \ . \tag{4}$$

3 Time discretization

3.1 Implicit method

First, we consider a one-step implicit finite differences method for the time dependent neutron diffusion equation. This method consists of integrating the ordinary differential equations (3) over a series of time intervals $[t_n, t_{n+1}]$, and using a one-step backward method, [6], obtaining

$$\left[T^{n+1}\right] \psi^{n+1} = \left[R^n\right] \psi^n + \sum_{k=1}^{K} \lambda_k e^{-\lambda_k h} X \left[C_k^n\right] \ , \tag{5}$$

where the upper index n refers to the matrices obtained with the cross sections at time t_n, T and R are defined as follows

$$\begin{aligned} \left[T^{n+1}\right] &= \frac{1}{h}\left[v^{-1}\right] + L^{n+1} - (1-\beta)M^{n+1} - \left(\sum_{k=1}^{K} \lambda_k \beta_k b_k\right) M^{n+1} \ , \\ \left[R^n\right] &= \frac{1}{h}\left[v^{-1}\right] + \left(\sum_{k=1}^{K} \lambda_k \beta_k a_k\right) M^n \ , \end{aligned}$$

and the coefficients a_k and b_k are given by

$$a_k = \frac{(1+\lambda_k h)(1-e^{-\lambda_k h})}{\lambda_k^2 h} - \frac{1}{\lambda_k} \ , \quad b_k = \frac{\lambda_k h - 1 + e^{-\lambda_k h}}{\lambda_k^2 h} \quad .$$

3.2 Modal method

To solve equations (3) using a modal approximation [7], we assume that ψ can be expressed approximately as

$$\psi = \sum_{l=1}^{M_d} n_l(t)\psi_l \quad , \tag{6}$$

where ψ_l, $l = 1, \ldots, M_d$ are the dominant Lambda modes of a given configuration of the core. A small amount of the dominant Lambda modes and their corresponding adjoint modes can be efficiently computed using, for example, the implicit restarted Arnoldi method [5].

Multiplying equations (3) by the adjoint modes $\psi_m^\dagger$, writing

$$L = L_0 + \delta L \ , \quad M = M_0 + \delta M \ ,$$

and making use of expansion (6) and the biorthogonality of the modes and their corresponding adjoint modes, it is obtained

$$\sum_{l=1}^{M_d} \Lambda_{ml} \frac{d}{dt} n_l = (\rho_m - \beta) N_m n_m + (1-\beta) \sum_{l=1}^{M_d} A_{ml}^M n_l - \sum_{l=1}^{M_d} A_{ml}^L n_l + \sum_{k=1}^{K} \lambda_k C_{mk} \ ,$$

$$\frac{d}{dt} C_{mk} = \beta_k N_m n_m + \beta_k \sum_{l=0}^{M_d} A_{ml}^M n_l - \lambda_k C_{mk} \ , \quad k = 1, \ldots, K \ , \tag{7}$$

where

$$\Lambda_{ml} = \left\langle \psi_m^\dagger, [v^{-1}] \psi_l \right\rangle \ , \qquad A_{ml}^L = \left\langle \psi_m^\dagger, \delta L \psi_l \right\rangle \ ,$$
$$A_{ml}^M = \left\langle \psi_m^\dagger, \delta M \psi_l \right\rangle \ , \qquad C_{mk} = \left\langle \psi_m^\dagger, X C_k \right\rangle ,$$

and the mode m reactivity is defined as $\rho_m = (k_m - 1)/k_m$

These equations can be rewritten in the following matrix form

$$\frac{d[n]}{dt} = [\Lambda]^{-1} \left([\rho - \beta I][N][n] + (1-\beta)[A^M][n] - [A^L][n] + \sum_{k=1}^{K} \lambda_k [C_k] \right) \ ,$$

$$\frac{d[C_k]}{dt} = \beta_k [N][n] + \beta_k [A^M][n] - \lambda_k [C_k] \ , \quad k = 1, \ldots, K \ . \tag{8}$$

To solve this system we have used the method implemented in the subroutine LSODE [8].

For realistic transients, the nuclear cross-sections are time dependent functions and to use the modal method a large amount of modes are necessary. This is numerically prohibitive from the computational point of view. Thus, instead of this, we use a small number of modes together with an updating modes strategy that is performed each certain *updating time step* [7].

4 Numerical Results

To test the methods presented above in 2-D geometries, a transient for a bidimensional VVER 440 reactor core has been studied. This transient is based on the 3-dimensional transient benchmark AER-DYN-001 proposed in [9].

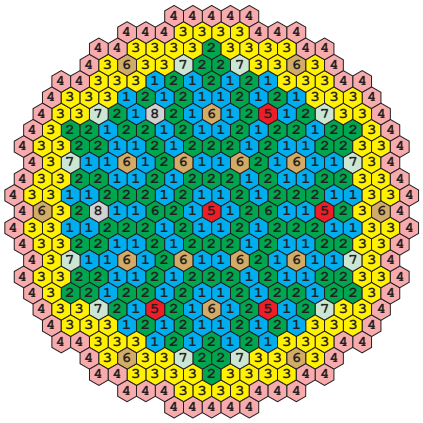


Figure 1: VVER 440 2D

The nuclear cross section given in the 3D benchmark have been collapsed in a single plane. Afterward, materials of the bidimensional reactor have been defined as is shown in Figure 1. The cross sections ot the materials and the neutron precursors parameters for the transient are shown in Tables 1 and 2.

Table 1: Cross sections for the 2-D VVER 440 reactor.

	D_1	D_2	Σ_{a1}	Σ_{a2}	Σ_{12}	$\nu\Sigma_{f1}$	$\nu\Sigma_{f2}$
1	1.346557	0.370075	0.008312	0.064282	0.016976	0.004413	0.072784
2	1.337728	0.367411	0.008745	0.079145	0.016000	0.005491	0.104256
3	1.332264	0.363171	0.009411	0.099536	0.014974	0.006990	0.147261
4	1.447520	0.251741	0.000933	0.033037	0.032215	0.000000	0.000000
5	1.231711	0.240027	0.012120	0.118846	0.020782	0.001345	0.027352
6	1.337727	0.367479	0.008747	0.079153	0.015996	0.005492	0.104316
7	1.346561	0.370177	0.008317	0.064282	0.016968	0.004416	0.072846
8	1.231640	0.239942	0.012123	0.118870	0.020785	0.001342	0.027299

A transient that simulates the movement of two control rods has been defined by means of the time evolution of the absorption cross section Σ_{a2}

Table 2: Neutron precursors parameters for the reactor VVER 440.

	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6
β_i	0.000247	0.0013845	0.001222	0.0026455	0.000832	0.000169
λ_i (s^{-1})	0.0127	0.0317	0.115	0.311	1.4	3.87

for the material 8 as follows,

$$\Sigma_{a2}(t) = \begin{cases} 0.118870 \cdot (1-t) + 0.016917 \cdot t & \text{si} \quad 0 \leq t \leq 1 \ , \\ 0.118870 \cdot (t-1) + 0.016917 \cdot (2-t) & \text{si} \quad 1 \leq t \leq 2 \ , \\ 0.118870 & \text{si} \quad 2 \leq t \leq 3 \ . \end{cases}$$

The mean power evolution computed using the implicit method with $\Delta t = 0.0001$ s is shown in Figure 2. This curve is taken as a reference for the average power evolution.

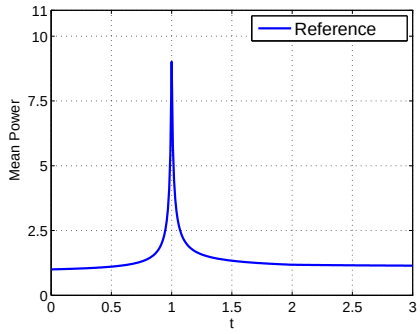


Figure 2: Reference result.

To study the dependence of the obtained solution on the spatial accuracy, in Figure 3 the mean power evolution and the local power distribution on the axis $y = 0$ at time $t = 1$ s are shown for different values of the number of polynomials used in the finite element method k. These results are computed with the implicit method and $\Delta t = 0.0001$ s. We observe that using $k = 3$ polynomials the obtained results are quite accurate. Thus, henceforth all the calculations are performed setting $k = 3$.

For the Modal method presented above, the first five dominant modes of the reactor have been considered and these modes have been updated each certain time step Δt_u. Table 3 shows the relative errors in the determination of the peak of power obtained with the implicit method and the modal method with different values of the time steps Δt and Δt_u. Also the CPU

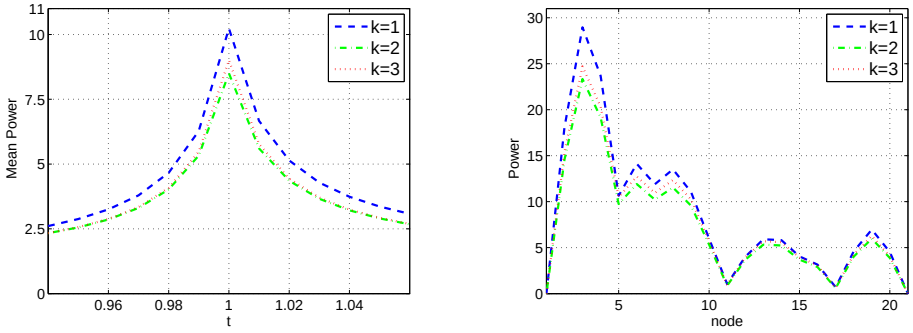


Figure 3: Mean power evolution and local power distribution for $t = 1$, $\Delta t = 0.0001$s, over the axis $y = 0$ for different values of k.

times necessary to calculate the transient in the different cases has been presented in Table 4. The implicit and the Modal method have been implemented in a Fortran program an have been run in a Intel Core 2 Duo 1.86 GHz computer with 1 Gb RAM .

Table 3: Percentage of relative error to the peak at $t = 1$.

	Backward Method Time step			
	$\Delta t = 0.05s$	$\Delta t = 0.01s$	$\Delta t = 0.001s$	$\Delta t = 0.0001s$
	4.57	0.37	0.01	reference
	Modal Method Updating time step			
Num. of Modes	No updates	$\Delta t_u = 1s$	$\Delta t_u = 0.1s$	$\Delta t_u = 0.01s$
1	86.62	14.93	2.23	7.68
3	86.68	13.14	5.56	1.80
5	86.24	13.06	5.81	2.17

Table 4: Errors and CPU time for transient calculations.

	Backward Method Time step			
	$\Delta t = 0.05s$	$\Delta t = 0.01s$	$\Delta t = 0.001s$	$\Delta t = 0.0001s$
	1 m 44 s	7 m 22 s	66 m 20 s	10 h 12 m 02 s
	Modal Method Updating time step			
Num. of Modes	No updates	$\Delta t_u = 1s$	$\Delta t_u = 0.1s$	$\Delta t_u = 0.01s$
1	7 m 55 s	8 m 27 s	10 m 50 s	27 m 20 s
3	8 m 01 s	8 m 28 s	11 m 41 s	43 m 16 s
5	8 m 23 s	8 m 50 s	11 m 27 s	39 m 02 s

It is important to remark that if the number of modes used in the ex-

pansion of the modal method is low the modes updating strategy has an important effect on the accuracy of the results.

5 Conclusions

Two methods for the solution of the time dependent neutron diffusion equation based on hexagonal spatial mesh have been studied. A one-step implicit method and a modal method. The implicit method needs to solve a large system of linear equations for each time step, thus, it is an expensive method from the computational point of view. With the aim of reducing the computational time a modal method has been proposed based on the dominant Lambda modes of the reactor core. The modal method succeeds at reducing the computational time maintaining a reasonable accuracy.

Acknowledgments

This work has been partially supported by the Spanish Ministerio de Educación y Ciencia under projects ENE2008-02669 and MTM2007-64477-AR07.

References

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

[2] S. González-Pintor, D. Ginestar, G. Verdú, High Order Finite Element Method for the Lambda Modes problem on hexagonal geometry. Annals of Nuclear Energy, 36 (2009) 1450–1462.

[3] A. Hébert, A Raviart-Thomas-Schneider solution of the diffusion equation in hexagonal geometry, Annals of Nuclear Energy 35 (3) (2008) 363–376.

[4] G. EM. Karniadakis, S. Sherwin, Spectral/hp Element Methods for Computational Fluid Dynamics, Oxford University Press, Oxford, 2005.

[5] G. Verdú, R. Miró, D. Ginestar, V. Vidal, The implicit restarted Arnoldi method, an efficient alternative to solve the neutron diffusion equation, Annals of Nuclear Energy 26 (1999) 579-593.

[6] D. Ginestar, G. Verdú, V. Vidal, R. Bru, J. Marín, J.L. Muñoz-Cobo, High Order Backward Discretization of the Neutron Diffusion Equation, Annals of Nuclear Energy 25 (1-3) (1998) 47-64.

[7] R. Miró, D. Ginestar, G. Verdú, D. Hennig, A nodal modal method for the neutron diffusion equation. Application to BWR instabilities analysis. Annals of Nuclear Energy 29 (2002) 1171-1194.

[8] A.C. Hindmarsh, ODEPACK, A Systematized Collection of ODE Solvers in Scientific Computing, North-Holland, 1983.

[9] A. Keresztúri, M. Telbisz, A Three Dimensional Hexagonal Kinetic Benchmark Problem, 2nd AER Symposium, Paks Hungary, 1992.

Measuring Performance in the Banking Sector.

F. García, F. Guijarro, and I. Moya *

Facultad de Administración y Dirección de Empresas,

Universidad Politécnica de Valencia,

Edificio 7J, 46022 Valencia, España.

December 10, 2009

1 Introduction

The aim of the present work is to propose a goal programming-based multicriteria methodology to provide a measure of company performance and then apply it to the ranking of Spanish savings banks according to their performance in 2007. This sector was chosen for its importance in the Spanish financial system at a time when radical changes are expected to take place in its structure. It is therefore of interest to determine how potential mergers would affect the relative positions of the savings banks from a global standpoint rather than from a single-criterion, bearing in mind that these banks have no share quotations on the stock market.

2 Selection of Representative Performance Variables

In the studies on financial performance, there are notable differences in the methodologies, criteria and information used. However, it is possible to identify certain recurring general criteria that determine strategic bank policy. If we concentrate on studies on the Spanish financial system and savings banks,

*imoya@esp.upv.es

we can group these criteria into: productivity, (Grifell-Tatjé and Lovell, 1997), costs (Prior, 2003), profitability (Lozano-Vivas, 1997), management of different types of risk, and size (Marco and Moya, 2000).

To select the variables used in the present study, those used in the above-cited works were considered, as were all the different business aspects (productivity, profits, risk management and size). Table 1 shows these variables together with the dimensions they represent.

The data base for the present study was compiled from the annual accounts published by the Spanish savings banks for the financial year 2007.

The ranking of business companies, or in this case savings banks, is usually done on the basis of a single variable. This type of ranking refers only to the situation with reference to the criterion used and gives no information on the overall situation of an individual company within the sector. In order to carry out a multi-criteria ranking, various explanatory variables or single-criterion rankings must be available.

Table 1. Variables used for the multi-criteria ranking

Variable	Dimension
RLS: Lending / Staff	Productivity
RDB: Deposits / Number of branches	Productivity
ROA: Results before taxes / Total assets	Profitability
ROE: Results before taxes/ Equity	Profitability
IDR: Inverse default rate	Credit risk
SOLV: Coefficient of solvency	Credit risk
PBD: Provision for bad debts	Credit risk
TA: Total assets	Size

The first problem to be overcome is how to organise all this information, minimising the impact of the least important factors and emphasising the most important or most representative of the general tendency.

The second problem is how to weight the variables used in the multi-criteria performance rating, minimising as far as possible the subjectivity of the person who decides the weightings.

The proposal made in this paper differs from previous studies (Diakoulaki *et al.*,1995; Deng *et al.*, 2000) , in the method by which it obtains multi-criteria performance. Using the multi-criteria goal-programmming technique, weights are calculated in such a way that the similarity is maximum between

values standardised by the range of the different criteria and the multi-criteria performance, which is the value which will later be used to rank the companies.

$$
\begin{aligned}
&Min \quad \lambda \sum_{i=1}^{n} \sum_{j=1}^{c} (n_{ij} + p_{ij}) + (1-\lambda)\, D \\
&s.t. \\
&\sum_{j=1}^{c} w_j v_{ij} + n_{ij} - p_{ij} = v_{ij} \quad i = 1 \ldots n \quad j = 1 \ldots c \\
&\sum_{i=1}^{n} (n_{ij} + p_{ij}) \leq D \quad j = 1 \ldots c \\
&\sum_{j=1}^{c} w_j = 1 \qquad [1] \\
&\sum_{j=1}^{c} w_j v_{ij} = V_i \quad i = 1 \ldots n \\
&\sum_{i=1}^{n} (n_{ij} + p_{ij}) = D_j \quad j = 1 \ldots c \\
&\sum_{j=1}^{c} D_j = Z
\end{aligned}
$$

Where:

w_j = weight to be estimated for the jth criterion.

$n_{ij}\,(p_{ij})$ = negative deviation variable (positive). Quantifies the difference by excess (defect) between the value of the ith company in the jth criterion and the multi-criteria value obtained by applying the weights w_j. This is, $n_{ij} - p_{ij} = v_{ij} - \sum_{j=1}^{c} w_j v_{ij}$, with $n_{ij}, p_{ij} \geq 0$. The objective function of [1] ensures that only one of the deviation variables can have a value greater than zero: $n_{ij} \times p_{ij} = 0$

D_j= degree of disagreement between the jth criterion and the multi-criteria value.

Z = magnitude of global disagreement.

3 Ranking of Spanish Savings Banks for Financial Year 2007

This section describes the use of the methodology presented in the previous Section to obtain a multi-criteria performance ranking of the Spanish savings banks for 2007.

The 8 criteria shown in Table 1 are used as a starting point. The values are standardised by range, according to normal procedure. The extended

goal programming model is applied to these data [1].

On solving [1] for different values of λ we obtain firstly the weighting or relative importance of each individual criterion in the overall ranking and secondly the multi-criteria value which ranks the banks according to performance.

The model for λ=1 obtains non-null coefficients for all variables, that of before-tax results on equity (ROA) being noteworthy for its greater weight. We can thus conclude that the profits criterion is representative of the general performance tendency. Indeed, if we add together the weights obtained by the variables which represent this dimension, a value of 43.5% is obtained. The second most important dimension is credit risk (23.9%), followed closely by productivity (22.6%). The least important, at 10% for its only variable, is total assets.

Figure 1 reports the weight of each of the dimensions contained in the analysis, obtained as the sum of the individual weights of each criterion. It can be clearly seen that as the value of λ diminishes, Profitability loses part of its weight to Productivity and Credit Risk, while Size remains around a discreet 10% or even lower for the entire range of values analysed.

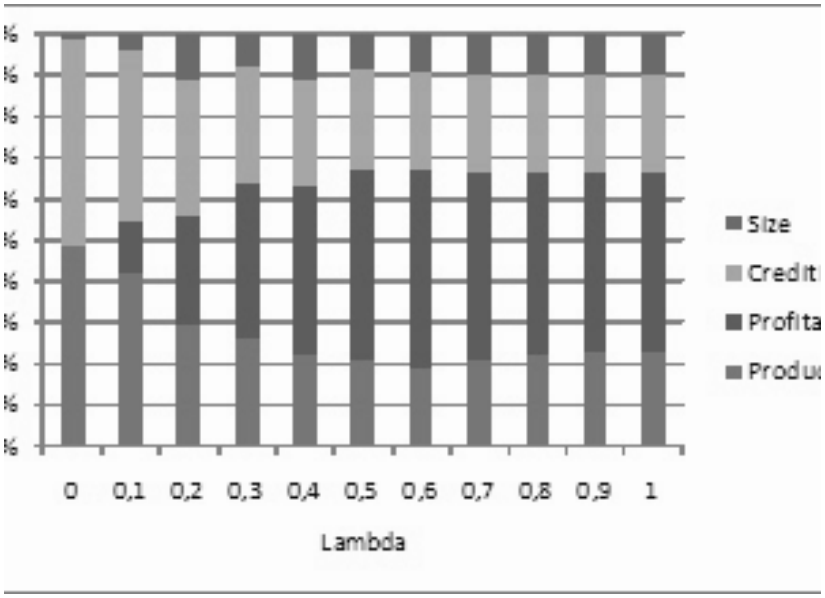
Figure 1. Results of applying the extended goal programming model

Although the weight of each criterion, or the set calculated for the dimension, offers an idea of the relative importance of each measurement in calculating multi-criteria performance, a Spearman's correlation analysis must be carried out to analyse the correlation between each of the single-criterion measurements and final performance.

The highest ranked saving banks are: BBK, Cajastur, La Caixa and Caja Murcia, the worst ranked ones are: Caixa d'Estalvis de Sabadell, Caixa de Girona, Caixa Penedés, Caixa Manlleu, Caixa Tarragona, Caja España, Cajasol, Caja Canarias and Cajasur.

4 Conclusions

Performance analysis can provide a great deal of important imformation on business companies. The objective of the present study was to propose a goal-programming based multi-criteria methodology which would provide a global estimation of the performance of a business company, combining the individual criteria in such a way as to include all the dimensions that affect



its performance. The proposed methodology was then used to obtain a multi-criteria ranking of the Spanish savings banks for the year 2007.

The methodology proposed in this paper differs from others by the way in which global performance is estimated. By means of the goal programming multi-criteria technique weightings are calculated so as to maximise similarities between values standardised by the range of the different criteria and multi-criteria performance, which is the value which will subsequently be used to rank the companies according to performance. Applying different versions of the goal programming model, a collective approach and an individualistic approach are considered. As a compromise solution between the

two approaches, a parametric version is developed.

REFERENCES

[1] Deng, H., Yeng, C-H. y Willis, R.J., 2000. Inter-company comparison using modified TOPSIS with objective weights. *Computers & Operations Research*, 27(10), 963-973.

[2]Diakoulaki, D., Mavrotas, G. y Papayannakis, L., 1995. Determining objective weights in multiple criteria problems: the CRITIC method. *Computers & Operations Research*, 22(7), 763-770.

[3] Grifell-Tatjé, E., y Lovell, C.A.K., 1997. The sources of productivity change in Spanish banking. *European Journal of Operational Research*, 98, 365-381.

[4] Marco, M.A., Moya, I., 2000. Factores que inciden en la eficiencia de las entidades de crédito cooperativo. *Revista Española de Financiación y Contabilidad*, 105, 781-808.

[5] Prior, D., 2003. Long and short-run nonparametric cost frontier efficiency: an application to Spanish savings banks. *Journal of Banking and Finance*, 27, 107-123.

[6] Zeleny, M., 1982. *Multiple Criteria Decision Making*. New York: McGraw-Hill.

Positive solutions of discrete dynamic Leontief input-output model with possibly singular capital matrix

L. Jódar*and P. Merello†

Instituto Universitario de Matemática Multidisciplinar,

Universidad Politécnica de Valencia,

Edificio 8G, 2º, 46022 Valencia, Spain

December 10, 2009

1 Introduction

A dynamic Leontief model of a multisector economy has the form

$$x_n = Lx_n + C[x_{n+1} - x_n] + d_n\,, \tag{1}$$

where x_n is the vector of output levels, d_n is the vector of final demands (excluding investment), L is the Leontief input-output matrix, and C is the capital coefficient matrix. The matrices L and C are assumed known and time invariant what means that market and technology do not change under considered time period. If the dimension of the vector x_n is r, and the system (1) is assumed to operate over the period $n = 0, 1, \ldots, N-1$, then for a specified set of final demands the set of equations represents a set of rN equations in $r(N+1)$ unknowns. A solution to this set is called a trajectory of the Leontief system.

*ljodar@imm.upv.es

†pamegi@ade.upv.es

If the matrix C were invertible, the trajectory of the model (1) is computationally easy to obtain by means of the expression

$$x_{n+1} = C^{-1}\left\{[I - L + C]x_n - d_n\right\},$$

starting with an arbitrarily specified initial x_0.

Unfortunately, the assumption of C being nonsingular is usually not justified. The element c_{ij} of matrix C represents the amount of stock of commodity i, as a capital good, that sector j must have on hand for each unit of production. Since not every sector produces significant capital goods (agriculture being a typical example in many models), it is common for some rows of the matrix C to contain only zero elements. Thus, the very structure of capital requirements in a multisector economy often dictates that C be singular. In fact, the rank of C may be much smaller than r, the number of sectors.

The problem of computing Leontief trajectories when C is singular has been treated in [1], [2], [3], [4], but in these papers the authors do not pay attention to the positive character of solutions. An algebraic-geometric approach in order to guarantee positive solutions for the closed Leontief dynamic input-model may be found in [5]. Positive solutions for a different Leontief price model have been treated in [6] using an algebraic approach based on the concept of nonnegative generalized inverse.

In this paper we construct trajectories of the Leontief model (1) by using firstly the solution of singular systems of difference equations based on the Drazin inverse approach developped in [7], [8]. Then, nonnegative solutions are obtained by using matrix and vector inequalities.

2 Preliminaries and matrix inequalities

One of the main goals of this paper is to guarantee that solutions of systems of difference equations modeling discrete and dynamic Leontief economic problems are positive. This is a crucial requirement in order to have realistic solutions.

For the sake of clarity in the presentation of results of the next sections we include here some definitions and results related to matrix inequalities.

Let $A = (a_{ij})$ be a real square matrix of size n, element of $\mathbb{R}^{r\times r}$. We say that A is positive, if all its entries a_{ij} are positive real numbers. We say that A is nonnegative, if $a_{ij} \geq 0$ for all $1 \leq i, j \leq n$. If A is positive, we write

$A > 0$, and if A is nonnegative we will write $A \geq 0$. The matrix A is said to be negative, denoted $A < 0$, if $a_{ij} < 0$ for all $1 \leq i, j \leq n$. If a is a real number, we say $A \leq a$, if all entries of A are equal or smaller than a.

If A, B are square nonnegative matrices of size r, such that

$$A \leq a\,,\ B \leq b\,,$$

then from the definition of the matrix product one gets

$$AB \leq r A_{max} B_{max}\,, \tag{2}$$

where $A_{max} = \max\{a_{ij};\ 1 \leq i, j \leq r\}$; $B_{max} = \max\{b_{ij};\ 1 \leq i, j \leq r\}$.

A vector $x \in \mathbb{R}^r$ is said to be positive, denoted $x > 0$, if all its components x_i are positive. The vector x is said to be nonnegative if $x_i \geq 0$ for all $1 \leq i \leq n$, and in this case we write $x \geq 0$. The notation $x \geq y$ is equivalent to $x - y \geq 0$. If $x \geq y$, $A \geq 0$, then it is easy to check that $Ax \geq Ay$. In particular, if $x \geq 0$ and $A \geq 0$, then $Ax \geq 0$.

Let A be a matrix in $\mathbb{R}^{r\times r}$, then we denote

$$d_{min}(A) = \min\{a_{ii};\ 1 \leq i \leq n\}\ ;\ \ d_{max}(A) = \max\{a_{ii};\ 1 \leq i \leq n\}\ .$$

The Euclidean norm of a vector $x \in \mathbb{R}^r$ will be denoted by $\|x\|$ with $\|x\| = \left(x^T x\right)^{1/2} = \left(|x_1|^2 + \cdots + |x_n|^2\right)^{1/2}$. If A is a matrix in $\mathbb{R}^{r\times r}$, its 2-norm is denoted by $\|A\|$, and from [9, p.57] one gets

$$\max_{i,j} |a_{ij}| \ \leq \|A\| \leq r\ \max_{i,j} |a_{ij}|\ .$$

3 Positive solutions of singular systems of difference equations

Consider the inhomogeneous system of difference equations

$$A\,x_{n+1} = B\,x_n + f_n\,, \tag{3}$$

where A, B are matrices in $\mathbb{R}^{r\times r}$ and f_n are vectors in $\mathbb{R}^r$.

Assume that the homogeneous system is tractable in the sense of [8, p.182], i.e., there exists $\lambda \in \mathbb{C}$ such that $(\lambda A - B)^{-1}$ exists, and let $k = \mathrm{Ind}(\widehat{A})$

where $\widehat{A} = (\lambda A - B)^{-1}A$. Let $\widehat{B} = (\lambda A - B)^{-1}B$ and $\widehat{f}_i = (\lambda A - B)^{-1}f_i$. Then from theorem 9.3.2 of [8], the general solution of (3), is given by

$$x_n = (\widehat{A}^D\widehat{B})^n\widehat{A}\widehat{A}^Dq + \widehat{A}^D\sum_{i=0}^{n-1}(\widehat{A}^D\widehat{B})^{n-i-1}\widehat{f}_i - (I - \widehat{A}\widehat{A}^D)\sum_{i=0}^{k-1}(\widehat{A}\widehat{B}^D)^i\widehat{B}^D\widehat{f}_{n+i}\,,$$

and a vector $c \in \mathbb{R}^r$ is a consistent initial vector if and only if c lies in $\left\{\widehat{\omega} + R(\widehat{A}^k)\right\}$, where

$$\widehat{\omega} = -(I - \widehat{A}\widehat{A}^D)\sum_{i=0}^{k-1}(\widehat{A}\widehat{B}^D)^i\widehat{B}^D\widehat{f}_i\,.$$

Taking norms in the expression of vector $z_k(n)$ defined by

$$z_k(n) = -(I - \widehat{A}\widehat{A}^D)\sum_{i=0}^{k-1}(\widehat{A}\widehat{B}^D)^i\widehat{B}^D\widehat{f}_{n+i}\,,$$

it follows that,

$$\|z_k(n)\| \leq \left\|I - \widehat{A}\widehat{A}^D\right\|\sum_{i=0}^{k-1}\left\|\widehat{A}\widehat{B}^D\right\|^i\left\|\widehat{B}^D\right\|\left\|\widehat{f}_{n+i}\right\|\,,$$

and

$$z_k(n) \leq \left\|I - \widehat{A}\widehat{A}^D\right\|\sum_{i=0}^{k-1}\left\|\widehat{A}\widehat{B}^D\right\|^i\left\|\widehat{B}^D\right\|\left\|\widehat{f}_{n+i}\right\|\,. \tag{4}$$

Theorem 3.1 *Let A, B be matrices in $\mathbb{R}^{r\times r}$ such that*

- *(i) All the diagonal elements of matrices $\widehat{A}$, $\widehat{A}^D$ and $\widehat{B}$ are nonzero.*
- *(ii) $\widehat{A} \leq 0$, $\widehat{A}^D \leq 0$ and $\widehat{B} \leq 0$.*
- *(iii) Vector $\widehat{f}_i \geq 0$ and $q \geq 0$.*

Then the vector expression

$$x_n = L_n(q) + z_k(n)\,, \tag{5}$$

with

$$L_n(q) = (\widehat{A}^D\widehat{B})^n\widehat{A}\widehat{A}^Dq + \widehat{A}^D\sum_{i=0}^{n-1}(\widehat{A}^D\widehat{B})^{n-i-1}\widehat{f}_i\,, \tag{6}$$

and $z_k(n)$ *given by (3), is nonnegative, if for* $1 \le n \le N$ *vector* q *satisfies*

$$q \ge \frac{\left\|I-\widehat{A}\widehat{A}^D\right\|\sum_{i=0}^{k-1}\left\|\widehat{A}\widehat{B}^D\right\|^i\left\|\widehat{B}^D\right\|\left\|\widehat{f}_{n+i}\right\| - \sum_{i=0}^{n-1}\left[\left(\widehat{A}^D\right)_{min}\right]^{n-i}\left[\left(\widehat{B}\right)_{min}\right]^{n-i-1} r^{n-i+1}\widehat{f}_i}{\left(d_{max}\left(\widehat{A}^D\right)\right)^{n+1}\left(d_{max}\left(\widehat{B}\right)\right)^n d_{max}\left(\widehat{A}\right)}\,. \tag{7}$$

4 Positive solutions of the discrete dynamic Leontief model

Let us start this section by writing model (1) in its equivalent form

$$C\,x_{n+1} = (I - L + C)x_n - d_n\,,\ \ 0 \le n \le N-1\,, \tag{8}$$

where C and L are matrices in $\mathbb{R}^{r\times r}$ and d_n lies in $\mathbb{R}^r$.

Let us assume that the Leontief matrix L satisfies:

$$\left.\begin{array}{ll} 0 \le \underset{r}{l_{ij}} \le 1\,, & 1 \le i,\, j \le r\,, \\ \displaystyle\sum_{i=1}^{r} l_{ij} \le 1\,, & 1 \le j \le r\,, \\ \displaystyle\sum_{i=1} l_{ij_0} < 1\,, & \text{for some } j_0 \text{ with } 1 \le j_0 \le r\,, \end{array}\right\} \tag{9}$$

and

- All the diagonal entries $\widehat{c}_{ii}$ of $\widehat{C}$ are nonzero and different from one,
- All the diagonal elements of $\widehat{C}^D$ are nonzero,

(10)

and

$$\widehat{C}^D \le 0\,. \tag{11}$$

By [10, p.318-319], under condition (9) the matrix $I - L$ is invertible and

$$(I-L)^{-1} > 0\,.$$

Taking $\lambda = 1$, note that under hypothesis (9), the system

$$Cx_{n+1} = (I - L + C)x_n\,, \text{ is tractable.}$$

Let k be the index of $\widehat{C}$, $k = \mathrm{Ind}(\widehat{C})$, and

$$\widehat{\omega} = \left(I - \widehat{C}\widehat{C}^D\right)\sum_{i=0}^{k-1}\left(\widehat{C}\left(\widehat{C}-I\right)^D\right)^i\left(\widehat{C}-I\right)^D(L-I)^{-1}d_i\,.$$

A vector $q \in \mathbb{R}^r$ is a consistent initial vector for (8) if and only if q lies in $\left\{\widehat{\omega} + R\left(\widehat{C}^k\right)\right\}$ and the general solution of (8) for $n \geq 1$ is given by

$$\begin{aligned} x_n &= \left(\widehat{C}^D\left(\widehat{C}-I\right)\right)^n\widehat{C}\widehat{C}^D q \\ &- \widehat{C}^D\sum_{j=0}^{n-1}\left(\widehat{C}^D\left(\widehat{C}-I\right)\right)^{n-j-1}(I-L)^{-1}d_j \\ &+ \left(I-\widehat{C}\widehat{C}^D\right)\sum_{i=0}^{k-1}\left(\widehat{C}\left(\widehat{C}-I\right)^D\right)^i\left(\widehat{C}-I\right)^D(I-L)^{-1}d_{n+i}\,,\ n \geq 1\,. \end{aligned} \tag{12}$$

Once the solution x_n of model (8) given by (12) has been obtained, the next task is to find conditions on the data problem so that x_n is nonnegative, for $n = 1, 2, \ldots, N$. Let us denote

$$\widehat{d}_i = -(I-L)^{-1}d_i\,,\ 0 \leq i \leq N\,,$$

and let α_n be defined as

$$\begin{aligned} \alpha_n & \\ &= \frac{\|I-\widehat{C}\widehat{C}^D\|\sum_{i=0}^{k-1}\left\|\widehat{C}(\widehat{C}-I)^D\right\|^i\left\|(\widehat{C}-I)^D\right\|\left\|\widehat{d}_{n+i}\right\|}{\left(d_{max}(\widehat{C}^D)\right)^{n+1}\left(d_{max}(\widehat{C}-I)\right)^n d_{max}(\widehat{C})} \\ &+ \frac{\sum_{i=0}^{n-1}\left[\left(\widehat{C}^D\right)_{min}\right]^{n-i}\left[\left(\widehat{C}-I\right)_{min}\right]^{n-i-1}r^{n-i+1}\widehat{d}_i}{\left(d_{max}(\widehat{C}^D)\right)^{n+1}\left(d_{max}(\widehat{C}-I)\right)^n d_{max}(\widehat{C})}\,, \end{aligned} \tag{13}$$

$$\alpha = \max\left\{\alpha_i\,;\ 0 \leq i \leq N\right\}\,. \tag{14}$$

If the affine subspace $\left\{\widehat{\omega} + R\left(\widehat{C}^k\right)\right\}$ contains a vector $q \in \mathbb{R}^r$ having all its components $q_i \geq \alpha$, under hypotheses (9), (10) and (11), the expression (12) is nonnegative. A vector q can be always chosen. In fact, let k be the index of $\widehat{C}$. If k is odd, the $\widehat{C}^k < 0$ because $\widehat{C} < 0$. Let v_0 be a column of $\widehat{C}^k$ and let α_0 be a negative real number with $|\alpha_0|$ big enough so that

$$\alpha_0 v_0 > \alpha - \widehat{\omega}\,,$$

then $q = \alpha_0 v_0 + \widehat{\omega}$ satisfies $q > \alpha$.

If k is even, then $\widehat{C}^k > 0$ and taking a column v_1 of $\widehat{C}^k$ and a positive real number β_1 big enough so that

$$\beta_1 v_1 > \alpha - \widehat{\omega}\,,$$

then $q = \beta_1 v_1 + \widehat{\omega}$ satisfies $q > \alpha$.

5 Conclusions

In this paper we construct trajectories of the Leontief model (1) with possibly singular capital matrix by using Drazin inverses and a singular discrete difference systems approach.

Then, under hypotheses (9), (10) and (11) and with previous notation, there exists $q \in \mathbb{R}^r$ having all its components $q_i \geq \alpha$, where α is defined by (13)-(14), and the corresponding Leontief trajectory x_n given by (12) is nonnegative for $0 \leq i \leq N$. Also, we show that such a vector q can be always chosen and how to construct it.

References

[1] D. Kendrik, On the Leontief Dynamic-inverse, Quarterly Journal of Economics 86 (1972) 693-696.

[2] H.G. Bergendorff, Application of Control Theory to Leontief-planning Models, VIII International Symposion on Mathematical Programming, Stanford, California, 1973.

[3] D.A. Livesey, The Singularity Problem in the Dynamic Input-output Model, International Journal of Systems Science 4 (1973) 437-440.

[4] D.G. Luenberger, A. Arbel, Singular Dynamic Leontief Systems, Econometrica, Vol.45 (1977) No.4.

[5] D.B. Szyld, L. Moledo, B. Sauber, Positive Solutions of the Leontief Dynamic Input-output Model, in Input-output Analysis, M. Ciaschini editor, Chapman and Hall, New York, 1988, pp.91-98.

[6] M.S. Silva, T.P. de Lima, Looking for Nonnegative Solutions of a Leontief Dynamic Model, Linear Algebra Appls. 364 (2003) 281-316.

[7] S.L. Campbell, Singular Systems of Differential Equations, Pitman Pub. Co., London, 1980.

[8] S.L. Campbell, C.D. Meyer, Generalized Inverses of Linear Transformations, Dover, New York, 1979.

[9] G. Golub, C.F. Van Loan, Matrix Computations, John-Hopkins Univ. Press, Baltimore, MA. 1989.

[10] R. Bellman, Introduction to Matrix Analysis, McGraw-Hill, New York, MA. 1965.

Moving mesh strategies for the simulation of direct injection engines

Xandra Margot, Sergio Hoyas [*],
Pablo Fajardo and Stavrouna Patouna [†]

CMT-Motores Térmicos,

Universidad Politécnica de Valencia,

December 10, 2009

1 Introduction

In several situations it is convenient for the computational grid to follow moving boundaries. In the past decade, there have been some efforts in understanding the theoretical backgrounds of moving mesh methods [1], [2]. The basic idea of moving mesh methods is to construct a transformation from a computational domain to the physical one [3], [4]. However, the extension to transient problems does not seem simple.

In diesel engines, the nozzle internal flow greatly affects the fuel atomization characteristics and so the subsequent engine combustion. The transient nature of the flow is greatly affected by the needle movement. There are studies referring to moving mesh CFD studies but limited information is published about the mesh generation and moving grid methodology used to simulate the needle motion [5], [6].

The work starts with a mathematical description of the models. In section 3, the methodology for the moving mesh generation and movement approach is presented. In section 4, the numerical case of the diesel injection simulation will be presented followed by some concluding remarks.

[*]Corresponding author

[†]xmargot,serhocal,pabfape,stapa1@mot.upv.es

2 Models description

The mass and momentum conservation equations are solved using the Navier-Stokes equations in non-steady coordinates, in Cartesian tensor notation [7]:

$$\frac{1}{\sqrt{g}}\frac{\partial}{\partial t}\left(\sqrt{g}\rho\right)+\frac{\partial}{\partial x_j}\left(\rho\tilde{u}_j\right) = s_m, \quad (1)$$

$$\frac{1}{\sqrt{g}}\frac{\partial}{\partial t}\left(\sqrt{g}\rho u_i\right)+\frac{\partial}{\partial x_j}\left(\rho\tilde{u}_j u_i-\tau_{ij}\right) = -\frac{\partial p}{\partial x_i}+s_i, \quad (2)$$

Where t is time, x_i the Cartesian coordinate ($i=1,2,3$), u_i the absolute fluid velocity component in direction x_i, $\tilde{u}_j$ the relative velocity (u_j-u_{cj}) between fluid and local (moving) coordinate frame that moves with velocity u_{cj}, p the piezometric pressure, ρ the density, τ_{ij} stress tensor components, s_m the mass source, s_i the momentum source components and $\sqrt{g}$ is the determinant of the metric tensor.

For this type of problem, an additional equation, the 'space conservation law' is solved for the moving coordinate velocity components (Eq. 14).

$$\frac{1}{\sqrt{g}}\frac{\partial}{\partial \tau}\left(\sqrt{g}\right)+\frac{\partial}{\partial x_j}\left(\rho\tilde{u}_j\right)=s_m \quad (3)$$

This relates the change in cell volume to the cell-face velocity. The simultaneous satisfaction of the space conservation law and all other equations of fluid motion facilitates the general moving mesh operations performed [8]. This problem has been solved using a conventional two-equation $k-\epsilon$ model. The cavitation model implemented in this code is based on the Rayleigh equation [10] which links the rate of change of the bubble radius with the local pressure.

3 Methodology

To apply the method proposed, the first step is to define the initial mesh which is not subject to motion. As long as the initial mesh remains the same, by extracting coordinate information from edge lines and defining vertices, the desired geometry on the whole domain is achieved. These vertices will be subjected to motion and also will be used to generate the mesh which maintains uniformity during mesh motion.

Since the needle lift law as a function of time is not known and could not be measured, the following technique was followed in order to extract it. Six calculations were performed, each at a different needle lift. Thus, the curve injection rate as a function of needle lift was obtained. A correspondence between this curve and the experimentally measured injection rate yields the needle lift law as a function of time The interpolated points between the curve injection rate as a function of needle lift and the measured injection rate were plotted and the slopes found between them programmed into the calculations.

Thus, the velocity is estimated as a function of time using the formula $lift = lift(t_0) + slope\,(Time - t_0)$ which is used to move the needle in the transient calculations.

4 Example

In this problem, the moving area is the annulus formed between the needle and the needle seat while, the bottom part of the injector is considered as initial mesh and is not subject to motion. Once the initial geometry is meshed ($10\mu m$), the vertices that define the contour of the geometry at full lift ($250\mu m$) have to be specified. These vertices then enable to create patches that represent the solution domain The patches are defined in terms of vertices that are not subject to motion and vertices that are subject to motion and then extruded along a specified direction and coordinate system in order to obtain the mesh corresponding to the full needle lift .The extrusion consists of defining the number of cells that the mesh will have in each direction.The meshes of the annulus between the needle and the needle seat stretch and shrink with the needle movement.

The other patches are moved following direction 1. The nozzle seating forms an angle of 30^o with the axis of the nozzle, so the displacement of some patches assumes the value of (needle displacement * $sin(30^o)$).

The simulation was realized for different operating conditions. In particular, the effect of cavitation on the nozzle flow was studied. Additional results of the diesel injector flow behavior of this study are documented in [11] and [12].

5 Conclusions

In this paper, we have presented a simple moving mesh strategy for solving a multi-phase flow in a diesel injector nozzle. The approach for the needle movement required an interpolation between the experimental results and the results with calculations at fixed needle lifts. The moving mesh computation was satisfactory in terms not only of accuracy but also efficiency.

References

[1] T. Linß, Uniform pointwise convergence of finite difference schemes using grid equidistribution, Computing 66 (2001), 27–39

[2] T. Dupont and Y. Liu, Symmetric error estimates for moving mesh Galerkin methods for advection-diffusion equations, SIAM J. Numer. Anal. 40 (2002), 914–927

[3] W.Z. Huang and R.D. Russell, Moving mesh strategy based on a gradient flow equation for two-dimensional problems, SIAM J. Sci. Comput. 20 (1999), 998–1015

[4] R. Li, T. Tang and P. Zhang, A moving mesh finite element algorithm for singular problems in two and three space dimensions, J. Comput. Phys. 177 (2002), 365-393

[5] B. Argueyrolles, D. Passerel and D. Maligne, Influence of the nozzle Geometry on Diesel Injector Internal Flow: a Computational Approach, THIESEL 2004 Conference on Thermo-and fluid dynamics processes in Diesel Engines, Valencia, Spain, (2004).

[6] M. Gavaises, D. Papoulias, E. Gianadakkis, A. Andriotis, N. Mitroglou and A. Theodorakakos, Comparison of cavitation formation and development in Diesel VCO nozzles with cylindrical and converging tapered holes, THIESEL 2008 Conference on Thermo-and fluid dynamics processes in Diesel Engines, Valencia, Spain, (2008).

[7] Z.V.A. Warsi, Conservation form of the Navier-Stokes equations in general nonsteady coordinates, AIAA Journal, 19 (1981), 240-242.

[8] I. Demirdzic and M Peric, Space conservation law in finite volume calculations of fluid flow, Int. J. Numer. Methods in Fluids 8 (1988), 1037-1050.

[9] STAR-CD Methodology, version 4.06, CD adapco, 2008

[10] L. Rayleigh. On the pressure developed in a liquid during the collapse of a spherical cavity. Phil. Mag. 34 (1917), 94-98.

[11] F. Payri, X. Margot, S. Patouna, F. Ravet and M. Funk, A CFD study of the effect of the needle movement on the cavitation pattern of Diesel Injectors, Conference on Society of Automotive Engineers International.

[12] Xandra Margot, Sergio Hoyas , Pablo Fajardo and Stavroula Patouna, Moving mesh strategy solving a Diesel injector needle movement problem. Submitted to Mathematical and Computer Modelling, 2010.

Multi-objective Particle Swarm Optimization Applied to Water Distribution Systems Design: an Approach with Human Interaction

I. Montalvo[a], J. Izquierdo[1a], S. Schwarze[b] , R. Pérez-García[a]

[a]Instituto de Matemática Multidisciplinar, Universidad Politécnica de Valencia
Camino de Vera s/n, 46022, Valencia, Spain

[b]*Institute of* Information *Systems, University of Hamburg,*
Von-Melle-Park 5, D-20146, Hamburg, GERMANY

1 Introduction

Particle swarm optimization (PSO) has been used successfully to solve different problems in the water industry [1-3] as well as other fields [4-7]. Nevertheless, the integration of various disciplines, differing points of view, and the participation of various experts during decision-making processes increasingly requires a more multi-objective approach to the algorithm for tackling complex optimization problems. The criteria used in water distribution network (WDN) design are numerous: economical, technical, political, etc.

Regardless of the problem to be solved, some changes must be made when using PSO as a multi-objective alternative in decision-making.

Firstly, leadership in the swarm should be redefined. The big question is how can values resulting from a multi-objective function be compared and how can the leader in the swarm be chosen. To work with a multi-objective approach, a representation of the Pareto front is used to discriminate between solutions.

A solution *A* is said to dominate another solution *B* when *A* is better than *B* in at least one objective, and not worse in the others. Two solutions are called indifferent or incomparable if neither dominates the other. The aim of multi-objective optimization is to find the Pareto-optimal set or front (or an approximation sample). The front of non-dominated solutions produced by a Pareto-based multi-objective algorithm can be used in a subsequent phase to select one or more solutions according to different criteria.

The most natural option is to select as leader the closest particle to the so-called utopia point in the search space. This utopia point is defined as the point in the search space where components give the best values found for every single objective. In this problem, the utopia point is an unknown dynamical point since the best value for every objective is something unknown at the very beginning (and even during the whole

[1] *e-mail:* imontalvo@gmmf.upv.es, jizquier@gmmf.upv.es, schwarze@econ.uni-hamburg.de, rperez@gmmf.upv.es

process). Accordingly, we will use an approximation of this utopia point, which we call a singular point, and which is updated during the evolution of the algorithm.

Arguably, the most interesting solutions are located close to that singular point and not too far from the extreme ends of the Pareto front. Consequently, instead of seeking a complete and detailed Pareto front, we are more interested in more precise details around the singular point. Nevertheless, situations of non-symmetric Pareto front with respect to the singular point can occur; and, consequently, poorly detailed sections on the Pareto front may appear. It seems plausible that due to problem complexity this is the most frequent case in multi-objective water distribution system design.

In this work, the user can specify additional points where the algorithm should focus the search, and to specify how much detail a region should contain. This is achieved in real time during the execution of the algorithm. Human proposed solutions could even become leaders of the swarm. At this point, human behaviour begins to have a proactive role during the evolution of the algorithm.

2 The problem of optimal water distribution system design

WDN design is a wide problem in hydraulic engineering that involves the addition of new elements in a system; the rehabilitation or replacement of existing elements; decision-making on operation; reliability and protection of the system; among other actions. In most cases, economic reasons condition the difference between a solution and a better solution. In any case, all technical and non-technical constraints of the problem must be fulfilled.

The key point in water distribution system design is to properly assess the value of a solution. As different actors (politicians, hydraulic engineers, environmental engineers, economists, etc.) are normally involved in water projects, it is frequent that more than one objective will arise and they may even compete among themselves.

In this paper, two objectives are used to evaluate the algorithm in a benchmarking problem case. The real-world problem considered later will additionally use a third objective. The inclusion of additional objectives is straightforward.

The first objective is the minimization of the initial investment cost (for the pipes required for the new system),

$$F_1(D) = \sum_{i=1}^{L} c_i l_i . \tag{1}$$

L is the total number of pipes; $D = (D_1, \ldots, D_L)^t$ is the vector of the pipe diameters; the cost per meter, depending on the diameter of pipe i, D_i, is given by c_i; l_i represents the length of pipe i. Note that D_i is chosen from a discrete set of commercially available diameters and c_i is a non-linear function of diameter.

The second objective is the minimization of the lack of pressure at every consumption node. This objective is also a function of the selected pipe diameters (through the hydraulic model). The lack of pressure is calculated as the difference between the required pressure and the existent pressure in a node. When the existent pressure is greater than the required pressure the lack of pressure at the node is zero:

$$F_2(D) = \sum_{j=1}^{N} H(p_{\min} - p_i) \cdot (p_{\min} - p_i). \quad (2)$$

The function $H(p_{\min} - p_i)$ is the Heaviside function. Hydraulic modelling software for water distribution system analysis is used to evaluate the actual pressure at consumption nodes for a specific solution. The integration of such software to run different analyses or simulations for potential solutions of the problem is performed during the optimization process that is developed within the evolutionary algorithms [1-3] – such as the algorithm presented in this paper.

The benchmarking problem used in this paper for evaluating the algorithm is known as *The Hanoi Network* [8]. Figure 1 represents a scheme of the network.

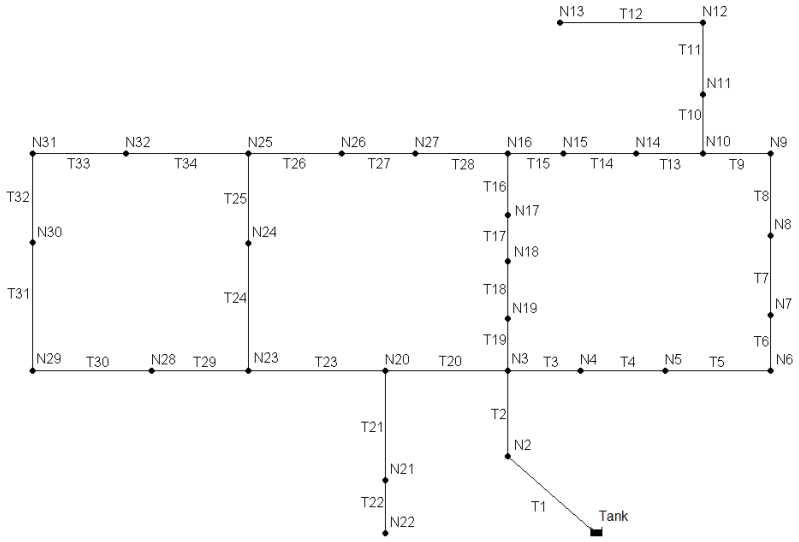


Figure 1. The Hanoi network

The multi-objective function for the Hanoi network will have, as a consequence, two dimensions:

$$MF_1(D) = (F_1(D), F_2(D)).$$

In addition, a real world network is also considered in this paper, Figure 2. This corresponds to one water distribution sector in Lima, Peru [9].

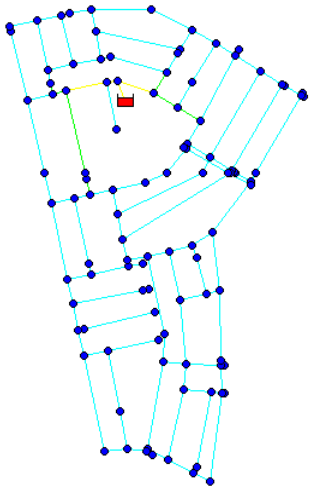


Figure 2. Hydraulic sector in a network in Lima

In this case a third objective is included: a reliability assessment of the network. There are various approaches to assess the reliability of a water distribution system [10-12]. This paper uses a formulation proposed in [13] that indirectly assesses reliability from an economic point of view. This formulation considers the costs of the water not delivered due to disruptions in the system, and associated repair costs. These costs have a close relationship with the network model, and which again depend on the pipe diameters:

$$F_3(D) = \sum_{k=1}^{L} w_k \cdot l_k \cdot D_k^{-u}. \quad (3)$$

w_k is a coefficient associated with each pipe that accounts for the number of expected failures per year of one pipe, its daily average cost of repair work, and the average cost of the water supplied to affected consumers due to the loss of water.

The objective in this case will have three dimensions:

$$MF_2(D) = (F_1(D), F_2(D), F_3(D)).$$

Lower values of the third objective are desirable because they correspond to more reliable water distribution systems. This paper is not intended to go deeper into the design of real cases; but aims to present a general algorithm that could be used in water distribution system multi-objective optimization tasks. Nevertheless, to incorporate any additional objective or reliability assessment to deal with real cases does not require major effort. In the following section, it can be seen how the algorithm works.

3 Description of the algorithm

PSO is an evolutionary computation technique that was first developed by Kennedy and Eberhart [14]. In the algorithm, a swarm (population) consists of a set of an integer number, M, of particles, X_i, moving within the solution space, $S \subset \mathrm{R}^d$, each representing a potential solution of the problem:

$$\text{Find } \underset{X \in S}{best}\, F(X), \text{ subject to appropriate constraints,} \quad (4)$$

where F is the fitness function associated with the problem. The optimization is aimed at the minimization or maximization of the value of F, depending on the nature of the problem.

In addition to the position vector for every particle i,

$$X_i = (x_{i1}, x_{i2}, ..., x_{id}),$$

an individual velocity vector

$$V_i = (v_{i1}, v_{i2}, ..., v_{id}),$$

and the position at which the best fitness was encountered by the particle

$X_{i\_\text{pbest}} = (x_{i1\_\text{pbest}}, x_{i2\_\text{pbest}}, \ldots, x_{id\_\text{pbest}})$,

are computed and stored. Also, a record of the position of the best particle of the swarm,

$X_{\text{global\_best}} = (x_{1\text{global\_best}}, x_{2\text{global\_best}}, \ldots, x_{d\text{global\_best}})$,

is maintained. In each generation, the velocity of each particle is updated based on its trajectory, its best-encountered position, and the best position of the swarm:

$$V_i = \omega \times V_i + c_1 \times rand(\,) \times (X_{i-pbest} - X_i) + c_2 \times rand(\,) \times (X_{global\_best} - X_i). \quad (5)$$

In equation (5), □ is a factor of inertia suggested by Shi and Eberhart [15] that controls the impact of the velocity history into the new velocity. Here, we use

$$\omega = 0.5 + \frac{1}{2(\ln(k)+1)}, \quad (6)$$

where k is the iteration number [4].

Parameters c_1 and c_2 in (5) are called acceleration parameters. The term *rand*( ), represents a function that creates random numbers between 0 and 1.

On each dimension, particle velocities are restricted to minimum and maximum velocities:

$$V_{\min} \le V \le V_{\max}, \quad (7)$$

which are user-defined parameters, to control excessive roaming of particles outside the search space. Usually, it is assumed that $V_{\min} = -V_{\max}$.

The position of each particle is updated at every generation:

$$X_i = X_i + V_i. \quad (8)$$

In the approach presented in this paper the updating of a particle's best position is performed when a new position dominates the previous best position of the particle. As a result, particles will continue moving in the direction of the leader – but now they will try to do so over the Pareto front. Sometimes neither position dominates the other; in this case, the particle will decide that the best position is the one closest to the position of the leader. When the best previous position of a particle is located on the Pareto front formed by the whole swarm, and its new position is also on the Pareto front, it then produces a clone, provided a higher density in the Pareto front is possible.

Two more observations must be made before providing a description of the proposed algorithm dealing with human interaction.

Firstly, to tackle discrete variables, this algorithm takes the integer parts of the flying velocity vector's discrete components into account; hence, the new discrete component velocities V_i are integer. Accordingly, velocity for discrete variables is calculated by using the expression:

$$V_i = fix(\omega \times V_i + c_1 \times rand(\,) \times (X_{i-pbest} - X_i) + c_2 \times rand(\,) \times (X_{global\_best} - X_i)), \quad (9)$$

where *fix*(·) implies that we only take the integer part of the result.

Secondly, in [3], PSO was endowed with a re-generation-on-collision formulation which further improves the performance of discrete PSO. The random regeneration of

the many birds that tended to collide with the best birds was shown to avoid premature convergence, as it prevented clone populations from dominating the search. The inclusion of this procedure into the discrete PSO algorithm produces greatly increased diversity, improved convergence characteristics, and yields higher-quality final solutions.

Humans can interact with the algorithm by adding new singular points, but with fixed values. These values will be specified by the user during runtime. Once a new singular point is added, a new swarm is created with the same characteristics of the swarm created first. Swarms will run in parallel but they share (and can modify) the information related to the Pareto set. Particles from any swarm can be added to the Pareto set. If the user changes the fixed values for a singular point then the corresponding swarm selects a new leader considering the new location of the singular point.

In addition, a user can devise a solution and ask the algorithm in real time to analyze and evaluate it. Eventually, that solution can be incorporated to the Pareto front or lead the behaviour of a group of particles. User solutions will always be evaluated in the first swarm created. If a particle is being evaluated then the user request waits until the evaluation of the particle is finished. If a solution proposed by the user is being evaluated then any particle belonging to the first swarm should wait for evaluation. Once any solution is evaluated, the algorithm checks if it could be incorporated in the Pareto front. Synchronization is made among all the swarms in order to open access for managing the Pareto front.

In this study, an initial population size of $M = 100$ particles has been used. Also, among the various termination conditions that may be stated, if there is no improvement after 20 iterations, the process is stopped.

The performance of the approach herein introduced can be observed from the results reported in the next section.

4 Results

Optimal results for the Hanoi network are presented first. At the beginning only one swarm was working using a singular point formed by the best value for each objective. This singular point behaves dynamically because the best values for every objective are not known a priori. Firstly, the best value for each objective is selected from the random population of particles generated initially. Then, whenever a better value for any objective is found, the corresponding component of the singular point is updated.

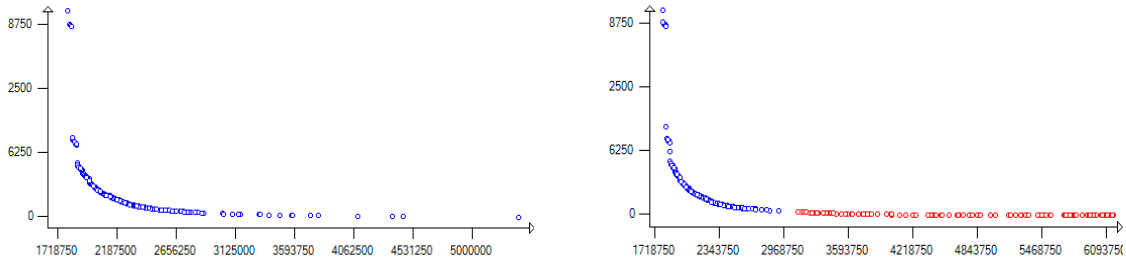


Figure 3. Results for the Hanoi network

Figure 3 (left) shows the approximated Pareto set found by the particles. By observing the graph, it can be noted that the lack of pressure in the network decreases rapidly at first with the increase of the investment cost (vertical asymptotic branch). However, the relation between the investment cost and the lack of pressure in the network is not always the same: new additional reductions of lack of pressure are increasingly costly. As a result, the leader of the swarm will exhibit a tendency to be located near the region where the increase of the investment cost reduces the lack of pressure significantly – or at least substantially. At this point, the minimum value of pressure in the network could be reconsidered. If it is decided to maintain the same value of minimum required pressure – then it is better to send a second swarm to the region where the lack of pressure is zero and the investment cost is as low as possible.

The value of the minimum required pressure was not changed and a second swarm was sent to work in parallel with the first. The approximated Pareto set obtained by both swarms is represented in Figure 3 (right). Solutions from the second swarm increase the density of the horizontal asymptotic branch.

Results for the second network used in this paper, corresponding to a sector of the water distribution system in Lima, are represented in Figure 4. In this case two swarms were also sent to search for solutions. These swarms had the same characteristics as the swarms used for the Hanoi network. Three objectives were considered for this problem instead. Because of the bi-dimensional representation, some solutions seem to be dominated but actually they are not: their values for the third objective are good enough to belong to the Pareto set.

During this research several tests were made exploring the possibility that users could add potential solutions. Nevertheless, for comparison purposes, the results presented in this paper were obtained without the intervention of users adding potential solutions.

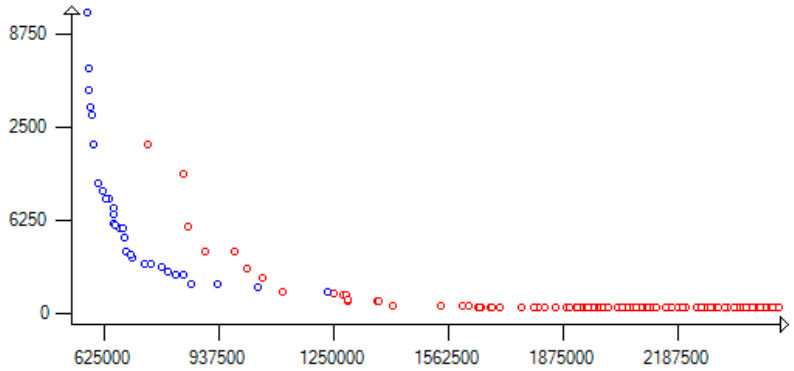


Figure 4. Results for sector in Lima

5 Conclusions

The use of evolutionary algorithms or any meta-heuristic has been widely applied to solve engineering optimization problems. Global optimal solutions cannot be guaranteed but good solutions can be found when applying the right method to the right problem. Nevertheless, these algorithms cannot assume all of the responsibility for

building an engineering solution, especially in cases of multi-objective or multi-criteria optimization problems that require the participation and integration of experts from different fields. Working as a team and profiting from the collective intelligence is the best way to face current challenges. This work represents one step towards the development of algorithms that could be integrated into a human-computer collaborative system. Integrating the search capacity of these algorithms and the ability of specialists to redirect the search towards specific interest points – based on their experience in solving problems – results in a powerful collaborative system for finding solutions to engineering problems. Additionally, the ability of the algorithm to redirect its own search by considering good solutions proposed by users is a great advantage.

Extensions to this work are perfectly possible for successfully solving more complex problems – especially in the field of water. Finally, the results herein presented can be applied to other engineering fields or wherever the search for a solution implies the use of a computer.

Acknowledgments

Thanks to the support of the project IDAWAS, **DPI2009-11591,** of the Dirección General de Investigación del Ministerio de Educación y Ciencia, and the grant MAEC-AECI 0000202066 awarded to the first author by the Ministerio de Asuntos Exteriores y Cooperación of Spain. The use of English in this paper was revised by John Rawlins; and the revision was funded by the Universidad Politécnica de Valencia, Spain.

References

[1] J. Izquierdo, I. Montalvo, R. Pérez. V.S. Fuertes, Design optimization of wastewater collection networks by PSO, Computers & Mathematics with Applications 56(3) (2008) 777-784.

[2] J. Izquierdo, I. Montalvo, R. Pérez, M. Tavera, Optimization in water systems: a PSO approach, in Proc. Business and Industry Symposium (BIS), Ottawa, Canada, 2008.

[3] I., Montalvo, J. Izquierdo, R. Pérez, P.L. Iglesias, A diversity-enriched variant of discrete PSO applied to the design of Water Distribution Networks, Engineering Optimization 40(7) (2008) 655-668.

[4] Y.X. Jin, H.Z. Cheng, J.Y. Yan, L. Zhang, New discrete method for particle swarm optimization and its application in transmission network expansion planning, Electric Power Systems Research 77(3-4) (2007) 227-233.

[5] X. H. Shi, Y.C. Liang, H.P. Lee, C.Lu, Q.X. Wang, Particle swarm optimization-based algorithms for TSP and generalized TSP, Information Processing Letters 103(5) (2007) 169-176.

[6] J. Izquierdo, R. Minciardi, I. Montalvo, M. Robba, M. Tavera, Particle swarm optimization for the biomass supply chain strategic planning, in Proc. 4th Biennial Meeting International Conference on Environmental Modelling and Software (iEMSs), Barcelona, España, 2008.

[7] J. Izquierdo, I. Montalvo, R. Pérez, V.S. Fuertes, Forecasting pedestrian evacuation times by using swarm intelligence, Physica A: Statistical Mechanics and its Applications, 388(7) (2009) 1213-1220.

[8] O. Fujiwara, D.B. Khang, A two-phase decomposition method for optimal design of looped distribution networks, Water Resour. Res., 26(4) (1990) 539-549.

[9] I. Montalvo, Diseño óptimo de sistemas de distribución de agua mediante Particle Swarm Optimization. Departamento de Ingeniería Hidráulica. Universidad Politécnica de Valencia, (2008).
[10] L.W. Mays, Review of reliability analysis of water distribution systems. Stochastic Hydraulic, Rotterdam, The Netherlands (1996).
[11] A. Ostfeld, U. Shamir, Design of Optimal Reliable Multiquality Water-Supply Systems, Journal of Water Resources Planning and Management 122(5) (1996) 322-333.
[12] T.T. Tanyimboh, Y. Setiadi, Joint layout, pipe size and hydraulic reliability optimization of water distribution systems, Engineering Optimization 40(8) (2008) 729-747.
[13] J.B. Martínez, Quantifying the economy of water supply looped networks, Journal of Hydraulic Engineering-ASCE 133(1) (2007) 88-97.
[14] J. Kennedy, R.C. Eberhart, Particle swarm optimization, in Proc. IEEE International Conference on Neural Networks, Perth, Australia, IEEE Service Center, Piscataway, NJ, 1995.
[15] T. Shi, R.C. Eberhart, A modified particle swarm optimizer, in Proc. IEEE international conference on evolutionary computation, Piscataway, NJ, 69-73, 1998.
[16] Y. Shi, R.C. Eberhart, Empirical study of particle swarm optimization, in Proc. IEEE Congress on Evolutionary Computation, Washington, DC, USA 1999.

Cost analysis of a vaccination strategy for respiratory syncytial virus (RSV) in a network model.

J.-A. Moraño*, L. Acedo†,
and J.Díez-Domingo§

*, †, Instituto de Matemática Multidisciplinar,

Universidad Politécnica de Valencia,

Edificio 8G, Piso 2, 46022 Valencia, España.

§ Centro Superior de Investigación en Salud Pública,

CSISP, Valencia, España.

December 10, 2009

The spread of epidemic diseases has been traditionally simulated by means of systems of differential equations [1, 2, 3]. Typically in these models we consider the fraction infected (I), susceptible (S) and recovered (R) individuals and propose a compartmental model for the transitions between these states. The resulting SIR model has been widely studied [3, 4] but, albeit it is a good approximation in some cases, it is clear it cannot be the final word in the epidemiology of any real disease.

However, the continuous approach cannot, by its own nature, distinguish among individuals and, consequently, the effect of age, sex, previous illnesses and any other parameters influencing the propagation of the epidemic under study are difficult to implement. In the differential equations approach we only consider continuous functions, $I(t)$, $S(t)$ and $R(t)$ and insurmountable difficulties are faced out when the interest is upon the evolution of single individuals instead of an average over the full population. The vaccination

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

†e-mail: acedo@imm.upv.es

programs are an example of a situation in which the network approach shows its advantages. In the network we can easily monitor the age of any individual and implement vaccination doses at a given age for children or catch-up policies. In continuous models we use a vaccination probability but, this way, we cannot avoid counting the same inviduals two or more times and to obtain reliable costs of the diseases is difficult. Moreover, taking into account the local and discrete character of epidemic spread also allows to include variable susceptibility or recovery rates of individuals, mobility and long-range infections. For instance, the debate about targeted or mass vaccination in the control of smallpox has also been addressed within the context of network models [5, 6].

In this work, we consider a complete network model for the propagation of the respiratory syncytial virus seasonal epidemic. This pandemic is the direct cause of around $1,300$ pediatric hospitalizations in the Spanish region of Valencia and $15,000$-$20,000$ visits to primary care every year [7]. The cost to the Valencian Health System is estimated in 3.5 million euros per year. We have retrieved hospitalization data for children less than one year old in the region of Valencia as a consequence of bronchiolitis or pneumonia developed by RSV infection. Unfortunately, prevalence data is still not available but we will be able to compare with the models by an adequate scaling of the predicted infected children. On the other hand, most of the hospitalizations correspond to children less than one year old and, consequently, we have to single out this age group both in the continuous differential equation model and the network model.

Modelling of the RSV epidemic has also been carried out within the framework of the standard SIR differential model. For example, Weber et al. [8] developed SIRS (susceptible-infected-recovered-susceptible) mathematical model with four possible reinfections and applied it to explain the data curves for Gambia, Singapore, Florida and Finland. They found that the seasonal component depends on the local climate of the country under study and even the period of the epidemic can be different. This paper leads us to conclude that the propagation of RSV is still not understood properly because the sharp peaks at the outbreaks are not adequately fitted by continuous models. Similar approach has also been adopted by White et al. [9] in a nested RSV model.

On the other hand, continuous models are more reliable concerning the application of optimizing techniques as shown below. Consequently, the objective in this work has been to propose a two-age group generalization of We-

ber's model [8] and use it to fit the seasonal parameters. Seasonality of the infection probability is modelled by a single harmonic $\beta(t) = b_0 + b_1 \cos(2\pi t + \varphi)$ where b_0, b_1 and φ are parameters obtained by fitting the data for the region under consideration. This kind of functions has been considered not only for RSV [8] but also for other seasonal epidemics such as measles [10]. The transition rates between the compartments R and S, I and R have been obtained from literature [11] and [8] respectively. In order to fit the population parameters of the continuous model we have resorted to the Valencian Institute for Statistics [12]. Processing this database we find that the average population in the Valencian region is $4,252,386$ inhabitants during the period of interest from January 2001 to December 2004 where data was harvested. We use numerical integration to fit the real data and the fitting is carried out by means of the downhill simplex method due to Nelder and Mead [13].

Once the seasonal parameters are fitted we apply it to the evolution of a complete network model, with a Forster-Mckendrick population model, for the Valencian region. The network includes a node for every person in the region of Valencia.

Moreover a PIV-vectored vaccine is already under development and clinical studies have been carried out since past year [14]. This vaccine could be available in the near future and, consequently, it is an urgent task to anticipate vaccination strategies. To the best of our knowledge, vaccination strategies for RSV have not been studied and the imminence of the application of PIV-vectored vaccines demands such an study. A previous work on the cost-effectiveness of immunoprophylaxis with palivizumab has been recently reviewed [15].

These jobs have allowed to develop a cost analysis for a vaccination strategy for RSV. The main purpose of vaccination is to avoid the first RSV infection which is the most acute, affects children younger than a year old, and sometimes, the overreaction of the child immune system leads to very severe situations that require hospitalization. Newborns are too young to be vaccinated and the emerging strategy to be considered by the doctors, is the vaccination of non-infected children in three doses: at 2 months, 4 months and 1 year old.

Differents fractions $85\%, 90\%$ and 95% of non-infected are vaccinated. In this fraction we consider those children that have been vaccinated the three times and, therefore, the vaccine protection is complete. Then, taking the next 5 years, costs are taken as the mean cost of these 5 years and the total cost of RSV healthcare is calculated taking into account the hospitalization

cost (6.28 hospitalization days for every acutely infected child [7] and 500 euro per day and child hospitalized), vaccination cost (100 euros per dose with three doses programmed during the first year of life, at 2, 4 months and a year old) and the parent work loss with 2,3,4 lose days per each case of infected children with milder symptoms (the labor cost in Spain is 75.21 euro per day [16]).

The results for the global cost for a vaccination of the 85%, 90% and 95% of the non-infected children, for $d = 2, 3$ or 4 days of parent work loss in the case of children that do not develop sufficiently acute symptoms to become hospitalized, are the following

Estimated cost per year (in euros) without vaccination or vaccinating several fractions of non-infected at 2, 4 and 12 months			
Parent work lose days	2	3	4
Without vaccination	11,761,196	15,585,474	19,409,752
Vaccinating 85%	14,770,010	16,044,020	17,318,030
Vaccinating 90%	14,971,806	16,105,021	17,238,235
Vaccinating 95%	15,155,538	16,137,625	17,119,712

A reduction of more than 2 million euros of total cost is predicted for an estimation of 4 days of parent work loss on average for infected children. In the case of 2 or 3 days of parent work loss, the increasing of the total cost is around $3,400,000$ and $550,000$ euros, respectively, but however, the hospitalization and parent loss work costs decrease dramatically at the expense of vaccination cost. These reductions avoid the saturation in the hospital casualty departments. Moreover, we have not taken into account the long-term effects of RSV infections. In particular, there is an agreement among pediatrics about a connection among RSV at early ages and asthma episodes of children and adolescents. This has been confirmed by recent studies in mice [17]. Therefore, even of the assumption that parents only lose two or three working days for caring children which develop mild symptoms of RSV a positive balance for the implementation of the vaccine is obtained.

References

[1] W. O. Kermack, A. G. McKendrick, Contributions to the mathematical theory of epidemics, Part I, Proc. R. Soc. A 115 (1927) 700.

[2] L. Edelstein-Keshet, Mathematical Models in Biology, Random House, New York, 1988.

[3] J. D. Murray, Mathematical Biology, Springer-Verlag, Heidelberg, 1993.

[4] H. W. Hethcote, The mathematics of infectious diseases, SIAM Review 42-4 (2000) 599.

[5] M.E. Halloran, I.M. Longini Jr., A. Nizam, et al., Containing Bioterrorist Smallpox, Science, 298 (2002) 1428.

[6] J. Koopman, Controlling smallpox, Science, 298 (2002) 1342.

[7] J. Díez-Domingo, M. Ridao-López, I. Úbeda-Sansano, et al., Incidencia y costes de la hospitalización por broquiolitis de las infecciones por virus respiratorio sincitial en la Comunidad Valenciana. Años 2001 y 2002, Anales de Pediatría 65-4 (2006) 325.

[8] A. Weber, M. Weber, P. Milligan, Modeling epidemics caused by respiratory syncytial virus (RSV), Mathematical Biosciences 172 (2001) 95.

[9] L. J. White, J. N. Mandl, M. G. M. Gomes, et al., Understanding the transmission dynamics of respiratory syncytial virus using multiple time series and nested models, Mathematical Biosciences 209-1 (2007) 222.

[10] B. Grenfell, B. Bolker, A. Kleczkowski, Seasonality, demography and the dynamics of measles in developed countries, in: D. Mollison (Ed.), Epidemic Models – Their Structure and Relation to Data, Cambridge University, 1995, pp. 248-268.

[11] C. B. Hall, Respiratory syncytial virus and human metapneumovirus, in: R. D. Feigin, J. D. Cherry, G. J. Demmler, S. L. Kaplan (Eds.)., Textbook of Pediatric Infectious Diseases, 5th Edition, Saunders, Philadelphia, PA, 2004, pp. 2315-2341.

[12] Instituto Valenciano de Estadística, [on-line]. Available from http://www.ive.es.

[13] W.H. Press, B.P. Flannery, S.A. Teukolsky, et al., Numerical Recipes: The Art of Scientic Computing, Cambridge Univ. Press, 1986.

[14] R. S. Tang, R. R. Spaete, M. W. Thompson, et al., Development of a PIV-vectored RSV vaccine: Preclinical evaluation of safety, toxicity, and enhanced disease and initial clinical testing in healthy adults, Vaccine 2008: 26: 6373-6382.

[15] D. Wang, C. Cummins, S. Bayliss, et al., Immunoprophylaxis against respiratory syncytial virus (RSV) with palivizumab in children: a systematic review and economic evaluation, Health Technology Assessment 2008: 12: 36, Available from: http://www.hta.ac.uk.

[16] Encuesta Trimestral de Coste Laboral, Instituto Nacional de Empleo, Spain [on-line]. Available from http://www.ine.es [Accessed April 21, 2009]

[17] H. S. Jafri, S. Chavez-Bueno, A. Mejías, et al., Respiratory syncytial virus lower respiratory tract infection induces acute pneumonia, cytokine response, airway obstruction and chronic inflammatory infiltrates associated with long-term airway hyperresponsiveness in a murine model, Journal of Infectious Diseases 2004: 189: 1856-65.

Obstacle detection in object tracking based on fuzzy controllers.

E. Parrilla *, J. Riera, J.-R. Torregrosa

Instituto de Matemática Multidisciplinar.

Universidad Politécnica de Valencia.

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

December 10, 2009

1 Introduction

In previous works [1], we have studied a fast and robust object tracking system based on optical flow. Optical flow is an approximation to the 2-d motion field of an image sequence, that is a projection of the 3-d velocities of surface points onto the imaging surface [2]. To solve the occlusion problem, we have developed a combined tracking system based on optical flow and adaptive filters. In this article, we use the recursive least squares (RLS) filter [3]. The critical point of the system is the coupling between optical flow and predictive algorithms. This coupling is governed by parameters such as tolerance, the absolute value of the difference between the value of velocity calculated by the optical flow and the estimated value. In this paper, we propose the use of a fuzzy control system to solve this coupling problem between the different velocities. This technique will provide great robustness to the tracking algorithm.

In order to predict the velocities of the objects, we follow the next steps:

*Corresponding author e-mail: edparber@fis.upv.es, Tel.:+34-963877000; Fax: +34-963879009.

1. We calculate velocities for N_{in} frames of each sequence by using Lucas and Kanade algorithm and we use this samples to inicialize the filter coefficients.

2. For the N-th frame, we calculate the velocity v_N in the following way:

   a. We calculate v_N^{of} by using optical flow.

   b. We estimate v_N^{af} by using an adaptive filter.

   c. If $|v_N^{of} - v_N^{af}| < tol_N$, then $v_N = v_N^{of}$. Else $v_N = v_N^{af}$

In a previous work [4], we have demonstrated that an optimum value of tolerance is

$$tol_N = k\,|v_{N-1}|\ . \tag{1}$$

Tolerance values are a critical factor for the correct operation of the proposed algorithms. A very small value of the tolerance will cause the methods to select predicted velocities when there is no obstacle. On the other hand, a large value of the tolerance will make that the methods can not detect any occlusion. For this reason, the choice of parameter k is very critical. In this paper, we propose to use a variable value for parameter k controlled by a fuzzy system, instead of selecting its value a priori.

2 Fuzzy controller

A fuzzy control system is a control system based on fuzzy logic [5]. Fuzzy control provides a formal methodology for representing, manipulating and implementing a human's heuristic knowledge about how to control a system [6]. The fuzzy controller block diagram is given in figure 1, where we can see a fuzzy controller embedded in a closed-loop control system.

To design a fuzzy controller, we have to specify the different blocks of the system:

- Input variables LHS
- Output variables RHS
- Input fuzzification method

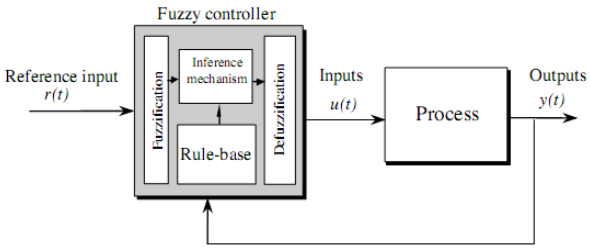


Figure 1: Fuzzy controller architecture

- Fuzzy rules
- Inference engine parameters
- Defuzzification method

3 Obstacle detection

In our system, we want the value of *tol* to be small when there is an obstacle and large in any other case. The value of *tol* can be controlled by the variable k. In this way, we have designed a fuzzy system to control the value of k.

The system has two inputs, $\Delta\varepsilon$ and qV, and one output, k.

The variable $\Delta\varepsilon$ is the increment error (EI). If we call error ε to the absolute value of the difference between the velocity calculated by optical flow and the velocity estimated by the adaptive filter, we have that:

$$\varepsilon_n = \left| v_n^{of} - v_n^{af} \right| , \tag{2}$$

$$\Delta\varepsilon_n = \varepsilon_n - \varepsilon_{n-1} . \tag{3}$$

The variable qV is a binary variable that indicates if we are using the velocities calculated by optical flow in the previous frames ($qV = 1$) or the ones estimated by the adaptive filter ($qV = 0$).

In figure 2, we can see the set of membership functions used to the fuzzification of the input $\Delta\varepsilon$. Fuzzy values for this input can be negative big (NB), medium (M) or positive big (PB).

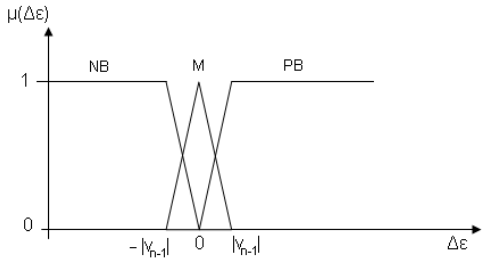


Figure 2: Membership functions for the input $\Delta\varepsilon$

In this system, the value of variable k must be high (H), to select the velocity calculated by optical flow, when there is no occlusion (EI=M, qV=1) and when the target is in the output of the obstacle (EI=NB, qV=0). In all the other cases, the value of k will be low (L) to select the velocity estimated by the adaptive filter.

In this way, the system fuzzy rules are the following:

- IF EI is NB and qV=0 THEN k is H
- IF EI is M and qV=0 THEN k is L
- IF EI is PB and qV=0 THEN k is L
- IF EI is NB and qV=1 THEN k is L
- IF EI is M and qV=1 THEN k is H
- IF EI is PB and qV=1 THEN k is L

The inference engine calculates the global output $\mu_{out}(k)$ by using these conditions and the set of output membership functions that we can observe in figure 3. In this way, the value of the output variable k can oscillate between 0.5 and 1.

Finally, the defuzzificator calculates from the global output $\mu_{out}(k)$ the real value for the variable k by using the center of sums technique, that provides fast and satisfactory results.

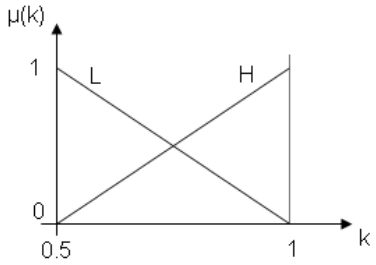


Figure 3: Membership functions for the output k

4 Results

The algorithm exposed in this paper has been tested in different video sequences, studying synthetic and real traffic videos. We have used windows of 7×7 pixels for Lucas and Kanade algorithm, and a value of $N_{in} = 7$ frames, an order of 2 and a forgetting factor of $\lambda = 0.99$ for the adaptive filter.

In figure 4, we can observe differents frames of a synthetic video sequence designed with 3D Studio Max ® that consists in an object that follows a curved trajectory passing under an obstacle.

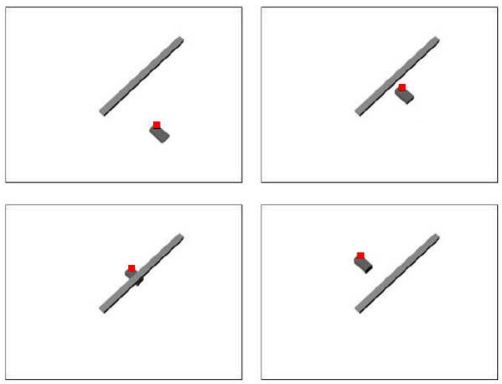


Figure 4: Object tracking in a synthetic sequence

In this example, the occlusion is produced in frames 37-43. As we can observe in figure 5, the value of k decreases in those frames. In this way, the

predicted velocity is selected while the target is hidden by the obstacle and the object is correctly tracked along all the sequence.

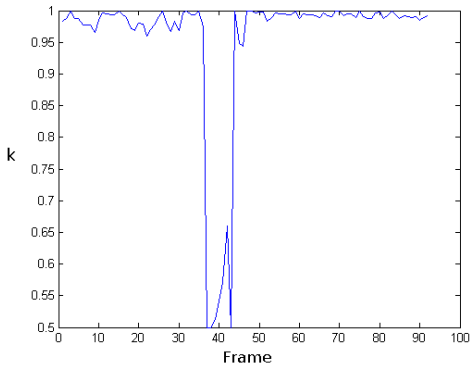


Figure 5: k values in synthetic sequence

The same result is obtained analyzing the example shown in figure 6. In this figure, we can see an urban route where there is a partial vehicle occlusion because of a street light.

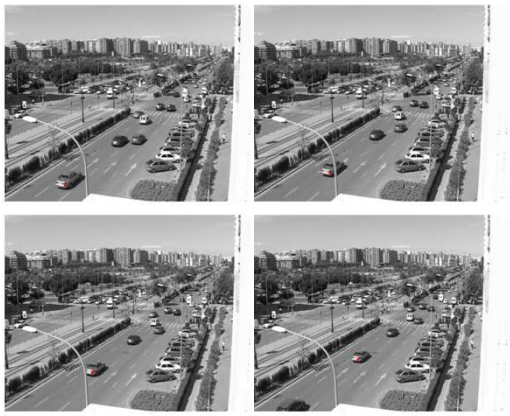


Figure 6: Object tracking in a real sequence

In this case, the occlusion occurs in frames 21-27 and, as we can see in figure 7, k decreases in this position. On the other hand, in this figure, we

can observe that the value of k also decreases in frames 51-54. This change is due to the fact that there is an error in the optical flow calculation. The performance of the system in this point is very interesting since it is able to correct possible errors in the optical flow calculation, besides handling occlusion.

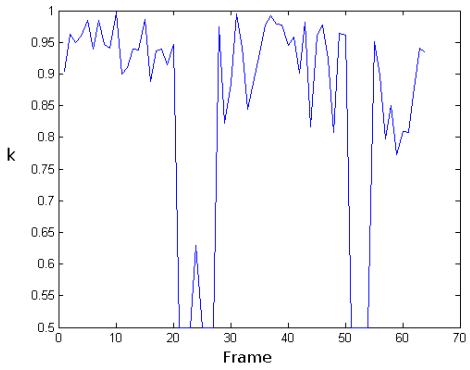


Figure 7: k values in real sequence

5 Conclusions

In this article, we have analyzed the occlusion problem in object tracking in a video sequence. We have proposed the use of a fuzzy controller to combine the velocities calculated by an optical flow algorithm and the ones estimated by an adaptive filter in order to predict the target movement and approximate its trajectory when the target disappears.

We have studied the effect that parameters values produce in the performance of the method, obtaining an optimal system to control the tolerance value. This fuzzy system provides a great robustness and a low computational cost.

Finally, we have shown two examples that verify the efficiency of the algorithms, proving that this technique also can detect and correct errors in the object tracking.

References

[1] E. Parrilla, J. Riera, J.R. Torregrosa, J.L. Hueso, Handling occlusion in object tracking in stereoscopic video sequences, Mathematical and Computer Modelling **50**(5-6) (2009) 823-830.

[2] B.K.P. Horn, B.G. Schunck, Determining optical flow,Artificial Intelligence **17**(1-3) (1981) 185-203.

[3] L. Ljung, System Identification. Theory for the User, Prentice-Hall, Upper Saddle River NJ. (1999).

[4] E. Parrilla, D.Ginestar, J.L. Hueso, J. Riera, J.R. Torregrosa, Handling occlusion in optical flow algorithms for object tracking, Computers and Mathematics with Applications **56**(3) (2008) 733-742.

[5] Lotfi A. Zadeh, Fuzzy sets, Inf. Control **8** (1965) 338-353.

[6] K. M. Passino, S. Yurkovich, Fuzzy control, International Journal of General Systems **29** (2) (2000) 341-342.

Internal Flow Modeling in Diesel Nozzles using Large Eddy Simulation. *

R. Payri[†], B.Tormos [‡], J.Gimeno[§] and G.Bracho[¶]

CMT - Motores Térmicos,

Universidad Politécnica de Valencia,

Edificio 6D, 46022 Valencia, España.

December 10, 2009

1 Introduction

In diesel engines, spray evaporation is the first step in combustion processes and it depends strongly on fuel injection and atomization. Then, it is important to know the flow pattern inside the injector holes, because it affects directly the diesel spray behaviour, determining the combustion. Usually, flow in diesel injection nozzles is characterised by high pressure drops, producing very high velocities and, in spite of the small size of the nozzles, these velocities will produce a semi-turbulent or, and more generally, a turbulent flow regime. Turbulence, as is known, is characterised by chaotic motion and fluctuating flow in both time and space, containing eddies of a wide range of scales from the largest one characterised by the geometry, to the smallest Kolmogrov dissipation scales. In diesel injection the experimental procedure to determine the specific pattern of turbulent motion in internal flow is very complicated (not possible so far), due to the small temporal and dimensional

*This research was funded by TRA2006-13782 from MICINN Spain

[†]rpayri@mot.upv.es

[‡]betormos@mot.upv.es

[§]jaigigar@mot.upv.es

[¶]gbracho@mot.upv.es

scales involved in the process. Therefore, it is necessary to employ accurate theoretical and computational approaches in order to model the fuel flow development, predicting quantities that are indispensable in turbulent processes and not possible to obtain experimentally. Hence, in the present work the main idea is to use the Large Eddy Simulation (LES) computational technique [1], in order to determine if it is capable to reproduce the different turbulent patterns that appear in the internal flow of diesel injectors that are very hard to obtain experimentally, evaluating the potential of the code. Results are compared with the classical numerical RANS method and validated with experimental data. The main code was developed by OpenSource Ltd for OpenFoam, which is a flexible software that allows to incorporate more models and expressions.

2 LES methodology

The equations that are solved in a LES are filtered versions of the governing equations [2, 3, 4]. In this study the filtering operation is applied to the Navier-Stokes equation, for an incompressible flow of a Newtonian fluid. Application of the filtering operation to the continuity and momentum equations yields:

$$\nabla \cdot \bar{u} = 0 \tag{1}$$

$$\frac{\partial \bar{u}}{\partial t} + \nabla \cdot \bar{u}\bar{u} = -\frac{1}{\rho}\nabla \bar{p} + \nu \nabla^2 \bar{u} - \nabla \tau \tag{2}$$

where u is velocity field, t is time, p is pressure, ρ is fuel density, ν is the uniform kinematic viscosity and τ is the stress-like tensor. Equations (1) and (2) govern the evolution of the large (energy-carrying) scales of motion [5, 6]. For the sub-grid scale model the Smagorinsky Model is employed, which is the most widely used for internal flows [4, 7, 8], in where it is supposed that the turbulent eddy viscosity is proportional to the sub-grid scale characteristic l length and to a characteristic SGS velocity:

$$\nu_t = (C_s \Delta)^2 (2|\bar{S}|^2)^{1/2} \tag{3}$$

where C_S is a theoretical value ($0.1 \sim 0.2$) and $\Delta = (\Delta x \Delta y \Delta z)^{1/3}$ is a measure of the local grid length scale, which varies spatially in this study [9, 10]. The presence of a solid wall modifies the turbulence dynamics, inducing inhomogeneity and anisotropy in the flow [4, 5]. The model used for these

Table 1: Physical properties and simulation conditions

Pinlet $[MPa]$	Poutlet $[MPa]$	Dynamic Viscosity $[kg/m.s]$	Density $[kg/m^3]$
120	5	2.68e-3	819

regions is the van Driest damping model, where the turbulent mixing length $l = C_S\Delta$ is modified using:

$$l = C_S\Delta\left[1 - exp\left(-\frac{y^+}{A^+}\right)\right]^{1/2} \tag{4}$$

where y^+ is the distance from the wall in viscous wall units $y^+ = yu_\tau/\nu$, the friction velocity is $u_\tau = \sqrt{\tau/\rho}$ and A^+ is an empirical dimensionless constant.

3 Numerical Technique

The governing equations presented in the previous section are solved using the finite volume CFD code OpenFOAM [11]. The solution procedure employs the PISO algorithm, commonly used in this kind of studies [12]. The calculations are performed in an axi-symmetric nozzle manufactured specially for research purposes. The nozzle has a convergent shape, thus cavitation is avoided [13, 14], consequently the simulation involves just one phase (liquid). The exit hole has a diameter of 112 μm and a length of 1mm. The fluid is winter diesel fuel used in previous experimental works [13]. The physical properties of the fluid and the simulation conditions are summarised in Table 1.

The nominal inlet pressure and backpressure chosen values correspond to a typical operational condition in diesel engines, with available experimental data for further comparisons and validation. The high pressure value chosen issues high velocity flow and could guarantee a turbulent regime, (Reynolds number is $\sim 17.5 \times 10^3$). The first calculation is a two-dimensional Reynolds Average Simulation (RANS), that serves to compare and highlight the advantages of LES. Next, two LES simulations are done, that differs only in respect of the mesh resolution and sector size. The three cases specifications are summarised in Table 2. The reason of apply these geometry simplifications is the reduction of the grid size, saving computational costs. Besides,

Table 2: Cases specifications

Case Name	Transversal Section	Cell Size Resolution	Cell Number
RANS	5°	3μm	32×10^3
LES_90	90°	0.75μm	582×10^3
LES_360	360°	0.12μm	609×10^3

this will allow making comparisons between the different grids setup for LES results. The mesh resolution shown in Table 2 indicates the size of the smallest cells near the wall. They are also depicted in Figure 1.

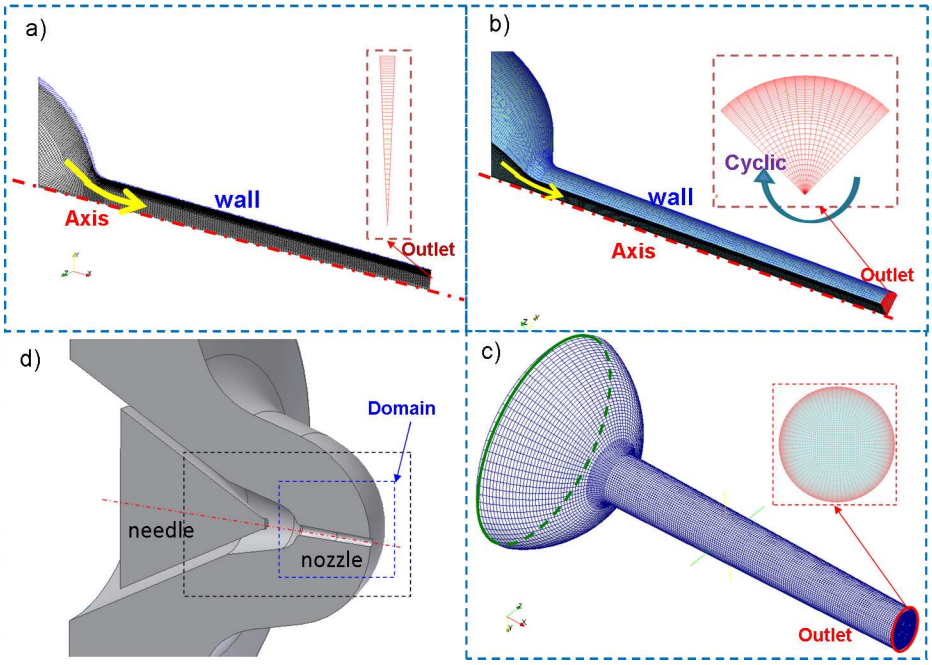


Figure 1: Calculation domain and boundary conditions. a) RANS case, b) LES_90case, c) LES_360case, d)Geometry Scheme

4 Results

Figure 2 compares the RANS and LES results for the streamwise instantaneous velocity U in three different planes along the nozzle ($S1$, $S2$ and $S3$ defined in Figure 3. It can be seen how the core velocity in the central zone

is similar in both approaches, but important differences are observed in the boundary layer near the wall, this might be because the wall model employed in LES reproduce better this zones of the flow. On the other hand, LES_360 case is able to capture more scales than LES_90 case, simulating the unstable (or chaotic) shape of the instantaneous velocity profile. The fluctuating pattern is also confirmed in Figure 3, where U contour is depicted for RANS (upper part) and LES_360 (lower part); there the chaotic motion of the flow and the turbulent structures presence is evident. Concerning LES_90 case, the calculation is not able to capture all the vortices evolution; so, it is not depicted because its behaviour in general is quite similar to RANS case. The no development of eddies could be because of the partial sector configuration, that restricts their growth and maintenance, resulting in almost non existing large structures.

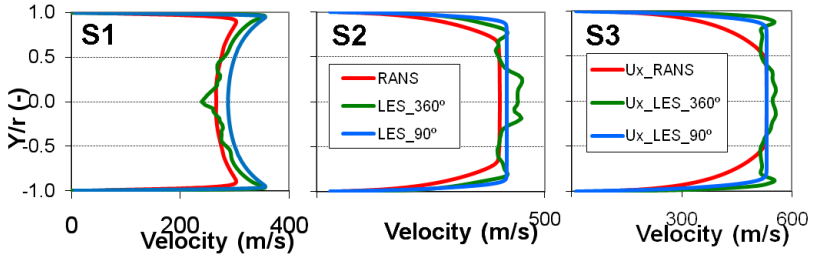


Figure 2: Velocity Profile at three different sections of the hole. Red Line: RANS case, Blue Line: LES_90case, Green Line: LES_360case

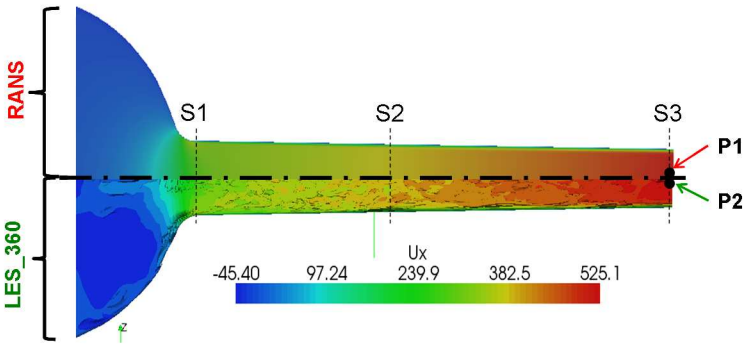


Figure 3: Instantaneous velocity contours. Upper part: RANS case, Lower part: LES_360case. $S1$,$S2$,$S3$ Sections for U profiles

Additionally, one of the objectives of LES is to capture the vortical structures produced by turbulence. Figure 4 depicts the vortex cores visualised via isosurfaces of $Q > 10^{14}$ and coloured by the velocity magnitude, where Q is the second invariant of ∇u [5, 6]. These contours facilitate the analysis and the description of the flow motion inside the hole. It can be seen how the solid boundary (wall) interacts with the fluid flow by retarding the motion tangentially to the surface, due mainly to the viscous shear [15]. For that reason, the structures close to the wall begin to grow longitudinally. The streamwise vortices nearby the solid boundary are ejected from the wall (note that these structures have an inclination respect to the wall), and they are drifted towards the core flow; then, they continue moving forward the flow direction, and finishing at the exit of the hole as elongated streamwise vortices, enlarging the small structures of the hole centre.

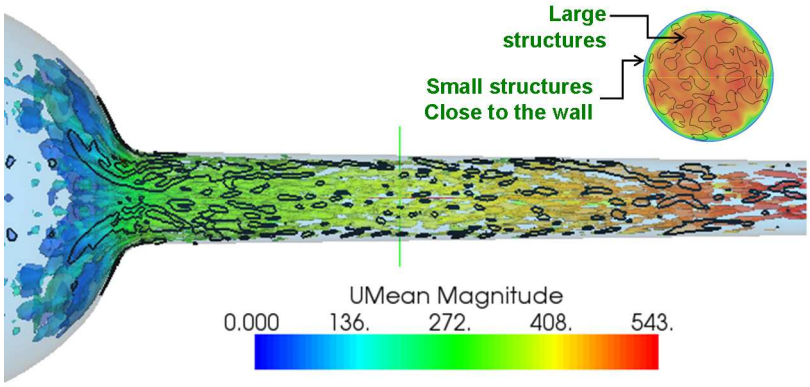


Figure 4: Instantaneous turbulence morphology of the internal flow. Resolved vortical structures visualised via isosurfaces of the second invariant of $\nabla \bar{u}$, and coloured by the streamwise velocity

Besides, results are compared against equivalent experimental data performed in [13]. In Figure 5 it can be seen how the simulation results are close to the experimental ones, being the closest one the full geometry LES case. The parameter used for comparison purposes is the effective velocity $u_{ef} = \dot{M}/\dot{m}$, (where $\dot{M}$ is the flow momentum and $\dot{m}$ is the mass flow), since it does not depend on the outlet diameter, then possible errors due to differences between real nozzle diameter and simulated one are avoided.

Finally, the LES_360 case is compared to DNS simulations performed by Hoyas and Jimenez [17, 16], since those results can be considered an

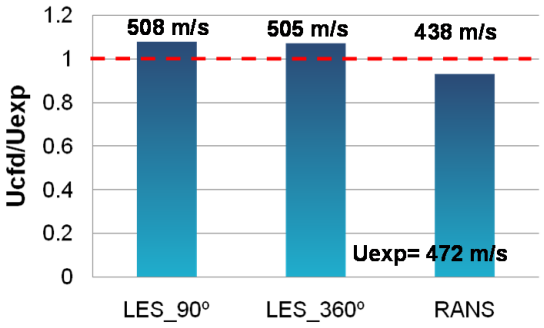


Figure 5: Experimental - Simulated Effective velocity values comparison

exact solution of the Navier-Stokes equations for purposes at the present comparison. Once the statistically steady state is reached, an averaging of the calculated parameters was done over a time interval equal to 10 flow-through times of the domain. The parameters were normalised by means of wall units, as states Equation 5.

$$U^+ = \frac{U}{u_\tau} \tag{5}$$

The dimensionless variables are depicted against the distance from the wall in viscous wall units $y^+ = yu_\tau/\nu$, as shows Figure 6, where the trend of the *law of the Wall* is also represented. It can be seen that results provide a good match, especially in the core region (log region), which is a good indicator, since it is the zone dominated from the large scales structures, and calculated explicitly by LES. Nevertheless, the LES model seems to under-predict the velocity in zones close to the wall (buffer and viscous layer) which is precisely the modelled region of the LES method. This means that more efforts should be done modelling this region, trying to refine more the zone, and/or applying other SGS available models.

5 Conclusions

A study based on numerical simulations of internal flow in diesel injectors, evaluating the skills of the Large Eddy Simulation, has been done. The calculations were made using special diesel injector geometry, applying high

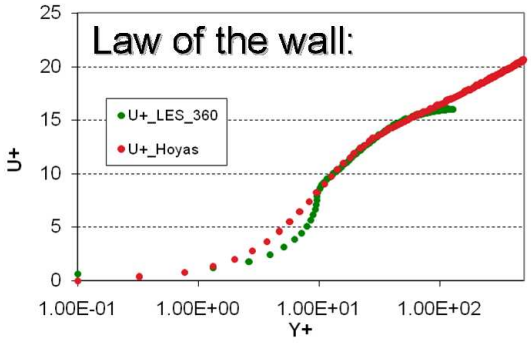


Figure 6: Normalised Mean streamwise velocity vs the log of the distance from the wall in viscous wall units. Red line: DNS Data [17, 16], Green Line:LES_360

injection pressure condition and so ensuring a turbulent regime. The simulations results have been compared to standard numerical RANS method results, and simultaneously validated with experimental values and DNS data. In general, results showed good agreement with the experimental values and DNS data. It was found that LES is capable to reproduce the turbulent structures in diesel nozzles, and as expected is more accurate than RANS, mainly in the boundary layer.

References

[1] M. Rudman, H.M. Blackburn, Large Eddy Simulation of turbulent pipe flow. Second International Conference o CFD in the Minerals and Process Industries CSIRO, Melbourne, Australia, 6-8 December 1999.

[2] A. Moene, Swirling pipe flow with axial strain Experiment and Large Eddy Simulation. Cip-data library Technische Universiteit Eindhoven, 2003.

[3] A. Aldama, Filtering Techniques for Turbulent Flow Simulation, In Springer Lecture Notes in Eng. Volume 56, Springer, Berlin, (1990).

[4] U. Piomelli, Large-eddy simulation: achievements and challenges, Progress in Aerospace Sciences (1999) 335-362.

[5] P. Sagaut, Large Eddy Simulation for incompressible flows, Springer, Berlin, 2001.

[6] M. Lesieur, O. Mtais, P. Comte, Large Eddy Simulation of turbulence, Cambridge University Press, 2005.

[7] J. P. Roop, Numerical comparison of nonlinear subgridscale model via adaptative mesh refinement, Mathematical and Computer Modelling 46 (2007) 1487-1506.

[8] W. J. Layton, Energy Dissipation Bounds for Shear Flows for a Model in Large Eddy Simulation, Mathematical and Computer Modelling 35 (2002) 1445-1451.

[9] A.Yoshizawa, K. Horiuti, A statistically-Derived Subgrid-Scale Kinetic Model for the Large-Eddy simulation of Turbulent Flows. Journal of the Physical Society of Japan. 54 (8) (1985) 2834-2839.

[10] P. Sullivan, J. McWilliams, C. Moeng, A subgrid-scale model for Large-Eddy simulation of planetary boundary layer flows, Boundary-Layer Meteorology 71 (1994) 247-276.

[11] OpenCFD Ltd., FOAM - The complete Guide, http://www.opencfd.co.uk

[12] V. Macian, R. Payri, X. Margot, F.J. Salvador, A CFD analysis of the influence of diesel nozzle geometry on the inception of cavitation, Atomization and Sprays 13 (2003) 579-604.

[13] R. Payri, F. J. Salvador, J. Gimeno, G. Bracho, Understanding diesel injection characteristics in winter conditions, SAE paper 2009-01-0836 (2009).

[14] J.M. Desantes, R. Payri, J.M. Pastor, J. Gimeno, Experimental characterization of internal nozzle flow and diesel spray behavior. Part I: Nonevaporative conditions. Atomization and Sprays 5 (2005) 489-516.

[15] W. Schoppa, F. Hussain, Coherent structure dynamics in near-wall turbulence, Fluid Dynamics Research 26 (2000) 119-139

[16] S. Hoyas, J. Jimenez, Reynolds number effects on the Reynolds-stress budgets in turbulent channels, Phys. Fluids 20 (2008) 101511.

[17] S. Hoyas, J. Jimenez, Scaling of the velocity fluctuations in turbulent channels up to Re=2003, Phys. of Fluids 18 (2006) 011702.

Parameters to choose leaders on Social Network Sites

F. Pedroche*
Institut de Matemàtica Multidisciplinària
Universitat Politècnica de València
Camí de Vera s/n. 46022 València. Spain.
{pedroche@imm.upv.es}

December 10, 2009

1 Introduction

In recent years Social Network Sites (SNSs) have been caught the attention of researchers from different disciplines; see, [1], [2], [3], [4], [5], [6], [7],[8],[9],[10], [11].

In this communication we use the definition of SNS given in [12]. We present a model to classify the users of an SNS based on the PageRank algorithm. In more detail, we use the so-called *personalization vector*. This vector was originally introduced to bias the PageRank to personal preferences of the users [13]. However this makes the PageRank of Google to be query-dependent and therefore computationally impossible [14]. However some authors have used the personalization vector to obtain some variations of the PageRank algorithm. In [15] there is an algorithm that uses topics of the queries to bias the PageRank. In [16] there is another approach computing the PageRank using different personalization vectors. Our main idea consists in using the personalization vector to bias the PageRank to users that are important according some specific features of the SNS, such as number of friends or activity. Here we define the *Competitive groups* (sets of nodes).

*Supported by Spanish DGI grant MTM2007-64477

2 Definitions

Let $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ be the directed graph representing a Social Network Site. Users are represented by the set of nodes $\mathcal{N} = \{1, 2, \ldots, n\}$ and the hyperlinks are represented by the set of directed links $\mathcal{E} \subseteq \mathcal{N} \times \mathcal{N}$. The link (also called arc or edge) represented by the pair (i, j) belongs to the set $\mathcal{E}$ if and only if there exists a hyperlink connecting node i to node j.

A link (i, j) is said to be an *outlink* for node i and an *inlink* for node j. We denote d_i the number of outlinks of a node i; this is called the *outdegree* of node i by some authors.

In an SNS we assume that each node has at least one outlink; i.e., there are no *dangling nodes*. This is a natural assumption: in an SNS each user has, at least, one friend. Therefore we have $d_i \neq 0$ for all $i \in \mathcal{N}$.

We use the PageRank vector [13] as the main classification tool. Since there are no *dangling nodes* we can define the row stochastic matrix $P = (p_{ij}) \in \mathbb{R}^{n \times n}$, in the form

$$p_{ij} = \begin{cases} d_i^{-1} & \text{if } (i,j) \in \mathcal{E} \\ 0 & \text{otherwise} \end{cases} \qquad 1 \leq i, j \leq n.$$

Let $0 < \alpha < 1$ be the so-called damping factor (that we use as $\alpha = 0.85$). Let $\mathbf{e} \in \mathbb{R}^{n \times 1}$ be the vector of all ones and let $\mathbf{v}$ be the personalization (or teleportation) vector, i.e., $\mathbf{v} = (v_i) \in \mathbb{R}^{n \times 1} : v_i > 0$ for all $i \in \mathcal{N}$ and $\mathbf{v}^T\mathbf{e} = 1$. Then the Google matrix is defined as

$$G = \alpha P + (1 - \alpha)\mathbf{e}\mathbf{v}^T,$$

and is an stochastic and primitive (irreducible and aperiodic) matrix [14]. The PageRank vector is defined as the unique left Perron vector of G

$$\pi^T = \pi^T G,$$

with $\pi^T\mathbf{e} = 1$. Denoting $\mathbf{e}_i$ the i-th column of the identity matrix of order n, the PageRank of a node i is $\pi_i = \pi^T\mathbf{e}_i$.

Since we shall use different personalization vectors in the computation of the PageRank vector we shall write $\pi = \pi(\mathbf{v})$ to express this fact. We recall here that given the Jordan canonical form of P the analytical expression of $\pi(\mathbf{v})$ is known; see [17]. We also recall that the analytical expression of the derivative $d\pi/d\mathbf{v}$ is known; see [14] p. 63.

3 Competitivity groups

Definition 1. *Given a directed graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$, let $0 < \epsilon < 1$ and let $\mathbf{v}_i = [v_{ij}] \in \mathbb{R}^{n\times 1} : v_{ii} = 1 - \epsilon$, $v_{ij} = \epsilon/(n-1)$ if $i \neq j$. For each $i \in \mathcal{N}$, let*

$$PR_i = \pi(\mathbf{v}_i).$$

and we denote as $(PR_i)_j$ the j-th entry of PR_i.

Note that $\mathbf{v}_i$ is a personalization vector (i.e. a positive probability vector) and therefore PR_i is well defined since the properties of G are preserved[1]. Therefore we can apply the same numerical methods that we use to compute the usual PageRank; see [18] for a review of numerical methods.

Definition 2. *Given a directed graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ and $0 < \epsilon < 1$, for each node $j \in \mathcal{N}$ we define the* Competitivity interval *$S_C(j)$ as*

$$S_C(j) = [\min_{i\in\mathcal{N}} (PR_i)_j, \max_{i\in\mathcal{N}} (PR_i)_j].$$

Definition 3. *Given a directed graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$, and $0 < \epsilon < 1$ we define the* Competitivity matrix *of the graph, $C = [C_{ji}] \in \mathbb{R}^{n\times 2}$, as follows*

$$C_{j,1} = \min_{i\in\mathcal{N}} (PR_i)_j, \qquad C_{j,2} = \max_{i\in\mathcal{N}} (PR_i)_j.$$

Example 1. *Let $\mathcal{G}$ be the graph shown in Fig. 1.*

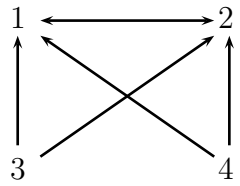


Figure 1: Graph of example 1.

The stochastic matrix of this graph is

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \end{bmatrix}.$$

[1] In [15] personalization vectors with null entries are used, but G is modified to ensure irreducibility.

For the sake of simplicity we take $\epsilon = 0$*; in this example irreducibility of* G *is not needed to compute the PageRank vector. Therefore we have* $\mathbf{v}_i = \mathbf{e}_i$*. Note that* $PR_1 = \pi([1,0,0,0])$*,* $PR_2 = \pi([0,1,0,0])$*, etc. Computing the* Competitivity matrix *we obtain:*

$$C = \begin{bmatrix} 0.425 & 0.54 \\ 0.425 & 0.54 \\ 0.0 & 0.15 \\ 0.0 & 0.15 \end{bmatrix}.$$

Note that the *Competitivity interval* of node i is shown in the i-th row of matrix C. We now define the *Competitivity group*.

Definition 4. *Given a directed graph* $\mathcal{G} = (\mathcal{N}, \mathcal{E})$*, and* $0 < \epsilon < 1$*, a* Competitivity group *is a subset of* $\mathcal{N}$*. Nodes* $i \in \mathcal{N}$ *and* $j \in \mathcal{N}$ *belong to the same* Competitivity group *if* $S_C(i) \cap S_C(j) \neq \emptyset$*.*

Example 2. *For the graph of example 1 we have only two Competitivity groups:* $\{1,2\}$ *and* $\{3,4\}$*. The utility of these groups is the following. If we use the personalization vector to enhance the importance of some nodes we could make a difference between nodes* $\{1,2\}$ *and also between nodes* $\{3,4\}$ *but, clearly, these two groups of nodes compete in different levels (*leagues *or* markets *using sports term or economical term, respectively). Note that this technique to produce groups is different from the usual techniques to study community structure, see [8].*

By using the personalization vector we could never have node 3 *or node* 4 *in a better position (i.e. with a greater PageRank) than nodes* 1 *and* 2*. The* Competitivity group *is a simple idea and it can be plotted easily: see Fig. 2.*

The application to an SNS can be made enhancing the personalization vector (and thus the PageRank of a particular node) attending to some features of the user; e.g., using the number of visits as a sign of popularity, the number of friends, the activity of the user, etc. Therefore by using these features we can modify the ranking inside each Competitivity group. Of course we are open to introduce the possibility to change from one Competitivity group to another attending to other features of the SNS. In this case the present model gives a lot of possibilities to future modeling.

Regarding the visualization of the graph, note that the *Competitivity matrix* of a graph gives an idea of to what extend we can use $\mathbf{v}$ to modify the PageRank of each node and this can be plotted as in Fig. 2.

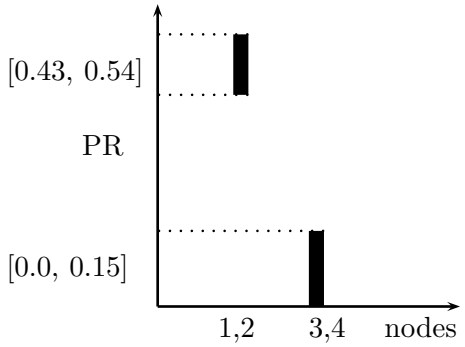


Figure 2: *Competitivity groups* of example 1.

We remark here that the computational cost of this algorithm to obtain the competitivity matrix is $(n+1)$ times the cost of the usual PageRank. Since we are dealing with SNSs this cost is not so high; matrices of SNSs are very small compared with the entire WWW.

Definition 5. *Given a directed graph* $\mathcal{G} = (\mathcal{N}, \mathcal{E})$*, and* $0 < \epsilon < 1$*, the* Leadership group *is a subset of* $\mathcal{N}$*. Node* $j \in \mathcal{N}$ *belong to the* Leadership group *if, for some* $i \in \mathcal{N}$ *it holds that* $(PR_i)_j \geq (PR_i)_k$ *for all* $k \neq j$*. i.e. for some personalization vector* $\mathbf{v}_i$ *node* j *has the greatest PageRank.*

Example 3. *For the example 1 we have:* $PR_1 = [\ 0.54 \quad 0.46 \quad 0.0 \quad 0.0\ ]^T$*,* $PR_2 = [\ 0.46 \quad 0.54 \quad 0.0 \quad 0.0\ ]^T$*,* $PR_3 = [\ 0.425 \quad 0.425 \quad 0.15 \quad 0.0\ ]^T$*,* $PR_4 = [\ 0.425 \quad 0.425 \quad 0.0 \quad 0.15\ ]^T$*. Therefore the* Leadership group *is* $\{1, 2\}$*.*

References

[1] K. A. Fredericks, M. M. Durland, The Historical Evolution and Basic Concepts of Social Network Analysis, New Directions For Evaluations, No. 107, 15-23, 2005.

[2] M. Thelwall, Interpreting social science link analysis research: A theoretical framework, Journal of the American Society for Information Science and Technology, Volume 57 , Issue 1, pp. 60-68, 2006.

[3] F. Stutzman, D. Boyd, S. Golder, R. Recuero, A. Zollers, Research directions in social network websites, Proceedings of the American Society

for Information Science and Technology, Volume 44, Issue 1, pp. 1-4, 2008.

[4] N. Henry, J-D. Fekete, MatLink: Enhanced Matrix Visualization for Analyzing Social Networks, Lecture Notes in Computer Science (Proceedings of the 13th IFIP TC13 International Conference on Human-Computer Interaction, INTERACT'07), 4663, pp. 288-302, Springer. 2007.

[5] Dil M. A. Hussain, Destabilization of Terrorist Networks through Argument Driven Hypothesis Model, Journal of Software, Volume: 2, Issue: 6, pp. 22-29, 2007.

[6] M. Thelwall. Social Networks, Gender and Friending: An Analysis of MySpace Member Profiles, Journal of the American Society for Information Science and Technology. Volume 59 , Issue 8 (2008), pp. 1321-1330.

[7] E. Hargittai, Whose space? Differences among users and non-users of social network sites, Journal of Computer-Mediated Communication, 13 (2008) 276-297.

[8] M.E.J. Newman and M. Girvan Finding and evaluating community structure in networks Physical Review E 69 (2004) 026113.

[9] Community structure identification L. Danon, J. Duch, A. Arenas and A. Daz-Guilera, in "Large Scale Structure and Dynamics of Complex Networks: From Information Technology to Finance and Natural Science", World Scientific, 93-113 (2007).

[10] A-L. Barabasi, Linked: How Everything Is Connected to Everything Else and What It Means for business, Science and Everyday Life, Plume Editions, New York, 2003.

[11] R. Albert, A.L. Barabsi, Statistical mechanics of complex networks, Rev. Mod. Phys., Vol. 74, No. 1, 2002.

[12] D. M. Boyd, N. B. Ellison, Social Network Sites: Definitions, History, and Scholarship, Journal of Computer-Mediated Communication, 13 (2008) 210-230.

[13] L. Page, S. Brin, R. Motwani, T. Winograd, The PageRank Citation Ranking: Bringing Order to the Web, Stanford Digital Library Technologies Project, 1999.

[14] A. N. Langville, C. D. Meyer. Google's Pagerank and Beyond: The Science of Search Engine Rankings, Princeton University Press, 2006.

[15] T. H. Haveliwala, Topic-sensitive PageRank: A context-sensitive ranking algorithm for web search, IEEE Transactions on knowledge and data engineering, vol. 15, No. 4. 2003.

[16] G. Jeh, J. Widow, Scaling personalized web search, Technical Report, Standford University, 2002.

[17] S. Serra-Capizzano, Jordan Canonical Form of the Google Matrix: A Potential Contribution to the PageRank Computation, SIAM Journal on Matrix Analysis and Applications, Volume 27 , Issue 2 , pp. 305 - 312 , 2005.

[18] F. Pedroche, Métodos de cálculo del vector PageRank (in spanish), Bol. Soc. Esp. Mat. Apl., vol. 39, pp. 7-30, 2007.

QR-decomposition for large and sparse linear systems in computed tomography *

M.-J. Rodríguez-Alvarez, †

Instituto de Matemática Multidisciplinar, Universidad Politécnica de Valencia

Edificio 8G, 2°,Camino de Vera, 46022 Valencia, (Spain)

Filomeno Sánchez ‡, Antonio Soriano, §

Instituto de Física Corpuscular (IFIC)

Edificio Institutos de Investigación, Paterna, Valencia E46071 (Spain)

Amadeo Iborra ¶

Instituto de Matemática Multidisciplinar, Universidad Politécnica de Valencia,

Edificio 8G, 2°, Camino de Vera, 46022 Valencia, (Spain)

December 10, 2009

1 Introduction

We present a general method for solving large and sparse linear systems. The method is particularly efficient for systems with multiple right-hand sides. The proposed method is based on Givens QR factorization of the original matrix A, assuming that A has full rank. We propose the QR decomposition with Givens algorithm as a direct method to solve the linear system. QR decomposition can be a large computational procedure. However, once it has

*This work is partially supported by Generalitat Valenciana GVPRE/2008/303 and the Spanish M.E.C. Grant MTM2009-08587.

†e-mail: mjrodri@imm.upv.es

‡e-mail: filomeno@ific.uv.es

§e-mail: Antonio.Soriano@ific.uv.es

¶e-mail: amibcar@fiv.upv.es

been calculated for a specific system, matrices Q and R are stored and used for any right-hand side of the system. Then. the triangular system is solved. Implementation of the QR decomposition with pivoting techniques in order to take more advantage of the sparsity of the system matrix and the numerical stability is also discussed here. In medicine, computed tomographic images are reconstructed from a high number of measurements of X-ray transmission through the patient (projection data). The mathematical model used to describe a computed tomography device is a large system of linear equations of the form $AX = B$, where matrix A describes the systems geometry, B is the projection data and X is the image we want to construct . The proposed method was implemented to reconstruct computed tomography images.

We study the QR decomposition techniques for sparse matrix used to described computed tomography system geometry [1], [2]. Nevertheless results are applicable to any linear sparse systems of equations which have the same requirements: full-rank matrices. It is desirable to have no more than 3% of non-zero elements, because fill-in could cause problems. Sparse matrix are one of the main data structures used in large-scale scientific and engineering applications for representing linear systems of equations. Many linear systems have thousands of variables, but each individual variable only depends on a few variables. This leads to equations where most of the coefficients are zero. Sparse matrix data structures exploit this feature and try to minimize the amount of memory used by only allocating memory for the non-zero elements.

One way of solving a linear system involves decomposition into QR [3] of a $m \times n$ matrix A. We assume that matrix A verifies $m \geq n$. For $m < n$ we compute QR decomposition of A^T. We assume that A has full rank. Orthogonal decompositions have been used extensively for small and dense matrices, because it is well known that such decomposition is numerically stable. The storage and the operation count requirements in the decomposition are $O(mn)$ and $O(mn^2)$ respectively [4].

2 Sparse matrix representation

Data structure for sparse matrix is a very important issue. Data structure must be compact and easily accessed by the applications it has been designed. Sparse row-wise representation of a matrix A with m rows and n columns is given by three one-dimensional arrays.

For instance consider the sparse matrix

$$A = \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & 0 & a_{23} \\ a_{31} & a_{32} & 0 \\ a_{41} & 0 & 0 \end{bmatrix} \tag{1}$$

If nz is the number of non-zero elements, (five in this case), for a sparse row-wise storing scheme it is necessary two arrays of nz elements (RA and JA), and one one array of $m+1$, (five in this case), elements IA. RA contains the matrix non-zero elements $RA = a_{11}, a_{23}, a_{31}, a_{32}, a_{41}$, ordered by increasing row index and then by column. JA contains the column index of the elements stored in RA, in this example $JA = 1, 3, 1, 2, 1$. Finally, the IA array contains the pointer element to the first element of each row in RA (and also in JA). In this example $IA = 1, 2, 3, 5, 6$. Note that element (m+1) points to the first empty element.

$$\begin{array}{lcccccc} IA & = & 1 & 2 & 3 & & 5 & 6 \\ & & \downarrow & \downarrow & \downarrow & & \downarrow & \downarrow \\ RA & = & a_{11} & a_{23} & a_{31} & a_{32} & a_{41} & \\ JA & = & 1 & 3 & 1 & 2 & 1 & \end{array} \tag{2}$$

Notice that the column indices in JA are ordered within a given row. It is worth to emphasize that certain applications don't require an ordered row-wise format, e.g some sparse matrix multiplications [5]. In our case, we use sparse ordered row-wise, because we must "find" quickly non-zero elements placed in the inferior half triangle of the A matrix.

In the same way it is also used the sparse order column-wise:

$$\begin{array}{lcccccc} JA & = & 1 & & & 4 & 5 & 6 \\ & & \downarrow & & & \downarrow & \downarrow & \downarrow \\ RA & = & a_{11} & a_{31} & a_{41} & a_{23} & a_{32} & \\ IA & = & 1 & 3 & 4 & 2 & 1 & \end{array} \tag{3}$$

Notice tat JA and IA interchange their roles with respect to the row-wise representation. A is stored in sparse row-wise scheme.

3 Givens rotations and row ordering

We propose to solve the linear system $AX = B$ by QR factorization based on Givens rotations [3], [6], [7], [8]. Givens-based QR decomposition factorizes

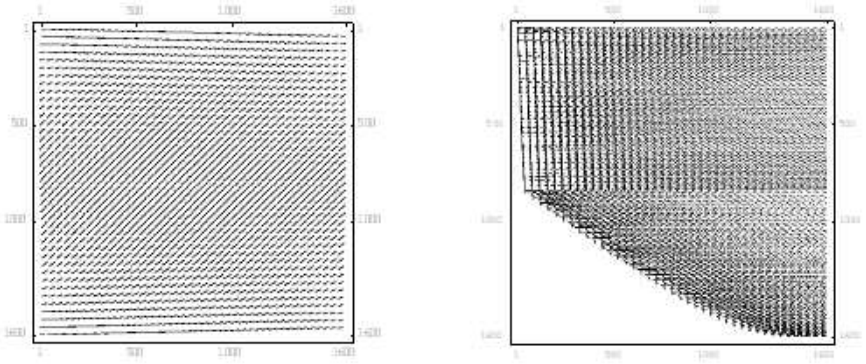


Figure 1: Effect of row ordering algorithm in a (1600 x 1600) matrix.

the $m \times n$ matrix A into the product QR, where Q is orthogonal and R is upper triangular (or upper trapezoidal if $m < n$).

In order to reduce the number of Givens rotations and consequently the number of fill-in, is important to use a new row ordering strategy [9].We propose a very simple algorithm which reduces considerably the number of eliminations and consequently the number of fill-in.

The proposed ordering algorithm assign to each row a label with the column number of the first non zero element in that row. Then it orders rows by increasing label number. By permuting rows, we can reduce 70% the number of eliminations. Eliminations could produce fill-in which needs more rotations and so on. Permuting rows is a quasi-instantaneous process.

Fig. 1 shows the final result effect obtained with the proposed ordering algorithm for a sparse matrix of 1600×1600 with a 2% of non-zero elements.

The new order for rows is stored in an array P of m elements, which is stored in an external file to save RAM.

4 Algorithms

Two new algorithms for sparse matrix are decomposition are developed. The first algorithm, computes R matrix and stores data in order to compute Q matrix. The second algorithm computes Q matrix. These algorithms are implemented in two separated programs, although both algorithms share a

common data structure.

Finally a backward substitution strategy has been used to solve upper the triangular system and obtain the final solution for

$$R \cdot X = Q^T B \tag{4}$$

4.1 Discussion and Conclusions

We are tested the proposed algorithm with $LAPACK$ [10] library $SGELSS$ to compared results. We have compared result for a matrix A of 10^8 elements in (20000×5000) with a 2% of non zero elements. It has been used a 32 bits 2GHz microporocessor with 2 GB of RAM.

Differences between both results are negligible within the significant digits, but computing time, and maximum size of system matrices are quite differents.

It isn't possible to solve bigger systems with our PC with $LAPACK$ dense format, but computing time are very good; $LAPACK$ only need ≈ 75 min to solve a (20000×5000) system. Our algorithm works more slowly (600 min for the same system) but we haven't limit in the size of matrix A. The real limit (and could be improved) is the number of the non-zero elements. We have managed up to 10^9 non-zero elements. This implies that assuming a 1% of sparsity, our matrix in dense format have 10^{11} elements.

We have used the algorithm to reconstruct images of computed tomography (CT) obtaining results of similar quality than classical techniques as MLEM (Maximum Likelihood Expectation Maximization) [11], [12], and FBP (Forward Back Projection) [13], [14]. It should be noticed that although QR factorization is a very slow process, this must be made only one time for a specific geometry of a CT system. Then it could be stored and all we need to reconstruct the CT image is to solve an upper triangular systems and to make a product (quasi-instant procedure).

References

[1] C. Mora, M. J. Rodríguez-Alvarez, J.V., Romero: New pixellation scheme for CT algebraic reconstruction to exploit matrix symmetries. Comput. Math. Appl. **56**(3)(2008) 715–726.

[2] C. Mora, M. J. Rodríguez-Alvarez, I. Baeza, Blobs-bases algebraic reconstruction methods using polar grids. In Modelling for engineering and medicine, 2008.

[3] G. H. Golub, C.F. Van Loan. Matrix Computations 3 ed, The Johns Hopkins University Press, 1996.

[4] E. G.Y., Ng, Row elimination in sparse matrices using rotations. Thesis University of Waterloo, 1983.

[5] F. G. Gustavson, Two fast algortihm for sparse matrices: Multiplication and permuted trnaspositions. ACM Trans. Math. Softw. 4(3)(1978)250-269.

[6] T. A Davis, Direct methods for sparse linear systems, Siam, , 2006, Chapter 7.

[7] A. George and M. T. Heath, Solution of sparse linear-least squares problems usign Givens rotation, Linear Algebra Appl., 34 (1980) 69–83.

[8] W. Givens, Computation of plane unitary rotations transforming a general matrix to triangular form, J. Soc. Indst. Appl. Math., 6 (1958) 26–50.

[9] A. George and E. Ng, On row and column orderings for sparse least square problems ysystems., SIAM J. Svci. Statist. Comput., 8 (1983) 390–409.

[10] http:www.netlib.orglapack, last update 2009.

[11] G. T. Herman, Image Reconstruction from Projections: The Fundamentals of Computed Tomography,9(3), (1982), 446-448.

[12] E. Levitan, G.T. Herman, A maximum a posteriori probability expectation maximization algorithm for image reconstruction in emission tomography, IEEE Transactions on Medical Imaging MI-6:(1987)185-192,

[13] L. A. Shepp, Y. Vardi: Maximum likelihood reconstruction for emission tomography. IEEE Trans. Med. Imaging, **1**(2), (1982) , 113–122.

[14] S. Basu, Y. Bresler, $O(N^2 log_2 N)$ Filtered backprojection algortithm for tomography. IEEE Trans. Image Processing, **9**(10), (2000), 1760–1773.

Validation of a Code to Model Cavitation Phenomena in Diesel Injector Nozzles*

F. J. Salvador†, J.-V. Romero*,
M.-D. Roselló‡ and J. Martínez§

†, § CMT-Motores Térmicos,

Universidad Politécnica de Valencia,

Camino de Vera, s/n, 46022 Valencia, Spain

∗, ‡ Instituto de Matemática Multidisciplinar,

Universidad Politécnica de Valencia,

Edificio 8G, Piso 2, 46022 Valencia, España.

December 10, 2009

1 Introduction

The behavior of the internal nozzle flow has a strong influence in the spray and its atomization characteristics [1]. A good understanding of the internal flow physics inside the nozzle is fundamental to predict spray development. This is particularly true when cavitation occurs within the nozzle passage.

Cavitation can occur when a fluid with high velocity passes through a contraction. Due to the abrupt change in flow direction, the flow tends to

*This work was partly sponsored by "Vicerrectorado de Investigación, Desarrollo e Innovación" of the "Universidad Politécnica de Valencia" in the frame of the project "Estudio del flujo en el interior de toberas de inyección Diesel", Reference N^o 3150, by "Generalitat Valenciana" in the frame of the project with the same title and by "Spanish M.C.Y.T. and FEDER" grant TRA2007–68006–C02–02. This support is gratefully acknowledged by the authors.

†e-mail: fsalvado@mot.upv.es

separate from the wall at the inlet section. As a consequence, a recirculation phenomenon appears accompanied by a pressure fall due to the acceleration of the fluid. If conditions in the nozzle are such that static pressure decreases under saturation pressure of the working fluid, local change of state from liquid to vapour takes place, this phenomenon is called hydrodynamic cavitation.

Under the injection conditions in modern Diesel engines (where injection pressure can reach up 180 MPa) cavitation often occurs in fuel injector nozzles, introducing vapour bubbles into the flow, increasing the maximum velocity in the center (liquid core). The velocity is increased for two reasons when the fluid is cavitating. Firstly, if there is vapour along the wall the liquid will not have a no-slip condition boundary thus allowing the velocity of the liquid to increase [2, 3]. Moreover, due to the formation of vapour bubbles the liquid cannot fill the entire channel (geometrical diameter), and so, the diameter (effective) is reduced with regard to the geometrical one [2, 4]. Furthermore, cavitation increases spray cone angle and so it is expected to improve the air-fuel mixing process [2, 5, 6].

Up to now, the numerical simulations performed to study cavitation phenomena have shown serious difficulties when simulating the internal nozzle flow in Diesel engine conditions. Nevertheless, a new code has been developed by OpenCFD ® Ltd [7], which has been validated for small pressure drop conditions [8, 9], but not at high injection pressures in small orifices, which are both typical characteristics of Diesel injector nozzles. The goal of this study has been to evaluate the potential of the code to treat cavitation in such as conditions by comparing the results from numerical simulations with experimental measurements.

2 Description of the CFD approach

Due to high pressures and velocities that occur in diesel injectors, the use of a homogeneous equilibrium model with a barotropic equation of state is the most suitable method to model cavitation. In this code, it is assumed that liquid and vapour are always perfectly mixed in each cell (homogenous equilibrium) and the compressibility of both phases is taken into account.

To calculate the growth of cavitation, a common barotropic equation of

state, which relates pressure and density, is used:

$$\frac{D\rho}{Dt} = \Psi \frac{Dp}{Dt}, \tag{1}$$

where Ψ is the compressibility of the mixture, which is the inverse of the speed of sound squared:

$$\Psi = \frac{1}{a^2}. \tag{2}$$

This equation can be used directly in the continuity equation to formulate a pressure equation. The barotropic equation of state should be consistent with the liquid and vapour equations of state both at the limits when there is pure liquid or pure vapour, and at intermediate states when there is a mixture of them. Both phases can be defined with a linear equation of state:

$$\rho_v = \Psi_v \, p, \tag{3}$$

$$\rho_l = \rho_l^0 + \Psi_l \, p. \tag{4}$$

The amount of vapour in the fluid is determinated using the parameter γ. It is worked out as:

$$\gamma = \frac{\rho - \rho_{l\,sat}}{\rho_{v\,sat} - \rho_{l\,sat}}, \tag{5}$$

where

$$\rho_{v\,sat} = \Psi_v \, p_{sat}. \tag{6}$$

It can be seen that in a flow with no cavitation $\gamma = 0$, whereas for fully cavitated flow $\gamma = 1$. In equation (6), Ψ_v is the compressibility of the vapour.

These equations together form the mixtures equilibrium equation of state:

$$\rho = (1-\gamma)\,\rho_l^0 + (\gamma\,\Psi_v + (1-\gamma)\,\Psi_l)\;p_{sat} + \Psi(\gamma)\,(p - p_{sat}), \tag{7}$$

with

$$\rho_l^0 = \rho_{l\,sat} - \Psi_l \, p_{sat}. \tag{8}$$

As far as the mixture's compressibility is concerned, it is modeled by a simple linear model:

$$\Psi = \gamma\,\Psi_v + (1-\gamma)\,\Psi_l, \tag{9}$$

with Ψ_l equal to the compressibility of the liquid.

As done for compressibility, it is possible to obtain the viscosity of the mixture with a linear model:

$$\mu = \gamma\,\mu_v + (1-\gamma)\,\mu_l. \tag{10}$$

The methodology used by the solver starts solving the following continuity equation for ρ:

$$\frac{\partial \rho}{\partial t} + \nabla\,(\rho\,u) = 0. \tag{11}$$

The spatial discretization of ρ in the divergence term $\nabla(\rho\,u)$ is made by using a Gauss upwind numerical scheme. It is well known that a first-order scheme leads to a much stable numerics, but it is more sensitive to mesh coarseness and it normally presents more numerical diffusion than a second-order scheme. Nevertheless, because of large gradients in pressure and density present in Diesel nozzles, stability is a main argument and so, an upwind scheme has been used. Numerical diffusion and grid-independence have been controlled by an adequate-mesh refining done in a previous set up of the model, where a mesh sensitivity study was performed.

It's important to remark that the numerical schemes used by default have been changed to improve the numerical convergence, trying several ones, such as the MUSCL scheme or others options available in the code to solve the different terms included in the equations above, allowing to chose the most suitable.

The value of ρ obtained, is used to determine preliminary values for γ and Ψ by means of equation (5) and equation (9), and also, for solving momentum equation (equation (12)) which is used to get the matrices used to calculate the pressure-free velocity, u:

$$\frac{\partial \rho\,u}{\partial t} + \nabla\,(\rho\,u\,u) = -\nabla p + \nabla\,(\mu_f\,\nabla u)\,. \tag{12}$$

The same Gauss upwind numerical scheme is used for the velocity divergence scheme and a Gauss linear corrected scheme (second order) is used for the Laplacian term discretization.

Following, an iterative PISO algorithm is used to solve for p and correct the velocity to achieve continuity. The equation solved for the PISO loop is the continuity equation transformed into a pressure equation by use of the equation of state (Eq. (7)):

$$\frac{\partial \Psi\,p}{\partial t} - \left(\rho_l^0 + (\Psi_l - \Psi_v)\;p_{sat}\right)\;\frac{\partial \gamma}{\partial t} - p_{sat}\,\frac{\partial \Psi}{\partial t} + \nabla\,(\rho\,u) = 0. \tag{13}$$

Once continuity has been reached, the properties ρ, γ, and Ψ are updated by means of equations (7), (5) and (9) respectively which are taken into account to solve again momentum equation, and so, repeating the algorithm until convergence.

The time step is limited by both the Courant number and the acoustic Courant number, defined as:

$$Co = \max\left(\frac{|u|}{\Delta x}\right)\Delta t, \tag{14}$$

$$Co_{acoustic} = \max\left(\frac{1}{\sqrt{\Psi}\Delta x}\right)\Delta t. \tag{15}$$

The Courant number was chosen to be limited to 0.125, while the acoustic Courant number was limited to 12.5.

Taking into account the important velocities that are normally found in Diesel applications (due to severe pressure drops) and the small size of the cells, as a result of a preliminary study that allowed the appropriate mesh fineness to be chosen, time steps used are around 10^{-9} s in both simulations.

3 Simulated geometries

Firstly, the results included in the report "Comprehensive hydraulic and flow field documentation in model throttle experiments under cavitation conditions" [10] were used as a reference to validate the code implemented for OpenFOAM, where the behavior of the flow is studied in a throttle eroded into 0.3 mm thick steel sheet (see Fig. 1).

On the other hand, Fig. 2 shows the one-hole nozzle geometry used in the second validation, whose diameter at the orifice inlet is 163 μm and whose outlet diameter is 165 μm. In order to compare with stabilized flow conditions, the geometry representing of fully needle lift conditions (i.e. 250 μm) has been modeled.

The calculation for the one-holed nozzle and simple contraction nozzle was done for a 1/4 of the whole geometry, thanks to the nozzle symmetry in both cases, breaking up the physical domain into 15789 and 121820 cells respectively.

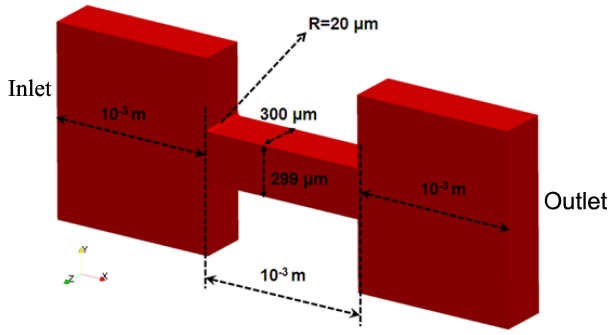


Figure 1: Geometry used in the first validation.

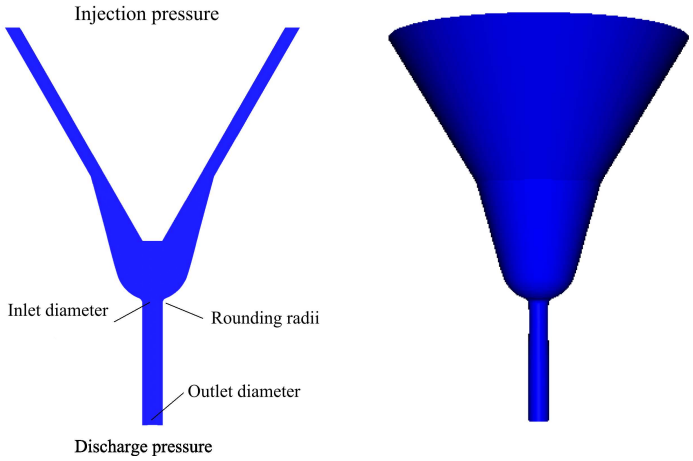


Figure 2: Geometry of the nozzle used in the second validation.

4 Results

Despite of large volumes of information obtained with the CFD code explained before, only mass flow and velocity at the outlet section have been compared in the simple contraction nozzle validation. The operating conditions have been fixed to reproduce critical cavitation conditions, adjusting the pressure inlet to 100 bars and the pressure outlet to 25.65 bar.

The mass flow, defined as

$$\dot{m}_f = \int \rho \, u \, dA \tag{16}$$

is compared using a time-averaged value of the last 25 μs, although the oscil-

lations of the numerical results time step to time step are perfectly negligible. Numerical simulation predicts a mass flow of 7.76 kg/s, whereas the report used as reference indicates 7.72 kg/s.

On the other hand, the second parameter to be evaluated, the velocity at the outlet section of the channel, has been obtained integrating the velocity in the direction of the flow using the following equation:

$$u_{x_{mean}} = \frac{1}{A} \int u_x \, dA. \tag{17}$$

The outlet velocity was calculated for several time steps and averaged in time over the last 125 μs, predicting a value of 104 m/s. The cavitation model shows again a negligible overestimation of the velocity about 0.01% relative to the value given in the report (103.3 m/s).

As far as the one-hole nozzle is concerned, measurements of mass flow and momentum flux (impact force of the spray) and effective velocity were also used to validate the model.

As can be seen in Figure 3, the results obtained with numerical simulations using three different levels of injection pressure (30, 70 and 110 MPa) and adjusting the discharge pressure to 4 MPa, are very closed to experimental data. These differences increase at high injection pressures, where the deviation between both values of momentum flux reaches up 8%.

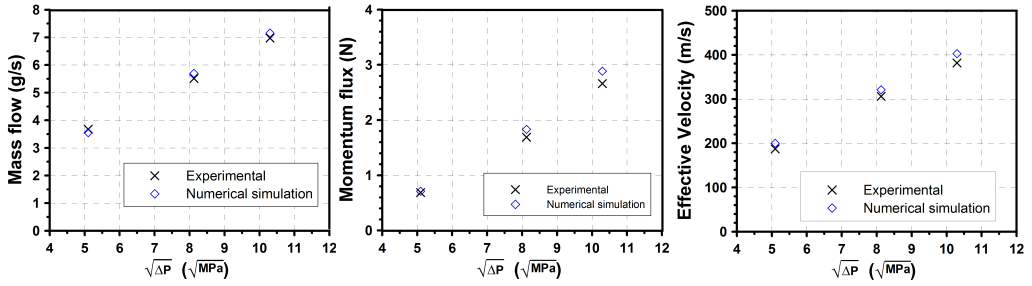


Figure 3: Comparison of experimental and numerical results in terms of mass flow, momentum flux and effective injection velocity as a function of $\sqrt{\Delta P}$.

5 Conclusions

From the realized study the following main conclusions can be drawn:

- A code to model cavitation phenomena has been applied for the simulation of 3D cavitating flows assuming the nozzle flow as a homogeneous mixture of liquid and vapour.
- An extended validation of the code has been performed in two different geometries: a simple contraction nozzle and a real Diesel injector nozzle.
- With regard to the simple contraction nozzle, the code has been able to predict variables as important as mass flow or velocity at the outlet.
- As far as the one-holed nozzle is concerned, the code showed a fairly good agreement with respect to experimental results, in terms of mass flow, momentum flux and effective velocity at the orifice exit.

NOMENCLATURE

a: speed of sound
A: area
$\dot{m}_f$: mass flow / mass flux
p: pressure
P_{back}: discharge back pressure
P_{inj}: injection pressure
p_{sat}: vaporisation pressure
t: time
u: velocity
u_x: velocity in the x direction
$u_{x_{mean}}$: mean velocity in the x direction
Greek symbols:
γ: vapour fraction
ΔP: pressure drop, $\Delta P = P_{inj} - P_{back}$
Δt: time step
Δx: cell size
μ: fluid viscosity
ρ: fluid density
$\rho_{l\,sat}$: liquid density at saturation
ρ_l^0: liquid density at a given temperature condition
$\rho_{v\,sat}$: vapour density at saturation
Ψ: fluid compressibility
Subscripts:
l: liquid
v: vapour

References

[1] G.M. Faeth, L.P. Hsiang and P.K. Wu, Structure and Breakup Properties of Sprays *International Journal of Multiphase Flow* **21** 99–127 (1995).

[2] H. Chaves and F. Obermeier, *Correlation between Light Absorption Signals of Cavitating Nozzle Flow within and outside of the Hole of a Transparent Diesel Injection Nozzle*, in Proc. ILASS-EUROPE. Manchester, UK. 1998, July 6–8.

[3] R. Payri, J.M. García, F.J. Salvador and J. Gimeno, Using spray momentum flux measurements to understand the influence of diesel nozzle geometry on spray characteristics. *Fuel* **84** 551–561 (2005).

[4] D.P. Schmidt and M.L. Corradini, The internal Flow of Diesel Fuel Injector Nozzles: a Review. *International Journal of Engine Research* **2 (1)** 295–324 (1995).

[5] F. Payri, V. Bermudez, R. Payri and F.J. Salvador, The influence of cavitation on the internal flow and the spray characteristics in Diesel injection nozzles. *Fuel* **83** 419–431 (2004).

[6] R. Payri, F.J. Salvador, J. Gimeno and J. de la Morena, Study of cavitation phenomena based on a technique for visualizing injected bubbles in a liquid pressurized chamber. *Internal Journal of Heat and Fluid Flow.* In press, doi:10.1016/j.ijheatfluidflow.2009.03.011.

[7] OpenCFD® is a registered trade mark, http://www.opencfd.co.uk/

[8] F. Peng Kärrholm, *Numerical Modelling of Diesel Spray Injection, Turbulence Interaction and Combustion*, PhD. Thesis, Chalmers University of Technology, 2008.

[9] F. Peng Kärrholm, H. Weller and N. Nordin, *Modelling injector flow including cavitation effects for Diesel Applications*, Proceedings of FEDSM2007, 5th Joint ASME/JSME Fluids Engineering Conference, July 30 – August 2, San Diego, California, USA.

[10] E. Winklhofer, E. Kull, E. Kelz and A. Morozov, *Comprehensive hydraulic and flow field documentation in model throttle experiments under cavitation conditions*, ILASS-Europe 2001.

Modeling of Drinking Behavior in Spain.

E. Sánchez†, F. J. Santonja‡,
M. Rubio* and J. L. Morera*

† Unidad de Conductas Adictivas de Catarroja

Agencia Valenciana de Salud

‡ Departamento de Estadística e Investigación Operativa

Universidad de Valencia

∗ Instituto de Matemática Multidisciplinar

Universidad Politécnica de Valencia

December 10, 2009

1 Introduction

In this work, we present an epidemiological-type mathematical model to study the transmission dynamics and evolution of the alcohol consumption in Spanish population (See Table 1 [1]). This type of epidemiological models also have been used in the study of another social epidemics (ecstasy or heroin addiction [2, 3]) and in the approach to another topics that spread by social contact like obesity or extreme behaviors [4, 5].

2 Mathematical model

2.1 Building the model

In this paper we assume the proposal showed in [6, 7] and treat alcohol consumption as a disease that spreads by social contact. We suppose that these contacts influence in the probability of transmission of the consumption habits.

Table 1: Evolution of the proportion of $A(t)$ (Non-consumers), $M(t)$ (Non-risk consumers) and $R(t)$ (Risk consumers) subpopulation for different years.

	$A(t)$	$M(t)$	$R(t)$
1997	36.2%	58.1%	5.7%
1999	38.3%	57.8%	3.9%
2001	36.3%	58.1%	5.6%
2003	35.9%	58.8%	5.3%
2005	35.4%	59.1%	5.5%
2007	40.0%	56.6%	3.4%

For model building, 15-64 years old Spanish population is divided into three subpopulations [8]:

- $A(t)$: Non-consumers, individuals that have never consumed alcohol or they infrequently have alcohol consumption.
- $M(t)$: Non-risk consumers, individuals with regular low consumption. To be precise, men who consume less than 50 cubic centimeters of alcohol every day and women who consume less than 30 cubic centimeters of alcohol every day.
- $R(t)$: Risk consumers, individuals with regular high consumption, i.e., men who consume more than 50 cubic centimeters of alcohol every day and women who consume more than 30 cubic centimeters of alcohol every day.

Furthermore, we consider the following assumptions:

1. We assume population homogeneous mixing. That is, each individual can contact with any other individual [9].
2. The transitions between the different subpopulations are determined as follows:
    - We consider that the new recruited 15-years-old individuals become members of the $A(t)$ subpopulation.

- Once an individual starts regular alcohol consumption he/she becomes a non-risk consumer, $M(t)$. If this person increases his/her consumption habit he/she can become a risk consumer, $R(t)$.
- Individuals of subpopulation $R(t)$ becomes a member of subpopulation $A(t)$ if the alcohol consumption is reduced at an appropriate rate.

3. The transitions described above can be modeled as follows:

   - An individual in $A(t)$ transits to $M(t)$ because people in $M(t)$ or $R(t)$ transmit the alcohol consumption habit by social contact at rate β. Therefore, this is a nonlinear term modeled by $\beta A(t)(M(t)+R(t))/P(t)$. We consider $P(t)= A(t)+M(t)+R(t)$.
   - An individual in $M(t)$ transits to $R(t)$ at rate α proportionally to the size of $M(t)$ if his/her alcohol consumption increases. Hence, this is a linear term modeled by $\alpha M(t)$.
   - An individual in $R(t)$ transits to $A(t)$ when decides to give up the alcohol consumption and to go into therapy. An individual in $R(t)$ transits to $A(t)$ at rate γ proportionally to the size of $R(t)$. Hence, this is a linear term modeled by $\gamma R(t)$.

Under the above assumptions, dynamic alcohol consumption model for Spanish population is given by the following nonlinear system of ordinary differential equations:

$$A'(t) = \mu P(t) + \gamma R(t) - d_A A(t) - \beta A(t) \frac{[M(t) + R(t)]}{P(t)} \quad (1)$$

$$M'(t) = \beta A(t) \frac{[M(t) + R(t)]}{P(t)} - dM(t) - \alpha M(t) \quad (2)$$

$$R'(t) = \alpha M(t) - \gamma R(t) - dR(t) \quad (3)$$

$$P(t) = A(t) + M(t) + R(t) \quad (4)$$

where the constant parameters of the model are:

- μ, birth rate in Spain.
- γ, rate at which a risk consumer becomes a non-consumer.

- d_A, death rate in Spain.
- β, transmission rate due to social pressure to increase the alcohol consumption (family, friends, marketing, TV, etc.).
- d, augmented death rate due to alcohol consumption. Accidents at work, traffic accidents and diseases derived by alcohol consumption are considered.
- α, rate at which a non-risk consumer moves to the risk consumption subpopulation.

2.2 Scaling the model

Data obtained in Table 1 is related to the percentages of population meanwhile model $(1)-(4)$ is related to the number of individuals. It leads us to transform (by scaling) the model into the same units as data, because one of our objetives is to fit data with the model. Hence, following ideas developed in [5] about how to scale models where the population is varying in size we obtain:

adding equations $(1)-(3)$ one gets

$$P'(t) = \mu P(t) - d_A A(t) - dM(t) - dR(t) \tag{5}$$

Dividing both members of (5) by $P(t)$ we have that

$$\frac{P'(t)}{P(t)} = \mu \frac{P(t)}{P(t)} - d_A \frac{A(t)}{P(t)} - d\frac{M(t)}{P(t)} - d\frac{R(t)}{P(t)} \tag{6}$$

If we define the rates (depending on time)

$$a = \frac{A}{P}, m = \frac{M}{P}, r = \frac{R}{P} \tag{7}$$

equation (6) can be transformed into

$$\frac{P'}{P} = \mu - d_A a - dm - dr \tag{8}$$

On the other hand, we compute the derivative of a, defined in (7). Using (8) we obtain that,

$$a' = \frac{A'P - AP'}{P^2} = \frac{A'}{P} - \frac{A}{P}\frac{P'}{P} = \frac{A'}{P} - a[\mu - d_A a - dm - dr] \tag{9}$$

In an analogous way, we also have that,

$$\begin{aligned} m' &= \frac{M'}{P} - m[\mu - d_A a - dm - dr] \\ r' &= \frac{R'}{P} - r[\mu - d_A a - dm - dr] \end{aligned}$$

Now, consider equation (1) . If we multiply it by $1/P$, we have

$$\frac{A'}{P} = \mu\frac{P}{P} + \gamma\frac{R}{P} - d_A\frac{A}{P} - \beta\frac{A}{P}\frac{(M+R)}{P}$$

using (9) and substituting by the corresponding rates defined in (7) one gets

$$a' = \mu + \gamma r - d_A a - \beta a(m+r) - a[\mu - d_A a - dm - dr] \tag{10}$$

Remainder equations can be scaled in the same way to obtain

$$\begin{aligned} m' &= \beta a(m+r) - \alpha m + d_A am - dam - \mu m & (11) \\ r' &= \alpha m - \gamma r + d_A ar - dar - \mu r & (12) \end{aligned}$$

2.3 Estimation of parameters

Now, the estimation of the parameters, for time t in years, is presented:

- $\mu = 0.01\ years^{-1}$ is the average Spanish birth rate between years 1997-2007 [10].
- $d_A = 0.008\ years^{-1}$ is the average Spanish death rate between years 1997-2007 [10].
- $d = 0.009\ years^{-1}$ is the average Spanish alcohol consumption death rate between years 1997-2007.We considered that approximately 4% of mortality is due to the alcohol consumption [11, 12].

- $\gamma = 0.00144\ years^{-1}$ From [13] it can be obtained that around 32% of risk consumers begin a therapy program every year. Furthermore, using data from Table 1, corresponding to National Drug Observatory Reports [1], we obtain that the mean value of population with risk consumption is 5.2%. Moreover, the conclusion obtained in [14] is that a risk consumer takes around ten years before to go into therapy. Therefore, the percentage of risk consumers in therapy per year is 0.16% (0.052*0.32*1/10=0.0016). On the other hand, [15] concludes that around 45% of the individuals on therapy recover in six months. Then, γ =0.0016*0.45*1/0.5=0.00144. Hence, we can consider $\gamma = \gamma_1 * \gamma_2 * \gamma_3 * \gamma_4 * 1/0.5$, where γ_1= 0.052, γ_2= 0.32, γ_3 =1/10 and γ_4 =0.45.

Additionally, taking as the initial condition of the scaled model (year 1997, *i.e.*, t=0), $A(t=0) = 0.362$, $M(t=0) = 0.581$ and $R(t=0) = 0.057$, the parameters β and α have been estimated by fitting the model with data from Table 1, and we obtained β =0.0284534 and α =0.000110247.

In order to compute the best fitting, we carried out computations with *Mathematica* [16] and we implemented the function

$$\begin{array}{rccc} \mathbb{F} & : & \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ & & (\beta, \alpha) & \longrightarrow & \mathbb{F}(\beta, \alpha) \end{array}$$

which variables are β and α and such that:

1. Solve numerically (*NDSolve[]*) the system of differential equations (10)−(12) with initial values ($A(t = 0)$ =0.362, $M(t = 0)$ =0.581 and $R(t=0)$ =0.057),

2. For t = 1997, 1999, 2001, 2003, 2005 and 2007 evaluate the computed numerical solution for each subpopulation $A(t)$, $M(t)$, $R(t)$.

3. Compute the mean square error between the values obtained in Step 2 and the data from Table 1.

Function $\mathbb{F}$ takes values in $\mathbb{R}^2$ (β and α) and returns a positive real number. Hence, we can try to minimize this function using the Nelder-Mead algorithm [17, 18], that does not need the computation of any derivative or gradient, impossible to know in this case.

In order to find a global minimum the feasible chosen domain is

$$D = [0,1] \times [0,1] \subset \mathbb{R}^2,$$

and it is divided in disjoint subdomains where, in each one, Nelder-Mead algorithm is applied. We stored all the minima obtained and, among them, the values of β and α that minimize the function $\mathbb{F}$ are

$$\begin{array}{c} \beta = 0.0284534 \\ \alpha = 0.000110247 \end{array} \tag{13}$$

3 Numerical simulation

In Table 2, some of the predictions are presented. It can be noted an increasing trend in non-risk consumers population ($M(t)$) and a decreasing of risk consumers population ($R(t)$).

Table 2: Evolution of proportion of risk consumer ($R(t)$) and non-risk consumers ($M(t)$) subpopulations for the next few years. Percentages are defined by the deterministic model.

$Year$	$R(t)$	$M(t)$
2011	4.9%	58.7%
2013	4.8%	58.8%

References

[1] Spanish Ministry of Health (2008). National Drug Observatory Reports. Retrieved 13th November 2008 from: *http://www.pnsd.msc.es/Categoria2/observa/estudios/home.htm.*

[2] B. Song, M. Castillo-Garsow, K.R. Ros-Soto, M. Mejran, L. Henso, C. Castillo-Chávez, Raves, clubs and ecstasy: the impact of peer pressure, Mathematical Biosciences and Engineering 3 (1) (2006) 249-266.

[3] E. White, C. Comiskey, Heroin epidemics, treatment and ODE modeling, Mathematical Biosciences 208 (1) (2007) 312-324.

[4] L. Jódar, F.J. Santonja, G. González-Parra, Modeling dynamics of infant obesity in the regin of Valencia, Spain, Computer and Mathematics with Applications 56 (3) (2008) 679-689.

[5] F.J. Santonja, A.C. Tarazona, R.J. Villanueva, A mathematical model of the pressure of an extreme ideology on a society, Computer and Mathematics with Applications 56 (3) (2008) 836-846.

[6] S. Galea, C. Hall, G.A. Kaplan, Social epidemiology and complex system dynamic modelling as applied to health behaviour and drug use research, International Journal of Drug Policy 20 (3)(2009) 209-216.

[7] D.M. Gorman, J. Mezic, I. Mezic, P.J. Gruenewald, Agent-based modeling of drinking behavior: a preliminary model and potential applications to theory and practice, American Journal of Public Health 96 (11) (2006) 2055-2060.

[8] R. Altisent, R. Córdoba, J.M. Martín-Moros, Operative criteria for the prevention of the alcoholism, Medicina Clínica 99 (1992) 584-588.

[9] J.D. Murray, Mathematical Biology, third ed., Springer, New York, 2002.

[10] Spanish Statistic Institute. Retrieved 13th November 2008 from: *http://www.ine.es.*

[11] T. Stockwell, P.J. Gruenewald, J. Toumbourou, W. Loxley (ed.), Preventing harmful substance use: the evidence base for policy and practice, Wiley, Chichester, 2005.

[12] M.T. Gómez-Talegón, C. Prada, M.C. Del Río, F.J. Álvarez, F.J, Evolution of the alcohol consumption of the Spanish population between 1993, 1995 and 1997, from the database of the National Health Survey, Adicciones 17 (1) (2005) 17-28.

[13] Valencian Health Department, Profile of drug user to treatment in Valencian Health Department. Report 2006. Retrieved 13th November 2008 from: *http://www.sp.san.gva.es.*

[14] S. Tomás-Dols, J. C. Valderrama-Zurián, A. Vidal-Infer, T. Samper-Gras, M.C. Hernández-Martínez, M.J. Torrijo-Rodrigo, Barriers of accessibility to treatment in alcohol consumers of the region of Valencia, Adicciones 19 (2) (2007) 169-179.

[15] N. Heather, Beyond alcoholism: modern perspectives on alcohol dependence and problems, Adicciones 11 (2) (1999) 463-474.

[16] *http://www.wolfram.com.*

[17] J.A. Nelder, R. Mead, A simplex method for function minimization, The Computer Journal 7 (1964) 308-313.

[18] W.H. Press, B.P. Flannery, S.A. Teukolsky, W. Vetterling, Numerical recipes: the art of scientific computing, Cambridge University Press, 1986.

Stochastic network modelling of the pressure of extreme groups in a society.

F.-J. Santonja*, A.-C. Tarazona†, R.-J. Villanueva ‡
and F.-J. Villanueva-Oller §

⋆ Departamento de Estadística e Investigación Operativa,

Universidad de Valencia, Valencia (Spain)

†, ‡ Instituto de Matemática Multidisciplinar,

Universidad Politécnica de Valencia,

§ Edificio 8G, Piso 2, 46022 Valencia, España.

Centro de Estudios Superiores Felipe II,

Aranjuez, Madrid (Spain)

December 10, 2009

Extreme behaviour is produced by small groups but their actions have an impact in a large amount of people. The fear is the strategy developed by these groups to influence the decisions of the whole population in order to achieve their political goals.

The understanding of the transmission dynamics of such a type of extreme behaviours increases the knowledge of the mechanisms behind the evolution of cultural norms and values. Moreover, it can give us tools to know their evolution a priori, eventually disappearing or getting their objectives.

To our knowledge, the antecedents of mathematical modelling approaches where the spread of fanatical behaviour is considered are [1, 2, 3, 4]. Recently, the reference [5] about mathematical methods in counterterrorism has been published.

*e-mail: francisco.santonja@uv.es
†e-mail: actarazona@asic.upv.es
‡e-mail: rjvillan@imm.upv.es
§e-mail: jvillanueva@cesfelipesegundo.com

In [1] the dynamics of the spread of extreme behaviours is studied as a type of communicable social contact process (recruitment) that may be under the influence of friends, mates, environment, fear, menaces, terrorism, propaganda, law enforcement, etc. Thus, a mathematical model is built and its equilibrium points, thresholds and bifurcations are studied. One of the most interesting conclusions obtained in [1], is that the eradication of these groups may be a long time task and before they begin to decay, can still experience grow and expand in finite time.

In [2] and [3] the authors consider a discretization of the continuous model in [1] over some classes of scale-free networks obtaining similar conclusions exploiting features easily applicable to networks but not in the continuum, as the range of interactions between people.

Other interesting reference is [4] where it is developed a quasi-predator-prey model to be applied to Colombia scenario, with insurgent groups that kidnap, traffic with drugs, etc.

In [6] we propose a type-epidemiological mathematical model applied to the situation in the Basque Country [7], a northern Spanish region where the armed Marxist-Leninist nationalist organization ETA (Basque for "Basque Homeland and Freedom") [8] acts for getting Basque independence. ETA was founded in 1959 and evolved from a group advocating traditional cultural ways to an armed group using the violence (murders, kidnapping, vandalism, etc.) to demand independence. From [6] we recall some statistical studies that will permit us to build the network mathematical model. These results are described in Section 2.

The parameters of the network mathematical model are unknown and our goal is to find the best parameters in such a way that the network model fits real data. To do that an intensive computing procedure is designed. Note that fitting network models is a difficult issue that depends strongly on available computational resources.

Usually, the model sensitivity analysis is done locally, and only for a parameter and the others fixed. However, these methods do not accurately assess uncertainty and sensitivity in the system. Thus, Latin Hypercube Sampling (LHS) appears as a technique that allows a global study of multi-dimensional parameter spaces in order to identify uncertainties [9, 10] and measure the variations in the model output.

This paper is organized as follows. In Section 2 we build the stochastic network model recalling some relevant results obtained in [6]. In Section 3 we present a method to fit the network model with the available electoral data.

The result of the fitting procedure and a simulation to show the next few years trend are also presented. Model sensitivity using the LHS technique is studied in Section 4. Finally, in Section 5, results and conclusions are discussed.

1 The stochastic network mathematical model

In [6], using as a source data results of the general elections to the Spanish Parliament in the Basque Country, we applied statistical techniques of correspondence analysis to justify the ideological division of the population into the following four subpopulations:

1. Subpopulation E: people non-nationalist, against independence and the violence of ETA.

2. Subpopulation N: people nationalist agreeing with the idea of independence but disagreeing the use of the violence to get this goal.

3. Subpopulation V: people agreeing with the use of the violence to achieve the independence of the Basque Country.

4. Subpopulation A: the rest of the people. This population includes people who do not share any of the above ideas or people who abstain.

Grouping votes, in Table 1 we can see the percentage of votes of each subpopulation in each election. Taking into account that, in Spain, only people older than 18 can vote and supposing that children and teenagers have the same ideology as their parents, let us assume that data in Table 1 is an ideological landscape of the whole population in the Basque Country.

For the sake of simplicity we assume that the total population $T = E + N + V + A$ is constant in the Basque Country, i.e., nobody borns, nobody dies and there are not migration movements. Now, let us determine the transition terms among subpopulations. In [6] we carried out the computation of the partial correlation coefficients. These coefficients study the linear relation between two variables under the influence of a third variable [11]. Then, we took data from Table 1 corresponding to elections from March 1st 1979 to March 3rd 1996 (where the major part of time the Socialist Party (PSOE) was in the Government of Spain and it can be supposed the same

Table 1: Pecentage of votes per subpopulations E, N, V and A in each election.

Election date	E	N	V	A
Jun 15th, 1977	0.435392	0.266825	0.0316636	0.26612
Mar 1st, 1979	0.278466	0.233303	0.0967287	0.391502
Oct 28th, 1982	0.347978	0.306366	0.114331	0.231325
Jun 22nd, 1986	0.294959	0.245271	0.117587	0.342183
Dec 17th, 1989	0.246631	0.283516	0.111871	0.357982
Jun 6th, 1993	0.330461	0.234575	0.100969	0.333994
Mar 3rd, 1996	0.364871	0.236203	0.0872077	0.311718
Mar 12th, 2000	0.364974	0.239676	0.	0.39535
Mar 14th, 2004	0.382756	0.299592	0.	0.317652

policy against terrorism) and we found that there is a linear inverse relation between E and A under V, that is, under V an increasing of subpopulation E implies a decreasing of subpopulation A and vice versa. Moreover, the linear correlation coefficient between E and A without the presence of V is not significant. Therefore the transition between E and A is modelled by the non-linear term $\beta_1 EV$, where the transmission rate $\beta_1 > 0$ indicates that the transition is due to the pressure of violent acts and $\beta_1 < 0$ indicates a law strict enforcement.

Analogously, a similar situation occurs between subpopulations N and A under the influence of E, and A and V under the pressure of V. Then, the transition between subpopulations N and A and A and V are modelled by $\beta_2 NE$ and $\beta_3 AV$, respectively. Note that the model parameters β_1, β_2 and β_3 are unknown.

Under the above assumptions, the diagram of the compartmental model is depicted in Figure 1.

Now, let us discuss the implementation of a social network model for the ideological evolution of the people in the Basque Country under the pressure of ETA. We consider that every person occupies a node of a network of relations among individuals. These individuals could be in any of the four states: E, N, V or A. Our approach relies upon the complete graph, i. e., every person is potentially connected with any other person, and the graph starts with T people (constant along the time) at $t_0 =$ March 1st, 1979. 27.84% of the nodes have the state E, 23.33% of nodes have the state N, 9.67% of nodes have the state V and 39.15% have the state A. Time steps

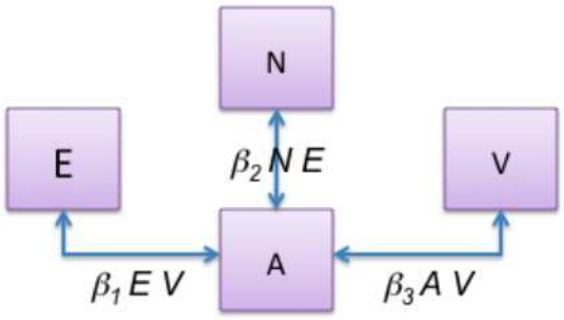


Figure 1: Flow diagram of the compartmental network model for the evolution of ideological subpopulations in the Basque Country. The boxes represent the subpopulations and the arrows represent the transitions between the subpopulations. The signs of β_1, β_2 and β_3 determine the sense of the arrows.

are set to one month. Thus, evolution rules are as follows:

- The transition from E to A under the presence of V is simulated by a mean-field procedure. The probability of transition at a given time step is:

$$\begin{aligned} &\text{if } \beta_1 > 0, \quad P(E \to A) = 1 - (1 - \beta_1)^V, \\ &\text{if } \beta_1 < 0, \quad P(A \to E) = 1 - (1 + \beta_1)^V, \end{aligned}$$

  where β_1 is the transition rate between subpopulations E and A.

- Analogously, the transitions from N to A under the presence of E and from A to V under the presence of V are given by

$$\begin{aligned} &\text{if } \beta_2 > 0, \quad P(N \to A) = 1 - (1 - \beta_2)^E, \\ &\text{if } \beta_2 < 0, \quad P(A \to N) = 1 - (1 + \beta_2)^E, \end{aligned}$$

$$\begin{aligned} &\text{if } \beta_3 > 0, \quad P(A \to V) = 1 - (1 - \beta_3)^V, \\ &\text{if } \beta_3 < 0, \quad P(V \to A) = 1 - (1 + \beta_3)^V. \end{aligned}$$

The mean-field approach has been successfully applied in other network models [12] and it yields good results in comparison with the correct, but

very computational intensive, procedure of visiting every pair of nodes with different states to determine the spread of each ideological group at the next time step. The required condition to this simplifying procedure to be valid is that the transmission rate $\beta << 1$.

The transition processes have been simulated by a constant probability of transition from a state to the next independent of the time that the individual has remained in the initial state. This is equivalent to the standard exponential distribution of remaining times which plays a main role in the traditional classical continuous mathematical epidemics models [13].

2 Fitting the stochastic network model with data

Once defined the network model and the evolution rules, for a given β_1, β_2 and β_3, we run the model (carrying out a realization) obtaining, for each time step $0, 1, 2, \ldots, n$, the number of people in subpopulations E, N, V and A. Due to the randomness of the network evolution and, in order to avoid possible extreme results, several realizations should be done, the number of individuals in each subpopulation are computed for each realization and the mean of the realizations are calculated. In our case, after some empirical tests, we decided to take $T = 10000$ individuals and to carry out 4 realizations.

Now, our objective is to find β_1, β_2, β_3 such that the network model fits electoral data in Table 1. To do that, let us take β_1, β_2 and β_3 from -0.25 to 0.25 with steps 0.02 in order to generate 17576 3-tuples $(\beta_1, \beta_2, \beta_3)$. For each 3-tuple we compute the mean of the 4 realizations, as we described above, for time steps $0, 1, 2, \ldots, 204$, i.e., from March 1st, 1979 to March 3rd, 1996. Thus, the best fitting parameters $(\beta_1^*, \beta_2^*, \beta_3^*)$ satisfy that the electoral data at time steps 43 (Oct 28th, 1982), 84 (Jun 22nd, 1986), 129, (Dec 17th, 1989), 171 (Jun 6th, 1993) and 204 (Mar 3rd, 1996) in Table 1, multiplied by $T = 10000$ are the closest (in least square sense) to the mean of the four model realizations carried out for parameters $(\beta_1^*, \beta_2^*, \beta_3^*)$. Thus, the best fitting has been achieved with

$$(\beta_1^*, \beta_2^*, \beta_3^*) = (-0.15, -0.03, 0.19). \tag{1}$$

In Figure 2 it is depicted the results of the fitting for each subpopulation

and the trend for the next few years until December 2015. The model predicts an increasing in subpopulations E (until 37.19%), N (until 27.01%) and V (until 14.2%) at expense of A (until 21.6%).

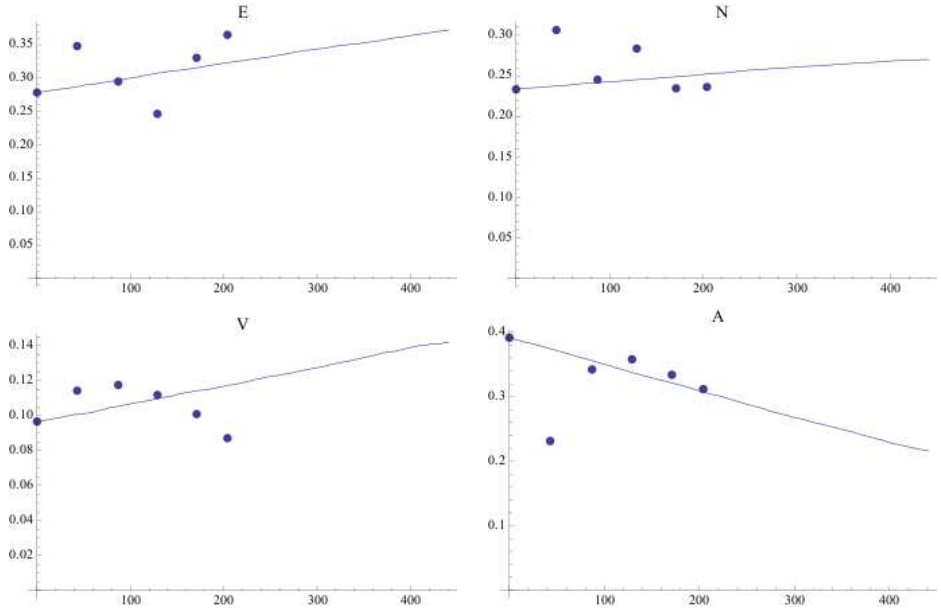


Figure 2: Best fitting for the network model for a total population of $T = 10000$ individuals. Points are the electoral data of Table 1 and the lines the number of individuals in each subpopulation for each time step. Also, in the graphs appear the prediction trends of the evolution of the four subpopulations until December 2015. Note the increasing of subpopulations E, N and V at subpopulation A expense.

3 Sensitivity analysis

We performed several simulations varying the parameters β_1, β_2, β_3 of the model around the fitted values $(\beta_1^*, \beta_2^*, \beta_3^*) = (-0.15, -0.03, 0.19)$ in order to find out what the influence of the changes on the final solution is, and therefore, the effect on the considered subpopulations. We carry out these variations (sensitivity analysis) to analyze the political strategies of violence (pressure of V) in opposition to the law enforcement (pressure of E). Take into account that a total pressure of violence implies $\beta_1 > 0$, $\beta_2 < 0$, $\beta_3 > 0$,

and a total law enforcement $\beta_1 < 0$, $\beta_2 > 0$, $\beta_3 < 0$. But, as in this case, it may have mixed combinations where the pressure of V or E is enough for certain transitions but not for others.

In order to perform the sensitivity analysis, let us use the technique called Latin Hypercube Sampling (LHS) to vary parameter values in the proposed model. Latin Hypercube Sampling, a type of stratified Monte Carlo sampling, is a sophisticated and efficient method for achieving equitable sampling of all input parameters simultaneously [14, 15].

Each parameter for a model can be defined as having an appropriate probability density function associated with it. It is usual to use the uniform distribution centred at deterministic parameter estimators in absence of data to inform on the distribution for a given parameter [10, 16]. Then, the model can be simulated by sampling a single value from each parameter distribution. Many samples should be taken and many simulations should be run, producing variable output values.

To vary β_1, β_2 and β_3, we assume that all of them follow a uniform probability distribution with support on the intervals $[-0.225, -0.075]$, $[-0.045, -0.015]$ and $[0.095, 0.285]$ respectively. The intervals are chosen assuming that the value of the parameter may have a perturbation not greater than 50%.

LHS was used to generate 5000 different values of the parameters β_1, β_2 and β_3. Then we used these samples to run 5000 evaluations of the model. The results of these evaluations allow us to determinate the 90% confidence interval (90% c.i.), the mean value of the 5000 realizations (Mean realz.) and compare them to the model predictions (Model pred.). The results obtained for December of year 2015 after the variation of the parameters are (the values in the following tables are in percentages):

1. Variation in the transition between E and A at December of 2015, i.e., $\beta_1 \in [-0.225, -0.075]$. $\beta_2 = -0.03$ and $\beta_3 = 0.19$ remain constant.

	E	N	V	A
90% c.i.	$[33.08, 40.41]$	$[26.96, 27.69]$	$[13.24, 14.44]$	$[18.99, 25.15]$
Mean realz.	36.85	27.33	13.82	21.99
Model pred.	37.19	27.01	14.20	21.58

2. Variation in the transition between N and A at December of 2015, i.e., $\beta_2 \in [-0.045, -0.015]$. $\beta_1 = -0.15$ and $\beta_3 = 0.19$ remain constant.

	E	N	V	A
90% c.i.	$[36.36, 37.50]$	$[25.55, 29.10]$	$[13.40, 14.23]$	$[20.67, 23.17]$
Mean realz.	36.94	27.33	13.81	21.91
Model pred.	37.19	27.01	14.20	21.58

3. Variation in the transition between V and A at December of 2015, i.e., $\beta_3 \in [0.095, 0.285]$. $\beta_1 = -0.15$ and $\beta_2 = -0.03$ remain constant.

	E	N	V	A
90% c.i.	$[36.44, 37.40]$	$[26.91, 27.73]$	$[11.68, 16.43]$	$[19.29, 24.14]$
Mean realz.	36.92	27.32	13.91	21.83
Model pred.	37.19	27.01	14.20	21.58

4. Variation of β_1, β_2 and β_3 (all together), i.e., $\beta_1 \in [-0.225, -0.075]$, $\beta_2 \in [-0.045, -0.015]$ and $\beta_3 \in [0.095, 0.285]$,

	E	N	V	A
90% c.i.	$[33.10, 40.48]$	$[25.53, 29.11]$	$[11.66, 16.62]$	$[17.72, 26.29]$
Mean realz.	36.85	27.32	13.92	21.89
Model pred.	37.19	27.01	14.20	21.58

The mean of each of the 5000 realizations are very similar to the model predictions in all cases. On the other hand, large variations in confidence intervals imply more margin to influence on a subpopulation and the prediction is more uncertain and unexpected. In this case is when apropriate policies may lead to get the desired ideological objective. Under the point of view of this reasoning, we have that,

1. The variation of β_1 produces changes of 7.32% and 6.16% on subpopulations E and A, and very small variations in the other subpopulations.

2. The variation of β_2 produces changes of 3.55% and 2.49% on subpopulations N and A. The variations due to changes on β_2 are smaller than the ones of β_1.

3. Variations on β_3 cause noticeable changes of 4.74% and 4.85% on V and A, respectively.

4. The variations in all the parameters show changes of 7.38%, 3.58%, 4.95% and 8.57% in subpopulations E, N, V and A.

Therefore, the subpopulations more affected by the changes in the parameters are A and E. There is a noticeable effect on V when β_3 moves and N is the subpopulation less sensitive to parameter changes.

4 Conclusions

In this paper, we propose a type-epidemiological social network mathematical model to study the ideological evolution of a society where a group pressures to get political goals. This model considers the ideology as a communicable proccess that is spread by social transmission. The population is divided into four subpopulations of interest, E, N, V and A, where certain parameters determine the transitions between these subpopulations.

We also present a method to estimate the unknown model parameters with intensive computation and without using any equivalent or similar continuous model.

Model sensitivity has been studied and the effect of the changes in the parameters on the model output for December of 2015 tells us that the most affected subpopulations are A and E, in this order, and N hardly suffers the parameter variations.

References

[1] C. Castillo-Chávez, B. Song, Models for the transmission dynamics of fanatic behaviors, in Bioterrorism: Mathematical Modeling Applications in Homeland Security, SIAM Frontiers in Applied Mathematics, ed.: H.T. Banks and C. Castillo-Chávez, SIAM, Philadelphia, 28 (2003) 155 – 172.

[2] D. Stauffer, M. Sahimi, Discrete simulation of the dynamics of spread of extreme opinions in a society, Physica A: Statistical Mechanics and its Applications 364 (Mayo 15, 2006): 537-543, doi:10.1016/j.physa.2005.08.040.

[3] D. Stauffer, M. Sahimi, Can a few fanatics influence the opinion of a large segment of a society?, The European Physical Journal B Condensed Matter and Complex Systems 57, no. 2 (Mayo 1, 2007): 147-152, doi:10.1140/epjb/e2007-00106-7-

[4] J.A. Adam, J.A. Sokolowski, C.M. Banks, A two-population insurgency in Colombia: Quasi-predator-prey models-A trend towards simplicity, Mathematical and Computer Modeling 49, no. 5-6 (Marzo 2009): 1115-1126, doi:10.106/j.mcm.2008.03.017.

[5] Mathematical Methods in Counterterrorism, Memon, N.; Farley, J.D.; Hicks, D.L.; Rosenorn, T. (Eds.) 2009, ISBN: 978-3-211-09441-9

[6] Francisco J. Santonja, Ana C. Tarazona, y Rafael J. Villanueva, A mathematical model of the pressure of an extreme ideology on a society, Computers & Mathematics with Applications 56, no. 3 (August 2008): 836-846, doi:10.1016/j.camwa.2008.01.001.

[7] *http://en.wikipedia.org/wiki/Basque_Country_(autonomous_community)*

[8] *http://en.wikipedia.org/wiki/ETA*

[9] McKay, M., Meyer, M., 2000. Critique of and limitations on the use of expert judgements in accident consequence uncertainty analysis. Radiat. Prot. Dosim. 90 (3), 325-330.

[10] Simeone Marino et al., A methodology for performing global uncertainty and sensitivity analysis is system biology, Journal of Theoretical Biology 254, no. 1 (Septiembre 7, 2008): 178-196, doi:10.1016/j.jtbi.2008.04.011.

[11] M.H. DeGroot, Probability and statistics, Addison-Wesley, 1986.

[12] L. Acedo, A second-order phase transition in the complete graph stochastic epidemic model, Physica A 370 (2006) 613.

[13] F. Brauer and C. Castillo-Chavez, Mathematical Models in Population Biology and Epidemiology, Springer Verlag, 2001.

[14] Blower, S.M. and Dowlatabadi, H. Sensitivity and Uncertainty Analysis of Complex Models of Disease Transmission: and HIV Model, as an Example. International Statistical Review, 62(2), 229-243, 1994.

[15] Olsson, A., Sandberg, G. and Dahlblom, O., On Latin hypercube sampling for structural reliability analysis. Structural Safety, 25, 47-68, 2003.

[16] Hoare, A., Regan, D.G. and Wilson, D.P. Sampling and sensitivity analyses tools (SaSAT) for computational modelling. Theoretical Biology and Medical Modelling, 5:4, 2008.