



ELSEVIER

Available online at www.sciencedirect.com



Expert Systems with Applications xxx (2006) xxx–xxx

Expert Systems
with Applications

www.elsevier.com/locate/eswa

Predicting service request in support centers based on nonlinear dynamics, ARMA modeling and neural networks [☆]

Emili Balaguer ^a, Alberto Palomares ^a, Emilio Soria ^{b,*}, Jose David Martín ^b

^a Tissat S.A., R&D Department, Av. Leonardo Da Vinci, 5. 46980 Paterna, Valencia, Spain

^b G.P.D.S, Digital Signal Processing Group, Electronic Engineering Department, Escuela Técnica Superior de Ingeniería, Universitat de València, C/ Dr. Moliner 50, 46100 Burjassot, Valencia, Spain

9 Abstract

10 In this paper, we present the use of different mathematical models to forecast service requests in support centers (SCs). A successful
11 prediction of service request can help in the efficient management of both human and technological resources that are used to solve these
12 eventualities. A nonlinear analysis of the time series indicates the convenience of nonlinear modeling. Neural models based on the time
13 delay neural network (TDNN) are benchmarked with classical models, such as auto-regressive moving average (ARMA) models. Models
14 achieved high values for the correlation coefficient between the desired signal and that predicted by the models (values between 0.88 and
15 0.97 were obtained in the out-of-sample set). Results show the suitability of these approaches for the management of SCs[†].
16 © 2006 Elsevier Ltd. All rights reserved.

17 *Keywords:* Time series analysis; Neural networks; Call centers

19 1. Introduction

20 Support centers (SCs) usually deal with all the requests
21 reported by either external customers/citizens or internal
22 users of an organization/company. Although telephone
23 has been the traditional way to provide support in the call
24 centers (CC^s), nowadays the Internet provides new mech-
25 anisms of communication that overcome some of the limi-
26 tations associated with telephone support, and it is widely
27 used by SCs. The formal contract between the service pro-
28 vider and the service recipient is known as the service level
29 agreement (SLA). SLA must define the quality measures or
30 service level measurements (SLMs) that are used to evalu-
31 ate the quality of the support service. SLA tends to contain
32 clauses that economically incentive/penalize the SC
33 depending on its degree of fulfillment according to SLMs.
34 However, SLMs is one of the most crucial and complex

tasks in the management of SCs and is out of the scope 35
of this study. 36

Different system performance and business parameters 37
can be considered for SLMs: cumulative time-based 38
parameters, service availability, number of affected users, 39
metrics based on a particular business process, etc. These 40
measurements must be automatically collected, maintained 41
and analyzed in order to manage the service support pro- 42
cess. HelpDesk consists of several software applications 43
that allow to record, track and report all the information 44
involved in a SC, including the SLAs and SLMs manage- 45
ment². In general, the information than can be obtained 46
from a HelpDesk is considerably wide: general informa- 47
tion, purchases, complaints, etc. Operators usually answer 48
many calls of different nature: sales, relational marketing, 49
customer services, citizen services, technical support, and 50
in general, any specialized activity related to business or 51
public administration. Moreover, some contracts require 52
a minimum fulfillment degree, and apply penalizations if 53
this degree is not achieved. 54

There are several aspects that should be covered by a 55
methodology which tries to help in the management of a 56

[☆] This work has been partially supported by the Spanish Ministry of Industry, Tourism and Commerce, with the project FIT-340001-2004-11.

* Corresponding author.

E-mail address: Emilio.soria@uv.es (E. Soria).

certain SC. First, SCs are especially interested in managing the forthcoming events as good as possible. Second, operators are interested in searching optimal and fast solutions for the events they have to solve. Third, and as a general rule, it is preferable that problems are solved by operators at the first levels of the system hierarchy, otherwise, the participation of qualified technical personnel is required.

In this work, we analyze different eventualities managed by an official SC, called CETESI.¹ CETESI is part of a global project, which includes many other subprojects (Infoville, Infocentre, ...), and whose goal is to achieve a complete integration of Valencian Region within the Information Society.

Since there is much information that should be taken into account for improving the management of CETESI, the use of data mining (DM) techniques seems to be an appropriate choice to tackle the problem. In particular, time series analysis and prediction has been successfully used in a wide range of knowledge fields (Box & Jenkins, 1976; Makridakis, Wheelwright, & Hyndman, 1998; Brockwell & Davis, 2002). The aim of this work is to predict the number of forthcoming service requests as well as the time when they will occur in order to optimize the resources used to solve these events. Different classical models and advanced neural network models are tested in order to obtain a robust and reliable method to predict future events.

The remainder of the paper is outlined as follows. Section 2 presents the methods used in this work, Section 3 shows the obtained results, ending up the paper with the conclusions that can be extracted from this work in Section 4.

2. Methods

In this work, three different techniques are used for time series prediction:

1. Linear models: Auto-regressive moving average (ARMA) models (Box & Jenkins, 1976; Makridakis et al., 1998).
2. Time processing neural models, and in particular, the time delay neural network (TDNN) (Haykin, 1999).
3. Non-linear time series analysis (Kantz & Schreiber, 1999).

2.1. ARMA modeling

ARMA models of order (n,m) can be viewed as linear filters from the point of view of digital signal processing.

The time structure of these filters is the following models (Box & Jenkins, 1976; Makridakis et al., 1998):

$$y(k) = a_1 \cdot y(k-1) + \dots + a_n \cdot y(k-n) + e(k) + b_1 \cdot e(k-1) + \dots + b_{m-1} \cdot e(k-m+1) + C \quad (1)$$

where $y(k)$ is the variable to be predicted using previous values of the time series, $e(i)$ is a sequence of i.i.d.² terms which have zero mean, and C is a constant. As it can be observed from (1), ARMA modelling is made up of two parts:

- Those terms involving coefficients $\{a_i\}(i = 1, \dots, n)$ carry out a linear relationship between the value predicted by the model at time k and the past values of the time series; this part is known as auto-regressive (AR).
- Those terms involving coefficients b_i carry out a linear relationship between the value predicted by the model at time k and a Gaussian distribution of i.i.d. samples. This part is known as moving average (MA), and can be considered as the autoregressive structure of the residues (errors committed in the time series prediction).

Several steps must be followed in order to obtain the coefficients that appear in (1) (Makridakis et al., 1998):

- (a) *Series stationarity*. ARMA modeling can only be used correctly if the time series is *wide sense stationary* (WSS). Exponential decreasing of autocorrelation function across lags is a strong indication of time series stationarity (Makridakis et al., 1998). Therefore, the first step is based on transformations of the time series in order to achieve the desired effect in the autocorrelation. Usually, a differentiation if the time series is sufficient. If the differentiation is carried out between consecutive values, it is called *simple differentiation*. Otherwise, if the differentiation is carried out between periods of the time series, it is known as *seasonal differentiation (or integration)*.
- (b) *Identification*. The order of both the AR and the MA part must be estimated. Autocorrelation and partial autocorrelation computation is used to obtain these orders.
- (c) *Optimization*. The model coefficients are calculated using a training group by means of the autocorrelation computed in the previous step.
- (d) *Selection*. The most suitable models are chosen. These models should offer adequate predictions. In order to evaluate models, data are split into two groups: training and validation. The training group is used to build the model (as it was mentioned in the previous item), and the validation group to evaluate it. Moreover, the model should have a few number of parameters (as few as possible). Akaike information criterion (AIC) based on information theory is used

¹ CETESI is the acronym of “Centro de Telecomunicaciones y Sistemas de Información de la Generalitat Valenciana (Center of Telecommunications and Information Services of Generalitat Valenciana)”. Generalitat Valenciana is the name of the autonomous government of the Valencian Region (Spain).

² Independent and identically distributed.

155 to achieve a trade-off between both requirements (an
156 adequate prediction and a few number of parameters)
157 (Weigend & Gershenfeld, 1996).

158 (e) *Residues*. Autocorrelation and partial autocorrelation
159 of the residues are calculated to test whether they are
160 statistically relevant. Ideally, residues should be i.i.d.
161

162 ARMA processing has shown to be the most effective
163 tool to model a wide range of time series (Makridakis
164 et al., 1998). However, these models do not work properly
165 when there are elements of the time series that show a non-
166 linear behavior. In this case, other models, such as time
167 processing neural networks, must be applied.

168 2.2. Neural networks

169 A neural network is a mathematical model formed by
170 elementary processing units, the so-called neurons. Neu-
171 rons are defined by the following parts, as shown in
172 Fig. 1 (Haykin, 1999):

- 173 • Connections known as synaptic weights, which deter-
174 mine the neuron behavior. These connections may be
175 either excitatory or inhibitory.
- 176 • A sum function which carries out a weighted sum of the
177 neuron inputs.
- 178 • A nonlinear function, known as activation function.
- 179 • A bias, λ .

181 Neurons can be arranged in different architectures. The
182 most widely used is the so-called multilayer perceptron
183 (MLP), shown in Fig. 2. In an MLP, neurons are arranged

184 in fully-connected layers, and the values of the synaptic
185 weights are usually obtained by means of the backpropaga-
186 tion algorithm (Haykin, 1999). There are numerous modi-
187 fications of the MLP that are used in time series prediction.
188 In this work, we used the TDNN, a neural network that
189 includes a time memory in the input layer by means of time
190 delays (Haykin, 1999). There are also other approaches,
191 such as FIR, IIR, Elman and Jordan neural networks,
192 but they are not considered in this study due to the excel-
193 lent results provided by the simpler TDNN. The evaluation
194 of the model is carried out using a data set different from
195 that used to build the model. Likewise, the number of
196 parameters of the model should be low in order to guaran-
197 tee generalized results with unknown data sets (Vapnik,
198 1999). Therefore, neural networks used in this study are
199 chosen by analyzing both the results achieved and the num-
200 ber of parameters (weights) involved.

201 2.3. Nonlinear dynamics

202 Dynamic systems analyze temporal evolution of either
203 systems of continuous differential equations or discrete
204 maps defined in a phase space, in which each axis repre-
205 sents a dynamic variable of the system (Wiggins, 2003).
206 Last years, this analysis has been applied to stochastic
207 experimental data. In this work, we try to determine which
208 kind of time series we are dealing with. There are several
209 possibilities:

- (a) *Multiperiodical*: values are repeated following a cer-
210 tain combination of frequencies. In this case, it is
211 not necessary to carry out an analysis based on non-
212

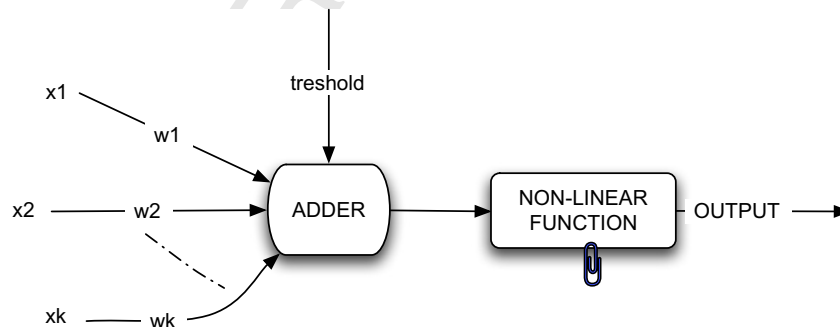


Fig. 1. Schematic of a neuron.

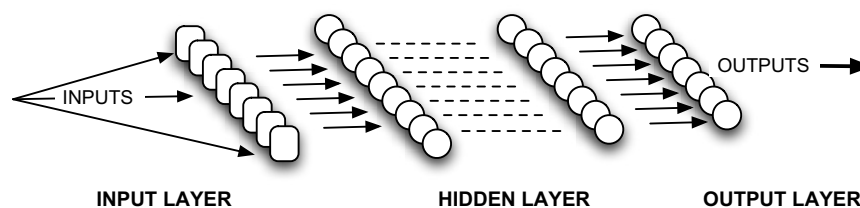


Fig. 2. Schematic of a Multilayer Perceptron.

linear dynamic systems. Either Fourier's analysis or correlation analysis is sufficient to describe time series.

(b) *Chaotic*: values do not follow a periodical evolution. Fourier's analysis and correlation analysis do not show any periodically repeated behavior. The evolution of the system can be rather fast, and the phase space shows a strange attractor (Kantz & Schreiber, 1999). Time series of this kind are strongly nonlinear and Fourier theory does not provide the necessary tools to analyze the time series.

The main problem to overcome is the likely presence of noise in the time series. This is because this noise may cause both Fourier's and autocorrelation analyses to show a wrong information since these mathematical tools cannot distinguish whether the nonperiodical part of the time series is caused by unpredictable noise or by the chaotic behavior of the time series. Nonlinear time series analysis can distinguish between the two possibilities. If there is noise but nonlinearities do not appear, ARMA models provide the best possible modeling, since the nonlinear part of the time series is purely stochastic, and therefore, unpredictable. However, if nonlinearities do appear, it is possible to improve ARMA predictions by using other models, such as neural networks.

In order to analyze system dynamics, the first step to do consists of embedding the time series in a suitable phase space. The criterion used in this work is based on considering that each axis of the phase space represents a delay of the time series. This way, the autocorrelation between axes shows its lowest value (Kantz & Schreiber, 1999). A two dimensional embedding is proposed; once the series attractor is obtained, the following parameters should be determined:

Correlation dimension (D_2). It is the attractor dimension that is easiest to compute. A fractal dimension shows that a "strange" attractor, and leads to a nonlinear behavior. The value of D_2 gives an idea about the suitable dimension of the phase space necessary to embed the series (typically, $\geq 2 \cdot D_2 + 1$) (Takens, 1981), although this criterion assumes the series to be deterministic and can be "relaxed" empirically. In order to calculate this parameter, and also the other chaos indicators, it is necessary to calculate the correlation sum (Kantz & Schreiber, 1999):

$$C(\varepsilon) = \frac{2}{(N - n_{\text{sec}}) \cdot (N - n_{\text{sec}} - 1)} \cdot \sum_{i=1}^N \sum_{j=i+n_{\text{sec}}}^N \theta(\varepsilon - \|\vec{X}_i - \vec{X}_j\|) \quad (2)$$

where N is the quantity of data, ε is the radius of the reticulum in which the phase space is divided, and \vec{X}_i is the representative vector of a point in the phase space. For instance, in the two-dimensional case, $\vec{X}_i = [x(i)x(i-1)]$, where $x(i)$ is the value of the time series

at instant i . Function θ is the so-called Heaviside function (step function which equals 1 for positive arguments, and 0 for negative arguments). Factor n_{sec} is the time chosen to avoid an excessive autocorrelation among the different vectors which made up the attractor, because it might distort the value of D_2 . This sum gives information about the ratio of points of the phase space that are closest to ε , i.e., it gives information about the population density of the space (Kantz & Schreiber, 1999).

Nonlinearity test. A non parametric statistical test must be computed to check whether the origin of the nonlinearity of the time series is (Kantz & Schreiber, 1999). The null hypothesis considers that the time series is produced by an ARMA process with Gaussian inputs. An infinite number of artificial data sets with the same magnitude spectrum but different phase spectrum can be produced (Kantz & Schreiber, 1999). These "surrogate" data sets are actually time series with the same length and statistical moments that the original one, but produced by a linear ARMA process with normal inputs. Then, the characteristic parameter (D_2) of the nonlinear dynamics of the series is calculated. This nonlinear test consists of determining whether all the correlation dimensions of the artificial series are lower (or higher) than D_2 . If they are lower, it can be rejected that the series was not generated by a linear process. In this work, we produced 20 artificial data sets, having a 95% bilateral confidence level (Kantz & Schreiber, 1999).

3. Results

In this section, we show the results achieved in the analysis of the events managed by CETESI during years 2003 and 2004. Three different kinds of events are analyzed, namely:

- Calls received at the CC. They represent 48.9% of the total registered events.
- Automatically generated calls (AGC): 40.6%.
- Requests from web forms (WF): 7%.

Events are usually related to projects focused on critic public services for citizens and public administration. The most interesting projects are shown in Table 1. They are also the oldest projects with many records and well-known characteristics.

According to CETESI criteria, a useful model must provide robust 6-h ahead predictions. Therefore, this was the sample period used by our prediction models. Statistical analyses were carried out using SPSS®, whereas models (ARMA, TDNN and nonlinear analysis) are implemented in a Matlab® environment.

The methodology presented below was followed in all the projects. Therefore, it is shown only the detailed analysis related to the most relevant project in authors' opinion

Table 1

Most important projects analyzed in this work

Project name	Percentage of the total	Majority kind of events
RED (<i>Network</i>)	18.4	AGC (64.2%)
INFOCENTRE (<i>Infocenter</i>)	16.1	AGC (52.0%)
LEGAL (<i>Legal</i>)	11.8	CC (79.1%)
TARJETA SANITARIA (<i>Health Care System Card</i>)	11.1	CC (87.4%)
MAIL (<i>E-Mail</i>)	6.7	CC (96.1%)
112 (<i>Emergency phone number</i>)	6.5	AGC (65.0%)
INFOVILLE XXI (<i>Citizen web portal</i>)	4.4	AGC (91.5%)

The English translation or description of the project names appears in italics. It is shown which percentage of the total number of events are related to each project (2nd column), and which is the majority kind of events in each case (3rd column).

319 (Health Care System Card). The same methodology was
 320 also applied to the other projects with similar results and
 321 conclusions.

322 3.1. Descriptive analysis

323 The most important statistical parameters related to this
 324 project were obtained, as shown in Table 2. A first conclu-

sion is that the number of outliers is rather scarce. More-
 325 over, as the values of skewness and kurtosis indicate a
 326 behavior close to normality, we can consider as outliers
 327 those values beyond 3.5 standard deviations from the
 328 mean, with a 99.9% of confidence. Using this criterion,
 329 we found two outliers, which has been replaced by the max-
 330 imum value of the remaining series (93 events/6 h).
 331

Fig. 3 shows the weekly values of the mean and standard
 332 deviation in years 2003 and 2004. The homogeneity than
 333 can be observed is considered as an empirical indicator of
 334 the time series stationarity (Makridakis et al., 1998).
 335

ANOVA shows that there are not significant differences
 336 in the means of the series ($F = 0.39$ with significance equal
 337 to 1.000). However, standard deviations did vary signifi-
 338 cantly, as it is shown by the Levene Test (3.0 in 2003 and
 339 5.1 in 2004 with significance equal to 0.000). A decrease
 340 in the standard deviation is observed around week 34,
 341 due to summer holidays. Therefore, modeling was carried
 342 out in two ranges: before and after summer holidays.
 343

Summarizing, there are four modeling periods (two per
 344 year) in which the time series showed a quasi-steady behav-
 345 ior with regard to mean and standard deviation. Moreover,
 346 CETESI ruled out weeks 1, 48 and 49 of year 2003, and
 347 also weeks 1 and 6 of year 2004 (a 4.7% of the total quan-
 348

Table 2

Statistical parameters of the events related to the project "Health Care System Card"

Year	Maximum	Mean	Standard deviation	Percentile 5	Percentile 50	Percentile 95	Skewness	Kurtosis
2003	93	21	25	0	13	56	1.04	-0.18
2004	93	24	21	0	9	65	1.04	-0.18

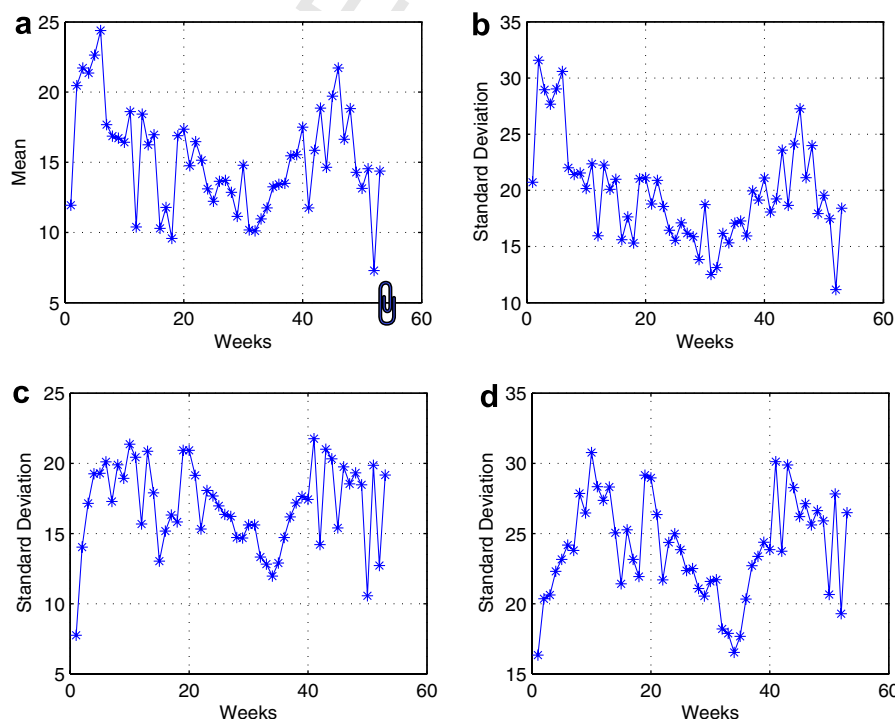


Fig. 3. Representation of the weekly values of mean and standard deviation for the project "Health Care System Card". (a): mean-year 2003; (b): standard deviation-year 2003; (c): mean-year 2004; (d): Standard deviation-year 2004.

349 tity of data) because there were some problems in the man-
 350 agement of events.

351 A daily analysis of statistical parameters showed no sig-
 352 nificant differences. Moreover, since no events were
 353 recorded on weekends, weekends were not taken into
 354 account in the analysis.

355 Once the series stationarity was analyzed, the likely peri-
 356 odical structure was studied. Daily periodicity was ana-
 357 lyzed by means of both total and partial autocorrelations
 358 (Fig. 4). The daily period (each four samples) was much
 359 more relevant than the weekly period (each 20 samples).
 360 A four-order differentiation was proposed in order to elim-
 361 inate this periodicity.

362 3.2. Linear models

363 Once the time series has been differentiated, ARIMA
 364 modeling is carried out. After the daily differentiation,
 365 TAC and PAC show that it is not necessary any other dif-
 366 ferentiation, being a possible candidate either the model
 367 ARMA(0,1) or ARMA(1,0). This is because it was
 368 observed a fast decrease of TAC (and PAC in the second
 369 model) after the first delay and multiples of four delays

(Box & Jenkins, 1976). Algorithm stopping criterion is 370
 given by the sum of least-squares (if the change between 371
 consecutive iterations is lower than 0.001%). Using this 372
 threshold, all the models used in this study achieve conver- 373
 gence in 3–10 iterations. The best ARMA structure, 374
 according to AIC (Box & Jenkins, 1976; Makridakis 375
 et al., 1998) (Table 3, second column) was an ARMA(0,1). 376
 Table 3 also shows predictions obtained by the winner 377
 model. Correlation coefficients between the desired signal 378

Table 3
 Results obtained with an ARMA(0,1) model in the project “Health Care System Cardit should be pointed out that the gap between weeks 22 and 37 corresponds to summer period

Period	Stan error	Correlation coefficient (CC)	CC for the predictions (validation)
Weeks 2–22 (2003)	10.8	0.88 ± 0.01	0.91 ± 0.01
Weeks 37–48 (2003)	8.7	0.91 ± 0.01	0.78 ± 0.01
Weeks 2–22 (2004)	9.6	0.93 ± 0.01	0.97 ± 0.01
Weeks 37–48 (2004)	9.4	0.94 ± 0.01	0.81 ± 0.01

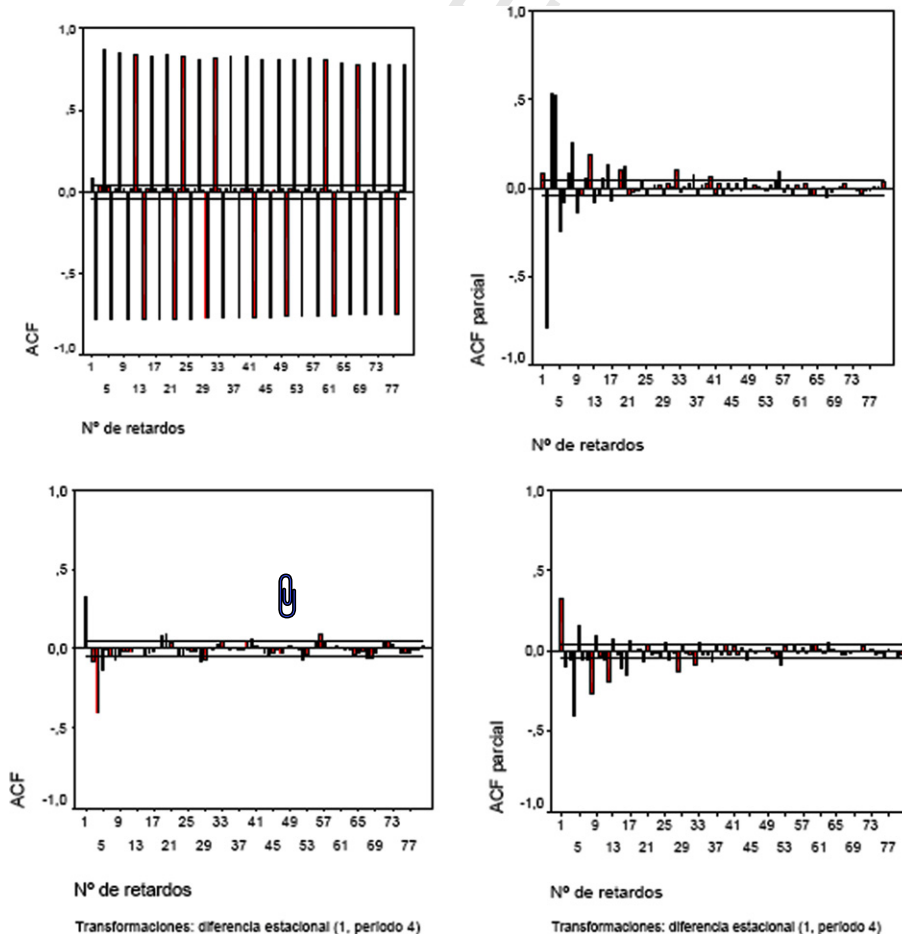


Fig. 4. Total autocorrelation (TAC) (top-left figure) and partial autocorrelation (PAC) (top-right figure) of the time series “Health Care System Card”. The same information, but after carrying out daily differentiation, is shown in bottom figures.

379 and the model output in the validation groups were very
380 appropriate in the first months of years 2003 and 2004 (val-
381 ues of 0.91 and 0.97, respectively). Validation groups had a
382 length of four weeks in both years 2003 and 2004.

383 Fig. 5 shows the autocorrelation analysis of the residues.
384 There are not clear tendencies, being the residues quite
385 close to white noise. Therefore, the identification carried
386 out by ARMA model was correct.

387 Summarizing, the predictor showed an excellent behav-
388 ior for the first weeks of the year and moderate for the last
389 weeks.

390 3.3. Neural networks

391 The results shown in previous subsections suggest the
392 existence of nonlinearities that ARMA models could not
393 manage. Therefore, a nonlinear structure, such as a TDNN
394 (Haykin, 1999), was used for series modeling. The main
395 characteristics of the TDNN used in this work were the fol-
396 lowing (Bishop, 1995):

- 397 (a) Weights were initialized following a Gaussian distri-
398 bution of zero mean and unity variance.
399 (b) The used learning algorithm was a simple variant of
400 the classical backpropagation, in which the learning
401 rate and the momentum coefficient increased
402 (decreased) depending on the error committed by
403 the network (if the error decreased, then the learning
404 rate and the momentum coefficient increased, and
405 viceversa) (Haykin, 1999). Increase/decrease factors
406 were taken equal to 1.1 and 0.6, respectively. The ini-
407 tial learning rate was varied between 0.01 and 0.1 and
408 the initial momentum coefficient between 0.01 and
409 0.2.
410 (c) Only one hidden layer formed by a number of neu-
411 rons between 1 and 4 was used. The number of delays
412 in the input layer was set to 4 as suggested by auto-
413 correlation analyses (Fig. 4).
414 (d) Cross-validation strategies were not used as stopping
415 criterion due to the short length of the time series.

Table 4

Correlation coefficients obtained by TDNN and ARMA in the out-of-
sample time series predictions (i.e., predictions during the validation
period)

Period (training)	Correlation coefficient (validation)	
	TDNN	ARMA
(Weeks 49–52; 2003) (Weeks 37–48; 2003)	0.88 ± 0.01	0.78 ± 0.01
(Weeks 49–52; 2004) (Weeks 37–48; 2004)	0.90 ± 0.01	0.81 ± 0.01

Parentheses indicate the training period while the validation period is
indicated without parentheses.

416 Instead, overtraining was avoided by selecting that
417 TDNN which improved the ARMA predictions using
418 the smallest number of neurons, in order to maximize
419 the generalization ability of such perceptron-like
420 architecture (Vapnik, 1999; Haykin, 1999).
421

422 Table 4 shows correlation coefficients between the actual
423 time series and the ARMA and TDNN out-of-sample pre-
424 dictions (i.e., using data not previously shown to the mod-
425 els). Models are obtained during the training period, and
426 then evaluated during the corresponding validation period.

3.4. Nonlinear analysis

427
428 Since TDNN did increase prediction performance, it
429 could be assumed the existence of nonlinearities in the ana-
430 lyzed time series. In order to reassert this claim, the nonlin-
431 ear behavior of the time series between weeks 37 and 52 of
432 years 2003 and 2004 was determined. First of all, time ser-
433 ies should be embedded in the phase space. Since autocor-
434 relation analyses indicated a strong single-periodic
435 behavior (daily), it was expected a low dimensional attrac-
436 tor (Kantz & Schreiber, 1999). Therefore, we used a two-
437 dimensional phase space since it was expected that this
438 dimension would be sufficient for a good embedding
439 (Takens, 1981). The two axes were the actual and the pre-
440 vious values of the series, since the first minimum of the

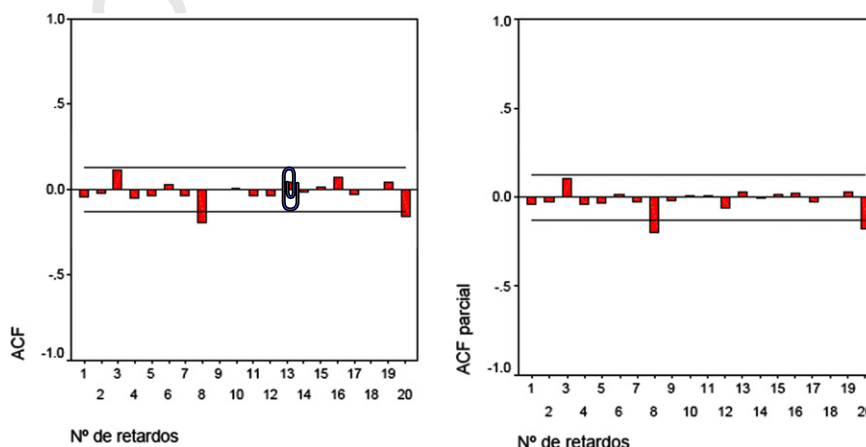


Fig. 5. TAC and PAC of the residues of the model ARIMA(0,4,1).

Table 5

Correlation coefficients obtained by TDNN and ARMA in the out-of-sample time series predictions (i.e., predictions in the validation period) in the projects ‘Legal’ and ‘Mail’

Project name	Validation period (training period)	Correlation coefficient in out-of-sample predictions (best model)
LEGAL	Weeks 23–26, 2003 (Weeks 2–22, 2003)	0.94 ± 0.01 (ARMA, TDNN)
	Weeks 49–52, 2003 (Weeks 37–48, 2003)	0.87 ± 0.01 (TDNN)
	Weeks 23–26, 2004 (Weeks 2–22, 2004)	0.93 ± 0.01 (ARMA, TDNN)
	Weeks 49–52, 2004 (Weeks 37–48, 2004)	0.86 ± 0.01 (ARMA, TDNN)
	Weeks 23–26, 2003 (Weeks 2–22, 2003)	0.88 ± 0.01 (ARMA, TDNN)
Mail	Weeks 49–52, 2003 (Weeks 37–48, 2003)	0.86 ± 0.01 (TDNN)
	Weeks 23–26, 2004 (Weeks 2–22, 2004)	0.91 ± 0.01 (ARMA, TDNN)
	Weeks 49–52, 2004 (Weeks 37–48, 2004)	0.75 ± 0.01 (ARMA, TDNN)

Parentheses in the second column indicate the training period while the validation period is indicated without parentheses. In the third column, the best model is enclosed in parentheses.

autocorrelation function corresponded to the first delay. Six-hundred and forty vectors in the phase space were used to calculate D_2 with a security time equal to 5 in order to avoid an excessive autocorrelation between vectors; the radius of the reticulum was varied between $\sigma/20$ and $\sigma/2$, being σ the standard deviation of data. As it was expected, D_2 showed a value equal to [1.5, 1.6], i.e. according to Takens Theorem, our embedded space should be between 2 and 3 (Takens, 1981; Kantz & Schreiber, 1999).

Twenty data sets were generated from the original series, thus obtaining values for D_2 between 1.79 and 1.86. Each of the D_2 surrogates values were higher than the original time series D_2 value (1.6), hence, the time series can be considered as significantly nonlinear with a confidence of 95%. Details of this nonlinearity test can be found in Kantz and Schreiber (1999).

Summarizing, time series of the project “Health Care System Card” showed a behavior which was mainly periodic, having nonlinear effects in the last months of years 2003 and 2004. Therefore, linear models offered a suitable prediction in the number of events 6 h ahead, but TDNN improved considerably prediction performance in the last months of both years. This improvement was due to the significantly-deterministic nonlinear behavior of the series, i.e. the origin of the nonlinearities was not only noise.

3.5. Other projects

The analyses shown in previous subsections were also carried out in all the projects shown in Table 1. Table 5 presents the most successful models in terms of the prediction accuracy in other two relevant CETESI projects, namely, ‘Legal’ and ‘Mail’. Correlation coefficients

between the targets and the model outputs were also very adequate in out-of-sample predictions, except in the case of the project ‘Mail’ during the weeks 49–52 of 2004, in which the correlation coefficient was equal to 0.75. In both projects, deterministic nonlinear effects were only relevant during the weeks 49–52 in 2003, in which TDNN significantly improved the ARMA model predictions. In the rest of the projects shown in Table 1, it was possible to build accurate prediction models in several periods, but the standard of high accuracy required by CETESI (out-of-sample prediction coefficient >0.8) was not always achieved.

4. Conclusions

Time series related to the number of events in projects managed by CETESI have been analyzed in this study. Stochastic modeling and analysis has been used, as well as time processing neural networks and nonlinear time series analysis tools. The proposed approach allows to know whether time series show a deterministic behavior, and hence, it is possible to obtain a prediction model.

Results have shown the suitability of our approach in predicting the number of events 6 h ahead in the project “health care system card”. This series is essentially linear, and therefore, ARMA modeling has offered robust predictions in many cases. However, there is a significant nonlinear effect in the last months of both years, being the prediction improved by the use of neural networks in these periods. Moreover, analysis of nonlinear parameters has proved this nonlinearity, and moreover, has guaranteed a prediction horizon that can be extended until one day ahead with an 88–97% prediction accuracy (correlation between out-of-sample observed data and model output). The proposed approach can be successfully applied to other CETESI projects, with 86–94% prediction accuracy in many of the year periods. We feel that this new method is a useful tool for management of Support Centers.

5. Uncited reference

[Froyland, 1992],

References

- Bishop, M. (1995). *Neural networks for pattern recognition*. Clarendon Press. 510–511
- Box, G. E. P., & Jenkins, G. M. (1976). *Time series analysis: forecasting and control*. Wiley. 512–513
- Brockwell, P. J., & Davis, R. A. (2002). *Introduction to time series and forecasting*. Springer. 514–515
- ~~Froyland, J. (1992). *Introduction to chaos and coherence*. Institute of Physics Publishing. 516–517~~
- Haykin, S. (1999). *Neural networks: a comprehensive foundation*. Prentice Hall. 518–519
- Kantz, H., & Schreiber, T. (1999). *Nonlinear time series analysis*. Cambridge University Press. 520–521
- Makridakis, S., Wheelwright, S. C., & Hyndman, R. J. (1998). *Forecasting: methods and applications*. Prentice Hall. 522–523

- 524 Takens, F. (1981). Detecting strange attractors in turbulence. *Springer*
525 *lecture notes in mathematics*, 898, 366–381. 528
- 526 Vapnik, V. N. (1999). *The nature of statistical learning*. New York:
527 Springer. 529
- Weigend, A. S., & Gershenfeld, N. A. (1996). *Time series prediction: forecasting the future and understanding the past*. Addison-Wesley. 530
- Wiggins, S. (2003). *Introduction to applied nonlinear dynamical systems and chaos*. New York: Springer. 531
532

UNCORRECTED PROOF