Information Sciences 178 (2008) 3288–3299

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Time series forecasting with a non-linear model and the scatter search meta-heuristic

Carlos Gomes da Silva *

Escola Superior de Tecnologia e Gestão de Leiria, Morro do Lena, Alto Vieiro, 2401-951 Leiria, Portugal
INESC-Coimbra, Rua Antero de Quental, 199, 3000-033 Coimbra, Portugal

A R T I C L E   I N F O

Article history:
Received 3 March 2007
Received in revised form 25 March 2008
Accepted 27 March 2008

Keywords:
Forecast
Non-linear models
ARMA
Scatter search
Meta-heuristics

A B S T R A C T

Forecasting the behavior of variables (e.g., economic, financial, physical) is of strategic value for organizations, which helps to sustain practical interest in the development of alternative models and resolution procedures. This paper presents a non-linear model that combines radial basis functions and the ARMA(p,q) structure. The optimal set of parameters for such a model is difficult to find. In this paper, a scatter search meta-heuristic is used to find this optimal set. Five time series are analyzed to assess and illustrate the pertinence of the proposed meta-heuristic method.

1. Introduction

Many planning activities require predicting the behavior of variables (e.g., economic, financial and physical). These predictions support the strategic decisions of organizations [10], which in turn sustains practical interest in forecasting methods. Several methods – both causal and non-causal – have been proposed to forecast such variable behavior. The causal methods try to explain one variable in terms of a set of other variables. This category includes the widely studied classification problem, which consists of assigning objects into a predefined group given a number of observed attributes of these objects (see [18] for a review of neural networks applied to this problem). The non-causal methods, on the other hand, study the historic values of a variable in order to forecast its future value. This category includes the time series methods.

Time series methods are generally used when there is not much information about the generation process of the underlying variable and when other variables provide no clear explanation of the studied variable [19]. Gooijer and Hyndman [4] have published a review of the literature for the methods used to model time series over the last 25 years. The most popular methods are the moving average, exponential smoothing and ARIMA (autoregressive integrated moving average) methods. These methods all assume linear relationships among the past values of the forecast variable.

Although the linear assumption makes it easier to manipulate the models mathematically, it can lead to inappropriate representations of many real-world patterns in which non-linear relationships are prevalent. Non-linearities can also be used to show abrupt changes in the real data. Nonetheless, incorporating non-linearities into models can lead to very difficult mathematical problems, in which the optimal set of parameters may be difficult to know. This difficulty could explain why there have been fewer studies devoted to non-linear time series than to linear time series.

* Tel.: +351 244 820 300/239 851 040; fax: +351 244 820 310/239 824 692.
E-mail address: cgsilva@estg.ipleiria.pt

0020-0255/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.ins.2008.03.024

Specific non-linear patterns can be taken into account using TAR (threshold autoregressive) and ARCH (autoregressive conditional heterocedastic) models [19]. The usefulness of alternative resolution methods, approximate in nature but less restrictive, for general non-linearities has been noted by several authors. For instance, [7] used a genetic fuzzy predictor ensemble; [14] used probability instructions combined with genetic programming; [15] applied genetic algorithms to tune radial basis functions; [1] proposed a neural model; [21] used neural networks for quarterly time series forecasting; [19] combined the ARIMA structure with neural networks; [13] also proposed the use of hybrid ARIMA and artificial neural networks models; [5,16] presented fuzzy time series models; [20] proposed a neural ensemble model that incorporates noise into the data used to build different training sets; and [3] applied an enhanced polynomial artificial neural networks.

Although neural networks constitute one of the most popular models for dealing with non-linear time series, the method's general applicability is now being questioned [4], and the model's over-parameterization remains a drawback. (A comparison of neural network forecasting applications with the classical ARMA(p,q) structures can be found in [6].) This paper investigates the potential of a non-linear model, whose parameters are estimated with the scatter search meta-heuristic, to forecast a given time series. This model is based on the linear ARMA(p,q) model, expressed mathematically as:

$$Y(t) = w_0 + \sum_{i=1}^{p} w_i Y(t-i) + \sum_{i=p+1}^{p+q} w_i u(t-i) + u(t), \quad t = p+1, \ldots, n, \tag{1}$$

where $Y(t)$ is the value, in period t, of the variable Y to be forecast, w_i ($i = 0, \ldots, p + q$) are parameters of the model, $u(t)$ is the error term and n is the number of available/used periods of time.

The non-linear model NARMA(p,q,l) can thus be defined as:

$$Y(t) = \alpha_0 + \sum_{j=1}^{l} \alpha_j h_j\left(w_0 + \sum_{i=1}^{p} w_i Y(t-i) + \sum_{i=p+1}^{p+q} w_i u(t-i)\right) + u(t), \quad t = p+1, \ldots, n, \tag{2}$$

where $h_j(x)$ is a non-linear function, with x being an input vector, and (p,q,l) is the order of the model. This NARMA(p,q,l) model uses l non-linear functions to transform the ARMA(p,q) output, and then $Y(t)$ is approximated by a linear combination of those transformations. This model is also related to the radial basis function models [9,15] which incorporate an approximation property, provided that a large number of radial basis functions are used. This paper focuses on the non-linear Gaussian function (3), usually used in studies of neural networks [17,9,15,1], for example). The logistic function, $h_j(x) = \frac{1}{1+\exp(-x)}$, can also be used as alternative, as [19] has done; however, the computational experiments conducted for this paper produced worse results for this function. For this reason, only the results for the Gaussian function

$$h_j(x) = \exp\left(-\left(\frac{x - c_j}{r_j}\right)^2\right) \tag{3}$$

are presented.

Let $w = (w_0, w_1, \ldots, w_p, w_{p+1}, \ldots, w_{p+q})$, $\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_l)$, $c = (c_1, \ldots, c_l)$, $r = (r_1, \ldots, r_l)$ be the parameters to be estimated, represented in a compact form by $X \equiv (w, \alpha, c, r)$. Using a Gaussian function and the NARMA(p,q,l) model, there are a total of $p + q + 3l + 2$ parameters to be found. The aim of this paper is to find the set of parameters X, also called the solution set that minimizes the mean sum square of the errors (MSE), which is used to evaluate the results:

$$MSE(X) = \frac{\sum_{t=1}^{n}\left(Y(t) - \widehat{Y}(t)\right)^2}{n}, \tag{4}$$

where $\widehat{Y}(t)$ is the estimated value of $Y(t)$ using the found parameters, X.

The problem of finding such a set of parameters for (2) is complex, and alternative methods for deriving these parameters is of significant interest in time series studies. When dealing with non-linear models, two approaches are usually followed. The first seeks to develop exact methods for deriving the optimal parameters for the model. The second tries to find a good "near-optimal" set of parameters using meta-heuristic procedures. The meta-heuristic methods attempt to lead the parameter search to an interesting region discovered by an evolutionary process. Due to the complexity of the underlying problem, the second approach is appealing because meta-heuristic procedures require no assumptions about the parameters and offer the possibility of obtaining a global optimum of (4). Proposed by Fred Glover [2], the scatter search meta-heuristic used in this paper is based on the progressive combination of interesting solutions. This method has been successfully applied to many optimization problems. (See http://www.uv.es/~rmarti for an updated list of studies based on the scatter search meta-heuristic.)

The rest of the paper is organized as follows. Section 2 describes how the scatter search is applied to the estimation problem. Section 3 shows the computational experiments conducted to test the proposed method. Section 4 summarizes the main conclusions drawn from the research presented in the paper.

2. The meta-heuristic estimation of the model parameters

The scatter search meta-heuristic is used to estimate the parameters for the NARMA(p,q,l) model described above. The order of the model is considered previously defined. To find an appropriate model order, several combinations of the parameters may be tested in order to avoid selecting an inappropriate a priori model.