Designing Morphological/Rank/Linear Filters via Modified Genetic Algorithm for Time Series Forecasting

Ricardo de A. Araújo[1], **Tiago A. E. Ferreira**[1]

[1]Center for Informatics – Federal University of Pernambuco
Av. Prof. Luiz Freire, s/n, CDU, 50732-970, Recife - PE - Brazil

{raa,taef}@cin.ufpe.br

Abstract. *This paper presents an evolutionary approach for designing Morphological/Rank/Linear (MRL) filters for time series forecasting. It consists of an evolutionary model composed of a MRL filter and a modified Genetic Algorithm (GA) with optimal genetic operators to accelerate the search convergence. The proposed method performs an evolutionary search for the minimum number of time lags (and their corresponding specific positions) to represent the time series, as well as the parameters of the MRL filter, defined by mixing parameter (λ), rank (r), linear Finite Impulse Response (FIR) filter ($\underline{b}$) and Morphological/Rank (MR) filter ($\underline{a}$) coefficients. An experimental analysis is conducted with the proposed method using two real world time series and five well-known performance measurements, demonstrating good performance of MRL filtering systems for time series forecasting.*

1. Introduction

The popular statistical technique of Box & Jenkins [Box et al. 1994] (ARIMA models) is the most common choice for the prediction of time series. However, the Box & Jenkins models are linear and most real world applications can involve nonlinear problems. This fact may introduce a limitation to the accuracy of the generated predictions with Box & Jenkins models. In order to overcome this limitation, many nonlinear statistical approaches have been developed, such as the bilinear models [Rao and Gabr 1984], the threshold autoregressive models [Ozaki 1985], the exponential autoregressive models [Priestley 1988], the general state dependent models [Rumelhart and McCleland 1987], amongst others. However, these nonlinear statistical models have a high mathematical complexity and, in the practical, similar performance to linear statistical models [Clements et al. 2004].

In order to overcome the limitation of linear and nonlinear statistical models, approaches based on Artificial Neural Networks (ANNs) have been successful applied for nonlinear modeling of time series [Zhang et al. 1998]. In this context, a relevant work was presented by Ferreira [Ferreira 2006], where was defined the Time-delay Added Evolutionary Forecasting (TAEF) method for time series prediction, which performs an evolutionary search for the minimum necessary number of time lags adequate for representing the time series, based on the Takens theorem [Takens 1980]. The TAEF method [Ferreira 2006] finds the most fitted predictor model for representing a time series, and then performs a behavioral statistical test in order to adjust time phase distortions that may appear in the representation of the time series.

Nonlinear filters are widely applied in signal processing. An important class of nonlinear filter systems is based on the framework of Mathematical Morphology (MM) [Maragos 1989, Serra 1982]. In the literature, many works have focused on the design of discrete increasing morphological systems and rank or stack filters. Salembier [Salembier 1992, P. Salembier 1992] designed Morphological/Rank (MR) filters via gradient-based adaptive optimization. Pessoa and Maragos [Pessoa and Maragos 1998] proposed a new hybrid filter, referred to as Morphological/Rank/Linear (MRL) filter, which consists of a linear combination of an MR filter and a linear Finite Impulse Response (FIR) filter. In the morphological systems context, another work was presented by Araújo et al. [Araújo et al. 2006]. It consists of an evolutionary morphological approach definition for financial time series forecasting, which provides a mechanism to design an evolutionary model based on increasing and non-increasing translation invariant morphological operators.

This paper presents an evolutionary approach for designing MRL filters for time series forecasting. It consists of an evolutionary model composed of a MRL filter [Pessoa and Maragos 1998] and a modified Genetic Algorithm (GA) [Leung et al. 2003] having optimal genetic operators to accelerate the search convergence. The proposed method performs an evolutionary search for the minimum number of time lags (and their corresponding specific positions) to represent the time series, as well as the parameters of MRL filter (mixing parameter (λ), rank (r), linear FIR filter ($\underline{b}$) and MR filter ($\underline{a}$) coefficients). An experimental analysis is conducted with the proposed method using a financial time series (National Association of Securities Dealers Automated Quotation (Nasdaq) Index) and a natural phenomena time series (Brightness of a Variable Star Series). Five well-known performance measurements are used to assess the performance of the proposed method: Mean Square Error (MSE), Mean Absolute Percentage Error (MAPE), Normalized Mean Square Error (NMSE), Prediction Of Change In Direction (POCID) and Average Relative Variance (ARV).

2. Forecasting Problem

A time series is a set of points, generally time equidistant, defined by,

$$X_t = \{x_t \in \mathbb{R} \mid t = 1, 2, \ldots, N\}, \tag{1}$$

where t is the temporal index and N is the number of observations. Therefore, X_t is a sequence of temporal observations orderly sequenced and equally spaced.

The main objective of the forecasting techniques is to identify certain regular patterns present in the data set in order to create a model capable of generating the next temporal patterns. In this context, a crucial factor for a good forecasting performance is the correct choice of the time lags. Such relationship structures among historical data constitute a d-dimensional phase space, where d is the dimension capable of representing such relationship. Takens [Takens 1980] proved that if d is sufficiently large, such built phase space is homeomorphic to the phase space which generated the time series.

The crucial problem in reconstructing the original state space is the correct choice of the d, or more specifically, the correct choice of the time lags. The proposed method in this paper tries to reconstruct the phase space of a given time series by carrying out a search for the minimum dimension necessary (i.e. important time lags) to reproduce the phenomenon that generates the times series and its subsequent values.

3. Modified Genetic Algorithm

The modified Genetic Algorithm (GA) used here is based on the work of Leung et al. [Leung et al. 2003], where specials crossover and mutation operators are applied to accelerate the search convergence.

The crossover operator [Leung et al. 2003] is used for exchanging information from two parents (p_1 and p_2) obtained in the selection process by a roulette wheel approach [Leung et al. 2003]. The recombination process to generate the offspring (C_1, C_2, C_3 and C_4) is done by four crossover operators, which are defined by the following equations [Leung et al. 2003]:

$$C_1 = \frac{p_1 + p_2}{2}, \tag{2}$$

$$C_2 = p_{max}(1-w) + max(p_1, p_2)w, \tag{3}$$

$$C_3 = p_{min}(1-w) + min(p_1, p_2)w, \tag{4}$$

$$C_4 = \frac{(p_{max} + p_{min})(1-w) + (p_1 + p_2)w}{2}, \tag{5}$$

where $w \in [0, 1]$ denotes the crossover weight (the closer w is to 1, the greater is the direct contribution from parents), $max(p_1, p_2)$ and $min(p_1, p_2)$ denote the vector whose elements are the maximum and the minimum, respectively, between the gene values of p_1 and p_2. The terms p_{max} and p_{min} denote the maximum and minimum possible gene values, respectively. After the offspring generation by crossover operators, the son with the best evaluation (greater fitness value) will be chosen as the son generated by the crossover process and denoted C^{best}.

After the crossover operator, C^{best} is selected to have a mutation process, where three new sons are generated and defined by the following equation [Leung et al. 2003]:

$$M_j = C_k^{best} + \gamma_k \Delta M_k, \ j = 1, 2, 3 \ \text{ and } \ k = 1, 2, \ldots, n, \tag{6}$$

where γ_k can only take the values 0 or 1, ΔM_k are randomly generated numbers such that $p_{min} \leq C_k^{best} + \Delta M_k \leq p_{max}$ and n denotes the number of genes in the chromosome.

The first mutation son (M_1) is obtained according to (6) using only one term γ_k set to 1 (k is randomly selected within the range $[1, n]$) and the remaining terms γ_k are set to 0. The second mutation son (M_2) is obtained according to (6) using some γ_k randomly chosen and set to 1 and the remaining terms γ_k are set to 0. The third mutation son (M_3) is obtained according to (6) using all γ_k set to 1.

4. Morphological/Rank/Linear Fundamentals

Next, will be presented some preliminary theoretical concepts and notations of Morphological/Rank/Linear (MRL) filter, which are used in the proposed approach.

4.1. Morphological/Rank/Linear Filter Preliminaries

Definition 1 – Rank Function: the r-th rank function of the vector $\underline{t} = (t_1, t_2, \ldots, t_n) \in \mathbb{R}^n$ is the r-th element of the vector $\underline{t}$ sorted in decreasing order ($t_{(1)} \geq t_{(2)} \geq \ldots \geq t_{(n)}$). It is denoted by [Pessoa and Maragos 1998],

$$\mathcal{R}_r(\underline{t}) = t_{(r)}, \quad r = 1, 2, \ldots, n. \tag{7}$$

For example, given the vector $\underline{t} = (3, 0, 5, 7, 2, 1, 3)$, its 4-th rank function is $\mathcal{R}_4(\underline{t}) = 3$.

Definition 2 – Unit Sample Function: the unit sample function is given by [Pessoa and Maragos 1998]

$$q(v) = \begin{cases} 1, & \text{if } v = 0, \\ 0, & \text{otherwise.} \end{cases} \tag{8}$$

where $v \in \mathbb{R}$.

Applying the unit sample function to a vector $\underline{v} = (v_1, v_2, \ldots, v_n) \in \mathbb{R}^n$ yields a vector unit sample function ($Q(\underline{v})$), given by [Pessoa and Maragos 1998]

$$Q(\underline{v}) = [q(v_1), q(v_2), \ldots, q(v_n)]. \tag{9}$$

Definition 3 – Rank Indicator Vector : the r-th rank indicator vector $\underline{c}$ of $\underline{t}$ is given by [Pessoa and Maragos 1998]

$$\underline{c}(\underline{t}, r) = \frac{Q\left((z \cdot \underline{1}) - \underline{t}\right)}{Q\left((z \cdot \underline{1}) - \underline{t}\right) \cdot \underline{1}'}, \tag{10}$$

where $z = \mathcal{R}_r(\underline{t})$, $\underline{1} = (1, 1, \ldots, 1)$ and the symbol $'$ denotes transposition.

For example, given the vector $\underline{t} = (3, 0, 5, 7, 2, 1, 3)$, its 4-th rank indicator function is $\underline{c}(\underline{t}, 4) = \frac{1}{2}(1, 0, 0, 0, 0, 0, 1)$.

Definition 4 – Smoothed Rank Function: the smoothed r-th rank function is given by [Pessoa and Maragos 1998]

$$\mathcal{R}_{r,\sigma}(\underline{t}) = \underline{c_\sigma}(\underline{t}, r) \cdot \underline{t}', \tag{11}$$

with

$$\underline{c_\sigma}(\underline{t}, r) = \frac{Q_\sigma\left((z \cdot \underline{1}) - \underline{t}\right)}{Q_\sigma\left((z \cdot \underline{1}) - \underline{t}\right) \cdot \underline{1}'}, \tag{12}$$

where c_σ is an approximation for the rank indicator vector $\underline{c}$ and $Q_\sigma(\underline{v}) = [q_\sigma(v_1), q_\sigma(v_2), \ldots, q_\sigma(v_n)]$ is a smoothed impulse function (where $q_\sigma(v)$ is like $sech^2(v/\sigma)$ or $exp\left[-\frac{1}{2}(v/\sigma)^2\right]$) and $\sigma \geq 0$ is a scale parameter.

Thus, c_σ is an approximation for the rank indicator vector $\underline{v}$. Using ideas of fuzzy set theory, $\underline{c_\sigma}$ can also be interpreted as a membership function vector [Pessoa and Maragos 1998]. For example, if the vector $\underline{t} = (3, 0, 5, 7, 2, 1, 3)$, $q_\sigma(v) = sech^2(\frac{v}{\sigma})$ and $\sigma = 0.5$ then its smoothed 4-th rank indicator function is

$$\underline{c_\sigma}(\underline{t}, 4) = \tfrac{1}{2}(0.9646, 0, 0.0013, 0, 0.0682, 0.0013, 0.9646),$$

whereas $\underline{c}(\underline{t}, 4) = \frac{1}{2}(1, 0, 0, 0, 0, 0, 1)$.

4.2. Morphological/Rank/Linear (MRL) Filter

The MRL filter [Pessoa and Maragos 1998] is a linear combination between a Morphological/Rank (MR) filter and a linear Finite Impulse Response (FIR) filter.

Definition 5 – MRL Filter [Pessoa and Maragos 1998]: Let $\underline{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ represent the input signal inside an n-point moving window and let y be the output from the filter. Then, the MRL filter is defined as the shift-invariant system whose local signal transformation rule $x \rightarrow y$ is given by [Pessoa and Maragos 1998]

$$y = \lambda\alpha + (1 - \lambda)\beta, \tag{13}$$

with

$$\alpha = \mathcal{R}_r(\underline{x} + \underline{a}) = \mathcal{R}_r(x_1 + a_1, x_2 + a_2, \ldots, x_n + a_n), \tag{14}$$

and

$$\beta = \underline{x} \cdot \underline{b}' = x_1 b_1 + x_2 b_2 + \ldots + x_n b_n, \tag{15}$$

where $\lambda \in \mathbb{R}$, $\underline{a}$ and $\underline{b} \in \mathbb{R}^n$. The terms $\underline{a} = (a_1, a_2, \ldots, a_n)$ and $\underline{b} = (b_1, b_2, \ldots, b_n)$ represent the coefficients of the MR filter and the coefficients of the linear FIR filter, respectively. The term $\underline{a}$ is usually referred to structuring element because for $r = 1$ or $r = n$ the rank filter becomes the morphological dilation and erosion by a structuring function equal to $\pm\underline{a}$ within its support [Pessoa and Maragos 1998]. If $1 < r < n$, it uses $\underline{a}$ to generalize the standard unweighted rank operations to filters with weights [Pessoa and Maragos 1998].

5. The Proposed Approach

The proposed approach consists of an evolutionary model composed of a MRL filter [Pessoa and Maragos 1998] combined with the modified GA [Leung et al. 2003] (Section 3). It is based on the definition of the two main elements necessary for building an accurate forecasting system according to Ferreira [Ferreira 2006]: (a) minimum number of time lags adequate for representing the time series, and (b) structure of the model capable of representing such underlying information. It is important to consider the minimum possible number of time lags in the correct representation of the series because the model must to be as parsimonious as possible.

Following this principle, the proposed method, referred to as Morphological/Rank/Linear design by GA training (MRL-GA), uses the modified GA [Leung et al. 2003] for training the MRL filter [Pessoa and Maragos 1998], whose main goal is to define the following important parameters: (1) the minimum number of time lags and their corresponding specific positions to represent the series (initially, a maximum number ($MaxLags$) is defined and then the GA can choose any value in the interval $[1, MaxLags]$ for each individual of the population), and (2) the parameters of MRL filter (mixing parameter (λ), rank (r), linear FIR filter coefficients and MR filter coefficients). The proposed prediction scheme in this paper also uses the phase fix procedure idea from Ferreira [Ferreira 2006] where a two step procedure is introduced, which tries to adjust time phase distortions that may appear in the financial time series.

The GA individuals are evaluated by the fitness function defined by,

$$fitness = \frac{POCID}{1 + MSE + MAPE + NMSE + ARV} \tag{16}$$

where MSE, $MAPE$, $NMSE$, $POCID$ and ARV will be formally defined on Section 6.

The termination conditions for the GA are:

1. Minimum value of fitness function: $fitness \geq 40$, where this value mean the accuracy to predict direction around 80% ($POCID \gtrsim 80\%$) and the sum of the other errors around one ($MSE + MAPE + NMSE + ARV \cong 1$). If a model (best individual of GA population) reaches this fitness value, then the system tests new models by 1000 GA iterations ahead. If there is not fitness improvement in these 1000 GA interactions ahead, then the system stops;
2. The increase in the validation error or generalization loss (Gl) [Prechelt 1994]: $Gl > 5\%$;
3. The decrease in the training error process training (Pt) [Prechelt 1994]: $Pt \leq 10^{-6}$.

5.1. Modeling of GA Individuals

Each individual of the GA population is a MRL filter. The individuals are represented by chromosomes that have the following genes (MRL filter parameters):

- $\underline{a}$: MR filter coefficients (in the range $[-1, 1]$);
- $\underline{b}$: linear FIR filter coefficients (in the range $[-1, 1]$);
- ρ: variable used to determine the rank r, which is given by [Pessoa and Maragos 1998]

$$r = round\left(n - \frac{n-1}{exp(-\rho)}\right), \tag{17}$$

  where $round(\cdot)$ denotes the usual symmetrical rounding operation and n is the dimension of the input vector;
- λ: mixing parameter (in the range $[0, 1]$);
- $\underline{lag}$: a vector having size $MaxLags$, where each position has a real-valued codification, which is used to determine if a specific time lag will be used ($lag_i \geq 0$) or not ($lag_i < 0$).

6. Performance Evaluation

Many performance evaluation criteria is found in literature. However, most of the existing literature on time series prediction frequently employ only one performance criterion for model evaluation. The most widely used criterion is the Mean Squared Error (MSE), given by

$$MSE = \frac{1}{N}\sum_{j=1}^{N}(target_j - output_j)^2, \tag{18}$$

where N is the number of patterns, $target_j$ is the desired output for pattern j and $output_j$ is the predicted value for pattern j.

The MSE measure may be used to drive the prediction model in the training process, but it cannot be considered alone as a conclusive measure for comparison of different prediction models [Clements et al. 2004]. For this reason, other performance criteria should be considered for allowing a more robust performance evaluation.

How the measure MSE doesn't offer clearly the forecasting model performance, it's necessary to use a measure that is capable of to identify accurately the model deviation. A measure that presents such behavior is the Mean Absolute Percentage Error (MAPE), given by

$$MAPE = \frac{100}{N}\sum_{j=1}^{N}\left|\frac{target_j - output_j}{x_j}\right|, \tag{19}$$

where x_j is the time series value at point j.

A naive strategy used to make predictions is define the last time series observation as the best future prediction value ($X_{t+1} = X_t$). This kind of behavior is generally found in financial time series prediction, and it's known as the random walk model [Mills 2003]. Thus, a way to evaluate the model based on this behavior is using the Normalized Mean Squared Error (NMSE), which associates the model performance with a random walk model, and given by

$$NMSE = \frac{\sum_{j}(target_j - output_j)^2}{\sum_{j}(output_j - output_{j+1})^2}, \tag{20}$$

where, if the NMSE is equal to 1, the predictor has the same performance of a random walk model. If the NMSE is greater than 1, then the predictor has a performance worse than a random walk model, and if the NMSE is lesser than 1, the predictor is better than a random walk model. In the perfect model, the NMSE tends to zero.

Another interesting measure to evaluate the model consists in the accuracy to predict direction or, more specifically, if the future value (prediction target) is going to up or to down, regarding the last observed time series value. It's known as Prediction Of Change In Direction (POCID), and given by

$$POCID = \frac{100}{N}\sum_{j=1}^{N} D_j, \tag{21}$$

where

$$D_j = \begin{cases} 1 & \text{if } (target_j - target_{j-1})(output_j - output_{j-1}) > 0 \\ 0 & \text{otherwise} \end{cases} \tag{22}$$

The last measure used associates the model performance with the mean of the time series. The measure is the Average Relative Variance (ARV), and it is given by

$$ARV = \frac{\sum_{j=1}^{N} (output_j - target_j)^2}{\sum_{j=1}^{N} (output_j - \overline{target})^2}, \tag{23}$$

where $\overline{target}$ is the mean of the time series. If the ARV is equal to 1, the predictor has the same performance of the time series average. If the ARV is greater than 1, then the predictor has a performance worse than the time series average, and if the ARV is lesser than 1, the predictor is better than the time series average. In the ideal model, ARV tends to zero.

7. Experimental Results

The GA parameters used in MRL-GA method are a maximum number of GA generations corresponding to 10^4, weight $w = 0.9$ (used in the crossover operator), mutation probability equals to 0.1 and maximum number of lags $MaxLags = 10$. The MR filter coefficients and the linear FIR filter coefficients ($\underline{a}$ and $\underline{b}$, respectively) were normalized in the range $[-1, 1]$. The MRL filter parameters λ and ρ were in the range $[0, 1]$ and $[-MaxLags, MaxLags]$, respectively.

A set of two time series was used as a test bed for evaluation of the proposed method: a financial time series (National Association of Securities Dealers Automated Quotation (Nasdaq) Index) and a natural phenomena time series (Brightness of a Variable Star Series). All series investigated were normalized to lie within the range $[0, 1]$ and divided in three sets according to Prechelt [Prechelt 1994]: training set (50% of the points), validation set (25% of the points) and test set (25% of the points).

Next, will be presented the simulation results involving the MRL-GA model with and without the phase fix procedure [Ferreira 2006], referred to as MRL-GA model out-of-phase and MRL-GA model in-phase, respectively. These two procedures (in-phase and out-of-phase) were used to study the possible performance improvement, in terms of fitness function, of the phase fix procedure applied to MRL-GA Model.

To build a reference performance level, also will be presented results by random walk (RW) model [Mills 2003] and by ANN (Artificial Neural Network - multilayer perceptron) model. The ANN was trained by Levenberg-Marquardt (ANN-LM) [Hagan and Menhaj 1994] algorithm with architecture $X - 10 - 1$, which X denotes the number of units in the input layer, 10 units in the hidden layer and 1 unit in the output layer (prediction horizon of one step ahead). The term X represents the relevant time lags, and for each time series we use the same time lags chosen by the proposed model (MRL-GA). The termination conditions for the ANN-LM [Hagan and Menhaj 1994] algorithm are the maximum number of epochs (10^4), the increase in the validation error or generalization loss ($Gl > 5\%$) and the decrease in the error of the process training ($Pt < 10^{-6}$). For each time series, was made ten experiments, where the experiment with the largest validation fitness function is chosen to represent the prediction model.

7.1. National Association of Securities Dealers Automated Quotation (NASDAQ) Index Series

The National Association of Securities Dealers Automated Quotation (NASDAQ) Index series corresponds to daily observations from February 2nd 1971 to June 18th 2004, constituting a database of 8428 points.

For the NASDAQ Index series prediction (with one step ahead of prediction horizon), the proposed method automatically chose the lags 1, 3, 4, 6, 8, 9 and 10 as the relevant lags ($n = 7$) and defined the parameters $\rho = -2.3355$ and $\lambda = 0.0897$. The Table 1 shows the results for all performance measures for RW model, ANN-LM model and MRL-GA model with and without phase fix procedure.

Table 1. Results for the NASDAQ Index series.

	RW Model	ANN-LM Model	MRL-GA Model In-Phase	MRL-GA Model Out-Of-Phase
MSE	$1.8747 \cdot 10^{-5}$	$2.0511 \cdot 10^{-5}$	$2.0628 \cdot 10^{-5}$	$1.8731 \cdot 10^{-5}$
MAPE	0.39%	0.42%	0.41%	0.39%
NMSE	1.0000	1.0941	1.0985	1.0000
POCID	52.70%	53.23%	52.94%	53.65%
ARV	$2.9877 \cdot 10^{-3}$	$3.2689 \cdot 10^{-3}$	$3.2839 \cdot 10^{-3}$	$2.9812 \cdot 10^{-3}$
fitness	21.9541	21.0646	21.0127	22.4195

According to Table 1, the prediction of MRL-GA model out-of-phase (with the phase fix procedure) obtained performance slightly better (in terms of fitness function) than RW model, ANN-LM model and MRL-GA model in-phase (without the phase fix procedure). The obtained NMSE measure value (1.0000) indicates that the MRL-GA model out-of-phase is a random walk like model [Mills 2003]. The POCID measure (53.65%) shows that the MRL-GA model out-of-phase has a similar behavior to a heads or tails experiment, but was slightly better than RW model (52.70%), ANN-LM model (53.23%) and MRL-GA model in-phase (52.94%).

The phase fix procedure was applied in intention to correct the distortions of the forecast phase [Ferreira 2006]. The MRL-GA model out-of-phase had a forecast performance slightly better than the model version without the phase fix procedure, but the prediction for the NASDAQ Index series is still dislocated one step ahead the original values, characterizing the results obtained by the NMSE. Figure 1(b) shows the actual NASDAQ Index values (solid line) and the predicted values generated by the MRL-GA model out-of-phase (dashed line) for the last 100 points of the test set.

7.2. Brightness of a Variable Star Series

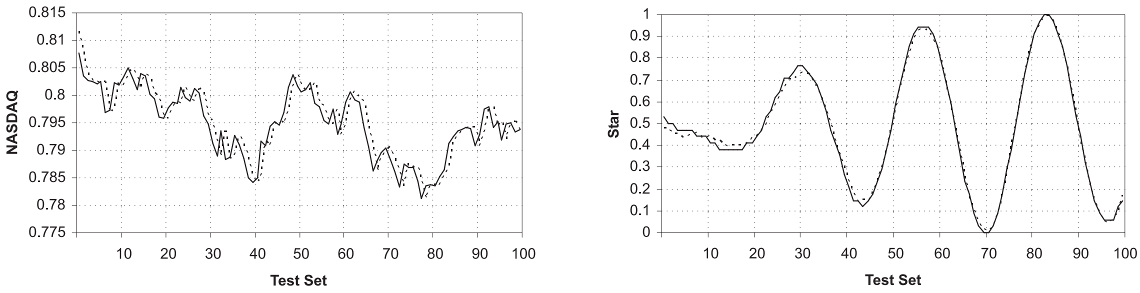
The Brightness of a Variable Star series, or Star series, corresponds to daily observations in the same place and hour of an oscillating shine star, constituting a database of 600 points.

For the prediction of the Star series (with one step ahead of prediction horizon), the proposed method automatically chose the lags 1, 2, 3, 8, 9 and 10 as the relevant lags ($n = 6$) and defined the parameters $\rho = 6.7532$ and $\lambda = 0.4914$. The Table 2 shows the results for all performance measures for RW model, ANN-LM model, MRL-GA model in-phase and MRL-GA model out-of-phase.

Table 2. Results for the Star series.

	RW Model	ANN-LM Model	MRL-GA Model In-Phase	MRL-GA Model Out-Of-Phase
MSE	$3.7191 \cdot 10^{-3}$	$1.3041 \cdot 10^{-3}$	$6.3223 \cdot 10^{-4}$	$3.5584 \cdot 10^{-4}$
MAPE	16.14%	7.35%	6.88%	4.83%
NMSE	1.0000	0.3510	0.1679	$9.6329 \cdot 10^{-2}$
POCID	65.98%	75.51%	76.43%	82.70%
ARV	$5.4643 \cdot 10^{-2}$	$1.9161 \cdot 10^{-2}$	$9.2891 \cdot 10^{-3}$	$5.1522 \cdot 10^{-3}$
fitness	3.6257	8.6545	9.4851	13.9417

According to Table 2, the prediction of the MRL-GA model out-of-phase obtained better performance (in terms of fitness function) than RW model, ANN-LM model and MRL-GA model in-phase. The obtained NMSE measure ($9.6329 \cdot 10^{-2}$) indicates that the MRL-GA model out-of-phase has a better performance than a random walk like model [Mills 2003]. The POCID measure (82.70%) shows that the MRL-GA model out-of-phase has a better performance than a heads or tails experiment, but it was slightly better than RW model (65.98%), ANN-LM model (75.51%) and MRL-GA model in-phase (76.43%). The Figure 1(a) shows the actual Star values (solid line) and the predicted values generated by the MRL-GA model out-of-phase (dashed line) for the last 100 points of the test set.



(a) (b)

Figure 1. Prediction results for the analyzed time series (test set): actual values (solid line) and predicted values (dashed line).

8. Conclusions

An evolutionary approach for designing Morphological/Rank/Linear (MRL) filters for time series forecasting was presented in this paper. It consists of an evolutionary model composed of a MRL filter and a modified Genetic Algorithm (GA) having optimal genetic operators to accelerate the search convergence. The proposed method, called Morphological/Rank/Linear design via Genetic Algorithm training (MRL-GA), searches for the minimum number of time lags (and their corresponding specific positions) to represent the time series, as well as the parameters of MRL filter (mixing parameter (λ), rank (r), linear Finite Impulse Response (FIR) filter ($\underline{b}$) and Morphological/Rank (MR) filter ($\underline{a}$) coefficients) to solve the time series forecasting problem.

Five different metrics were used to measure the performance of the proposed method for time series forecasting. The method was applied to a real world time series from the financial market with all their dependence on exogenous and uncontrollable variables (Nasdaq Index) and a natural phenomena time series (Brightness of a Variable Star Series). The experimental results demonstrated slightly better performance of the proposed MRL-GA model out-of-phase (with the phase fix procedure), when compared to RW model, ANN-LM model and MRL-GA model in-phase (without the phase fix procedure). Furthermore, the MRL filtering structure is quite attractive with its modest computational complexity for designing the model when compared to the model proposed in Araújo et al. [Araújo et al. 2006] and other statistical nonlinear models [Rao and Gabr 1984, Ozaki 1985, Priestley 1988, Rumelhart and McCleland 1987].

It was observed that the proposed model obtained a better performance than a random walk model [Mills 2003] for the analyzed natural phenomena series, whereas it obtained a similar performance to a random walk model for all the analyzed financial time series, where the predicted values were shifted one step ahead the original values, indicating a random walk like model [Mills 2003]. This observation is also in agreement with the work of Sitte and Sitte [Sitte and Sitte 2002] and Araújo et al. [Araújo et al. 2006], which have shown that the predictions of financial time series exhibit a characteristic one step shift with respect to the original data. However, Ferreira [Ferreira 2006] shown that this behavior (random walk like model) can be corrected with the phase fix procedure applied to a evolve Artificial Neural Network (ANN). Ferreira [Ferreira 2006] also shown that the phase fix procedure successful is strongly dependent of the ANN parameters adjustment and the model used to forecasting.

The phase fix procedure proposed by Ferreira [Ferreira 2006] was not able to correct the prediction phase when applied to the proposed model, but it was capable to improve the prediction performance. Why the proposed model has this behavior is a mystery for us, and studies are being accomplished to try to explain such behavior.

References

Araújo, R. A., Madeiro, F., Sousa, R. P., Pessoa, L. F. C., and Ferreira, T. A. E. (2006). An evolutionary morphological approach for financial time series forecasting. In *Proceedings of the IEEE Congress on Evolutionary Computation*, Vancouver, Canada.

Box, G. E. P., Jenkins, G. M., and Reinsel, G. C. (1994). *Time Series Analysis: Forecasting and Control*. Prentice Hall, New Jersey, third edition.

Clements, M. P., Franses, P. H., and Swanson, N. R. (2004). Forecasting economic and financial time-series with non-linear models. *International Journal of Forecasting*, 20(2):169–183.

Ferreira, T. A. E. (2006). *A New Hybrid Intelligent Methodology for Time Series Forecasting*. PhD thesis, Federal University of Pernambuco, Av. Prof. Luz Freira, s/n, Cidade Univercitaria, Recife, Pernambuco - Brazil - 50732-970.

Hagan, M. and Menhaj, M. (1994). Training feedforward networks with the marquardt algorithm. *IEEE Transactions on Neural Networks*, 5(6):989–993.

Leung, F. H. F., Lam, H. K., Ling, S. H., and Tam, P. K. S. (2003). Tuning of the structure and parameters of the neural network using an improved genetic algorithm. *IEEE Transactions on Neural Networks*, 14(1):79–88.

Maragos, P. (1989). A representation theory for morphological image and signal processing. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 11:586–599.

Mills, T. C. (2003). *The Econometric Modelling of Financial Time Series*. Cambridge University Press, Cambridge.

Ozaki, T. (1985). *Nonlinear Time Series Models and Dynamical Systems*, volume 5 of *Hand Book of Statistics*. Noth-Holland, Amsterdam.

P. Salembier (1992). Adaptive rank order based filters. *Signal Process.*, 27(1):1–25.

Pessoa, L. F. C. and Maragos, P. (1998). MRL-filters: A general class of nonlinear systems and their optimal design for image processing. *IEEE Transactions on Image Processing*, 7:966–978.

Prechelt, L. (1994). Proben1: A set of neural network benchmark problems and benchmarking rules. Technical Report 21/94.

Priestley, M. B. (1988). *Non-Linear and Non-Stacionary Time Series Analysis*. Academic Press.

Rao, T. S. and Gabr, M. M. (1984). *Introduction to Bispectral Analysis and Bilinear Time Series Models*, volume 24 of *Lecture Notes in Statistics*. Springer, Berlin.

Rumelhart, D. E. and McCleland, J. L. (1987). *Parallel Distributed Processing, Explorations in the Microstructure of Cognition*, volume 1 & 2. MIT Press.

Salembier, P. (1992). Structuring element adaptation for morphological filters. *Journal for Visual Communication and Image Representation*, 3(2):115–136.

Serra, J. (1982). *Image analysis and mathematical morphology*. Academic Press, London.

Sitte, R. and Sitte, J. (2002). Neural networks approach to the random walk dilemma of financial time series. *Applied Intelligence*, 16(3):163–171.

Takens, F. (1980). Detecting strange attractor in turbulence. In Dold, A. and Eckmann, B., editors, *Dynamical Systems and Turbulence*, volume 898 of *Lecture Notes in Mathematics*, pages 366–381, New York. Springer-Verlag.

Zhang, G., Patuwo, B. E., and Hu, M. Y. (1998). Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting*, 14:35–62.