Journal of Forecasting, *J. Forecast.* (2015)
Published online in Wiley Online Library (wileyonlinelibrary.com) **DOI**: 10.1002/for.2354

Forecasting Daily and Monthly Exchange Rates with Machine Learning Techniques

VASILIOS PLAKANDARAS, THEOPHILOS PAPADIMITRIOU* AND PERIKLIS GOGAS
Department of Economics, Democritus University of Thrace, Komotini, Greece

ABSTRACT
In this paper we propose and test a forecasting model on monthly and daily spot prices of five selected exchange rates. In doing so, we combine a novel smoothing technique (initially applied in signal processing) with a variable selection methodology and two regression estimation methodologies from the field of machine learning (ML). After the decomposition of the original exchange rate series using an ensemble empirical mode decomposition (EEMD) method into a smoothed and a fluctuation component, multivariate adaptive regression splines (MARS) are used to select the most appropriate variable set from a large set of explanatory variables that we collected. The selected variables are then fed into two distinctive support vector machines (SVR) models that produce one-period-ahead forecasts for the two components. Neural networks (NN) are also considered as an alternative to SVR. The sum of the two forecast components is the final forecast of the proposed scheme. We show that the above implementation exhibits a superior in-sample and out-of-sample forecasting ability when compared to alternative forecasting models. The empirical results provide evidence against the efficient market hypothesis for the selected foreign exchange markets.

KEY WORDS exchange rate forecasting; support vector regression; multivariate adaptive regression splines; ensemble empirical mode decomposition

INTRODUCTION

Following the breakdown of the Bretton Woods fixed exchange rate system, a large number of models were proposed to outperform the random walk (RW) model used as the usual benchmark in exchange rate forecasting. Forecasting the evolution of exchange rates is, of course, important for (a) investors that make portfolio decisions and (b) macroeconomic variable forecasting, since exchange rates play an important role in their future evolution and vice versa (Rime *et al*., 2010). Overall, we can distinguish the existing empirical approaches into two broad categories: (a) models exploiting the influence of macroeconomic variables on the exchange rate markets and (b) models that build on a microstructural and atheoretical approach of the market. The former include the so-called monetary exchange rate models, which attempts to establish a link between the evolution of exchange rates and the variability of fundamentals based, for example, on the theory of purchasing power parity, the uncovered interest rate parity and the money demand (for a detailed exposition see Mark and Sul, 2001). Building on the general relationship between macroeconomic variables and exchange rate evolution, Mundell (1968) and Fleming (1962) propose one of the first analytical model forms, later extended by Dornbusch (1976) to the overshooting (or sticky prices) model. On a similar perspective, the flexible price monetary model has been for many years the workhorse of exchange rate economics (Bilson, 1978; Stockman, 1980; Lucas, 1982) extended by Frankel (1979) to its real interest rate differential variant.[1] These models are mainly used for medium and long-term forecasting.

A substantial part of the critique against the ability of fundamentals to forecast exchange rates stems from the seminal paper of Messe and Rogoff (1983). In a forecasting exercise considering various monetary exchange rate models against a RW with a drift, they find that regardless of the version of the monetary model used none outperforms the RW model in terms of mean square error in out-of-sample forecasting. Nevertheless, later studies challenge the findings of Meese and Rogoff. Mark and Sul (2001), using panel regression models, reject the univariate framework of Messe and Rogoff (1983), concluding that fundamentals do possess significant forecasting ability on cross-sectional data. Cheung *et al*. (2005) study a variety of monetary exchange rate models for the period 1990–2000. They conclude that monetary models, in many cases, outperform the RW model. Molodtsova and Papell (2009), by developing models that depend on the Taylor rule on fundamentals, report superior out-of-sample performance on short forecasting horizons for 11 exchange rates as compared to the RW. Overall, the review of the existing literature leads to the conclusion that fundamentals can be used successfully for exchange rate forecasting.

Focusing on medium-term forecasting, Karemera and Kim (2006) develop autoregressive integrated moving average (ARIMA) models that outperform RW models for a number of currencies on a monthly forecasting horizon.

*Correspondence to: Theophilos Papadimitriou, Department of Economics, Democritus University of Thrace, 69100 Komotini, Greece. E-mail: papadimi@ierd.duth.gr

[1]A complete reference of all monetary exchange rate literature is beyond the scope of this paper. The interest reader is referred to Macdonald (2010).

Nevertheless, the forecasting ability is strongly tied to the time period under evaluation. Cheung (1993) detects long-run memory in exchange rates and proposes the use of autoregressive fractionally integrated moving average (ARFIMA) models, with limited success. On the other hand, generalized autoregressive conditional heteroskedasticity (GARCH) models are a popular selection among practitioners for volatility modeling, although they lack clear-cut validation of their forecasting ability. Galbraith and Kisinbay (2005) compare GARCH, Exponential GARCH and simple AR models in forecasting the DEM/USD and the JPY/USD exchange rates. Projections based on GARCH-in-the-mean models provide better one-step-ahead forecasts than GARCH and fractionally integrated GARCH models for a period of up to 30 trading days.

Machine learning methodologies, and more specifically NNs and SVMs, gained significant merit based on their ability to model nonlinear systems with minimal initial assumptions and high forecasting accuracy. Dunis and Williams (2002) compare numerous alternative models: NN, RW, ARCH, moving average convergence/divergence, ARMA and a logit model on the EUR/USD exchange rate. Their sample spans from May 2000 to July 2001 and uses a daily forecasting horizon. They report the superiority of the NN-based models over all alternatives used. Brandl *et al.* (2009) use genetic algorithms for variable selection and set their model using the SVR methodology. The initial dataset includes variables suggested by the purchasing power parity theory, the covered interest rate parity and the uncovered interest rate parity. The proposed model outperforms a NN, an OLS regression and an ARIMA model on monthly out-of-sample forecasting for the EUR/USD, the USD/JPY and the USD/GBP rates. On a similar framework, Ince and Trafalis (2006) couple ARIMA and SVM in order to improve the directional forecasting of the EUR/USD exchange rate. Their results show that the proposed methodology outperforms the logit/probit models. Overall, the forecasting ability of the SVM/R methodology in time series forecasting has attracted significant interest in the relevant economics literature (e.g. Härdle *et al.*, 2009; Khandani *et al.*, 2010; Rubio *et al.*, 2011; Öğüt *et al.*, 2012; Papadimitriou *et al.*, 2015).

The use of the empirical mode decomposition (EMD) technique as part of a hybrid forecasting methodology is common to the engineering literature. EMD originates from the field of signal processing and it decomposes a given signal into its basis functions (Huang *et al.*, 1998). Then, using an SVR model, we can forecast each component separately. The summation over all SVR models produces the final forecast value of the method. Specifically in exchange rate forecasting, Fu (2010) applies the EMD-SVR methodology for the weekly EUR/RMB exchange rate and a one-period-ahead forecasting window, producing forecasts for the period 25 June 2005 to 26 June 2009. He concludes that the combined EMD-SVR model forecasts better than the simple SVR model. Lin *et al.* (2012) apply the EMD-SVR method in forecasting the USD/NTD, the JPY/NTD and the RMB/NTD exchange rates. Overall the EMD-SVR methodology outperforms the SVR, the NN, the ARIMA and the EMD-ARIMA models.

In this paper we propose a novel forecasting methodology using three steps:

Step 1. Extraction of the long- and short-run dynamics of the series based on the components of the EEMD method.
Step 2. Feature selection using the MARS approach.
Step 3. Employing an SVR model to produce the in-sample and out-of-sample forecasts.

The innovations of our approach in comparison to previous EMD-SVR studies are as follows: (a) We decompose the original time series of the exchange rates into the long-run trend and short-run fluctuations, exploiting the superior EEMD decomposition method instead of the earlier EMD (for more details see Wu and Huang, 2009). According to the theory the two components of the decomposition are driven by an entirely different data-generating process. As a result, we avoid decomposing the exchange rates into several components that often lack a theoretical and practical justification as it is done in previous studies in the literature. (b) We train only two instead of training many forecasting models (one forecasting model for each of the EEMD decomposition basis functions), and thus we achieve a lower computational cost. (c) To the best of our knowledge, no previous study considers a structural approach to the EMD-SVR forecasting methodology.

METHODOLOGY AND DATASET

Methodology overview

Macdonald (2010) argues that after the collapse of the Bretton Woods fixed exchange rate system, the high volatility of exchange rate time series makes accurate forecasting difficult. Smoothing techniques provide a less volatile representation of the exchange rate series, removing short-run dynamics and noise which are extremely volatile and thus harder to forecast. In the literature regarding time series smoothing, several methodologies are proposed.[2] A key issue to all smoothing methodologies is the definition of the optimal point, beyond which smoothing produces distortion rather than a long-run trend representation. The optimal percentage of the information that should be discarded and its potential contribution to forecasting are unknown beforehand. Many smoothing procedures select ad hoc the smoothness

[2]For a detailed survey on the field of trend extraction and smoothing on time series the interested reader is referred to Alexandrov *et al.* (2009).

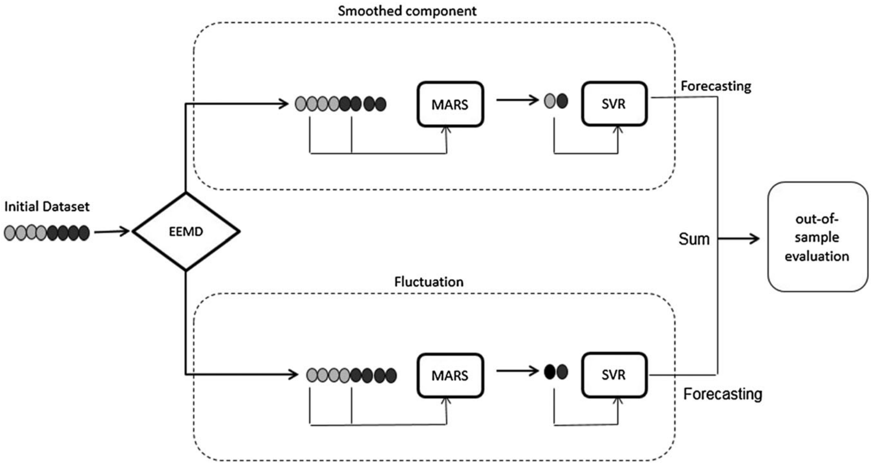

Figure 1. Overview of the proposed EEMD-MARS-SVR methodology

degree imposed: see, for example, the commonly used Hodrick–Prescott (1997) filter and discussion on the optimum setting of the smoothness parameter λ according to the sampling frequency.

To overcome information loss and mitigate proper model selection dilemmas, some researchers propose that instead of removing the fluctuating components we should forecast the initial time series in different sampling frequencies (daily, monthly, yearly, etc.) and then combine the results (Kourentzes *et al*., 2014; Silvestrini and Veredas, 2008). In high-frequency sampling short-run dynamics are prominent, and as the sampling frequency decreases long-run characteristics such as a trend and seasonality are observed. In this way, different time series characteristics can be identified and the combined results can be more accurate in forecasting and bias reduction (Abraham, 1982; Mohammadipour and Boylan, 2012).

In this paper we combine the suggestion of smoothing methodologies concerning the extraction of short-run dynamics of the series with the proposition of combining short- and long-run dynamics (Kourentzes *et al*., 2014). Both these approaches suggest an improvement on the overall forecasting accuracy. An overview of the method is depicted in Figure 1. The long-run component of each variable is evaluated as a potential regressor in the forecasting model for the long-run component of the exchange rate; the same concept is followed for the short-run component of each variable.[3] We call the proposed method EEMD-MARS-SVR from the methodologies that compose it.

Time series decomposition: ensemble empirical mode decomposition

The EEMD is a data-driven algorithm that decomposes a time series into finite additive oscillatory components called intrinsic mode functions (IMF). Proposed by Wu and Huang (2009), the main advantage of the EEMD is the lack of initial assumptions on the dataset, i.e. linearity or stationarity. The decomposition into IMFs is achieved through an iterative scheme of the following generic procedure:

1. Add white noise with a predefined amplitude to the time series under consideration.
2. Detect local minima and maxima.
3. With cubic interpolation compute the lower and upper envelopes from the local minima and maxima respectively.
4. Compute the mean value of the lower and the upper envelope. If (a) the number of the local minima, the local maxima and the zero crossing points vary at most by one and (b) the mean is approximately zero, we subtract the mean from the initial time series. The residual is the first IMF. If the criteria are not met we go back to step 2 and repeat the procedure until criteria fulfillment.
5. Repeat step 4 until the residual has no local maxima and minima.
6. Go back to step 1 and repeat the process.
7. Take the mean of the ensemble of all decompositions for each IMF as the final output (i.e. the mean of the first IMFs of all decompositions is the final first IMF and so on). The ensemble of the decomposed IMFs is relieved from the initially added noise.[4]

Unlike other traditional decomposition methods such as the Fourier transform or wavelet decompositions, the EEMD is less restrictive as it does not use a priori determined basis functions and allows for time-varying envelopes extracted from the data. In this way it is capable of capturing intrinsic physical features hidden in nonlinear and non-

[3]Pearson correlation coefficient results for the selection of the long-run component can be accessed at http://utopia.duth.gr/~vplakand/Supplementary_material_on_EEMD_MARS_SVR_paper/
[4]For more details the interested reader is referenced to Wu and Huang (2009).

stationary time series that are often smoothed out by other methods (Hou *et al*., 2009). An example of the decomposition method applied to the daily EUR/USD exchange rate is depicted in Figure 2.

The EEMD decomposition method results by construction in successive IMFs of lower frequency. Many researchers argue that IMFs with high frequency represent short-run dynamics of the initial series and the less volatile ones represent long-run trends, with the final residual often representing the long-run trend of the phenomenon (Wu *et al*., 2007). For instance, in Figure 2, we could argue that the first IMF (second sub-graph) represents daily or even tick-by-tick patterns and the final residual depicts the upward trend of the exchange rate over the period 1999–2011.

Moghtaderi *et al*. (2013) build on the aforementioned framework, proposing a trend extraction technique based on decomposition. They argue that long-run trends should be examined as summations rather than independent IMFs, since individual characteristics are dispersed between IMFs and are not exclusively isolated in only one. In other words, they suggest that the long-run trend is a smoothed representation of the initial series resulting from the summation of lower-frequency IMFs and of the final residual of the EEMD decomposition. Highly volatile IMFs are considered as irrelevant components or noise and are discarded. As a result, the smoothing problem breaks down to selecting the IMF index which defines the limit between short- and long-run dynamics. The authors conclude that the tipping point between the highly volatile irrelevant IMFs and the initial point of the summation is the first IMF where the energy in comparison to its former IMF rises. In general, the mathematical notation of a signal's energy is given by

$$E^i \triangleq \sum_{t=1}^{n} \left|\mathrm{IMF}_t^i\right|^2, \ 1 \le i \le L \tag{1}$$

where IMF_t^i is the tth observation of the ith IMF, E^i is the energy of the ith IMF and L is the total number of IMFs. Thus, with mathematical notation, the smoothing function is expressed as

$$e_{i^*} = \sum_{i=i_*}^{L} \mathrm{IMF}^i + r, \ 1 \le i \le L \tag{2}$$

where e_{i*} is the smoothed variant of the initial series for index i_*, r is the final EEMD residual and i_* is the index of the IMF from which we start the summation. For instance, in Figure 2, after the decomposition of the daily exchange rate time series into 10 IMFs and the residual, energy rises on the fourth and sixth IMF. By summing up from the fourth up to the tenth IMF plus the residual, we obtain the smoothed function representation in Figure 3, while Figure 4 depicts the corresponding implementation on monthly data.

In contrast to Moghtaderi *et al*. (2011), we propose a different approach. The suggestion that the fluctuating (short-run representation) part is irrelevant to forecasting the actual exchange rate is rather strong and arbitrary. We propose a methodological framework where the previously discarded highly volatile IMFs are summed to compose a fluctuating component that represents the short-run dynamics of the exchange rate. Then, this fluctuating component is modeled and forecasted independently from the long-run trend. This allows us to identify and model alternative data generation processes for the short- and long-run dynamics of the exchange rates, exactly as

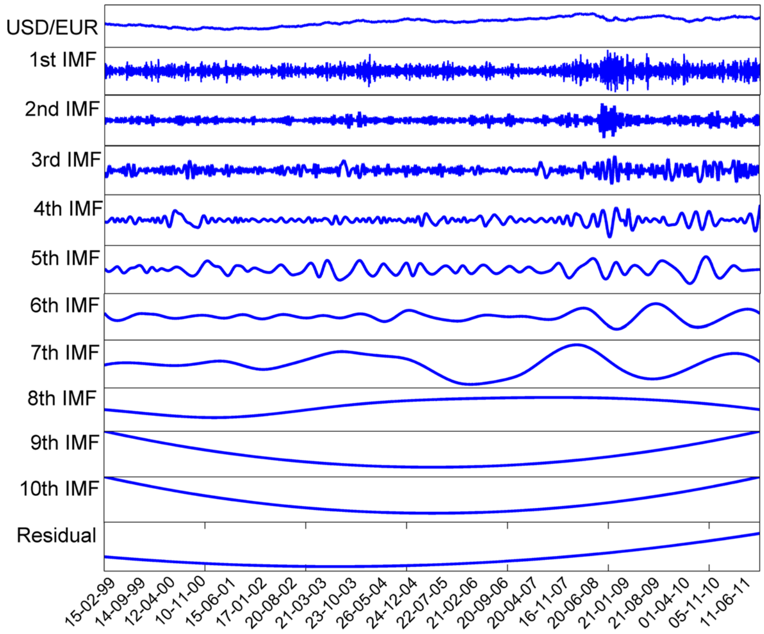

Figure 2. Decomposition of the original (first row) daily EUR/USD exchange rate into 10 IMFs. The last row (straight line) is the residual of the EEMD process.

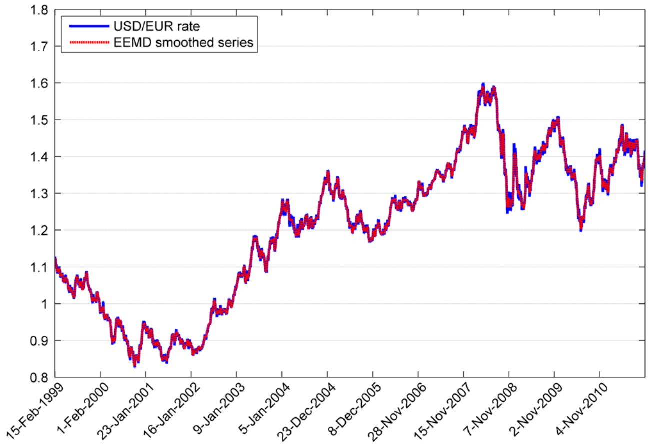

Figure 3. Daily EUR/USD exchange rate and smoothed series.

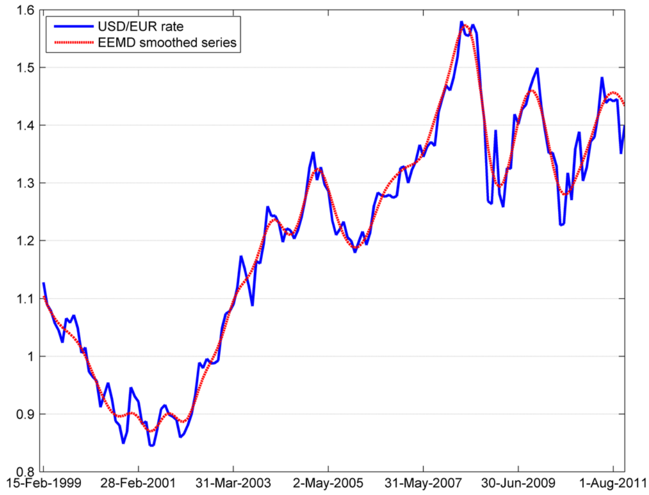

Figure 4. Monthly EUR/USD exchange rate and smoothed series.

international economics theory suggests. Our approach n using both short and long run dynamics in forecasting is supported by the empirical results.[5]

Variable selection: Multivariate Adaptive Regression Splines (MARS)

A crucial step in any forecasting model is the selection of the input variables since they determine the overall accuracy of the model. A large part of the existing literature in exchange rate forecasting selects the input variable set based on a theoretical framework, e.g. the aforementioned sticky or flexible price monetary exchange rate model. Instead of 'hand-picking' the input variables, we use a large set of relevant variables in our sample and implement MARS for variable selection based on statistical loss metrics. The choice of MARS for variable selection is based on the relevant literature which provides empirical evidence on the ability of the method to select the most appropriate set of explanatory variables in terms of high forecasting accuracy of the resulting models (Sephton, 2001; Andres *et al*., 2011; Kao *et al*., 2013). In our empirical section alternative methodologies such as the elastic net (a variant of ridge regression and LASSO) have also been used. Nonetheless, the variables selected by the elastic net approach lead to forecasting models that provide a lower forecasting accuracy as compared to the EEMD-MARS-SVR model.[6]

The MARS, proposed by Freidman (1991), is a non-parametric form of piece-wise nonlinear regression. In global parametric methods such as linear regression, the relationship between a depended and a set of explanatory variables is described using a global parametric function fitted universally to the full range of the dataset. In a piece-wise approach the available range is separated into sub-regions and a model is fitted locally to each sub-region. The points introducing the sub-regions are called knots. Using mathematical notation and starting from a training dataset $D = [(\boldsymbol{x}_1, y_1), (\boldsymbol{x}_2, y_2), \ldots, (\boldsymbol{x}_n, y_n)], \boldsymbol{x}_i \in \mathbb{R}^m, y_i \in \mathbb{R}, i = 1, 2, \ldots, n$, where for each observation pair $\boldsymbol{x}_i$ are the observation

[5]The limited space even of a journal paper is not enough to present all the produced simulations. For simulation treating only the smoothed part of each variable as potential regressor of the exchange rate forecasting model see http://utopia.duth.gr/~vplakand/Supplementary_material_on_EEMD_MARS_SVR_paper/One_Component_model.pdf
[6]Owing to space restrictions analytical comparison results between the two methods are available at http://utopia.duth.gr/~vplakand/Supplementary_material_on_EEMD_MARS_SVR_paper

samples and y_i is the dependent variable (the target of the regression system), MARS builds a model of the form

$$y_i = f(\boldsymbol{x}_i) + e = \beta_0 + \sum_{i=1}^{m} \beta_i B_i(x) + e \tag{3}$$

where β_o is a constant, β_i are local regression model coefficients of each sub-region, m the number of the sub-regions and $B_i(\boldsymbol{x})$ are spline basis functions of the form

$$B_i(\boldsymbol{x}) = \max(0, \boldsymbol{x_i} - \boldsymbol{h_i}) \tag{4}$$

where $\boldsymbol{h_i}$ is the corresponding knot of sub-region i and $\boldsymbol{x_i}$ the data sample of the sub-region i. MARS models are developed through a two-stage forward/backward stepwise regression procedure. In the forward stage, the entire D is split arbitrarily into overlapping sub-regions, considering a large number of knots (often treating almost all points as knots) and model parameters are selected by minimizing a lack-of-fit criterion. On the backward stage, spline basis functions (variables) and knots that no longer contribute to the accuracy of the fit are removed. A representation of a MARS model compared to a parametric linear regression model is depicted in Figure 5.

Forecasting: support vector regression (SVR)

The SVR is considered a state-of-the-art forecasting methodology; see Sapankevych and Sangar (2009) for a generalized comparison and Abraham and Yeung (2003) for a comparison specifically in exchange rate forecasting.

The SVR is a direct extension of the classic support vector machines algorithm (Cortes and Vapnik, 1995). When it comes to regression, the basic idea is to find a function that has at most a predetermined deviation from the actual values of the dataset. The support vector (SV) set which bounds this 'error tolerance band' is located in the dataset through a minimization procedure. The model is built in two steps: the training and the testing step. In the training step, the largest part of the dataset is used for the estimation of the function (i.e. the detection of the SVs that define the band); in the testing step, the generalization ability of the model is evaluated by checking the model's performance in the small subset that was left aside during training.

Using mathematical notation the linear regression function takes the form $f(\boldsymbol{x}) = \boldsymbol{w}^T\boldsymbol{x} + b$ by solving:

$$\min\left(\frac{1}{2}\|\boldsymbol{w}\|^2 + C\sum_{\imath=1}^{n}(\zeta_i + \zeta_i^*)\right)$$
$$\text{subject to} \begin{cases} y_i - (\boldsymbol{w}\boldsymbol{x}_i + b) \leq \varepsilon + \zeta_i \\ (\boldsymbol{w}\boldsymbol{x}_i + b) - y_i \leq \varepsilon + \zeta_i^* \\ \zeta_i,\ \zeta_i^* \geq 0 \end{cases} \tag{5}$$

where ε defines the tolerance band around the regression, and ζ_i, ζ_i^* are slack variables controlled through a penalty parameter C (see Figure 6). All the points inside the tolerance band have $\zeta_i,\ \zeta_i^* = 0$.Using the Lagrange multipliers from system (5) the solution is given by

$$\boldsymbol{w} = \sum_{i=1}^{n}(a_i - a_i^*)\boldsymbol{x}_i \tag{6}$$

and

$$y = \sum_{i=1}^{n}(a_i - a_i^*)\boldsymbol{x}_i^T\boldsymbol{x} \tag{7}$$

where a_i, a_i^* are Lagrange multipliers.

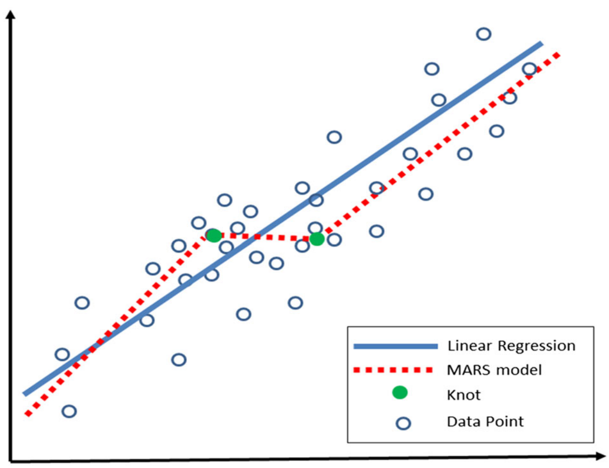

Figure 5. A MARS and a parametric linear regression model representation.

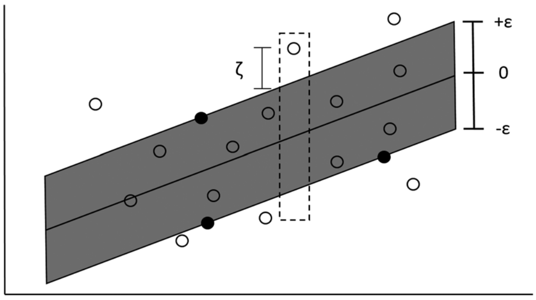

Figure 6. Upper and lower threshold on error tolerance indicated by the symbol ε. The boundaries of the error tolerance band are defined by SVs, denoted by black filled points. Forecast values greater than ε get a penalty ζ according to their distance from the tolerance accepted band.

A possible solution to treat nonlinear real phenomena datasets would be to project them into a higher-dimensional space where the transformed dataset can be described by a linear function. Nonlinear kernel functions have evolved the SVR mechanism to a nonlinear regression model, capable of approximating nonlinear phenomena. In our simulations we tested four kernels: linear, radial basis function (RBF), sigmoid and polynomial. The mathematical representation of each kernel is as follows:

Linear $$K_1(\boldsymbol{x}_1, \boldsymbol{x}_2) = \boldsymbol{x}_1^T \boldsymbol{x}_2 \tag{8}$$

RBF $$\boldsymbol{K}_2(\boldsymbol{x}_1, \boldsymbol{x}_2) = \boldsymbol{e}^{-\gamma\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|^2} \tag{9}$$

Polynomial $$\boldsymbol{K}_3(\boldsymbol{x}_1, \boldsymbol{x}_2) = \left(\gamma \boldsymbol{x}_1^T \boldsymbol{x}_2 + r\right)^d \tag{10}$$

Sigmoid $$\boldsymbol{K}_4(\boldsymbol{x}_1, \boldsymbol{x}_2) = \tanh\left(\gamma \boldsymbol{x}_1^T \boldsymbol{x}_2 + r\right) \tag{11}$$

with factors d, r, γ representing kernel parameters.

Forecasting: artificial neural networks

Inspired by the human neural system, an artificial neural network is an interconnected ‘neuronal’ network that models the relationship between a set of independent regressors and a dependent variable. In this paper, we develop a two-layer feed-forward network with a hidden and an output layer, respectively. We evaluate the sigmoid transfer function for the training of the neurons and the linear activation function in the output layer.

We train the network based on the Levenberg–Marquardt (LM) back-propagation training algorithm as it is faster and more reliable than other back-propagation techniques (Jeong and Kim, 2005). Moreover, since the LM training algorithm is sensitive to initial conditions, we perform a series of Monte Carlo simulations to provide robustness to our results.

The dataset

We compile an extensive set of variables that correspond to the predictors reported in the literature. The data are daily and span the period from 1 January 1999 to 30 October 2011, not including weekends and holidays. These variables include: (i) eight spot exchange rates; (ii) major stock indices from the U.S., Europe, the U.K. and Japan; (iii) 10 precious and non-precious metals spot prices; (iv) 18 commodities including crude oil; (v) moving averages of 3, 5, 10 and 30 days for all exchange rates; (vi) an index than represents the times that the exchange rate has appreciated over the last five trading days (to include technical analysis instruments); (vii) three indices of the USD exchange rate according to U.S. trade relationships: the Broad index, the Major Currencies index and the Other Partners index, as reported by the Federal Reserve Bank of Saint Louis (FRED); (viii) EURIBOR interest rates of various maturities; (ix) interest rate spreads between commercial paper, the federal funds and the effective federal funds rate;[7] and (x) basic macroeconomic variables from the U.S., Europe, the U.K., Japan, Australia, Norway, South Africa, Brazil and New Zealand. These result in the 193 variables of our initial dataset that are presented in Table I.

We choose 1-day-ahead forecasting as suggested by Rime *et al*. (2010) because it is a relevant horizon for exchange rate practitioners (e.g. most currency hedge funds). Then, using monthly data, we change the forecasting horizon to a monthly frequency.

The exchange rates that we use in our empirical analysis include (a) the high-volume exchange rates of EUR/USD and USD/JPY and (b) the medium-volume rates of AUD/NOK and NZD/BRL and the low-trade ZAR/PHP exchange rate. By doing this, we evaluate the forecasting ability of the proposed model in currencies of different trading interests and thus of different expected forms of market efficiency.

[7]For the EMU we used Marginal Lending Rate, EONIA and EURIBOR 6 months for the construction of the interest rate spreads.

Table I. Input variables

Commodities	Metals	Interest rates	Macroeconomic Variables for all countries
Crude oil	Gold	T-bill 6 months	Consumer price index
Cotton	Copper	T-bill 10 years	Productivity index
Lumber	Palladium	Spread MLP-EURIBOR 3M	Gross domestic product
Cocoa	Platinum	Spread MLR-Eonia	Trade balance
Coffee	Silver	Spread FF-CP	Unemployment rate
Orange juice	Aluminum	Spread FF-EFF	Central bank discount rate
Sugar	Zinc	EONIA	Long-term interest rate
Corn	Nickel	EURIBOR 1 week	Short-term interest rate
Wheat	Lead	ECB Interest rate	Aggregate money M3
Oats	Tin	EURIBOR 1 month	Public debt
Rough rice		FED rate	Deficit/surplus of government budget
Soybean meal	**Stock indices**		
Soybean oil	Dow Jones	**Technical analysis variables**	
Soybeans	Nasdaq 100	5-day index	**Exchange rates**
Feeder cattle	S&P 500	Moving average 3-day	JPY/EUR
Lean hogs	DAX	Moving average 5-day	JPY/USD
Live cattle	CAC 40	Moving average 10-day	USD/GBP
Pork bellies	FTSE 100	Moving average 30-day	BRL/NZD
Iron ore	FTSE 100		NOK/AUD
			PHP/ZAR
	USD trade weighted indices		EUR/GBP
	Major partners		EYR/USD
	Broad Index		
	Other Partners		

EMPIRICAL RESULTS

Forecasting models

In order to test the ability of the proposed EEMD-MARS-SVR methodology to forecast exchange rates, we compare it with several alternative forecasting models that are used in the relevant literature: (a) a RW model that serves as a benchmark for comparing forecasting methodologies; b) an ARIMA (p, d, q) representation; and (c) a GARCH(p, q) model. Moreover, from the field of machine learning we also use NNs as an alternative technique to the SVR. The number of neurons included in the hidden layer are determined based on a five-fold cross-validation scheme. We develop different models for each exchange rate: (a) a simple autoregressive NN (AR-NN); (b) an autoregressive NN after the EEMD decomposition (EEMD-AR-NN); (c) a structural NN with variables selected by MARS without the EEMD decomposition step (MARS-NN); and finally (d) a structural NN model with both the EEMD decomposition and the MARS step for variable selection.

In order to evaluate the individual contribution of each technique used to the overall performance of the EEMD-MARS-SVR methodology we follow the same approach with the NN models. We develop: (a) a simple autoregressive SVR (AR-SVR) and an autoregressive model with the EEMD preprocessing step coded EEMD-AR-SVR; (b) a structural MARS-SVR model without the EEMD preprocessing step; and (c) the full EEMD-MARS-SVR model. In this way we can observe the improvement in the overall forecasting ability of the full model provided in each step.

Table II reports the number of variables selected by MARS for monthly and daily data in panels A and B respectively. In the first two columns of each panel we report the number of selected variables for both the short-run and the long-run components for the models that use the EEMD decomposition. The third column of each panel (labeled MARS) reports the number of variables selected when no EEMD pre-filtering is applied. As we observe from Table II,

Table II. Number of input variables selected by MARS

	Panel A: Monthly data			Panel B: Daily data		
Exchange rate	Fluctuating component	Smoothed component	MARS	Fluctuating component	Smoothed component	MARS
EUR/USD	25	10	16	24	2	1
USD/JPY	7	13	19	25	2	1
AUD/NOK	9	11	10	21	2	1
NZD/BRL	24	11	14	22	2	1
ZAR/PHP	17	10	7	24	2	1

Table III. Number of input variables for ARIMA and GARCH models

	Panel A: Monthly data		Panel B: Daily data	
Exchange rate	ARIMA	GARCH	ARIMA	GARCH
EUR/USD	5	10	11	5
USD/JPY	5	5	4	4
AUD/NOK	5	3	4	5
NZD/BRL	2	5	9	2
ZAR/PHP	1	4	5	4

there are significant differences on the number of regressors selected for monthly and daily data and the short- and long-run components of all five exchange rates used. In general, the long-run component can be modeled with fewer explanatory variables than the short-run fluctuating part. This difference is especially noticeable with daily data as only two variables can model the smoothed component. Additionally, with daily data we detect the same difference when no decomposition is used on the original series. In this case, only one variable, namely the previous value of the exchange rate, can best model it, reducing the structural model to a simple AR(1).

The variables selected in Table II are next used to train the best forecasting MARS-SVR and EEMD-MARS-SVR models. The AR-SVR and EEMD-AR-SVR models are trained with the autoregressive terms only. Training is performed using the four alternative kernels as described above. For the selection of the optimum cost (C) and the optimum kernel parameters for the SVR models we follow a ten-fold and a four-fold cross-validation procedure for the daily and monthly dataset, respectively. The selection of the optimum parameters is achieved by performing a five-level coarse-to-fine grid search.

According to the standard ADF (Dickey and Fuller, 1981) and the Phillips–Perron tests (Phillips and Perron, 1988) that are used to determine the order of integration, all variables are found to be $I(1)$ according to the Engel and Granger (1987) notation.[8] Thus, in the ARIMA(p, d, q) and GARCH(p, q) models, where non-stationarity can lead to spurious results, all variables are expressed in first differences. In developing the ARIMA(p, d, q) model we iteratively determine the orders of the autoregression (p) and the moving average (q) using the minimized Schwarz information criterion (SIC)[9] (Schwarz, 1978). For the specification of the ARIMA(p, d, q) model we gradually introduce all exogenous variables using a forward iterative procedure: (a) we develop an initial ARIMA (p, d, q) model that includes the exchange rate and we introduce, one at a time, the rest of the variables; (b) from all these models we keep the one with the lowest SIC; (c) we start over with the two fixed variables (exchange rate and the one selected in step (b)) and introduce, one at a time, the rest of the variables, selecting the one with the lowest SIC; and (d) we continue until the introduction of new variables does not result in lower SIC values. The above configuration is repeated for the GARCH(p, q) model as well.[10] Table III reports the number of variables selected in the ARIMA(p, d, q) and GARCH(p, q) models; Panel A refers to monthly and panel B to daily data.

Statistical evaluation

Thus far we have constructed the best models in terms of in-sample forecasting accuracy for the proposed EEMD-MARS-SVM scheme and the nine alternative methodologies for all five exchange rates used and for both daily and monthly data. For this step we have only used 80% of our full sample data: 2600 observations spanning from January 4, 1999 to March 30, 2009 for the daily frequency, and 124 observations spanning January 4, 1999 to April 30, 2009 for the monthly frequency. In order to evaluate the generalization ability of these best constructed models, in this section we use them to produce one-period-ahead out-of-sample forecasts in the remaining 20% of the sample: 648 daily observations for the period April 2, 2009 to October 30, 2011 and 30 monthly observations for the period June 2009 to October 2011.

We evaluate and compare the out-of-sample forecasts of all ten methodological approaches using the mean absolute percentage error (MAPE) and directional symmetry (DS) metrics as follows:

$$\mathrm{MAPE} = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{\hat{y}_i - y_i}{y_i}\right| \tag{12}$$

[8]Unit root tests results are available at http://utopia.duth.gr/~vplakand/Supplementary_material_on_EEMD_MARS_SVR_paper/
[9]Application of the Akaike information criterion (Akaike, 1974) and Hannan–Quinn information criterion (Hannan and Quinn, 1979), as expected, selected higher values of p and q. We follow the SIC, which results in more parsimonious models.
[10]Owing to space restrictions, AIC, SIC and HQ values and the variables selected during the development of the ARIMA and GARCH models and those selected by MARS are available at http://utopia.duth.gr/~vplakand/Supplementary_material_on_EEMD_MARS_SVR_paper/

Table IV. Forecasting results on monthly exchange rates

Model	In-sample MAPE(%)	In-sample DS(%)	Out-of-sample MAPE(%)	Out-of-sample DS(%)
EUR/USD				
RW	2.290	52.846	2.952	43.333
ARIMA(0,1,0)	1.991	60.484	3.901	50.000
GARCH(2,1)	1.905	66.129*	3.749	43.333
AR-NN(1)	2.301	52.846	2.892	43.333
MARS-NN(1)	3.941	49.417	7.353	42.100
AR-SVR(RBF)	3.845	50.407	3.159	43.333
MARS-SVR(Linear)	5.158	51.639	8.068	50.000
EEMD-AR-NN(1)	2.317	52.033	2.886	46.667
EEMD-AR-SVR(Linear)	2.090	52.033	2.387*	56.667*
EEMD-MARS-NN(2)	1.840*	56.911	3.037	53.333
EEMD-MARS-SVR(Linear)	2.089	55.738	3.148	50.000
USD/JPY				
RW	2.250	51.220	2.207	40.000
ARIMA(3,1,3)	3.549	58.871	3.824	36.667
ARCH(2)	2.095	57.258	2.299	40.000
AR-NN(1)	2.298	51.219	7.728	40.000
AR-SVR(linear)	2.263	51.639	3.540	40.000
MARS-NN(9)	2.692	64.800	18.805	44.700
MARS-SVR(Poly)	1.374	72.311	18.433	63.333
EEMD-AR-NN(2)	2.237	57.724	7.536	50.000
EEMD-AR-SVR(Linear)	2.214	56.911	1.816*	65.517*
EEMD-MARS-NN(19)	1.341*	86.992*	7.542	53.333
EEMD-MARS-SVR(Linear)	2.186	55.738	3.052	46.667
AUD/NOK				
RW	2.383	39.837	1.508	56.667*
ARIMA (4, 1, 1)	2.091	51.613	1.651	43.333
GARCH (2, 2)	2.372	43.548	1.573	53.333
AR-NN (2)	2.321	36.667	4.392	56.667*
AR-SVR (linear)	4.428	52.459*	2.613	56.667*
MARS-SVR (1)	4.618	47.508	13.067	50.467
MARS-SVR (linear)	2.021*	45.082	4.202	50.000
EEMD-AR-NN (2)	2.218	48.781	4.468	53.333
EEMD-AR-SVR (poly)	2.230	40.650	1.433*	56.667*
EEMD-MARS-NN (1)	4.618	49.975	15.773	46.867
EEMD-MARS-SVR (linear)	2.102	41.803	10.223	43.333
NZD/BRL				
RW	4.101	46.342	2.195	33.333
ARIMA (0,1,0)	3.776	48.387	2.854	33.333
GARCH (1,1)	3.693	59.677*	4.012	30.000
AR-NN (4)	4.396	45.000	2.928	33.333
AR-SVR (RBF)	4.120	46.721	2.094	33.333
MARS-NN (4)	7.166	54.642	23.028	44.967
MARS-SVR (linear)	4.327	50.000	5.574	30.000
EEMD-AR-SVR (4)	4.026	46.342	2.258	40.000
EEMD-AR-SVR (sigmoid)	5.662	45.529	1.577*	53.333*
EEMD-MARS-NN (1)	4.627	52.846	3.553	40.000
EEMD-MARS-SVR (linear)	3.386*	55.738	3.651	36.667
ZAR/PHP				
RW	4.107	52.846	2.696	36.667
ARIMA (1,1,0)	3.880	56.452	3.214	40.000
GARCH (2,1)	3.846	58.871	2.800	40.000
AR-NN (19)	3.257	56.342	3.246	33.800
AR-SVR (linear)	4.076	52.459	2.669	36.667
MARS-NN (7)	7.887	53.742	17.725	42.533
MARS-SVR (linear)	4.065	52.459	3.361	30.000
EEMD-AR-NN (19)	3.440	57.724	3.203	33.333
EEMD-AR-SVR (poly)	4.063	51.220	2.131*	50.000*
EEMD-MARS-NN(15)	2.454*	72.358*	5.475	36.667
EEMD-MARS-SVR(RBF)	2.462	66.393	5.063	40.000

Note: Best forecast values are denoted by an asterisk. Best kernel of the SVR models and selected number of neurons for the NN model reported in parentheses.

$$\mathrm{DS} = \frac{100}{n}\sum_{i=1}^{n} d_i, \text{ where } d_i = \begin{cases} 1 \ \textit{if} \ (y_i - y_{i-1})(\hat{y}_i - \hat{y}_{i-1}) > 0 \\ 0 \ \textit{otherwise} \end{cases} \tag{13}$$

where $\hat{y}_i$ is the forecast of the actual y_i rate and n is the number of observations. MAPE measures the percentage of the absolute error in forecast, while DS measures the percentage of the times we correctly forecast the future direction of the exchange rate. These two metrics are independent of the variables' magnitude, thus permitting comparisons between models of different exchange rates. Directional forecasting is of key interest to market participants, since

Table V. Forecasting results on daily exchange rates

	In-sample		Out-of-sample	
Model	MAPE(%)	DS(%)	MAPE(%)	DS(%)
EUR/USD				
RW	0.469	45.231	0.554	45.567
ARIMA (0, 1, 0)	0.471	46.749	0.568	50.154
GARCH (1, 1)	0.468	47.364	0.554	48.611
AR-NN (1)	0.505	45.241	0.555	48.382
AR-SVR (linear)	0.494	45.231	0.555	48.457
EEMD-AR-NN (10)	0.398	59.923	0.455	58.398
EEMD-AR-SVR (linear)	0.687	62.077	0.512	51.543
EEMD-MARS-NN (1)	0.301	67.423	0.344*	67.180*
EEMD-MARS-SVR (RBF)	0.287*	69.423*	0.348	66.718
USD/JPY				
RW	0.487	45.962	0.500	43.301
ARIMA (0, 1, 0)	0.473	49.250	0.509	48.228
GARCH (1, 1)	0.481	48.019	0.506	47.920
AR-NN (2)	0.582	45.868	1.528	45.347
AR-SVR (linear)	0.487	45.962	0.525	45.370
EEMD-AR-NN (3)	0.445	49.359	0.589	47.920
EEMD-AR-SVR (linear)	0.445	49.500	0.417	56.944
EEMD-MARS-NN (3)	0.354*	62.346*	0.879	58.706
EEMD-MARS-SVR (linear)	0.362	61.192	0.397*	59.322*
AUD/NOK				
RW	0.538	46.846	0.488	47.920
ARIMA (0, 1, 1)	0.538	47.865	0.479	48.074
GARCH (2, 2)	0.537	48.492	0.484	47.304
AR-NN (1)	0.544	46.769	0.773	47.920
AR-SVR (linear)	0.541	46.846	0.496	47.840
EEMD-AR-NN (1)	0.472	54.962	0.464	54.083
EEMD-AR-SVR (linear)	0.540	59.500	0.421	53.704
EEMD-MARS-NN (1)	0.344*	71.462*	0.369	74.268*
EEMD-MARS-SVR (linear)	0.344*	70.462	0.312*	73.652
NZD/BRL				
RW	0.856	48.308	0.637	48.228
ARIMA (0, 1, 0)	0.854	49.327	0.643	47.612
GARCH (2, 2)	0.856	49.327	0.646	48.690
AR-NN (1)	0.869	48.154	0.638	48.228
AR-SVR (linear)	0.858	48.308	0.636	48.303
EEMD-AR-NN (12)	0.731	59.539	0.653	56.395
EEMD-AR-SVR (linear)	2.171	62.615	0.574	56.173
EEMD-MARS-NN (8)	0.581*	67.039*	0.455	63.945
EEMD-MARS-SVR (linear)	0.593	65.115	0.446*	65.331*
ZAR/PHP				
RW	0.844	50.077	0.848	49.615
ARIMA (0, 1, 1)	0.827	51.404	0.808	49.923
GARCH (1, 1)	0.827	51.404	0.825	50.231
AR-NN (4)	0.847	50.077	0.850	49.615
AR-SVR (linear)	0.846	50.077	0.846	49.537
EEMD-AR-NN (3)	0.722	57.500	0.740	55.316
EEMD-AR-SVR (linear)	0.734	59.385	0.697	58.025
EEMD-MARS-NN (1)	0.568*	70.769*	0.580*	73.960
EEMD-MARS-SVR (linear)	0.572	70.269	0.582	74.576*

Note: Best forecast values are denoted by an asterisk. Best kernel of the SVR models and number of neurons for the NN model noted in parentheses.

trading decisions on whether to go long or short on a currency depend on whether an appreciation or depreciation is expected rather than the exact future value of the exchange rate. The forecasting accuracy of each model is reported in Tables IV and V for monthly and daily data respectively.

In Table IV and in-sample forecasting the EEMD-MARS-NN model has the lowest MAPE in in-sample forecasting in three out of the five exchange rates. The EEMD-MARS-SVR model is the most accurate for the NZD/BRL and the MARS-SVR model for AUD/NOK. Overall, no model outperforms consistently the rest in in-sample forecasting. Nevertheless, only the EEMD-AR-SVR model outperforms the RW model in all exchange rates. Moreover, in the last two columns of Table IV, in out-of-sample forecasting the EEMD-AR-SVR model outperforms all alternative models for all five exchange rates in terms of both the MAPE and the DS metrics. The EEMD-AR-SVR model produces 13–25% better directional forecasts than the RW model. With the exception of the AUD/NOK, where we observe the same DS ratio in three cases, the rest of the models are all inferior to RW.

In Table V we report the same results for daily data. In terms of MAPE, the EEMD-MARS-SVR is superior in in-sample forecasting for the EUR/USD and the AUD/NOK, while the EEMD-MARS-NN forecasts more accurately than all models the three remaining exchange rates. With respect to DS, the EEMD-MARS-NN is the most accurate for all the five exchange rates with the exception of the EUR/USD, where the EEMD-MARS-SVR outperforms the other models. Nevertheless, the true forecasting ability of a model lies with out-of-sample forecasting. EEMD-MARS-SVR has the lowest MAPE for the USD/JPY, the AUD/NOK and the NZD/BRL with EEMD-MARS-NN being only slightly more accurate for the EUR/USD and the ZAR/PHP. Another interesting observation is that only the EEMD-AR-SVR and the EEMD-MARS-SVR consistently outperform the RW model both in in-sample and out-of-sample forecasting. In terms of DS the results are qualitatively the same. Both EEMD-MARS-ML methods achieve a 16–26% better directional forecast than the RW model, reaching up to 74.6% of accurate out-of-sample directional forecasting for the ZAR/PHP rate. Overall, both machine learning methodologies coupled with EEMD and MARS exhibit far more accurate forecasts than the alternative econometric approaches.

According to the results presented in Tables IV and V, the structural EEMD-MARS-SVR and EEMD-MARS-NN models have the highest out-of-sample forecasting performance in daily frequency, while the autoregressive EEMD-AR-SVR model is consistently more efficient in monthly frequency. This distinction may well reflect the underlined dynamics of exchange rate determination. Exchange rates follow different dynamics in the long and the short run. Long-run exchange rates are more dependent on macroeconomic dynamics: relative price levels (the theory of purchasing power parity), interest rate differentials (uncovered and covered interest rate parity), current account imbalances, public debt, etc. Short-run exchange rate determination adheres more closely to the forces of demand and supply in international exchange rate markets. This difference in the data-generating processes is probably captured by our two distinct models for daily and monthly exchange rate forecasting. Moreover, these results provide evidence that rejects even the weak-form of efficiency in both frequencies, as using only own historical values may be adequate to forecast the corresponding exchange rates.

In conclusion, our empirical findings are in line with those of Meese and Rogoff (1983), who conclude that in long-run forecasting horizon no structural exchange rate model can outperform in out-of-sample forecasting a RW model. Following the microstructural aspect of the exchange rate market, the role of exogenous variables in the determination of exchange rate value is imminent in short-run forecasting, while the influence of other exogenous variables is diminished in longer forecasting horizons. Moreover, coupling machine learning methodologies with the EEMD pre-processing step as proposed in this paper increases the overall forecasting ability of the method. The rejection of the EMH on both monthly and daily data points out the potential of developing profitable trading strategies based on forecasts of the proposed methodology.

CONCLUSION

We propose a hybrid combination of the EEMD time series decomposition into short- and long-run trend components that are then used to train an SVR model to produce out-of-sample forecasts of exchange rates. This setup combines a signal processing technique with a machine learning methodology. Motivated by the relevant theory, by decomposing the exchange rate time series and forecasting each component independently we are able to accommodate and focus on the different short- and long-run data generation processes as implied by the theory. Our specified models outperform the benchmark random walk model in out-of-sample forecasting for all five exchange rates employed; EUR/USD, JPY/USD, NOK/AUD, NZD/BRL and PHP/ ZAR. We implement two versions of this hybrid methodological setup; an autoregressive and a structural one. The former exhibits consistency in outperforming all other models used on forecasting out-of-sample monthly data. The latter (structural) produces the highest out-of-sample daily forecasting accuracy for three out of five exchange rates, with a close performance to the alternative EEMD-MARS-NN methodology that outperforms for the other two. Nevertheless, only the EEMD-MARS-SVR methodology consistently outperforms the benchmark random walk model in out-of-sample forecasting. The above findings corroborate with the microstructural aspect of the exchange rate market on the short run and the effect of macroeconomic dynamics (such as PPP, UIP, etc.) on the long run, validating exchange rate economics theory. Overall, our models seem to capture the different data generating processes between short and long horizons. The weak form of exchange rate market efficiency is rejected for both sampling frequencies due to the fact that the autoregressive model outperforms the RW benchmark.

ACKNOWLEDGEMENTS

This research has been co-financed by the European Union (European Social Fund – ESF) and Greek national funds through the Operational Program 'Education and Lifelong Learning' of the National Strategic Reference Framework (NSRF) – Research Funding Program: THALES (MIS 380292). Investing in knowledge society through the European Social Fund.

We would also like to thank the participants of the 75–76th International Atlantic Economic conference held in Vienna, Austria, in April 2013 and in Philadelphia, MA, USA, in October 2013, for their helpful comments on an earlier version of this paper.

REFERENCES

Abraham A, Yeung A. 2003. Integrating ensemble of intelligence systems for stock exchange prediction. *Notes on Computational Science* **2687**: 774–781.

Abraham B. 1982. Temporal aggregation and time series. *International Statistical Review* **50**: 285–291.

Akaike H. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* **19**(6): 716–723.

Alexandrov T, Bianconcini S, Bee Dagum E, Maass P, McElroy T. 2009. A review of some modern approaches to the problem of trend extraction. *Research Report Series*: *Statistics*, Vol. 3, US Census Bureau, Washington, DC.

Andres J-D, Lorca P, Juez F-J, Lacheras F. 2011. Bankruptcy forecasting: a hybrid approach using fuzzy c-means clustering and multivariate adaptive regression splines. *Expert Systems with Applications* **38**(3): 1866–1875.

Bilson J. 1978. The monetary approach to the exchange rate: some empirical evidence. *IMF Staff Papers* **25**(1): 48–75.

Brandl B, Wildburger U, Pickl S. 2009. Increasing of the fitness of fundamental exchange rate forecast models. *International Journal of Contemporary Mathematical Sciences* **4**(16): 779–798.

Cheung Y-W. 1993. Long memory in foreign-exchange rates. *Journal of Business and Economic Statistics* **11**: 93–101.

CheungY-W CM, Pascual AG. 2005. Empirical exchange rate models of the nineties: are any fit to survive? *Journal of International Money and Finance* **24**(7): 1150–1175.

Cortes C, Vapnik V. 1995. Support-vector networks. *Machine Learning* **20**: 273–297.

Dickey D, Fuller W. 1981. Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica* **49**: 1057–1072.

Dornbusch R. 1976. Expectations and exchange rate dynamics. *Journal of Political Economy* **86**: 1161–1176.

Dunis C, Williams M. 2002. Modeling and trading the EUR/USD exchange rate: do neural network models perform better? *Derivatives Use, Trading and Regulation* **8**(3): 211–239.

Engle R, Granger C. 1987. Cointegration and error correction: representation, estimation and testing. *Econometrica* **55**: 251–294.

Fleming JM. 1962. Domestic financial policies under fixed and floating exchange rates. *IMF Staff Papers* 369–379.

Frankel JA. 1979. On the mark: a theory of floating exchange rates based on real interest differentials. *American Economic Review* **69**: 610–622.

Freidman J. 1991. Multivariate adaptive regression splines. *Annals of Statistics* **19**: 1–141.

Fu C. 2010. Forecasting exchange rate with EMD-based support vector regression, in *International Conference on Management and Service Science* (*MASS*), Wuhan, China.

Galbraith JW, Kısınbay T. 2005. Content horizons for conditional variance forecasts. *International Journal of Forecasting* **21**: 249–260.

Hannan EJ, Quinn BG. 1979. The determination of the order of an autoregression. *Journal of the Royal Statistical Society B* **41**: 190–195.

Härdle W, Lee W-J, Schäfer D, Yeh Y-R. 2009. Variable selection and oversampling in the use of smooth support vector machines for predicting the default risk of companies. *Journal of Forecasting* **28**(6): 512–534.

Hodrick RJ, Prescott EC. 1997. Postwar US business cycles: an empirical investigation. *Journal of Money, Credit, and Banking* **29**: 1–16.

Hou T, Yan M, Wu Z. 2009. A variant of the EMD method for multi-scale data. *Advances in Adaptive Data Analysis* **1**(4): 483–516.

Huang N-E, Wu S-R, Long S-S, Shen W-D, Qu P, Gloersen, Fan KL. 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London A* **454**: 903–993.

Ince H, Trafalis T. 2006. A hybrid model for exchange rate prediction. *Decision Support Systems* **42**(2): 1054–1062.

Jeong D, Kim Y. 2005. Rainfall-runoff models using artificial neural networks for ensemble streamflow prediction. *Hydrological Processes* **19**: 3819–3835.

Kao L-J, Chiu C-C, Lu C-J, Chang C-H. 2013. A hybrid approach by integrating wavelet-based feature extraction with MARS and SVR for stock index forecasting. *Decision Support Systems* **54**(3): 1228–1244.

Karemera D, Kim B. 2006. Assessing the forecasting accuracy of alternative nominal exchange rate models: the case of long memory. *Journal of Forecasting* **25**(5): 369–380.

Khandani AE, Kim AJ, Lo AW. 2010. Consumer credit-risk models via machine-learning algorithms. *Journal of Banking and Finance* **34**(11): 2767–2787.

Kourentzes N, Petropoulos F, Trapero J. 2014. Improving forecasting by estimating time series structural components across multiple frequencies. *International Journal of Forecasting* **30**(2): 291–302.
Lin C-S, Chiu S-H, Lin T-Y. 2012. Empirical mode decomposition-based least square support vector regression for foreign exchange rate forecasting. *Economic Modelling* **29**: 2583–2590.
Lucas RE. 1982. Interest rates and currency prices in a two-country world. *Journal of Monetary Economics* **10**: 335–360.
Macdonald R. 2010. Exchange Rate Economics: Theory and Evidence(1st edn). Routledge: New York.
Mark NC, Sul D. 2001. Nominal exchange rates and monetary fundamentals: evidence from a small post-Bretton Woods panel. *Journal of International Economics* **53**(1): 29–52.
Meese R, Rogoff K. 1983. Empirical exchange rate models of the seventies: do they fit out of sample? *Journal of International Economics* **14**: 3–24.
Moghtaderi A, Flandrin P, Borgnat P. 2011. Trend filtering: empirical mode decompositions versus L1 and Hodrick–Prescott. *Advances in Adaptive Data Analysis* **3**: 41–61.
Moghtaderi A, Flandrin P, Borgnat P. 2013. Trend filtering via empirical mode decomposition. *Computational Statistics and Data Analysis* **58**: 114–126.
Mohammadipour M, Boylan JE. 2012. Forecast horizon aggregation in integer autoregressive moving average (INARMA) models. *Omega* **40**(6): 703–712.
Molodtsova T, Papell DH. 2009. Out-of-sample exchange rate predictability with Taylor rule fundamentals. *Journal of International Economics* **77**: 167–180.
Mundell RA. 1968. International Economics. Macmillan: New York.
Öğüt H, Doğanay MM, Ceylan NB, Aktaş R. 2012. Prediction of bank financial strength ratings: the case of Turkey. *Economic Modelling* **29**(3): 632–640.
Papadimitriou T, Gogas P, Matthaiou M, Chrysanthidou E. 2015. Yield curve and recession forecasting in a machine learning framework. *Computational Economics* **45**: 635–645.
Phillips PCB, Perron P. 1988. Testing for a unit root in time series regression. *Biometrika* **75**: 335–346.
Rime D, Sarno L, Sojli E. 2010. Exchange rate forecasting, order flow and macroeconomic information. *Journal of International Economics* **80**: 72–88.
Rubio G, Pomares H, Rojas I, Herrera LJ. 2011. A heuristic method for parameter selection in LS-SVM: application to time series prediction. *International Journal of Forecasting* **27**(3): 725–739.
Sapankevych NI, Sangar R. 2009. Time series prediction using support vector machines: a survey. *IEEE Computational Intelligence Magazine* **4**(2): 24–38.
Schwarz GE. 1978. Estimating the dimension of a model. *Annals of Statistics* **6**(2): 461–464.
Sephton P. 2001. Forecasting recessions: can we do better on MARS? *Review*, *Federal Reserve Bank of St Louis* March; 39–49.
Silvestrini A, Veredas D. 2008. Temporal aggregation of univariate and multivariate time series models: a survey. *Journal of Economic Surveys* **22**(3): 458–497.
Stockman A. 1980. A theory of exchange rate determination. *Journal of Political Economy* **88**: 673–698.
Wu Z, Huang NE. 2009. Ensemble empirical mode decomposition: a noise-assisted data analysis method. *Advances in Adaptive Data Analysis* **1**(1): 1–41.
Wu Z, Huang NE, Long S, Peng C-K. 2007. On the trend, detrending and variability of nonlinear and nonstationary time series. *Proceedings of the National Academy of Sciences* **4**(38): 14889–14894.

Authors' biographies:

Vasilios Plakandaras is a PhD candidate at the Department of Economics, Democritus University of Thrace. His research interests include nonlinear dynamics, macroeconomics, econometrics, financial economics and machine learning. He has published in Applied Economics, Economic Modelling, Journal of Economic Asymmetries and others.

Theophilos Papadimitriou received the Diploma degree in mathematics from the Aristotle University of Thessaloniki, Greece, and the D.E.A. A.R.A.V.I.S (Automatique, Robotique, Algorithmique, Vision, Image, Signale) degree from the University of Nice-Sophia Antipolis, France, both in 1996 and the Ph.D. degree from the Aristotle University of Thessaloniki in 2000. In 2001, he joined the Department of Economics (International Economic Relations and Development Department until 2012), Democritus University of Thrace, Komotini, Greece, where, he served as a lecturer (2002–2008), assistant professor (2008–2013). Currently he holds the position of Associate Professor in the same department. Dr. Papadimitriou co-authored more than 80 journal papers, conference papers and book chapters. He served as a reviewer for various publications and as a member to scientific committees for Conferences and Workshops. Theophilos Papadimitriou current research interests include complex network, machine learning, and data analysis.

Periklis Gogas is an Associate Professor of Economic Analysis and International Economics at the Democritus University of Thrace and was recently a Visiting Scholar at the Ross School of Business of the University of Michigan. He received his Ph.D. degree from the University of Calgary, his Master's degree from the University of Saskatchewan and his B.A. from the University of Macedonia. His research interests include macroeconomics, financial economics, chaotic and non-linear dynamics and machine learning methods applied to macro and finance, published in the Journal of Money Credit and Banking, Journal of Banking and Finance, Physica A, Open Economies Review, etc.

Authors' addresses:

Vasilios Plakandaras, Theophilos Papadimitriou and **Periklis Gogas,** Department of Economics, Democritus University of Thrace, Komotini, Greece.

