A hybrid slantlet denoising least squares support vector regression model for exchange rate prediction

He, Kaijian; Lai, Kin Keung; Yen, Jerome

Published in:
Procedia Computer Science

Published: 01/01/2010

Document Version:
Final Published version, also known as Publisher's PDF, Publisher's Final version or Version of Record

Published version (DOI):
10.1016/j.procs.2010.04.270

Publication details:
He, K., Lai, K. K., & Yen, J. (2010). A hybrid slantlet denoising least squares support vector regression model for exchange rate prediction. *Procedia Computer Science*, *1*(1), 2407-2415. https://doi.org/10.1016/j.procs.2010.04.270

Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 1 (2012) 2397–2405

Procedia Computer Science

www.elsevier.com/locate/procedia

International Conference on Computational Science, ICCS 2010

A hybrid slantlet denoising least squares support vector regression model for exchange rate prediction

Kaijian He[a], Kin Keung Lai[b,*], Jerome Yen[c]

[a]*Department of Management Sciences, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong*
[b]*School of Business Administration, North China Electric Power University, Beijing 102206, China*
[c]*Department of Finance and Economics, Tung Wah College, Wylie Road, Kowloon, Hong Kong*

Abstract

Despite the active exploration of linear and nonlinear modeling of exchange rates, there is no consensus on the optimal forecasting model other than the traditional random walk and ARMA benchmark models in the literature. Given the increasing recognition of heterogeneous market structure, this paper proposes an alternative Slantlet denoising based hybrid methodology that attempts to incorporate the linear and nonlinear data features. The recently emerging Slantlet analysis is introduced to separate the linear data features as it constructs filters with varying lengths at different scales and has more appealing time localization features than the normal wavelet analysis. Meanwhile, the Least Squares Support Vector Regression (LSSVR) is used to model and correct for the empirical errors nonlinear in nature. As empirical studies were conducted in the Euro exchange rate market, the performance of the proposed algorithm was compared with those of benchmark models including random walk, ARMA and LSSVR models. The proposed algorithm outperforms the benchmark models. More importantly the proposed methodology explores and offers deeper insights as to the underlying data generating process.

Keywords: Slantlet analysis, Denoising algorithm, ARMA model, Random walk model, Least squares support vector regression model

1. Introduction

As one of the most important economic factors for the national economy in the increasingly open and globalized economic system, the accurate and reliable forecasting of the exchange rate movement has profound impacts throughout different levels of the economy. In practice, it affects various monetary and foreign exchange decision making process of the government. It also influences different operational decision making process and the hedging behaviors of individual firms. For academics, it is closely related to some important theoretical issues in the economics and finance fields such as Efficient Market Hypothesis (EMH) and the Derivative Pricing, etc.

Despite the significant interests received, exchange rate forecasting remains one of the most challenging issues in the field of economics and finance. Since the seminal work by Meese & Rogoff, traditional linear structural models

*Corresponding author. Tel.: +852-2788-8563.
Email address: mskklai@cityu.edu.hk (Kin Keung Lai)

doi:10.1016/j.procs.2010.04.270

have largely failed to demonstrate competent forecasting performance in empirical studies [1]. Part of the reason is attributed to the nonlinear data characteristics uncovered in the recent empirical studies, which shed light on the research of more accurate characterization and prediction of exchange rate movement. The use of nonlinear models is now perceived as the key to the further performance improvement of exchange rate forecasting models. This approach is mainly led by the adoption of different distribution free artificial intelligence techniques such as neural network and support vector regressions. However, they are mainly black box in nature and offer little insights into the underlying patterns as well as supporting theories behind. Recently the computational harmonic theory and the denoising theory have provided the alternatives as they both take data driven approach and have tractable theoretical foundations. Besides, they provide alternative view consistent with the emerging Heterogeneous Market Hypothesis (HMH), where heterogeneous agents are recognized market forces in contrast to the homogeneous agents assumption behind majority of time series models in the literature [2].

Therefore, this paper proposes a hybrid methodology, incorporating both linear and nonlinear data characteristics during the modeling process. The novelty is how these features are separated and modeled. The linear data feature is extracted using the Slantlet analysis recently introduced in the harmonic analysis field and modeled by Autoregressive Moving Average (ARMA) model. The nonlinear data feature, represented by the empirical forecast errors, is modeled by Least Squares Support Vector Regression (LSSVR). Therefore, the proposed Slantlet denoising based hybrid model improves the forecasting accuracy as a result. Empirical studies in the representative Euro market were conducted to support the superior performance of the proposed models against conventional benchmark models.

The contributions of this paper are documented in the following three aspects: Firstly, this paper argues that inappropriate or inadequate removal of noises could result in lower goodness of fit and previous failure of mainstream exchange rate models. The parameters for the denoising algorithm is critical to the performance of the proposed hybrid methodology and deserves special attentions. Secondly, to the best of our knowledge, there are few researches identified exploring the potential of Slantlet transform, an improved wavelet functions that employs more appealing features but receives relatively less attentions, in the forecasting literature [3, 4, 5]. This paper represents one of few attempts to explore application of Slantlet based denoising algorithm and use it as the basis for constructing hybrid algorithm to further improve the modeling accuracy in the field of exchange rate forecasting. Thirdly, previous hybrid approaches mostly uses nonlinear models to correct for the error specification of the linear models fitted to the original data directly [6]. However, the linear models may be biased during the fitting process as their assumptions are violated by the original data with mixture of both linear and nonlinear data features, which leads to distortions in the empirical errors calculated and further biases in the nonlinear models estimated. The proposed model in this paper removes the noises using Slantlet analysis before the linear models are fitted to improve the validity of linear models in this setting. The nonlinear LSSVR is used to model the nonlinear components from the empirical modeling errors. This is a paradigm shift that improves the fitting accuracy of different models and further the forecasting accuracy.

The rest of the paper proceeds as follows: section 2 covers relevant theories and literatures, including Slantlet analysis, time series models and LSSVR models. Section 3 proposes a novel hybrid Slantlet Denoisong Least Squares Support Vector Regression (SDLSSVR) methodology. Major findings and performance evaluations results are reported in section 4. Section 5 summarizes and provides some concluding remarks.

2. Relevant Theories

2.1. Slantlet Analysis

Slantlet transform is an orthonormal transform that defines a continuous function over L^2 space with shorter support and retains the same level of vanishing moment [7]. This is achieved by employing a filterbank approach than the traditional tree based approach, with filters of different lengths at different scales. The methodology gives it more flexibility in designing filters that target different data features, in the spirit of multi wavelet analysis. In the case of Slantlet analysis, the desirable feature is shorter support, which gives it the improved time frequency localization upon the Haar wavelet counterpart.

The filter banks in the Slantlet analysis is determined by solving variables (parameters) in (1)

$$g_i(n) = \begin{cases} a_{0,0} + a_{0,1}n \;\; for \;\; n = 0, .., 2^i - 1 \\ a_{1,0} + a_{1,1}n \;\; for \;\; n = 2^i, ..., 2^{i+1} - 1 \end{cases} \tag{1}$$
$$h_i(n) = \begin{cases} b_{0,0} + b_{0,1}n \;\; for \;\; n = 0, .., 2^i - 1 \\ b_{1,0} + b_{1,1}n \;\; for \;\; n = 2^i, ..., 2^{i+1} - 1 \end{cases}$$
$$f_i(n) = \begin{cases} c_{0,0} + c_{0,1}n \;\; for \;\; n = 0, .., 2^i - 1 \\ c_{1,0} + c_{1,1}n \;\; for \;\; n = 2^i, ..., 2^{i+1} - 1 \end{cases}$$

The solution to (1) is solved by subjecting (1) to several constraints. All three functions need to satisfy the unit form condition and orthogonality to their shifted time reverse condition. Both $g_i(n)$ and $f_i(n)$ functions need to annihilates linear discrete time polynomials [7].

2.2. Time Series and Least Squares Support Vector Regression Models

The random walk model is closely related to the Efficient Market Hypothesis. It is based on the notion that all past information is reflected in the current price, which is the only needed information to forecast future movement. Random walk is a special case of more general ARMA with (0,1,0) structure. The ARMA model is based on the autocorrelation data feature widely recognized in the financial market. Mathematically ARMA model is defined as in (2).

$$\phi(L)r_t = \mu + \theta(L)u_t \tag{2}$$
$$Where \quad \phi(L) = 1 - \phi_1 L - \phi_2 L^2 - ... - \phi_r L^r$$
$$\theta(L) = 1 + \theta_1 L + \theta_2 L^2 + ... + \theta_m L^m$$

Where both ϕ and θ are lag operators for the autoregressive coefficients and moving average coefficients respectively. μ is the white noise (innovation).

Least Squares Support Vector Regression (LSSVR) is the least squares version of Support Vector Regression, which is an emerging nonlinear regression model based on statistical learning theory. Compared to traditional neural network based approach, its solution is more stable and global optimum. This results from the adoption of the same structural risk minimization principle during the data training and learning process. Thus it would alleviate the overfitting and local minima issue with the traditional neural network models based on the empirical risk minimization principle [8]. Compared to the quadratic programming approach used in SVR, LSSVR solves a system of linear equations with equality constraints and achieves higher level of computational efficiency when dealing with large scale data.

Using kernel functions that satisfy the Mercer condition, the nonlinear data are mapped into higher dimensions using kernel tricks and a typical regression problem is formulated accordingly as in (3).

$$y = f(x) = sign[\omega^T \phi(x) + b, \phi : R^n \rightarrow F, \omega \in F] \tag{3}$$

Where $\phi(x)$ is the feature space mapped from the nonlinear input space x. ω and b are coefficients.

A slack variable e_i is introduced to model the estimation error, thus the LSSVR is formulated as in (4).

$$\min_{w,b,e} \;\; J_i(w, e) = \frac{1}{2}\omega^T \omega + \lambda \frac{1}{2} \sum_{i=1}^{l} e_i^2 \tag{4}$$
$$s.t. \;\; y_i = \omega^T \phi(x_i) + b + e_i, i = 1, ..., l$$

Where $e_i = (e_1; e_2; ...; e_l) \in R^l$, λ is the penalty parameters. It serves to control for two objectives, i.e. the minimization of estimation error and the function smoothness, during the optimization process.

The Lagrangian function for the dual problem is formulated as in (5).

$$L(w, b, e; a) = j(w, e) + \sum_{i=1}^{l} a_i(\omega^T \phi(x_i) + b + e_i - y_k) \quad (5)$$

Where $a = (a_1; a_2; ...; a_l)$ is the Lagrange multipliers. Optimality is achieved by solving the linear system obtained by differentiating L in (5) with the variable $\omega, b, \epsilon, \alpha$, as in (6).

$$\begin{aligned}
\frac{\partial L}{\partial \omega} &= 0 \rightarrow \omega = \sum_{i=1}^{l} \alpha_i \phi(x_i) \\
\frac{\partial L}{\partial b} &= 0 \rightarrow \omega = \sum_{i=1}^{l} \alpha_i = 0 \\
\frac{\partial L}{\partial \epsilon_i} &= 0 \rightarrow \omega = \sum_{i=1}^{l} = \lambda \epsilon_i, i = 1, ..., N \\
\frac{\partial L}{\partial \alpha_i} &= 0 \rightarrow \omega = \omega^T \phi(x_i) + b + \epsilon_i - y_i = 0, i = 1, ..., N
\end{aligned} \quad (6)$$

3. A Slantlet Denoisong Least Squares Support Vector Regression Hybrid Methodology

The Slantlet denoising hybrid methodology is based on the 'Divide and Conquer' principle. It uses the Slantlet analysis to disentangle the underlying Data Generating Processes (DGPs) in the time scale domain, consistent with the Heterogeneous Market Hypothesis (HMH). The complex problem of modeling mixture of linear and nonlinear data characteristics in the original time series data are divided into smaller issues with the help of orthonormal transformation and denoising techniques, which are conducted in a distribution free manner with data integrity retained.

SDLSSVR hybrid methodology involves the following key steps.

1. The deterministic data are extracted from the original time series using Slantlet analysis at the chosen decomposition level, as in (7).

$$r_{t-1} = r_{t-1,data} + r_{t-1,noises} \quad (7)$$

Where $r_(t-1)$, $r_{t-1,data}$ and $r_{t-1,noises}$ refer to the original, denoised and noises return data series respectively. During the denoising process, the Slantlet analysis is used to project the original data into several sub data series in the time scale domain. The coefficients at different scales from the Slantlet analysis are filtered using the shrinkage strategy. During the shrinkage process, the threshold level is calculated with the threshold selection strategy, when the boundaries between data and noises at different scales are set. The denoised data are then reconstructed from the projected data points.

There are different denoising strategies proposed in the literature. The dominant ones include Universal, Minimaxi and Steins Unbiased Risk Estimate (SURE) [9]. The difference among these threshold selection rules is the trade off set between smoothness and accuracy to achieve for the denoised data. E.g. the Universal threshold selection rule minimizes the noise levels given that the noises are normally distributed. The minimaxi threshold selection rules minimizes the function fitness criteria such as Mean Square Error (MSE). Shrinkage algorithm are also proposed to govern the noise removal strategies once the noises are separated following specific threshold selection strategies. There are mainly two approaches: hard and soft shrinkage methods [9]. Both hard and soft shrinkage methods remove the projected data points smaller than the threshold value for that scale. However, the soft shrinkage method subtracts the set threshold value from the remaining projected data points while the hard shrinkage method leaves them intact. The Mean Absolute Deviation (MAD) is a critical parameter to the calculation of threshold selection rules as it measures and controls the volatility level of noises. It is calculated as $\sigma_{mad} = \frac{median(|c_0|,|c_1|,...,|c_{2^{n-1}-1}|)}{0.6745}$, where c is the Slantlet coefficients [10]. There are two approaches to calculate MAD parameter, i.e. sln or mln. The sln estimates the threshold level using the coefficients at the first scale only while the mln estimate the threshold level in a level dependent manner [11].

2. The denoised deterministic data component is supposed to follow ARMA process. Parameters are estimated using Maximum Likelihood Estimation (MLE) method, as in (8).

$$\widehat{r_{t,data}} = \delta_t + \sum_{i=1}^{m} \phi_i r_{(t-i),data} + \sum_{j=1}^{n} \theta_j \varepsilon_{(t-j),data} \tag{8}$$

Where $\widehat{r_t, data}$ is the forecasted denoised data at time t, $r_{(t-i),data}$ is the lag m returns with parameter ϕ_i , and $\varepsilon_{(t-j),data}$ is the lag n residuals in the previous period with parameter θ_j. δ_i is the constant coefficient. During the model fitting process, the most parsimonious model orders are determined by minimizing the information criteria, i.e. Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) [12, 13].

3. The LSSVR model is used to model the innovation series, i.e. the empirical forecast errors, as in (9).

$$\widehat{r_{t,noises}} = \sum_{i=1}^{p} \omega_i r_{(t-i),noises} \tag{9}$$

$$Where \quad r_{t,noises} = r_t - \widehat{r_{t,data}}$$

Where $\widehat{r_{t,noises}}$ is the forecasted noises data at time t. $r_{\widehat{(t-i),noises}}$ is the lag p returns with parameters ω_i.

4. The final forecasts are aggregated from individual forecasts of both deterministic and noises data component, as in (10).

$$\widehat{r_t} = \widehat{r_{t,data}} + \widehat{r_{t,noises}} \tag{10}$$

4. Empirical Studies

4.1. Data Set and Descriptive Statistics

We use the data set containing 1494 daily observations for Euro against US dollar, from 1 July, 2002 to 24 March, 2008, which starts from the stabilization of Euro since its official introduction. The data set covers some turbulent time period with major global events such as the second gulf war, SARS, bird flu, etc. so that the predictive accuracy of the model is stress tested through different market scenarios for its generalizability. The data set for the original daily price quotes is prepared and obtained from Global Financial Data. Following convention in the literature, the data set is further divided into the training set and test set by 60-24-16 ratio. The first 60% data is used as the in-sample training set for ARMA model. The next 24% is reserved as the training set for LSSVR model. And the final 16% data, which amounts to statistically significant 239 out-of-sample observations, is reserved as the test set to evaluate the predictive accuracy of different models out-of-sample. As Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) Analysis suggest there are indications of uprising trend, the original daily observations are transformed into daily returns as in $y_t = ln\frac{p_t}{p_{t-1}}$. The descriptive statistics are listed in table 1.

Table 1: Descriptive Statistics and Statistical Tests

Data	r	$r_{denoisedData}$	r_{noises}
Mean	-1.8676	0	0.0001
SD	0.0062	0.0025	0.0047
Skewness	0.0748	-0.2916	0.0123
Kurtosis	3.3908	14.7504	2.7343
JB Test	0.0384	0	0.4468
BDS Test	0.1236	0	0.0049

There are some stylized facts from the descriptive statistics in table 1. It indicates that the Euro market is a relatively efficient market, with moderate level of standard deviation and kurtosis. The rejection of Jarque-Bera test of normality show that the market deviates from the normal distribution [14]. The acceptance of BDS test of independence suggests that prices in the market is independently distributed [15]. However, the statistical significance of the results of these two tests are weak and inconclusive. When both data and noises components are extracted from the original data, the descriptive statistics shows some interesting findings. The denoised data show significant leptokurtic, i.e. fat tail, behavior. The null hypothesis of both Jarque-Bera and BDS test are rejected at statistically significant levels. Meanwhile, the distribution of noises is normal, as indicated by the acceptance of the null hypothesis of BDS test at the medium level of statistical significance. However, the null hypothesis of BDS test is rejected, indicating that there is unknown serial correlations.

4.2. Experiment Results

Following convention in the literature, the performances of different models are evaluated based on statistical measure tested on the out-of-sample data set. These include Mean Square Error (MSE), Clark-West test of predictive accuracy, Pesaran-Timmermann test of directional predictive accuracy [16, 17, 18]. The null hypothesis for the Clark-West test is equal predictive accuracy of the benchmark and competing models [16, 17]. The null hypothesis of the Pesaran-Timmermann test is that the actual and predicted values of variables are independent of each other [18]. The benchmark models used include Random Walk model, ARMA and LSSVR models [1, 19].

The lag order for ARMA model is determined based on the AIC and BIC minimization principle. The lag order for ARMA model fitted for the original data series is (1,1). For Slantlet analysis, the model specification and parameters are determined using trail and error method as follows: the decomposition level is set to 9, the rolling window is set to 512, the lag order for the fitted ARMA model is (1,1), the threshold selection rule chosen is Minimaxi, the shrinkage algorithm chosen is hard. For LSSVR model, model specification and parameters chosen are determined using the Grid Search method as follows: For LSSVR fitted for the original data, the kernel chosen is Radial Basis Function (RBF) kernel, the penalizing parameter λ is 0.02905, the σ for RBF is 0.074445. For LSSVR fitted for the separated noises data, the kernel chosen is Radial Basis Function (RBF) kernel, the penalizing parameter λ is 0.3629, the σ for RBF is 0.8833

Table 2 lists the predictive accuracy for different models, in terms of both MSE and Clark-West test against ARMA and RW model. The subscript sln refers to the denoising algorithm with threshold estimated based on the first level coefficients. The subscript mln refers to the denoising algorithm with threshold estimated in a level dependent manner.

Table 2: Performance Comparison of Different Models

Models	$MSE_{\times 10^{-5}}$	CW^{ARMA}_{pvalue}	CW^{RW}_{pvalue}
ARMA	2.1618	N/A	0.5777
RW	2.1574	0.5777	N/A
LSSSVR	2.1490	0.0853	0.1547
SDARMA$_{mln}$	2.1649	0.2878	0.3566
SDLSSVR$_{mln}$	2.1393	0.0306	0.0732
SDARMA$_{sln}$	2.1554	0.1652	0.2041
SDLSSVR$_{sln}$	2.1327	0.0288	0.0572

Experiment results in table 2 show that the proposed Slantlet based hybrid approach, i.e. both SDLSSVR$_{mln}$ and SDLSSVR$_{mln}$ outperform the traditional benchmark models, in terms of forecasting accuracy measured by MSE. Compared to ARMA, RW and LSSVR model, SDLSSVR achieves lower MSE. The performance gap is statistically significant as suggested by the rejection of the null hypothesis for Clark-West test, i.e. the null hypothesis of equal predictive accuracy of benchmark models against competing models is rejected. The p value for the performance of both SDLSSVR$_{sln}$ and SDLSSVR$_{mln}$ against ARMA are less than the cutoff value 0.05 while their performance against RW model is near 0.05. Among models tested, SDLSSVR$_{sln}$ achieves the best performance. Due to the nested nature of the proposed model with the benchmark models, the performance improvement is attributed to the use of Slantlet analysis to separate the data from noises in a finer manner and the use of different model specifications, both linear and nonlinear, for separated data of different characteristics.

Meanwhile, although the Slantlet desnoising ARMA model have already led to the improved performance, the introduction of LSSVR into SDLSSVR algorithm to model the noises component lead to further performance improvement. It can be seen from experiment results in table 2 that SDARMA fails to beat the RW model while SDLSSVR performs much better and beats the RW model at statistically significant level. This implies that the denoising process through Slantlet transform alone failed to capture and passed over some nonlinear data into the filtered noises components. The proper modeling of these nonlinear data patterns is critical to the further performance improvement. Meanwhile, this approach is also motivated from the analysis on the descriptive statistics of noises data component. The descriptive statistics and Jarque Bera test of normality both support the normality assumptions of the noises distribution. However, the finding that the BDS test of independence is rejected at the statistically significant level suggests that there is unspecified serial correlation among data observations, which is nonlinear in nature.

Experiment results in table 2 also confirms that the performance of the conventional ARMA demonstrate inferior performance to the forecasting accuracy of the Random Walk based approach, as reported in the literature [1]. However, the superior performance of the ARMA model with appropriate denoising algorithm, i.e. SDARMA model, suggests that the inferior performance is caused by the violation of assumptions of ARMA by the noises contaminated data. The superior performance of SDLSSVR suggest that there are patterns in the exchange rate market to be exploited with the utilization of appropriate models for certain data features. E.g. the linear data pattern can be modeled with ARMA when nonlinear noises are properly removed. This observation casts doubts on the EMH and provide the supporting evidence to the alternative market hypothesis such as Homogenous Market Hypothesis (HMH) and Fractal Market Hypothesis (FMH), etc.

Meanwhile, we also investigate the model performance in forecasting the direction of exchange rate movement, which has significant implications in trading activities. Table 3 lists the directional predictive accuracy for different models, in terms of both predictive success ratio, i.e. Dstat, and the results of Pesaran-Timmermann test.

Table 3: Pesaran-Timmermann Directional Test

Models	Dstat $_{\%}$	PT
ARMA	54.39	0.2152
LSSVR	52.52	0.5000
SDARMA$_{mln}$	51.88	0.2740
SDLSSVR$_{mln}$	51.46	0.8699
SDARMA$_{sln}$	54.81	0.0708
SDLSSVR$_{sln}$	52.30	0.7366

Experiment results in table 3 show that the Slantlet denoising algorithm could lead to significant performance improvement in directional forecasting accuracy. The SDARMA$_{sln}$ improves the performance against ARMA model. The directional forecasting accuracy is statistically significant with p value less than 10%, where the null hypothesis for Pesaran-Timmermann test, i.e. actual and predicted values of variables are independent of each other, can be rejected. However, the other models only achieve inferior performance.

Interestingly, instead of the best performing model SDLSSVR$_{sln}$ in terms of predictive accuracy, the best performing model overall is SDARMA$_{sln}$, in terms of both level predictive accuracy and directional predictive accuracy. All other models including LSSVR, SDARMA$_{mln}$, SDLSSVR$_{mln}$ and SDLSSVR$_{sln}$ achieve inferior directional predictive accuracy. This suggests that the data may contain more complicated structure that merits further development of theories and techniques to understand.

Another interesting finding is that experiment results show that generally the performance of sln based models are superior to mln based ones. One possible explanation for the inferior performance of level dependent based denoising algorithm is that the thresholds set at higher scales are inappropriate and may jeopardize important data features. The thresholds for each scale are 0.0057, 0.0069, 0.0050, 0.0044, 0.0114, 0.0147, 0.0102, 0.0123 and 0.0044. It's obvious that the threshold for the first scale 0.0057 is significantly lower than thresholds at 4 higher scales, i.e. scale 5, 6, 7 and 8. Thus there are fewer coefficients at higher scales processed to remove noises using the threshold level estimated from first scale coefficients than those level dependent ones. The superior performance for sln based algorithm suggest that the denoising algorithm with threshold estimated on first scale coefficients is preferred, and imply that the threshold levels set at higher scales may be inappropriate. Therefore, we argue that the performance of the Slant denoising algorithm is sensitive to the estimation method of threshold level, which is less explored issue in the literature. The widely used parameters set such as level dependent threshold method mln may not always lead to superior performance when the basic distribution assumption of noise is violated in the higher scale domain.

5. Conclusions

In this paper, a novel Slantlet denosing Least Squares Support Vector Regression model has been proposed for exchange rate prediction. Following the HMH, we found that using Slantlet analysis to combine the linear modeling of fundamental trend and nonlinear modeling of forecast noises data series lead to a hybrid algorithm with superior performance against other alternatives. In the proposed SDLSSVR model, the denoising algorithm is used to separate the underlying Data Generating Process (DGP) of different characteristics, based on the fundamental noises and data distributional assumptions. The Slantlet analysis is used to conduct the denoising process in the time scale domain, incorporating the multi scale heterogeneous structure and the removal of noises nonlinear nature. The LSSVR model is used to correct for the error specifications during the forecasting process. The performance comparison results show that the proposed SDLSSVR algorithm outperforms the traditional benchmark models in terms of predictive accuracy while overall the proposed SDARMA achieves the best performance in terms of both level and directional predictive accuracy. As the proposed algorithm is nested with the benchmark models, the performance improvement can be attributed to the proposed hybrid strategy that extracts distinct data features and the improves the modeling accuracy for them as more appropriate model specifications are chosen with better behaved sub data set. In contrast to the results from the seminal work by Meese & Roguff, our results show that statistically significant evidence of out-of-sample predictability exists in the exchange rate market.

Acknowledgements

The work described in this paper was supported by a Research Grant of City University of Hong Kong (No. 9610058).

References

[1] R. Meese, K. Rogoff, Empirical exchange rate models of the seventies - do they fit outof sample, Journal Of Interantional Economics 14 (1-2) (1983) 3–24.
[2] S. Manzan, F. H. Westerhoff, Heterogeneous expectations, exchange rate dynamics and predictability, Journal of Economic Behavior & Organization 64 (1) (2007) 111–128.
[3] M. Maitra, A. Chatterjee, Hybrid multiresolution slantlet transform and fuzzy c-means clustering approach for normal-pathological brain mr image segregation, Medical Engineering & Physics 30 (5) (2008) 615–623.
[4] M. Maitra, A. Chatterjee, A slantlet transform based intelligent system for magnetic resonance brain image classification, Biomedical Signal Processing and Control 1 (4) (2006) 299–306.
[5] A. Chatterjee, M. Maitra, S. K. Goswami, Classification of overcurrent and inrush current for power system reliability using slantlet transform and artificial neural network, Expert Systems with Applications 36 (2) (2009) 2391–2399.
[6] L. A. Diaz-Robles, J. C. Ortega, J. S. Fu, G. D. Reed, J. C. Chow, J. G. Watson, J. A. Moncada-Herrera, A hybrid arima and artificial neural networks model to forecast particulate matter in urban areas: The case of temuco, chile, Atmospheric Environment 43:35 (2008) 8331–8340.
[7] I. Selesnick, The slantlet transform, Signal Processing, IEEE Transactions on 47 (5) (1999) 1304–1313.
[8] V. N. Vapnik, The nature of statistical learning theory, 2nd Edition, Statistics for engineering and information science, New York : Springer, 2000.
[9] D. L. Donoho, I. M. Johnstone, Adapting to unknown smoothness via wavelet shrinkage, Journal Of The American Statistical Association 90 (432) (1995) 1200–1224.
[10] D. B. Percival, A. T. Walden, Wavelet methods for time series analysis, Cambridge series in statistical and probabilistic mathematics, Cambridge University Press, Cambridge, 2000.
[11] S. R. Messer, J. Agzarian, D. Abbott,Optimal wavelet denoising for phonocardiograms, Microelectronics Journal 32 (12) (2001) 931 – 941.
[12] H. Akaike, A new look at the statistical model identification, Automatic Control, IEEE Transactions on 19 (6) (1974) 716–723.
[13] G. Schwarz, Estimating dimension of a model, Annals Of Statistics 6 (2) (1978) 461–464.
[14] C. Jarque, A. Bera, A test for normality of observations and regression residuals, International Statistical Review 55 (2) (1987) 163–172.
[15] W. A. Brock, D. A. Hsieh, B. D. LeBaron, Nonlinear dynamics, chaos, and instability : statistical theory and economic evidence, MIT Press, Cambridge, Mass., 1991.

[16] T. E. Clark, K. D. West, Approximately normal tests for equal predictive accuracy in nested models, Journal Of Econometrics 138 (1) (2007) 291–311.
[17] T. E. Clark, K. D. West, Using out-of-sample mean squared prediction errors to test the martingale difference hypothesis, Journal Of Econometrics 135 (1-2) (2006) 155–186.
[18] M. H. Pesaran, A. Timmermann, A simple nonparametric test of predictive performance, Journal Of Business & Economic Statistics 10 (4) (1992) 461–465.
[19] T. A. Knetsch, Forecasting the price of crude oil via convenience yield predictions, Journal Of Forecasting 26 (7) (2007) 527–549.