This paper presents financial time series forecasting with multistage wavelet transform (WT). First, the time series data is processed through WT with different mother wavelet functions to extract high frequency and low frequency coefficients. Later, standard particle swarm optimization (PSO) algorithm is utilized to find optimal regression models in order to predict future samples. Mean square error (MSE) is opted as cost function for PSO to find optimal coefficients of the regression model. This study further extended to various mother wavelet functions and their decomposition levels to investigate their impacts on time series prediction. These investigations help to data scientists for selection of process parameters and variables. Further, the impact of control parameters of PSO is also discussed to show the importance in the search mechanism especially in regression problems.

Keywords Linear regression · Particle swarm optimization · Wavelet transform

1 Introduction

To enhance the output results of the estimation problems in financial sector, proper forecasting is required. Since the data is large, traditional approaches are failed and even the application produce premature results. Further, it affects the planning, operation, scheduling of future pricing. Therefore, usage of intelligent approaches yields comparatively close forecasting [1]. These forecasting studies are popular in all domains of the engineering fields.

Forecasting models implemented based on direct time data may not provide an accurate prediction outputs with limited information in time domain [1]. To reduce the error between actual and forecast values, multiple features extraction from normal time data is required. For this purpose, time–frequency transformations (TFT) are better choice to extract wide variety of features. Further, artificial and swarm intelligent techniques are used for estimation [2]. In literature, several linear and nonlinear approaches are available to forecast/predict future samples with the help of available information. Linear regression models are widely used when the input data correlates with output [3–7]. A regression approach for learning linear time-invariant dynamic models from time-series data is presented in Ref. [3] to reduce the errors in between the actual and forecasted samples. Further, segmented linear regression model used in Ref. [4] on healthcare data. It fits complex linear sections to the time series. When a greater number of dependent variables exist to forecast the output variable, multiple linear regression (MLR) models are helpful. Such applications are available in literature. In Ref. [5], MLR approach is used to forecast the short-term load of electricity based on time intervals variation information as input variables. To identify linear regression and MLR model optimal coefficients to fit the data exactly, traditional mathematical approaches framed based on regression errors are not suited for the applications with vague data. Therefore, intelligent population search-based algorithms are applied to select the regression model coefficients to fit the time data. PSO algorithm is used for such
curve fitting models where the coefficients of models are achieved by PSO [6]. In case of multiple inputs, MLR models are framed with PSO [7] and invasive weed optimizer (IWO) [8] assistance. These algorithms iterative mechanism helps to find optimal coefficients of regression models. Apart from applicability, these models are simple and easily fit to any time data. For forecasting studies, artificial neural networks (ANN) are also extensively used in different domains. ANN and adaptive neuro-fuzzy inference systems (ANFIS) are applied in Ref. [9] for maximum flood daily flow and compared with MLR and multi nonlinear regression models (MNLR). This paper concluded that MNLR models produces better statistical measures (forecasting measures) than MLR, ANN and ANFIS. Similar forecasting studies using ANN is applied for currency exchange rate in Ref. [10]. Accuracy of prediction studies enhanced further with intelligent search algorithm in addition to ANN. Due to volatile nature of financial time values, wind speeds or stock market metrics, the weights of ANN architectures are tuned with PSO [11], genetic algorithm (GA) [12] and twin support vector regression (TSVR) [13] for time series forecasting,Later, kernel mapping and high-order fuzzy cognitive maps [14], and fuzzy-c-means [15] approaches are also applied for time series forecasting studies. Apart from ANN and fuzzy, machine learning techniques are also available in literature related to forecasting [16]. Support vector regression (SVR) is most preferable approach in forecasting as per reports [16]. In Ref. [17], chaos-based firefly algorithm assisted SVR is introduced to forecast financial data. Further, SVR is implemented to predict recent problems like COVID 19 etc. and found the results are accurate compared to earlier approaches [18, 19]. In SVR, parameter selection is important, and this selection can be made with populations search based techniques to enhance the outputs of prediction variables [20]. Twin support vector regression (TSVR) models are proposed in Ref. [21] to enhance the financial series prediction results over SVR. The SVR based models relatively produce good results over radial basis function (RBF) and back-propagation (BP) neural network models. Recently deep neural and convolution neural network mechanisms are using by researchers in the forecasting studies [22–25]. In Ref. [22], convolutional neural network (CNN)-long short-term memory (LSTM) architecture is implemented to predict the volatility of gold which falls under forecasting area. Further, a systematic review is presented in [23] which gives basic idea of application of deep neural networks (DNN). To apply DNN based models, methodologies and data representations are provided in Refs. [24] and [25]. The architectures and identification of precise parameters is challenging and need extensive investigations prior to apply. Such drawbacks can eliminate with simple approaches with swarm assistance algorithms.

In this paper, a simple prediction models are developed with a hybrid framework by using wavelet transform (WT) and PSO. Initially, the data is processed through WT to extract multiple features (approximated and detailed coefficients) and a regression model is developed with decomposed coefficients. The coefficients of the regression model are identified with PSO to fit the model to data using MSE performance measure. This architecture is simple and yields good results. Wide comparisons are provided with different WT decompositions using levels. The paper organized as follows: Sect. 2 includes time series data processing through WT followed by standard PSO for identification of fitting models with WT coefficients in Sect. 3. Complete simulation results along with comparisons are provided in Sect. 4. Finally, conclusions are presented in last section.

2 Time series data processing through wavelet transform

Wavelet transform (WT) enhance the time information into multi scale frequency information with advantage of the locality of the analysis. This WT is extensively used in numerous studies especially for image processing, detection, and classification problems [26, 27]. Recently, WT application found in forecasting and prediction studies such as wind speed and solar power forecasting, financial series forecasting and stock market analysis etc. Earlier, Fourier transform (FT) is used for extracting frequency components. The main drawback of FT is that it doesn’t provide time resolution. However, WT has both time and frequency resolutions. According to time frequency transformation concept, the time domain signal multiplied with transformation function over an interval converts it into wavelets. By using wavelets, function in time domain can be analyzed at various frequency levels of resolutions. The one-dimensional discrete wavelet transform is expressed as,

$$w_f(a,b) = \int_{-\infty}^{\infty} f(t) \cdot \phi_{a,b}(t) \cdot dt$$  \hspace{1cm} (1)

In Eq. (1), $f(t)$ is a time domain signal and $\phi_{a,b}(t)$ is the transformation function. In literature, several basis functions are available with specific features. In detail, $\phi_{a,b}(t)$ depends on mother wavelet family choice. These basis functions consist length of window, interval of frequency etc. as parameters which influences the process of
decomposition. For example, the signal \( f(t) \) decomposed using ‘Haar’ wavelet family for specific application. The process of extraction of detailed and approximation coefficients for further process is shown in Fig. 1. According to the Fig. 1, the reconstructed signal \( f(n) \) is given by

\[
f(n) = A_1 + D_1
\]

Further, the level-1 approximation coefficients are processed through WT provides level-2 and the process carried up to level-n. Mathematically,

\[
f(n) = A_2 + D_2 + D_1
\]

\[
f(n) = A_n + D_n + D_{n-1} + \ldots + D_2 + D_1
\]

On the other hand, over the time interval \((-\infty, +\infty)\), the parameters used in wavelet family are fixed. There are several wavelet families, decomposition takes place. In the process of decomposition, at each level \( l \), components of high and low frequencies of the signal are reanalyzed using the approximation coefficients \( a_l \) and the detail coefficients \( d_l \) decomposed with the help of low pass filter \( w_0 \) and high pass filter \( w_l \) and is expressed mathematically as in (5) and (6),

\[
a_l(r) = \sum_k w_0(l_k - 2k)a_{l-1}(l_k)
\]

\[
d_l(r) = \sum_k w_1(l_k - 2k)a_{l-1}(l_k)
\]

For time series, selection of mother wavelet and level of decomposition are key points when it is processed through WT. Based on literature survey, Daubechies family (db) is opted in this paper since the fine extraction of components of signal is possible with this basis function. However, the information extracted after data processing from n-level helps to identify correct prediction models in forecasting studies. In this paper, the models are framed based on different decomposition levels and conclusions are made from that function identified using PSO with WT coefficients.

### 3 Standard PSO for identification of fitting models with WT coefficients

After extracting approximated and detailed coefficients of WT with particular basis function, particle swarm optimization (PSO) algorithm is utilized to identify the mathematical function to fit the existing data and to predict the future samples. For this purpose, a linear equation with input and output variables is chosen. The optimal coefficients of these fitting models are obtained from PSO [28–31]. Therefore, number of decision variables of PSO depends on the level of decomposition and data size. For example, when single input information is taken for prediction then 2 decision variables are required for fitting the data with linear equation. When the same data is processed through WT and level-1 decomposed coefficient are utilized for fitting model, then the number of decision variables are increased to 3. This additional information extracted from WT enhances the accuracy of the model. In order to fit the data correctly to regression models, the coefficients of models are generated by PSO. PSO is a population search-based optimization algorithm where the initial solutions of the problem are generated randomly within the boundaries of the variables subjected to constraints. These initial solutions represent initial positions of the particles represented by \( p_i \). In this position \( n \) represents particle number and \( i \) represents iteration number and velocity of each particle is \( v_i \). After calculation of fitness value at each position, the local best position of each particle is represented by \( p_{best} \) and the overall global best among all positions is \( g_{best} \). For new solutions, the velocity of \( n \)th swarm is updated for \((i + 1)\)th iteration from the \( ith \) iteration is expressed as

\[
v_n^{i+1} = \omega v_n^i + c_1 r_1 (p_{best} - p_n^i) + c_2 r_2 (g_{best} - p_n^i)
\]

In Eq. (7), \( c_1 \) and \( c_2 \) are the acceleration constants, \( r_1 \) and \( r_2 \) are the randomly generated numbers for updating velocity in limits 0 and 1. The inertia weight factor is represented by symbol \( \omega \). Using (7) the new position is updated as

\[
p_n^{i+1} = p_n^i + v_n^{i+1}
\]

If the new position or solution of the swarm particle for a particular iteration yields best value than it’s previous, then the corresponding solution is updated and this process continues until the optimal position is achieved. The objective function for PSO is framed with the help of actual data and predicted data with fitting models.

---

Fig. 1 Decomposition of time signal using discrete wavelet transform
Fitness function \[ \text{Fitness function} = \frac{1}{N-k} \sum_{i=k}^{N} (y_{\text{pred}}(i) - y_{\text{act}}(i))^2 \] \hspace{1cm} (9)

Using (9), the solutions of PSO are driven towards their optimal values to fit actual data perfectly. However, the data is coefficients generated after WT processing.

### 4 Simulation results

The data is processed through WT with different db families to find the suitable dbN in level 1. As PSO is used for finding regression equations, several case studies are performed at different control parameter values of PSO to avoid premature solutions of regression models. In case, the actual and forecasting graphs along with regression equations at different inertia weights are presented along with errors between actual and predicted to show the applicability of the particular dbN. For this purpose, level-1 decomposition is carried by WT with db1/Haar basis function. Initially, the time data information is processed through WT and approximated and detailed coefficients are extracted known as \( A_1 \) and \( D_1 \). For forecasting, a multi linear regression model is considered as follows

\[ y(k+1) = \alpha A_1(k) + \beta D_1(k) + \gamma \] \hspace{1cm} (10)

These coefficients of regression model are identified by using PSO so that the error between predicted and actual information is minimized since this MSE is used as cost function and it should be minimized. The simulations are carried out in MATLAB for 2 data sets related to stock market available in the Yahoo financial website [21]. These data sets are processed through WT to extract various decomposed levels (based on frequency) and corresponding plots are presented in Fig. 2. For Fig. 2, 4-level decomposed coefficients are plotted along with actual data. Out of 2 data sets, overall comparisons are provided for first set and final optimum WT based results are provided for both sets. In all simulation studies, the extracted feature from actual data is used for data fitting with the help of PSO. For this purpose, population size of 50 and iterations of 100 are taken as common parameters of PSO. At \( \omega = 0.7 \), the coefficients obtained by PSO are 0.9955, 0.1650 and 2.2952. Figure 3a shows both actual and predicted data for the entire time scale and corresponding errors at each sample is plotted in Fig. 3b for data set 1.

When the basis function is changed to db2, the coefficients by keeping other parameters as same as db1, the coefficients of regression model are 0.9962, -0.1928 and 1.9607. Figure 4a shows both actual and predicted data for the entire time scale and corresponding errors at each sample is plotted in Fig. 4b. Compared to Fig. 2, the overall errors of predictions are reduced in case of db2.

Similar analysis is carried out for other basis functions of Daubechies family known as db3, db4, db5 and db6. In case, regression models are identified with data by fitting after extracting the level-1 detailed and approximated coefficients. Table 1 shows the regression coefficients obtained by PSO. For each function, the errors between actual and predicted data samples are plotted in Fig. 5. Among all these estimators obtained by PSO with different wavelet families, db4 is more suitable where the fitness value for entire data cycle is less compared to other families. This conclusion is based on both final fitness value identified using PSO and overall average errors of entire samples of the data.

The fitness values of objective function of PSO at the end of final iteration are plotted in Fig. 6. Instead of providing MSE, squared errors are presented in Fig. 6 for fair comparison. For each WT function, five inertia weights are tested since the global optimum values are achieved at a particular value of inertia weight. This analysis also avoids the prematurity of PSO. In this case, db4 yields better solution in terms of its fitness value evidence shown in Fig. 6.

Finally, the comparison between actual and predicted information is provided in Fig. 7. Together, db4 is the best choice for extraction of detailed and approximated coefficients in order to predict the time series data accurately and it is evident from fitness function values, error plots and statistical measures.

So far, the level-1 decomposition with different Daubechies family basis functions is considered to identify the best dbN for forecasting purpose. However, identification of optimal level at which decomposition of time series provides better prediction model is also significant. Therefore, the data is processed through WT with db4 function to decompose the frequency coefficients at various levels starting from 1 to 6. In each case the regression equations are identified by using PSO algorithm given by

\[ y = 0.9992A_2 - 0.4704D_1 + 0.5303D_2 + 0.5626 \] \hspace{1cm} (11)

\[ y = 1.0006A_3 - 0.4697D_1 + 0.5308D_2 + 0.8812D_3 - 0.1255 \] \hspace{1cm} (12)

\[ y = 1.0006A_4 - 0.4697D_1 + 0.5307D_2 + 0.8808D_3 + 0.9983D_4 - 0.1231 \] \hspace{1cm} (13)

\[ y = 1.0021A_5 - 0.4694D_1 + 0.5318D_2 + 0.8806D_3 + 0.9986D_4 + 0.9485D_5 - 0.8027 \] \hspace{1cm} (14)

\[ y = 1.0011A_6 - 0.4716D_1 + 0.5318D_2 + 0.8814D_3 + 0.9989D_4 + 0.9485D_5 + 1.0089D_6 - 0.3438 \] \hspace{1cm} (15)

For each level of decomposed coefficients, the fitness values achieved by PSO are presented in Fig. 8. When the
Fig. 2 Approximated and detailed coefficients of WT for (a) Data set 1, (b) Data set 2.

Fig. 3 Db1 assisted results, (a) Actual and predicted data set 1 with model shown in Eq. (10), (b) Errors.
number of decomposition levels increases then the overall relative errors decreases. However, this task increases number of regression model components. Therefore, level-3 is opted as marginal solution from which much change is not observed in fitness values. Further, data set 2 is processed through WT with db4 and extracted level-3 decomposed coefficients for regression models identification. Using optimized multistage WT-PSO model, time data and predicted data is presented in the Fig. 9. Errors corresponding to predictions are presented in Fig. 10. Similar to data set-1, the regression model obtained for set-2 provides better forecasting evident from the forecasting and error plots. The regression equation of the set is given by

\[ y = 0.9992A_3 - 0.3959D_1 + 0.4886D_2 + 0.8576D_3 + 0.1068 \]  

PSO algorithm fitness value (squared value) for this new data set is 1.7403 and therefore the errors are low compared to data-1 as shown in Fig. 10. Finally, the approach produces accurate models with less computational burden irrespective of nature of data, length, and other parameters of the data.

5 Conclusion

To improve the accuracy of the prediction, wavelet based PSO approach is proposed in this paper. Initially, the time data is processed through WT and extracted detailed and approximated coefficients are used for forecasting purpose. Since mother wavelet function and level of decomposition are two major factors in WT, extensive study is carried out in this paper for their selection. After identification of db4 and level-4 as accepted features for better forecasting, data sets are processed through WT and PSO based regression models are developed with detailed and approximated coefficients. To fit data correctly to MLR models with WT coefficients, PSO is used. This approach reduces computational burden and acceptable for N-D data. For tested data sets, average errors of all data points are in
between ± 2% shows the effectiveness of the proposed models. In future, other advanced signal processing techniques can be used for better feature extraction and to reduce the prediction errors. New meta heuristic algorithms are also useful to find the regression models with better convergence.
Fig. 7 Actual vs. predicted data analysis after identification of PSO-WT model

Fig. 8 Fitness function values for various levels of db4

Fig. 9 Actual and predicted data set 2 using final optimized WT-PSO approach
Fig. 10 Errors between actual and predicted values over entire time scale

**Funding** No funding is available.

**Conflict of interest** Authors declared that no conflicts of interest.

**References**

1. Altay E, Satman MH (2005) Stock market forecasting: artificial neural network and linear regression comparison in an emerging market. J Financ Manage Anal 18(2):18
2. Hurvich CM, Tsai CL (1989) Regression and time series model selection in small samples. Biometrika 76(2):297–307
3. Alaeddini A, Alemzadeh S, Mesbahi A, Mesbahi M (2018) Linear model regression on time-series data: non-asymptotic error bounds and applications. In: 2018 IEEE conference on decision and control (CDC). IEEE, pp 2259–2264
4. Valsamis EM, Husband H, Chan GK (2019) Segmented linear regression modelling of time-series of binary variables in healthcare. Comput Math Methods Med. https://doi.org/10.1155/2019/3478598
5. Amral N, Ozveren CS, King D (2007) Short term load forecasting using multiple linear regression. In: 2007 42nd International universities power engineering conference. IEEE, pp 1192–1198
6. Rao PS, Varma GP, Prasad CD (2020) Identification of linear and non-linear curve fitting models using particle swarm optimization algorithm. In: AIP Conference Proceedings, vol 2269, no 1, pp 030040. AIP Publishing LLC
7. Nagaraju TV, Prasad CD (2020) Swarm-assisted multiple linear regression models for compression index (Cc) estimation of blended expansive clays. Arab J Geosci 13(9):1–11
8. Nagaraju TV, Prasad CD, Murthy NGK (2020) Invasive weed optimization algorithm for prediction of compression index of lime-treated expansive clays. In: Soft computing for problem solving. Springer, Singapore, pp 317–324
9. Rezaeianzadeh M, Tabari H, Yazdi AA, Isik S, Kalin L (2014) Flood flow forecasting using ANN, ANFIS and regression models. Neural Comput Appl 25(1):25–37
10. Kamruzzaman J, Sarker RA (2003) Forecasting of currency exchange rates using ANN: a case study. In: Proceedings of the 2003 international conference on neural networks and signal processing, vol 1, no 793–797. IEEE
11. Pradeep Kumar D, Ravi V (2017) Forecasting financial time series volatility using particle swarm optimization trained quantile regression neural network. Appl Soft Comput 58:35–52
12. Hassan MR, Nath B, Kirley M (2007) A fusion model of HMM, ANN and GA for stock market forecasting. Expert Syst Appl 33(1):171–180
13. Khashei M, Hejazi SR, Bijari M (2008) A new hybrid artificial neural networks and fuzzy regression model for time series forecasting. Fuzzy Sets Syst 159(7):769–786
14. Yuan K, Liu J, Yang S, Wu K, Shen F (2020) Time series forecasting based on kernel mapping and high-order fuzzy cognitive maps. Knowl-Based Syst 206:106359
15. De Andrés J, Lorca P, de Cos Juez FJ, Sánchez-Lasheras F (2011) Bankruptcy forecasting: a hybrid approach using Fuzzy c-means clustering and Multivariate Adaptive Regression Splines (MARS). Expert Syst Appl 38(3):1866–1875
16. Shen S, Jiang H, Zhang T (2012) Stock market forecasting using machine learning algorithms. Department of Electrical Engineering, Stanford University, Stanford, pp 1–5
17. Kazem A, Sharifi E, Hussain FK, Saberi M, Hussain OK (2013) Support vector regression with chaos-based firefly algorithm for stock market price forecasting. Appl Soft Comput 13(2):947–958
18. Singh V, Poonia RC, Kumar S, Dass P, Agarwal P, Bhatnagar V, Raja L (2020) Prediction of COVID-19 coronavirus pandemic based on time series data using Support Vector Machine. J Discrete Math Sci Cryptogr 23:1583–1597
19. Sahoo BB, Jha R, Singh A, Kumar D (2019) Application of support vector regression for modeling low flow time series. KSCE J Civ Eng 23(2):923–934
20. Cao LJ, Tay FEH (2003) Support vector machine with adaptive parameters in financial time series forecasting. IEEE Trans Neural Networks 14(6):1506–1518
21. Gupta D, Pratama M, Ma Z, Li J, Prasad M (2019) Financial time series forecasting using twin support vector regression. PLoS ONE 14(3):e0211402
22. Vidal A, Kristjanpoller W (2020) Gold volatility prediction using a CNN-LSTM approach. Expert Syst Appl 157:113481
23. Sezer OB, Gadelek MU, Ozbayoglu AM (2020) Financial time series forecasting with deep learning: a systematic literature review: 2005–2019. Appl Soft Comput 90:106181
24. Chong E, Han C, Park FC (2017) Deep learning networks for stock market analysis and prediction: methodology, data representations, and case studies. Expert Syst Appl 83:187–205
25. Duan J (2019) Financial system modeling using deep neural networks (DNNs) for effective risk assessment and prediction. J Franklin Inst 356(8):4716–4731
26. Hsieh TJ, Hsiao HF, Yeh WC (2011) Forecasting stock markets using wavelet transforms and recurrent neural networks: an integrated system based on artificial bee colony algorithm. Appl Soft Comput 11(2):2510–2525
27. Maheswaran R, Khosa R (2012) Comparative study of different wavelets for hydrologic forecasting. Comput Geosci 46:284–295
28. Eberhart R, Kennedy J (1995) Particle swarm optimization. In: Proceedings of the IEEE international conference on neural networks, vol 4, pp 1942–1948

29. Pant M, Kumar S (2021) Particle swarm optimization and intuitionistic fuzzy set-based novel method for fuzzy time series forecasting. Granul Comput 7:285–303

30. Gundu V, Simon SP (2021) PSO–LSTM for short term forecast of heterogeneous time series electricity price signals. J Ambient Intell Humaniz Comput 12(2):2375–2385

31. Rao PS, Varma GP, Prasad C (2022) PSO-WT-based regression model for time series forecasting. In: Applied information processing systems. Springer, Singapore, pp 227–233