American Journal of Applied Mathematics and Statistics, 2019, Vol. 7, No. 1, 25-31
Available online at http://pubs.sciepub.com/ajams/7/1/4
Published by Science and Education Publishing
DOI:10.12691/ajams-7-1-4

Modified 1D Multilevel DWT Segmented ANN Algorithm to Reduce Edge Distortion

WMND Basnayake[1,*], MDT Attygalle[1], Liwan Liyanage-Hansen[2], KDW Nandalal[3]

[1]Department of Statistics, University of Colombo, Colombo, Sri Lanka
[2]School of Computing, Engineering and Mathematics, Western Sydney University, Western Sydney, Australia
[3]Department of Civil Engineering, University of Peradeniya, Peradeniya, Sri Lanka
*Corresponding author: nadeeka@stat.cmb.ac.lk

Received October 10, 2018; Revised November 28, 2018; Accepted January 01, 2019

Abstract In spite of the ability of Artificial Neural Network (ANN) to handle nonlinear relationships in data, there are instances where ANNs have not been able to predict accurately in the presence of non-stationarity. A novel algorithm that has the ability to treat the nonstationary and nonlinearity in a time series had been presented in [1]. This paper presents a modification done to the algorithm via addressing the edge distortion that arises in the real time execution. The proposed algorithm in [1] was named as "1D Multilevel DWT Segmented ANN Algorithm" where the modified algorithm presented in this paper will be called as "Denoised 1D Multilevel DWT Segmented ANN Algorithm".

Keywords: *edge distortion, nonlinear, non-stationary, wavelet, NAR-ANN*

Cite This Article: WMND Basnayake, MDT Attygalle, Liwan Liyanage-Hansen, and KDW Nandalal, "Modified 1D Multilevel DWT Segmented ANN Algorithm to Reduce Edge Distortion." *American Journal of Applied Mathematics and Statistics*, vol. 7, no. 1 (2019): 25-31. doi: 10.12691/ajams-7-1-4.

1. Introduction

The main intuition behind the 1D Multilevel DWT Segmented ANN Algorithm is building expert models to different data segments in a way that the forecasting error that materializes due to non-stationarity is reduced. Data segmentation is carried out using one dimensional multilevel discrete wavelet transformation. This transformation was used for the purpose as it returns a data vector of the same length of the input while capturing information at various resolution levels that addresses non-stationarity.

There are three main advantages of the 1D Multilevel DWT Segmented ANN Algorithm. First advantage is that the data are divided horizontally without losing the time domain information and separate expert models are built for each subseries wherein the final forecast will be the aggregated forecast from each expert model. This will introduce the ability of capturing information at various resolution levels, thus extracting even the hidden features. Secondly, each expert model, considers its lagged inputs so that the time dependent nature can also be contemplated. Thirdly, this is a one dimensional modeling approach where the forecasting is carried out using the lags of the response variable, without the use of any explanatory power of any other supporting variable. This aspect is somewhat related to deep learning concepts of the ANN as NAR-ANN considers feedback connections enclosing several layers of the network.

In the usual ANN model fitting procedure, actual values (as they are) of the test set will be used for calculating the errors in the test set. Since the wavelet procedure is subjected to decomposition of each actual value in the real time execution, the values in the test set need to be handled differently. However, when proposing the algorithm at first [1], the full length of the series was transformed at first and divided into train, validation and test sets. This strategy thereabout is a "hypothesized" approach whereas this paper considers a calibrated decomposition mechanism to incorporate the actual value of the new day to the decomposition. The main finding of this paper is the improvement done to the 1D Multilevel DWT Segmented ANN Algorithm coupling a de-noising strategy to reduce forecasting error that occurs due to the edge distortion of wavelet transformation when decomposing the series in the real time.

2. Literature

Non-stationary time series consist of events occurring for varying durations which can be ascertained through time segmentation [2]. Although artificial neural network (ANN) and fuzzy methods have been used extensively as useful tools for prediction of hydrological variables, dealing with nonstationary data through such methods have proved many drawbacks [3,4]. Thus, some mechanism need to be adopted that can improve the performance of ANN in the presence of non-stationarity. This can mainly be ascertained via certain transformations to the inputs of the ANN.

Various researchers in the context of hydrology have successfully applied discrete wavelet transformation based

ANNs ([5,6,7,8,9]). The initial approach that uses a horizontal divide and conquer approach using NAR-ANN in the entire context so far is published as our original work in [1]. The paper has not addressed the issue of edge distortion that is present in discrete wavelet transformation when executing the 1D multilevel DWT segmented ANN algorithm.

The wavelet transform is calculated as shifting the wavelet function in time along the input signal and calculating the convolution of them. In most practical applications, the signals of interest have finite support. As the wavelet gets closer to the edge of the signal, computing the convolution requires the non-existent values beyond the boundary ([10,11,12]). This creates boundary effects caused by incomplete information in the boundary regions. Thus, the results of wavelet transform in these boundary effects regions have questionable accuracy. Actually, the particular impacts of boundary effects become increasingly significant for some systems that may possess longer period sequence and thus require higher frequency resolutions [13].

This is called as edge distortion and can affect the accuracy of 1D multilevel DWT segmented ANN algorithm in [1]. The circumstance is such that the algorithm uses separate NAR-ANN models to the sub series resulted from discrete wavelet decomposition and aggregate the forecasts from each ANN model to produce the final forecast. Since the NAR-ANN model uses the lags of the same variable to forecast, the accuracy of the forecast heavily depend on the most recent observations. In that regard the sub series associated in the 1D multilevel DWT segmented ANN algorithm, may act differently in the test set as opposed to the training set. Therefore, remedial measures need to be taken in order to reduce the adverse effect on the model performance due to edge distortion. This paper presents one such mechanism to modify the 1D multilevel DWT segmented ANN algorithm in a way that the adverse effect of edge distortion to the model performance is reduced.

3. Methodology

In order to make adjustments to reduce the effect of edge distortion, the boundaries should be treated differently from the other parts of the signal. According to [12], two alternatives to deal with boundary effects can be found. The first one is to accept the loss of data and truncate those unfavorable results at boundaries after convolution between signal and wavelet. However, simply neglecting these regions in analysis yields to a considerable loss of data which is not allowed in many situations where the edges of the signal contain critical information. The other is making artificial extensions of the boundaries before processing signals. In fact, there is another approach that employs the usual wavelet filters for the interior of the signal and constructs different boundary wavelets at the ends of the signal.

This paper considers a method where a truncating mechanism is applied to the time series in the form of wavelet de-noising. The paper has proven effective in forecasting the daily catchment flow to the Kotmale reservoir in Mahaweli river basin. It should be noted that the series considered is a highly volatile nonlinear and non-stationary time series. The modification proposed to the 1D multilevel DWT segmented ANN algorithm is indicated in bold font in the algorithm presented in Figure 1.

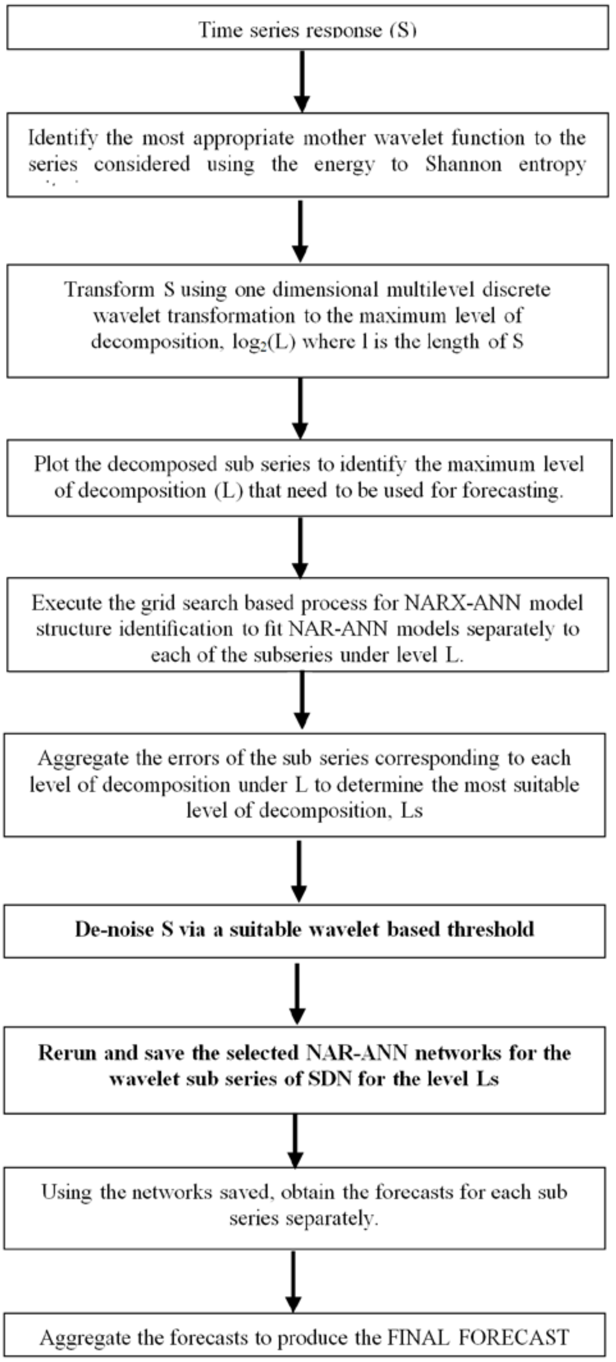

Figure 1. De-noised 1D multilevel DWT segmented ANN algorithm

In wavelet de-noising, the noise series is more generally represented as the following one dimensional model (Equation 1).

$$S(n) = f(n) + \sigma e(n). \tag{1}$$

In practice, S(n) is often a discrete time signal with equal time steps corrupted by additive noise and the attempts are being taken to recover that signal. Thus, in the time series context, the de-noising equation can be viewed as a T-dimensional random vector as given in Equation 2.

$$\begin{bmatrix} f(0)+\sigma e(0) \\ f(1)+\sigma e(1) \\ f(2)+\sigma e(2) \\ . \\ . \\ . \\ f(T-1)+\sigma e(T-1) \end{bmatrix} = \begin{bmatrix} f(0) \\ f(1) \\ f(2) \\ . \\ . \\ . \\ f(T-1) \end{bmatrix} + \begin{bmatrix} \sigma e(0) \\ \sigma e(1) \\ \sigma e(2) \\ . \\ . \\ . \\ \sigma e(T-1) \end{bmatrix}. \quad (2)$$

The general wavelet-based method for denoising and nonparametric function estimation is to transform the data into the wavelet domain, threshold the wavelet coefficients, and invert the transform. The simple steps to be followed in wavelet de-noising method are as stated below.

- Choose a wavelet and a level N. Compute the wavelet decomposition of the signals down to level N.
- For each level from 1 to N, threshold the detail coefficients.
- Compute wavelet reconstruction using the original approximation coefficients of level N and the modified detail coefficients of levels from 1 to N.

The threshold stated in the above step 2 can take different values depending on the situation. The widely used threshold rules are as stated below.

- Rigrsure-uses for the soft threshold estimator a threshold selection rule based on Stein's Unbiased Estimate of Risk (Stein, 1981) (quadratic loss function). An estimate of risk for a particular threshold value t is taken initially. A selection of threshold values can be taken by minimizing the risks in t.
- Minimax-uses a fixed threshold chosen to yield minimax performance for mean square error against an ideal procedure. The minimax principle which is used in statistics to design estimators is applied here. Since the denoised signal can be assimilated to the estimator of the unknown regression function, the minimax estimator is the option that realizes the minimum, over a given set of functions, of the maximum mean square error.
- Sqtwolog-uses a fixed form threshold yielding minimax performance multiplied by a small factor proportional to log (length (s)).
- Heursure-uses a mixture of the two options" rigrsure" and" Sqtwolog". Due to the same reason, if the signal-to-noise ratio is very small, the SURE estimate is very noisy. Under such circumstance, the fixed form threshold will be used.

Two other shrinkage rules can be applied for the thresholding depending on the requirement. The simplest scheme is hard thresholding. Let T denote the threshold and x your data. The hard thresholding is,

$$n(x) = \begin{cases} x & |x| \geq T \\ 0 & |x| < T \end{cases}. \quad (3)$$

The soft thresholding is,

$$n(x) = \begin{cases} x - T & x > T \\ 0 & |x| < T \\ x + T & x < -T \end{cases}. \quad (4)$$

4. Application

The effect of reducing the adverse effects of edge distortion to the application 1D Multilevel DWT segmented NAR-ANN algorithm is analyzed here in order to modify the algorithm. The algorithm was applied to forecast to the daily catchment flow to the upmost reservoir in the Mahaweli cascaded system in Sri Lanka. Daily catchment flow data for 21 years were used with a train, test, and validation split of 2: 1: 1.

Exploring the Pattern of Edge Distortion in the Daily Catchment Flow.

First and foremost, the nature of the edge distortion in wavelet decompositions was observed. For example, it was assumed that the last observation in the time sequence is the 6080th (i.e. 25th July 2011). This is the first observation considered as the edge. The wavelet transformation was then carried out taking observations only up to the 6080th and this was compared with the wavelet transformations obtained considering the full length of the series (7670). See Figure 1, Figure 2 and Figure 3 for the illustration of the edge distortion. Note that the addition of approximation 2, detail 1 and detail 2 constitutes the original series pf daily catchment flow.

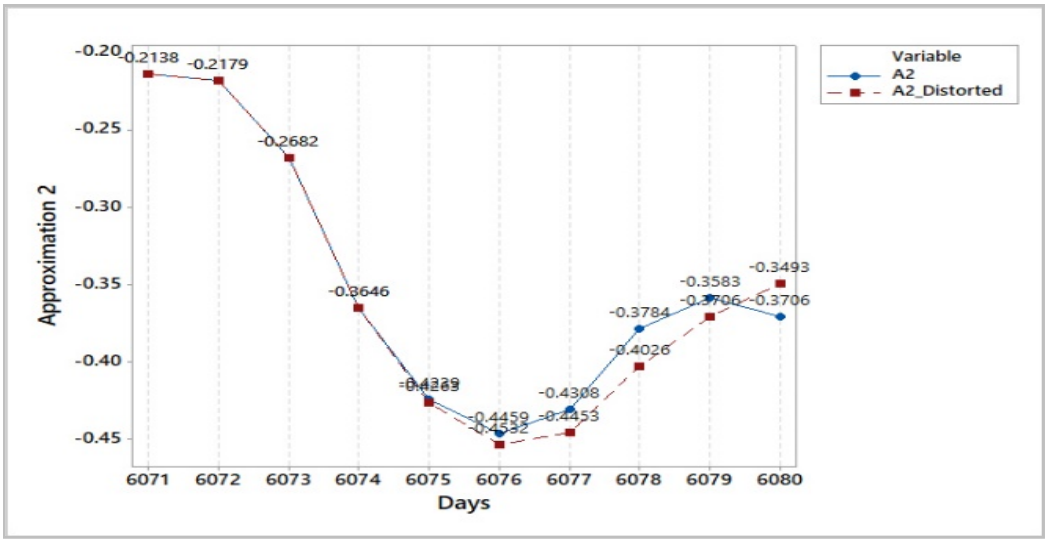

Figure 2. Random sample of Approximation 2 with Edge Distortion

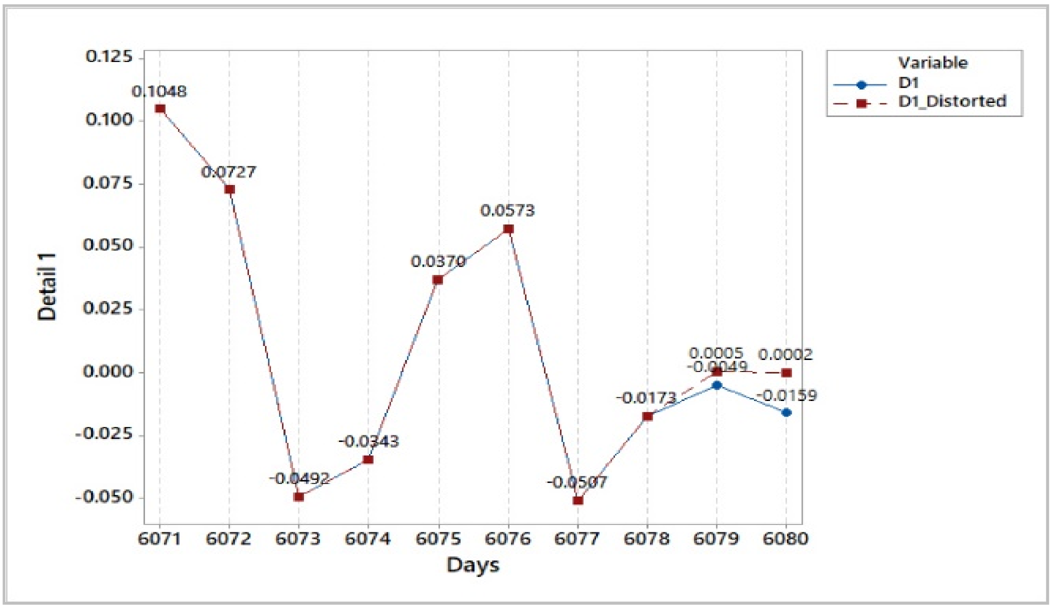

Figure 3. Random sample of Detail 1 with Edge Distortion

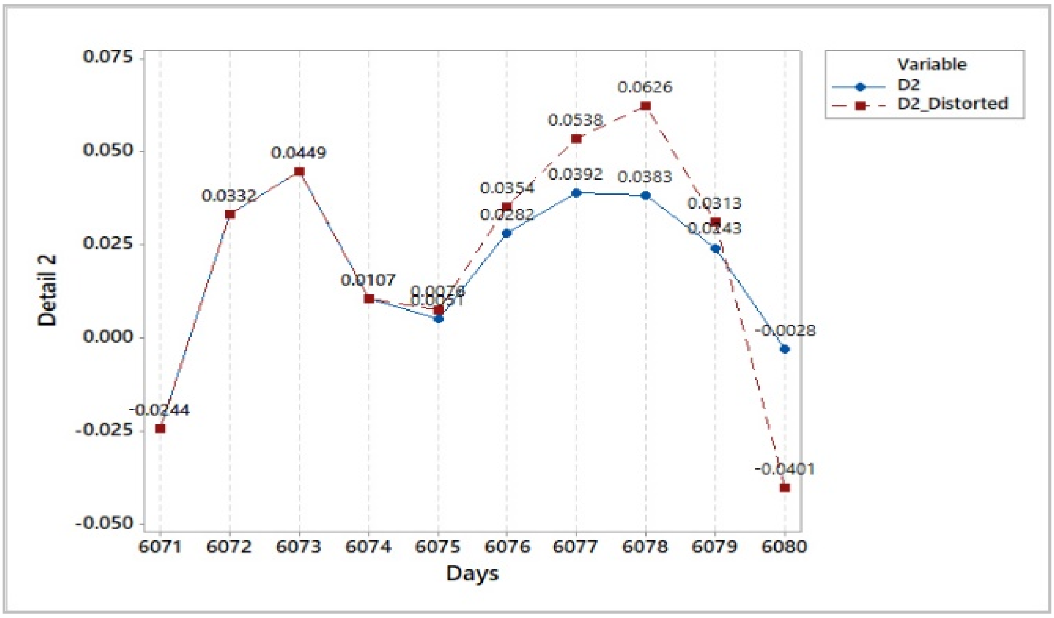

Figure 4. Random sample of Detail 2 with Edge Distortion

4.1. De-noising the Standardized Daily Catchment Flow

Several thresholds of de-noising for wavelet transformation were considered in order to see the effect of de-noising in reducing the forecasting error of the algorithm in [1]. The thresholds considered were fixed form, heuristic sure, minimax and rigorous sure. For the first four types of thresholding, there are three different forms with respect to rescaling using no rescaling and the two rescaling methods denoted by ”sln” and ”mln”. Altogether 15 types of thresholds were used.

For each form of de-noising, soft form thresholding was used, as the hard thresholding produces de-noised series that are not smooth. For the wavelet based de-noising, it is needed to determine the suitable mother wavelet and the level of decomposition. As such, the same mother wavelet function, “biorthogonal 3.1” and the level, 2, that was used in the discrete wavelet transformation in [1] will be considered here. Note the proper justifications in using these parameters are presented in [1].

Figure 5, Figure 6, Figure 7 & Figure 8 display 12 forms of de-noising as stated above. Note that the first 50 observations of the series are plotted for illustration. Observing the Figure 5, Figure 6,Figure 7 & Figure 8, it can be seen that the series de-noised using “Fixed form” threshold, “Minimax” threshold and using every other threshold with no rescaling have shredded the noise at a very high level. All these series in general have not considered detail components to the de-noised series, but rather have resulted in only the approximation series. Moreover, irrespective of the type of the threshold used, wavelet de-noising based on, no rescaling (one) has also produced similar de-noising patterns to that of the previously mentioned two thresholds, namely fixed form and minimax. As such, 9 threshold types are disregarded due to the excessive information loss in the de-noising process being unnecessarily high.

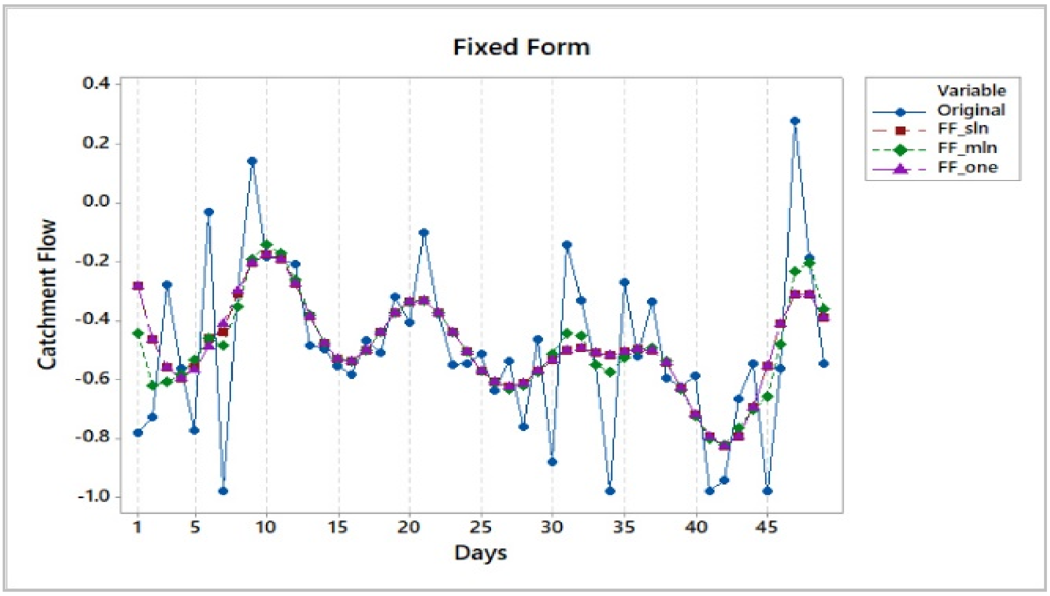

Figure 5. De-noising based on Fixed Form threshold

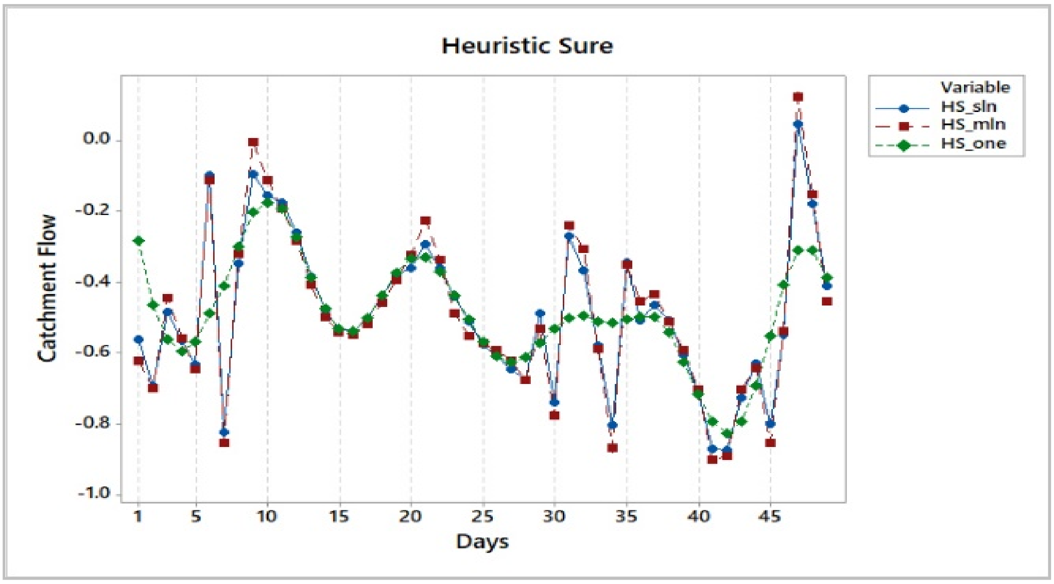

Figure 6. De-noising based on Heuristic Sure threshold

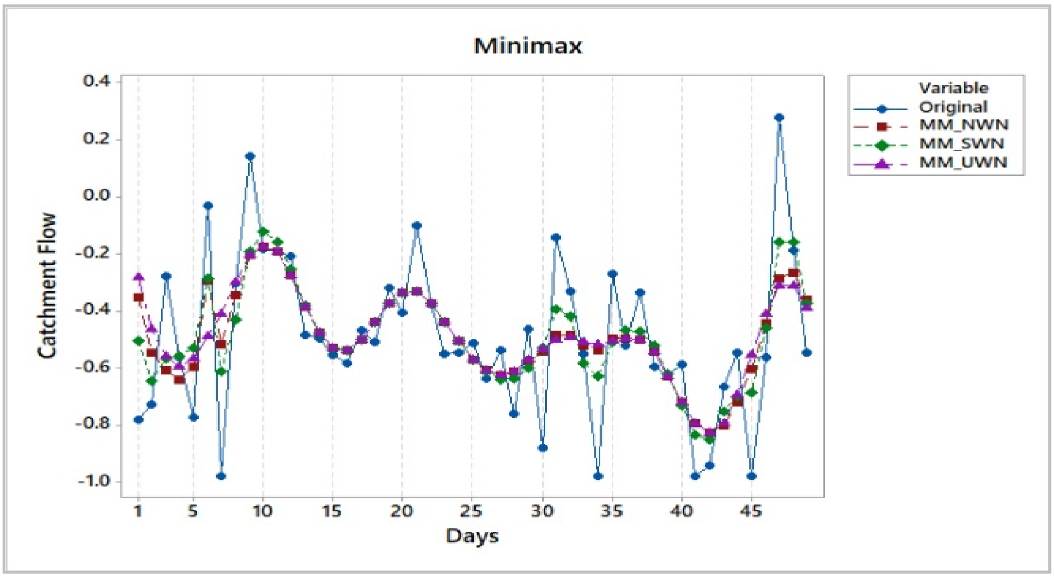

Figure 7. De-noising based on Minmax threshold

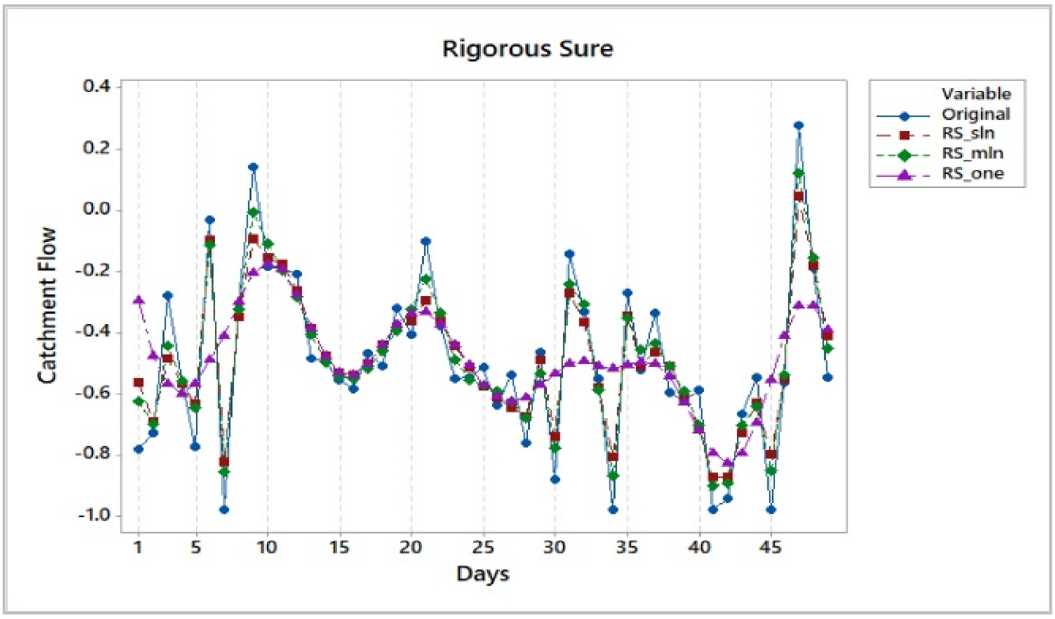

Figure 8. De-noising based on Rigorous Sure threshold

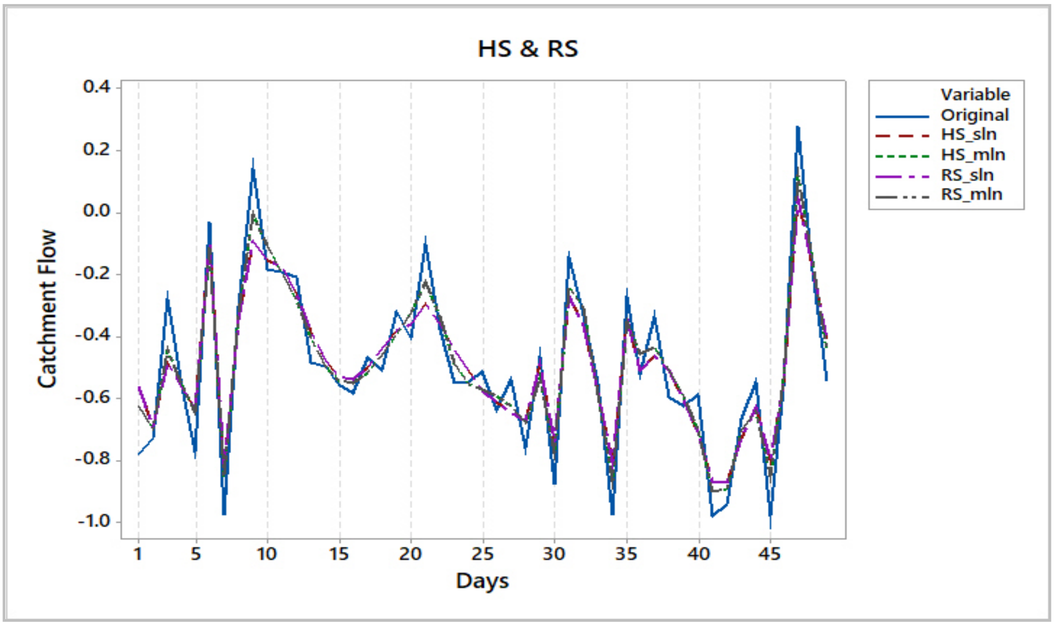

Figure 9. De-noising based on "Heuristic Sure-HS" and "Rigorous Sure-RS" thresholds for two rescaling types "sln" and "mln"

Rescaling based on "sln" seems to have produced somewhat similar results across the other two threshold types, "Heuristic Sure" and "Rigorous Sure". The remaining rescaling type "mln" was also observed as somewhat similar across the said two threshold types. Therefore, further graphical analysis for comparison were done for those threshold types to select the best suited form of de-noising. Figure 9 displays the corresponding plot.

It is observed in Figure 9 that the red line (HS_sln) overlaps with the purple line (RS_sln) and the green line (HS_mln) over laps with the black line (RS_mln). Thus, the comparison indicates that de-noising based on the two forms "sln" and "mln" using one type of thresholding will be adequate. Overlapped de-noised series depicts the fact that the two threshold types act same across each rescaling type. Thus, taking any threshold type out of the two seems adequate. Thus, the Heuristic Sure threshold was selected as for the purpose of de-noising.

4.2. Forecast Comparison for Wavelet Based NAR-ANN for Daily Catchment Flow

The forecasting error of the model that uses a wavelet transformation after a de-noising procedure was compared with that of the model with raw data (See Table 1). Model C was observed as the best, and model B was observed with a very similar and comparable performance to model C. However, the most suitable form of thresholding can be identified as "De-noised-Heuristic Sure (mln)".

Table 1. Performance comparison of the wavelet model formulations for one step ahead forecasts for three months

Model	Description	MAPE %
A	Raw data (Base)	33.78
B	De-noised-Heuristic Sure (sln)	31.05
C	De-noised-Heuristic Sure (mln)	29.79

5. Conclusion & Future Prospects

The analysis revealed that the most suitable wavelet threshold for de-noising is the "Heuristic Sure" threshold that uses level dependent estimation of level noise. Thus, the modified algorithm proposed in Section 3 can be applied with the selected threshold at the phase of de-noising. There is enough evidence in this paper to confirm the validity of the De-noised 1D Multilevel DWT Segmented ANN Algorithm. This algorithm is proven effective in forecasting nonlinear and nonstationary time series. It is worth mentioning that the accuracy of the modified algorithm can be further increased by exploring some other theoretical improvements proposed by various researchers. Especially, the work presented by [12] and [14] seem to be very promising in reducing edge distortion, but have not been the focus in this paper.

References

[1] Basnayake, W. M. N. D., Attygalle, D., & Hansen, L. L., Nandalal, K. D. W. (2017). Wavelet Based Nonlinear Autoregressive Neural Network to Predict Daily Reservoir Inflow. 1st International Conference on Machine Learning and Data Engineering (iCMLDE) (pp.69-75).

[2] Behnia, N., & Rezaeian, F. (2015). Coupling wavelet transform with time series models to estimate groundwater level. Arabian Journal of Geosciences, 8(10), 8441-8447.

[3] Cannas, B., Fanni, A., See, L., & Sias, G. (2006). Data preprocessing for river flow forecasting using neural networks: wavelet transforms and data partitioning. Physics and Chemistry of the Earth, Parts A/B/C, 31(18), 1164-1171.

[4] Cannas, B., Fanni, A., Sias, G., Tronci, S., & Zedda, M. K. (2005). River flow forecasting using neural networks and wavelet analysis. In Geophys. Res. abstr (Vol. 7, p. 08651).

[5] Adamowski, J., & Sun, K. (2010). Development of a coupled wavelet transform and neural network method for flow forecasting of non-perennial rivers in semi-arid watersheds. Journal of Hydrology, 390(1-2), 85-91.

[6] Kisi, O., & Cimen, M. (2011). A wavelet-support vector machine conjunction model for monthly streamflow forecasting. Journal of Hydrology, 399(1-2), 132-140.

[7] Okkan, U. (2012). Wavelet neural network model for reservoir inflow prediction. Scientia Iranica, 19(6), 1445-1455.

[8] Ramana, R. V., Krishna, B., Kumar, S., & Pandey, N. (2013). Monthly rainfall prediction using wavelet neural network analysis. Water resources management, 27(10), 3697-3711.

[9] Wei, S., Yang, H., Song, J., Abbaspour, K., & Xu, Z. (2013). A wavelet-neural network hybrid modelling approach for estimating and predicting river monthly flows. Hydrological Sciences Journal, 58(2), 374-389.

[10] Cohen, A., Daubechies, I., & Vial, P. (1993). Wavelets on the interval and fast wavelet transforms. Applied and computational harmonic analysis.

[11] De Queiroz, R. L. (1992). Subband processing of finite length signals without border distortions. In Acoustics, speech, and signal processing, 1992. Icassp-92., 1992 ieee international conference on (Vol. 4, pp. 613-616).

[12] Su, H., Liu, Q., & Li, J. (2012). Boundary effects reduction in wavelet transform for time frequency analysis. WSEAS Transactions on Signal Processing, 8(4), 169-179.

[13] Addison, P. S. (2002). Illustrated wavelet transform handbook. Introductory theory and applications in science, engineering, medicine and finance/paul s. addison. Inst. of Phys. Publ., Philadelphia, Pa.

[14] Williams, J. R., & Amaratunga, K. (1997). A discrete wavelet transform without edge effects using wavelet extrapolation. Journal of Fourier analysis and Applications, 3(4), 435-449.