Hybridization of Haar Wavelet Decomposition and Computational Intelligent Algorithms for the Estimation of Climate Change Behavior

Haruna Chiroma[1], Sameem Abdulkareem[1], Adamu I. Abubakar[2], Eka Novita Sari[3], Tutut Herawan[4], and Abdulsalam Ya'u Gital[5]

[1] Department of Artificial Intelligence, University of Malaya
50603 Pantai Valley, Kuala Lumpur, Malaysia
[2] Department of Information System
International Islamic University
Gombak, Kuala Lumpur, Malaysia
[3] AMCS Research Center, Yogyakarta, Indonesia
[4] Department of Information Systems, University of Malaya
50603 Pantai Valley, Kuala Lumpur, Malaysia
[5] Department of Computer Science
University of Technology Malaysia
Kampus Skudai, Johor Baru, Malaysia
{Freedonchi,asgital}@yahoo.com, 100adamu@gmail.com,
{sameem,tutut}@um.edu.my, eka@amcs.co

Abstract. We propose a hybrid of haar wavelet decomposition, relevance vector machine, and adaptive linear neural network (HWD-RVMALNN) for the estimation of climate change behavior. The HWD-RVMALNN is able to improve estimation accuracy of climate change more than the approaches already discussed in the literature. Comparative simulation results show that the HWD-RVMALNN outperforms cyclical weight/bias rule, Levenberg-Marquardt, resilient back-propagation, support vector machine, and learning vector quantization neural networks in both estimation accuracy and computational efficiency. The model proposes in this study can provide future knowledge of climate change behavior. The future climate change behavior can be used by policy makers in formulating policies that can drastically reduce the negative impact of climate change, and be alert on possible consequences expected to occur in the future.

Keywords: Haar Wavelet Decomposition, Relevance Vector Machine, Adaptive Linear Neural Network, Climate Change.

1   Introduction

In the last ten (10) to fifteen (15) years, the greater portion of the world experiences an unprecedented change in the temperature of sea water. For instance, the ice covering the Arctic is dramatically vanishing; the glaciers and green ice cap are

observed to be significantly melting; reduction in the volume of Antarctic ice sheets; unusual increase of sea level, and the seas are getting stormier; upsurge in the intensity, and frequency of rainfall due to precipitation variability; earlier occurrence of springtime, and high intensity of hurricane [1]. In addition, the world is experiencing an extraordinary warming. These changes experience throughout the globe are consequences of the climate change [2]. There is a growing report of numerous natural disasters triggered by climate change [3]. Therefore, the future knowledge of climate change is required for proper planning in order avoid the negative impact of natural disasters expect to be caused by the behavior of climate change. In the literature, several studies have been conducted using neural networks to predict climate change. For example, [4] proposed a neural network to build a model for predicting weather in Jordan. Goyal *et al*. [5] uses neural network to develop a downscaled model for predicting maximum and minimum temperature across 14 stations in Upper Thames River Basin Ontario, Canada. Holmberg *et al*. [6] proposed neural network to model the concentration of organic carbon, Nitrogen and phosphorus in a runoff stream water collected from Finland. The model is then used to predict future fluxes under climatic change. Tripathi [7] applied support vector machine (SVM) for statistical downscaling of precipitation. Subsequently, the SVM model is used to predict future climate changes in India. The results obtained from the SVM model was compared to the neural network and it was found that the SVM model performs better than the neural network. However, the studies are limited to the region where the research data were conducted.

The studies in the literature for predicting climate change using computational intelligent techniques, mainly focus on individual intelligent technique (IITs) whereas it is well known that hybrid intelligent technique is superior to the IITs when properly design. Hybrid techniques capitalized on their strengths to eliminate their weaknesses to build a synergistic model for effective prediction. Haar wavelet decomposition (HWD) has the ability to project data into time scale domain, and perform multi scale analysis to unveil useful hidden patterns in the historical data [8]. The sensitivity of the relevance vector machine (RVM) to parameters is less compared to the SVM, and RVM is faster than SVM [9]. We propose to apply RVM to different scales coefficients produce by HWD to predict global climate change in order to reduce estimation bias that might be introduced by the HWD. Adaptive linear neural network (ALNN) will be used to ensemble the results produce by each RVMs to produce more accurate ensemble result than the results produce by individual RVMs. We intend to investigate whether the propose hybrid of HWD, RVM and ALNN can yield superior novel results than the approaches in the literature.

2   Essential Rudiments

2.1   Haar Wavelet Decomposition

The HWD can be considered on the basis of the real line (IR) [10]: Let consider $L_2(IR)$ to be the space of complex value function (f) on IR in such a way that

$$\|f\| = \left(\sqrt{\int_{-\infty}^{\infty}|f(x)|^2 dx}\right) < \infty . \tag{1}$$

Eq. (1) is donated with scalar product, as such we have

$$f(f,g) = \int_{-\infty}^{\infty} f(x)g\overline{(x)}dx. \tag{2}$$

where the complex conjugate of $g(x)$ is given by $\overline{g(x)}$. If $(f,g) = 0$ then $f, g \in L_2(IR)$ can be definedl to each other, the functions can be define as

$$\{\varphi_k, k \in Z\}, \varphi_k \in L_2(IR). \tag{3}$$

Eq. (3) is referred to as an orthonormal system if Eq. (4) is satisfied

$$\int \varphi_k(x)\overline{\varphi_j(x)}dx = \delta_{jk}. \tag{4}$$

where δ_{jk} and $\{\varphi_k, k \in Z\}$ are Kronecker delta, and orthonormal basis in a subspace V of $L_2(IR)$ if any function $f \in V$ has a depiction

$$f(x) = \sum C_k \varphi_k(x). \tag{5}$$

The coefficient satisfy $\sum_k |C_k|^2 < \infty$

$$Z = \{...,-1,0,1...\}, \sum_k = \sum_{k=-\infty}^{\infty}, \int = \int_{-\infty}^{\infty}. \tag{6}$$

The subspace V_o of $L_2(IR)$ is define as $V_o = \{f, \in L_2(IR): f$
which is constant on $(k, k+1], k \in Z\}$, therefore

$$f \in V_0 \Leftrightarrow f(x) = \sum C_{k\varphi}(x-k), \tag{7}$$

where $\sum_k |C_k|^2 < \infty$, the series converges in Eqs. (8) and (9)

$$L_2(IR), \tag{8}$$

$$\varphi(x) = I\{x \in (0,1]\} = \begin{cases} 1, x \in (0,1] \\ 0, x \in (0,1] \end{cases}, \tag{9}$$

where $\varphi_{ok}(x) = \varphi(x-k), k \in \mathbb{Z}$ in which $\{\varphi_{ok}\}$ is the orthonormal basis in V_o, as such a new linear subspace of $L_2(IR)$ is defined as

$$V_1 = \{h(x) = f(2x): f \in V_o\}, \tag{10}$$

comprises of all functions in $L_2(IR)$ that are constant within the intervals of $\left(\dfrac{k}{2}, \dfrac{k+1}{2}\right], k \in \mathbb{Z}$. The $V_o \subset V_1$, orthonormal basis in V_1 is given by $\{\varphi_{1k}\}$ where $\varphi_{1k}(x) = \sqrt{2}\varphi(2x-k), k \in \mathbb{Z}$. This process continues and the space is generally defined as

$$V_j = \{h(x) = f(2^j x): f \in V_o\} \tag{11}$$

The V_j is a linear subspace of $L_2(IR)$ with the orthonormal basis in Eqs. (12) and (13)

$$\varphi_{jk}(x) = 2^{\frac{j}{2}}\varphi(2^j x - k), k \in \mathbb{Z}. \tag{12}$$

$$V_o \subset V_1 \subset ... \subset V_j \subset ... \tag{13}$$

In a similar manner we define V_j for $j < 0, j \in \mathbb{Z}$ to get the inclusions

$$... \subset V_{-1} \subset V_o \subset V_1 \subset ... \tag{14}$$

The process is infinitely continued until the entire space $L_2(IR)$ is approximated. The orthogonal sum decomposition of V_j is expressed as

$$L_2(IR) = V_o \bigoplus_{j=0}^{\infty} w_j \tag{15}$$

where $w_j = V_{j+1}\Theta V_j$ which is the orthogonal complement of V_j in V_{j+1}. Eq. (15) indicates that each $f \in L_2(IR)$ can be represented as series in the form

$$f(x) = \sum \alpha_{ok}\varphi_{ok}(x) + \sum_{j=0}^{\infty}\sum_k \beta_{jk}\Psi_{jk}(x) \tag{16}$$

where α_{ok}, β_{jk}, and Ψ_{jk} are representation of coefficients of the expansion.

2.2   Adaptive Linear Neural Network

The ALNN can mathematically be defined as [11]:

$$f(x) = \varphi\left(\sum_{i=1}^{m} w_i x_i + b\right). \tag{17}$$

where $x_i (i = 1,2,3...,n), f(x), b, w_i (i = 1,2,3,...,n), m,$ and $\varphi(\cdot)$ are the input variables, output, bias, weights, number of input neurons, and activation function for computation in the ALNN neurons. The architecture of the ALNN comprised of only one layer. The activation function and learning algorithm are pure linear and Widrow-Hoff, respectively. The learning algorithm minimizes mean square error (MSE) to optimized weights and bias as defined in Eq. (18)

$$MSE = \frac{1}{N}\sum_{j=1}^{N}\left[e_j(x)\right]^2 = \frac{1}{2}\sum_{j=1}^{N}\left[T_j(x) - p_j(x)\right]^2, \tag{18}$$

where $N, e_j(x), T_j(x),$ and $p_j(x)$ are the number of data points in the samples, error, target output, and network output. An adaptation of the weights and bias for multiple neurons can be defined as

$$w(x+1) = w(x) + 2\eta e(x)T^T(x). \tag{19}$$

$$b(x+1) = b(x) + 2\eta e(x), \tag{20}$$

where η is the learning rate.

2.3   Relevance Vector Machine

The RVM is typically used for solving classification and regression problems. In this study, we present regression relevance vector machine algorithm for our problem is a regression. Let $\{X_i, t_i\}_{i=1}^{N}$ be input-dethe datasetutput pairs of dataset. Each desired output t_i assumed a normal distribution with mean $y(x_i)$ and uniform variance δ^2 of the noise $\in$ define as

$$p(t/x) = N(t/y(x), \delta^2) \tag{21}$$

The desired output is also assumed joint normal distribution expressed as $N(\mu, \Sigma)$, where μ and Σ are unknown variables which can be determined by the RVM [13].

3   Experiments

3.1   Dataset

The global annual land-ocean temperature (GLOT) is chosen as the indicator of the climate change behavior because it significantly determines climate change more than other indicators such as atmospheric temperature, heat emanating from the sun, among others [14]. The GLOT data are provided in degrees Celsius (degree C) extracted from the National Aeronautic and Space Administration (NASA) website (www.nasa.gov), Goddard Institute for Space, for a period from 1880 to 2012 on a yearly basis. The GLOT data is a global data that is not restricted to a particular region, and were not normalized as the data are already provided in approximately equal proportion within -1 to 1.

3.2   The Propose Ensemble HWD-RVMALNN Framework

The propose ensemble HWD-RVMALNN paradigm is represented in the flowchart illustrated in Fig. 1. The major components of the entire process composed of the following stages: The original GLOT time series data are decomposed using HWD into five scale coefficient. For each of the scale coefficient, the RVM is employed as an estimation technique to model the decomposed scale coefficient and estimate for each of the five scale coefficient. The estimation results produce by each of the RVMs (HWD-RVM1, HWD-RVM2, HWD-RVM3, HWD-RVM4, and HWD-RVM5) is integrated by ALNN to generate an ensemble estimate of the GLOT which can be viewed as the estimation result of the original GLOT time series data. The optimal parameters of RVM were obtained through initial experimentation. The methodology can be summarized as a hybrid ensemble approach comprising of HWD (Decomposition) -RVM (estimate) -ALNN (ensemble) referred to HWD-RVMALNN. To verify the effectiveness of the our approach, we compared the
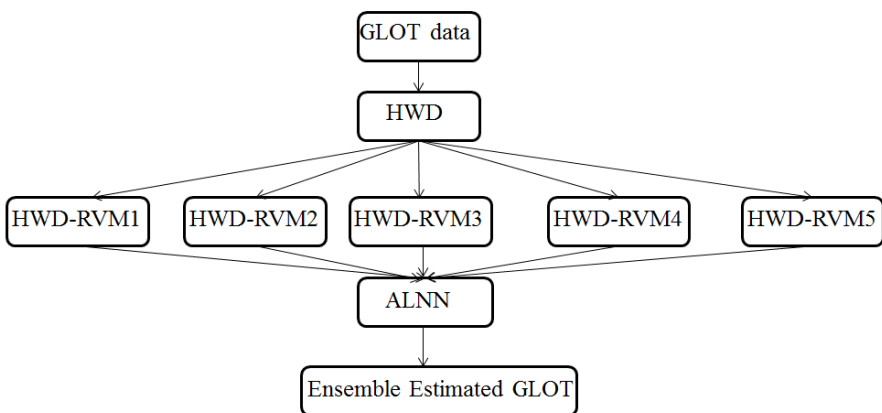
Fig. 1. The propose ensemble HWD-RVMALNN framework

proposed HWD-RVMALNN with the popular approaches of prediction such as Levenberg-Marquardt (LMQ), resilient backpropagation (RBP), Cyclical weight/bias rule (CWB), learning vector quantization (LVQ), and SVM. These algorithms are used as the benchmarks since they are the popular approaches in the literature for the estimation of climate change, readers can refer to [4-6] for details. The algorithms are also applied to estimate the GLOT for the purpose of comparing with the results generated by our proposal.

4   Results and Discussion

In this section, the simulated results are presented and discussed. The original GLOT signal is depicted in Figure 2. Decomposition of the GLOT time series data that significantly influence the fluctuation of climate is shown in Figure 3 displaying the multi resolution analysis that includes fifth approximation and five (d_1, d_2, d_3, d_4, and d_5) levels of details. The five levels of detail coefficients including amplitude of the scaling
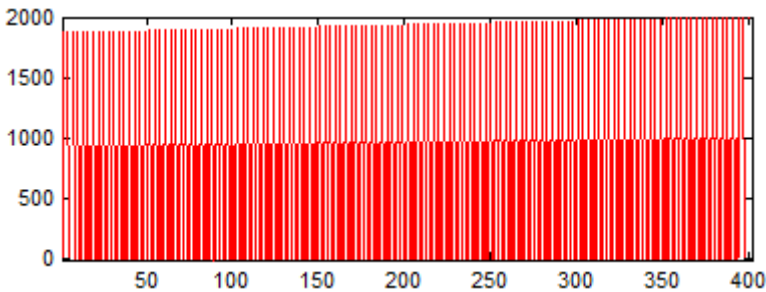

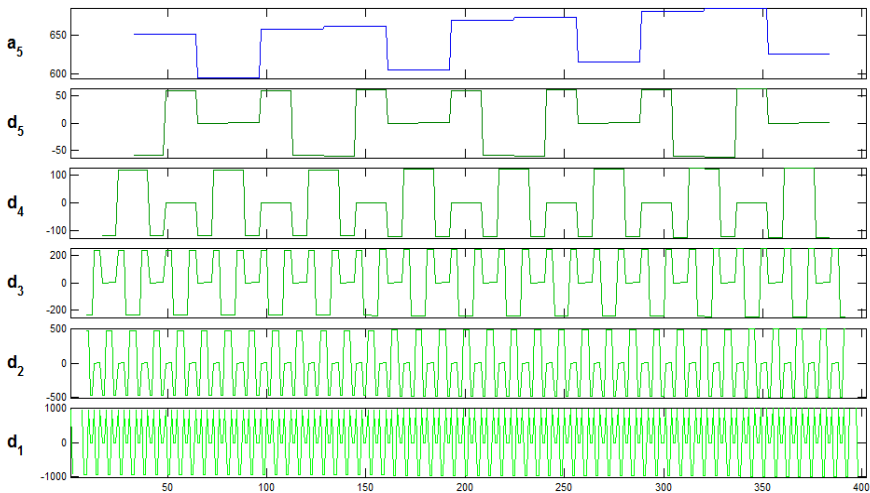**Fig. 2.** The original signal of the GLOT

Fig. 3. Haar Wavelet Decomposition of the GLOT Signal

coefficient are shown in Figure 4. The decomposition contributed to an understanding of the time varying patterns in the GLOT data. Some useful information is revealed from the analysis, and interpretation of the GLOT data. The original GLOT data (Figure 2) and the extracted portion (Figure 3) is in degree Celsius.
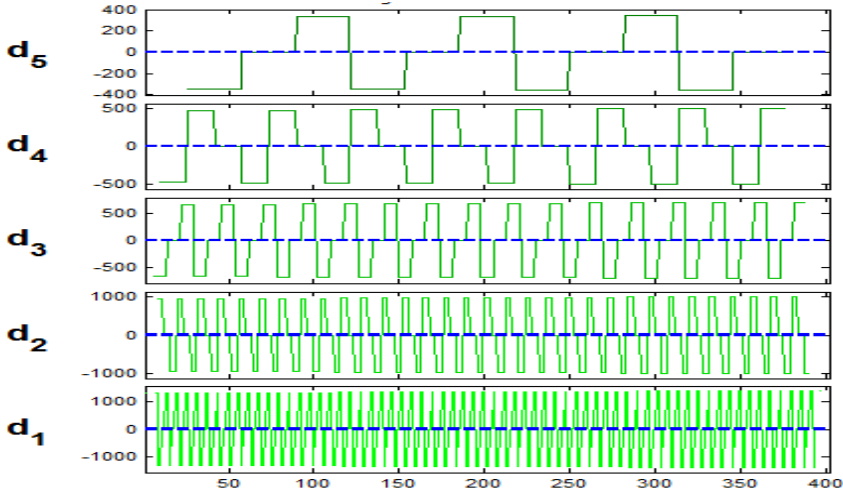

Fig. 4. Original details coefficients of the GLOT signal

The optimal parameters of RVM obtained after the initial experimentation were kernel function = radial basis, relevance vectors = 241, γ = 1.7, kernel length = 4.6, minimum delta scale = 0.0002, and maximum was 1×10^4.

Table 1. Performance metrics of the algorithms compared with the propose ensemble HWD-RVMALNN

Algorithm	MSE	MAE	SSE	R^2	RT(sec.)
LMQ	0.00413	0.0071	0.422	0.7047	2
CBW	0.00828	0.01923	0.51193	0.87268	932
RBP	0.015	0.0788	1.14	0.7932	3
SVM	0.002911	0.019884	0.195	0.88626	5
LVQ	0.091	0.234	8.07	0	2
HWD-RVMALNN	**0.00081**	**0.00721**	**0.00916**	**0.93087**	**0**

Mean absolute error (MAE), Sum of square error (SSE), Runtime (RT)

It is clear shown in Table 1 that the algorithms compared with our approach estimated GLOT with a good degree of accuracy as suggested by the simulated results. The SVM performs better than the LM, CBW, RBF, and LVQ in terms of R^2, MSE and MAE. The results corroborate with the study conducted by [7]. The possible reason for this performance exhibited by SVM could be attributed to the use of support vectors by the

SVM due to its ability to improve performance. The worst performance exhibited by LVQ in terms of R^2 which is merely random could likely be caused by the non-zero differentiability of the LVQ. The runtime of LM and LVQ outperform SVM, CBW, and RBP in terms of computational speed. The poorest convergence speed is observed to be exhibited by the CBW. The likely cause of this poor convergence can best be attributed to the cyclical nature of the CBW to present inputs to the network which might have caused the delay. An algorithm that converges to the optimal solution within the shortest time is considered the most efficient. The propose HWD-RVMALNN as reported in Table 1 based on the performance metrics was able to improve on the performance of the approaches typically use in the estimation of climate change in the literature. They propose an approach performs better than the commonly used approaches in both accuracy and computational efficiency. The likely reason for the performance of HWD-RVMALNN could be attributed to the fact that the results of several experts are ensemble to produce superior output. Estimates of GLOT obtained from simulating the propose ensemble HWD-RVMALNN on the independent test dataset is plotted in Figure 5.
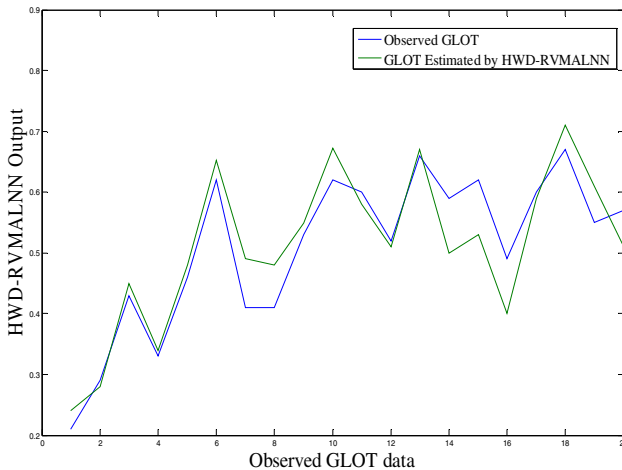

Fig. 5. Performance of the propose HWD-RVMALNN on the test dataset

5   Conclusion

In this research, a hybrid of wavelet transform and computational intelligent algorithms has been proposed for the estimation of climate change behavior. Specifically HWD, RVM, and ALNN were hybridized to propose ensemble HWD-RVMALNN. A series of comparative simulations were performed and it was found that the propose HWD-RVMALNN performs better in both accuracy and computational speed. The propose HWD-RVMALNN can provide better understanding of the future knowledge of climate change behavior than the methods already proposed in the literature. The model can assist policy makers in formulating global response policy in tackling the negative impact expected to be cause by the behavior of climate change. We intend to further this research by developing the real

life application of the propose HWD-RVMALNN model to help climate change experts and policy makers in the decision making process.

References

1. Hoegh-Guldberg, O., Bruno, J.F.: The impact of climate change on the world's marine ecosystems. Sci. 328, 1523–1528 (2010)
2. Belkin, I.M.: Rapid warming of Large Marine Ecosystems. Progr. Oceanogr. 81, 207–213 (2009)
3. Marvin, H.J.P., Kleter, G.A., et al.: Proactive systems for early warning of potential impacts of natural disasters on food safety: Climate-change-induced extreme events as case in point. Food Control 34(2), 444–456 (2013)
4. Matouqa, M., El-Hasan, T., Al-Bilbisi, H.: The climate change implication on Jordan: A case study using GIS and Artificial Neural Networks for weather forecasting. J. Taibah University Sci. 7, 44–55 (2013)
5. Goyal, M.K., Burn, D.H., Ojha, C.S.P.: Statistical downscaling of temperatures under climate change scenarios for Thames river basin, Canada. Int. J. Global Warm. 4(1), 13–30 (2012)
6. Holmberg, M., Forsius, M., Starr, M., Huttunen, M.: An application of artificial neural networks to carbon, nitrogen and phosphorus concentrations in three boreal streams and impacts of climate change. Ecol. Model. 195(1-2), 51–60 (2006)
7. Tripathi, S., Srinivas, V.V., Nanjundiah, R.S.: Downscaling of precipitation for climate change scenarios: A support vector machine approach. J. Hydrol. 330, 621–640 (2006)
8. Jammazi, R., Aloui, C.: Crude oil forecasting: experimental evidence from wavelet decomposition and neural network modeling. Energ. Econ. 34, 828–841 (2012)
9. Demir, B., Ertürk, S.: Hyperspectral image classification using relevance vector machines. IEEE. Geosci. Remote S. 4(4), 586–590 (2007)
10. Hardle, W., Kerkyacharian, G., Pikard, D., Tsybakov, A.: Wavelets, Approximation, and Statistical Applications. Lec. Notes Stat., vol. 129. Springer, New York (1998)
11. Hagan, M.T., Demuth, H.B., Beale, M.H.: Neural Network Design. PWS Publishing Company, Boston (2006)
12. Wipf, D., Palmer, J., Rao, B.: Perspective on spares Bayesian learning. In: Adv. Neural Inf. Processing Syst. 16. MIT Press, Cambridge (2004)
13. Tipping, M.E.: Sparse Bayesian learning and the relevance vector machine. J. Mach. Learn. Res. 1, 211–244 (2001)
14. Richardson, K., Steffen, W., Joachim, H.S., et al.: Climate change: Global Risk, Challenges & decisions. The United Nations Framework on Climate Change Synthesis Report. University of Copenhagen, Denmark (2009)