Multiple Complex Extreme Learning Machine using Holomorphic Mapping for Prediction of Wind Power Generation System

Ram Govind Singh
United Institute of Technology
Naini Allahabad

Akhil Pratap Singh
United Institute of Technology
Naini Allahabad

ABSTRACT
In this paper, a wind prediction system for the wind power generation using ensemble of multiple complex extreme learning machines (C-ELM) is proposed. The extreme learning machines is a single layer feed forward neural network having a fast learning and better generalization ability than the gradient-based learning methods. C-ELM is chosen as base classifier because it is very suitable for processing of non-linear data. For using the wind data in complex domain the wind speed and direction are represented as a complex number. This paper uses the elegant theory of conformal mapping to find better transformations in the complex domain for enhancing its prediction capability. Finally, to improve the generalization ability of the prediction system and to reduce the error encountered in single model predictors, an ensemble of multiple C-ELMS is used. The individual CELM model in the ensemble has different activation functions of the hidden layer neuron. Performance analysis proves that the predictions generated through our method are effective when compared to other complex valued neural network prediction systems.

Keywords
Classification; Complex-Valued Neural Networks; Extreme Learning Machine

1. INTRODUCTION
The current statistics on environmental pollution clearly depicts a global scenario with increasing pollution due to the high carbon dioxide emission [1]. This has lead to a new search for energy sources that are renewable [2]. Wind energy is considered to be a valuable source that does not face the threat of exhaustion along with its usage. However, the main problem in generation of wind power is that most of the wind turbine generators do not match with the speed and direction of the wind. Such a problem can even lead to the failure of a wind power plant. Therefore a prediction system for speed and direction of wind can play a significant role in the damage protection and vibration control of wind turbines.

Many works have been done in the field of wind prediction using neural network classifiers.However with the advance of many applications like radar processing, telecommunications etc where signals are inherently complex, much of the attention has been diverted to complex valued neural networks (CVNN). It has been proved by T.Nitta that CVNN has better generalization abilities [3]. The class of algorithms which uses the complex extreme learning machines, C-ELM [4] as a part of their learning mechanism show far better classification abilities than the existing methods. Recently, a new C-ELM named as Circular-Complex valued Extreme Learning Machine (CC-ELM) [5] proposed by Savitha *et al* has proved to exhibit superior classification abilities among the existing complex-valued classifiers. This is mainly due to the effective transformation function and a Gaussian-like activation function that they used in their classifier. This work uses variant of C-ELM which uses sech() activation function at hidden layer.

A conformal or holomorphic mapping is a complex mapping that has the capability of taking every point in a complex plane onto another complex plane. The conformal mapping defined as $w = f(z)$ has the property of preserving the local angle during mapping including the shape of very small figures. However, size cannot be preserved in such a mapping. *Thomas L. Clark* in [6] has mentioned that the elegant theory of conformal mappings can be used to find other neural and analytic functions for non-linear mapping in the complex plane. The advantage of using a holomorphic transformation function/activation function is that, we can utilize the conformal nature in total neural transform [7]. It has also been mentioned in [8] that the conformal mappings can also preserve the direction relationship of crossing two boundary curves in classification or prediction tasks.

An ensemble can be considered as an assembly of single model neural networks that exhibit diversity across the network by either varying the learning algorithm or varying the data provided to each model etc [9]. The first ensemble technique was introduced in literature by Hanson and Salmon [10]. A CVNN has better prediction making capabilities and ELMs have a fast learning rate, an ensemble based on multiple CELMs can reduce the error of the whole system and generate more accurate predictions. Even though various ensemble techniques are present, the dominating method remains to be Bagging along with its variants [11].

In this work, which we term as *Holomorphic based Multiple Complex Extreme Learning Machine (MH-CELM)* we have treated the wind information comprising of wind speed and direction as a complex number on the complex co-ordinates. This data is then transformed into another complex plane using the concept of conformal mapping with the function $w=\sin(z)$. The transformed data is then processed by the four complex valued extreme learning machines, with different activation functions.

The rest of the paper is organized as follows. In Section 2, the related work is described followed by Section 3 which explains the basic complex-valued extreme learning machine algorithm and the Bagging algorithm. Section 4 describes in detail the proposed work and its algorithm. The experimental evaluation and analysis is given in Section 5 and finally conclusion is presented in Section 6.

2. RELATED WORK

Various neural network methods of time series forecasting predicting have been utilized for the wind speed and direction. However, the work in complex-valued neural networks[12-15] is limited but effective than their real valued counterparts.

In [12], *Takahiro et al* presented a CVNN for predicting the wind speed and direction for a wind power generation system. The wind speed and direction were represented as a complex number on the complex co-ordinates for generating an input to the CVNN. The CVNN was trained using a complex back propagation algorithm during training for predictions. The paper vividly showed how the orthogonal decision boundaries of a complex plane can aid in predictions.

An Augmented Complex Least Square (ACLMS) method was proposed for wind prediction in [13]. It combines the uses of the augmented complex statistics with the complex-valued wind prediction methods. The algorithm brings together the concepts of linear adaptive forecasting in complex domain along with some new complex statistics.

Savitha *et al* proposed a Complex neuro-fuzzy inference system (CNFIS) for wind prediction in [14]. In this paper, a CNFIS was designed using a gradient descent learning algorithm using the Wirtinger calculus. They proposed a prediction system that realizes zero-order Takagi-Sugeno-Kang based fuzzy inference system. The CNFIS takes the advantages of the approximation abilities of neural networks and the data representability of the fuzzy inference systems. The method was proved to predict more accurately than the ACLMS prediction.

Another work worth to be mentioned in the wind forecasting methods using complex valued neural networks is the one which uses a Meta-Cognitive Fully Complex-valued neural network [15]. In this prediction method, a Fully Complex-valued Radial Basis Function (FC-RBF) [16] network works as the cognitive part and a self-regulatory learning method is the meta-cognitive component. In par with the knowledge acquired by the cognitive component and the information in each new sample, the meta-cognitive network decides in each epoch what the network should learn, when to learn and how to learn. The method proved to work better in predictions than that of FC-RBF predictions and real-valued ELM predictions.

3. SYSTEM FOR WIND SPEED & DIRECTION FORECASTING

3.1 Representation of wind speed and direction in MH-CELM

The inputs to the proposed MH-CELM should be purely in complex form. Since we have to predict the wind speed and direction this data must be represented in the form a complex number[13]. Let the wind speed and direction at *t*-th hour be v_t and φ_t respectively. The complex representation of this data, *z(t)* will be of the form as below:

$$z(t) = v_t \cos \varphi_t + v_t \sin \varphi_t \quad (1)$$

Real part of *z(t)* can be defined as $Re(z(t)) = v_t \cos \varphi_t$ and the imaginary part, $Im(z(t)) = v_t \sin \varphi_t$. As depicted in the figures 1 and 2, real part and imaginary parts of *z(t)* are nothing but the south-north and east-west components of the velocity of the wind.

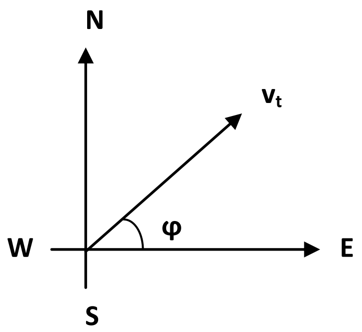

Figure 1: Speed and direction the the wind

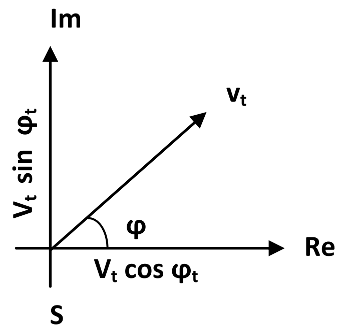

Figure 2: Complex representation of wind data

3.2 Variant of Complex-valued Extreme Learning Machine

The C-ELM is a single layer feed forward neural network (SLFN)[4] where the parameters between input and hidden layer are initialized randomly and the weights between hidden and output layer are obtained analytically

Let us assume that we have *N* training samples of the form (z_i, y_i) where $i = 1,2,3,\ldots\ldots, N$ and $z_i \in \boldsymbol{C}^n$ and $y_i \in \boldsymbol{C}^m$. In this paper, we use a variant of C-ELM with a fully non-linear complex-valued activation function $g_c(u,v,z)$ in the hidden layer of neurons.

$$h = g_c(\mathbf{u_j}^{\mathbf{T}}(\mathbf{z}_t - \mathbf{v}_j)); j = 1,2,\ldots, K \quad (2)$$

Here $\mathbf{u}_j$, $\mathbf{v}_j \in C^n$ are complex-valued scaling factor and complex-valued center of the *j*-th hidden neuron respectively.

The activation function used in the output layer of the C-ELM is linear in behavior. The variant of C_ELM differs from original CELM only in the activation used at hidden layer. Variant of CELM uses sech() activation function. The actual output o_n of the C-ELM network with *K*-hidden neurons and a non-linear activation function in the hidden layer *g(u,v,z)* can be defined as:

$$\widehat{\boldsymbol{o_n}} = \sum_{j=1}^{K} \boldsymbol{w_{nj} h_j} \,, \quad (3)$$

where w_{nj} is the weight vector containing the weights connecting the *n*-th output neuron to the *j*-th hidden neuron. This equation can be also represented as:

$$\hat{O} = WH \quad (4)$$

where W is the matrix containing the weights connecting the hidden layer and output layer neurons and H is the matrix with the outputs of the hidden layer which is defined below:

$$H(u_1,u_2,\ldots,u_k\ ,\ v_1,\ v_2,\ldots,\ v_k\ ,z_1,z_2,\ldots,z_N)= \begin{bmatrix} \mathrm{sech}(\boldsymbol{u_1}^T\|\boldsymbol{z}_1-\boldsymbol{v_1}\|) & . & . & \mathrm{sech}(\boldsymbol{u_1}^T\|\boldsymbol{z}_N-\boldsymbol{v_1}\|) \\ \vdots & \ddots & & \vdots \\ \mathrm{sech}(\boldsymbol{u_K}^T\|\boldsymbol{z}_1-\boldsymbol{v_K}\|) & . & . & \mathrm{sech}(\boldsymbol{u_K}^T\|\boldsymbol{z}_N-\boldsymbol{v_K}\|) \end{bmatrix}_{K\times N} \quad (5)$$

$$W = \begin{bmatrix} w_1^T \\ \ldots \\ w_N^T \end{bmatrix}_{N\times m} \qquad O = \begin{bmatrix} o_1^T \\ \ldots \\ o_N^T \end{bmatrix}_{N\times m} \quad \text{and } Y = \begin{bmatrix} y_1^T \\ \ldots \\ y_N^T \end{bmatrix}_{N\times m}$$

The $K \times N$ matrix H is known as the complex hidden layer output matrix.

4. PROPOSED WORK: Holomorphic based Multiple -Complex Extreme Learning Machine (MH-CELM)

4.1.Conformal/Holomorphic Mapping

Conformal mapping can be defined as a transformation function

$$w = f(z)$$

and wind data is represented in complex form as

$$z(t) = a+it = v_t\cos\varphi_t + v_t\sin\varphi_t$$

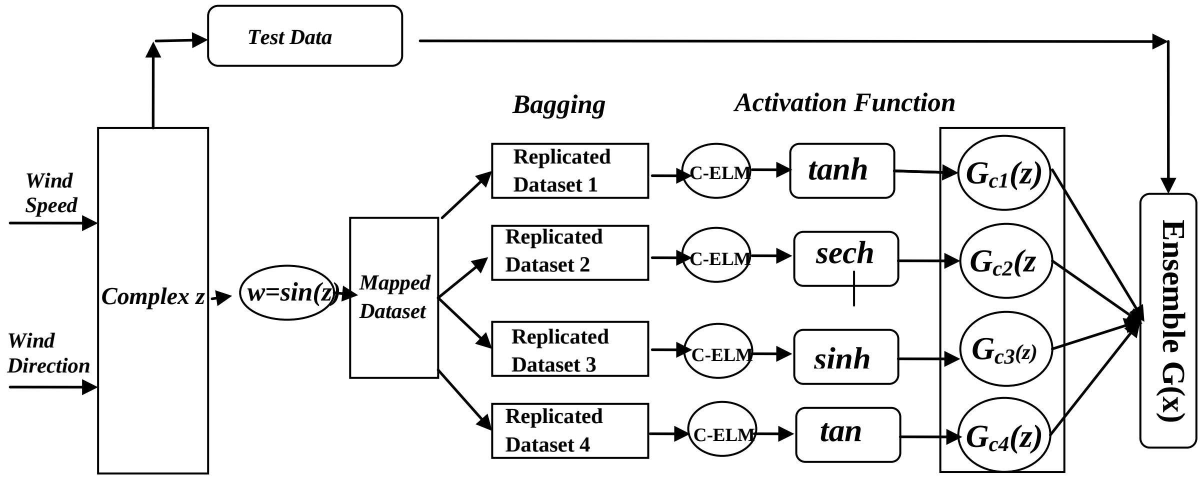

Figure 5: MH-CELM model

Now the task is to find out a complex function w=f(z) such that the complex wind data z will take a configuration consisting of points, lines, angles and regions in the complex z- plane and convert it into a simple and more readily analyzable configuration in the complex w-plane[17]. A holomorphic mapping w=f(z) in a Domain D is called conformal at $z=z_0$ when it satisfy following two properties. 1. z is analytic in the Domain D and $f'(z_0) \neq 0$ 2. The angle between any two intersections in the z-plane is equivalent to the angle of image arcs in the w-plane [17]. Let f:D→C be a holomorphic map on an open set. If a curve in D is parameterized by z=z(t) then w=f(z(t)) describes the image curve in the w-plane. Let $z_0 \in D$ and C_1:[-1,1]→D and C_2:[-1,1]→D be two path which meet $z_0=C_1(0)=C_2(0)$. The meeting point of original curve is defined as :

$$\Theta = arg\ C_2'(0)\ -\ arg\ C_1'(0) = arg\frac{C_2'(0)}{C_1'(0)}$$

The confomal mapping of the curve $f(C_1)$ and $f(C_2)$ meets at the $f(z_0)$ at the angle

$$\emptyset = arg\ (fC_2)'(0) - arg(fC_1)'(0)$$

$$= arg\frac{(fC_2)'(0)}{(fC_1)'(0)} = arg\frac{f'(C_2(0))C_2'(0)}{f'(C_1(0))C_1'(0)}$$

[By the Chain rule]

$$= arg\frac{f'(z_0)C_2'(0)}{f'(z_0(0))C_1'(0)} \quad [\text{as } C_2(0) = C_1(0) = z_0\ ]$$

$$= arg\frac{C_2'(0)}{C_1'(0)} = \Theta$$

So ,Conformal mapping of a complex variable is angle preserving and sense preserving which convert the input data of z-plane into w-plane.Due to this Conformal mapping can be used for complex transformationof data for further analysis. To apply this elegant theory a conformal function w=f(z)= sin(z) is selected which hold two essential property of holomorphic function and transform the input data into more analyzable form due to orthoganality of hyperplane in w-plane.the proofits orthogonality and analyticy isgiven below :

1. f'(z)=cos(z) , *f* is analytic at all points in the region where it has non zero derivative [17]. Since wind speed and directions are finite values so it must be analytic in the given region. Louivilles's theorem [18] proves that any analytical and almost bounded function can be used for transformation purpose in the complex domain
2. for image transformation in w-plane ,angle preserving for function w=Sin(z) can be described from the fundamental region of trigonometric functions[18]. For the function sin(z), a line x=a,exists inside the region within vertical strip $-\frac{\pi}{2} \leq x \leq \frac{\pi}{2}$ can be presented by
$z(t) = a+it = v_t\cos\varphi_t + v_t\sin\varphi_t$

From trigonometric properties we know that
sin(z)= Sin$x$Cosh$y$ +iCos$x$Sinh$y$ So,
u+iv = Sin(a+it) = SinaCosh$t$ +iCosaSinh$t$
= Sin $v_t cos\ \varphi_t$.Cosh $v_t sin\ \varphi_t$ +iCos $v_t cos\ \varphi_t$ .Sinh $v_t sin\ \varphi_t$
From the identity, $Cos^2ht - Sin^2ht = 1$

$$\frac{u^2}{Sin^2v_tCos\emptyset_t} - \frac{v^2}{Cos^2v_tCos\emptyset_t} = 1$$

The image of line x=a of w plane is therefore a hyperbola with $\pm$ Sin $v_t cos\ \varphi_t$ as an u-intercept.Similerly the horizontal line segment x=b of z-plane will map into w-plane as:

$$\frac{u^2}{Cosh^2v_t} + \frac{v^2}{Sinh^2v_t} = 1$$

Horizontal line of z-plane mapped as an elipse in w-hyperplane. Therefore hyperbolas and ellipses are orthogonal since they are the images of orthogonal families of horizontal and vertical lines. [3] has proved that a complex function which have orthogonal boundary in a complex plane have a better decesion making capability .so elegant orthogonal conformal mapping of the function w=Sin(z) can be used to enhance the weather prediction capability using the wind data set and C-ELMs. The transformation of complex function is depicted in figure 3 and figure 4[18].

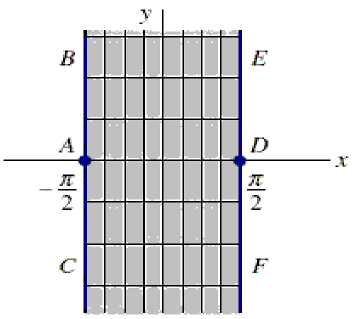

Figure 3.Image of Complex z as vertical strip

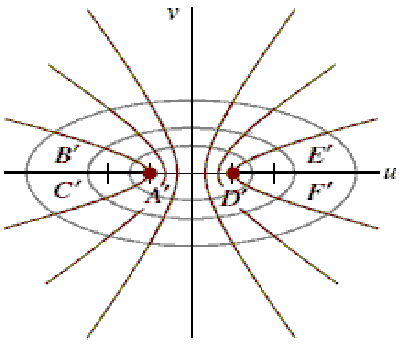

Figure 4.Conformal mapping w=Sin(z) Orthogonal families of hyperbolas and ellipses

4.2. Multiple complex activation functions in MH-CELM

For a CVNN, the activation function must be both analytic and bounded which are the essential properties of a CVNN activation function as per pointed out by T. Nitta in [3].

The choice of a non-linear complex activation function plays a very vital role and finding such a function is a very tedious job. This is because there occurs a conflict in the boundedness and differentiability of complex functions in a complex plane as in Louivilles's theorern [18]. As in [18], a bounded function in a complex domain is constant. Thus for a complex activation function the aim is find those functions in the complex domain that are almost bounded. This limits the number of non-linear activation functions for a CVNN.

It was then proved in [20] that complex hyperbolic functions like *tanh z*, *sinh z* and trigonometric function like can be used as a fully complex valued activation function. It outperforms many other non-linear activations used mainly because it's bounded and well defined derivates meet the Cauchy-Riemann equations almost everywhere in the complex domain. In a recent paper by *Savitha et al* [16]. *sech z* has been proved as a more better complex activation function to be employed in a CVNN. In general, using circular, inverse circular, hyperbolic and inverse hyperbolic functions are *almost* bounded and also analytic qualifying them for the job of non-linear complex activation functions to be used in a CVNN.

It has been already mentioned before that for better prediction accuracy, an ensemble must have a proper diversity across the individual models and also it must implement an appropriate technique to gather the predictions from the individual C-ELMs and give a proper output with accuracy. In MH-CELM, the output is predicted through an ensemble learning algorithm known as *Bagging*. *Bagging* (**B**ootstrapped **Agg**regati**ng**) trains the ensemble by providing a new randomly selected data for each C-ELM in the ensemble from the training dataset. This is choosing is done by randomly replicating the dataset for each C-ELM. The size of the replicated dataset can be same as that of the original or less than that. The selection is made with replacement. Apart from the diverse activation functions in each C-ELM, providing different data to each of them increases the diversity of the ensemble enabling it to predict the outputs more precisely.

The operation of a MH-CELM can be thus summarized as follows. To obtain superior ensembles the task is how to establish a diverse and precise prediction algorithms and how to combine their results. [24] has mentioned that four different type of activation function can increase the diversity of ELM's. So, With the activation functions *tan, sech, tan and sinh* , MH-CELM consists of four basic C-ELMs to which data is provided after converting into complex numbers followed by a conformal mapping. Each C-ELM in the ensemble has 10 hidden neurons which has been selected using cross fold validation method [5] and they employ a different non-linear complex activation function that is almost bounded and analytic. In the training phase, the dataset is randomly replicated for each C-ELM to train in a diverse manner. During the testing phase, unknown instances are provided for which each C-ELM predicts an output. The final output is predicted through averaging of each output from the ensemble. The pseudo code of the proposed MH-CELM ensemble is given in Table No. 1.

Algorithm 1: MH-CELM

Input: A training set N = $\{(\boldsymbol{z_n^{train}}, \boldsymbol{y_n^{train}}) \mid \boldsymbol{z_n^{train}}, \boldsymbol{y_n^{train}} \in \boldsymbol{C^N}\}$, testing set T = $\{(\boldsymbol{z_n^{test}}, \boldsymbol{y_n^{test}}) \mid \boldsymbol{z_n^{test}}, \boldsymbol{y_n^{test}} \in \boldsymbol{C^N}\}$,

the hidden node activation function *G(u,v,z)*, No. of hidden nodes *L*, No. of trial *K*

Output: final prediction by learning algorithm MH-CELM = avg{*Result*}

Training phase

1. Set k =1
2. Apply a conformal transformation upon the dataset, $w = sin(z)$
3. **while (**$k \leq K$**) do**
4. Set i = 1
5. **while (**$i \leq 4$**) do**
6. Randomly assign learning parameter $u_j(k,i)$, $v_j(k,i)$ $j=1,2, \ldots, L$ of the i-th C-ELM for the k-*th* ensemble
7. Calculate the hidden layer output matrix $H(k,i)$
8. Calculate the output weight $W(k,i)$, $W(k,i) = YH(k,i)^{\dagger}$, where Y is the target matrix
9. $i = i + 1$
10. **end while**
11. $k = k + 1$
12. **end while**

Testing phase

1. Apply a conformal transformation upon the test dataset T, $w = sin(z)$
2. Set k = 1
3. **while (**$k \leq K$**) do**
4. Set i = 1
5. **while (**$i \leq 4$**) do**
6. Using the $(k,i)^{th}$ trained basic C-ELM with learning parameter $(u_i(k,i), v_i(k,i), w_i(k,i))$ $j=1,2,\ldots, L$ to predict target output of testing sample.
7. $output_i$ = output of testing sampling z^{test} using $(k,i)^{th}$ C-ELM
8. i=i+1
9. **end while**
10. $Result_k = avg\ (output)$
11. $k=k + 1$
12. **end while**

Table 1: MH-CELM Algorithm

5. EXPERIMENTAL STUDY AND RESULT ANALYSIS

In this section, the results of the wind speed and direction are presented. Five different datasets obtained from Iowa (USA), Department of Transport [21] are used throughout the experiments. The datasets consist of wind speed and direction of five different wind power stations obtained from 1st January 2011 to 28th February 2011. For verifying the validity of results we have used different number of training and testing instances. The detailed descriptions about the datasets that are used in the experiments are given in Table 2.

The experiments conducted can be classified into four categories:

1. Comparison of the predicted results using MH-CELM with the original data series for both wind speed and direction. This is to show that the predictions of MH-CELM are stable with the original data series.

Station name	Training instances	Testing instances	Min Speed (m/sec)	Max Speed (m/sec)	Mean Speed (m/sec)	Std dev
Washington	719	648	0	7.5	2.4242	1.3222
Algona	410	300	0	7.778	2.18	0.41
Clarion	510	200	0	7.77	2.17	0.4203
Sheldon	190	129	0	1	0.3781	0.0742
Oskalossa	900	488	0	8.888	2.2562	0.4243

Table 2: Description of wind datasets of various wind power stations

2. Comparison of the predictions of MH-CELM with H-CELM (Holomorphic-CELM without ensemble) and C-ELM for various parameters including NMSE, RMSE, coefficient determinant and prediction gain. This experiment proves that the proposed ensemble can attain better results in comparison to single model complex valued neural networks.

3. Comparison of MH-CELM with various other prediction systems in literature. This helps to prove that the proposed method is better than many other already proposed methods.

4. Comparison of MH-CELM with H-CELM and C-ELM by varying the number of trials in terms of RMSE, NMSE and prediction gain.

For conducting the experiments the input to the MH-CELM is passed through a moving average. The dataset consists of hourly data and for analyzing the hourly data the input vector F is organized as consisting of five data points F1, F2, F3, F4, F5 as follows.

F = {F1, F2, F3, F4, F5} = {X(ti), MA2, MA3, MA5, MA10}

Here X(ti) is the current ith hour wind speed and MA2, MA3, MA5, MA10are the moving averages of 2, 3, 5 and 10 hours.The simple moving average can be described as the unweighted mean of the previous n points. Therefore if we want to get the moving average of previous n hours then the formula is as follows:

MAn = $\frac{p_{present}+p_1+p_2 \ldots\ldots + p_{n-1}}{n}$ Where $p_{present}$ is the value of current hour.

The moving average is taken rather than the simple average because it ensures that variations in the mean are aligned with the variations in the data rather than being shifted in time. Thus, the MH-CELM takes the input vector F and predicts the wind speed or direction of i+1th hour.

5.1 Comparison of MH-CELM prediction with the original wind speed and direction time series

In this section the MH-CELM is used to predict the wind speed and direction using the five different datasets that is given in Table 4.12. The MH-CELM is first trained using training instances and then the predictions are tested using the testing instances. Since Range normalization gives better prediction [5],The training and testing data is normalized in the range of [0,1].The number of training and testing instances is also specified for each dataset in Table 2. The predictions made by MH-CELM with the original values are compared and graphically presented in Figures 6 to 15 for all the five datasets.

The figures clearly indicate that the predictions have the same fluctuations as per the original wind speed and direction time series. This proves that the MH-CELM predictions can help the wind turbines to stabilize themselves against the various fluctuation based on the previous data values.

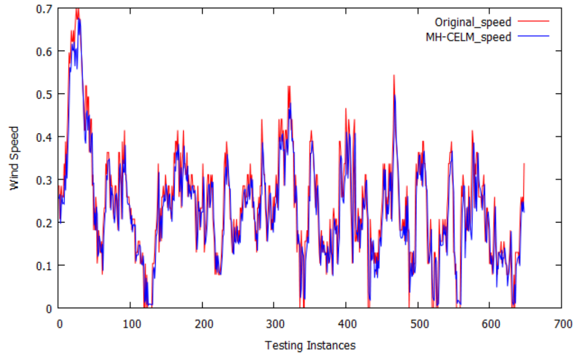

Figure 6. Wind speed prediction for Washington dataset

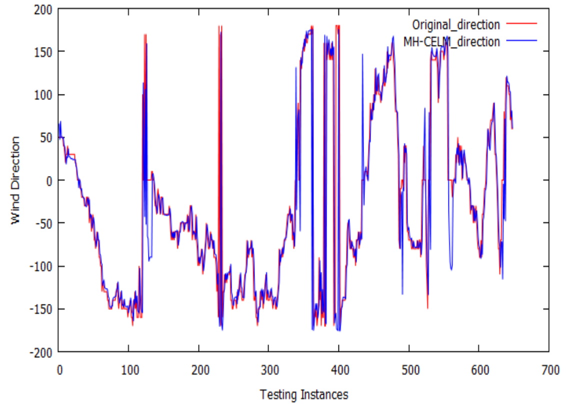

Figure 7: Wind direction prediction for Washington dataset

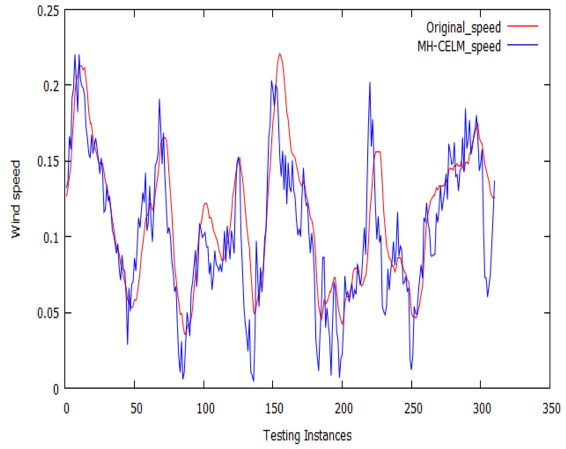

Figure 8: Wind speed prediction for Algona dataset

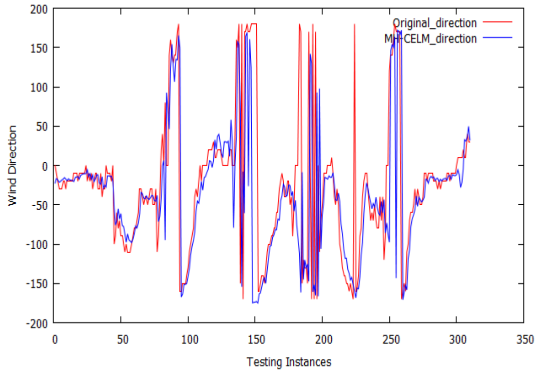

Figure 9: Wind direction prediction for Algona dataset

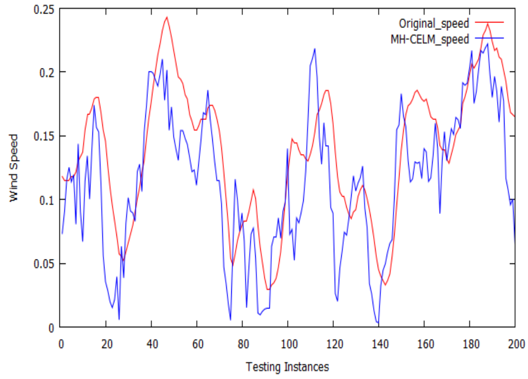

Figure 10: Wind speed prediction for clarion dataset

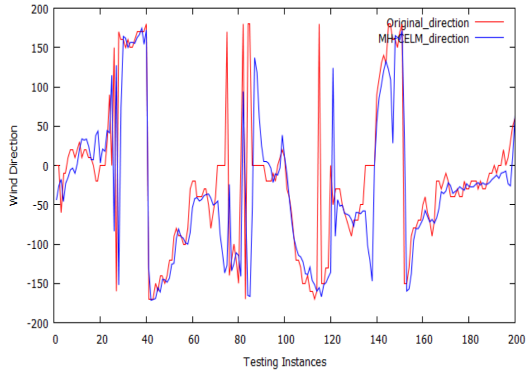

Figure 11: Wind direction prediction for clarion dataset

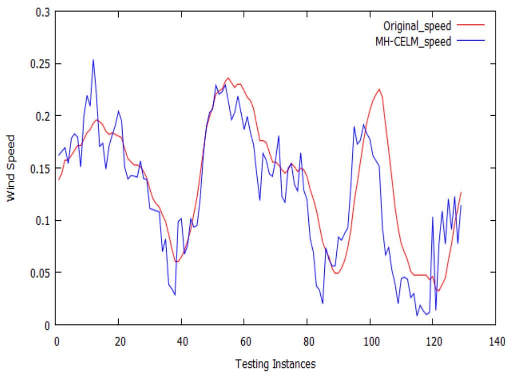

Figure 12: Wind speed prediction for Sheldon dataset

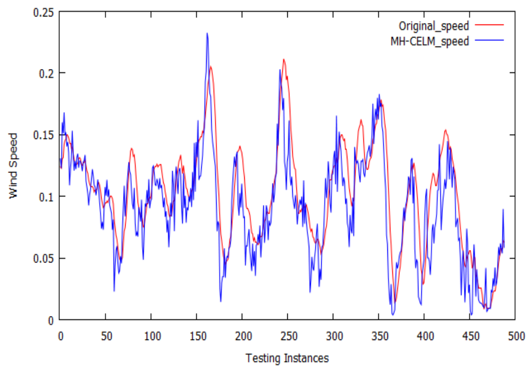

Figure 13: Wind speed prediction for Oskalassa dataset

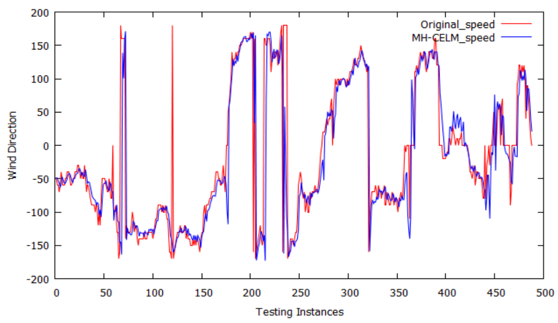

Figure 14: Wind direction prediction for Oskalassa dataset

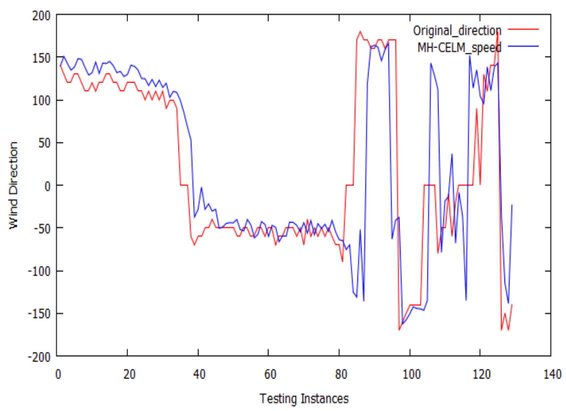

Figure 15: Wind direction prediction for Sheldon dataset

5.2 Comparison of MH-CELM with single model complex-valued neural networks

The proposed method for wind prediction consists of an ensemble of complex-valued neural network with conformal mapping. The basic idea that drives this method is that ensembles can generalize the predictions and minimize the risk of error in single model complex-valued neural network. To show that the ensemble method is far better for predictions in comparison to the single model neural network predictors experiments were conducted on the predictions of MH-CELM with Simple Complex valued Extreme learning machine (C-ELM) [4] and C-ELM with holomorphic mapping. For showing that the results for MH-CELM are far better than the single models five parameters [13, 15, 22] are chosen whose description is detailed below:

Normalized Mean Square Error (NMSE): The NMSE is a measure to check the deviation between actual and predicted value.

$\text{NMSE} = \frac{\sum_k (y_k-\widehat{y_k})^H (y_k-\widehat{y_k})}{\sum_k (y_k-\overline{y_k})^H (y_k-\overline{y_k})}$ Where k is the number of sample to be forecasted ,yk is the real signal value. $\widehat{y_k}$ is the predicted value and $\overline{y_k}$ is the mean of the data.

Coefficient of multiple determination: the value of r closed to unity indicates perfect prediction.

$$r^2 = 1 - \frac{\sum_k (y_k-\widehat{y_k})^H (y_k-\widehat{y_k})}{\sum_k (y_k-\overline{y_k})^H (y_k-\overline{y_k})}$$

MAE: It is an average of absolute error between original and predicted value and it is defined as

$$MAE = \frac{1}{N} |y_k - \widehat{y_k}|$$

RMSE: it is defined as an expected value of the quadratic or squared error loss and measures as the square root of the average of square of the errors.

$$\text{RMSE} = \sqrt{\frac{(y-\hat{y})^H (y-\hat{y})}{N \times n}}$$

Prediction gain: $R_p \triangleq 10 \log_{10} \left(\frac{\sigma_y^2}{\overline{\sigma_e}^2} \right)$ [dB]

Where σ_y^2 denotes the variance of the input signal y, $\overline{\sigma_y}^2$ denotes the estimated variance of the forward prediction error e.

With these parameters predictions were made with the five dataset for MH-CELM, H-CELM and C-ELM and the results are tabulated in the Tables 3-7. As for NMSE, all the five datasets show high value for C-ELM followed by H-CELM and then MH-CELM. This indicates that although conformal mapping can reduce the NMSE of CELM constructing an ensemble can further reduce the error. Also all the results have a lower NMSE value for H-CELM than for C-ELM. This further shows the significance of holomorphic mapping in arranging the data for predictions. The Washington dataset shows lower NMSE than any other dataset experimented followed by Oskalassa dataset and Clarison dataset shows a higher difference in the NMSE values between MH-CELM, H-CELM and C-ELM with MH-CELM outperforming both H-CELM and C-ELM.

The coefficient of multiple determinations has to have higher values for a far improved prediction capability and unity value is called perfect prediction. The MH-CELM exhibits values that are higher than H-CELM and C-ELM in all the datasets tested with H-CELM having better values than C-ELM. The Washington datasets have higher values *r*. however significant improvement in coefficient of multiple determinations can be seen in Clarion dataset.The MAE and RMSE values show similar trends throughout the experiments. The RMSE and MAE value is the lowest for the MH-CELM rather than H-CELM and C-ELM. This proves that an ensemble can always minimize the error of single model neural network models because the NMSE, RMSE and MAE of MH-CELM are significantly lower than H-CELM and C-ELM.Higher values of predictive gain indicate better predictions for a prediction model. From the results it can be seen that the gain for MH-CELM is better than that of H-CELM and C-ELM.

Thus it can be inferred that construction of an ensemble has significantly increased the prediction capability of the wind prediction model. Also the results further emphasize the importance of holomorphic mapping in making the data readily analyzable for the prediction system.

	NMSE	**Coff deter(r)**	**MAE**	**RMSE**	**Gain_Rp**
CELM	0.198	0.801	0.0390	0.0471	7.158
H-CELM	0.192	0.8070	0.0383	0.0465	7.305
MH-CELM	0.185	0.8144	0.0378	0.0334	7.686

Table 3:Performance result of Washington dataset

	NMSE	**Coff deter(r)**	**MAE**	**RMSE**	**Gain_Rp**
CELM	0.136	0.8632	0.0736	0.0975	8.735
H-CELM	0.123	0.8768	0.0728	0.092	9.138
MHCELM	0.113	0.8867	0.0713	0.0853	9.850

Table 4: Performance result of Algona dataset

	NMSE	**Coff deter(r)**	**MAE**	**RMSE**	**Gain_Rp**
CELM	0.3052	0.6947	0.0613	0.0746	5.6013
H-CELM	0.2939	0.7060	0.0602	0.0731	5.8006
MH-CELM	0.2728	0.7271	0.0586	0.0626	6.4514

Table 5: Performance result of Clarion dataset

	NMSE	**Coff deter(r)**	**MAE**	**RMSE**	**Gain_Rp**
CELM	0.1553	0.8446	0.0332	0.0411	8.1674
H-CELM	0.1598	0.8401	0.0337	0.0417	8.0444
MH-CELM	0.1540	0.8459	0.0333	0.0322	8.3645

Table 6: Performance result of Oskalassa dataset

	NMSE	**Coff deter(r)**	**MAE**	**RMSE**	**Gain_Rp**
CELM	0.2324	0.7675	0.0551	0.0715	6.9760
H-CELM	0.2362	0.7637	0.0566	0.0721	7.0390
MH-CELM	0.2526	0.7473	0.0594	0.0612	7.2265

Table 7: Performance result of Sheldon dataset

5.3 Comparison of MH-CELM with other learning algorithms in literature

The results in Section 5.2 were able to prove the performance overhand of MH-CELM ensembleover other single model complex valued neural network prediction systems like H-CELM and C-ELM. To further study the prediction capability of MH-CELM experiments were conducted to compare the RMSE of MH-CELM with other complex-valued neural network learning algorithms in literature. Table 8 depicts the learning algorithms used for comparison with the number of hidden layer neurons (NHN) and their corresponding RMSE. The learners compared with MH-CELM are ELM (Extreme Learning Machine) [23], FC-RBF () [16], Mc-FCRBF () [15], CELM [4] and H-CELM. The results of ELM, FC-RBF and Mc-RBF are reproduced from.

The results in Table 8 vividly demonstrate that the RMSE of MH-CELM is the minimum among all the other prediction models in comparison. The ELM presented is a real valued Extreme Learning Machine. The FC-RBF has a Gaussian type radial basis function as the activation function in the hidden layer and uses gradient descent based learning algorithm. The lower RMSE value of ELM in comparison with the FC-RBF demonstrates the efficiency of extreme learning machines in learning and their ability in minimizing error.

The Mc-FCRBF consists of two components: a cognitive and a meta- cognitive component. A Fully Complex-valued Radial Basis Function (FC-RBF) network is the cognitive component and a self-regulatory learning mechanism is its meta-cognitive component. In each epoch of the training, when the sample is presented to the Mc-FCRBF network, the meta-cognitive component decides what to learn, when to learn, and how to learn based on the knowledge acquired by the FC-RBF network and the new information contained in the sample. This approach is proved to have better prediction capability with minimum RMSE when compared to both ELM and FC-RBF. It is seen in Table 8 that C-ELM has a lower RMSE than FC-RBF, but still higher than the other models in comparison. But using a conformal mapping to configure the input data in H-CELM has brought the RMSE close to Mc-FCRBF. This further underlines the significance of conformal mapping. However, combining the concept of conformal mapping with that of ensemble has produced the minimum RMSE of all the prediction models as can be seen in Table 8.

Learning algorithm	**NHN**	**RMSE %**
ELM	15	9.3
FC-RBF	15	12.56
Mc-FCRBF	15	9.25
CELM	10	9.7514
H-CELM	10	9.2849
MH-CELM	10	5.501093

Table 8: Comparison with other learning algorithms

5.4 Comparison of MH-CELM with H-CELM and C-ELM in terms of RMSE, NMSE and prediction gain for varying number of trails

In this section, the main objective is to compare the performance metrics like RMSE, NMSE and prediction gain for varying number of trails. This is done for MH-CELM, H-CELM and C-ELM to further ensure the efficiency of MH-CELM. The experiments are carried out on Washington dataset. The Washington dataset because it uses a higher value for the number of instances trained and tested. The results are graphically represented in Figures 16-18.

The comparisons of MH-CELM, H-CELM and C-ELM for average RMSE can be seen in Figure 16. The graph shows that while C-ELM and H-ELM have instable variations throughout the number of trials the MH-CELM has significantly minimum values than H-CELM and C-ELM throughout the experiment.

The Figure 17 vividly demonstrates how large prediction gain is attained by MH-CELM in comparison with H-CELM and C-ELM. Also, the H-CELM has better prediction gain than C-

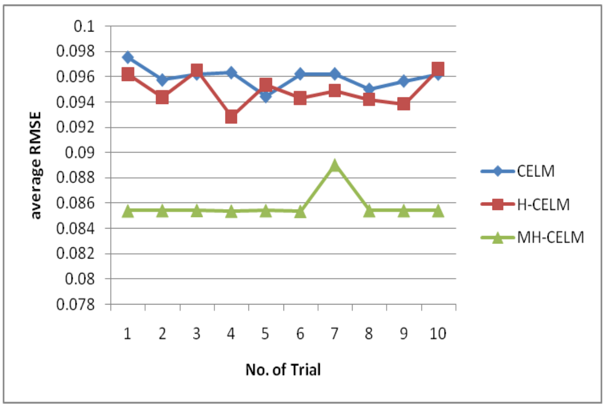

Figure 16: Average RMSE of Washington data set varying no. of trial

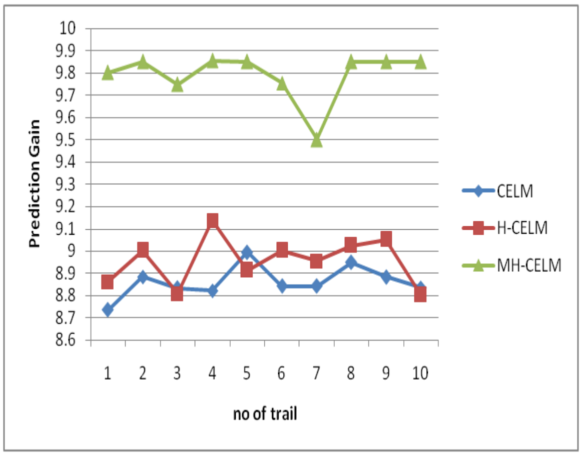

Figure 17: Forward prediction gain of Washington data set varying no. of trial

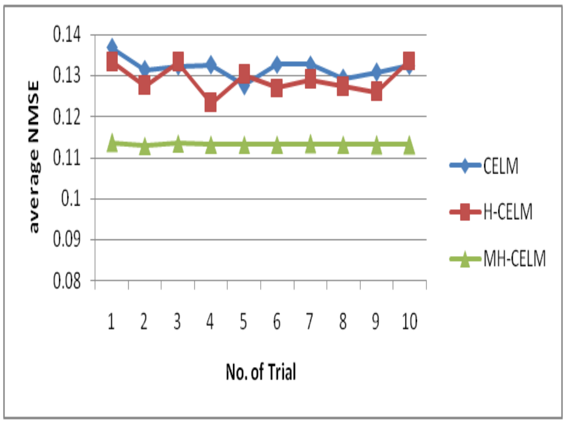

Figure 18: Average NMSE of Washington data set varying no. of trial

ELM in most of the points again reinstating the fact that conformal mapping of input data is important to make data more analyzable for prediction.The NMSE results are presented in Figure 18 and can be seen that almost follows the same pattern as that of the RMSE graph. The NMSE of MH-CELM is both minimum and almost constant along with the number of trials.

The results shown in the graph is a proof to the fast learning ability, lower error rates and generalization capability of ensembles and also the better prediction capabilities of complex valued neural networks. However still variation in the performance exists in the figure 16-18.this is because of random parameter used in each CELM for the initialization of scaling parameter, center of neuron and the weight between hidden and output layer.

6. CONCLUSION

In this paper a method Holomorphic based multiple complex extreme learning machine (MH-CELM) has been applied for wind's speed and direction forecasting. At first wind speed and direction has represented in complex coordinates and then a Conformal transformation has applied to map the complex valued data in more convenient and simple complex plane. By varying the activation function each sample is presented to different ELMs to achieve better generalized prediction. And finally Bagging method has applied to combine the result of ensembles. Performance of MH-CELM shows better prediction compared to real valued ELM, FC-RBF and a single model of C-ELM. This proves the significant use of conformal mapping in complex learning algorithm.

7. REFERENCES

[1] L. Xingpei, L. Yibing, X. Weidong, "Wind speed prediction based on Genetic Neural network", Industrial Electronics and Applications,. 4th IEEE Conference on ICIEA, 2009, pp. 2448-2451.

[2] R.S. Li, Wunsch DC, O'Hair EA, Giesselmann M., "Using neural networks to estimate wind turbine power generation". IEEE Transaction on Energy Conversion, Vol. 16(3), 2001, pp. 276–82.

[3] T. Nitta, "On the inherent property of the decision boundary in complex-valued neural networks", Neurocomputing, Vol. 50, 2003, pp. 291 – 303

[4] Ming-Bing Li, Guang-Bin Huang, P. Saratchandran, N. Sundararajan, 'Fully Complex Extreme Learning Machine', Neurocomputing, Vol. 68, 2005, pp. 306-314

[5] R. Savitha, S. Suresh, N. Sundararajan, 'Fast Learning Circular Complex-valued Extreme Learning Machine (CC-ELM) for real-valued classification problems', Elsevier Journal on Information Sciences, Vol. 187, 2012, pp. 277-290

[6] T. Clarke, Generalization of Neural Network to the Complex Plane, Proceedings of international jt. conference on neural networks, vol. 2; 1990. p. 435–440

[7] N. Ozdemir, B. B. Iskender ,N. Y. Ozgur, "Complex valued neural network with Möbius activation function", Commun Nonlinear Sci Numer Simulat vol. 16; 2011. pp. 4698–4703

[8] A. Hirose, "Complex-Valued Neural Networks", Springer 2006.

[9] M.P. Perrone, L.N. Cooper, When Networks Disagree: Ensemble Methods for Hybrid Neural Networks, Technical Report A121062, Brown University, Institute for Brain and Neural Systems, January 1993

[10] L.K. Hansen and P. Salmon, "Neural network ensembles", IEEE Transactions on Pattern analysis and machine intelligence, Vol. 12 (10), October 1990.

[11] L. Breiman, "Bagging predictors," Mach. Learn., Vol. 24, 1996, pp. 123–140.

[12] T. Kitajima, T. Yasuno, "Output Prediction of Wind Power Generation System Using Complex-valued Neural Network" SICE Annual Conference,Taipei, August 18-21, 2010

[13] D.P. Mandic, S. Javidi, S.L. Goh, A. Kuh, K. Aihara, 'Complex-valued prediction of wind profile using augmented complex statistics', Elsevier Journal on Renewable Energy, Vol 34, 2009, pp. 196-201.

[14] K. Subramanian, R. Savitha, S. Suresh, "Complex-valued Neuro-Fuzzy Inference System for Wind Prediction" IEEE World Congress on Computational Intelligence WCCI, Brisbane, June 10-15, 2012

[15] E. Sathish, M. Sivachitra, R. Savitha & S. Vijayachitra, 'Wind Profile Prediction using a Meta-cognitive Fully Complex-valued Neural Network', IEEE-Fourth International Conference on Advanced Computing, ICoAC 2012, Chennai, Dec 13-15.

[16] 16. R. Savitha, S. Suresh, N. Sundararajan, "A fully complex-valued radial basis function network and its learning algorithm", International Journal of Neural Systems 19 (4) (2009) 253 –267.

[17] R. Schinzinger, A.A. Patricio, 'Conformal Mapping: Methods and Applications , Courier Dover Publications, 2003

[18] J. Mathews, R. Howell, "Complex Analysis for Mathematics and Engineering" , Jones & Bartlett Publishers, 2010

[19] T. Kim, T. Adali, Fully-complex multilayer perceptron network for nonlinear signal processing, Journal of VLSI Signal Processing Vol. 32 (1–2), 2002, pp. 29–43

[20] T. Kim, T. Adali, "complex backpropagation neural network using elementary transcendental activation functions", IEEE transaction 2001.

[21] http://mesonet.agron.iastate.edu/request/awos/1hr.php Iowa (USA), Department of Transport

[22] J. Kamruzzaman & R.A. Sarker, 'Forecasting of currency exchange rates using ANN: A case study', IEEE International Conference on Neural Networks & Signal Processing, China, Dec 14-17, 2003.

[23] G.B. Huang, Q.Y. Zhu, C.K. Siew, Extreme learning machine: theory and applications, Neurocomputing, Vol. 70 (1–3), 2006, pp. 489–501.

[24] S.J. Lin, C. Chang, M.F. Hsu, Multiple extreme learning machines for a two-class imbalance corporate life cycle prediction, Knowledge-Based Systems, Vol. 39, 2013, pp. 214‑223