Wind Speed Forecasting using ANN, ARMA and AIC hybrid to Ensure Power Grid Reliability

Diksha Sharma Auckland University of Technology Auckland, New Zealand dikaur17@gmail.com

Tek Tjing Lie Auckland University of Technology Auckland, New Zealand tek.lie@aut.ac.nz

Nirmal-Kumar C Nair University of Auckland Auckland, New Zealand n.nair@auckland.ac.nz

Brice Vallès Inland Revenue Auckland, New Zealand b.valles@auckland.ac.nz

***Abstract—*Reducing carbon emissions has accelerated the use of various renewable resources for electricity generation. Wind generation, in this context has seen increasing installations globally. Managing the intermittency of wind towards existing power system operation and control therefore becomes crucial. One effective solution is to predict the future values of wind power production. This paper focuses on the improvement of the present forecasting methods and reduction of the forecasting error. This paper uses Auto Regressive Moving Average (ARMA) to predict wind speed. However, order estimation of ARMA is a crucial issue. Artificial Neural Network (ANN) has been used for parameter estimation, which is then combined with Akaike Information Criteria (AIC) for order estimation. A simulation study has been conducted by comparing the proposed hybrid results with Genetic Algorithm (GA) for parameter estimation and an exhaustive search for order estimation.**

Keywords—ARMA; ANN; AIC; MATLAB

# INTRODUCTION

Environmental considerations have led to a strong interest and an increase in the use of various renewable resources for electricity generation. Wind is a clean, renewable resource that has gained increased attention in the past decade.With this increasing popularity and utilization of wind power technology, the cost of new energy resources has already approached the conventional energy one and thus, it has the potential ability to compete with traditional power plants [1].

Managing the intermittency of wind towards existing power system operation and control therefore becomes crucial. This intermittent nature makes the output power of wind farms difficult to control. It becomes the greatest challenge while integrating wind power into the electric grid. This also results toward increasing needs for power systems regulation and reserves requirement to ensure stability and reliability. Thus, this issue has a great impact on grid security, system operation, and market economics. Future wind penetrations up to 20% of system peak loads, alongside the existing uncertainty in wind predictions, will result in large increase in system-operating costs such as unit commitment.

An effective solution is to predict the future values of wind power production, for which the most important factor responsible is the local wind speed. So, if more consistent and robust forecasting of wind speed/power can be established, from a few minutes to several hours ahead, the effective utilization of wind power and generation of electricity can be ensured. Since wind power penetration is increasing, more sophisticated and innovative approaches are required.

This paper focuses on the improvement of the present forecasting methods and the reduction of the forecasting error. The objective of this paper is to develop a novel wind speed forecasting technique, which produces more accurate predictions. A hybrid method composed of the Artificial Neural Network (ANN) and Akaike Information Criteria (AIC) along with the Auto Regressive Moving Average (ARMA) technique is proposed. Simulation studies were conducted to show the effectiveness of the proposed method. The simulation study results of the proposed hybrid have shown very good results in the forecasting error for the test data given in this paper.

One of the main challenges is to properly determine the order of ARMA that would optimize the results. To find the optimum ARMA order, another method has to be used. Instead of the most basic Box–Jenkins (BJ) identification (which requires all calculations to be repeated when new pieces of data arrive [16]) or similar methods, this paper comes out with a very novel identification method that uses ANN and AIC to determine the order of ARMA. The comparisons are made with the Genetic Algorithm (GA) method and are based on the value of the coefficients obtained. The results show that the ANN computes the coefficients of an ARMA system accurately. This paper uniquely combines ANN with AIC to determine the true order of ARMA and bring down the forecasting error. Consequently, the proposed method will not only benefit the wind forecasting community but also can be adapted towards load forecasting and price forecasting methods in power systems.

# LITERATURE REVIEW

Wind speed forecasting is a very active and specialized research and development area. Generally, the forecast method developed for one site or wind farm does not suit another due to several factors, e.g., change in terrain and different wind speed patterns. Extensive research has been reported to decrease error but there is still a lot of scope to improve the short term forecasting techniques.

The Ensemble Kalman Filter (EnKF) method is explained in [2]. Reference [3] uses EnKF to predict wind speed that gives an error of about 5.06% but only for 10 min ahead forecasts. It gives significant errors beyond that timeframe. Linear prediction model of the ARMA is shown in [4] where results from 1s to 5s only are reported. A hybrid between the ARMA model and the ARCH (Auto Regressive Conditional Heteroscedasticity) is used in [5], in which comparison is done with the classic Persistence model. The results however show high Mean Absolute Error (MAE) of up to 29.37%. This paper does not mention the true order of ARMA though. When the Wavelet Theory is combined with the neural network satisfactory results are obtained [3]. Reference [1] shows a similar result when the WT is combined with an empirical mode decomposition (EMD) method showing a 4.53% Mean Absolute Percentage Error (MAPE).

Grey Model is a time series model and is used in [6] to predict the next 3 hours. There is very little research done in this area and this paper only shows a pilot study, which does not give small errors. Developing a generalized, portable and easy-to-operate model is a big challenge. Such a method is presented in [7]. It has a strong robustness giving 40% better results than the Persistence model but fails to give a good MAPE.

Reference [8] shows that over fitting is a result of excessive complexity and lack of clear guidelines on how many parameters in a model should be taken, for a given sample size. A case study from Tasmania, Australia is shown in [9] for very short-term forecasting in which more importance is given to wind direction rather than wind speed. Numeric Weather Prediction (NWP) models are used for up to 6 hours ahead forecasts and show better results when weather conditions are stable. NWPs include global forecasting systems [10] and High Resolution Limited Area Model (HIRLAM) [11] for instance. The method in [12] uses a new statistical approach that combines artificial intelligence and fuzzy logic techniques. The errors vary for weather conditions and can still be drastically reduced.

Reference [13] shows a GA fitness value relying on the deviation between the actual plant output, with or without an additive noise, and the estimated plant output. Simulation results show in detail the efficiency of the proposed approach however it does not directly tell us the best order. Reference [15] combines the effectiveness of the Multi Model Partitioning theory with the robustness of Evolutionary Algorithms but gives a bit complicated approach with a 10% error.

Reference [17] uses Artificial Neural network (ANN) technique and compares it with some known and widely acceptable techniques. The comparisons are entirely based on the value of the coefficients obtained. The results show that the use of ANN also gives an accurate computation of the coefficients of an ARMA system. But it only stops to the determination of the parameters and not the true order.

Reference [19] shows that the ARMA model based on System Identification Toolbox of MATLAB is valid to forecast wind signal and can reflect the future characteristics of the signal. The average relative error of the model is found to be 6.9%, which can be drastically reduced. Reference [9] uses ANN but fails to give any explanation on the analysis of the algorithm used. Reference [18] proposes to utilise ANN for order identification but uses Extended Sample Auto Correlation Function (ESACF) method to determine the order of ARMA. However, the proposed method finds it hard to estimate the proper order when the *p* and *q* values become too great. In addition, it takes a long learning time. The present work proposes to combine ANN with AIC and the simulation studies are conducted to show that the proposed hybrid method gives much satisfactory results. The gaps in these papers are filled with the proposed hybrid.

There are several publications that report to predict ARMA and its true order but the combination of ANN- Akaike Information Criteria has not been used for order estimation. Therefore, in this paper the ARMA model is combined with the ANN and AIC to give good and accurate results.

# METHODOLOGY

## ARMA

An ARMA (*p*, *q*) model, with *ak* and *bk* the autocorrelation and the moving average coefficients respectively, is represented by

(1)

The order of the ARMA model is included in parentheses as ARMA (p, q), where *p* is the autoregressive order and *q* the moving-average order. In Equation (1), *{y (n)}* is the return series of original time series. In the process of constructing an ARMA model, it is imperative to confirm the order *p* and *q* [20].

ARMA parameters may be obtained from three-layer neural networks utilizing a polynomial representation of the activation function in the hidden units. Consider a nonlinear, time-invariant, discrete-time dynamic system represented by the following ARMA model shown in Fig. 1.

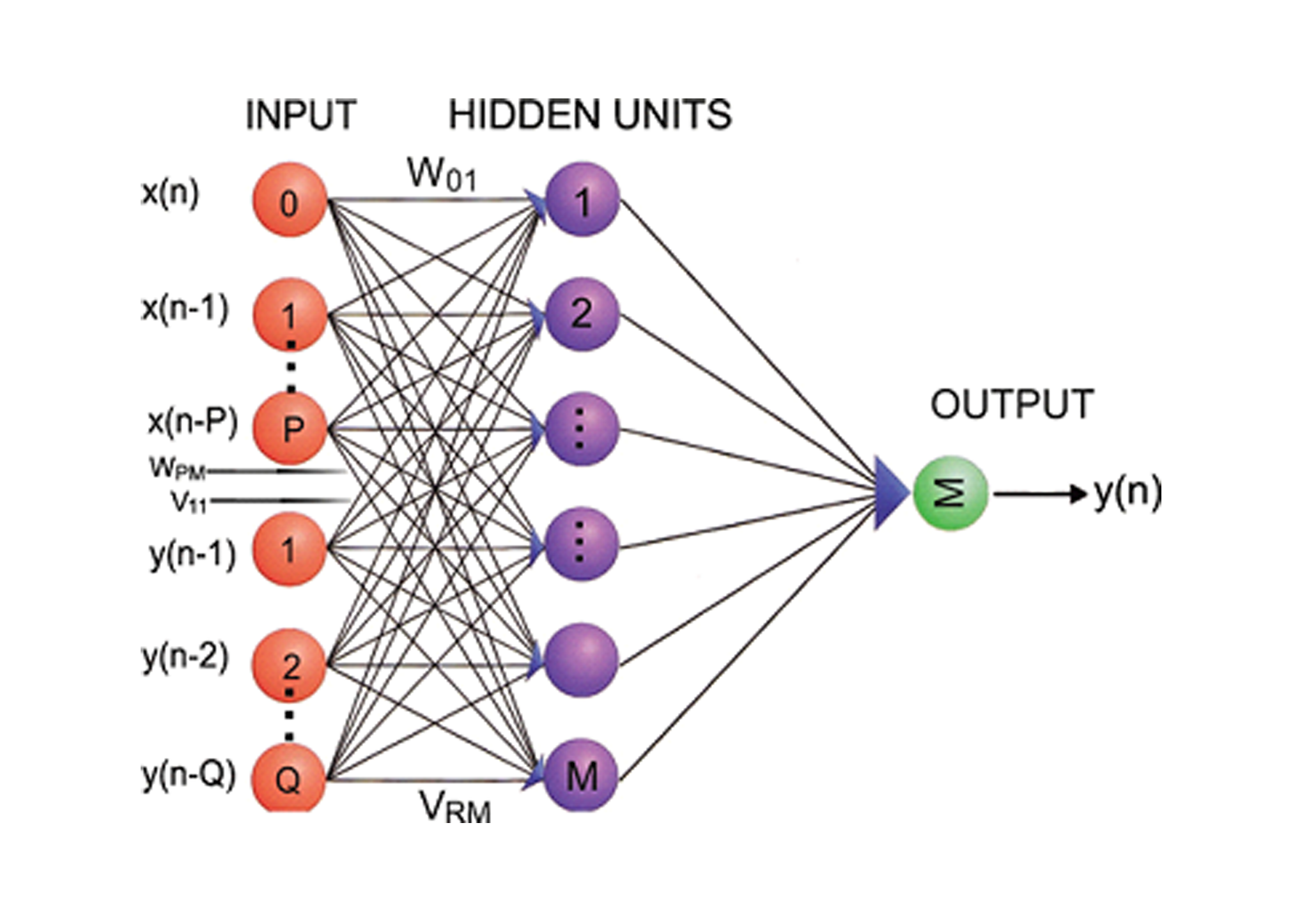
## ANN

Fig. 1 shows a three-layer ANN topology. Note that weights of *x* input neurons are W leads and weights of *y* input neurons are *V* leads. The coefficients are obtained from the neural network weights values and polynomial coefficients given by





where *M* is the number of hidden units, *i, k, m*, and *n* are indexes [17].



1. A three layer ANN topology. Note that the weights of x input neurons are W leads and the weights of y input neurons are V leads.

These model parameters are obtained using ANN, which are then fed into the Akaike Information Criteria to obtain the true order.

## AIC

The AIC criterion can be defined as  
 

where is the error variance of fitted model and *N* is the number of observations estimated under the assumption that *d* is the true order [14] [18].

For the ARMA (*p*, *q*) model, *d* is the model order where *d = p + q.* In practice, the optimal *d* is obtained by minimizing AIC (*d*) where

****

where,

   

and   is given by



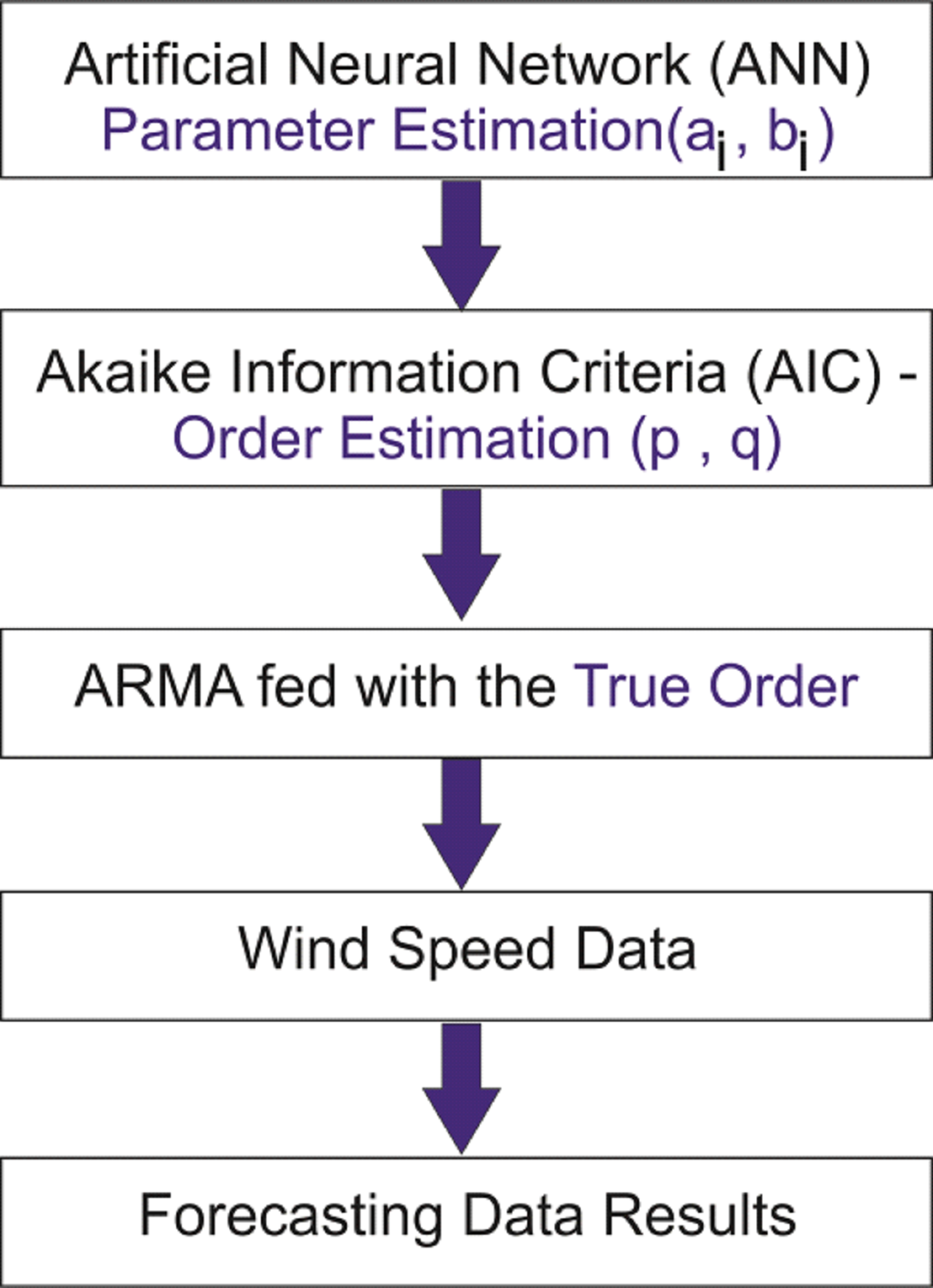
The value of AIC is obtained for different *p* and *q*. The one with the least value of AIC is considered to be the true order. It can be explained with the help of Fig. 2.

According to Fig. 2, the parameters are obtained using (2) and (3) of ANN. Then these model parameters are fed into (4) of Akaike to estimate the true order of ARMA. Then, ARMA takes the true order and forecasts wind speed data. For the sake of completion, it is imperative to account for the following steps

1) First, we fix the range of the ARMA model orders to be considered (*1 ≤ p ≤ pmax* and *1 ≤ q ≤ qmax*)

2) Then, estimation of the model parameters by an appropriate algorithm such as ANN. Compute the predicted signal and keeping the error variance equal to one mimicking what has been done in [17].

3) Evaluate AIC over *S=(1 ≤ p ≤ pmax),* and choose the value of *d* that minimizes AIC to find the true order of ARMA



1. Procedure of the ANN-ARMA-AIC hybrid

# RESULTS

In this research, wind data and the input-output patterns for training of a network are taken from North Island, New Zealand for the year 2006 on a hourly basis for station number STH1, as provided by NIWA (NIWA is a crown-owned research and consultancy company, with a global reputation as experts in water and atmospheric research, http://www.niwa.co.nz). The platform chosen to test the available set of data is MATLAB. The six hourly wind fields are extracted from the data. The test data has an error of 0.8%.

During the experiments, 1500 input/output data values are employed for the simulation. The first 750 data values are used for the training (these become the training data set), while the others are used as checking data for validating the ARMA model.

For the first part ANN is used to determine the model parameters. We use a polynomial function neural network (PFNN) model for training the simulated data to identify the coefficients of ARMA models. The max value of *p* and *q* have been taken equal to 6, giving a total of 36 possible combinations for *p* and *q*. The best order has been found to be *(p=3, q=2)* therefore the results for this order have been shown here. For this order the ARMA model was generated, with the Moving Average (MA) excitation being uncorrelated Gaussian white noise (GWN) with a variance of one. The following linear ARMA model was utilized:

y(n) = x(n) + 0.358x(n−1) + 0.212x(n−2) + 0.343x(n−3) +0.201y(n−1)+0.301y(n−2) 

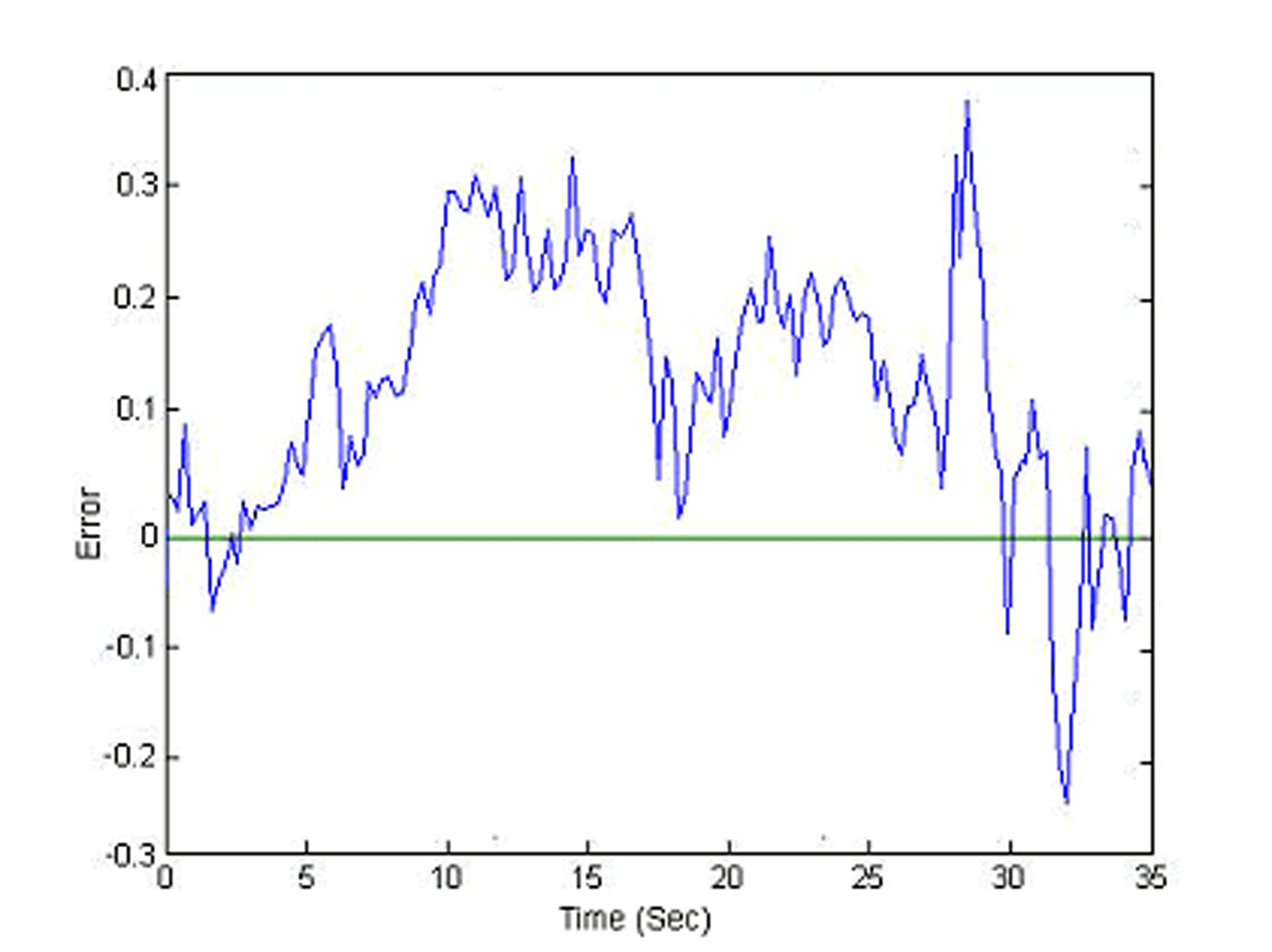
For the PFNN analysis, the input and output data pair was segmented into two 750-point data segments. The first half of the input–output data segment has been used to train the network, and the second half was used to test the predictive quality of the network. All of the simulations were carried out in this manner. Since the PFNN method utilizes a back-propagation learning algorithm, the time required to compute parameter estimates is less than GA. It should also be noted that the learning rate of the neural network is affected by the selection of the ARMA model order (or the selection of the memory length), that is, longer training of the network is required with larger ARMA model orders, and vice versa.

To examine how strongly the neural network topology depends on the assumed model order, the results obtained by the PFNN method and Genetic Algorithm have been compared. To check the validity of the results from (2) and (3), the simulations were also carried out for different combinations of *1 ≤ p ≤ 6* and *1 ≤ q ≤ 6*. The results for the order *p=3* and *q=2* were chosen, corresponding to the best case as it will be shown in the latter section. Table 1 shows that both methods gave very similar parameter values.

table 1. COMPARISON OF PARAMETERS USING ANN AND GA

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Parameters | Parameters for order (3,2) | | | | |
| a1 | a2 | a3 | b1 | b2 |
| ANN | 0.3 | 0.21 | 0.34 | 0.2 | 0.3 |
| GA | 0.35 | 0.21 | 0.3 | 0.21 | 0.3 |

Using Table 1, the time series were computed corresponding to the parameter values obtained for ANN and GA. The difference between the two time series is plotted in Fig. 3. This figure shows that there is no significant difference between the two outcomes thereby indicating the validity of the proposed approach when compared to GA.



1. Difference between the parameters for ANN and GA

Next, these parameter values are put in the equation of AIC to determine the true order of ARMA. An exhaustive search for different *p* and *q* has also been done to compare the results. It is shown in Table 2. The matrix gives the AIC values for different *p* and *q* values. It can be seen that the least values are obtained for *p=3* and *1 ≤ q ≤ 4.* The least value is obtained for *p=3, q=2.* The other combinations have a greater AIC value.

TABLE 2 THE EXHAUSTIVE SEARCH RESULTS FOR DIFFERENT *p* AND *q*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| q  p | 1 | 2 | 3 | 4 | 5 | 6 |
| 1 | 329.3 | 329.4 | 328.9 | 330.0 | 330.3 | 330.4 |
| 2 | 328.1 | 329.7 | 329.8 | 328.5 | 330.3 | 331.2 |
| 3 | 327.1 | 325.9 | 326.5 | 327.3 | 328.9 | 330.3 |
| 4 | 333.0 | 331.2 | 333.3 | 334.1 | 332.4 | 337.1 |
| 5 | 335.2 | 335.4 | 337.5 | 339.1 | 340.1 | 340.5 |
| 6 | 340.3 | 342.4 | 341.3 | 344.4 | 343.5 | 346.6 |

The procedure for true order determination by AIC is summarized as:

1. The different parameters are fed into AIC.
2. The different values of AIC are determined for different *p* and *q* orders.
3. The order that gives the least value of AIC is considered to be the best order.

The AIC values have been calculated for *1 ≤ p ≤ 6* and *1 ≤ q ≤ 6.* The range of AIC values obtained is between 326 and 345. The best values are observed to be between *p=3* and *1 ≤ q ≤ 4* and these AIC values are given in Table 3.

Table 3 AIC VALUES

|  |  |
| --- | --- |
| Model Order | AIC for p=3, 1 ≤ q ≤ 4 |
| AIC |
| (3,1) | 327.1 |
| (3,2) | 325.9 |
| (3,3) | 326.5 |
| (3,4) | 327.3 |

When ARMA is applied to the test data of station STH1 provided by NIWA with *p=3* and *q=2,* the forecasting error is close to 0.75% which is very close. If we propose a hybrid including ARMA with the correct *p* and *q* and another method then it may bring down the error even further, which opens doors for future researchers.

# CONCLUSION

The effective solution to accommodate intermittency of wind is to predict the future values of wind power production, which is totally dependent on wind speed. Changing the prediction model to decrease the forecast error is one way and has been chosen in this paper. This is alongside a hybrid method that can accurately predict wind speed from a few minutes to several hours. The ARMA model has been chosen which has proved to be very accurate for long term load forecasting and price forecasting. However, order estimation of ARMA is of utmost importance. Therefore, ANN has been chosen and has been combined with AIC to estimate ARMA true order. The proposed ANN-ARMA-AIC hybrid has been tested on real data from NIWA on MATLAB. An exhaustive search has been carried out for different combinations of *1 ≤ p ≤ 6* and *1 ≤ q ≤ 6*. The results have been compared with the Genetic Algorithm that shows no significant difference between the parameter estimation of ANN and GA thus validating the results. We showed that the time required to compute estimates of the parameters is less than GA. The results also showed that if a proper sequence is followed then ANN can be used to estimate the parameters and in turn AIC can be used to identify the model order correctly. This research may not only benefit forecasting of wind but also several other applications as well such as load forecasting or flood forecasting for instance.

# REFERENCES

1. Liu D, Li H, Ma Z “One hour ahead prediction of wind speed based on data mining,” Paper presented at the *2nd international symposium on the Advanced Computer Control*, 27-29 March 2010.
2. Evensen G (2009) Data Assimilation : The Ensemble Kalman Filter. Berlin: Springer Verlag.
3. Wei W, Wu G, Yang M, Zhang Y, Qiu S, Sun A “Short-term forecasting for wind speed based on wavelet decomposition and LMBP neural network,” Paper presented at the *4th international symposium on Electric Utility Deregulation and Restructuring and Power Technologies*, 6-9 July 2011.
4. Negnevitsky M, Johnson P, Santoso S “Short term wind power forecasting using hybrid intelligent systems,” In:*Abstracts of the symposium on the IEEE Power Engineering Society General Meeting*, 24-28 June 2007.
5. Gao S, He Yu, Chen H “Wind speed forecast for wind farms based on ARMA-ARCH model,” Paper presented at the *international symposium on the Sustainable Power Generation and Supply*, 6-7 April 2009.
6. Huang C Y, Liu Y W, Tzeng W C, Wang P Y “Short Term Wind Speed Predictions by Using the Grey Prediction Model Based Forecast Method,” Paper presented at the *international symposium on the Green Technologies* -15 April 2011.
7. Fan S, Liao J R, Yokoyama R, Chen L, Lee W J “Forecasting the Wind Generation Using a Two-Stage Network Based on Meteorological Information,” J Mol Med 24(2): 474-482, 2009 .
8. Hippert H S, Pedreira C E, Souza R C “Neural networks for short-term load forecasting: a review and evaluation,” J Mol Med 16(1): 44-55, 2001.
9. Potter C, Negnevitsky M “Very short-term wind forecasting for Tasmanian power generation,” In: *Abstracts of the international symposium on the IEEE Power Engineering Society*, 2006
10. Senjyu T, Yona A, Urasaki N, Funabashi T “Application of Recurrent Neural Network to Long-Term-Ahead Generating Power Forecasting for Wind Power Generator,” Paper presented at the *international symposium on the IEEE PES Power Systems and Exposition*, 29 Oct- 1 Nov 2006.
11. Brusch S, Lehner S, Schulz-Stellenfleth J “Synergetic Use of Radar and Optical Satellite Images to Support Severe Storm Prediction for Offshore Wind Farming," J Mol Med 1(1) 57-66, 2008.
12. Sideratos G, Hatziargyriou N D “An Advanced Statistical Method for Wind Power Forecasting,” J Mol Med, 22(1): 258-265, 2007.
13. Akramizadeh, A., Farjami, A. A., Khaloozadeh, H. “Nonlinear Hammerstein model identification using genetic algorithm,” *Symposium conducted at the meeting of the Artificial Intelligence Systems*, 2002.
14. Aksasse, B., Radouane, L. “Two-dimensional autoregressive (2-D AR) model order estimation,” IEEE Transactions on Signal Processing*,* J Mol Med *47*(7) 2072-2077, 1999.
15. Beligiannis, G. N., Likothanassis, S. D., Demiris, E. N. “A robust hybrid evolutionary method for ARMA model identification,” *Symposium conducted at the meeting of the Image and Signal Processing and Analysis, 2001*. ISPA 2001.
16. Chatterjee, S., Bard, J. F. “A comparison of Box-Jenkins time series models with autoregressive processes,” Systems IEEE Transactions on Man and Cybernetics*,* J Mol Med *15*(2) 252-259, 1985.
17. Chon, K. H., Cohen, R. J. “Linear and nonlinear ARMA model parameter estimation using an artificial neural network,” IEEE Transactions on Biomedical Engineering*,* J Mol Med *44*(3), 168-174, 1997.
18. Wang, L., Libert, G. A. “Combining pattern recognition techniques with Akaike's information criteria for identifying ARMA models,” Signal models., IEEE Transactions on Signal Processing*,* J Mol Med 42(6), 1388-1396, 1994.
19. Shiqiong, Z., Jixuan, Y., Zhumei, S., Jun, T., Longyun, K. “Wind Signal Forecasting Based on System Identification Toolbox of MATLAB,” *Symposium conducted at the meeting of the Intelligent System Design and Engineering Applications (ISDEA*), 2013.
20. Cao L, Li R “Short-term wind speed forecasting model for wind farm based on wavelet decomposition,” Paper presented at the *3rd international symposium on the Electric Utility Deregulation and Restructuring and PowerTechnologies*, 6-9 April, 2008.