REAL TIME PREDICTION OF SHIP MOTIONS AND ATTITUDES USING ADVANCED PREDICTION TECHNIQUES

Ameer Khan, RMIT University, Melbourne, Australia,
Cees Bil, RMIT University, Melbourne, Australia,
Kaye Marion, RMIT University, Melbourne, Australia,
Malcolm Crozier BAE Systems Australia Melbourne, Australia

Keywords: *Ship, Neural networks, Modeling, Combinations*

Abstract

The deployment and the landing of vehicles from a ship in an open water environment can often be difficult and even dangerous due to the random nature of the ship's motion. The ability to predict reliably the motion will allow improvements in safety on board ships and facilitate more accurate deployment of vehicles off ships. This paper presents an investigation into the application of the auto-regressive moving average method and the artificial neural network methods for the prediction of ship motion. It is shown that the artificial neural network is superior to auto-regressive moving average techniques and is able to predict the ship motion satisfactorily for up to 10 seconds. Also presented in this paper is the combination approach that combines multiple time series prediction techniques potentially to deliver better overall results than the individual techniques alone.

1 Introduction

An algorithm capable of predicting the motion of a ship is required for the successful deployment of a ship system currently used on ships that operate in open sea environments. The predicted motion and attitude of the ship will be transmitted to the ship system to ensure successful activation. With the predicted ship motion, the correct flight conditions can be calculated thereby allowing successful and safe deployment of the ship system.

A key requirement of the algorithm is to predict in advance by x, x+2 and x+4 seconds (where x=3-4 seconds) whether the angles are likely to exceed the "launch lock-out value" which is the condition where the system cannot be activated. When it is predicted that the angles exceed the launch lockout value the algorithm will automatically select another ship system that is not in lockout. Otherwise, the algorithm allows activation of the ship system. Therefore it is necessary to predict the maximum and minimum roll angles or the turning points in the motion. It is important that the predicted angles are of a high accuracy as the batteries for the system are "one shot" batteries, which means that the process of deployment once activated cannot be reversed.

The motion of a ship in an open water environment is the result of complex hydrodynamic forces between the ship, the water and unknown random processes. This leads to the necessity to use statistical prediction methods for the prediction of this motion rather then a deterministic analysis, which would lead to a ship specific model that involves highly complex calculations [1].

2 Rationale for Performing the Research

The ability to predict the ship motion reliably in any sea state will enable better control of systems operated off ship platforms. For example, the landing and take off of helicopters and aircraft whether manned or unmanned from ship decks in rough sea conditions can be difficult and at times dangerous.

If the motion of a ship can be predicted with reasonable error bounds and communicated to the aircraft or helicopter, touchdown dispersion can be improved on landing and a smoother aircraft trajectory can be achieved on take off. Prediction of ship motion is also important for the deployment of missiles and remote piloted vehicle from ship platforms as shown Fig. 1 and Fig. 2 for the correct trajectory calculation [2]. In some cases there is a launch “lock out” condition where the missile or remote piloted vehicle cannot be launched safely.

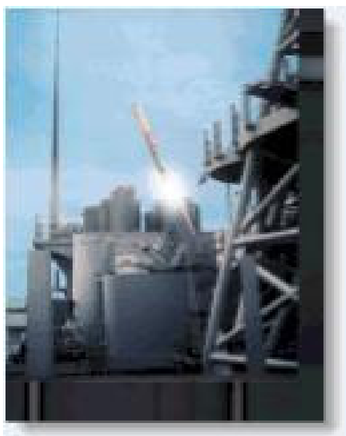

Fig. 1. Deployment of a remote piloted vehicle

Fig. 2. A Harpoon surface-to-surface missile is fired from the USS New Jersey.

It is envisaged that the predictive paradigm developed will have high accuracy as well as high efficiency so there is also the possibility that the methodology developed can be applied to other scenarios that require the prediction of a stochastic harmonic process in real time. In particular the paradigm could be used in areas such as economics, acoustics and other areas of interest where there is stochastic harmonic motion present.

3 Background investigations of ship motion prediction techniques

There are numerous theories available that are related to time series analysis and predictive algorithms. Many investigations have sought to apply these prediction models to predict the motion of ships in real time and have only obtained a limited degree of success. The following sections give a description of the various techniques used for the purposes of time series prediction and give the results of their application to ship motion.

4 Auto-Regressive Moving Average (ARMA) Models

The ARMA model is widely used in the field of forecasting and there are a large number of variations of the model proposed in various literatures over the years [3]. The ARMA method, as defined by Box and Jenkins [4] essentially combines the autoregressive and moving average methods. It is a statistical method that finds the time domain model that best fits an input wave sensor time history to the ship response time history. The ARMA method basically finds the best statistical fit to the time history values using time domain coefficients for a stationary process.

To describe these methods it is first necessary to have a time series with measurements taken at equally spaced time intervals (t, t-1, t-2,…) with the associated values (z_t, z_{t-1}, z_{t-2},…). Now, we define $\tilde{z}_t, \tilde{z}_{t-1}, \tilde{z}_{t-2}$ as the deviations from the mean μ. (i.e. $\tilde{z}_t = z_t - \mu$). The Auto-Regressive model becomes:

$$\tilde{z} = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \ldots + \phi_p z_{t-p} + a_t \qquad (1)$$

where:

$$\phi(B) = 1 - \phi_1 B - \phi_1 B^2 - \ldots - \phi_1 B^p \qquad (2)$$

Therefore, the current value of the ship motion in this model is expressed as a finite, linear aggregate of previous values of the motion and a shock or error term a_t [4].

The Moving Average model as defined by Box and Jenkins [4] is:

$$\tilde{z}_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \ldots - \theta_q a_{t-q} \qquad (3)$$

where:

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q \qquad (4)$$

Combining the two, one obtains the mixed Auto-Regressive, Moving Average model:

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \ldots + \phi_p \tilde{z}_{t-p} + a_t - \theta_1 a_{t-1} - \ldots - \theta_q a_{t-q} \qquad (5)$$

A difference of the process is taken to ensure stationary behavior and somewhat eliminate trends. The *d*th difference of the process is referred to as the autoregressive integrated moving average model abbreviated as ARIMA (*p, d, q*) and is described in detail by Box and Jenkins [4].

An extension of this method is to use a multivariate AR process. The multivariate AR process uses interdependence that may exist between two of the available readings on the ship motion. Much of the theory used for the univariate analysis extends to the multivariate case as discussed in Brockwell and Davis [5].

5 Artificial Neural Networks (ANN)

Artificial neural networks (ANNs) form a class of systems that are inspired by biological neural networks [6]. A neural network is simply a series of neurons that are interconnected to create a network. They are a class of non-linear systems and there are a wide variety of different approaches that can be used. The use of ANNs in time series prediction relate to the application of ANNs for the nonlinear system identification. The use of ANNs is particularly appealing due to the ability of the ANNs to learn and adapt which will be important for this investigation as one of the underlying goals is to create an algorithm that is able to work in all conditions and environments.

The following description and analysis of ANNs follows Masters [7]. The ANN architecture that will be used to create the ANN for time series prediction will be the multi-layer feed-forward ANN. This type of architecture has a minimum of two layers consisting of the input layer and the output layer. Commonly there are a number of hidden layers between the input and the output layers. Neurons in the input layer are hypothetical in the sense that they have no inputs and do no processing. In a feed-forward ANN the inputs for each layer come from the preceding layer. A single neuron is shown below:

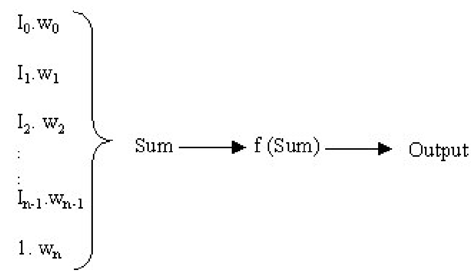

Fig. 3. A representation of a single neuron

It has n inputs including a bias term, which has been set to 1 in this investigation. The inputs are each multiplied by their corresponding weight value, which are summed together and subsequently entered into an activation function. The output of the activation function will correspond to the output of the neuron.

Mathematically, the output of a neuron is given as

$$out = f\left(net\right) = f\left(\sum_{i=0}^{n-1} x_i w_i + w_n\right) \qquad (6)$$

where the inputs are $\{x_i, i = 0, \ldots, n-1\}$.

Generally the neuron's operation is not effected significantly by the activation function $\left(f\left(net\right)\right)$ but the training speed is effected somewhat. The activation function is usually a non-linear function that will determine the output of the neuron. Its domain is generally all real numbers. The range of the output for an activation function is usually limited between 0 to 1 and sometimes –1 to 1. The majority of activation functions use a sigmoid (S-shaped) function. Any continuous, real valued function whose domain is the set of real numbers, whose derivative is always positive and whose range is bounded can be approximated using a three-layered feed-forward ANN. In this investigation the activation function shown below was primarily used:

$$f\left(net\right)=\tanh\left(net\right) \tag{7}$$

The use of ANNs for time series prediction has a number of distinct advantages. Firstly, any amount of information pertinent to the prediction can be incorporated into the ANN. There is also no need to choose any particular model for the ANN. A validation process is included to ensure that the ANN is working correctly. It is to ensure that the ANN has not *overfitted* the data. If the architecture of the ANN is poorly designed, the ANN may be able to learn irrelevant details specific to the training set which will lead to an ANN that is only relevant to the training set. Conversely, the ANN may have a deficient architecture where the ANN is not able to learn the subtleties required for accurate outputs. The validation process should reveal these problems.

To validate the ANN a simple procedure is used. The training set is divided into two. One set of the training set is designated for training purposes only while the second set is designated as the validation set. The ANN is only trained using the training set and no data from the validation set is used while training. Once trained, the validation set is inputted into the ANN and the resulting predictions based on the validation set are used to measure the effectiveness of the ANN.

The basic model for time series prediction is shown in Fig. 4:

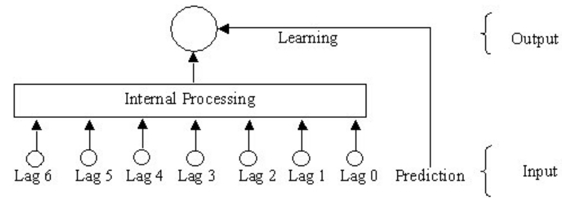

Fig. 4. Basic model for time series prediction using ANNs

Any neural network that is capable of accepting real valued vectors as input and producing real valued outputs may be used for time series prediction. In the above diagram it can be seen that there are seven points. Lag 0 represents the current sample while the past six values are represented by lags 1-6.

The prediction of the ANN will be the output and will be used for training the network. It can be noticed that there is only one output shown. For every lead prediction interval it is advisable to use a single ANN. If multiple predictions are required then for every prediction interval a separate ANN is used. The basis for the presumption is that the weights for an optimal prediction will vary according to the prediction interval desired. By having the ANN create multiple predictions, the overall optimal prediction could not be made. By having separate ANNs create separate predictions, the optimal weight configuration can be obtained for each prediction and therefore, higher accuracy can be expected.

5.1 Training the network

The training of the network can be viewed as a minimization process where the weights in the ANN are systematically adjusted in a manner that reduces the error between the output of the ANN and the desired output. The algorithm used to determine the minimum must ensure that global minimum is achieved and has not merely discovered the local minimum.

The back-propagation algorithm and the conjugate gradient methods are very capable of finding the local minimums but there is no guarantee that the global minimum will be found using these techniques. The methods of simulated annealing and genetic algorithm for minimizing functions are algorithms that are designed for global minimum searches. The following sections describe all of the minimization techniques mentioned above.

5.2 Back-Propagation

The back-propagation (BP) algorithm for training was first introduced by Rumelhart and McClelland [8]. Although other techniques have been developed that are more robust and efficient, the BP algorithm is used in many current ANNs and remains an effective training method.

The BP algorithm has simple mechanics that can be easily implemented. It works by progressively propagating the errors backward through the ANN starting from the output layer error. The BP algorithm works by determining the gradient of the multivariate function to determine how to vary the weights, which will achieve a reduction in the overall error in the output layer of the ANN. Therefore, the BP algorithm will determine the gradient and then alter the weights in the direction of the negative gradient. For this reason, the BP algorithm is also referred to as the gradient decent method [7]. This process is continued until a local minimum is reached.

Unfortunately, this method can easily find a local minimum rather than the global minimum, which is the value that is desired. Another associated problem with the BP algorithm is that it is difficult to determine how far the algorithm should step in the chosen direction. Small step size will ensure that the minimum will be found but the computational time may be too great. If larger step sizes are used, the minimum may be overshot and the error may actually increase as a result.

Commonly a momentum term is introduced into the BP algorithm to ameliorate the above problems somewhat by having each new search direction computed as a weighted sum of the current gradient and the previous search direction [7]. This modification does have the effect of increasing the convergence rate but it still relies on a user introduced momentum term. If the momentum term is too low then the algorithm will still be lethargic but if too great then the BP algorithm will not be able to effectively follow a complex function.

5.3 Conjugate-Gradient Algorithm

Conjugate-Gradient (CG) algorithm shares some commonality with the BP algorithm but is superior to the BP algorithm because it does not have the drawbacks outlined for the BP algorithm. The premise behind the conjugate gradient method is that the algorithm moves not down the new gradient but rather in a direction that is constructed to be conjugate to the old gradient and to previous directions traversed [9].

The algorithm created for the ANN in this investigation was based on the Polak-Ribiere algorithm. The mathematical justifications for the algorithm are beyond the scope of this investigation but a detailed description can be found in Polak [10]. In a general sense, the algorithm generates a sequence of vectors and search directions. It can be shown that the exact minimum will be attained if the multi-dimensional function can be expressed as a quadratic. The ANN error function is quadratic close to the minimum so it is expected that once close to a minimum, convergence to the local minimum will be very rapid [9].

5.3 Simulated Annealing

Simulated annealing is a Monte Carlo approach for minimizing multivariate functions. The inspiration for simulated annealing comes from the physical process of heating and then slowly cooling a substance to obtain a strong crystalline structure.

When applied to finding the global minimum of a given function the minima of the cost function corresponds to this ground state of the substance. In the simulated annealing process, the temperature is lowered in stages until the system reaches the zero temperature of freezing temperature at which point, no further changes occur. At every temperature in the annealing process, the system is allowed to attain a steady state or equilibrium, known as thermalization.

When implementing the simulated annealing procedure to ANNs it is preferable that the smallest number of trials is conducted for every weight to reduce computational time. In the simulated annealing process, after the iteration at a particular temperature is completed, the best weights become the focal point for the subsequent temperature.

Although the simulated annealing process is capable of finding the values for the minimum, it is more efficient to use simulated annealing to establish a rough approximation of the location of the global minimum and use the conjugate gradient method to ascertain the exact values of the weights. The reasoning behind this approach is that the simulated annealing method is inefficient in determining the exact values of the weights and would require the temperature to be arduously dropped at a slow rate. The conjugate gradient can rapidly find a minimum and if weights that generally approximate to the global minimum are inserted into the CG algorithm, then rapid progress to the global minimum is likely.

5.4 Genetic Algorithm

The genetic algorithm (GA) is a part of a rapidly growing area of artificial intelligence called evolutionary computing. The term 'evolutionary computing' is based on Darwin's theory of evolution, which states that problems are solved by an evolutionary process resulting in a best solution. It is basically survival of the fittest where the 'fittest' (best) 'survivor' (solution) evolves to create the next population [11].

Solution to a problem solved by the GA uses an evolutionary process. The algorithm begins with a population of solutions analogous to the chromosomes described previously. Solutions from one population are taken and used to form the next population. The expectation is that the new population will be better than the old one. Solutions or individuals are then selected to form new solutions or 'offspring' according to their fitness. As mentioned previously, the more suitable they are, the more chances they have to reproduce. This is repeated until a predefined stop criterion is satisfied. A flow chart of the process is given below in Fig. 5.

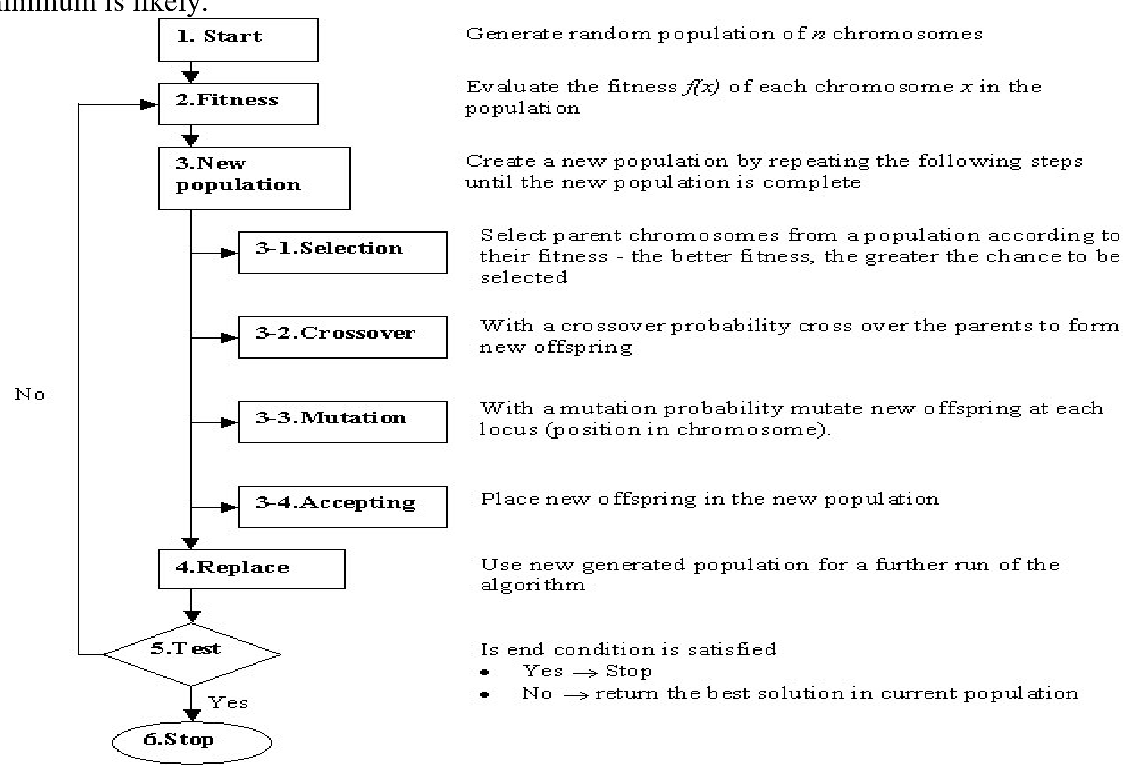

Fig. 5. Flow chart of the GA process

To rapidly converge to a solution, a combination of the GA and CG methods were used in a similar fashion to the way that the SA and CG methods were combined. As with the SA algorithm, the GA can efficiently locates the general set of optimum weights but the CG algorithm must be applied to quickly refine the weights to their precise value.

5.5 Prediction accuracy

Prediction tests were performed for a variety of different ANNs and ARIMA models. The architecture of the ANN can vary significantly depending upon the number of input neurons and hidden neurons used.

The univariate and multivariate ARIMA methods based on the Hannan Rissanen (HR) method and the Yule-Walker (YW) method respectively, which are both discussed in Brockwell and Davis [5], were applied to the ship motion data. The univariate ARIMA model had the parameter setting of ARIMA (15,0,1) and the multivariate ARIMA model had the parameter settings of ARIMA (10,1,0). All predictions were made on the roll motion while the multivariate analysis utilized the pitch and wind speed.

There were two performance criteria. The first statistic measures the number of predictions that were within the 95% confidence interval. The second measure gives a measure of the number of turning points (TP) or maximum and minimum roll angles accurately predicted within the 95% confidence interval at the actual TP. The performance statistics are shown in Fig. 6 and Fig. 7 for the predictive techniques applied.

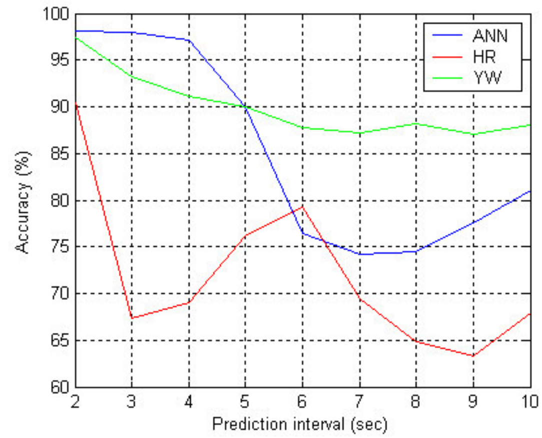

Fig. 6. Comparison of points within the 95% confidence interval

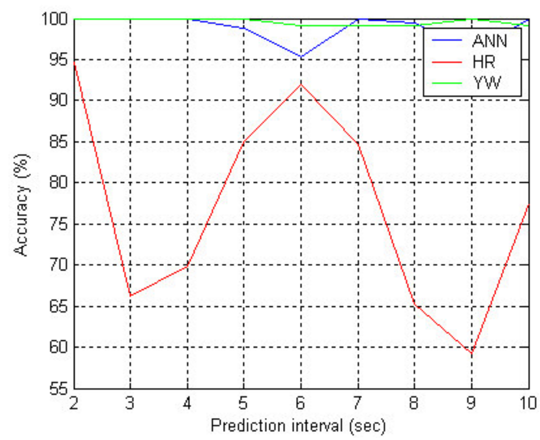

Fig. 7. Comparison of TP within the 95% confidence interval at the actual TP

6 Discussion on results

It can be clearly seen that both the multivariate YW and ANN prediction method give superior results to the univariate HR predictions. On closer inspection it is found that the statistics are slightly misleading. Below are the predictions made by the YW and ANN methods respectively.

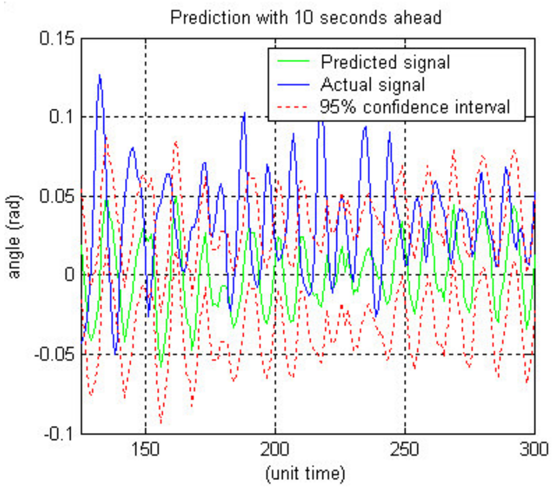

Fig. 8. 10 second YW prediction

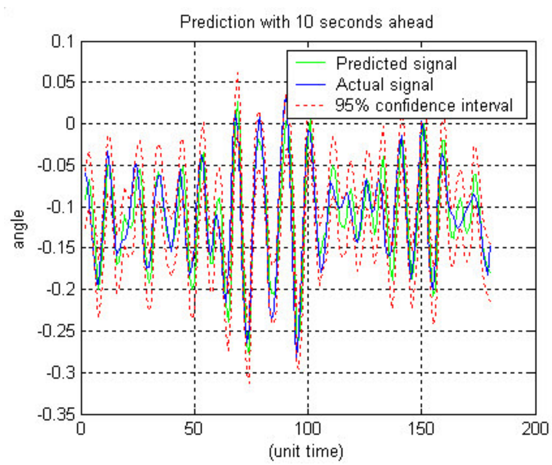

Fig. 9. 10 second ANN prediction

Visual inspection of the predictions leaves no doubt that the ANN prediction leads to a better representation of the actual roll angle then the predictions of the YW method. It can be seen that there is a slight phase difference between actual roll angle and the predicted roll angle in the ANN, which leads to a reduction in the percentage of points correctly predicted within the 95% confidence interval.

This explains the reason for the ANN accurately predicting the TPs in the data whereas the YW predictions were considerably worse.

7 Combination/Composite predictions

There is a growing emergence of combination forecasts being employed to further improve the accuracy and speed of predicted values based on time history data. The idea behind the composite predictive regime is to combine a number of forecasts developed through different predictive techniques in a manner that will enable superior forecasts to be made compared to those produced by the individual forecasts. The single algorithm is capable of producing superior forecasts compared to the individual forecasts.

Once trained and validated the ANN is capable of obtaining highly accurate results but the training process can take time and if the current motion of the ship is not similar to the training data, the accuracy of the ANN predictions could be poor. For that reason, it would be advantageous to use the ANN in combination with another predictive technique. This would ensure that if the ANN is performing poorly another predictive technique could be used which would offer greater accuracy. Therefore it would be theoretically feasible to obtain higher accuracy on a greater basis then if only one predictive technique was used alone.

Confidence that the composite predictive regime will result in superior performance is gained through observations of past investigations. In an investigation by Cox and Popken [12] a two level approach to short term forecasting was used. The way the two level approach operates is that information from the lower level forecast algorithms is combined with the upper level via a classification tree algorithm. The claim made by the authors is that the high-order interactions among error patterns from different predictive schemes are exploited by the tree classification approach. In essence, a hybrid forecast is developed and was capable of outperforming the predictions of the individual forecast systems.

Wedding and Cios [13] used another approach exploiting composite predictions. In their investigation, a composite predictive scheme was developed that combined Radial

Basis Function Neural Networks (RBFNN), certainty factors and the Univariate Box-Jenkins (UBJ) model. Three hybrid prediction combinations are presented in the paper.

The underlying reason that this combination worked effectively was that the RBFNN and the UBJ models both had weaknesses. However, when they were used in combination these weaknesses dissipate as each of the predictive schemes account for each other's weaknesses. The UBJ was not able to find subtle patterns in the time series data due to its simplistic nature. It is capable of obtaining high accuracy in many situations and hence it is a widely used method.

The RBFNN is capable of being trained to find the subtle changes in the time series data and is capable of discovering complex relationships in the data. The RBFNN work well when the predictions are based on data similar to the data on which they were trained. However, if the data is not similar to the training data very poor predictions are created as the RBFNN has difficulty extrapolating an answer.

When they are combined the weaknesses are somewhat nullified and the overall predictions are of a higher accuracy then the individual predictive schemes. The decision on whether the UBJ or the RBFNN is used will depend upon the certainty factors that indicate the similarity of the data to the data used to train the RBFNN.

There are several different methods that could be used to combine the predictions from the different schemes. Xiong, Shamseldin and O'Connor [14] used four different combination methods in their study into the prediction of rainfall-runoff. They found that the first order Takagi-Sugeno combination method [15] was superior to the simple average method, weighted average methods and the neural network methods for combining predictions because it was as easy to implement and efficient.

In a separate study conducted by Fiordaliso [16] the Takagi-Sugeno combination methods was used to combine individual forecasts. The Takagi-Sugeno was then compared with the traditional linear combining models and demonstrated that there are advantages in using the non-linear approach. The linear combination methods used for comparison were the weighting of forecasts [17], Bayesian methods [18, 19] the minimum-variance method [20] and regression based methods [21].

8 Conclusions

In this investigation a number of different time series prediction techniques based on ARIMA and ANN theories were applied to ship motion data. It was found that the ANN based method utilizing a combination of the GA and CG minimization algorithms for the prediction of the ship motion was superior to the univariate and multivariate ARIMA techniques. A combination approach that combines multiple predictive techniques was presented. A discussion was given which gave strong indications that a combination of different predictive techniques will yield higher overall results than the individual techniques alone.

9 Acknowledgements

The authors would like to thank BAE Systems for providing funding and support for the project.

10 References

[1] Price W G and Bishop R E D. *Probabilistic theory of ship dynamics.* Chapman and Hall Ltd, 1974.

[2] Sidar M M and Doolin B F. On the feasibility of real-time prediction of aircraft carrier motion at sea. *IEEE Transactions on Automatic Control*, 28, 3, pp 350-356, 1983.

[3] Stoica P, Mari J and McKelvey T. Vector ARMA estimation: an enhanced subspace approach. *38th IEEE Conference on Decision and Control*, 4, pp 3665-3670, 1999.

[4] Box G E P and Jenkins G M. *Time series analysis : forecasting and control.* Holden-Day, 1976.

[5] Brockwell P J and Davis R A. *Introduction to time series and forecasting.* Springer-Verlag, 1996.

[6] Suykens J A K, Vandewalle J P L and Moor B D. *Artificial neural networks for modeling and control of non-linear systems.* Kluwer Academic Publishers, 1996.

[7] Masters T. *Practical neural network recipes in C++.* Academic Press, 1993.

[8] Rumelhart D, McClelland J and Group P R. *Parallel distributed processing.* MIT Press, 1986.

[9] Press W H, Flannery B P, Vetterling W T and Teukolsky S A. *Numerical recipes in Fortran: The art of scientific computing.* Cambridge University Press, 1992.

[10] Polak E. *Computational methods in optimisation.* Academic Press, 1971.

[11] Obitko M. *Genetic algorithms*. Viewed 16 March 2004, http://cs.felk.cvut.cz/~xobitko/ga/, 1998.

[12] Cox L A and Popken D A. A hybrid system-identification method for forecasting telecommunications product demands. *International Journal of Forecasting*, 18, pp 647-671, 2002.

[13] Wedding D K and Cios K J. Time series forecasting by combining RBF networks, certainty factors, and the Box-Jenkins model. *Neurocomputing*, 10, pp 149-168, 1996.

[14] Xiong L, Shamseldin A Y and O'Connor K M. Non-linear combination of the forecasts of rainfall-runoff models by the first-order Takagi-Sugeno fuzzy system. *Journal of Hydrology*, 245, pp 196-217, 2001.

[15] Takagi T and Sugeno M. Fuzzy identification of systems and its application to modeling and control. *IEEE Trans. Syst. Man. Cybern*, 15, pp 116-132, 1985.

[16] Fiordaliso A. A nonlinear forecasts combination method based on Takagi-Sugeno fuzzy systems. *International Journal of Forecasting*, 14, pp 367-379, 1998.

[17] Doyle P and Fenwik I A. Sales forecasting using a combination of approaches. *Long-Range Planning*, 9, pp 60-69, 1976.

[18] Bunn D W. A Bayesian approach to the linear combination of forecasts. *Operational Research Quarterly*, 26, pp 325-329, 1975.

[19] Winkler R L. Combining probability distributions from dependent information sources. *Management Science*, 27, pp 479-488, 1981.

[20] Dickinson J P. Some comments on the combination of forecasts. *Operational Research Quarterly*, 26, pp 205-210, 1975.

[21] Granger C W J and Ramanathan R. Improved methods of combining forecasts. *Journal of Forecasting*, 3, pp 197-204, 1984.