Hybrid Artificial Intelligence Methods in Oceanographic Forecast Models

Juan M. Corchado and Jim Aiken

***Abstract*—An approach to hybrid artificial intelligence problem solving is presented in which the aim is to forecast, in real time, the physical parameter values of a complex and dynamic environment: the ocean. In situations in which the rules that determine a system are unknown or fuzzy the prediction of the parameter values that determine the characteristic behaviour of the system can be a problematic task. In such a situation it has been found that a hybrid artificial intelligence model can provide a more effective means of performing such predictions than either connectionist or symbolic techniques used separately. The hybrid forecasting system that has been developed consists of a case-based reasoning system integrated with a radial basis function artificial neural network. The results obtained from experiments in which the system operated in real time in the oceanographic environment, are presented.**

***Index Terms*—Case-based reasoning systems, Forecasting, Hybrid Artificial Intelligence systems.**

I. Introduction

Forecasting the behaviour of a dynamic system is, in general, a difficult task, especially if the prediction needs to be achieved in real time. In such a situation one strategy is to create an adaptive system which possesses the flexibility to behave in different ways depending on the state of the environment. This paper presents the application of a novel hybrid artificial intelligence (AI) model to a real time forecasting problem. The approach which is discussed is capable of producing satisfactory results in situations in which neither artificial neural network nor statistical models have been sufficiently successful.

The oceans of the world form a highly dynamic system for which it is difficult to create mathematical models [1], [2]. Although some statistical models have been formulated to describe partial oceanographic water masses, there are, as yet, no accurate general models. It is well known that the behaviour and characteristics of the oceans change seasonally and spatially. However, current knowledge of the ocean structure is still too weak to create a comprehensive model. Ocean water masses are extremely heterogeneous; each water mass has certain properties that differentiate it from other water masses. Moreover, the convergence area between different water masses can be very noisy; for example, both the Arctic and Antarctic convergence zones are extremely heterogeneous and very variable.

An artificial intelligence approach to the problem of forecasting in the ocean environment offers potential advantages over alternative approaches, because it is able to deal with uncertain, incomplete and even inconsistent data [2]. Several types of standard artificial neural network (ANN) have been used to forecast time series. In these experiments it has been discovered that it is very difficult to train a neural network to forecast successfully over the whole time series, especially if the data relate to a dynamic system. Statistical models such as Auto-Regressive Integrated Moving Averages (ARIMA) have been applied, but the results so far obtained have indicated that neural networks have a greater facility for forecasting using this type of time series data.

An attempt was also made to create an ANN able to train itself, in real time, with a number of consecutive values from a time series in order to forecast future parameter values [3], [4]. The aim was to build a system having the ability to create a local model, thus capturing the characteristics of each particular section of the time series. Although this network produced better results that any of the other models, it was considered not to be sufficiently accurate.

Following earlier experiments using case-based reasoning (CBR) as a problem solving strategy in other domains [5], a case-based approach to the forecasting problem was considered worthy of investigation. Consequently a prototype system based on this approach was developed in the belief that a CBR mechanism might, as a data mining strategy, make better use of the vast database of oceanographic data held at PML. The results of subsequent experiments employing a CBR approach indicated that case-based reasoning methods could facilitate the organisation of data, the recovery of relevant data necessary to make an accurate forecast and incremental system learning. The adaptation of the recovered data is a crucial factor in obtaining an accurate result. With the aim of providing a more effective case adaptation procedure it was decided to investigate a hybrid systems approach in which a neural network would be employed as a means of case adaptation. As a result, a Radial Basis Function (RBF) network has been found to be effective in the case adaptation stage of the CBR system.

An important aim in the current work is to develop a universal forecasting mechanism, in the sense that it might operate effectively anywhere, at any point on the surface of

J. M. Corchado is with Dpto. de Informática y Automática, Universidad de Salamanca, Facultad de Ciencias, Plaza de la Merced, s/n., 37008, Salamanca, SPAIN (e-mail: corchado@usal.es).

J. Aiken is with Plymouth Marine Laboratory, Prospect Place, West Hoe, Plymouth, PL1 - 3DH, UK (e-mail: JA@wpo.nerc.ac.uk).

the ocean, and at any time of the year without human intervention. The results obtained to date suggest that the approach to be described in this paper appears to fulfil this aim. The structure of the paper is as follows. First a brief overview of the basic concepts of case-based reasoning is given. Then the oceanographic problem domain is briefly outlined. Following this, a review of the application of CBR elsewhere as a strategy for forecasting is presented. The hybrid neural network enhanced CBR system is then explained, and, finally, an outline of some of the results obtained to date is presented.

II. Case-based Reasoning Systems Overview

The idea underlying CBR is that people frequently rely on previous problem-solving experiences when solving new problems. This assertion may be verified in many day to day problem solving situations by simple observation or by psychological experimentation [6], [7]. A case-based reasoning system solves new problems by adapting solutions that were used to solve previous problems [8]. The case base holds a number of cases, each of which represents a problem together with its corresponding solution. Once a new problem arises, a possible solution to it is obtained by retrieving similar cases from the case base and studying their recorded solutions.

A CBR system is dynamic in the sense that, in operation, cases representing new problems together with their solutions are added to the case base, redundant cases are eliminated and others are created by combining existing cases. A CBR system analyses a new problem situation, and by means of indexing algorithms, retrieves previously stored cases, together with their solution, by matching them against the new problem situation, then adapts them to provide a solution to the new problem by reusing knowledge stored in the form of cases in the case base. All of these actions are self contained and may be represented by a cyclic sequence of processes, in which human interaction may possibly be needed. Case-base reasoning can be used by itself or as part of another intelligent or conventional computing system. Furthermore, case-based reasoning can be a particularly appropriate problem solving strategy when the knowledge required to formulate a rule-based model of the domain is difficult to obtain, or when the number or complexity of rules relating to the problem domain is too great for conventional knowledge acquisition methods.

A typical CBR system is composed of four sequential steps which are called into action each time that a new problem is to be solved [7], [9], [10]. Figure 1 outlines the basic CBR cycle. This cyclic process of CBR involves four major steps, represented by the ellipses in figure 1: (i) Retrieve the most relevant case(s), (ii) Reuse the case(s) to attempt to solve the problem, (iii) Revise the proposed solution if necessary, and (iv) Retain the new solution as a part of a new case.

The purpose of the retrieval step is to search the case base and to select from it one or more previous cases that most closely match the new problem situation, together with their solutions. The selected cases are reused to generate a solution appropriate to the current problem situation. This solution is revised if necessary and finally the new case (i.e. the problem description together with the obtained solution) is stored in the case base. Cases may be deleted if they are found to produce inaccurate solutions, they may be merged together to create more generalised solutions, and they may be modified, over time, through the experience gained in producing improved solutions. If an attempt to solve a problem fails and it is possible to identify the reason for the failure, then this information should also be remembered in order to avoid the same mistake in the future. This corresponds to a common learning strategy employed in human problem solving. Rather than creating general relationships between problem descriptors and conclusions as is the case with rule-based reasoning, or relying on general knowledge of the problem domain, CBR systems are able to utilise the specific knowledge of previously experienced concrete problem situations. A CBR system provides an incremental learning process because each time that a problem is solved a new experience is retained, thus making it available for future reuse.

In the CBR cycle there is normally some human interaction. Whilst case retrieval and reuse may be automated, case revision and retention are often undertaken by human experts. This is a current weakness of CBR systems and one of their major challenges. In this paper a method of automating the process of case adaptation (revision) is presented for the solution of problems in which the cases are characterised predominantly by numerical information.

Several researchers [11], [12] have used k-nearest-neighbour algorithms for time series predictions. Although a k-nearest-neighbour algorithm does not, in itself, constitute a CBR system, it may be regarded as a very basic and limited form of CBR operation in numerical domains. In [11] a relatively complex hybrid CBR-ANN system is used. In contrast in [12] a data set is forecasted just by searching in a given sequence of data values for segments that closely match the pattern of the last n measurements and then by supposing that similar antecedent segments are likely to be followed by similar consequent segments. [5] presents a compilation of CBR based systems which have been used in forecasting problems during the past decade and which have been successfully implemented.

III. The Oceanographic Problem

A. *The Ocean Environment*

Oceanography is a science that concentrates on the understanding of the physical principles that drive the oceans, and which uses the tools of mathematics and theoretical fluid dynamics to forecast their behaviour [1]. Oceanic waters are divided into provinces (also called water masses) which are moderately homogenous. The boundaries between provinces or water masses are known as fronts. These areas are very dynamic and their properties depend on the nature of the two water masses converging to create each front. Some frontal areas, e.g. the Arctic and Antarctic convergence zones, are extremely heterogeneous and are very variable; therefore forecasting the temperature of the water in these areas can be difficult.

The movement of water masses makes the ocean's water temperature change in a complex manner in both spatial and temporal domains [1]. The analysis and interpretation of large volumes of oceanographic data have traditionally been achieved through the use of statistical software tools. However the processing speed when using such tools is limited by the need for frequent user intervention.

B. *The Application of Artificial Intelligence Methods*

Knowledge-based systems have been developed to assist in weather prediction. However, apart from the work in [2], there appears to be little evidence of the application of knowledge based approach as a means of predicting the location and movements of large water masses.

With the aim of providing an improved method for analysing the large masses of available oceanographic data, the application of knowledge-based (in particular, rule-based) methods was investigated in an earlier research project [13]. The motivation for that work was derived from the oceanographers' need for a better understanding of the ocean environment, for which, because of the complexities of the ocean, it is very difficult to formulate adequate and complete mathematical models. The results, using rule-based methods, provided some insight into the nature and complexities of the ocean. Building on this experience, it was decided to carry out further investigations into the application of alternative artificial intelligence problem solving approaches. Detecting oceanographic features, their boundaries and being able to predict their evolution was the goal of the Knowledge Based Oceanographic System (KBOS) [14] and subsequent projects: the On line Real time Knowledge based Analysis (ORKA) system [15], and the Simulated Tactical Environmental Bubble (STEB) system [4, 16].

The focus in recent research has been to investigate ways to forecast the thermal structure of the water ahead of an ongoing vessel, in real time. Forecasting in such an environment is a difficult task due to the nature and behaviour of the ocean waters, which are in a continuous state of movement. The scales of physical motion of the oceans and the atmosphere range from being ocean-wide down to tiny eddies, which are present in the neighbourhood of fronts.

In order to obtain acceptable predictions an autonomous universal methodology, capable of forecasting changes in the water temperature as an expert oceanographer might do, is desirable. In addition, the system should be able to analyse, in real time, the variation in the temperature of the water on a local basis, to analyse and select the most relevant local knowledge from the huge database of information available (in the form of satellite images and thermal data profiles) to produce a forecast, taking into account factors such as the season of the year and the geographical location of the vessel from which the forecast is made.

IV. Forecasting Systems

In the current work the aim is to investigate a hybrid case-based approach to forecasting as a methodology for predicting the values of physical parameters (in particular, sea temperature) at a given depth around a sea going vessel from data acquired in real time, and also from past records of sea temperature (and possibly other oceanographic parameters) surrounding the vessel at some point ahead which will be reached in the immediate future. This information may also then be used to provide a forewarning of an impending oceanographic front. The approach builds on the methods and expertise previously developed in the earlier research referred to above.

The problem of forecasting, which is currently being addressed, may be simply stated as follows:

Given: *a sequence of data values (which may be obtained either in real-time, or from stored records) relating to some physical parameters,*

Predict: *the value of that parameter at some future point(s) or time(s).*

The raw data (on sea temperature, salinity, density and other physical characteristics of the ocean) which are measured in real time by sensors located on the vessel consist of a number of discrete sampled values of a parameter in the form of a time series. These data values are supplemented by additional data derived from satellite images, which are received weekly. In the present work the parameter used is the temperature of the water mass at a fixed depth. Values are sampled along a single horizontal dimension, thus forming a set of data points.

This data must be pre-processed in order to eliminate noise, to enhance interesting features, to smooth stable areas and to transform the data set into a form which may be represented on an absolute scale. There are several techniques [2] that can be applied to transform the original data set to reduce noise, sharpen data and aid in the detection of fronts. The approach adopted employs a Sobel Filter, the operation of which is based on the idea that local variations, corresponding to edge transitions, occur at a slower rate than those corresponding to noise.

A. *Hybrid CBR - Neural Network System*

In order to produce a forecast, in real-time, of ocean temperature a certain distance ahead of a vessel as it traverses the ocean, a problem case is generated every 2 km. A problem case consists of a sequence of the N sampled data values (after suitable filtering and pre-processing) immediately preceding the data value corresponding to the current position of the vessel. A value of 40 for N has been found empirically to produce satisfactory results. The problem case also includes various other numerical values, including the current geographical location of the vessel and the time and date when the case was recorded.

The set of N data values forms an input vector, which is then used to produce a forecast of the ocean temperature, several km ahead. In outline, this process is depicted in figure 2. (Note that, in practice, it is the set of differences ($T_i - T_{i-1}$, $T_i - T_{i-2}$ etc.) between the temperature Ti at the current point and the temperature at successive earlier points which is used as the input vector.

The forecasted values are created using a neural network enhanced case-base reasoning system. The CBR mechanism allows the experience recorded in previous forecasting situations to be reused. The role of the neural network lies in the case adaptation process. The relationships between the processes and components of the hybrid system are illustrated in figure 3. The cyclic CBR process shown in the figure has been inspired by work in [9]. The four basic phases in the CBR cycle are shown as ellipses. Superimposed on the fundamental CBR cycle is a cycle of neural network operations during which the network parameters are retrieved from a neural network knowledge base, employed in case adaptation, and then are revised, with their updated values being stored back in the knowledge base. The full cycle of operations of the hybrid system is explained in the following section.

The particular type of neural network of interest in the current research is the Radial Basis Function [17], in which the input layer is a receptor for the input data, whilst the hidden layer performs a non-linear transformation from the input space to the hidden layer space. The hidden neurons form a basis for the input vectors; the output neurons merely calculate a linear combination of the hidden neurons' outputs. Activation is fed forward from the input layer to the hidden layer where a Basis Function is calculated. The weighted sum of the hidden neurons' activations is calculated at the single output neuron. Radial Basis Functions (RBF) are better at interpolating that at extrapolating. Furthermore, RBFs are less sensitive to the order in which data is presented to them than is the case with other neural network models, such as Multi-Layer Perceptrons. Radial Basis Functions are of potential use in hybrid systems because of their fast learning capability.

B. Hybrid System Operation

The forecasting system uses data from two sources: (i) the real-time data are used to create a succession of problem cases, characterising the current forecasting situation; (ii) data derived from satellite images are stored in a database (which, for clarity, is not shown in figure 3). The satellite image data values are used to generate cases, which are then stored in the case base and subsequently updated during the CBR operation.

The cycle of forecasting operations (which is repeated every 2 km) proceeds as follows. First a new problem case is created from the pre-processed real-time data. A set of k cases, which most closely matches this current problem case, is then obtained from the case base during the CBR retrieve phase, using nearest neighbour matching.

In the reuse phase, the values of the weights and centres of the neural network used in the previous forecast are retrieved from the neural network knowledge base. These network parameters together with the k closest matching cases are then used to create a forecast of the temperature a distance 5 km, say, ahead. At this point the parameters of the network are modified by taking into account the information contained in the retrieved cases. The effect of this is to allow the system to learn from all these k cases (rather than simply using the single adjudged closest matching case) in making a new forecast.

During each forecasting cycle the RBF network is retrained, using the retrieved weights and centres, with the input vectors contained in the k matching cases applied as inputs to the network. This process adapts the network, by accommodating the retrieved cases, thus updating the values of the network parameters (empirically, a value for k of between 500 to 1000 has been found to be appropriate). The input vector from the problem case is then fed into the trained network to produce a proposed forecast.

In the revise phase, the proposed forecast is modified by taking into account the accuracy of the previous forecasts, which were reused in obtaining the new forecast. Each case has associated with it a measure of the average error over the previous forecasts for which that particular case was used to train the neural network. Error limits are calculated by averaging the average error of the k cases used to train the ANN in producing the current forecast. The revised forecast is then expressed, using the error limits, as an interval between upper and lower limits rather than as a single value.

The revised forecast is then retained in a temporary store – the forecasts database. When the vessel has travelled a further 5 km the actual value of the water temperature at that point is measured. The forecasted value for the temperature at this point can then be evaluated, by comparison of the actual and forecasted values, and the error obtained. A new case, corresponding to this forecasting operation, is then entered in the case base. Knowledge of the forecasting error is also, at this point, used to update the average error of all the k cases that were reused to obtain that forecast.

C. Radial Basis Function Operation

The RBF network uses nine input neurons, between twenty and thirty five neurons in the hidden layer and a single neuron in the output layer. Input vectors (explained earlier) form the input to the network; the output of the network is the difference between the temperature at the present point and the temperature a fixed distance ahead. Initially, twenty vectors are randomly chosen from the first training data set and used as centres in the middle layer of the RBF network. All the centres are associated with a Gaussian function, the width of which, for all the functions, is set to the mean value of the Euclidean distance between the two centres that are separated the most from each other.

Training of the network is done by presenting pairs of corresponding input and desired output vectors. After an input vector activates every Gaussian unit the activations are propagated forward through the weighted connections to the output units which sum all incoming signals. The comparison of actual and desired output values enables the mean square error (the quantity to be minimised) to be calculated.

The closest centre to each particular input vector is moved toward the input vector by a percentage α of the present distance between them. By using this technique the centres are positioned close to the highest densities of the input vector data set. The aim of this adaptation is to force the centres to be as close as possible to as many vectors from the input space as possible. The value of α is initialised to a value of 20 each time that the network is retrained, and its value is linearly decreased with the number of iterations until its value becomes zero; then the network is trained for a number of iterations (between 10 and 30 iterations for the whole training data set, depending on the time left for the training) in order to obtain the best possible weights for the final value of the centres.

A new centre is inserted into the network when the average error in the training data set does not fall by more than 10% after 10 iterations (using the whole training set). In order to determine the most distant centre C, the Euclidean distance between each centre and each input vector is calculated and the centre whose distance from the input data vectors is largest is chosen. A new centre is inserted between C and the centre closest to it. Centres are also eliminated when they do not contribute significantly to the output of the neural network. Thus, a neuron is eliminated if the absolute value of the weight associated with that neuron is smaller than twenty per cent of the average value of the absolute value of the five smallest weights. The number of neurons in the middle layer is maintained above 20.

V. Results

The hybrid forecasting system has been tested in the Atlantic Ocean on the AMT (Atlantic Meridional Transect) research cruise going from the UK to the Falkland Islands [2]. The cruises crossed several water masses and oceanographic fronts (areas where two water masses with different characteristics converge). The system used in this experiment was set up to forecast the temperature of the water 5 km ahead of the ongoing vessel. Figure 4 illustrates the error in the forecasts over a total distance traversed of 10500 km. The strategy adopted was to create an accurate successful method which was able to forecast a short distance ahead, and then to extend it so as to produce forecasts a further distance ahead.

The average error in the forecast was found to be 0.02 ºC . Only 4.5% of the forecasts have an error higher than 0.5 ºC, 8.3% higher than 0.04 ºC, 32% higher than 0.02 ºC. These figures indicate that the hybrid system is able to produce a forecast with an average error of 0.02 ºC and with a probability of 0.96 that the error in the forecast is smaller than 0.05 ºC Although the experiment was carried out using a limited data set (11000 km between the latitudes 50º North and 50º South), eleven water masses with different characteristics were crossed, six fronts were traversed. The Falkland Front (km 10000) in particular is one of the most chaotic oceanographic areas in the world. It is believed that these error value results are significant enough to be extrapolated over the whole Atlantic Ocean.

Figure 4 shows the absolute value of the difference between the actual temperature value of the water and the forecast value obtained using the RBF neural network for the case adaptation in the hybrid system. This graph does not take into account any improvement that may be obtained using error limits during the review phase of the CBR cycle.

The use of error limits can substantially improve the accuracy of the forecast. The error limit values are determined and modified dynamically from information relating to the past forecasting performance of the system and which is contained in the stored cases. These error limits indicate the range of forecast values than may be expected to be produced through the adaptation of particular stored cases. Figure 5 shows the value of the error limits used during the experiment and figure 6 the forecast error outside the error limits. If, using error limits, the forecasted temperature value at a particular location is, say, 6 ± 0.5 ºC - i.e. from 5.5 ºC to 6.5 ºC - then, if the actual temperature value is found to be 6.8 ºC, the value of the forecast error outside the error limits will be 6.8 - 6.5, i.e. 0.3 ºC. Although 45.5% of the predictions were outside the limits of the error band, only 3.4% of the predictions were more than 0.005 ºC outside the error limits. The average error of the predictions using error limits is 0.00099 ºC. The shape of the error limit plot (figure 6) is very similar to the error in the forecast presented in figure 4; this means that the error limits adapt themselves to the pattern of temperatures in the different water masses. Both the error and the error limits are, on average, higher in frontal water masses than in homogeneous water masses. This was to be expected due to the higher dynamic nature and heterogeneity of such areas. A similar experiment was carried out using the data recorded by the vessel during the cruise, but this time using cases obtained from satellite images recorded more than one week previously. Table 1 shows how the average error, the average error limits and the average error outside the error limits were found to vary when satellite images of different ages were used.

Table 1 shows that the forecasting error is only slightly changed when using satellite images which are one or even two weeks old. The table also shows that when using satellite pictures collected exactly one year back the error in the forecast may be similar or smaller than the error obtained using pictures that are three or more weeks old. This is the reason why data up to one year old is kept in the database and in the case base; if for technical reasons (e.g. clouds covering a certain area or problems with the data telecommunications) recent satellite images can not be obtained, data recorded one year back can be used by the system and may in fact produce better results that data recorded three or four weeks previously. This is due to the annual cycle of most of the water masses. However, such results can not be guaranteed as there are also other factors that determine the pattern of ocean temperature variations.

Further experiments have been carried out to compare the performance of the CBR-ANN hybrid forecasting system with several other forecasting approaches. These include standard statistical forecasting algorithms and the application of several neural network methods. The results obtained from these experiments are listed in table 2. The table shows the average error obtained with a Finite Impulse Response ANN [4], a standard Radial Basis Function network [4], a Linear Regression model, an ARIMA (Auto-regressive Integrated Moving Averages) model and a CBR system without the neural network component which creates cases in real time using the temperature recorded from the sea surface [18]. Using the hybrid system the forecasting error is less than 20% of the corresponding value produced by any of the other forecasting methods.

VI. Conclusions

This paper has presented a hybrid problem solving method that combines a case-based reasoning system integrated with an artificial neural network, which is able to identify trends present in a large data set in order to create forecasts in real time. The forecasting task which is addressed in this work is difficult for two reasons: the unpredictable and complexity of the media in which the forecast must to be done, and the fact that the forecast must be done in real time. Essentially, the cases stored in the case base are created through the selection of data values which correspond to the current location and track of the ship. The time series data values obtained in real time are then matched by the CBR mechanism against the data patterns of the stored cases in order to produce the required forecast. The data are derived from the extensive database of earlier historical values, or from satellite images. The evaluation of the forecasts produced and the consequent modifications of the stored cases enable the system to learn and to improve its performance over time.

The hybrid forecasting system is able to produce a forecast with an acceptable degree of accuracy and within the time constraints imposed by the real time nature of the problem. Although the accuracy of the forecast depends, to a great extent, on the quality of the cases and on the actual date when the data from which the cases were created was collected, it has been demonstrated that good quality forecasts may be obtained even with data collected one year before the forecast was made.

The method combines the ability of case-based reasoning to index, organise and retrieve relevant data with the generalisation, learning and adaptation capabilities of the radial basis function neural network. The resulting hybrid system thus combines complementary properties of both connectionist and symbolic AI methods. The neural network plays an important role in the system; it adapts the cases selected during the case-based operations, combines aspects of the knowledge contained in several cases and supports the generation of the prediction. The Radial Basis Function network adapts its structure, in an unsupervised way, to the characteristics of the environment in which the system is operating and acts as a function that facilitates system learning by extracting the relevant characteristics of a number of closely matching cases and combining them in the form of a representative case.

The results obtained may be extrapolated to provide forecasts further ahead using the same technique, and it is believed than successful results may be obtained. However, the further ahead the forecast is made, the less accurate the forecast may be expected to be.

The limitations of this method of forecasting in its present form are as follows.

- Forecasts can only be produced while the vessel is proceeding in a straight line. The present system is not able to forecast while the vessel is changing its direction (or has changed it in the last 40 km). However, it would be possible to adapt the way in which the data is extracted from the database during the case creation process, to enable it to overcome this limitation.
- The present system operates satisfactorily only if there are no discontinuities in the data greater than 2 km in length. However, after a discontinuity the forecast can be resumed by interpolating the missing data with data from both ends of the case vector. This strategy has been found to be successful if the discontinuity is no longer than 5 km.
- The system can not function in a particular area if there are no stored cases from that area. In this situation, the only solution is to use a back-up mechanism to prime the system; for this, the experimental results obtained in comparing neural network methods suggest that a Finite Impulse Response network may be the most appropriate method to use. Once the system is in operation and is producing forecasts, a succession of cases will be generated, thus enabling the hybrid forecasting mechanism to function autonomously.

In conclusion, the hybrid problem solving approach provides an effective strategy for forecasting in an environment in which the raw data is derived from three distinct sources: a large database of historical data, satellite image data, and time series data obtained in real-time.

References

[1] M. Tomczak and J. S. Godfrey, *Regional Oceanography: An Introduction*. Pergamon, New York, 1994.

[2] J. M. Corchado, J. Aiken and N. Rees, *Artificial Intelligence Models for Oceanographic Forecasting*. Plymouth Marine Laboratory, Plymouth, 2001.

[3] J. M. Corchado, C. Fyfe and B. Lees, "Unsupervised Neural Method for Temperature Forecasting", in Proc. *EIS*, 1998, pp. 312-327.

[4] J. M. Corchado C. and Fyfe, "Unsupervised Neural Network for Temperature Forecasting" Artificial Intelligence in Engineering, vol. 13, no. 4, pp. 351-357, Oct. 1999.

[5] J. M. Corchado and B. Lees, "Adaptation of Cases for Case-based Forecasting with Neural Network Support" in Soft Computing in Case Based Reasoning, S.K Pal, T.S. Dillon and D.S. Yeung Ed. London: Springer Verlag, 2000, pp. 341-372.

[6] G. Klein and L. Whitaker, "Using Analogues to Predict and Plan" in Proc. WCBR, 1988, pp. 224-232.
[7] J. Kolodner, *Case-Based Reasoning*. Morgan Kaufmann. San Mateo, CA, 1993.
[8] C. K. Riesbeck and R. C. Schank, *Inside Case-Based Reasoning*. Lawrence Erlbaum Ass., Hillsdale, 1989.
[9] A. Aamodt and E. Plaza, "Case-Based Reasoning: foundational Issues, Methodological Variations, and System Approaches" AICOM, vol. 7, no. 1, pp. 39-59, Mar. 1994.
[10] I. Watson, *Applying Case-Based Reasoning: Techniques for Enterprise Systems*. Morgan Kaufmann. San Mateo, CA, 1997.
[11] G. Nakhaeizadeh, "Learning prediction of time series. A theoretical and empirical comparison of CBR with some other approaches", in Proc. EWCBR, 1993, pp. 65-76.
[12] G. G. Lendaris and A. M. Fraser, "Visual Fitting and Extrapolation" in Time Series Prediction, Forecasting the Future and Understanding the Past, A. S. Weigend and N. A. Gershenfield Ed., Santa Fe, Addison Wesley, 1994, pp. 341-368.
[13] B. Lees, N. Rees J. and Aiken, "Knowledge Based Oceanographic Data Analysis", in Procs. EXPERSYS, 1992, pp. 561-566.
[14] N. Rees, J. Aiken, and B. Lees, "Artificial Intelligence Techniques Applied to the Analysis of Oceanographic Data Sets" in Poc. OSATES, 1992, pp. 24-27.
[15] N. Rees and J. Aiken, "Detection of Oceanographic Fronts: A knowledge Base" in Proc. Expert Systems, 1995, pp 135-146.
[16] J. M. Corchado, B. Lees, C. Fyfe, N. Rees, J. Aiken, "Neuro-Adaptation Method for a Case Based Reasoning System", in Proc. IJCNN, 1998, pp 589-601.
[17] S. Haykin, *Neural Networks: A Comprehensive Foundation*. Prentice Hall, New York, 1994.
[18] N. Rees, B. Lees, and J. M. Corchado, "Oceanographic Modeling: using a CBR system" Plymouth Marine Laboratory, UK, Tech. Rep. OM-III, Dec. 1998.

Juan M. Corchado (Ph.D.) received a PhD. in Computer Science from the University of Salamanca in 1998 and a PhD. in Artificial Intelligence from the University of Paisley (UK) in 2000. At present he is *Profesor titular de Universidad* at the University of Salamanca (Spain), previously he was *Sub-director of the Escuela Superior de Ingeniería Informática* of the University of Vigo (Spain, 1998-00) and Researcher at the University of Paisley (UK, 1995-98). He has been a research collaborator with the Plymouth Marine Laboratory (UK) since 1993. He has leaded and worked on several Artificial Intelligence (AI) Research projects sponsored by Spanish and European public and private Institutions. He is the co-author of over 80 scientific publications, most of these present practical and theoretical achievements of Hybrid AI Systems.

Jim Aiken (Ph.D.) is a Group Leader at the Plymouth Marine Laboratory (UK) working in the field of “Upper Ocean Dynamics, Provinces and Biogases” incorporating Marine Optics, Remote Sensing, Atlantic Meridional Transect (AMT), Physical Oceanography and modelling teams. He is a visiting Professor of Optical Oceanography, at the University of Plymouth (1996-02). He has been a biological oceanographer for over 30 years and has published over 300 scientific papers and numerous technical reports, over half in peer reviewed journals. In the past few years, he has been a convenor of several international conferences and symposia on optical oceanography and remote sensing and an invited lecturer at 12 national or international meetings. He has been a consultant and contractor to ESA (MERIS Expert Systems Laboratory), EU JRC (Ispra), MOD, DERA, HO, MAFF, EA, BNSC, NASA, ONR and NASDA on biological oceanography, atmospheric correction and RS algorithms.