2014 International Joint Conference on Neural Networks (IJCNN)
July 6-11, 2014, Beijing, China

Estuarine flood modelling using Artificial Neural Networks

Seyyed Adel Alavi Fazel, Michael Blumenstein, Hamid Mirfenderesk and Rodger Tomlinson

Abstract—Prediction of water levels at estuaries poses a significant challenge for modelling of floods due to the influence of tidal effects. In this study, a two-stage forecasting system is proposed. In the first stage, the tidal portion of the available records is used to develop a tidal prediction system. The predictions of the first stage are used for flood modelling in the second. Experimental results suggest that the proposed flood modelling approach is advantageous for forecasting flood levels with more than 1 hour lead times.

I. INTRODUCTION

An estuary is referred to as a "semi-enclosed coastal body of water which has a free connection with the open sea and within which sea water is measurably diluted with fresh water derived from land drainage" [1]. Estuaries are commonly densely populated establishing the importance delivering real-time flood warnings in such areas. During a flood emergency, authorities rely on models that forecast water levels for a particular lead time. These models map observable variables such as rainfall to upcoming flood levels. Due to differences between catchments, models have parameters that are particular to the catchment. There are three modelling approaches available for this purpose [2]:

- Physically inspired models: these models are based on governing laws of physics. Parameters of these models deterministically evaluated by laboratory observations.
- Conceptual models: these models are based on simplified physics of the problem. The parameters of these models are determined by fitting parameters based on a 'small' subset of observations.
- Data-driven models: these models are based on relationships between the quantity of interest and available observations. Physical aspects of the phenomenon are used as constraints and also as the initial formulation of the problem. The parameters of these models are evaluated by calibration base on large number of observations.

This work is supported in part by an ARC industry Linkage grant, project ID LP110200832, The authors of the paper would like to thank the Australian Research Council and the Gold Coast City Council for funding this study. They also would like to thank the Gold Coast City Council for the provision of data that have been used in this study. Seyyed Adel Alavi Fazel is with the school of Information and Communication Technology, Griffith University, Gold Coast Campus, Australia (seyyedadel.alavifazel@griffithuni.edu.au). Michael Blumenstein is with the school of Information and Communication Technology, Griffith University, Gold Coast Campus, Australia (m.blumenstein@griffith.edu.au). Hamid Mirfenderesk is with the Gold Coast City Council, Gold Coast, Australia (hmirfenderesk@goldcoast.qld.gov.au). Rodger Tomlinson is with the school of environment, Griffith University Gold Coast Campus, Australia (r.tomlinson@griffith.edu.au)

This research is concerned with development of data-driven modelling for estuarine. To the best authors' knowledge, apart from study by Chang et al [3] no data-driven model has been suggested for flood modelling in estuarine and solutions developed for flood modelling are not readily applicable for estuarine flood modelling.

Developing a decision support system requires several steps that are demonstrated in Figure 1 ([4]).

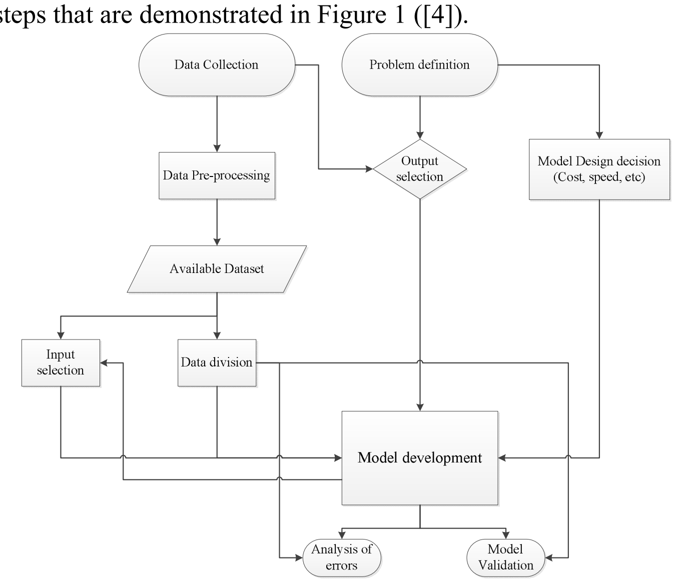

Figure. 1. Generic steps in the development of a data driven model for rainfall run-off simulation

In flood modelling context, data-driven models exploit patterns in available observation to predict flood levels at a particular forecasting horizon. In most common data-driven flood modelling solutions, available dataset consists of a set of observation of water levels or stream flow values sampled in equal time differences, forming a time series problem [5]. In the simplest form time series modelling involves a single random process. In univariate time series analysis, the future value of the time series is treated as dependent variable and observations up until the prediction point as regressors, which are formulated as:

$$x(t+h) = f(x(t), x(t-1), \ldots) \quad (1)$$

where 'x' is the time series process, and 't' denotes the observation index, 'h' defines the forecasting horizon and 'f' indicates the function that maps the two. According to conventional time series framework (e.g. [5]) dependency of future values of a process to past values can be categorized into short-term and long term 'memory'. For example, determination of expected monthly rainfall can be performed by pattern analysis of most recent observed rainfall which incorporates current atmospheric conditions as well as historical records of monthly rainfall for the specific season. The former is example of short memory and the latter is a form of long-term memory. It is noteworthy that direct inclusion of all relevant inputs in a long memory process as inputs would result in over-parameterization and poor generalization.

For many hydrological problems, commonly two types of long-term time dependencies are included in modelling namely seasonal and cyclic components. In seasonal time series (such as above-mentioned monthly rainfall example), data contains several seasons, each of which, exhibit distinct behavior. In previous research, deseasonalizatin of dataset [6] and development of distinct predictive models for each season [7] along other formulations have been suggested as possible solutions for seasonal time series modelling.

Another form of long-term process is existence of cyclic components, where one or more periodic components affect all observations. Astronomical tidal effect is example of such datasets. Abrahart *et al.* [8] suggest the inclusion of a phase of time with respect to significant periodic events as regressors for such datasets as a potential solution. This approach has been adopted by Chang *et al.* [3] for estuarine flood modelling. To avoid over-parameterization, Chang *et al.* limited number of dominant periodic to three dominant frequencies as input to the forecasting model. In a separate study, for sole tidal prediction (no flood modelling), Lee [9] suggest identification of dominant frequencies using conventional Fourier transform and using them as only regressors of model allowing incorporation of more periodic components for modelling.

For determination of flood levels, apart from short-term and long term memory of the process, researches commonly include the effect of rainfall (see Mins and Hall [10]). Such problem formulation includes more than one variable, hence forming a multivariate time series problem. In multivariate time series analysis, exogenous variables are included in time series formulation:

$$x(t+h) = f(x(t), x(t-1),.., y(t), y(t-1)...) \tag{2}$$

where 'y' is the exogenous variable. In flood modelling, one of the challenges in the development of a multivariate solution is the lack of existence of the exogenous variable in significant portions of the dataset. That is, in many flood modelling problems, available datasets contain proportionally few data points that contain flood related data. Thus developing a model that is trained on the entire dataset inevitably causes bias towards the non-flood portion (contrary to the purpose of flood modelling). To address this issue, modelers often limit calibration of parameters of model to training examples drawn from one or more flood events [11].

Estuarine flood modelling poses a unique challenge. As mentioned earlier, flood modelling should be 'event' based due to the lack of influence of rainfall in most of the dataset; meanwhile astronomical cyclical components influence the entire dataset requiring training examples from entire dataset.

The contribution of this research is exploration of two different aspects of flood modelling for estuarine namely developing exploring formulation that allow learning joint short-term, long term and cyclic components. Additionally, practical considerations in utilizing proposed modelling solutions are given.

II. ESTUARINE STUDIED AREA

The estuary dataset from Oyster's Creek, Gold Coast, Australia together with rainfall from 5 gauging stations upstream of the catchment are considered in this study. A summary of the dataset utilized for this study is presented in Table I. The recording intervals for each of the series are 15 minutes.

TABLE I
DATASETS USED FOR THIS STUDY

Station name	Recording Quantity	Recording Start Date	Recording Finish Data
Oysters Creek	Water Levels	1999 Dec 01, 06:00:00	2013 Apr 16, 08:30:00
Oysters Creek	Rainfall	1997 May 25, 12:00:00	2013 Apr 16, 08:30:00
Coplicks Bridge	Rainfall	2005 Aug 03, 07:00:00	2013 Apr 16, 08:30:00
Tallebudgera Creek Rd	Rainfall	2005 Aug 24, 13:00:00	2013 Apr 16, 07:45:00
Tallebudgera Creek Dam	Rainfall	2005 Jan 13, 07:00:00	2013 Apr 16, 07:30:00
Springbrook	Rainfall	1992 Jul 24, 20:00:00	2013 Apr 16, 07:30:00

III. METHODOLOGY

A. Overview of ANNs and Back propagation

Artificial Neural networks (ANNs) are among the most commonly utilized methods for function approximation in hydrological modelling [12]. The architecture of an ANN is shown in Figure 2. ANNs are layered information processing systems that allow identification of mapping between regressors and desired variables. One characteristic of ANNs is that they have little assumptions in regards to the structure of the problem and the inter-relationship between variables, providing a framework for both univariate and multivariate time series analysis. To solve a generalized problem of the form of equation (2) for the choice of "f", the modeler identifies key inputs that describe the output variable and creating 'training examples'. Each training example consists of a pair in the form of equation (2). The building blocks of an ANN are the information processing units referred to as neurons [13]. Each neuron receives information from inner network layers and performs a weighted summation over all information received in the inner layer. A function of this summation is then given to an outer layer (in a two layer network); this procedure continues until sensory information goes to the output layer. The output of neural networks is calculated by the following equation:

$$O = \varphi\left(\sum W_{32} * \phi(W_{21} * In)\right) \quad (3)$$

where φ is an activation function in the output layer, W_{32} is the weight matrix between the hidden and outer layer, ϕ is the activation function at the hidden layer, W_{21} is the weight matrix between the input and the hidden layer and finally In is the input sensory information.

By giving training examples to the network, synaptic weights of the network are adjusted in a way to best mimic the relationship between the input and output.

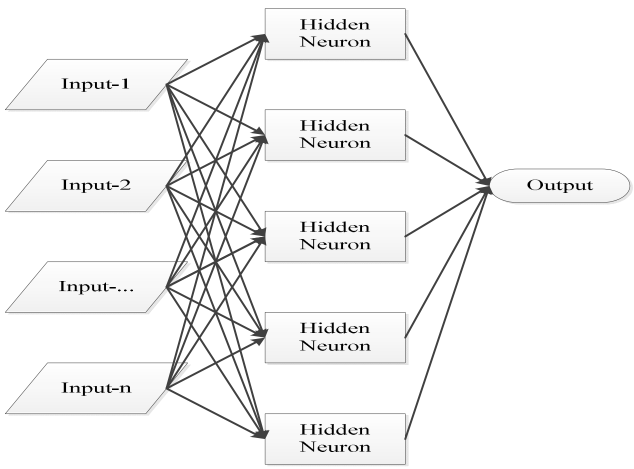

Figure. 2. Architecture of an ANN

In the back-propagation algorithm, the input and output variables are given to an arbitrarily initialized neuro-structure that maps input variables to output variables. Based on errors made by the model, the weights that connect input-output variables are adjusted in a way that penalizes a user-defined cost function. After reaching a user defined stopping criterion, the training concludes and the ANN can be deployed for prediction. For Development of an ANN, the number of layers, the transfer function between layers, the number of neurons in each layer, an objective function and the stopping criterion need to be chosen by the modeler. Despite extensive successful application of ANNs, no concrete rules for each of these steps can be found in the literature. Most commonly, no more than one hidden layer is employed in application of ANNs; one common transfer function used between the input and hidden layers is the tangent hyperbolic and between the hidden to output layer a linear function can be used; the objective function is the Root Mean Square error and the number of neurons is identified by trial and error. The most stable and commonly utilized optimization method for optimizing the weight space in "small" neural networks is the Levenberg-Marquardt (LM) method [14]. The LM method consists of the following steps:

1) Evaluation of the Jacobean matrix, defined by the derivative of the objective function with respect to each of the weights of the network

2) Approximation of the Hessian Matrix defined as

$$H \cong (J^T * J) \quad (4)$$

3) Adjustment of weights:

$$W \leftarrow W + [H + \sigma I]^{-1} * J \quad (5)$$

where σ is the adaptive quantity that controls the speed of convergence

4) Iterate until the stopping criterion is achieved

Definition of the stopping criterion is a design decision by the modeler. Ideally, training should be stopped to avoid over-fitting and also maximizing the utility from the available training examples. In this work, suggested by Moradkhani *et el.* the available training examples are randomized and divided into 3 portions, 2 are given for training. After each epochs of training, the performance of the predictive model is evaluated on the remaining portion. Subsequent to 10 successive failures in improvement of performance, training is stopped and the weights are assigned to the best identified iteration.

B. Approaches for rainfall run-off modelling

Armed with ANN for mapping random variables, the remaining task is identification of input random variables and initial formulation of the problem. As mentioned previously, rainfall-runoff modelling necessitates separation of flood event portions and non-flood portion, accordingly four formulations are considered:

1) The explicit tidal effect is ignored and it is assumed that the auto-regressive property of water levels as of equation (2) provides sufficient information about auto-regressive component of estuarine time series.

2) Suggested by Chang *et al*. [3], time with respect to significant harmonics are included as exogenous variables, which is of form of equation (2).

3) Assumes additive interaction between tidal levels and rainfall effects, the expected tidal levels are pre-subtracted from the water levels time series and the remaining series is formulated as equation (2).

4) Expected tidal values at the prediction time are given as exogenous variables.

Approaches and corresponding symbols for problem formulation of this study are provided in Table II.

TABLE II
SUMMARY OF FORMULATIONS

Symbol	Approach
A1	Ignoring the effect of time dependent astronomical tide
A2	Direct inclusion of time (with respect to dominant frequency components) as an exogenous variable
A3	Identifying the cyclic components – pre-subtractions of expected tidal levels from the water level time series
A4	Identifying the cyclic components – inclusion of expected tidal levels as exogenous variables

C. Input Selection

In this paper, 3 types of inputs are considered namely short-term auto-regressive component (i.e. dependency of future values of water levels on current water levels and last few hours), exogenous variables (i.e. effect of incoming rainfall to future water levels) and long term memory (tidal effects). In this section, overview of each of the components for formulation of flood levels at estuarine is provided.

a) Identification of dominant frequencies

Fourier analysis provides a practical tool for identification of significant harmonic components for a periodic time series analysis. Large magnitude points in Fourier domain are corresponded to dominant cyclic components of the time series data. The Fourier transform of a signal X(t) can be determined by the following equation:

$$X(f) = \sum_{i=1}^{N} X(t) e^{\frac{-2\pi jft}{N}} \quad (6)$$

where "f" is the index for the cyclic component, X(f) is a complex number referring to the strength of the component "f" spanning from 0 to $\frac{N}{2}$, "N" is the number of samples and "j" is the unit imaginary point. The Fourier transform of the water levels at Oysters Creek is shown in Figure 3. Time with respect to the dominant periodic (sorted by magnitude) are selected as input variables for approaches A2-A4.

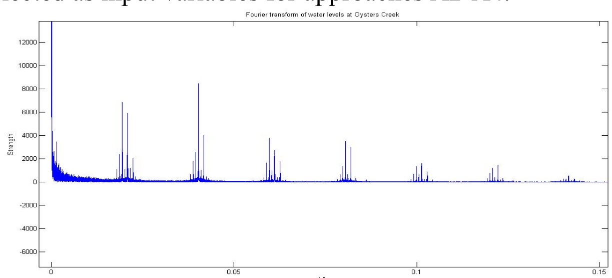

Figure. 3. Fourier transform of the water levels at Oysters Creek

By a simple evaluation of mangnitue of values in Fourier domain, dominant frequencies were identified. The list of 20 strongest frequencies is given at Table III.

TABLE III

LIST OF 20 DOMINANT FREQUENCIES IN ESTUARINE (1/HOUR)

0.0805	0.0833	0.0790	0.0418	0.0387
0.1667	0.1595	0.0433	0.0448	0.0777
0.0417	0.0403	0.1610	0.2415	0.1223
0.0818	0.0832	0.1250	0.1192	0.1638

b) Rainfall exogenous variables

Cross-Correlation between water level time series and water levels provides a simple practical tool for identification of the lag between rainfall and subsequent changes in water levels. The cross-correlation between available averaged rain gauge and water levels at Oysters Creek is demonstrated at Fig 4.

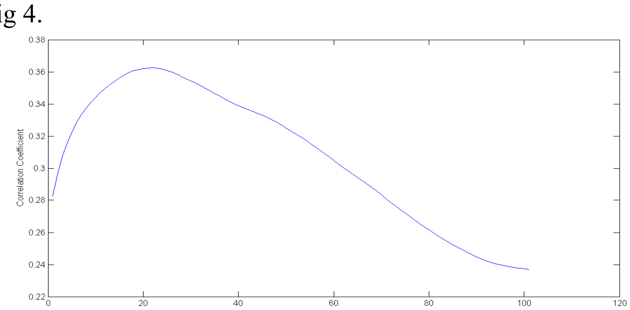

Figure. 4. Cross-Correlation between spatially averaged rainfall over entire catchment time series and water level.

As can be seen in Figure 4, subsequent to nearly 22 lags, the cross-correlation function peaks hence antecedent rainfall from 30 lag prior to the event are considered as inputs.

c) Antecedent water levels

Commonly examination of auto-correlation function (ACF) is used for evaluation of short term memory of a time series problem [15]. However ACF is meaningful for evaluation of stationary time series hence is not applicable to estuarine. Therefore, an empirical approach, by performance analysis of several different predictive models with different auto-regressive components is adopted herein.

IV. EXPERIMENTAL SETUP

The aim of this study is not the development of an optimal model or through investigation of all possible input space; rather it is the investigation of the difference between approaches. Knowing that BP is an iterative learning process and is sensitive to initial conditions, 20 models with different initial conditions are trained to reinforce that the observed difference between each model's performances is not due to effect of initial conditions. The forecasting horizon for all of the experiments is fixed at two and four hours. Notably, initial observations suggest that the variation of water levels is significant in such span of time, making such predictions useful for emergency warning systems.

One of the design decisions in the development of a flood forecasting model is the choice of the data-split for testing and calibration. This is particularly relevant as stakeholders wish to know the response of the system to a potentially record-breaking event. In this work, for flood forecasting components, the available flooding events are ranked based on the maximum water level reached during the event. Starting from the highest ranking, every third ranking event is preserved for testing and the remaining portion is utilized for calibration. Matlab 2012 and the associated neural network toolbox are used as a platform for designing the experiments.

V. RESULTS

A. Tidal Forecasting system

Realization of the approaches A3-A4 requires a dedicated tidal forecasting system. Such a flood forecasting system receives time with respect to the dominant periods. Networks with different numbers of inputs and neurons were tested; the network with the lowest Root Mean Square Error is selected as the tidal forecasting model for the remaining part of this paper. The results of different tidal forecasting models are summarized in Table IV.

TABLE IV

QUANTITATIVE RESULTS FOR THE TIDAL PREDICTION MODEL

Number of Frequency Components included as input variable	Number of neurons	Root Mean Square Error
17	5	0.0787
23	5	0.0784
17	10	0.0770
23	10	0.0717
17	15	0.0708

23	**15**	**0.0684**

To further illustrate the dynamic performance of the tidal prediction model, predictions of the model are demonstrated in Figure 5.

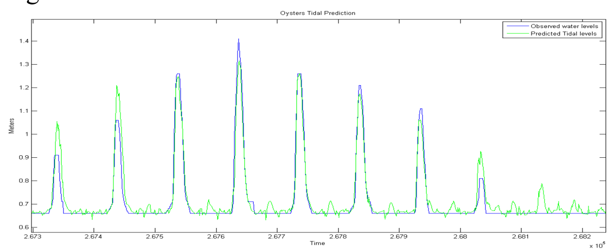

Figure. 5. Example of forecasted tidal levels

B. Flood forecasting model

Upon completion of an empirical process for the development of a reasonable representation of the input space, we noticed that the inclusion of rainfall more than 6 hours prior to the event does not enhance the results. Also, due to autoregressive nature of rainfall, available rainfall time series is averaged over different periods. It was found that averaging over a period of one hour does not negatively affect performance of predictive models. Hence, in all experiments, rainfall no more than 6 hours prior to the prediction point is included and rainfall for each gauging station is averaged over one hour interval.

To visually assess the difference between different modelling solutions across different approaches, a boxplot of each experiment for two hour forecasting horizon is illustrated in Figure 6 and Figure 7. A boxplot uncovers the effect of initialization and provides a visual representation of the experimental outcomes according to one performance metric. In a boxplot, x-axis refers to the experiment index (A1-A4) and y-axis represents experimented Root mean square error of each experimental outcome.

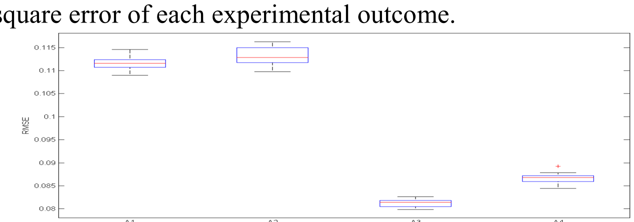

Figure. 6. Boxplot of different trials for cross-validation performance of different modelling approaches, two hours forecasting horizon

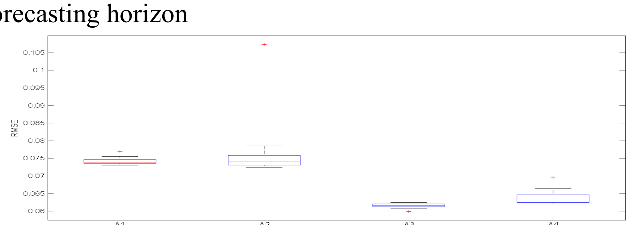

Figure. 7. Boxplot of different trials for independent test set performance of different modelling approaches, two hours forecasting horizon

As can be seem, according to both metrics, forecasts made by A3 and A4 approach significantly outperform A1-A2 approaches and more surprisingly, no significant difference between A1 and A2 is observable. The difference between approaches is evident both using cross-validation and during independent testing.

To reinforce these clams, statistical two tail u-test [16] of different experimental outcomes is performed using ranksum Matlab command. The results are summarized in Table V.

TABLE V
RESULTS OF STATISTICAL U-TEST FOR 20 OUTCOMES OF RMSE OVER CROSS VALIDATION IN TWO HOURS PREDICTION

Experiment Index	A1	A2	A3	A4
A1	-	0.0071	6.7956e-08	6.7956e-08
A2	-	-	6.7956e-08	6.7956e-08
A3	-	-	-	6.7956e-08
A4	-	-	-	-

TABLE VI
RESULTS OF STATISTICAL U-TEST FOR 20 OUTCOMES OF RMSE OVER INDEPENDENT TEST-SET IN TWO HOURS PREDICTION

Experiment Index	A1	A2	A3	A4
A1	-	0.6949	6.7956e-08	6.7956e-08
A2	-	-	6.7956e-08	6.7956e-08
A3	-	-	-	1.5757e-06
A4	-	-	-	-

As can be seen, the null hypothesis that these experimental setups have come from identical distributions can be rejected with ($p < 0.01$) between performances of all predictive systems apart from A1 and A2, when experiment is conducted on independent test sets. In all other cases significant difference between the outcomes of experiments can be observed.

Similar experiments were conducted for forecasting 4 hours into the future and subsequent boxplots are given in Figures 8 and 9.

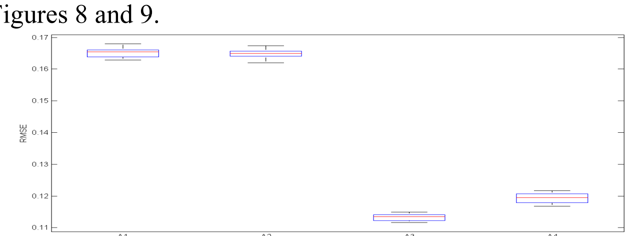

Figure. 8. Boxplot of different trials for cross-validation performance of different modelling approaches, four hours forecasting horizon

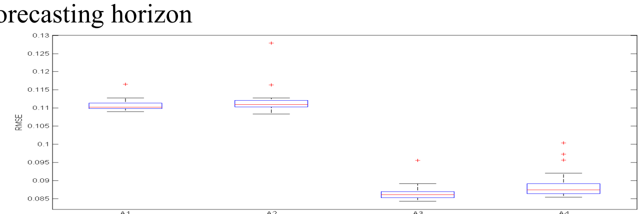

Figure. 9. Boxplot of different trials for cross-validation performance of different modelling approaches, four hours forecasting horizon

The results appear similar and to ensure scalability of results, similar statistical tests are conducted for Four hour forecasting horizon.

TABLE VII
RESULTS OF STATISTICAL U-TEST FOR 20 OUTCOMES OF RMSE OVER CROSS VALIDATION IN FOUR HOURS PREDICTION

Experiment Index	A1	A2	A3	A4
A1	-	0.6949	6.7956e-08	6.7956e-08

A2	-	-	6.7956e-08	6.7956e-08
A3	-	-	-	6.7956e-08
A4	-	-	-	-

TABLE VIII

RESULTS OF STATISTICAL U-TEST FOR 20 OUTCOMES OF RMSE OVER INDEPENDENT TEST-SET IN FOUR HOURS PREDICTION

Experiment Index	A1	A2	A3	A4
A1	-	0.1404	6.7956e-08	6.7956e-08
A2	-	-	6.7956e-08	6.7956e-08
A3	-	-	-	0.0031
A4	-	-	-	-

As can be seen, the apart from difference between A1-A2, the remaining approaches have statistically significant difference with ($p < 0.01$). This suggests the proposed A3 and A4 outperform existing approaches for flood modelling in estuarine.

To further investigate the application of the proposed models as decision support systems, forecasts of the best performing solution for historical peaks for two hours forecasting horizon is shown in Figure 10.

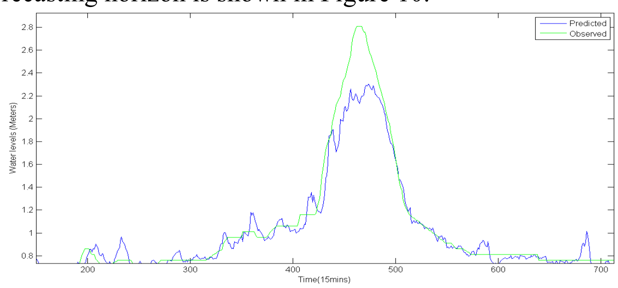

Figure. 10. Forecasts of the A3 model at historical peak at Oysters Creek, two hour forecasting horizon

It appears that the proposed modelling solution, although significantly outperforming other solutions investigated, still is not entirely satisfactory. Accordingly, further experiments are conducted to trace sources of error in the proposed solution. In this experiment, unseen rainfall from prediction point to forecasting horizon is assumed as a known quantity to arbitrary spatial and temporal resolution. In the first experiment, the unseen rainfall from the prediction time to the prediction point is given as input to the model (referred to as “a”); in the second one, rainfall from different gauging stations are given as distinct inputs (referred to as “b”). To simplify making conclusions, only the A3 approach for modelling is considered. The boxplots of the resulting experiments are shown in Figures 11 and 12.

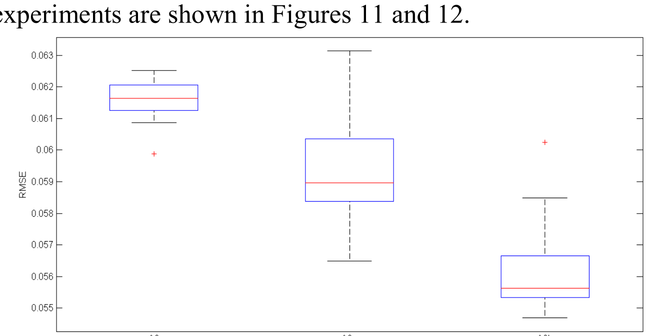

Figure. 11. Boxplots of experiments conducted for assessment of the effect of unseen rainfall on water levels.

As can be seen, the effect of unseen rainfall is significant; similarly providing rainfall from different stations as separate inputs. For comparison, the performance of this model for prediction of historical peaks is demonstrated in Figure 12.

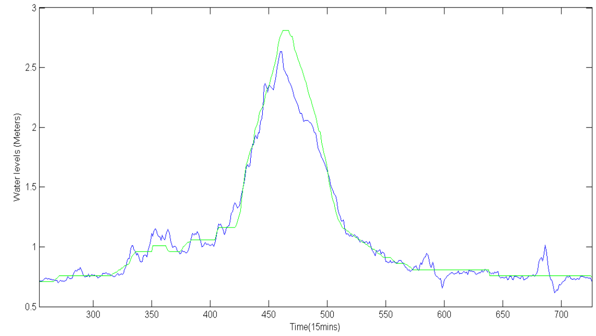

Figure. 12. Forecasts of the A3b model at the historical peak at Oyster’s Creek, two hour forecasting horizon

As can be seen, this model provides significantly better forecasts and the peak are underestimated only by 30cm and timing is correct. Statistical significant u-test also confirms that the difference between these settings is significant ($p < 0.01$).

VI. SUMMARY AND DISCUSSION

Four strategies for estuarine rainfall-runoff flood modelling using ANNs were compared, abbreviated as A1-A4 (see Table II). A1 and A2 approaches are so far only approaches suggested for estuarine flood modelling. In this work, A3 and A4 are realized using axillary tidal prediction system trained on non-flood portion of dataset. Commonly such portion is omitted from modelling altogether.

The results of this study suggest that proposed A3 and A4 significantly outperform existing data-driven estuarine flood analysis methods. Most basic A1 and suggested A2 were found to be insignificant for 2 and 4 hours forecasts. The main shortcoming of A1 and A2 models is attributed to training set bring solely on flood related portion of dataset, hence, the model is unable to capture the relationship between time with respect significant harmonics and future water levels. Despite the substantial advantage of A3, the model produced significant errors when testing the observed historical peak of the available observation for two hour forecasts. The second part of the experiments attempted to trace the source of error in proposed A3 system; it was found that the accurate prediction flood level requires accurate temporal and spatially distributed rainfall forecasts that are not available at time that forecasting is performed. Potential users may decide to consider either attempting shorter forecasting horizons or attempt alternative methods for direct prediction of unseen rainfall.

VII. CONCLUSION

In this work, the issue of the existence of cyclic components in an estuarine dataset has been directly tackled. Despite apparent significant gain in the proposed modelling solutions, the authors note that the results should be approached in context. In this case study, sources of errors and uncertainty were partially determined so that operators can fully appreciate the strength and weaknesses of the proposed model. It is essential for the modelers to understand effect of different variables and integrate them to the model and allocate sufficient training examples so that model learns the pattern. From end-users point of view, users should be aware of existence and the extent of the influence of hidden variables (in this case, unseen rainfall between prediction time and forecasting horizon and the spatial distribution of rainfall), structural limitation of the model, and the impact of parameter uncertainty, as well as potential errors in the recording of data. In this work, although not all of the abovementioned analysis of a decision support model was investigated, a method for integration of tidal effect on flood level was proposed as a first step towards a dedicated data-driven flood decision support for estuarine.

REFERENCES

[1] D. W. Pritchard, "What is an estuary: physical viewpoint," *Estuaries,* vol. 83, pp. 3-5, 1967.

[2] K. J. Beven, *Rainfall-Runoff Modelling: The Primer*, 2nd ed. Noida, India: Wiley-Blackwell, 2012.

[3] F.-J. Chang and Y.-C. Chen, "Estuary water-stage forecasting by using radial basis function neural network," *Journal of Hydrology,* vol. 270, pp. 158-166, 2003.

[4] J. C. Refsgaard, H. J. Henriksen, W. G. Harrar, H. Scholten, and A. Kassahun, "Quality assurance in model based water management–review of existing practice and outline of new approaches," *Environmental modelling & software,* vol. 20, pp. 1201-1215, 2005.

[5] B. Schelter, M. Winterhalder, and J. Timmer, *Handbook of Time Series Analysis*. Weinheim: Wiley-Vch, 2006.

[6] M. Nelson, T. Hill, W. Remus, and M. O'Connor, "Time series forecasting using neural networks: Should the data be deseasonalized first?," *Journal of forecasting,* vol. 18, pp. 359-367, 1999.

[7] W. Wang, P. H. V. Gelder, J. Vrijling, and J. Ma, "Forecasting daily streamflow using hybrid ANN models," *Journal of Hydrology,* vol. 324, pp. 383-399, 2006.

[8] R. J. Abrahart and L. See, "Comparing neural network and autoregressive moving average techniques for the provision of continuous river flow forecasts in two contrasting catchments," *Hydrological Processes,* vol. 14, pp. 2157-2172, 2000.

[9] T.-L. Lee, "Back-propagation neural network for long-term tidal predictions," *Ocean Engineering,* vol. 31, pp. 225-238, 2004.

[10] A. W. Minns; and M. J. Hall, "Artificial neural networks as rainfallrunoff models," *Hydrological Sciences Journal,* vol. 43, pp. 399-417, 1996.

[11] P.-A. Chen, L.-C. Chang, and F.-J. Chang, "Reinforced Recurrent Neural Networks for Multi-Step-Ahead Flood Forecasts," *Journal of Hydrology,* 2013.

[12] H. R. Maier, A. Jain, G. C. Dandy, and K. P. Sudheer, "Methods used for the development of neural networks for the prediction of water resource variables in river systems: Current status and future directions," *Environmental modelling & software,* vol. 25, pp. 891-909, 2010.

[13] S. Haykin, "Multilayer Perceoptrons," in *Neural networks and learning machines*, 3rd ed: Prentice Hall, 2009, pp. 123-217.

[14] A. P. Piotrowski and J. J. Napiorkowski, "Optimizing neural networks for river flow forecasting – Evolutionary Computation methods versus the Levenberg–Marquardt approach," *Journal of Hydrology,* vol. 407, pp. 12-27, 2011.

[15] K. P. Sudheer, A. K. Gosain, and K. S. Ramasastri, "A data-driven algorithm for constructing artificial neural network rainfall-runoff models," *Hydrological Processes,* vol. 16, pp. 1325-1330, 2002.

[16] H. B. Mann and D. R. Whitney, "On a test of whether one of two random variables is stochastically larger than the other," *The annals of mathematical statistics,* vol. 18, pp. 50-60, 1947.