Desalination and Water Treatment
www.deswater.com

32 (2011) 284–290
August

doi: 10.5004/dwt.2011.2712

Measuring treatment effectiveness of urban wetland using hybrid water quality — Artificial neural network (ANN) model

Gurmeet Singh[a], Jaya Kandasamy[a]*, H.K. Shon[a], J. Cho[b]

[a]*Faculty of Engineering and Information Technology, University of Technology, Sydney, New South Wales, Australia 2000*
Tel. +61 (2) 9514 2558; Fax +61 (2) 9514 7920 ; email: jaya.kandasamy@uts.edu.au
[b]*NOM Ecology National Research Laboratory, Department of Environmental Science and Engineering,*
Gwangju Institute of Science and Technology (GIST), 1 Oryong-dong, Buk-gu, Gwangju 500-712, Korea

Received 14 July 2010; Accepted in revised form 5 January 2011

ABSTRACT

Constructed wetlands are now commonly used as tertiary treatment for urban stormwater. The wetlands have primary advantage over other forms of treatment as they remove dissolved organics and heavy metals in conjunction with other pollutants. The effectiveness of a wetland is a primary concern for validating its compliance with design objectives and regulatory requirements. The treatment in a wetland is however complex and is dependent on input pollutants, hydraulics, physico-chemical balance and biota within the wetland. Several models are available for wetlands but have limitations in simulating the physico-chemical and biological processes within the wetland. The aim of this paper is to introduce a hybrid modelling approach that involves both a deterministic model and artificial neural network (ANN) for testing the effectiveness of a constructed wetland at Olympic Park, Homebush, Sydney, Australia. This novel approach allows a combination of calibrated water quality and neural based models to predict the water quality from the wetland. The models were calibrated and validated using water quality monitoring data measured for eight months in both influent and effluent streams of the wetland. The calibrated hybrid models were then tested for treatment effectiveness for range of wet, dry and median flows conditions within the catchments. A water quality index was developed and used to quantify the effectiveness of the wetland.

Keywords: Water quality modelling; Artificial neural networks; Wetlands; Water quality index

1. Introduction

Measuring stormwater treatment effectiveness is a major challenge for regulators. The effectiveness of a wetland is a primary concern for validating its compliance with design objectives and regulatory requirements. The complex physico-chemical and biological processes in a wetland renders it difficult to use mathematical models as these processes cannot be readily captured. Artificial neural network (ANN) models are becoming popular in water quality modelling as they are able to capture the dynamic response without the need to understand and model the process. Although there are some published papers on using ANN for modelling wetlands [1–3], there is not much work done on using hybrid models for measuring treatment effectiveness. A case study was undertaken to explore using hybrid models EPA SWMM and ANN to measure treatment effectiveness of urban wetland in Sydney Olympic Park.

* Corresponding author.

Presented at the Third International Conference on Challenges in Environmental Science & Engineering, CESE-2010
26 September – 1 October 2010, The Sebel, Cairns, Queensland, Australia

2. Study area

Sydney Olympic Park is a 640 ha urban park in the geographic centre of Sydney, Australia. The park is maintained as a lasting legacy to the Sydney 2000 Olympics. Out of the 640 ha the park has a built environment in a 'town centre' occupying about 215 ha and the remaining 425 ha are 'parklands'. The parklands comprise 175-ha of wetlands, 40 ha of woodlands and 210 ha of picnic areas and pathways. All of these are situated on the Parramatta River and in the vicinity of Homebush Bay.

Sydney Olympic Park is located within the greater Sydney Harbour Catchment on the Parramatta River and is affected by both upstream and downstream (intertidal) activities. Within the Park, however, there are four subcatchments; the Brickpit system; Haslam's Creek system; Powell's Creek system; and Nature Reserve Wetland (Wanngal Wetland) system. The wetland investigated in this study is Northern Water Feature (NWF) (Stormwater Pond-MP12), which lies on Haslam's Creek (Fig. 1).

The NWF wetland is about 2.60 ha and incorporates a series of wetland fingers with vegetation lining along the ponds edges to visually represent the Olympic Flame (Fig. 1).

3. Water quality monitoring

To test the effectiveness of treatment system, on site monitoring was conducted between October 2009 to May 2010 at the inlet and outlet from the wetland (Fig. 1). A total of 32 composite samples were used during the 8 months monitoring period covering both wet and dry weather flows regime. A number of common pollutants, heavy metals and organic matter in the treated water were then analysed (Table 1). The pollutants included those used for evaluating the water quality index (WQI) using the NSF (National Sanitation Foundation, USA) [4]. The NSF WQI was developed to provide a standardized method for comparing the water quality of various bodies of water. NSF Water Quality Index is a 100 point scale that summarizes results from nine different measurements as follows, temperature change, pH, dissolved oxygen (DO), turbidity, faecal coliforms, biochemical oxygen demand (BOD), total phosphorus, nitrates and suspended solids. A total of 32 composite samples during the monitoring period were then used for this study.

The monitoring results show significant biological activity within the wetland for removal of common pollutants and bacteriological parameters. Most of the pollutants are contributed from particulate loading whilst there is also significant dissolved matter washed off from catchments contributing to the wetland. Total dissolved solids are higher than typical urban catchments, Australian Rainfall Quality [5].

4. Methodology for evaluating effectiveness of wetland

The major challenge to the success of stormwater treatment is to quantify the effectiveness over its life cycle given that it demands high maintenance. This requires comprehensive water quality measurements both upstream and downstream of the treatment device for a range of inflow conditions.

The methodology presented in this paper is based on using water quality sampling data (8 months) based on representative storm events and calibrating water quality models EPA SWMM [6] over the range of inflow conditions observed during the sampling period.

An artificial neural network (ANN) model was then developed, trained and validated for the range of influent and recorded effluent data of the wetland. ANN is a powerful tool for multivariate and nonlinear analysis, and offers an alternative to traditional statistical methods [7]. The calibrated ANN model was used to test the effectiveness of the wetland over a range of wet, dry and median inflow conditions. A NSF Water Quality Index was developed and tested to gauge effectiveness of the wetland. The methodology involves the following steps:

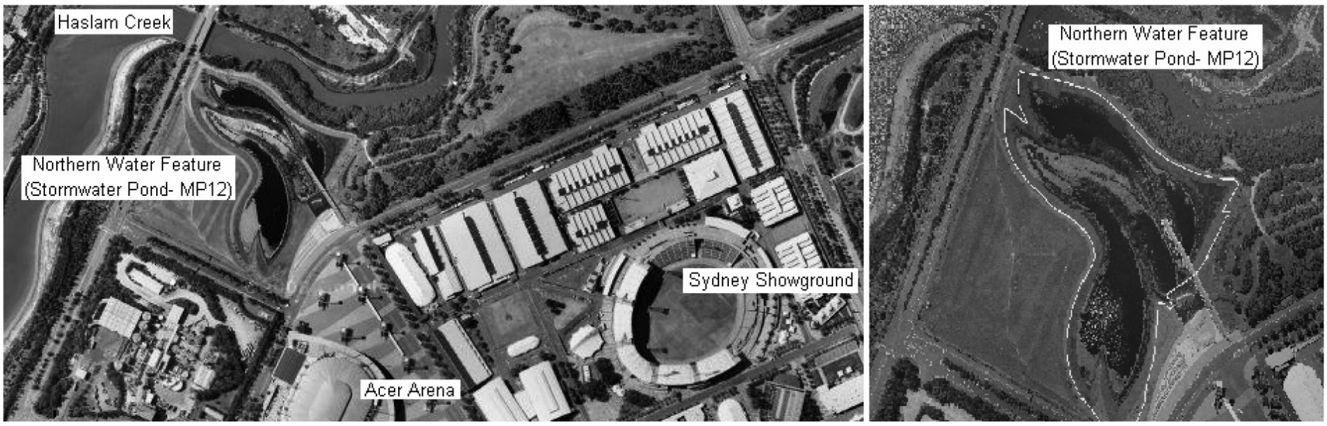


Fig. 1. Northern Water Feature (Stormwater Pond-MP12) and Haslam Creek.

Table 1
Sampling results (October 2009–May 2010)

	Minimum	Maximum	Mean	Std. deviation
Wetland level, m AHD	102.10	102.40	102.25	0.115
Influent parameters				
Temperature, °C	25.00	36.00	32.30	3.86
pH	6.90	9.76	8.52	0.97
DO, mg/L	5.00	9.40	7.05	1.62
BOD, mg/L	11.00	45.00	22.00	9.40
TDS, mg/L	276.8	754.1	534	166.1
Turbidity, NTU	1.89	6.00	3.65	1.82
Orthophosphate, mg/L	0.06	1.38	0.38	0.50
Nitrate, mg/L	0.01	2.08	0.47	0.81
Nitrite, mg/L	0.00	0.52	0.10	0.21
Temperature, °C	0.06	0.40	0.23	0.14
Total coliforms	9.00	6000	1123	2400
Faecal coliforms	9.00	4800.	858.	1933
Effluent parameters				
Temperature, °C	25.70	36.00	31.00	3.90
pH	6.81	9.52	8.69	1.08
DO, mg/L	3.96	8.50	4.60	1.70
BOD, mg/L	6.00	25.0	9.50	11.50
TDS, mg/L	225.8	661.0	429.7	157.7
Turbidity, NTU	1.41	10.00	5.26	2.95
Orthophosphate, mg/L	0.01	0.19	0.10	0.08
Nitrate, mg/L	0.00	1.82	0.31	0.74
Nitrite, mg/L	0.00	0.56	0.10	0.23
Ammonia	0.03	0.53	0.15	0.19
Total coliforms	9.00	350.0	86.2	136.4
Faecal coliforms	9.00	50.0	16.0	16.7

Stage 1. This stage involved calibrating the water quality variables to the observed field data. This was achieved by calibrating the EPA SWMM model by adjusting the hydrological parameters and pollutant generation rates. Flow calibration was undertaken by matching the range of recorded water levels within the wetland.

Stage 2. This stage involved training the neural network models using time series of input and output parameters from the water quality monitoring data for both influent and effluent streams in the wetland. The trained model was then tested and validated using the modelled input and output parameters and the best model was then adopted.

Stage 3. This involved running the calibrated EPA SWMM water quality model to generate pollutants for range of inflow conditions, wet, dry and median year.

Stage 4. This involved running the trained and tested neural model for the modelled inflow pollutants generated by EPA SWMM over the dry, median and wet weather conditions. This was called the production dataset and the results were then used to test the treatment effectiveness of the stormwater treatment train. The NSF Water Quality Index was also developed and reported for monthly variability in the index for range of dry, median and wet weather conditions in the catchment.

4.1. EPA SWMM model

The United Sates Environmental Protection Agencies (EPA) storm water management model (SWMM) [6] is a comprehensive computer model for the analysis of stormwater quantity and quality. SWMM can model continuous events and simulate all important aspects of the hydrological cycle. Rainfall and monthly evaporation values were obtained from the Bureau of Meteorology records for stations in the vicinity. Stormwater pollution is modelled through a buildup and washoff method, which includes dry-weather pollutant buildup over different land uses and pollutant washoff from specific land uses during storm events [6].

SWMM provides the user with a series of water quality inputs to model the accumulation of pollutants over the catchment and the removal of these pollutants. The event mean concentration (EMC) method was used to model the pollutant washoff. Table 2 gives typical ranges of common stormwater pollutants that were used in this study [5].

The following assumptions were made in the generation of pollutants that did not mask the overall objective of this study, which was to outline a viable modelling methodology to test the effectiveness of the wetland system. Temperature of the influent stream for modelled catchment conditions were estimated based on air temperature data recorded in the area and a correlation between observed air temperature and recorded water temperature (correlation coefficient 0.98). Dissolved oxygen was correlated to water temperature as a simplified approach. A high correlation coefficient of 0.92 was observed using the recorded data and the approach was deemed reasonable.

The EPA SWMM model was initially calibrated for hydrology by obtaining a good fit with the observed water levels within the wetland. The selected pollutants as specified in Table 3 were then calibrated in EPA SWMM using common export rates as specified in Australian Runoff Quality [5]. The model was adjusted iteratively for pollutant export rates until a good match was achieved with the recorded pollutants. The results of the EPA SWMM model are presented in Table 3. The objective of this method of calibration was not to attempt event calibration but to statistically match the distribution of event concentration over the sampled events. This method was deemed suitable for hybrid water quality modelling attempted in this study. Flow calibration was achieved by modelling wetland and the overflow weir structure to match the recorded water levels.

It was observed from Table 3 that the EPA SWMM model fits reasonably well to both the water level and selected pollutants observed in the influent stream to the wetland. The calibrated model was then used to develop the hybrid model in conjunction with an artificial neural network model.

Table 2
Range of pollutant concentration in urban stormwater [5]

Pollutants	Range
pH	4.7–8.5
Total suspended solids, mg/L	50–110
Total phosphorus, mg/L	0.15–0.25
Total nitrogen, mg/L	0.5–3
Turbidity, NTU	1.5–5
NO_2/NO_3-N, mg/L	0.86–2.2
BOD, mg/L	6–38
Total coliforms, cfu/100mL	5,000–80,000
Faecal coliforms, cfu/100mL	4–50

4.2. *Artificial neural network model*

An ANN is an adaptable system that can learn relationships through repeated presentation of data and is capable of generalizing to new, previously unseen data. Some networks are supervised, in that a human determines what the network should learn from the data. In this case, a network is provided a set of inputs and corresponding desired outputs, and the network

Table 3
EPA SWMM model results

	Minimum	Mean	Maximum		Minimum	Mean	Maximum
Water level, m AHD				pH			
Recorded	102.1	102.25	102.4	Recorded	6.9	8.52	9.76
Modelled	101.9	102.42	102.8	Modelled	7.2	8.15	9.2
Dissolved oxygen, mg/L				Turbidity, NTU			
Recorded	6.3	7.0	9.5	Recorded	1.89	3.65	6
Modelled	5.9	7.2	9.2	Modelled	2.3	3.82	8
Faecal coliform, Cfu/100mL				Total phosphorus, mg/L			
Recorded	9	858	4800	Recorded	0.06	0.38	1.38
Modelled	4	800	4200	Modelled	0.05	0.4	1.39
Nitrates, mg/L				Total dissolved solids, mg/L			
Recorded	0.01	0.47	2.08	Recorded	277	534	754
Modelled	0.01	0.48	2.2	Modelled	260	540	765
BOD, mg/L				Temperature, °C			
Recorded	11	22	45	Recorded	25	32	36
Modelled	10	24	42	Modelled	21	30	34

learns the input–output relationship by adapting its free parameters. Other networks are unsupervised, in that the way they organize information is hard-coded into their architecture. ANN is a powerful data modelling tool that is able to capture and represent complex input/output relationships. The use of ANN is growing rapidly with successful applications in many areas and has been also applied in prediction of water quality of wetlands [1–3].

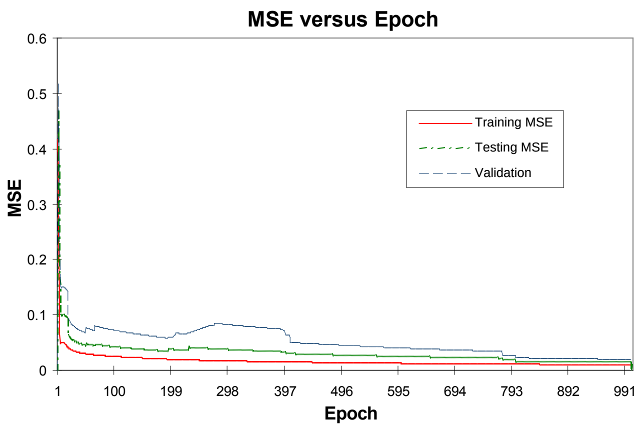
ANN model was developed using neural network toolbox in MATLAB [7]. A single hidden-layer, multi layered perceptron (MLP) feed forward neural network using error back propagation network, EBP was developed. All features of a feed forward neural model were investigated including training set creation, learning rate, number and layers of neurons, neural activation model predictions with test data sets. Most of the model configurations offered excellent predictive capabilities. Using either the logistic or the hyperbolic tangent neural activation function did not significantly affect predicted results. This was also true for the two learning algorithms tested the Levenberg-Marquardt and Polak-Ribiere conjugate-gradient descent methods. Hyperbolic tangent neural activation function was adopted for this study as this produced reasonable results. Cross validation was applied to the dataset with criteria set to terminate at either 100 epochs without improvement in the network or adopted mean square error (MSE).

The most important step during model development and training was the representative selection of data records for training of the model. A total of 16 samples were selected for training, 8 samples for validation and the remaining 8 samples for testing. A total of nine influent and effluent water quality variables (based on NSF Water Quality Index) along with recorded water levels in the wetland were used for developing the ANN model.

Network performance was estimated by linear regression between the actual and target (predicted) water quality parameters after post-processing the output to the original scalar variables. The results of training, testing and validation are presented in Table 4. The table shows that for trained network a mean square error of 0.05 was achieved in just 40 epochs with convergence around 200 epochs. The results show a good convergence in 1000 epochs for all the datasets and this trained model was then used for the production runs.

The tested ANN model was then run on input time-series (daily time step) from calibrated EPA SWMM models (nine water quality variables and water levels) for

Table 4
Results from ANN model runs



Best network	Training
Epoch #	1000
Minimum MSE	0.00984
Final MSE	0.00984

Performance	BOD	Dissolved oxygen	Temperature	pH	TDS	Turbidity
MSE	31.2259	0.0869	1.7819	0.1279	0.0201	4.7848
Min abs error	3.2925	0.0403	0.0989	0.0283	0.0176	0.2682
Max abs error	8.4287	0.5919	2.6732	0.6398	0.3011	5.0672
r^2	0.9174	0.9937	0.9317	0.9323	0.8455	0.9349

Performance	Ortho-phosphate	Nitrate	Faecal coliforms
MSE	0.0008	0.0041	2.1940
Min abs error	0.0103	0.0192	0.0620
Max abs error	0.0430	0.0948	3.0386
r^2	0.9272	0.9962	0.9975

MSE — mean square error

entire dry, median and wet year. The modelled effluent parameters from ANN were checked for wet, dry and median years to ensure that they were not statistically different from the recorded values distribution. The average monthly output from ANN model was then used for setting up NSF Water Quality index.

5. Results and discussions

The results of hybrid model for predicting long term pollutant removal for wet, dry and median flow conditions is presented in Table 5. Also, the overall performance of the wetland is presented in Table 6.

Results indicate the following:

1. In a typical wet year the performance of wetland is significantly lower than the average and dry conditions. The wetland is relatively small in size and therefore the detention time is significantly lower in a wet event which results in lower performance rates.
2. In a dry year, removal of total phosphorus and nitrogen is significantly higher. This is due to a higher intake of nutrients from aquatic growth. BOD and faecal coliform removal is significantly higher due to longer detention time in this flow regime.
3. In median year, the removal rates of pollutants are slightly better than the wet year, but the performance is still poor, and it appears that the wetland has a design issue most likely related to its size. It is also noted that the wetland does not have a dedicated sedimentation zone.

To gauge the relative performance of the wetland, a water quality index was developed based on the NSF. This index is a 100 point scale that summarises results from nine measurements, ie temperature change, pH, dissolved oxygen, turbidity, faecal coliform, bio-chemical oxygen demand, total phosphorus, nitrates and suspended solids.

Estimated NSF has been presented for dry, wet and median year flow conditions (Fig. 2). Based on the NSF Water Quality index, the range of score for this wetland is 30–70, which ranges from bad (25–50) and medium (50–70). There is high variability in the water quality index with seasonal variation. In a dry year, the performance is typically good for summer period (December–February period in Southern Hemisphere) that is due to macrophyte growth while deterioration in water quality is expected during pond overturning events (autumn and spring).

The hybrid modelling approach using both water quality (EPA-SWMM) and artificial neural network (ANN) models has demonstrated this approach as a powerful tool for evaluating the effectiveness of wetland. The methodology adopted compliments the limitations of both the models and provides a decision support tool that can be used with confidence. Based on the results of the hybrid modelling it can be concluded that the effectiveness of the wetland using the NSF Water Quality Index ranges from medium to bad and could be related to the size of the wetland. The results are preliminary in nature and further evaluation using more extensive monitoring data is required to properly assess the effectiveness of the

Table 5
Hybrid model results of selected pollutants for range of catchment conditions

	pH	DO	Turbidity	TSS	TP	Nitrate	BOD	FC	Temp
Dry year, rainfall 630 mm/y									
Influent, EPASWMM	8	6	6	18	1.2	1.1	53	1500	34
Effluent, ANN	9.4	5.1	2	3	0.04	0.1	8	480	30
Wet year, rainfall 1600 mm/y									
Influent, EPASWMM	6.5	9	8	32	0.23	0.68	15	220	28
Effluent, ANN	9.4	6.5	10.5	31	0.2	0.52	12	210	27
Median year, rainfall 1100 mm/y									
Influent, EPASWMM	7.5	7.1	6.8	23	0.89	0.9	24	410	30
Effluent, ANN	8	5.4	3.1	12	0.31	0.52	13	210	25

Table 6
Performance of wetland (in terms of percentage removal of pollutants)

	Turbidity	TDS	TP	Nitrate	BOD	FC
Dry year	67%	23%	97%	91%	85%	68%
Wet year	25%	17%	13%	24%	20%	5%
Median year	54%	22%	65%	42%	46%	49%

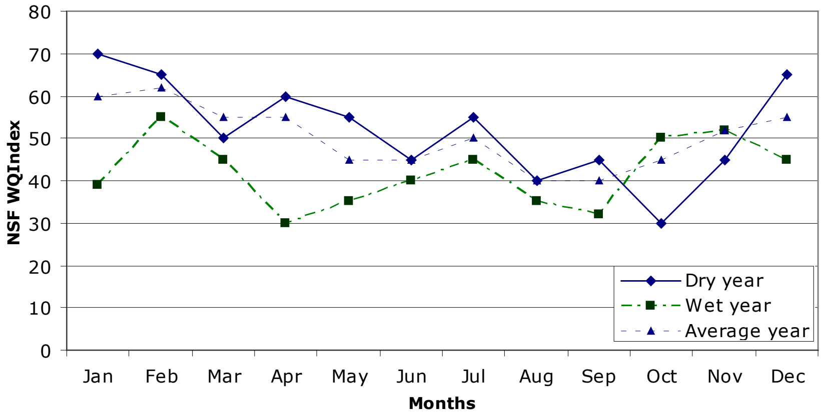


Fig. 2. NSF water quality index for wetland performance.

wetland using the hybrid modelling approach presented in this study.

Acknowledgements

This research was supported by a grant from the National Research Laboratory Program by the Korea Science and Engineering Foundation (NOM Lab: ROA-2007-000-20055-0).

References

[1] M. Minghao, X. Wang, H. Wu, S. Cai, X. Zhang and H. Wang, Ecosystem health assessment of Honghu Lake wetland of China using artificial neural network approach, Chin. Geogra. Sci., 19(4) (2009) 349–356.

[2] Y.Z. Shi and D.J. Xin, Applying artificial neural network model in assessing East Dongting Lake Wetland's ecosystem carrying capacity, ICICTA '08 Proc. 2008 International Conference on Intelligent Computation Technology and Automation, Vol. 1.

[3] V. Tomenko, S. Ahmed and V. Popov, Modelling constructed wetland treatment system performance, Ecol. Modelling, 205 (2007) 355–364.

[4] M.A. Brown, A water quality index: Do we dare?, Water Sewage Works, 117(10) (1970) 339–343.

[5] ARQ, Australian Runoff Quality, Institution of Engineers, Australia, 2005.

[6] EPA SWMM, Stormwater Management Model, Version 5, US EPA, Cincinnati, Ohio.

[7] J.W. Hines, MATLAB supplement to Fuzzy and Neural Approaches in Engineering, John Wiley & Sons Inc., 1997.