Lecture Notes in Computer Science, Vol. 3498, 2005, pp. 1034-1039

A Split-Step PSO Algorithm in Prediction of Water Quality Pollution

Kwokwing Chau

Department of Civil and Structural Engineering, Hong Kong Polytechnic University, Hunghom, Kowloon, Hong Kong, People's Republic of China
cekwchau@polyu.edu.hk

Abstract. In order to allow the key stakeholders to have more float time to take appropriate precautionary and preventive measures, an accurate prediction of water quality pollution is very significant. Since a variety of existing water quality models involve exogenous input and different assumptions, artificial neural networks have the potential to be a cost-effective solution. This paper presents the application of a split-step particle swarm optimization (PSO) model for training perceptrons to forecast real-time algal bloom dynamics in Tolo Harbour of Hong Kong. The advantages of global search capability of PSO algorithm in the first step and local fast convergence of Levenberg-Marquardt algorithm in the second step are combined together. The results demonstrate that, when compared with the benchmark backward propagation algorithm and the usual PSO algorithm, it attains a higher accuracy in a much shorter time.

1 Introduction

Over the past two decades, frequent algal blooms with occasional massive fish kills have been recorded in Tolo Harbour. It may largely be attributed to its intrinsic semi-enclosed nature and the extremely low tidal flushing rate. Moreover, most of the freshwater runoff in the catchment area is routed to reservoirs so that the river discharges to the harbour are much reduced. The condition is further deteriorated by the nutrient enrichment through municipal and livestock waste discharges in the harbour through the rapid economic development recently. Precise prediction of algal booms is beneficial to fisheries and environmental management since it allows the fish farmers to have more float time to take appropriate precautionary measures. However, the extremely complex dynamics of algal blooms are related to various pertinent physical and biochemical factors and are not well-comprehended.

Process-based mathematical models, such as finite element or finite difference methods, are conventionally used to forecast flow and water quality parameters in a water body. In general, they require exogenous input and embrace different assumptions. In numerical modeling, the physical problem is represented by a highly coupled, non-linear, partial differential equation set. The involving processes are highly complex and uncertain which may consume enormous computing cost and time. In this sense, mechanistic models are not totally satisfactory in representing the highly complex inter-relationships. Recently, soft computing (SC) techniques have been gradually becoming a trend to complement or replace the process-based models. The characteristics of these data-driven approaches include built-in dynamism, data-error tolerance, no need to have exogenous input and so on. Amongst others, artificial neural networks (ANN), in particular the feed forward back-propagation (BP) perceptrons, have been widely applied in water resources engineering [1]. However, slow training convergence speed and easy entrapment in a local minimum are inherent drawbacks of the commonly used BP algorithm [2]. Levenberg-Marquardt (LM) optimization technique [3] is a commonly used ANN that has attained certain improvements such as convergence rates over the BP algorithm. Swarm intelligence is another recent SC technique that is developing quickly [4]. This technique has been applied in hydrological problems and accomplished satisfactory results [5-6].

In this paper, a split-step PSO algorithm is employed to train multi-layer perceptrons for algal bloom prediction in Tolo Harbour of Hong Kong with different lead times and input variables. It is believed that, by combining the two algorithms, the advantages of global search capability of PSO algorithm in the first step and local fast convergence of LM algorithm in the second step can be fully utilized to furnish promising results.

2 Characteristics of PSO Algorithm

When PSO algorithm is initially proposed, it is considered a tool for modeling social behavior and for optimization of difficult numerical solutions [4,7]. This computational intelligence technique is intimately related to evolutionary algorithms and is an optimization paradigm that mimics the ability of human societies to process knowledge [8]. Its principle is based on the assumption that potential solutions will be flown through hyperspace with acceleration towards more optimum solutions. PSO is a populated search method for optimization of continuous nonlinear functions resembling the biological movement in a fish school or bird flock. Each particle adjusts its flying according to the flying experiences of both itself and its companions. During the process, the coordinates in hyperspace associated with its previous best fitness solution and the overall best value attained so far by other particles within the group are kept track and recorded in the memory.

Among other advantages, the more significant one is its relatively simple coding and hence low computational cost. One of the similarities between PSO and a genetic algorithm is the fitness concept and the random population initialization. However, the evolution of generations of a population of these individuals in such a system is by cooperation and competition among the individuals themselves. The population is responding to the quality factors of the previous best individual values and the previous best group values. The allocation of responses between the individual and group values ensures a diversity of response. The principle of stability is adhered to since the population changes its state if and only if the best group value changes. It is adaptive corresponding to the change of the best group value. The capability of stochastic PSO algorithm, in determining the global optimum with high probability and fast convergence rate, has been demonstrated in other cases [7-8]. PSO can be readily adopted to train the multi-layer perceptrons as an optimization technique.

3 Training of Perceptrons by PSO

Without loss of generality, a three-layered preceptron is considered in the following. $W^{[1]}$ and $W^{[2]}$ represent the connection weight matrix between the input layer and the hidden layer, and that between the hidden layer and the output layer, respectively. During training of the preceptron, the i-th particle is denoted by $W_i = \{W^{[1]}, W^{[2]}\}$ whilst the velocity of particle i is denoted by V_i. The position representing the previous best fitness value of any particle is denoted by P_i whilst the best matrix among all the particles in the population is recorded as P_b. Let m and n represent the index of matrix row and column, respectively, the following equation represents the computation of the new velocity of the particle based on its previous velocity and the distances of its current position from the best experiences both in its own and as a group.

$$V_i^{[j]}(m,n) = V_i^{[j]}(m,n) + r\alpha[P_i^{[j]}(m,n) - W_i^{[j]}(m,n)] + s\beta[P_b^{[j]}(m,n) - W_i^{[j]}(m,n)] \tag{1}$$

where j = 1, 2; m = 1, …, M_j; n= 1, …, N_j; M_j and N_j are the row and column sizes of the matrices W, P, and V; r and s are positive constants; α and β are random numbers in the range from 0 to 1. In the context of social behavior, the cognition part $r\alpha[P_i^{[j]}(m,n) - W_i^{[j]}(m,n)]$ denotes the private thinking of the particle itself whilst the social part $s\beta[P_b^{[j]}(m,n) - W_i^{[j]}(m,n)]$ represents the collaboration among the particles as a group. The new position is then determined based on the new velocity as follows:

$$W_i^{[j]} = W_i^{[j]} + V_i^{[j]} \tag{2}$$

The fitness of the i-th particle is determined in term of an output mean squared error of the neural networks as follows:

$$f(W_i) = \frac{1}{S}\sum_{k=1}^{S}\left[\sum_{l=1}^{O}\{t_{kl} - p_{kl}(W_i)\}^2\right] \tag{3}$$

where f is the fitness value, t_{kl} is the target output; p_{kl} is the predicted output based on W_i; S is the number of training set samples; and, O is the number of output neurons.

4 Split-Step PSO Algorithm

The combination of two different SC techniques could enhance the performance through fully utilization of the strengths of each technique. In this algorithm, the training process is divided into two stages. Initially the perceptron is trained with the PSO algorithm for a predetermined generation number to exploit the global search ability for near-optimal weight matrix. Then, after this stage, the perceptron is trained with the LM algorithm to fine tune the fast local search. This might be able to avoid the drawback of either entrapment in local minima in LM algorithm or longer time consumption in global search of PSO algorithm.

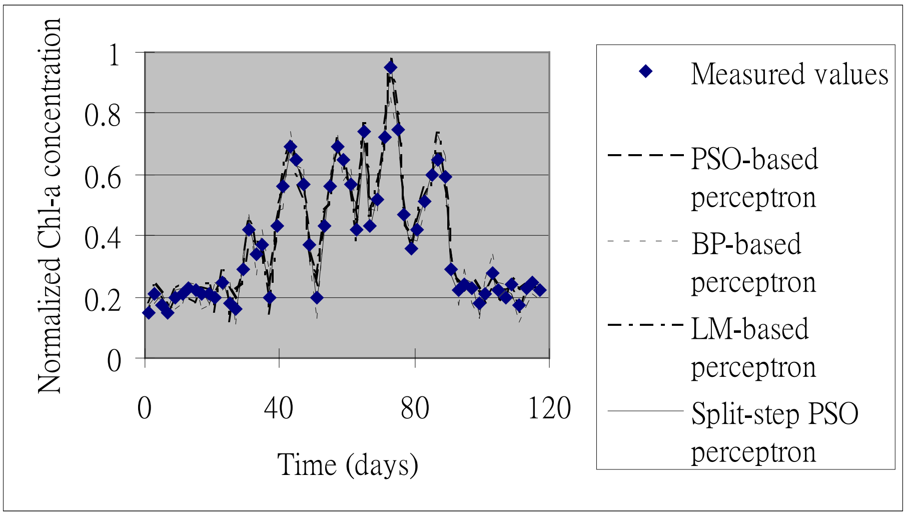

Fig. 1. 1 week lead time chlorophyll-a prediction for scenario 2 in the validation process

5 Algal Bloom Prediction in Tolo Harbour

In order to test the capability of the model to mimic a particular case study with accurate depiction of real phenomena, it has been employed to predict the algal bloom dynamics in Tolo Harbour of Hong Kong [9-10]. Observation data indicate the life cycle of algal blooms to be in the order of 1 to 2 weeks. As such, algal biomass, represented as chlorophyll-a, is forecasted with a lead time of 1 and 2 weeks based on a complete set of biweekly water quality data at Tolo Harbour from year 1982 to year 2002. The data of 1982-1995 and those of 1996-2002 are used for training and testing/validation, respectively. The division of data is tailored so as to include extreme frequency and intensity in both sets of data. Throughout the analysis, depth-averaged values from the surface, mean, and bottom of the water column are employed.

In this case, ten input variables, including the time-lagged chlorophyll-a, secchi disc depth, nitrogen, phosphorus, dissolved oxygen, rainfall, water temperature, solar radiation, wind speed and tidal range, are considered to be significant to the algal dynamics of Tolo Harbour [9]. Various perceptrons, having an input layer with one to ten neurons, a hidden layer with three to five neurons, and an output layer with one neuron, are tested. The single output node represents chlorophyll-a. Three scenarios are attempted with 10, 5 and 1 input variables for scenario 1, 2, and 3, respectively.

Other major PSO parameters adopted are as follows: number of population is 30; the maximum and minimum velocity values are 0.3 and -0.3, respectively. All source data are normalized into the range between 0 and 1, by using the maximum and minimum values of the variable over the whole data sets.

Table 1. Results for chlorophyll-a forecasting based on scenarios 1 to 3

Input data	Algorithm	Coefficient of correlation			
		Training		Validation	
		1 week ahead	2 weeks ahead	1 week ahead	2 weeks ahead
Scenario 1	BP-based	0.984	0.953	0.968	0.942
	PSO-based	0.989	0.980	0.973	0.962
	LM	0.986	0.958	0.970	0.958
	Split-step	0.991	0.983	0.975	0.969
Scenario 2	BP-based	0.974	0.934	0.956	0.934
	PSO-based	0.984	0.976	0.959	0.954
	LM	0.977	0.942	0.957	0.945
	Split-step	0.988	0.981	0.968	0.963
Scenario 3	BP-based	0.964	0.935	0.946	0.923
	PSO-based	0.983	0.973	0.965	0.952
	LM	0.969	0.938	0.951	0.928
	Split-step	0.985	0.978	0.969	0.958

Table 2. Steady-state fitness evaluation times during training for various algorithms

Algorithm	Steady-state fitness valuation time
BP-based	21,000
PSO-based	9,000
LM	5,000
Split-step	6,000

6 Analysis and Discussions

The performance of the split-step multi-layer ANN is evaluated in comparison with the benchmarking standard BP-based network, a PSO-based network and a LM network. In order to provide a fair and common initial ground for comparison purpose, the training process of the BP-based perceptron or LM network commences from the best initial population of the corresponding PSO-based perceptron or split-step network. Figure 1 shows the 1 week lead time normalized chlorophyll-a prediction for scenario 3 by all perceptrons in the validation process. Table 1 shows comparison of the results for chlorophyll-a forecasting with both 1 week and 2 weeks lead times for scenarios 1 to 3. It should be noted that the results do not exhibit a significant advantage of using more environmental variables as the network inputs and that 1 week lead time is better than its counterparts of 2 weeks. It can be observed that the split-step algorithm performs the best in terms of prediction accuracy. Table 2 shows the steady-state fitness evaluation times during training for various perceptrons. It can be observed that the split-step perceptron, with rate comparable to that of LM algorithm, exhibits much faster convergence than those by the BP-based perceptron and the PSO-based network.

7 Conclusions

In this paper, a perceptron based on a split-step PSO algorithm is employed for real-time prediction of algal blooms at Tolo Harbour in Hong Kong with different lead times and input variables. The results do not exhibit any advantage of using more environmental variables as the network inputs. The chlorophyll-a output from the 1 week time-lagged chlorophyll-a input is shown to be a robust forewarning and decision-support tool. The results also show that the split-step PSO-based perceptron outperforms the other commonly used optimization techniques in algal bloom prediction, in terms of both convergence and accuracy.

8 Acknowledgement

This research was supported by the Research Grants Council of Hong Kong (PolyU5132/04E).

References

1. Chau, K.W., Cheng, C.T.: Real-time Prediction of Water Stage with Artificial Neural Network Approach. Lecture Notes in Artificial Intelligence, **2557** (2002) 715-715
2. Rumelhart, D.E., Widrow, B., Lehr, M.A.: The Basic Ideas in Neural Networks. Communications of the ACM **37(3)** (1994) 87-92
3. Hagan, M.T., Menhaj, M.B.: Training Feedforward Networks with the Marquardt Algorithm. IEEE Transactions on Neural Networks **5(6)** (1994) 989-993
4. Kennedy, J., Eberhart, R.: Particle Swarm Optimization. Proceedings of the 1995 IEEE International Conference on Neural Networks. Perth (1995) 1942-1948
5. Chau, K.W.: River Stage Forecasting with Particle Swarm Optimization. Lecture Notes in Computer Science **3029** (2004) 1166-1173
6. Chau, K.W.: Rainfall-Runoff Correlation with Particle Swarm Optimization Algorithm. Lecture Notes in Computer Science **3174** (2004) 970-975
7. Kennedy, J.: The Particle Swarm: Social Adaptation of Knowledge. Proceedings of the 1997 International Conference on Evolutionary Computation. Indianapolis (1997) 303-308
8. Clerc, M., Kennedy, J.: The Particle Swarm—Explosion, Stability, and Convergence in a Multidimensional Complex Space. IEEE Transactions on Evolutionary Computation **6(1)** (2002) 58-73
9. Chau, K.W., Jin, H.S.: Eutrophication Model for a Coastal Bay in Hong Kong. Journal of Environmental Engineering ASCE **124(7)** (1998) 628-638
10. Chau, K.W., Jin, H.S., Sin, Y.S.: A Finite Difference Model of Two-Dimensional Tidal Flow in Tolo Harbor, Hong Kong. Applied Mathematical Modelling **20(4)** (1996) 321-328