Journal of Hydrology, Vol. 346, No. 3-4, 2007, pp 131-135

A Split-Step Particle Swarm Optimization Algorithm in river stage forecasting

K.W. Chau
Department of Civil & Structural Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, China (e-mail: cekwchau@polyu.edu.hk)

Abstract

An accurate forecast of river stage is very significant so that there is ample time for the pertinent authority to issue a forewarning of the impending flood and to implement early evacuation measures as required. Since a variety of existing process-based hydrological models involve exogenous input and different assumptions, artificial neural networks have the potential to be a cost-effective solution. In this paper, a split-step particle swarm optimization (PSO) model is developed and applied to train multi-layer perceptrons for forecasting real-time water levels at Fo Tan in Shing Mun River of Hong Kong with different lead times on the basis of the upstream gauging station (Tin Sum) or at Fo Tan. This paradigm is able to combine 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. The results demonstrate that it is able to attain a higher accuracy in a much shorter time when compared with the benchmarking backward propagation algorithm as well as the standard PSO algorithm.

Keywords: River stage forecasting; split-step; particle swarm optimization; Levenberg-Marquardt algorithm; artificial neural networks

Introduction

Throughout these years, prediction of river stages has long been an important research topic in hydrologic engineering because flooding is a type of natural disaster that has been occurring for centuries, but can only be mitigated rather than completely solved. An accurate water stage prediction allows the pertinent authority to issue a forewarning of the impending flood and to implement early evacuation measures when required. Mathematical models are conventionally used to forecast flow in a water body. In general, they require exogenous input and embrace different assumptions. Conventional numerical modeling addresses the physical problem by solving a highly coupled, non-linear, partial differential equation set. The involving processes affecting flooding occurrence are highly complex and uncertain which may consume enormous computing cost and time. In this sense, existing numerical models are not totally satisfactory in representing the highly complex inter-relationships.

In the past decade, soft computing (SC) techniques have been gradually becoming a trend to complement or replace the process-based models. Amongst others, artificial neural networks (ANN), in particular the feed forward back-propagation (BP) perceptrons, have been widely applied in different fields including water resources engineering. It is claimed that the multi-layer perceptrons can be trained to approximate and accurately generalize virtually any smooth, measurable function whilst taking no prior assumptions concerning the data distribution. Characteristics, including built-in dynamism in forecasting, data-error tolerance, and lack of requirements of any exogenous input, render it attractive for use in river stage prediction in hydrologic engineering.

Thirumalaiah and Deo [1] depict the use of a conjugate gradient ANN in real-time forecasting of water levels, with verification of untrained data. Liong et al. [2] demonstrate that a feed forward ANN is a highly suitable flow prediction tool yielding a very high degree of water level prediction accuracy in Bangladesh. Kumar et al. [3] employed the k nearest neighbors of the monthly spatial flow pattern to approximate nonparametrically the probability distribution of the vector of disaggregated flows conditional on the multisite monthly flows. Chau and Cheng [4] describe the sensitivity of various network characteristics for real-time prediction of water stage with the ANN approach in a river in Hong Kong. Raju et al. [5] adopted an integrated irrigation planning strategy for the case study of Jayakwadi irrigation project, Maharashtra, India, incorporating linear programming models, multiobjective optimization, Kohonen neural networks based classification algorithm and multicriterion analysis technique. Cheng et al. [6] perform a long-term prediction of discharges in Manwan Reservoir using several artificial neural network models. Although the BP algorithm is commonly used in recent years to perform the training task, some drawbacks are often encountered in the use of this gradient-based method. They include: the training convergence speed is very slow; it is easily to get stuck in a local minimum [7]. Different algorithms have been proposed in order to resolve these drawbacks, yet the results are still not fully satisfactory [8]. Levenberg-Marquardt (LM) optimization technique [9] is a commonly used ANN that has attained certain improvements such as convergence rates over the BP algorithm. Particle swarm optimization (PSO) is another recently developing SC technique that has been applied to different fields [10-12]. This technique has been applied in hydrological problems and accomplished satisfactory results [13-14]. Moreover, a combination of global and local search methods, such as [15-16] can be explored.

In this paper, a split-step PSO algorithm, coupled with LM technique, is employed to train multi-layer perceptrons for forecasting real-time water levels at Fo Tan in Shing Mun River of Hong Kong with different lead times on the basis of the upstream gauging station (Tin

Sum) or at Fo Tan. The split-step is the key improvement over [13]. 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.

Attributes of PSO algorithm

PSO is an algorithm tailored to optimize complicated numerical functions based on metaphor of human social interaction [8]. Although it is initially developed as a tool for modeling social behavior, the PSO algorithm has been recognized as a computational intelligence technique intimately related to evolutionary algorithms [17-18]. This optimization paradigm simulates the ability of human societies in processing knowledge [18]. It is a populated search method for optimization of continuous nonlinear functions resembling the biological movement in a fish school or bird flock. The basic principle of PSO algorithm is formed on the assumption that potential solutions will be flown through hyperspace with acceleration towards more optimum solutions. 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.

The most significant advantage of PSO algorithm is its relatively simple coding and hence low computational cost. It is quite similar to a genetic algorithm in aspects of 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 [17-18]. This algorithm can be readily adopted to train the multi-layer perceptrons as an optimization technique, as presented in the following section.

Adaptation to training of perceptrons

For this case for adaptation to training of perceptrons, a three-layered perceptron is considered here. $W^{[1]}$ 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.

$$V_{i,t+1}^{[j]}(m,n) = V_{i,t}^{[j]}(m,n)^t + r\alpha[P_{i,t}^{[j]}(m,n) - W_{i,t}^{[j]}(m,n)] + s\beta[P_{b,t}^{[j]}(m,n) - W_{i,t}^{[j]}(m,n)] \tag{1}$$

where j = 1, 2; t is the time step; t+1 is the new time step; $m = 1, \ldots, M_j$; $n = 1, \ldots, N_j$; 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 $r\alpha[P_{i,t}^{[j]}(m,n) - W_{i,t}^{[j]}(m,n)]$ denotes the private thinking of the particle itself whilst the social part $s\beta[P_{b,t}^{[j]}(m,n) - W_{i,t}^{[j]}(m,n)]$ represents the collaboration among the particles as a group. The new position is then determined based on the new velocity as follows:

$$W_{i,t+1}^{[j]} = W_{i,t}^{[j]} + V_{i,t}^{[j]} \tag{2}$$

The fitness of the i-th particle is determined in terms 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.

The split-step PSO paradigm

It is believed that 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 [9] to fine tune the fast local search. The drawbacks of either entrapment in local minima in LM algorithm or longer time consumption in global search of PSO algorithm might be avoided in this new paradigm.

Forecasting water stage in Shing Mun River

The capability of any model can only be validated through prototype application to mimic a particular case study with accurate depiction of real phenomena. In this case, the system is employed to study the potential flood hazards in terms of water levels in the Shing Mun River network, Hong Kong. Figure 1 shows the schematization of Shing Mun river channel together with the three tributary channels. The author has studied the tidal dynamics and potential flood hazards in this river network [19-22]. Details regarding the location of the Shing Mun River and its tributary nullahs can be found in [19-22] and are not repeated here. The maximum flow at the river for a 200-year storm is about 1500 m^3/s. The existing Shing Mun River has been trained for a length of about 2840m, from the bell-mouth outlet of Lower Shing Mun Dam to Sha Tin Tsuen. The three minor streams, i.e. the Tin Sum, Fo Tan and Siu Lek Yuen nullahs, form tributaries of the extended river. Surface water from an extensive catchment with an area of approximately 5200 ha flows into Sha Tin Hoi via the Shing Mun River.

The data available at the study area pertain to continuous stages from 1999 to 2002, in the form of daily water levels. In this study, water levels at Fo Tan is forecasted with a lead time of 1 and 2 days based on the measured daily levels there and at Tin Sum, which is located at 2500m upstream of Fo Tan. In total, 1095 pairs of daily levels were available, of which 730 were used for training and 365 were used to validate the network results with the observations. The division of data is tailored so as to include extreme frequency and intensity in both sets of data. In other words, it is ensured that the data series chosen for training and validation comprised both high and low discharge periods of the year and also rapid changes in water stages.

The perceptron has an input layer with one neuron, a hidden layer with three neurons, and output layer with two neurons. The input neuron represents the water stage at the current day whilst the output nodes include the water stages after 1 day and 2 days, 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. In both the PSO-based perceptron and split-step PSO-based perceptron, the number of population is set to be 40 whilst the maximum and minimum velocity values are 0.25 and -0.25 respectively. The learning rate is 0.9 and the number of epoch is 10,000. The computer code is developed under Visual Basic environment.

Results, analysis and discussions

Figure 2 shows the comparison between observed and predicted flow levels using spilt-step PSO. In order to gauge objectively the performance of the split-step multi-layer ANN, its results are compared 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. Since forecasts precision of high flows and extreme floods is more important than the precision of normal flows, special consideration is paid to testing the model performance in prediction of high floods. Two performance measures are employed in this study: (i) the correlation coefficient between the field and simulated data; and (ii) steady-state fitness evaluation times during training.

Table 1 shows comparison of the water stage forecasting results by different perceptrons based on data at the same station and at different station during high floods. It can be observed that, in terms of prediction accuracy, the split-step algorithm performs the best. For both data inputs at Fo Tan and at Tin Sum under training and validation processes, the order is consistent and is as follows: the split-step algorithm, PSO algorithm, LM algorithm and then BP algorithm. Moreover, it should be noted that 1 day lead time is better than its counterparts of 2 days and that forecasting at Fo Tan made by using the data collected at the upstream gauge station (Tin Sum) is generally better compared to the data collected at the same station. This result may be due to the average travel time of flow between the stations.

Table 2 shows the steady-state fitness evaluation times during training for various perceptrons. The fitness evaluation time is equal to the product of the population with the number of generations. It can be observed that the split-step PSO perceptron, with rate comparable to that of LM algorithm, exhibits much faster convergence than those by the BP-based perceptron and the standard PSO-based network. Although the improvement in performance by this novel algorithm over others is not substantial, the results are still encouraging because the improvement is observed to be consistently better over all cases (for different lead times, locations as well as training/validation processes).

Conclusions

In this paper, a perceptron based on a split-step PSO algorithm is employed for real-time prediction of water stage at Shing Mun River in Hong Kong with different lead times on the basis of the upstream gauging station or stage/time history at the specific station. The training and verification simulation results show that the split-step PSO-based perceptron outperforms the other commonly used benchmarking optimization techniques in water stage prediction, in terms of both convergence and accuracy. It is demonstrated that this novel hybrid optimization algorithm, which is able to provide model-free estimates in deducing the output from the input, is an appropriate forecasting tool. Moreover, forecasting at Fo Tan made by time-lagged water stage is shown to be a robust forewarning and decision-support tool.

Acknowledgements

This research was supported by the Central Research Grant of Hong Kong Polytechnic University (G-U162).

References

[1]. K. Thirumalaiah, and M. C. Deo, "River stage forecasting using artificial neural networks", Journal of Hydrologic Engineering, ASCE, Vol 3, No. 1, pp. 26-32, 1998.

[2]. S. Y. Liong, W. H. Lim, and G. N. Paudyal, "River stage forecasting in Bangladesh: neural network approach", Journal of Computing in Civil Engineering, ASCE, Vol 14, No. 1, pp. 1-8, 2000.

[3]. D. N. Kumar, U. Lall, and M. R. Petersen, "Multisite disaggregation of monthly to daily streamflow", Water Resources Research, Vol. 36, No. 7, pp. 1823-1833, 2000.

[4]. K. W. Chau, and C. T. Cheng, "Real-time prediction of water stage with artificial neural network approach", Lecture Notes in Artificial Intelligence, Vol 2557, pp. 715-715, 2002.

[5]. K. S. Raju, D. N. Kumar, and L. Duckstein, "Artificial neural networks and multicriterion analysis for sustainable irrigation planning", Computers & Operations Research, Volume 33, No. 4, pp. 1138-1153, 2006.

[6]. C. T. Cheng, K. W. Chau, Y. G. Sun, and J. Y. Lin, "Long-term prediction of discharges

in Manwan Reservoir using artificial neural network models”, Lecture Notes in Computer Science, Vol. 3498, pp. 1040-1045, 2005.

[7]. D. E. Rumelhart, B. Widrow, and M. A. Lehr, “The basic ideas in neural networks”, Communications of the ACM, Vol 37, No. 3, pp. 87-92, 1994.

[8]. R. Govindaraju, and A. Rao, Artificial Neural Networks in Hydrology, Kluwer Academic Publishers, Dordrecht, 2000.

[9]. M. T. Hagan, and M. B. Menhaj, “Training feedforward networks with the Marquardt algorithm”, IEEE Transactions on Neural Networks, Vol. 5, No. 6, pp. 989-993, 1994.

[10]. J. Kennedy, and R. Eberhart, “Particle swarm optimization”, Proceedings of the 1995 IEEE International Conference on Neural Networks, Perth, pp. 1942-1948, 1995.

[11]. D. N. Kumar, and M. J. Reddy, “Multipurpose reservoir operation using particle swarm optimization”, Journal of Water Resources Planning and Management, ASCE, Vol. 133, No. 3, pp. 192-201, 2007.

[12]. L. J. Li, Z. B. Huang, F. Liu, and Q. H. Wu, “A heuristic particle swarm optimizer for optimization of pin connected structures”, Computers & Structures, Vol. 85, No. 7-8, pp. 340-349, 2007.

[13]. K. W. Chau, “Particle swarm optimization training algorithm for ANNs in stage prediction of Shing Mun River”, Journal of Hydrology, Vol. 329, No. 3-4, pp. 363-367, 2006.

[14]. K. W. Chau, “Rainfall-runoff correlation with particle swarm optimization algorithm”, Lecture Notes in Computer Science, Vol 3174, pp. 970-975, 2004.

[15]. Q. J. Wang, “The genetic algorithm and its application to calibrating conceptual rainfall-runoff models”, Water Resources Research, Vol. 27, No. 9, pp. 2467-2471, 1991.

[16]. Q. Y. Duan, V. K. Gupta, S. Sorooshian, “Shuffled complex evolution approach for effective and efficient global minimization”, Journal of Optimization Theory and Applications, Vol. 76, No. 3, pp. 501-521, 1993

[17]. J. Kennedy, “The particle swarm: social adaptation of knowledge”, Proceedings of the 1997 International Conference on Evolutionary Computation, Indianapolis, pp. 303-308, 1997.

[18]. M. Clerc, and J. Kennedy, “The particle swarm — explosion, stability, and convergence in a multidimensional complex space”, IEEE Transactions on Evolutionary Computation, Vol 6, No. 1, pp. 58-73, 2002.

[19]. K. W. Chau, “A prototype knowledge-based system on unsteady open channel flow in water resources management”, Water International, Vol 29, No. 1, pp. 54-60, 2004.

[20]. K. W. Chau, and J. H. W. Lee, “Mathematical modelling of Shing Mun river network”, Advances in Water Resources, Vol 14, No. 3, pp. 106-112, 1991.

[21]. K. W. Chau, and J. H. W. Lee, “A microcomputer model for flood prediction with applications”, Microcomputers in Civil Engineering, Vol 6, No. 2, pp. 109-121, 1991.

[22]. K. W. Chau, and W. W. Yang, “Development of an integrated expert system for fluvial hydrodynamics”, Advances in Engineering Software, Vol 17, No. 3, pp. 165-172, 1993.

Table 1. Results for river forecasting at Fo Tan during high floods based on input data at different stations

Input data	Algorithm	Coefficient of correlation			
		Training		Validation	
		1 day ahead	2 days ahead	1 day ahead	2 days ahead
Tin Sum	BP	0.944	0.940	0.943	0.938
	PSO	0.975	0.971	0.973	0.969
	LM	0.966	0.960	0.959	0.953
	Split-step	0.991	0.984	0.986	0.979
Fo Tan	BP	0.936	0.922	0.931	0.915
	PSO	0.964	0.953	0.961	0.948
	LM	0.952	0.941	0.947	0.931
	Split-step	0.982	0.971	0.975	0.965

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

Algorithm	Steady-state fitness valuation time
BP	18,000
PSO	8,500
LM	4,500
Split-step	5,500

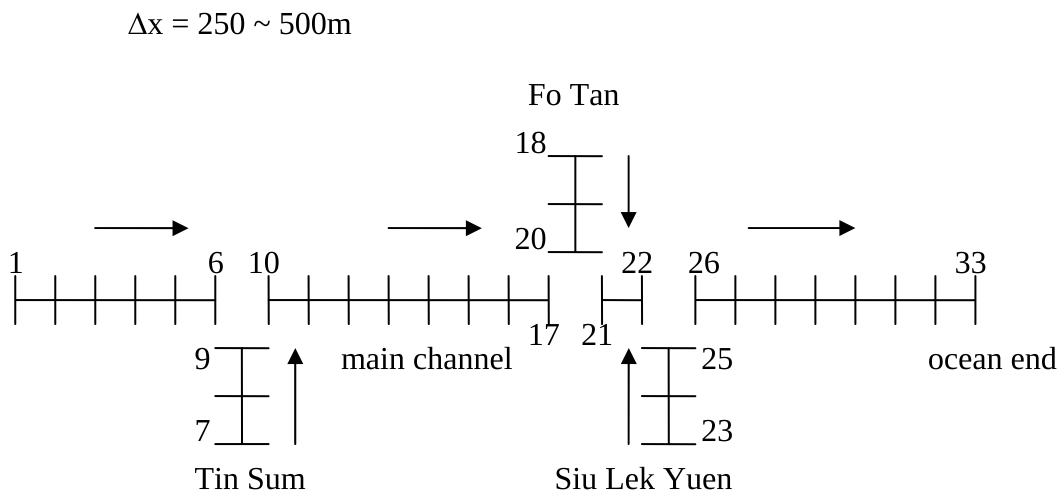

Figure 1. Schematization of Shing Mun river channel together with the three tributary channels

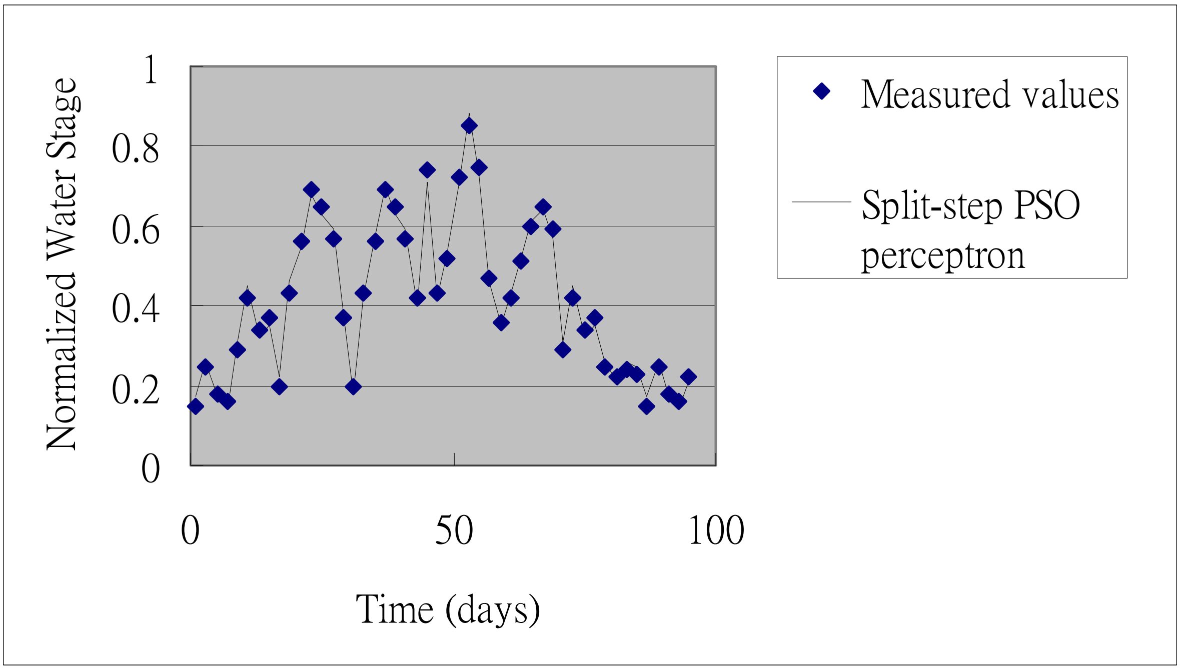

Figure 2. Comparison between observed and predicted flow levels using spilt-step PSO