Hydrological Model Derivation Using the Evolutionary Model Induction System

Hong, Yoon-Seok Timothy
Institute of Geological and Nuclear Sciences (GNS Science), New Zealand

ABSTRACT

Hydrological science and engineering attempts from trying to understand and predict the behaviour of environmental system with sustainable models. These models have traditionally used a mechanistic approach, which is driven by *first principles* (e.g. water or mass balance, chemical and physical laws, etc). This traditional mechanistic approach is typically a non-trivial task for many complex hydrological systems. The approach is time- and money- consuming to perform and can be simply impracticable due to a lack of adaptability and flexibility, lack of good data, model uncertainty, and unmodeled phenomena of unknown or partially known hydrological mechanisms.

The aim of this paper is to develop the evolutionary model induction system (Hong et al. 2003; Hong and Paik, 2007) using an evolutionary computational intelligence, called grammar-based genetic programming that is specially designed to automatically discover hydrological models that best fit observed hydrological data in the form of explicit mathematical formulas with high levels of accuracy. The evolutionary model induction system works on the population of individuals (mathematical models) applying the principle of survival-of-the-fittest to produce increasingly better models of observed data. The solution offered by each model is evaluated to give some measure of 'fitness', which indicates how 'fit' the solution is. At each generation, new sets of models, called 'offspring', are created by the process of selecting models according to their level of fitness and breeding them together using genetic operators such as reproduction, crossover, and mutation. These offspring (new models) then form the basis for the next generation. This process leads to the evolution of populations of models to produce still better models.

The evolution of hydrological models in the evolutionary model induction system is based on a repetitive computational process. The following steps summarize the evolutionary model induction system used in this work (Figure 1).

Initialization: Generate an initial population of P models using context-free grammar. Generation $K=0$.

(1) Optimize each model (variables and constants) in the population.
(3) Execute and evaluate the fitness of each model in the population.
(4) Genetic Loop: Repeat until termination criterion is met (maximum generation K_{max}).
- (4-a) Generate a new population by crossover, and mutation of models
  - (4-a-i) Select two models (parents) based on their fitness for breeding to the next generation.
  - (4-a-ii) Probabilistically perform crossover to randomly swap sub-trees of the two parents to two child models.
  - (4-a-iii) Probabilistically perform mutation to modify a random sub-tree of each child model.
  - (4-a-iv) Repeat steps (4-a-i) - (4-a-ii) to generate a total of P new offspring models.
  - (4-a-v) Old models from the parent population are replaced by new models generated.
- (4-b) Go to next generation: $K+1$.

More details of an evolutionary model induction system can be seen in the work of Hong et al. (2003) and Hong and Paik (2007).

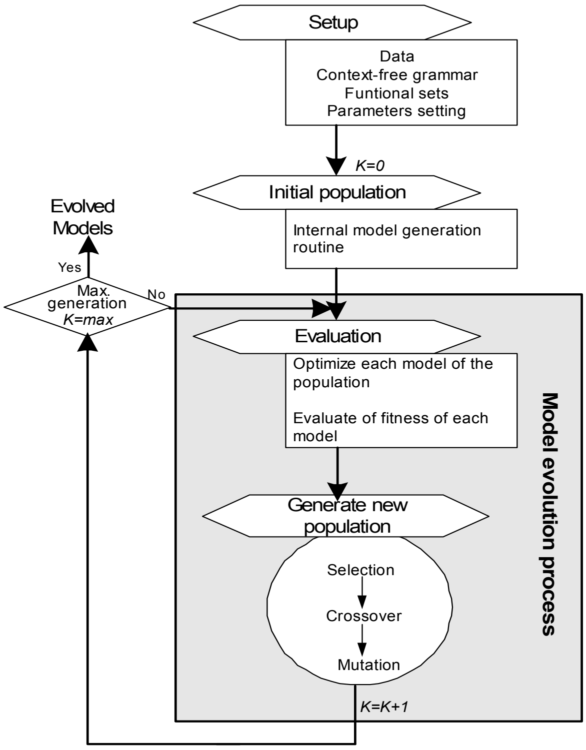

Fig. 1. Structure of the evolutionary model induction system.

The evolutionary model induction system is applied to derive models of rainfall recharge from observations of rainfall recharge and rainfall, calculated potential evapotranspiration (PET) and soil profile available water (PAW) at each of four monitoring sites over a four year period (1 May 1999 to 30 April 2003) in Canterbury, New Zealand. Rainfall recharge estimates provided by the best derived model from the evolutionary model induction system are compared with estimates of rainfall recharge provided by a soil water balance model and with observations at each of the four monitoring sites.

Results have shown that the evolutionary model induction system has evolved multivariate hydrological models automatically from the observed hydrological data in the form of understandable explicit mathematical formulas with high levels of accuracy by performing input selection, model form generation, and parameter estimation simultaneously of the automatic model generation system. The rainfall recharge model evolved by the evolutionary model induction system show a better prediction performance compared to a soil water balance model and the MLP-BP neural network model in the late spring, summer, and early autumn periods. The ‘‘best’’ rainfall recharge model derived from the evolutionary process model induction system provides estimates of cumulative sums of rainfall recharge that are closer than a soil water balance model to observations at all four sites. It is demonstrated that the evolutionary model induction system provides a good alternative to develop cost-effective hydrological models and is readily applicable to a variety of other complex hydrological processes.

REFERENCES

Hong, Y.-S.T., and Paik, B.C. (2007). "Evolutionary multivariate dynamic process model induction for a biological nutrient removal process." *Journal of Environmental Engineering,* ASCE, Vol. 133:12, pp.1126-1135.

Hong, Y.-S., and Rosen, M.R. (2002). "Identification of an urban fractured rock aquifer dynamics using an evolutionary self-organizing modelling." *J. Hydrology*, Vol. 259, pp.89–104.

Hong, Y.-S.T., White, P.A., and Scott, D.M. (2005). "Automatic rainfall recharge model induction by evolutionary computational intelligence." *Water Resources Research*, AGU, Vol. 41, No. 8, W08422, doi:10, 1029/2004WR003577.