

## **USE OF GENE EXPRESSION PROGRAMING FOR MULTI-MODEL COMBINATION OF RAINFALL- RUNOFF MODELS**

Achela K. Fernando<sup>1</sup>, Asaad Y. Shamseldin<sup>2</sup>, Robert J. Abrahart<sup>3</sup>

1. Senior Lecturer, Department of Civil Engineering, Unitec Institute of Technology,  
Carrington Road, Mt Alert, Auckland 1025, New Zealand. Email:

[afernando@unitec.ac.nz](mailto:afernando@unitec.ac.nz)

2. Associate Professor, Department of Civil and Environmental Engineering, University  
of Auckland, Auckland, New Zealand, E-mail [a.shamseldin@auckland.ac.nz](mailto:a.shamseldin@auckland.ac.nz)

3. Associate Professor, School of Geography, Nottingham University, Sir Clive Granger  
Building, University of Nottingham, University Park, Nottingham NG7 2RD, UK, E-mail

[bob.abrahart@nottingham.ac.uk](mailto:bob.abrahart@nottingham.ac.uk)

### **Abstract**

This paper deals with the application of an innovative method for combining estimated outputs from a number of rainfall-runoff models using Gene Expression Programming (GEP) to perform symbolic regression. The GEP multi-model combination method uses the synchronous simulated river flows from four conventional rainfall-runoff models to produce a set of combined river flow estimates for four different catchments.

The four selected models for the multi-model combinations are the Linear Perturbation Model (LPM), the Linearly Varying Gain Factor Model (LVGFM), the Soil Moisture

Accounting and Routing (SMAR) Model, and the Probability-Distributed Interacting Storage Capacity (PDISC) model. The first two of these models are 'black-box' models, the LPM exploiting seasonality and the LVGFM employing a storage-based coefficient of runoff. The remaining two are conceptual models. The data of four catchments with different geographical location, hydrological and climatic conditions, are used to test the performance of the GEP combination method.

The results of the model using GEP method are compared against original forecasts obtained from the individual models that contributed to the development of the combined model by means of a few global statistics. The findings show that a GEP approach can successfully used as a multi-model combination method. In addition, the GEP combination method also has benefit over other hitherto tested approaches such as an artificial neural network combination method in that its formulation is transparent, can be expressed as a simple mathematical function, and therefore is useable by people who are unfamiliar with such advanced techniques. The GEP combination method is able to combine model outcomes from less accurate individual models and produce a superior river flow forecast.

Keywords: (symbolic regression, gene expression programming, model combination, rainfall-runoff model)

## **Introduction**

In the context of rainfall-runoff modeling, the multi-model combination approach advocates the synchronous use of the simulated discharges of a number of rainfall-runoff models to produce an overall combined/integrated discharge which can be used as an alternative to that produced by a single rainfall-runoff model. The basic hypothesis made in the multi-model combination approach is that different models capture different aspects of the data and hence the combination of these aspects would produce better discharge estimates than those produced by any one of the individual models involved in the combination.

The use of the multi-model combination of rainfall-runoff models was advocated by Shamseldin(1997) and Shamseldin et al.(1997). Since then there have been several more studies which have dealt with multi-model combination of hydrological models (e.g. (Abrahart and See 2002, Ajami, et al. 2006, Coulibaly, et al. 2005, Hsu, et al. 2009, See and Openshaw 2000, Shamseldin and O'Connor 1999, Shamseldin, et al. 2007, Viney, et al. 2009, Xiong, et al. 2001)).As the nature of the combination function is unknown and no theory exists to analytically derive the combination function from a hydrological or physical point of view, previous studies have used empirical data-driven modeling to derive the combination function and such use is very appropriate.

In previous hydrological multi-model combination studies, both linear and non-linear soft computing (e.g. neural network and fuzzy-based) combination methods have been used to produce multi-model river flows (Coulibaly, et al. 2005, See and Openshaw 2000, Shamseldin 1997, Xiong, et al. 2001). The linear combination methods include the Simple Average Method (SAM) and the Weighted Average Method (WAM). In the

SAM, the combined output is simply the arithmetic average of the outputs of each individual model used in the combination. The SAM is a naïve forecast combination method, which can work very well when the constituent models have practically the same level of performance; it is more sensible to use it purely as a baseline against which the results of more sophisticated combination methods can be compared. In the WAM, the combined forecast is calculated as the weighted sum of the forecasts of the constituent rainfall-runoff models, thereby, allowing for the situation where the constituent models have significantly different levels of performance. Shamseldin et al. (1997) pioneered the use of an artificial neural network (ANN) for efficient multi-model combination of a group of rainfall-runoff models. In the previous studies ANN combination methods were found to be generally better than those of the linear weighting (WAM) and fuzzy-based methods (Xiong, et al. 2001) (See and Abrahart 2001, Shamseldin, et al. 1997, Xiong, et al. 2001). The Fuzzy-Based Combination Method (FBCM) uses a set of fuzzy IF-THEN rules to transform the synchronous individual model forecasts (operating as inputs) into a single combined output forecast. See and Openshaw (2000) and Xiong et al. (2001) introduced the use of fuzzy methods for the combination of river flow forecasts. Xiong et al.(2001) concluded that a first Order Takagi-Sugeno type FBCM can yield results that are better than those of the best of the constituent models, i.e. that it can function effectively as a hydrological multi-model combination technique. More recently guidelines for choosing an effective method of combination have been developed (Jeong and Kim 2009).

The above noted studies have demonstrated the potential capabilities of the multi-model combination approach in improving the accuracy and reliability of hydrological

modeling results and have laid the foundation for the further use of this approach in rainfall-runoff modeling. However, in these studies no attempts have been made to explore the nature of the combination function and its inner working. Furthermore, no explanation has been provided to clarify the improvements in the modeling results.

The use of Gene Expression Programming (GEP) (Ferreira 2006) to perform symbolic regression and develop a combination type rainfall-runoff model for a single river has been recently reported (Fernando, et al. 2009). The outcome of that preliminary study to develop a combined multi-model-output GEP model to simulate daily flow in a single river catchment concluded that the GEP model performance was superior in comparison to the individual models, the transparency of the GEP model was useful, and that there was a need to prevent over-fitting of the data to the GEP model.

The current paper extends that initial investigation to cover four river catchments with diverse characteristics and presents a set of mathematical equations that were evolved in GeneXproTools 4.0: a powerful software package that is used to accomplish the combinatorial operations. The paper focuses on further advancing our understanding about the inner working of the multi-model combination function which can hold the key for further improvements in modeling results as well as providing guidance about the effective development of multi-model combination systems. GeneXproTools 4.0 is used to develop multi-model combination functions based on daily estimates of four rainfall-runoff models for four catchments with different geographical location and climate conditions. In broad terms, symbolic regression is very similar to traditional parametric regression in the sense it attempts to derive a functional relationship/model which describes the relationship between dependent and

independent variables. In traditional parametric regression, the form of the function relating dependent and independent is specified a priori and the usual regression procedures are used to estimate the corresponding parameter values. Symbolic regression is a form of nonparametric regression in which the function relating dependent and independent variables is not specified a priori but the function is constrained to contain a number of mathematical or logical expressions to be chosen from a larger set of pre-selected expressions. GEP is used to simultaneously select the optimum set of expressions involving the appropriate input variables.

The four selected models for the multi-model combinations are: the Linear Perturbation Model (LPM), the Linearly Varying Gain Factor Model (LVGFM), the Soil Moisture Accounting and Routing (SMAR) Model and the Probability-Distributed Interacting Storage Capacity (PDISC) model. The first two models are 'black-box' models, the LPM exploiting seasonality and the LVGFM employing a linear variation of the runoff coefficient with discharge. The remaining two are conceptual rainfall-runoff models. Further details on these models and their applications are given by others (Ahsan and O'Connor 1994, Kachroo 1992, Moore 1985, Nash and Barsi 1983, Senbeta, et al. 1999, Shamseldin, et al. 1997, Tan and O'Connor 1996).

The rest of this paper is structured as follows. The next section gives an introduction to GEP and explains the basis of performing symbolic regression using GEP, as implemented in the software package GeneXproTools 4.0. The section after that is devoted to providing an explanation of the data used in the study. This is then followed by a section discussing the application and the results. The final section outlines the main conclusions of the paper and the limitations of this study.

## Gene Expression Programming

GEP is an evolutionary algorithm that is used to develop computer programs based on a search and optimization technique using analogies from natural selection and evolution. GEP captures the best properties of Genetic Algorithms (GA) and Genetic Programming (GP), but eliminating the constraints associated with implementing the genetic operators inherent to them. Although GA and GP are well known in the literature, application of GEP is only beginning to emerge (Steeb, et al. 2005). In GEP individual expressions are encoded in linear chromosomes which in turn make up expression trees whereas in GA it is done in the reverse, i.e., the individuals are parse trees, which can be somewhat cumbersome, that can thereafter be expressed as a linear string.

A detailed description of the operation of the GEP algorithm is not intended here; only a brief explanation of the underlying principles is instead presented by means of a simple example. For a detailed explanation of GEP theory and its operation refer to the extensive work of Ferreira (Ferreira 2001, Ferreira 2006, Ferreira 2009). GEP utilizes genotype (Linear Chromosome: LC) expressions which are translated to phenotype (Expression Tree: ET) expressions. Any algebraic and/or mathematical expression can be encoded as a genotype expression, i.e., as a gene of appropriate length made up of two main parts - a head and a tail. For instance, the algebraic expression involving two independent variables  $a$  and  $b$ ,  $\text{Sin}[(a+ab).a.\text{Sin}(b)]$  is represented by the genotype expression **S\*+\*a\*Sabababaa** where S stands for Sine Function. The length of this gene is 15, with the front part made up of the “head”, shown bold and underlined, of length 7 (usually chosen as appropriate to the problem) involving both functions and

variables, and a “tail” of length 8 involving only the variables. If the length of the head of the gene is  $h$ , the tail length  $t$  is then given by  $t = (n-1)h + 1$  where  $n$  is the number of variables (in this instance,  $n=2$ ). The ET for this gene, easily developed by following the coding sequence of the gene from left to right, is as shown in the Figure 1. As can be seen, although the gene itself is of length 15, only the functions and variables up to position 11 are required to express this algebraic equation. During evolution in GEP the gene undergoes standard genetic operations such as copying, cross-over and mutation. A mutation at, say, the 5th position in the gene from the variable “a” to division function “/” converts the expression from S\*+\*a\*Sabababaa to S\*+\*/\*Sabababab. This changes the expression to  $\text{Sin}[(b/a)+a.b).(a.\text{Sin}(b))]$  and the ET to the one shown in Figure 2.

The evolution of GEP is guided by the fitness of the functions generated during the evolution process. The fitness could be a measure of the error (i.e. the objective function) signifying the difference between the function outcome and the actual expected value. The actual GEP allows the concatenation of several functions (or genes), either through addition/subtraction or multiplication/division. The evolution of GEP can be terminated either when the fitness of the function reaches an optimum value or when the algorithm has evolved through a pre-specified number of evolutions.

GEP can, given the relevant data, find the mapping function between the independent variables (outputs from the individual models) and the dependent variable (the actual flow rate). The algorithm begins by setting up arbitrarily a gene based on user specified functions and input/independent variables to form the head of gene, the length of head, user specified or default values for genetic operators. The error between the program evaluated output and the expected output/dependent output (fitness of the



function) is calculated. In the following iterations, the gene is subjected to genetic operators at user-defined rates as a result of which a new equation is formed and the new fitness is evaluated. The evolution of the equation is terminated when a user-defined threshold fitness value or a number of iterations is achieved.

## **Independent Model Data Sets**

Five separate sets of data for each of the four major rivers were used for this study: the first set is the observed flow at the river gauge and the remaining four are the predicted daily flow values from four conventional models - Linear Perturbation (LPM), the Linearly Varying Gain Factor (LVGFM), the Soil Moisture Accounting and Routing (SMAR), and the Probability-Distributed Interacting Storage Capacity (PDISC). Each model was calibrated to the data of each catchment and the calibrated model parameters were used to produce the estimated discharges. In each catchment, the models and the combination methods were developed and verified using the same calibration and verification period. Full details of the procedures used in calibrating the four rainfall-runoff models can be found elsewhere (Ahsan and O'Conner 1994, Kachroo 1992, Moore 1985, Nash and Barsi 1983, Senbeta, et al. 1999, Shamseldin, et al. 1997, Tan and O'Connor 1996). LPM, LVGFM and PDISC models were calibrated on minimization of the least squares objective function. For SMAR model, the objective function reflected both the volumetric error and the mean square of the errors (Shamseldin et al, 1997).

Of the four rivers included in this study, three are situated in China and one in Ireland. The data for these river catchments have been previously used for combination model development research and therefore provide good prospects for comparison.

Table 1 summarizes some basic information regarding these catchments.

The Baihe catchment drains to a tributary of the Yangze River. This catchment, situated in a semi-arid to arid part of China, contributes extremely low flow inputs during most parts of the year but experiences peak flow rates of nearly ten times the dry weather flow during the flood season. The Brosna river catchment in Ireland is very flat and covered largely by grassland. The flow pattern in this river shows strong seasonal variation and a substantial base flow element.

The Shiquan catchment flows from a mountainous region and displays rapid rises and falls of daily flow rates. Nearly 70-80% of the annual rainfall in the catchment falls between May and October; this is evident from the flow records that confirm near-zero values during the remaining months of the year in this semi-arid catchment.

The upper reaches of the Yanbian catchment are relatively flat. Major flooding occurs during the wet season that runs from June to October and the flow rate recedes gradually to reach extremely low levels from November to May until the rains begin in June.

Of the four conventional models, SMAR is the individual model that best predicts the daily flow in all these catchments giving the highest correlation with the observed flow values. Full particulars are provided in a subsequent section.

## **Application of GenXProTools**

In this study, for each catchment the GenXProTools<sup>®</sup> tool was used to identify the mathematical relationship between the input variables - the individual rainfall runoff model estimated river flows - and the required output - the observed daily river flow.

This study is different to other reported applications of GEP in hydrology [e.g. (Aytek and Alp 2008, Barbulescu and Bautu 2009)] in that GEP is used to synthesize a transparent and more accurate model of predicted daily flow using a functional regression of combinations of predictions from other more conventional models. To the best of researchers' knowledge this has not yet been attempted to this extent.

### **Input/output Variables**

The four independent input variables used in this study comprised concurrent estimated discharge values obtained from LPM, LVGFM, PDISC and SMAR models. The expected output was the measured daily flow of the river. Table 1 summarizes the data sample sizes for training and testing.

### **Functions**

The basic functions used for the GEP in the software are given in Table 2. As the intention is to produce a simple combination function, the more complex options were excluded from this selection. The other important GEP model parameters and the notations of the inputs (i.e. individual model outputs) are shown in Table 3. As indicated in this table, the solution is expected to be the sum of three expressions (resulting from three genes) and the fitness of a population in one generation is calculated based on the Mean Square Error (MSE). The tool was used to evolve the GEP through 100,000 evolutions and the resulting final model for each catchment made up of three additive expression trees, resulting from three genes G1, G2 and G3, was identified.

### **Results**

Each evolved component and final mathematical multi-model for the four catchments is

shown in Figures 3-6. Each multi-model contains three Subordinate Expression Trees (indicated as Sub-ET 1, Sub-ET 2 and Sub-ET3) or genes (G1, G2 and G3) which in turn are constituted by a selection of input variables and functions. For instance, the first expression tree (Sub-ET1 or G1) for the Baihe River is the mathematical function  $G1 = 0.48(PDISC + 0.48 \sin 3.85) - LVGFM$ . Each model is made up of three genes or components, addition of which give the final mathematical expression for the combination of the individual rainfall-runoff models. Each gene (G1, G2 and G3) is not expected to contribute an equal amount to the predicted multi-model output. In order to understand the role each gene plays in the GEP combination model, the contribution by each gene/component towards the combined model output was then evaluated. The primary gene is the one that contributes most to forming the output and is also identified in the captions of the Figures 3-6 in which the expression for that gene is underlined.

Figure 7 (a -h) shows contributions from the primary component (on the primary y-axis with a larger range) and the two secondary components (on the secondary y-axis with a smaller range of +/- values) plotted for each catchment for training and testing sets respectively left to right. From these results it can be seen that one of the genes makes a dominant contribution (large variations depicted on the major primary scale) towards the combination model output and the other two act as supplementary components representing a relatively very small corrective signal (smaller variations depicted on the minor secondary scale).

Table 4a summarizes the individual models that participate in the primary and secondary components of each multi-model. The PDISC model is present as a primary contributor in three out of the four catchment models. As shown in Table 4b, out of the

four conventional models, the SMAR model is the individual model that has ranked highest in predicting the daily flow in all four catchments giving the highest correlation with the observed flow values. Comparing the information in these two tables, it can be seen that in most cases each GEP combined model involves the 1<sup>st</sup> and/or 2<sup>nd</sup> ranked individual in its primary component indicating preference of selection and/or elimination of redundant duplicate signals. Furthermore, the GEP combined models also use the lowest ranking individual models (PDISC and LVGFM) to form their primary components in all but one catchment; the Brosna catchment only uses one independent model in its primary component - SMAR.

The GEP multi-models for Baihe and Yanbian catchments do not include the highest ranking individual model SMAR at all. The GEP multi-models for Yanbian and Shiquan catchments contain all of the four individual rainfall-runoff models. Some explanation can be given for this selective participation of some of the individual models using how well these hydrologically different catchments were represented by the individual models, and how the individual model predictions correlate to the combined GEP model predictions.

Table 5 summarizes the correlation between the individual and the multi-model outputs and actual observations. The table shows that without exception, the GEP combined models deliver daily flow estimates that possess a higher correlation coefficient with the observed flow than the correlation coefficient values associated with the individual models.

As explained earlier, semi-arid Baihe catchment has widely varying flow and conceptual models such as SMAR and PDISC provide best estimates for individual

model predictions (Table 4b). The GEP model has involved the PDISC output as its primary component. For this catchment, the correlation between the SMAR model outputs and the other individual models, particularly, LPM ( $r=0.94$ ) and PDISC ( $r=0.935$ ) are significantly high implying that the patterns in the SMAR outputs are encapsulated in the latter two. Thus the participation of the PDISC and LVGFM in the primary component and LPM in the secondary had pre-empted the need for the SMAR models input. Although the individual SMAR model has the highest correlation with the observed flow ( $\sim 0.979$ ), the combined GEP model provides a better correlation ( $\sim 0.993$ ).

The Brosna catchment is located in a temperate climatic region and both the SMAR (conceptual) and the LPM (black box type) individual models provide high accuracy for daily flow estimates (See Table 4b). The PDISC model output has poor correlation with the observed data ( $r=0.643$ ) and has been completely left out of the combination model. The SMAR model, on the other hand, has the highest correlation with observed values ( $r=0.965$ ) and has participated in the dominant ET in the GEP model. The combined GEP model gives the best correlation with the observed flow ( $r=0.996$ ).

The semi-arid Shiquan catchment with near-zero flows in most part of the year is best represented by the conceptual models SMAR and PDISC (Table 4b). The combined model has used both of these models and also LVGFM in the primary component to produce a superior GEP model that has the highest correlation with the observed values ( $r=0.993$ ) (Table 5).

The Yanbian catchment that displays gradual recession between rainy seasons, has been fairly well represented by both the conceptual models as well as the black box

models, with correlation coefficients  $r > 0.935$  for all models (See Table 4b). The combined GEP combined model has chosen the LVGFM (black box type) and PDISC (conceptual) model outputs to synthesize a forecast superior to all the individual models with  $r = 0.994$  (see Table 5).

Table 6 summarizes the statistics for both the training and testing sets for all four catchments which highlight the superiority of the performance of the GEP models compared to the individual models. All of the individual models were calibrated and verified using the same calibration and verification periods. Summarizing the information in the tables 4 – 6, it can be concluded that the GEP combination model provides a superior solution; in addition, the model equation can be explicitly written as a simple mathematical function.

Figure 8 (a-h) shows the correlation between the observed and combined model predicted flow values for the four catchments for training and testing sets respectively from left to right. They show fairly tight scatter about the observed values.

Figure 9 (a-h) shows the flow duration curves for the catchments for the training and testing set, from left to right, which confirm that the GEP combined model predictions give the closest fit to the observed curve.

## **Conclusions**

The following conclusions can be made:

Compared to individual LVGFM, PDISC, LPM, and SMAR models, the GEP combination model that combines the synchronous flow forecast values from these four conventional models is able to provide a superior estimate of the flow forecast for all

four catchments considered in this study.

The GEP combination modeling approach offers a promising way to seek flow forecasts through a transparent model combination method as opposed to the “Black box” approaches used so far to combine a set of individual models. The mathematical equations that combine the synchronous outputs of individual models can be obtained and used further for forecasting without relying on sophisticated software. Thus this could be an alternative that can be provided to practitioners who do not necessarily have an in-depth understanding of novel techniques such as ANN or GEP concepts.

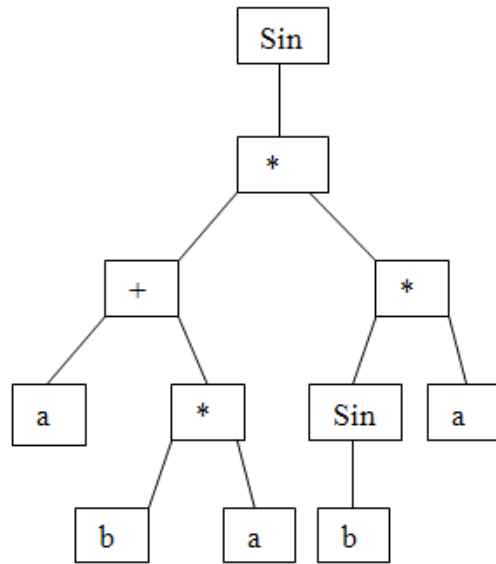
This research sought to establish that a superior and transparent combination model can be achieved using the novel GEP technique and using individual model outputs; further research is being carried out to compare the GEP technique with other transparent as well as black box combination techniques and to investigate ways to improve the performance of the GEP combined model to make it universally acceptable.

## References

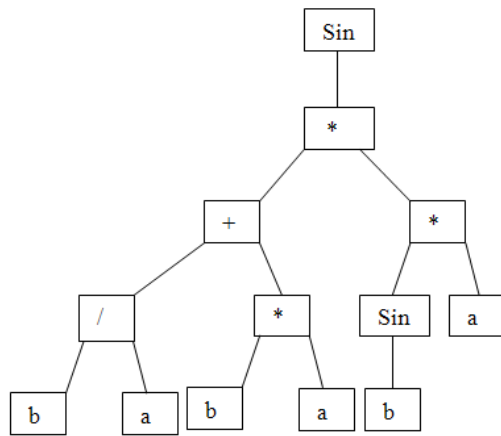
- Abrahart, R. J., and See, L. M. (2002). "Multi-model data fusion for river flow forecasting: an evaluation of six alternative methods based on two contrasting catchments." *Hydrology and Earth System Sciences*, 6(4), 655-670.
- Ahsan, M., and O'Conner, K. M. (1994). "A simple non-linear rainfall-runoff model with a variable gain factor." *Journal of Hydrology*, 155, 151-183.
- Ahsan, M., and O'Connor, K. M. (1994). "A simple rainfall-runoff model with variable gain factor." *Journal of Hydrology*, 155, 151-183.
- Ajami, N. K., Duan, Q., Gao, X., and Sorooshian, S. (2006). "Multimodel combination techniques for analysis of hydrological simulations: application to Distributed Model Intercomparison Project results." *Journal of Hydrometeorology*, 7, 755–768.
- Aytek, A., and Alp, M. (2008). "An application of artificial intelligence for rainfall-runoff modeling." *Journal of Earth System Science*, 87(2), 145-155.
- Barbulescu, A., and Bautu, A. (2009). "Meteorological time series modeling using an adaptive gene expression programming." *Proceedings of the 10th WSEAS international conference on evolutionary computing*, World Scientific and Engineering Academy and Society (WSEAS), Prague, Czech Republic, 17-22.
- Coulibaly, P., Hache, M., Fortin, V., and Bobee, B. (2005). "Improving Daily reservoir inflow forecasts with model combination." *ASCE Journal of Hydrologic Engineering*, 10(2), 91-99.



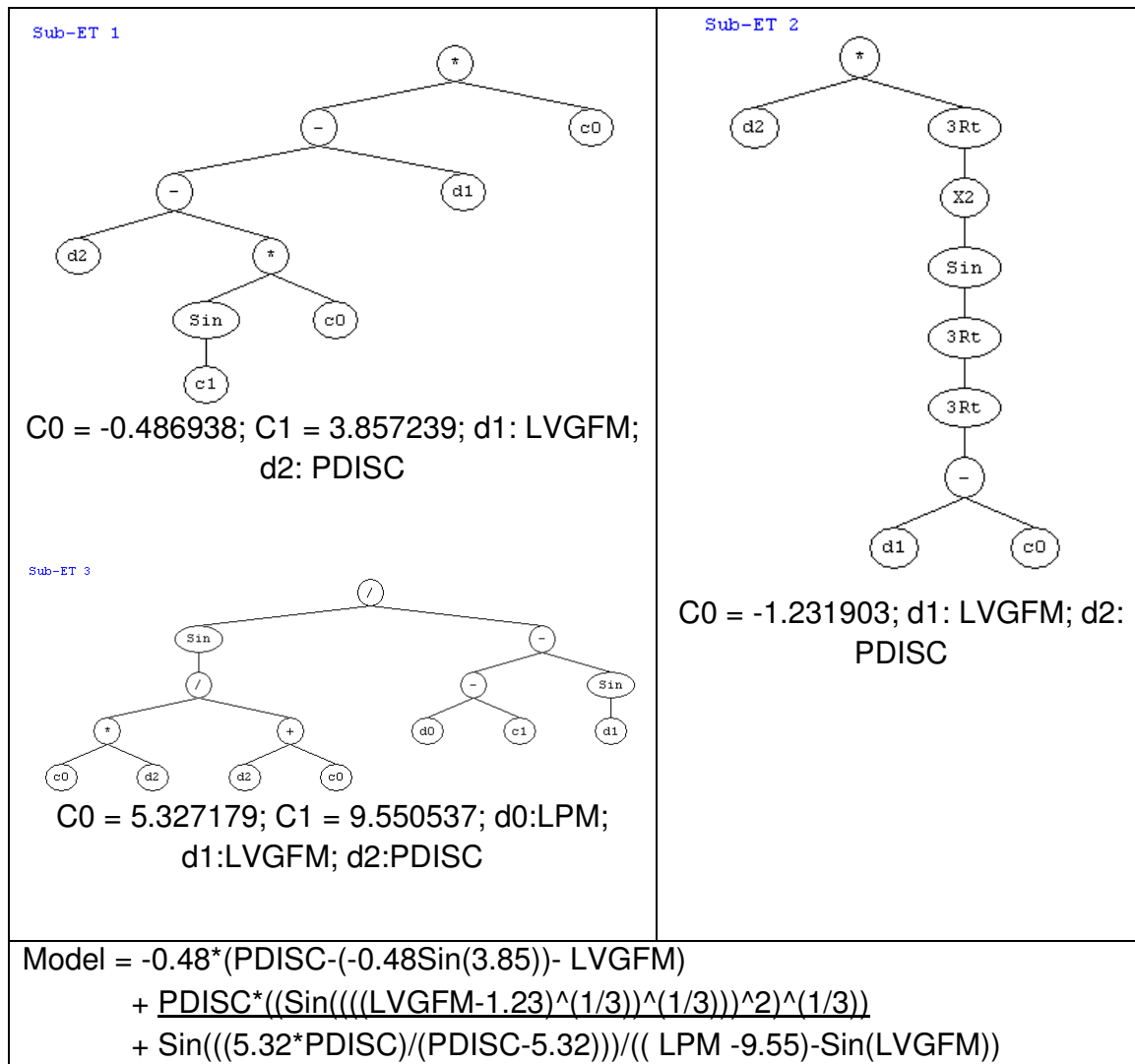
- Fernando, A. K., Shamseldin, A. Y., and Abrahart, R. J. "Using gene expression programming to develop a combined runoff estimate model from conventional rainfall-runoff model outputs." *Proc., 18th World IMACS Congress and MODSIM09 International Congress on Modelling and Simulation*, Modelling and Simulation Society of Australia and New Zealand and International Association for Mathematics and Computers in Simulation, 2377-2383.
- Ferreira, C. (2001). "Gene Expression Programming: A New Adaptive Algorithm for Solving Problems." *Complex Systems*, 13(2), 87-129
- Ferreira, C. (2006). *Gene Expression Programming: Mathematical Modeling by an Artificial Intelligence*, 2nd Edition, Springer-Verlag, Germany.
- Ferreira, C. (2009). ""What is GEP?" From GeneXproTools Tutorials - A Gepsoft Web Resource, <http://www.gepsoft.com/tutorial002.htm>." (10th February, 2009).
- Hsu, K. L., Moradkhani, H., and Sorooshian, S. (2009). "A sequential Bayesian approach for hydrologic model selection and prediction." *Water Resources Research*, 45 (W00B12).
- Jeong, D. I., and Kim, Y.-O. (2009). "Combining single-value streamflow forecasts – A review and guidelines for selecting techniques." *Journal of Hydrology* 377(3-4), 284–299.
- Kachroo, R. K. (1992). "River flow forecasting. Part 5. Application of a conceptual model." *Journal of Hydrology*, 133, 141-178.
- Moore, R. J. (1985). "The probability-distributed principle and runoff production at point and basin scales." *Hydrological Sciences Journal*, 30(2), 273-297.
- Nash, J. E., and Barsi, B. I. (1983). "A hybrid model for flow forecasting on large catchments." *Journal of Hydrology*, 65, 125-137.
- Nash, T., and Barsi, B. I. (1983). "A hybrid model for forecasting on large catchments." *Journal of Hydrology*, 65, 125-137.
- See, L., and Abrahart, R. J. (2001). "Multi-model data fusion for hydrological forecasting." *Computers and Geosciences*, 27(8), 987-994.
- See, L., and Openshaw, S. (2000). "A hybrid multi-model approach to river level forecasting." *Hydrological Sciences Journal*, 45(4), 523-536.
- Senbeta, D. A., Shamseldin, A. Y., and O'Connor, K. M. (1999). "Modification of the probability-distributed interacting storage capacity model." *Journal of Hydrology*, 224, 149-168.
- Shamseldin, A. Y. (1997). "Application of neural network technique to rainfall-runoff modelling." *Journal of Hydrology*, 199(3), 272-294.
- Shamseldin, A. Y., and O'Connor, K. M. (1999). "A Real-Time Combination Method for the Outputs of Different Rainfall-Runoff Models." *Hydrological Sciences Journal*, 44(6), 895-912.
- Shamseldin, A. Y., O'Connor, K. M., and Liang, G. C. (1997). "Methods for combining the outputs of different rainfall-runoff models." *Journal of Hydrology*, 197, 203-229.
- Shamseldin, A. Y., O'Connor, K. M., and Nasr, A. E. (2007). "A comparative study of three neural network forecast combination methods for simulated river flows of different rainfall-runoff models." *Hydrological Sciences*, 52(5), 896-916.
- Steeb, W.-H., Hardy, Y., and Stoop, R. (2005). *The nonlinear workbook*, World Scientific Publishing Company, Incorporated Singapore.
- Tan, B. Q., and O'Connor, K. M. (1996). "Application of an empirical infiltration equation in the SMAR conceptual model." *HJournal of Hydrology*, 185, 275-295.
- Viney, N. R., Bormann, H., Breuer, L., Bronstert, A., Croke, B. F. W., Frede, H. G., T., Hubrechts, L., Huisman, J. A., Jakeman, A. J., Kite, G. A., Lanini, J., Leavesley, G., Lettenmaier, D. P., Lindstrom, G., Seibert, J., Sivapalan, M., and Willems, P. (2009). "Assessing the impact of land use change on hydrology by ensemble modeling (LUCHEM). II: Ensemble combinations and predictions." *Advances in Water Resources*, 32, 147-158.
- Xiong, L., Shamseldin, A. Y., and O'Connor, K. M. (2001). "A non-linear combination of the forecasts of rainfall-runoff models by the first-order Takagi-Sugeno fuzzy system." *Journal of Hydrology* 245(1-4), 196-217.



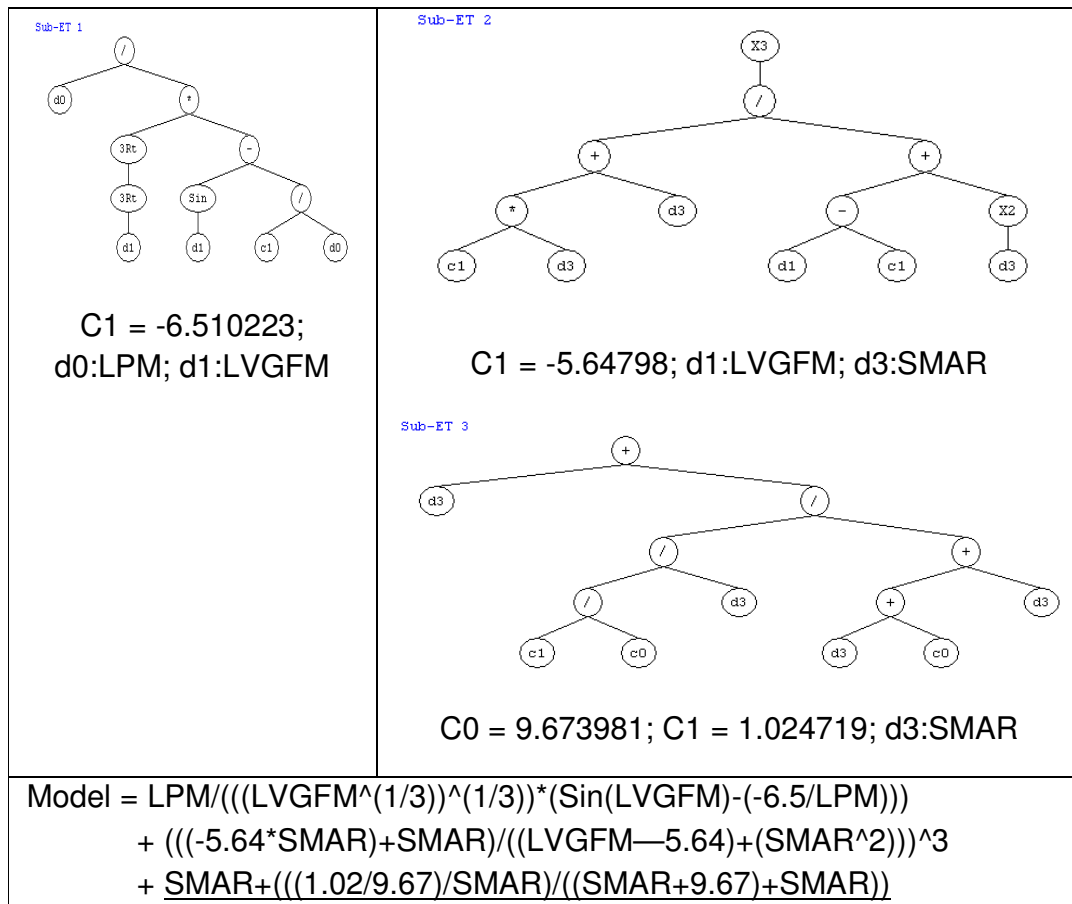
**Figure 1:** Expression tree



**Figure 2:** Expression tree after mutation

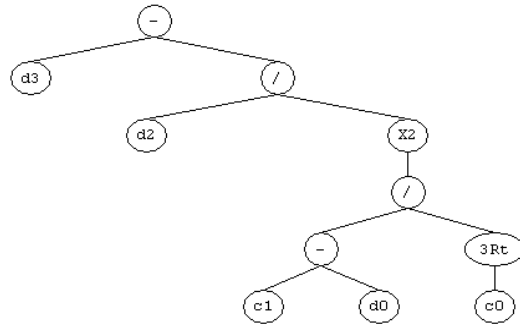


**Figure 3:** Expression Trees and model equation for combination model for Baihe River catchment (Primary Gene is Sub-ET 2, or G2 the expression for which is underlined.)



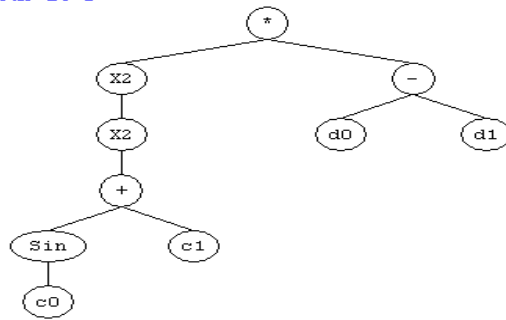
**Figure 4:** Expression Trees and model equation for combination model for Brosna River catchment (Primary Gene is Sub-ET3, or G3; expression underlined.)

Sub-ET 1



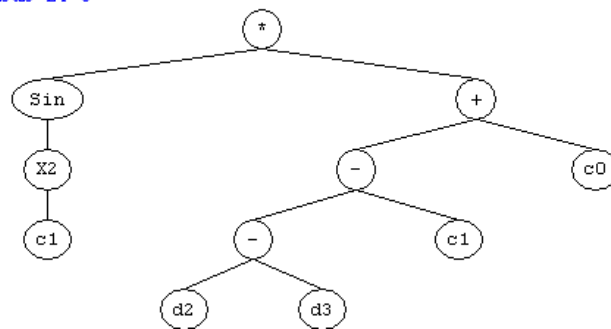
C0 = 5.791687; C1=-4.41333; d2:PDISC; d3:SMAR (Primary Gene)

Sub-ET 2



C0 = 2.715393; C1=0.198517; d0: LPM; d1:LVGFM

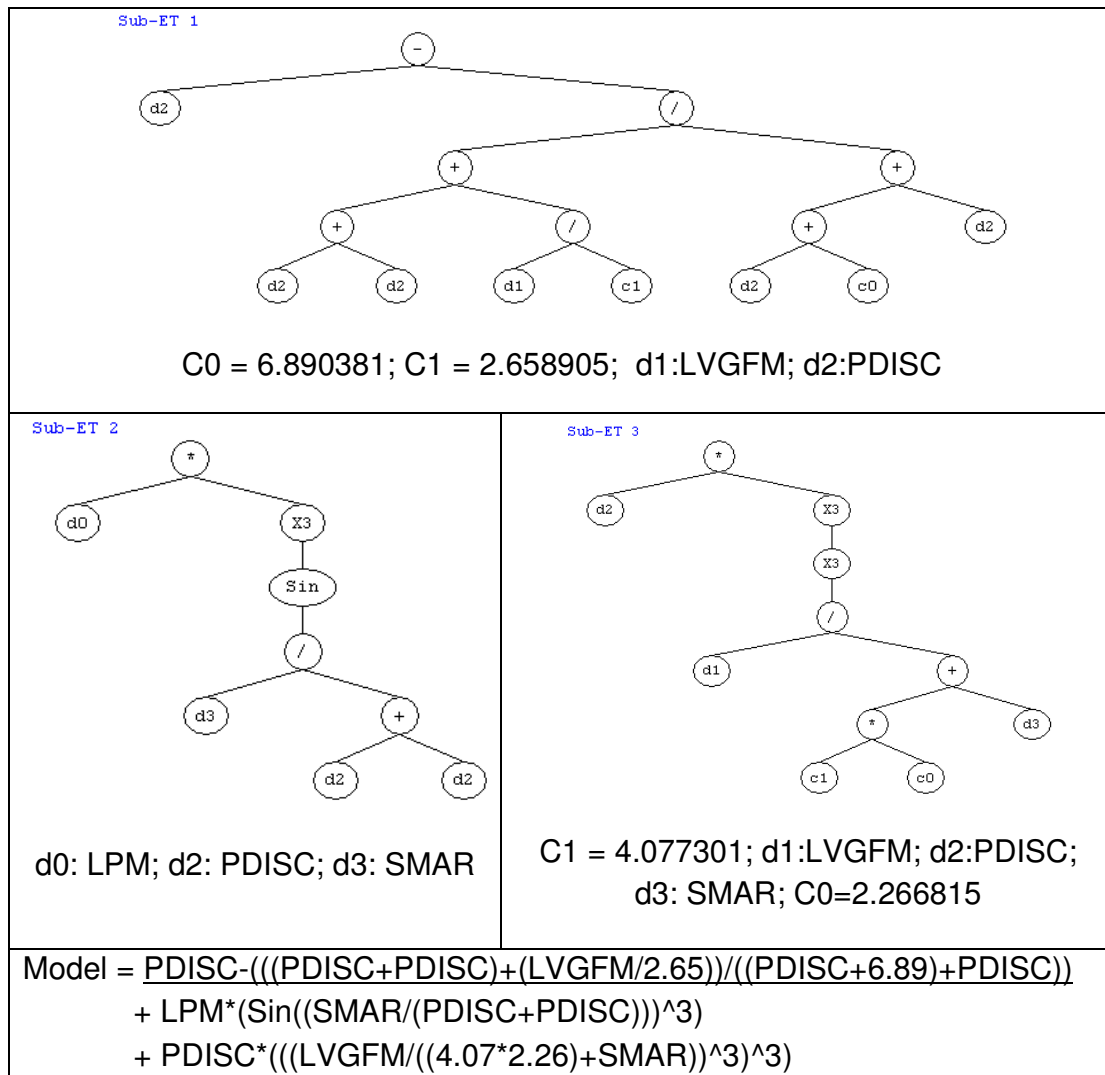
Sub-ET 3



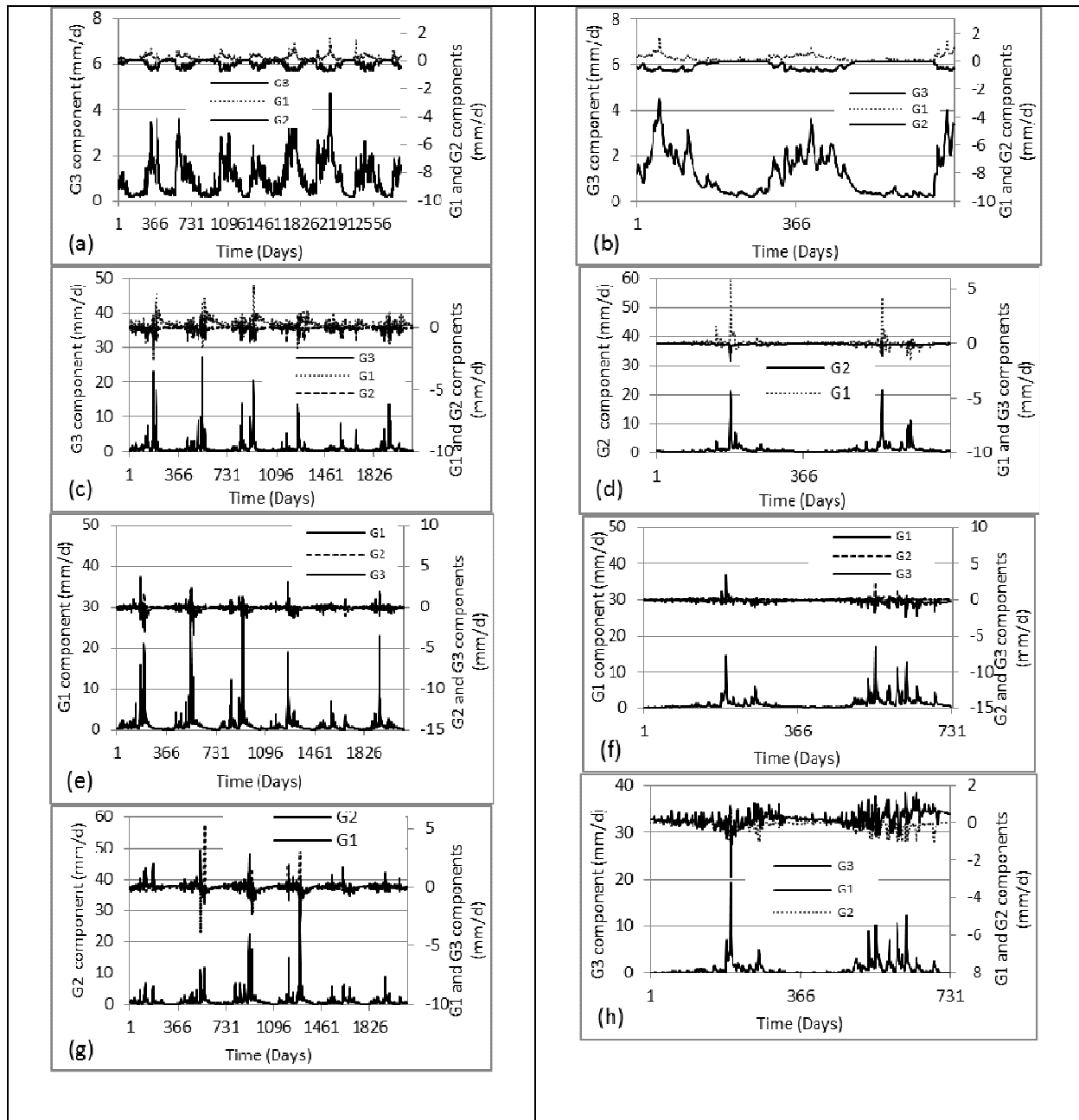
C0=6.213958; C1=6.182953;d2=PDISC; d3=SMAR

$$\text{Model} = \underline{\text{SMAR} - (\text{PDISC} / (((-4.41 - \text{LPM}) / (5.79^{(1/3)}))^2))} \\ + ((\text{Sin}(2.71 + 0.19)^2)^2) * (\text{LPM} - \text{LVGFM}) \\ + \text{Sin}((6.18^2) * (((\text{PDISC} - \text{SMAR}) - 6.18) + 6.21))$$

**Figure 5:** Expression Trees and model equation for combination model for Shiquan River catchment (Primary Gene is Sub-ET1, or G1; expression underlined.)

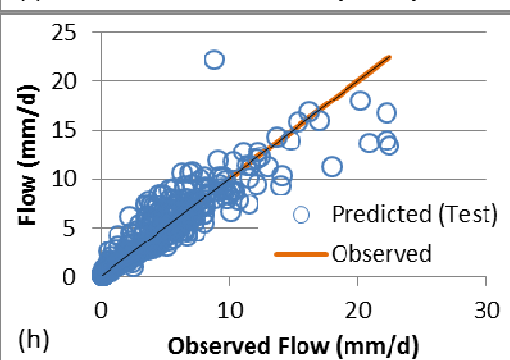
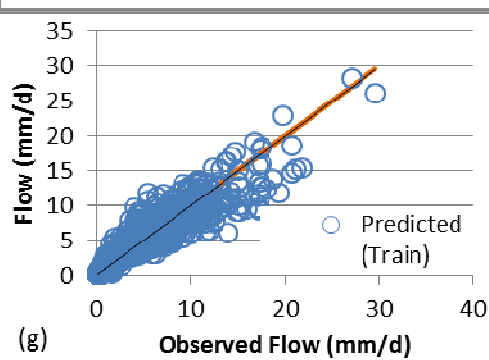
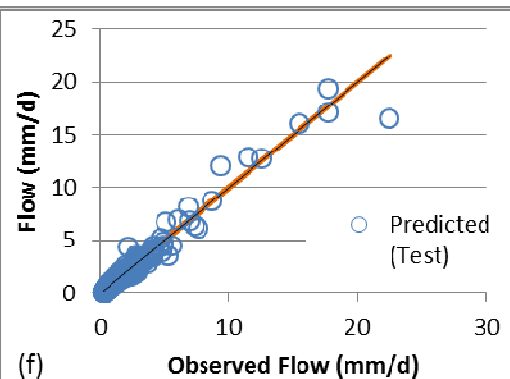
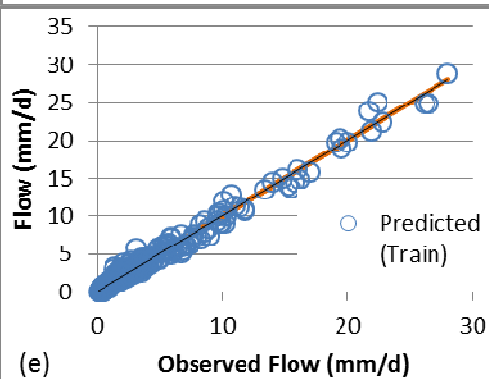
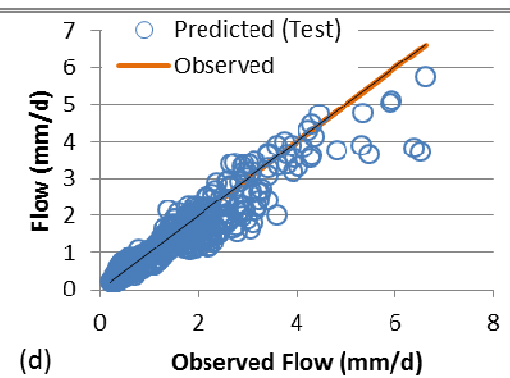
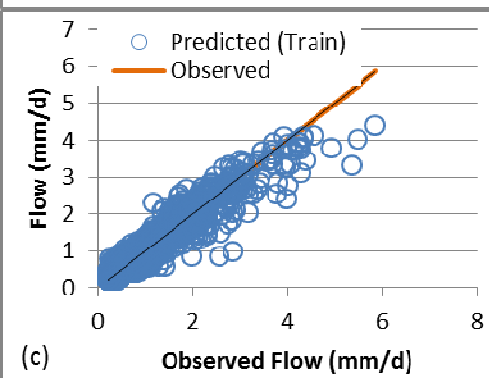
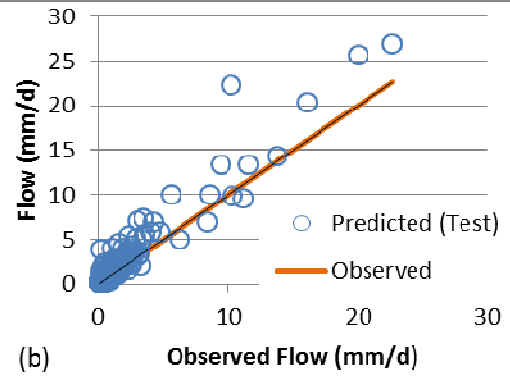
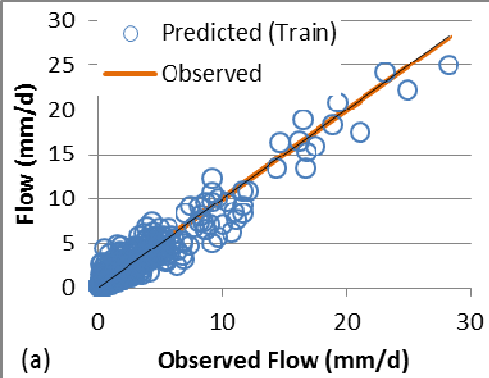


**Figure 6:** Expression Trees and model equation for combination model for Yanbian River catchment (Primary Gene is Sub-ET1, or G1;expression underlined.)

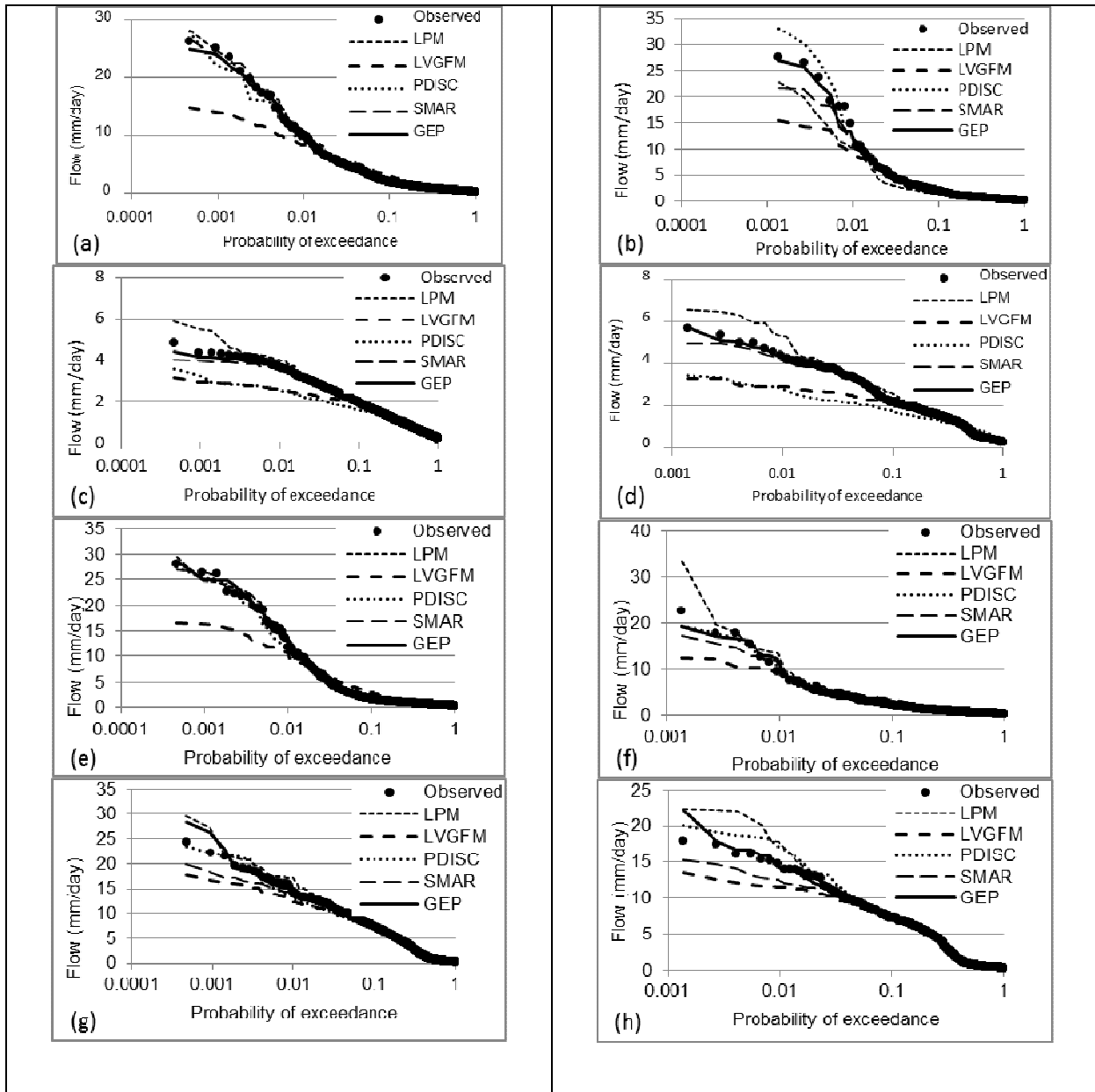


**Figure 7:** Components of the combined models (a) Bahie river Training set (b) Bahie river Testing set (c) Brosna river Training set (d) Brosna river Testing set (e) Shiquan river Training set (f) Shiquan river Testing set (g) Yanbian river Training set (h) Yanbian river Testing set





**Figure 8:** Predicted Vs. Observed flow (a) Bahie river Training set (b) Bahie river Testing set (c) Brosna river Training set (d) Brosna river Testing set (e) Shiquan river Training set (f) Shiquan river Testing set (g) Yanbian river Training set (h) Yanbian river Testing set



**Figure 9:** Flow duration curves (a) Bahie river Training set (b) Bahie river Testing set (c) Brosna river Training set (d) Brosna river Testing set (e) Shiquan river Training set (f) Shiquan river Testing set (g) Yanbian river Training set (h) Yanbian river Testing set

**Table 1.** Daily flow data used

Catchment	Country	Area (km <sup>2</sup> )	Climate	Daily flow record start	Training sample size	Testing set size
Baihe	China	61780	Semi-arid	01/01/1972	2117	730
Brosna	Ireland	1207	Temperate	01/01/1969	2832	730
Shiquan	China	3092	Semi-arid	01/01/1973	2116	730
Yanbian	China	2350	Humid	01/01/1978	2102	730

**Table 2.** Available inputs and functions for random selection

Function/Input	Symbol
Addition	+
Subtraction	-
Multiplication	*
Division	/
Square root	Sqrt
Exponential	Exp
Natural logarithm	Ln
x to the power of 2	x2
x to the power of 3	x3
Cube root	3Rt
Sine	Sin
LPM Model output as input	d0
LVGFM Model output as input	d1
PDISC Model output as input	d2
SMAR Model output as input	d3

**Table 3.** Parameters used for genes, genetic operators and assessment of fitness function

<b>Parameter</b>	<b>Value</b>	<b>Parameter</b>	<b>Value</b>
Number of chromosomes	30	Two point recombination	0.1
Head size	8	Gene recombination	0.3
Number of genes	3	Gene transposition	0.3
Fitness function	MSE	Numerical constants per gene	2
Mutation	0.044	Lower/Upper limits of constants	$\pm 10$
Inversion	0.1	RNC (Random Numerical constants) Mutation	0.01
IS (Insertion sequence) Transposition	0.1	Dc (additional gene domain) mutation	0.044
RIS (Root Insertion sequence) Transposition	0.1	Dc inversion	0.1
One point recombination	0.1	Dc Is transposition	0.1

**Table 4a.** Input variables chosen in combined model development

Catchment	Primary Gene	Models in primary component	Models in secondary components	Non-participating models
Baihe	G2	PDISC, LVGFM	PDISC, LVGFM, LPM	SMAR
Brosna	G3	SMAR	SMAR, LPM, LVGFM	PDISC
Shiquan	G1	PDISC, SMAR	LPM, LVGFM, PDISC, SMAR	-
Yanbian	G1	LVGFM, PDISC	SMAR, LVGFM, LPM, PDISC	-

**Table 4b.** Rank of individual models based on correlation coefficient (models used in the primary component of the combined model underlined)

Catchment	Rank 1	Rank 2	Rank 3	Rank 4
Baihe	SMAR (0.979)	<u>PDISC</u> (0.973)	LPM (0.961)	<u>LVGFM</u> (0.896)
Brosna	<u>SMAR</u> (0.965)	LPM(0.962)	LVGFM(0.870)	PDISC(0.664)
Shiquan	<u>SMAR</u> (0.970)	<u>PDISC</u> (0.967)	LPM(0.957)	<u>LVGFM</u> (0.886)
Yanbian	SMAR(0.989)	<u>LVGFM</u> (0.959)	LPM(0.949)	<u>PDISC</u> (0.935)

**Table 5:** Correlation coefficients between individual model results for training data (highest value in bold and dominant models in the primary component underlined)

Catchment	Model	LPM	LVGFM	PDISC	SMAR	GEP
Baihe	LPM	1	-	-	-	-
	<u>LVGFM</u>	0.862	1	-	-	-
	<u>PDISC</u>	0.936	0.923	1	-	-
	SMAR	0.940	0.898	0.935	1	-
	GEP	0.961	0.909	0.976	0.981	1
	Observed	0.961	0.896	0.973	0.979	<b>0.993</b>
Brosna	LPM	1	-	-	-	-
	LVGFM	0.839	1	-	-	-
	PDISC	0.643	0.722	1	-	-
	<u>SMAR</u>	0.928	0.851	0.683	1	-
	GEP	0.958	0.878	0.679	0.973	1
	Observed	0.962	0.870	0.664	0.965	<b>0.996</b>
Shiquan	LPM	1	-	-	-	-
	<u>LVGFM</u>	0.848	1	-	-	-
	<u>PDISC</u>	0.927	0.920	1	-	-
	<u>SMAR</u>	0.929	0.907	0.928	1	-
	GEP	0.826	0.898	0.974	0.980	1
	Observed	0.957	0.886	0.967	0.970	<b>0.993</b>
Yanbian	LPM	1	-	-	-	-
	LVGFM	0.911	1	-	-	-
	<u>PDISC</u>	0.888	0.940	1	-	-
	SMAR	0.939	0.950	0.920	1	-
	GEP	0.947	0.961	0.941	0.990	1
	Observed	0.949	0.959	0.935	0.989	<b>0.994</b>

**Table 6.** Summary statistics for model performance (best in bold)

	LPM	LVGFM	PDISC	SMAR	GEP	Observed
<b>Bahie (Training set)</b>						
Mean	1.07	1.06	0.94	1.14	1.07	1.07
Peak	28.25	14.77	27.94	26.67	25.03	26.33
RMSE	0.550	0.847	0.459	0.409	<b>0.226</b>	0
RMSE as a % Observed mean	51.6	79.5	43.0	38.4	<b>21.2</b>	0
RMSE as a % Observed peak	2.1	3.2	1.7	1.6	<b>0.9</b>	0
R <sup>2</sup> Value	0.9608	0.8958	0.9735	0.9787	<b>0.9930</b>	1
<b>Bahie (Testing set)</b>						
Mean	0.78	0.98	0.93	1.11	1.05	1.07
Peak	22.66	15.54	33.02	21.62	26.98	27.62
RMSE	0.905	1.075	0.502	0.540	<b>0.284</b>	0
RMSE as a % Observed mean	84.7	100.6	47.0	50.6	<b>26.6</b>	0
RMSE as a % Observed peak	3.3	3.9	1.8	2.0	<b>1.0</b>	0
R <sup>2</sup> Value	0.9552	0.9044	0.9844	0.9769	<b>0.9936</b>	1
<b>Brosna (Training set)</b>						
Mean	0.96	0.97	0.96	0.99	0.96	0.95
Peak	5.86	3.16	3.53	4.01	4.39	4.82
RMSE	0.212	0.357	0.529	0.205	<b>0.072</b>	0
RMSE as a % Observed mean	22.3	37.5	55.7	21.6	<b>7.6</b>	0
RMSE as a % Observed peak	4.4	7.4	11.0	4.3	<b>1.5</b>	0
R <sup>2</sup> Value	0.9619	0.8704	0.6645	0.9646	<b>0.9955</b>	1
<b>Brosna (Testing set)</b>						
Mean	1.22	1.06	1.04	1.19	1.12	1.12
Peak	6.62	3.30	3.43	5.00	5.75	5.71
RMSE	0.319	0.403	0.647	0.239	<b>0.087</b>	0
RMSE as a % Observed mean	28.4	35.9	57.7	21.3	<b>7.8</b>	0
RMSE as a % Observed peak	5.6	7.1	11.3	4.2	<b>1.5</b>	0
R <sup>2</sup> Value	0.9596	0.9124	0.7436	0.9694	<b>0.9957</b>	1
<b>Shichuan (Training set)</b>						
Mean	1.11	1.10	0.93	1.25	1.04	1.11
Peak	29.51	16.61	27.40	27.03	28.81	28.03
RMSE	0.670	1.023	0.594	0.572	<b>0.278</b>	0
RMSE as a % Observed mean	60.4	92.2	53.6	51.6	<b>25.1</b>	0
RMSE as a % Observed peak	2.4	3.7	2.1	2.0	<b>1.0</b>	0
R <sup>2</sup> Value	0.9570	0.8861	0.9672	0.9698	<b>0.9927</b>	1
<b>Shichuan (Testing set)</b>						
Mean	1.06	1.09	0.88	1.23	1.00	1.07
Peak	33.46	12.30	19.37	17.19	19.36	22.47
RMSE	0.960	0.923	0.527	0.572	<b>0.363</b>	0
RMSE as a % Observed mean	89.5	86.1	49.2	53.3	<b>33.9</b>	0
RMSE as a % Observed peak	4.3	4.1	2.3	2.5	<b>1.6</b>	0
R <sup>2</sup> Value	0.8901	0.8619	0.9634	0.9528	<b>0.9801</b>	1
<b>Yanbian (Training set)</b>						
Mean	2.64	2.64	2.39	2.64	2.64	2.64
Peak	29.56	17.78	23.60	19.85	28.25	24.31
RMSE	1.111	0.943	1.225	0.530	<b>0.367</b>	0
RMSE as a % Observed mean	42.1	35.7	46.4	20.1	<b>13.9</b>	0
RMSE as a % Observed peak	4.6	3.9	5.0	2.2	<b>1.5</b>	0
R <sup>2</sup> Value	0.9488	0.9592	0.9348	0.9892	<b>0.9939</b>	1
<b>Yanbian (Testing set)</b>						

Mean	2.64	2.62	2.43	2.63	2.65	2.65
Peak	22.43	13.58	20.09	15.30	22.22	17.90
RMSE	1.170	1.041	1.195	0.655	<b>0.477</b>	0
RMSE as a % Observed mean	44.2	39.3	45.1	24.7	<b>18.0</b>	0
RMSE as a % Observed peak	6.5	5.8	6.7	3.7	<b>2.7</b>	0
R <sup>2</sup> Value	0.9423	0.9515	0.9475	0.9838	<b>0.9896</b>	1