

The applicability of Generic Self-Evolving Takagi-Sugeno-Kang neuro-fuzzy model in modeling rainfall-runoff and river routing

Mohammad Ashrafi, Lloyd H. C. Chua and Chai Quek

ABSTRACT

Recent advancements in neuro-fuzzy models (NFMs) have made possible the implementation of dynamic rule base systems. This is in comparison with static applications commonly seen in global NFMs such as the Adaptive-Network-Based Fuzzy Inference System (ANFIS) model widely used in hydrological modeling. This study underlines key differences between local and global NFMs with an emphasis on rule base dynamics, in the context of two common flow forecast applications. A global NFM, ANFIS, and two local NFMs, Dynamic Evolving Neural-Fuzzy Inference System (DENFIS) and Generic Self-Evolving Takagi-Sugeno-Kang (GSETSK), were tested. Results from all NFMs compared favorably when benchmarked against physically based models. Rainfall-runoff modeling is a complex process which benefits from the advanced rule generation and pruning mechanisms in GSETSK, resulting in a more compact rule base. Although ANFIS resulted in the same number of rules, this came about at the expense of having the need for a large training dataset. All NFMs generated a similar number of rules for the river routing application, although local NFMs yielded better results for forecasts at longer lead times. This is attributed to the fact that the routing procedure is less complex and can be adequately modeled by static NFMs.

Key words | GSETSK, local learning, neuro-fuzzy, rainfall-runoff, river routing, rule base

Mohammad Ashrafi (corresponding author)

Lloyd H. C. Chua
School of Civil and Environmental Engineering,
Nanyang Technological University,
50 Nanyang Avenue, Singapore 639798,
Singapore
E-mail: mohammad021@e.ntu.edu.sg

Lloyd H. C. Chua
School of Engineering, Faculty of Science
Engineering & Built Environment,
Deakin University,
75 Pigdons Road, Waurn Ponds, VIC 3220,
Australia

Chai Quek
Computational Intelligence Laboratory, School of
Computer Science and Engineering,
Nanyang Technological University,
50 Nanyang Avenue, Block N4-02A-32, Singapore
639798,
Singapore

INTRODUCTION

The neuro-fuzzy model (NFM) is a data-driven approach that has become popular in flow forecasting. In an NFM, a rule base which consists of IF-THEN statements maps the input to the output space. The structure of the NFM is built based on this rule base which is determined by a clustering approach. In the first-order Takagi-Sugeno-Kang (TSK) approach, the consequent part of the conditional IF-THEN statement consists of a linear combination of all input variables. The parameters associated with the antecedent and consequent parts of the rule base are learnt using existing or training data.

Earlier applications of NFM to flood forecasting adopted the Adaptive-Network-Based Fuzzy Inference System or ANFIS (Jang 1993). These studies used data

from various parts of the world including India (Nayak *et al.* 2007; Mukerji *et al.* 2009), China (Chau *et al.* 2005; Wang *et al.* 2009), USA (Nayak *et al.* 2007), Indonesia (Aqil *et al.* 2007), Japan (Chidthong *et al.* 2009), UK (Remesan *et al.* 2008), Taiwan (Chang & Chen 2001), Lao PDR (Nguyen & Chua 2012), and Iran (Ghalkhani *et al.* 2013). It can be said that almost all applications used the global learning approach or batch training in ANFIS. More significantly, it is important to note that the rule base in ANFIS is fixed. Unlike ANFIS, modern NFMs, however, incorporate a flexible or dynamic model structure and hence rule base. The Dynamic Evolving Neural-Fuzzy Inference System or DENFIS (Kasabov & Song 2002) employs a local learning technique, allowing the model to respond to

changes through a modification of the rule base in an incremental fashion. Hong and White (2009) introduced the Dynamic Neuro-Fuzzy Local Modeling System or DNFLMS which is an online TSK NFM for flow forecasting.

One difficulty that has recently arisen in adopting local learning in NFM is that the model is able to adapt its rule base, which is done by adding new rules; however, the reverse is not possible, i.e. rules that are outdated are not removed. This means that as the model evolves with time, the rule base expands, and the model may become too complex leading to deterioration in interpretability and/or performance. In addition, current NFMs such as DENFIS (Eray et al. 2017) require information on the upper and lower bound of historical data. This is not always available, or even if it is available, the model will not be applicable if the test data range exceeds the historical upper/lower bounds. Recently, models such as the Generic Self-Evolving Takagi-Sugeno-Kang or GSETSK (Ashrafi et al. 2017) have been reported. GSETSK uses the Multidimensional-Scaling Growing Clustering (MSGC) method and the incorporation of Hebbian-based rule pruning algorithm (Ashrafi et al. 2017) with MSGC results in a dynamic NFM which can expand or shrink its rule base. In addition, MSGC does not need any prior knowledge of data which would be vital for applications where data are limited.

Given these advances in NFMs including the introduction of local learning, clustering and pruning mechanisms, there is now a need to critically assess these models. Therefore, the objectives of this paper were (i) to assess local versus global learning strategies in NFMs and (ii) to study the importance of rule base dynamics in relation to common hydrologic/hydraulic forecast applications. These objectives were studied in the context of rainfall–runoff and flow routing applications and results are compared against physically based and statistical models used as baselines. The rainfall–runoff application used precipitation, runoff, and temperature data for Klippan_2 Basin in Sweden and the river routing application used river stage data from three gauging stations in the Lower Mekong River. The novelty of this paper is the application of a state-of-the-art dynamic NFM, GSETSK. We compare GSETSK with current dynamic NFM approaches found in the literature highlighting its use and applicability to

rainfall–runoff and flow routing applications via an in-depth analysis of rule base dynamics.

STUDY AREA AND DATA

Two datasets were used in this study. The first dataset belongs to Klippan_2 Basin (area = 242.9 km²) located in Southern Sweden (Talei et al. 2013) which was used for the rainfall–runoff study. The main variables used for runoff forecasting in this basin are river discharge (Q), precipitation (P), and temperature (T). Precipitation including both rainfall and snow, and temperature was used as a surrogate for snowmelt. Thirty years of daily data available from the Swedish Meteorological and Hydrological Institute website (SMHI 2013) was used. This contains 10,947 data samples, ranging from 11 January 1961 to 31 December 1990. The second dataset from the Lower Mekong River in Southeast Asia (Nguyen & Chua 2012) was used as a river routing application. The water level at two upstream gauging stations Thakhek and Savannakhet (SV) were used to predict the stage at Pakse (PK) in this study. The data used in this study consist of 9 years (1994–1999 and 2009–2011) of data obtained for the flood season which occurs from June to October each year (MRC 2015).

METHODOLOGY

Neuro-fuzzy models

NFMs such as ANFIS, DENFIS, and GSETSK have a common six-layer architecture. The functions of these layers can generally be defined as follows. (i) *Input Layer* – this layer receives crisp inputs. (ii) *Fuzzification Layer* – fuzzification of the crisp values is done in this layer, usually adopting the Gaussian membership function (GMF). (iii) *Rule Layer* – the *firing strength* of each rule is computed in every node of this layer. (iv) *Normalization Layer* – the *normalized firing strength* is computed in this layer. (v) *Consequent Layer* – the contribution of each rule in the NFM final result is reflected in each consequent node in the Consequent Layer. (vi) *Summation Layer* – this layer sums up the result of each rule to provide the final

output. The ANFIS model is essentially a global model and has a fixed structure, but both DENFIS and GSETSK have a local learning approach which enables these models to have a dynamic structure. The results of a global model are more biased towards the average behavior of the system while local methods weigh the model towards current trends. Local and global models differ in the methodologies used to partition the input space. In a global model, the entire training dataset is used, whereas for a local model, an incremental partitioning method is adopted. Common global clustering approaches adopted in ANFIS (Jang 1993) are grid partitioning and subtractive clustering or SC (Chiu 1994), of which the latter is probably more well-known. The Evolving Clustering Method or ECM (Kasabov & Song 2002) is a local clustering method used in DENFIS. This method incrementally tracks changes in the input space. The ECM is initialized by adopting the location of the first data point in the unit hyperbox as the first cluster center with zero cluster radius. With the arrival of subsequent data, the Euclidean distance is calculated to determine the distance between the new data point with an existing cluster and depending on this distance, the algorithm decides if a new cluster should be formed or an existing cluster should be adapted. This process of cluster adaptation and/or creation continues as each data point is presented one at a time. The ECM allows for new cluster centers to be incrementally created and thereby increase the rule base. This is clearly an advantage over ANFIS which has a fixed rule base. However, ECM does not have the capability of removing cluster centers that may have become obsolete. GSETSK has the capability of removing outdated rules through a process called rule pruning. Rule pruning in GSETSK is based on and the MSGC algorithm, which is an online clustering method based on Hebbian learning (Ashrafi et al. 2017). Another important distinction between MSGC and SC or ECM is that both SC and ECM work on a unit hyperbox, which is created by data normalization. This means that prior data of min/max values are required. In MSGC, however, data are projected onto every single input space with the arrival of each data point. The data are thus evaluated in a one-dimensional space rather than multidimensional space, and thus preclude the need for historical data since data normalization is not required.

Evaluation indices

Three types of evaluation indices were used to evaluate the modeling results:

1. *Goodness-of-fit indices*, including the root mean square error (RMSE), mean absolute error (MAE), the coefficient of efficiency (Fortin et al. 1997) and coefficient of determination (R^2).
2. *Threshold-based indices* (Sene 2008) that compare against benchmarks, include CSI, POD and FAR:

$$CSI = TA/(TA + FA + FR) \quad (1)$$

$$POD = TA/(TA + FR) \quad (2)$$

$$FAR = FA/(TA + FA) \quad (3)$$

where TA or true acceptance is the number of data points greater than the threshold, FA or false acceptance is the number of measured data less than the threshold but were forecasted to be greater, and FR or false rejection is the number of measured data greater than the threshold but were forecasted to be less. For CSI and POD, values closer to 1 are desirable while FAR values closer to 0 are desired.

3. *Minimum size of rule base* is desired where an NFM is able to achieve the best results with the smallest rule base (Ashrafi et al. 2017).

Benchmark models

Four models used as benchmarks including a linear regression model, a nonlinear autoregressive model, and two physically based models. A brief description of these models is presented here, details can be found in the references provided.

1. Stepwise regression or SR (Draper & Smith 2014) is a linear model built by a systematic approach of fitting a regression model by adding or removing the predictive variable based on their statistical significance in regression.
2. Nonlinear autoregressive with external (exogenous) input or NARX (Remesan et al. 2008):

$$y_t = f(y(t-1), y(t-2), \dots, y(t-n_y), u(t-1), u(t-2), \dots, u(t-n_u)) \quad (4)$$

where y and u are the variable of interest and externally determined variable respectively.

- Hydrologiska Byråns Vattenbalansavdelning or HBV (Bergstrom & Forsman 1973) is a conceptual rainfall–runoff model:

$$Q = P - E - \Delta(S^* + SM + UZ + LZ) \quad (5)$$

where Q , P , and E indicate water fluxes for discharge, precipitation, and evaporation respectively. S^* and SM are snow and soil moisture water storage components and UZ and LZ are groundwater storages.

- Unified Runoff Basin Simulation or URBS (Carrol 2007) is a physically based modeling approach which uses a network of conceptual storages to represent the stream networks and reservoirs. A rainfall–runoff model converts gross rainfall to rainfall excess in the catchment and a runoff routing model converts the excess rainfall to flow.

Model inputs

A preprocessing procedure was conducted to determine the model inputs via a two-step procedure: (i) select the independent variables for dimensionality reduction and (ii) adopt a 10-fold cross-validation method (Kohavi 1995) to ascertain the best combinations of the independent inputs to be used for the models. The Monte Carlo Evaluative Selection (MCES) method (Quah & Quek 2006) was adopted to select the independent inputs. In this method, irrelevant and correlated inputs are identified based on weight ranking. Input variables which are redundant (irrelevant and correlated) are assigned negative weights, while the recommended inputs are weighted positively. Details of MCES can be found in Quah and Quek (2006). For the Klippan_2 dataset, the 1-day ahead forecast of discharge, Q_1 , is based on a selection of the following possible inputs:

$$Q_1 = f(Q_i, P_i, T_i), \quad i = -9, -8, \dots, -1, 0 \quad (6)$$

where the subscript i refers to time (days), and $i=0$ denotes the current time. For the routing application for the Lower Mekong River, the water level at Pakse Station

1, 3, and 5 days ahead, PK_j , is estimated based on:

$$PK_j = f(TK_i, SV_i, PK_i), \quad j = 1, 3, 5 \quad (7)$$

$$i = -7, -6, \dots, -1, 0$$

where $j=1, 3, 5$ days is the forecast lead time and $i = \text{time (day)}$, and $i=0$ denotes the current time.

In the second step of the input selection procedure, combinations of the independent inputs selected by MCES were generated and for each combination, a 10-fold cross-validation evaluation method was adopted to identify the best input–output combination set to be eventually adopted by the model. The cross-validation procedure first divides the entire dataset into 10 equally sized parts or folds, with nine folds adopted as the training dataset and the remaining fold for testing. This procedure was repeated 10 times, each time choosing a new training/testing dataset combination. Error statistics were computed for each of 10 testing dataset, and then they were averaged. Finally, the input combination that had the least error was selected to be the inputs to the model.

Modeling procedure

The GSETSK, DENFIS, and ANFIS models were run in the following manner:

- ANFIS is a global model and requires separate training and testing datasets (Table 1). The training data were used for rule base generation and optimization in ANFIS before being used on the testing data.
- As GSETSK is an online model, the model does not require separate training and testing datasets (Table 1). The first input–output data sample was for model initialization and predictions started from the second data point onwards. During the prediction stage, GSETSK first makes a forecast, after which the forecast value is compared against the measured data, once it becomes available. Based on a comparison between the predicted and actual values, the model is then updated. This process continues for the entire dataset with the arrival of each data point.
- The incremental version of the DENFIS was applied which allowed the model to be run as an online model, similar to GSETSK. The limitation of DENFIS, however,

Table 1 | Data used by different NFMs

NFM	Data function	Rainfall-runoff (Klippan_2)			River routing (Mekong River)		
		From	To	No. of data samples	From	To	No. of data samples
ANFIS	Training	11 January 1961	31 December 1979	6,929	1994	1999	918
	Testing	1 January 1980	31 December 1990	4,018	2009	2011	413
DENFIS	Model initialization	11 January 1961	10 January 1962	365	1994		153
	Continuous testing/training	11 January 1962	31 December 1990	10,582	1995–1999, 2009–2011		1,178
GSETSK	Model initialization	11 January 1961		1	1-June-1994		1
	Continuous testing/training	12 January 1961	31 December 1990	10,946	1994–1999, 2009–2011		1,330

is that the model requires training data for normalization and to initialize the rule base. Once the rule base is initialized, the dynamic feature of DENFIS enables more rules to be added to the rule base during the model run. The period of data used in the training and testing phases of the DENFIS model are shown in Table 1. Parameter values ranging from 0.1–0.3 (0.01 interval) were tested for D_{thr} in DENFIS and 0.1–0.5 (0.05 interval) were tested for the radii parameter in ANFIS. The optimal values achieved for rainfall-runoff were $D_{thr} = 0.23$ and radii = 0.1 and $D_{thr} = 0.15$ and radii = 0.2 for river routing.

The results from the NFMs were compared against available results obtained from HBV (Talei et al. 2013) and URBS (Nguyen & Chua 2012). The HBV model (Swedish dataset) results are available from 1 January 1980 to 31 December 1985 and the URBS model (Lower Mekong River) results are available for the monsoon seasons (June to October) of 2010–2011, for comparison. Also, SR and NARX models were built and tested using data concurrent with ANFIS training and testing period (Table 1), respectively.

RESULTS AND DISCUSSION

Rainfall-runoff modeling

For the rainfall-runoff application, nine variables returned positive weights (out of a total of 31 variables, Equation

(6)) when analyzed with MCES. Based on this analysis, six variables, Q_0 , Q_{-2} , Q_{-4} , P_0 , T_{-1} , P_{-1} , which had significant positive weightings were selected for cross-validation analysis. A total of 16 combinations of these six variables tested with the cross-validation method suggested that the following combination of inputs be chosen to forecast Q_1 :

$$Q_1 = f(Q_0, P_0, T_{-1}) \quad (8)$$

Figure 1 shows the results for 1 January 1980 to 31 December 1985 (the part of the test period which coincides with the period where HBV results are available) of forecasting with NFM models benchmarked against the physically based HBV model obtained from Talei et al. (2013) SR and NARX models. The results show that all three NFM models, SR and NARX models provide better results when compared against HBV with achieving CE ranging from 0.88 to 0.90, and RMSE from 1.19 to 1.38 m³/s, against CE = 0.80 and RMSE = 1.8 m³/s for the HBV model. Therefore, in terms of the NFMs being able to predict overall fits, these results are acceptable. Results from the threshold-based indices, adopting a threshold of 15 m³/s (Talei et al. 2013) are illustrated in Figure 1(b). The results show better performance of NFMs over SR, NARX and HBV models. As obvious from the figure, NFMs are achieving CSI from 0.56 to 0.57 which is at least 6% and 9% higher than CSI of SR and NARX models, respectively. Among the NMFs, GSETSK fared the best, with FAR 5% and 17% lower than SR and NARX models, respectively. This means that using GSETSK as a flood forecasting model is more reliable

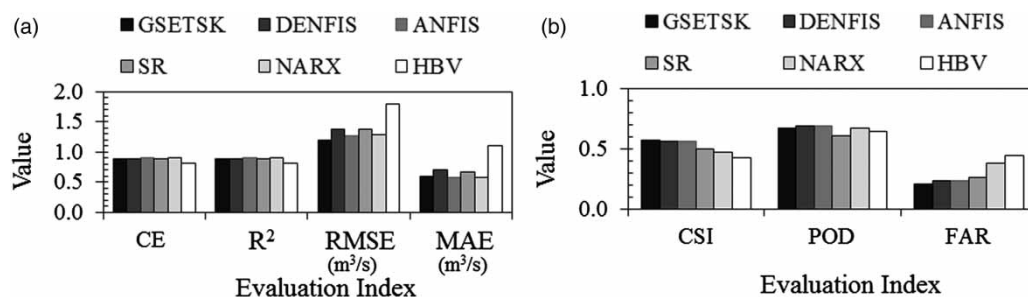


Figure 1 | Evaluation indices for different models for rainfall–runoff application. (a) Goodness-of-fit indices and (b) threshold-based indices.

compared to these models. An analysis of the relative merits of each of the NFM in rainfall–runoff modeling in terms of its rule base is provided later in discussions related to Figure 3.

River routing modeling

A similar analysis with MCES was carried out for the Lower Mekong River data. Following MCES, cross-validation was used to determine the best input–output combinations for different combinations of inputs from which the following results for 1, 3, and 5 step ahead forecasts for Pakse were obtained:

$$PK_1 = f(PK_0, SV_{-1}, SV_{-3}, TK_{-3}) \quad (9)$$

$$PK_3 = f(PK_{-2}, PK_{-3}, PK_{-4}, SV_{-3}, SV_{-5}) \quad (10)$$

$$PK_5 = f(PK_{-4}, PK_{-5}, PK_{-6}, SV_{-5}, SV_{-7}) \quad (11)$$

The time lags associated with the data at Thakhek and Savannakhet (SV) in Equations (9)–(11), referenced from Pakse are consistent with the travel times for a flood wave between these stations.

Figure 2 shows the results (computed over the entire test period, Table 1) for 1, 3 and 5 days ahead water level forecasting at Pakse. The results of 1-day ahead forecasting (Figure 2(a)) shows almost the same performance for all models in terms of general goodness-of-fit indices. However, in terms of threshold-based indices DENFIS has the best performance (CSI = 0.93, FAR = 0.03, POD = 0.95) and URBS worst (CSI = 0.84, FAR = 0.12, POD = 0.95). GSETSK, SR, and NARX have almost the same results, all of which performing better than ANFIS. For 3-day ahead

forecasting (Figure 2(b)), again all of the models have almost the same performance in terms of general goodness-of-fit indices. However, in terms of threshold-based indices, DENFIS and SR (CSI = 0.77, FAR = 0.08–0.11, POD = 0.83–0.85) have the best performance followed by GSETSK, ANFIS, and NARX, with URBS performing the worst (CSI = 0.55, FAR = 0.28, POD = 0.70). For 5-day ahead forecasting (Figure 3(c)), almost all the models have the same performance, and only slightly improved over NARX in terms of general goodness-of-fit indices. However, DENFIS and then GSETSK have the best performances in terms of threshold-based indices. This is followed by SR and ANFIS, followed by NARX and URBS which have the worst results. An analysis of the relative merits of each of the NFM in river routing in terms of its rule base is provided later in discussions related to Figure 4.

In general, the NFM, SR, and NARX are able to provide similar or better results compared to the URBS model, which requires more data and time to set up. However, using 11 m as the threshold (MRC 2015), the threshold-based indices indicate better results were achieved by the local NFM (DENFIS and GSETSK) over the global NFM (ANFIS), SR, NARX, and URBS models. These comparisons show that NFM and local NFM, in particular, are able to provide results that are at least comparable to the benchmarks. Most notably, both the local NFM require less data to train (Table 1) compared to the ANFIS, SR, NARX and physically based models. In addition, the local learning feature of DENFIS and GSETSK makes these two models most suitable as real-time models as they can be adapted continuously (online) without the need for re-calibration (as in the case of physically based models) or re-trained/refitted (in case of global NFM and NARX/SR).

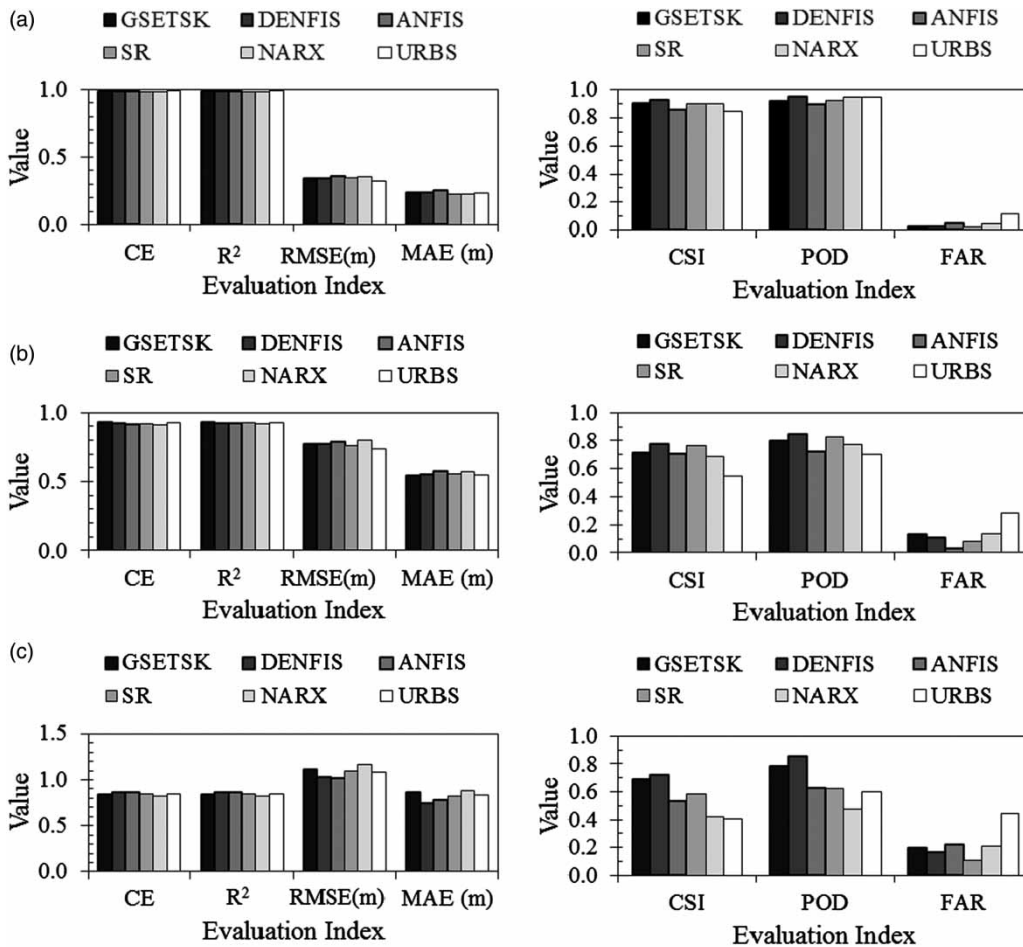


Figure 2 | Evaluation indices for river routing application. (a) 1-day ahead, (b) 3-day ahead, (c) 5-day ahead.

Differences in rule base dynamics between rainfall-runoff and river routing modeling

Figure 3(a) shows the complexity of the runoff time series owing to the complex process of transforming precipitation into runoff from a catchment. In addition, seasonal effects are present since the entire year's data are used. The rule base dynamics for the rainfall-runoff problem is illustrated in Figure 3(b). The rule base for ANFIS is static, with the number of rules fixed at nine rules, established during training using the SC clustering algorithm. Neither the number of rules nor their parameters change during the test phase. Even for such a complex process as in rainfall-runoff modeling, ANFIS assumes that the same rules can be applied after training. For the DENFIS model, 1 year of data was

used for model training (11 January 1961 to 10 January 1962) which created an initial rule base of eight rules. During the test phase (11 January 1962 to 31 December 1990), the number and parameters associated with the rules increased in response to the incoming data, increasing to 20 rules at the end of 1990. It is further observed that most of the new rules were generated during the first 8 years (roughly 1962–1970) as the model encountered new data and was further adapted online. After this period, however, the number of rules did not increase further until December 1980, October 1985 and September 1986 where one more rule was added each of these times. After this time, the model had a total of 20 rules with no more rules being added until the end of the test period. The rule base in DENFIS thus adapts online, generating new rules

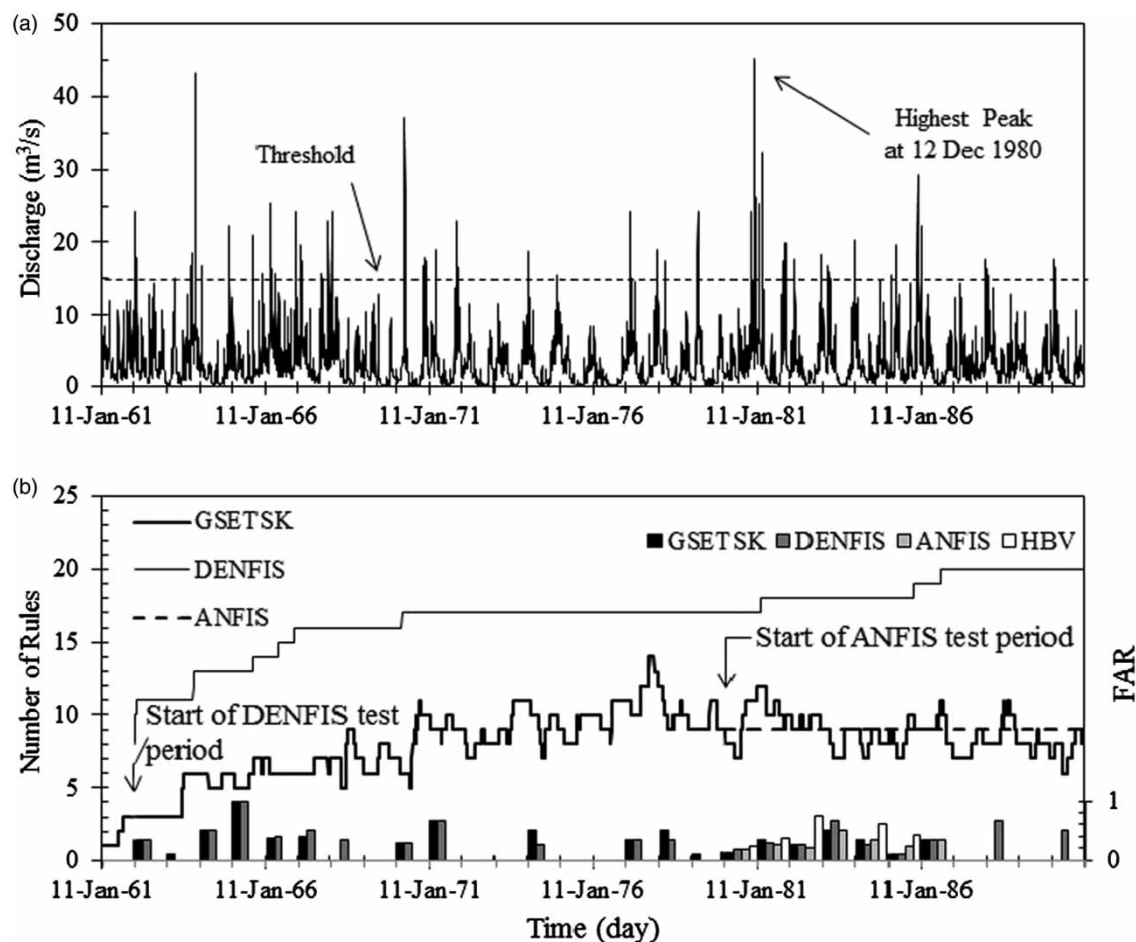


Figure 3 | (a) Time series of discharge and (b) change in the number of rules and yearly FAR during the model operation for different NFMs in the rainfall-runoff forecast application.

which describe ECM's reaction to the changes, whilst during periods where no new rules were added, ECM modifies the existing rules' parameters.

The GSETSK model has the most dynamic rule base system. As shown in Figure 3(b), the dynamic rule base arising from the rule pruning and rule updating procedures allows the model to either increase or decrease the number of rules in real time, as it responds to the complexities of the rainfall-runoff problem. Similar to DENFIS, GSETSK rule base has an increasing trend in the first decade. Starting from an empty rule base on 11 January 1961, the number of rules increased rapidly to 11 rules around September 1970. As expected, during this time, rule construction is observed to be more dominant (compared to rule pruning), after which both rule generation and pruning progress interchangeably, keeping the rule

base to a size of approximately eight rules. Interestingly, GSETSK achieves a similar number of rules as ANFIS at the end of the test period; however, ANFIS required a significant portion of the data to be used for batch training. It is also observed that the number of rules increased markedly during the extreme peak event in 1980, where the number of rules increases from 7 to 12 over a 7-month period. After this time, lower levels of runoff were encountered and there is less fluctuation in the rule base. This shows the dynamic nature of GSETSK, progressing from an initial stage of rule development, becoming highly active during extreme periods of runoff and stabilizing during less extreme runoff periods. This is in contrast to the DENFIS model, where the number of rules could only increase. Without the removal of rules that may have become redundant or obsolete, the model becomes overparameterized and

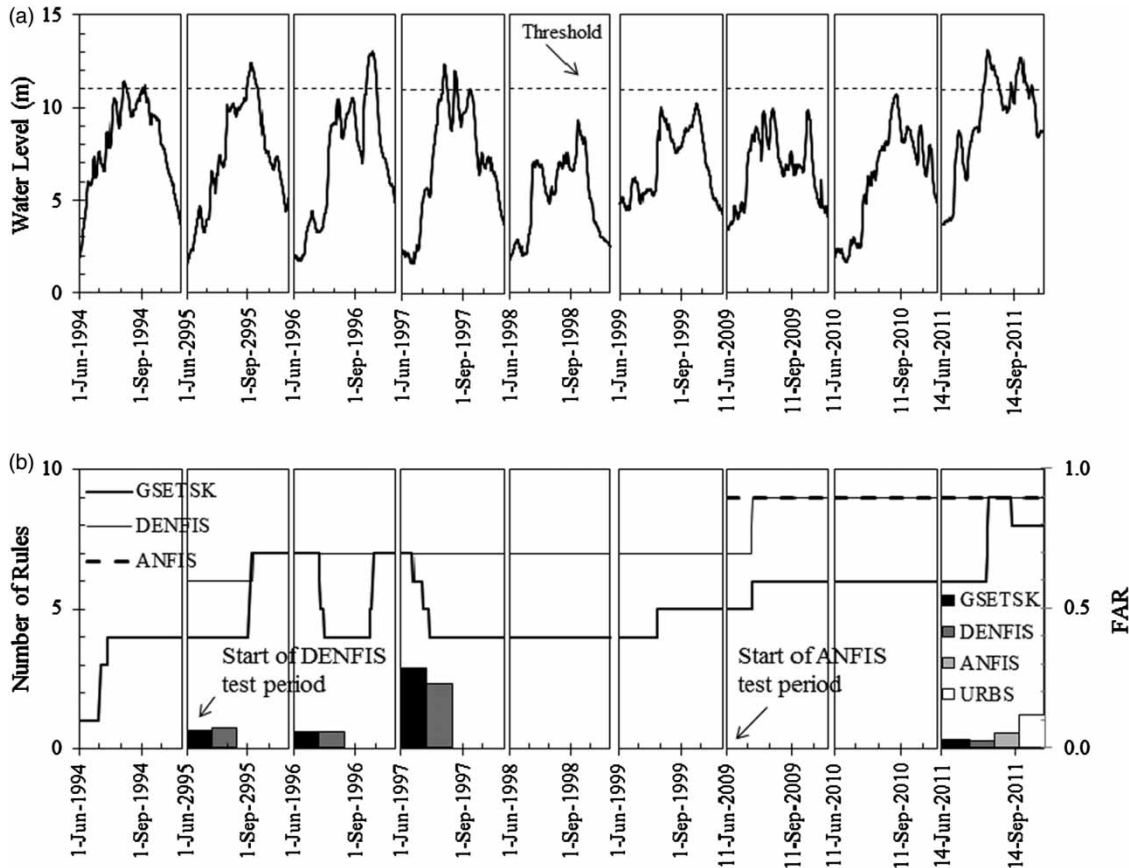


Figure 4 | (a) Time series of water level and (b) change in the number of rules and yearly FAR during model operation for different NFMs in the river routing forecast application.

deterioration in the results can be expected. This observation suggests that a more dynamic online model such as GSETSK is necessary in order to respond to the dynamic nature of the rainfall-runoff process.

The FAR (Figure 3(b)) computed on a yearly basis shows that FAR values for GSETSK are generally higher for the first half of the data series and generally reduces in the latter half. Large spikes in FAR for DENFIS, however, are evident, even in the second half of the data series (e.g. 1983, 1988, 1990) in spite of the online feature in DENFIS. As mentioned, the number of DENFIS rules rises to approximately 20 at the latter stages, about double the number of GSETSK rules. Presumably, this increase in the number of DENFIS rules makes the rule base unnecessarily complicated resulting in a deterioration in results.

Comparing to rainfall-runoff, the river routing problem is less complex since the translation of the flood wave from upstream to downstream is smoothly varying.

The rule base dynamics of the river routing problem shows a different response compared to the rainfall-runoff problem. As shown in Figure 4(b), an ANFIS model with nine rules was established using the training dataset after which the ANFIS model was used to provide the water level forecasts of the test dataset. Similar to the rainfall-runoff problem, both the number (nine rules) and parameters were determined during training and did not change during the test phase. The DENFIS model produced six rules at the end of the training period (1 June 1994 to 31 October 1994). This number of rules was found to be increased to seven on September 1995 and eight and nine on July 2009. At other times, the number of rules remained fixed throughout the remainder of the model run although model parameters were updated. For GSETSK, the model started with the generation of the first rule on 1 June 1994 after which the number of rules fluctuated between 4 and 10.

Significantly, the number of rules generated by the three NFMs are similar at the end of the test period for the river routing problem. The FAR (Figure 4(b)) values calculated on a yearly basis (FAR values for 1998, 1999, 2009, 2010 are not available since the water levels are <11 m) shows that generally FAR values for GSETSK and DNEFIS are very similar; however, both the local models have significantly smaller FAR values for the 2011 flood season compared to URBS.

The above analysis contrasted rule base dynamics of different NFMs for hydrologic and hydraulic forecast applications. For both the processes studied, it is evident that static models are limited since its rule base remains fixed during the test phase. This can also imply that a global model such as ANFIS will, in general, require a relatively large training dataset, to increase the likelihood that the training dataset will generate sufficient global rules. Among the two dynamic models studied, GSETSK accords an obvious advantage for the more complex rainfall-runoff problem, since the model has a more compact and up-to-date rule base. For the river routing problem, however, which is simpler, the benefits of GSETSK are more modest. In this case, both DENFIS and GSETSK have roughly the same number of rules at the end of the model operation.

CONCLUSIONS

The following can be concluded from this study:

1. All NFMs compared favorably against physical models adopted as benchmark.
2. The local NFMs (DENFIS and GSETSK) require considerably less data for training when compared against global model in order to achieve the same results.
3. Rainfall-runoff modeling is a more complex process (compared to river routing) and the GSETSK model was found to be best suited for this application as the model was able to generate a compact and up-to-date rule base. The DENFIS model resulted in the generation of a rule base that was much larger in size and presumably more complicated. Although ANFIS resulted in a rule base comparable in size to GSETSK, this came at a

cost of having the need for a large dataset for model training.

4. All NFMs produced a similar number of rules for the river routing application. River routing is a relatively simpler process to model and our study shows that both the local models were able to achieve similar results, i.e. the process is relatively simpler to model and does not require dynamic rule base features. Although ANFIS provided similar results to the local models, this came at a cost of having the need for a large dataset for model training.

REFERENCES

- Aqil, M., Kita, I., Yano, A. & Nishiyama, S. 2007 A comparative study of artificial neural networks and neuro-fuzzy in continuous modeling of the daily and hourly behaviour of runoff. *Journal of Hydrology* **337** (1–2), 22–34.
- Ashrafi, M., Chua, L. H. C., Quek, C. & Qin, X. 2017 A fully-online neuro-fuzzy model for flow forecasting in basins with limited data. *Journal of Hydrology* **545**, 424–435.
- Bergstrom, S. & Forsman, A. 1973 Development of a conceptual deterministic rainfall-runoff model. *Hydrology Research* **4** (3), 147–170.
- Carrol, D. G. 2007 *The Manual for URBS – A Rainfall Runoff Routing Model for Flood Forecasting and Design*. <http://www.urbs.com.au/>.
- Chang, F.-J. & Chen, Y.-C. 2001 A counterpropagation fuzzy-neural network modeling approach to real time streamflow prediction. *Journal of Hydrology* **245** (1–4), 153–164.
- Chau, K. W., Wu, C. L. & Li, Y. S. 2005 Comparison of several flood forecasting models in Yangtze River. *Journal of Hydrologic Engineering* **10** (6), 485–491.
- Chidthong, Y., Tanaka, H. & Supharatid, S. 2009 Developing a hybrid multi-model for peak flood forecasting. *Hydrological Processes* **23** (12), 1725–1738.
- Chiu, S. L. 1994 Cluster estimation method with extension to fuzzy model identification. In: *Proceedings IEEE International Conference on Fuzzy Systems*, Orlando, FL, 26–29 June 1994, Vol. 2, pp. 1240–1245.
- Draper, N. R. & Smith, H. 2014 *Applied Regression Analysis*. Wiley, Canada.
- Eray, O., Mert, C. & Kisi, O. 2017 Comparison of multi-gene genetic programming and dynamic evolving neural-fuzzy inference system in modeling pan evaporation. *Hydrology Research* **49** (4), 1221–1235.
- Fortin, V., Ouarda, T. B. M. J. & Bobée, B. 1997 Comment on ‘The use of artificial neural networks for the prediction of water quality parameters’ by H. R. Maier and G. C. Dandy. *Water Resources Research* **33** (10), 2423–2424.

- Ghalkhani, H., Golian, S., Saghafian, B., Farokhnia, A. & Shamseldin, A. 2013 Application of surrogate artificial intelligent models for real-time flood routing. *Water and Environment Journal* **27** (4), 535–548.
- Hong, Y. S. T. & White, A. 2009 Hydrological modeling using a dynamic neuro-fuzzy system with on-line and local learning algorithm. *Advances in Water Resources* **32** (1), 110–119.
- Jang, J.-S. R. 1993 ANFIS: adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man, and Cybernetics* **23** (3), 665–685.
- Kasabov, N. K. & Song, Q. 2002 DENFIS: dynamic evolving neural-fuzzy inference system and its application for time-series prediction. *IEEE Transactions on Fuzzy Systems* **10** (2), 144–154.
- Kohavi, R. 1995 A study of cross-validation and bootstrap for accuracy estimation and model selection. In: *Proceedings of International Joint Conference on AI*, Montreal, Quebec, Canada, 20–25 August 1995, pp. 1137–1145.
- MRC 2015 http://ffw.mrcmekong.org/historical_rec.htm.
- Mukerji, A., Chatterjee, C. & Singh Raghuvanshi, N. 2009 Flood forecasting using ANN, neuro-fuzzy, and neuro-GA models. *Journal of Hydrologic Engineering* **14** (6), 647–652.
- Nayak, C., Sudheer, K. P. & Jain, S. K. 2007 Rainfall-runoff modeling through hybrid intelligent system. *Water Resources Research* **43**, 7.
- Nguyen, K. T. & Chua, L. H. C. 2012 The data-driven approach as an operational real-time flood forecasting model. *Hydrological Processes* **26** (19), 2878–2893.
- Quah, K. H. & Quek, C. 2006 FITSK: online local learning with generic fuzzy input Takagi-Sugeno-Kang fuzzy framework for nonlinear system estimation. *IEEE Transactions on Systems, Man, and Cybernetics, Part B* **36**, 166–178.
- Remesan, R., Shamim, M. A., Han, D. & Mathew, J. 2008 ANFIS and NNARX based rainfall-runoff modeling. In: *IEEE International Conference on Systems, Man and Cybernetics, 2008 (SMC 2008)*, Singapore, 12–15 October 2008, pp. 1454–1459.
- Sene, K. 2008 *Flood Warning, Forecasting and Emergency Response*. Springer-Verlag, Berlin, Heidelberg.
- SMHI 2013 <http://www.smhi.se/Professionella-tjanster/Professionella-tjanster/Miljo-och-klimat/Vattenmiljo/referensguide-till-smhi-vattenwebb-1.22742>.
- Talei, A., Chua, L. H. C., Quek, C. & Jansson, E. 2013 Runoff forecasting using a Takagi-Sugeno neuro-fuzzy model with online learning. *Journal of Hydrology* **488**, 17–32.
- Wang, W.-C., Chau, K.-W., Cheng, C.-T. & Qiu, L. 2009 A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series. *Journal of Hydrology* **374** (3–4), 294–306.

First received 23 October 2018; accepted in revised form 6 March 2019. Available online 28 March 2019