



2008

Development of Neuro-Fuzzy Models to Account for Temporal and Spatial Variations in a Lumped Rainfall-Runoff Model

Ahmed Nasr

Dublin Institute of Technology, ahmedsiddig@hotmail.com

Michael Bruen

UCD, michael.bruen@ucd.ie

Follow this and additional works at: <http://arrow.dit.ie/engschcivart>

Recommended Citation

Nasr, A., and Bruen, M., 2008. "Development of neuro-fuzzy models to account for temporal and spatial variations in a lumped rainfall-runoff model". *Journal of Hydrology*, vol. 349, pp. 277-290. doi:10.1016/j.jhydrol.2007.10.060

This Article is brought to you for free and open access by the School of Civil and Structural Engineering at ARROW@DIT. It has been accepted for inclusion in Articles by an authorized administrator of ARROW@DIT. For more information, please contact yvonne.desmond@dit.ie, arrow.admin@dit.ie, brian.widdis@dit.ie.



This work is licensed under a [Creative Commons Attribution-NonCommercial-Share Alike 3.0 License](https://creativecommons.org/licenses/by-nc-sa/3.0/)



1 **Development of Neuro-Fuzzy Models to account for temporal and spatial**
2 **variations in a lumped rainfall-runoff model**

3 Ahmed Nasr^a and Michael Bruen^{a*}

4 a Centre for Water Resources Research, University College Dublin, Newstead Building, Clonskeagh,
5 Dublin 14, Ireland

6 * Corresponding author: Tel. +353 1 7163212, Fax: +353 1 7163297, E-mail: michael.bruen@ucd.ie
7

8
9 **Abstract**

10 For many good and practical reasons, lumped rainfall-runoff models are widely used
11 to represent a catchment's response to rainfall. However, they have some
12 acknowledged limitation, some of which are addressed here using a neuro-fuzzy
13 model to combine, in an optimal way, a number of lumped-sub-models. For instance,
14 to address temporal variation, one of the sub-models in the combination may perform
15 well under flood conditions and another under drier conditions and the neuro- fuzzy
16 system would combine their outputs for each time-step in a manner depending on the
17 prevailing conditions. Similarly to address spatial variation, one of the sub-models
18 may perform well for the upland parts of the catchment and another for the lowland
19 parts and again the neuro-fuzzy system is expected to combine the different outputs
20 appropriately. The proposed combination method can use any lumped catchment
21 model, but has been tested here with the Simple Linear model (SLM) and the Soil
22 Moisture and Accounting Routing (SMAR) models. Eleven catchments with different
23 hydrological and meteorological conditions have been used to assess the models with
24 respect to temporal variations in response while one catchment is used to address the
25 effect of spatial variation. The neuro-fuzzy combined-sub-models of SLM and SMAR
26 modelled the temporal and spatial variation in catchment response better than the

1 lumped version of each model. Also the SMAR model significantly outperformed the
2 SLM either as a lumped model or as a sub-model in any of the combinations.

3

4 **Keywords:** Neuro-fuzzy; lumped model; combined-sub-models; Simple Linear
5 model; Soil Moisture and Accounting Routing model; rainfall-runoff modelling

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

1 **1. Introduction**

2 Mathematical models are widely used in water resources applications despite
3 considerable difficulties arising from catchment heterogeneity, strong non-linearity in
4 its response to precipitation and uncertainties in parameter estimation. In many
5 practical cases, simple lumped models of either the black-box or conceptual type
6 often perform adequately, and compare well with the more complex distributed-
7 models, particularly for flood modelling in small catchments and for behaviour within
8 the range of the data used to calibrate its parameters (Beven, 2000). Larger
9 catchments can be modelled by associating different sub-lumped-models with
10 different spatial units within the catchment (e.g. Chen and Adams, 2006; Marechal
11 and Holman, 2005; Ajami et al., 2004). Similarly different sub-models could be used
12 to represent the various temporal patterns in the system's response (e.g. Shamseldin
13 and O'Connor, 1996; Ahsan and O'Connor, 1994; Kachroo and Natale, 1992). The
14 success of this approach is primarily because of its ability to capture some of the non-
15 linearity in the catchment behaviour resulting from its spatial heterogeneity and time-
16 varying character. The choice of a suitable lumped model for use in each of the sub-
17 catchments is critical to its success and practicality. It should have a small number of
18 parameters to reduce the total number to be estimated for the combined model thereby
19 reducing the computational requirements. This also is likely to reduce potential
20 problems caused by model over-parameterisation, such as ill-conditioning (Bruen and
21 Dooge, 1992), or equifinality (Beven, 1993) in which a number of different
22 combinations of parameter values give similar model fits and so a single optimal
23 parameter set is difficult to determine.

24

1 One obvious symptom of non-linearity is the very different responses of the
2 catchment to different flow regimes. The direct way of dealing with this is to build the
3 complicated non-linear physical relationships into the model. An alternative is to have
4 a different, but simple, sub-model for each different flow regime. For instance, Chen
5 and Adams (2006) used a number of sub-models to simulate spatial variation in the
6 rainfall-runoff relationship. The estimated runoffs from all sub-models were
7 combined together using an artificial neural network to estimate the total runoff.
8 Moreover, they investigated the suitability of using sub-models of three different
9 conceptual models including the Xinanjiang Model (Zhao and Liu, 1995), the Soil
10 Moisture Accounting and Routing (SMAR) Model (O'Connell et al., 1970) and the
11 Tank Model (Sugawara, 1995). A significant improvement was obtained when using
12 different sub-models compared to a single lumped model. Kachroo and Natale (1992)
13 also used three sub-models using the same Simple Linear Model (SLM) (Nash and
14 Foley, 1982) structure with different parameter sets to represent the response during
15 low, medium and high flow regimes. Although the total number of parameters is
16 tripled, all of the sub-lumped-model parameters could be calibrated using the least-
17 squares criterion. The choice of which of the sub-model to use at each time step is
18 guided by a type of wetness index taken as the current observed discharge in this case.
19 When no observed discharge is available at the current time step, (e.g. when either (a)
20 simulating or (b) forecasting beyond a single time step) the discharge simulated by the
21 lumped model is used for this index. The combined-sub-lumped models have shown
22 significant improvement over the lumped one.
23
24 Building on these efforts to improve the performance of combined-sub-lumped-
25 models, this paper reports the investigation of a fuzzy method proposed to combine

1 sub-lumped-models of two types, black box model and conceptual model. The former
2 is the Simple Linear Model (SLM) (Nash and Foley, 1982) and the latter is the Soil
3 Moisture Accounting and Routing model (SMAR) (O'Connell et al., 1970). Each of
4 the two models has been included into a framework of a special type of Neuro-Fuzzy
5 Model (NFM), called an Adaptive Neuro-Fuzzy Inference System. The first objective
6 is to produce a combined-lumped-model better able to represent the spatial and
7 temporal variability of the catchment's response to rainfall. The resulting NFM
8 addresses the temporal variations in response by using a number of sub-models for the
9 SLM and the SMAR models for different regimes (e.g. separate sub-models for floods
10 and low flow situations). Each of the sub-models describes a particular feature in the
11 temporal pattern of the catchment's response. The NFM is assessed by applying it to
12 eleven different catchments from around the world. In the second part of this study an
13 NFM (for the SLM and for the SMAR model) is developed that is able to identify
14 homogenous spatial units within a catchment on which the sub-models can be based.
15 In this, the NFM structure of the first part is further coupled to a subtractive fuzzy
16 clustering algorithm (Vernieuwe et al., 2005) to determine the homogeneous spatial
17 units using a number of spatial variables specified on a catchment grid. Finally, using
18 one of the catchments which has the required spatial database, namely, the Brosna, the
19 NFM developed in the second part of the study is tested and its results compared with
20 those of the corresponding model developed in the first part of the study.

21

22 The proposed method is described in section 2 and the NFM is reviewed in section 3.
23 The two rainfall-runoff models, SLM and SMAR, are briefly described in sections 4
24 and 5 respectively. In section 6, a detailed description is given of the two NFMs
25 applied in this study. In the final sections, 7 and 8, the results of the NFM applications

1 are presented and conclusions are drawn. Suggestions for further work are added in
2 section 9.

3

4 **2. Interpretation of the proposed sub-models combination method**

5 The method of sub-model combination used in this study is different from the flood
6 forecast model combination methods proposed in earlier work (e.g. Shamseldin et al.,
7 1997; See and Openshaw, 2000; See and Openshaw, 1999; Xiong et al., 2001;
8 Abrahart and See, 2002; Coulibaly et al., 2005; Fenicia et al. 2007). In those methods,
9 a number of models each with different internal structures were individually applied
10 to the entire study catchment and their simulated outputs were combined. Each model
11 was attempting the same task, to simulate the entire catchment. In contrast, in our
12 approach each model is truly a sub-model, assigned to simulate a particular part of the
13 catchment or a specific range of responses, e.g. for a particular flow regime.

14

15 Following the multi-linear model approach pioneered by Bruen (1985), Becker and
16 Kundzewicz (1987), Kachroo and Natal (1992), and Todini and Wallis (1997), our
17 proposed sub-model approach was previously used to build different rainfall-runoff
18 models. For instance, Bruen (1985) constructed a quasi-linear model from a
19 combination of linear sub-models. An illustration of the structure of this quasi-linear
20 model, with a single threshold, is given in Fig. 1. Note: (i) The input series (I) is
21 effectively divided into a number of separate series (e.g. I_1 , I_2 , etc.), each of the same
22 length as the original. The division procedure is performed in two steps. First, the
23 range of values in the input series is divided into a number of parts by threshold levels
24 (partitions) of fixed values. Then the magnitude of each input value determines the
25 band or division in which it lies, and the entire input in that band is then assigned to

1 the corresponding time series. (ii) The output from each of the separated input series
2 (e.g. O_1 , O_2) is obtained from a number of separate models (e.g. $model_1$, $model_2$). (iii)
3 The total output (e.g. O_f) is the sum of the outputs from each of the different models
4 applied to the corresponding separated inputs. This allows the overall model to
5 respond differently to low rainfall compared to high rainfall.

6

7 In essence a number of sub-models are constructed to describe the relationship
8 between the input and the output for different ranges of their values representing
9 different hydrologic conditions. This requires that each input value should be assigned
10 to a specific sub-set (e.g. low values, medium values, high values). Such an approach
11 assumes the inputs can be assigned to the sub-sets with certainty but there are times
12 where uncertainty might occur, such as when the magnitude of an input value is close
13 to a partition threshold value. The method proposed addresses this uncertainty using
14 fuzzy logic theory whereby different levels of memberships of input to all sub-sets are
15 estimated. These degrees of memberships can be taken as the weights given to the
16 outputs from the models corresponding to each of the input sub-sets.

17

18 To illustrate our proposed method Fig.1 has been extended in Fig. 2 which shows, still
19 for the case of a single threshold, how the concept of the membership of fuzzy sub-
20 sets is used to define weights given to the sub-models. Unlike in Bruen's method
21 (Bruen, 1985), the input series (e.g. I) is not separated here but alternatively it is
22 assumed that for certain hydrologic conditions there is a sub-model (e.g. $model_1$,
23 $model_2$) and a membership function (e.g. mf_1 , mf_2) associated with it. The former
24 produces the output (e.g. O_1 , O_2) from the sub-model while the latter calculates
25 membership values used to estimate the weight given to that output (e.g. w_1 , w_2). The

1 final output value (e.g. O_f) from the combination is the weighted average of the
2 outputs from the models used for each sub-set. It is worth mentioning that the method
3 described above is valid for the case of a lumped catchment. However, if the
4 catchment is divided into sub-catchments, then the method can be applied separately
5 to each sub-catchment and the final output can be estimated as the area-weighted
6 average of the outputs of each of the sub-catchments (where routing to the catchment
7 outlet is considered part of the sub-model).

8

9 **3. Neuro-Fuzzy Model (NFM)**

10 The Neuro-Fuzzy Model (NFM) used in this study implements the Takagi-Sugeno
11 fuzzy approach (Takagi and Sugeno, 1985) to obtain a direct crisp value for the output
12 variable(s) from fuzzy input variable(s). Jacquin and Shamseldin (2006) explored the
13 application of Takagi-Sugeno fuzzy inference systems to rainfall-runoff modelling.
14 They developed two different fuzzy models to account for the non-linearity in the
15 catchment response due to both antecedent catchment wetness and seasonality.
16 Vernieuwe et al. (2005) also investigated fuzzy rule-based models of the Takagi-
17 Sugeno type relating rainfall to catchment discharge. Their models differed in the
18 methods used to partition the spaces of the input and output variables and hence the
19 identification of the number of membership functions and their locations for each
20 variable. Earlier, the Takagi-Sugeno model was used by Xiong et al (2001) in the
21 multi-model output combination context. All these studies produced models by
22 combining different sub-models and this also has been followed in the present study.
23
24 A fuzzy number consists of a number of sub-sets each of which has an interval of
25 possible values between specified minimum and maximum limits. For every point in

1 the interval a corresponding membership function is defined that represents, within
2 the interval, the degree of confidence one might have for a particular value of the
3 fuzzy number (Ganoulis, 1994).

4

5 Generally the NFM consists of five layers configured analogously to any multi-layer
6 feed-forward neural network. Chen et al. (2006) named these five layers according to
7 their operative function, as ‘input nodes’, ‘rule nodes’, ‘average nodes’, ‘consequent
8 nodes’, and ‘output nodes’ respectively. Fig. 3 illustrates an NFM with two input
9 variables, x and y , each of which has two fuzzy sub-sets, A1 and A2 for x and B1 and
10 B2 for y . The first layer in the Figure thus has four nodes, one for each of the two
11 fuzzy sets of each of the two input variables. Each node in the first layer receives a
12 crisp value of one of the input variables (e.g. x) and, for each fuzzy sub-set of this
13 input variable, it uses a membership function (e.g. A1) to generate a membership
14 grade (e.g. u_{A1}). Different shapes for the membership function, such as Gaussian,
15 Generalised bell shaped, trapezoidal shaped, and triangular, can be used.

16

17 Although the second, third and fourth layers have different functions each has the
18 same number of nodes. Each node in these three layers is assigned to a certain IF-
19 THEN rule, called “the antecedent part” of the NFM. The total number of IF-THEN
20 rules is determined by the number of possible combinations of the fuzzy sub-sets of
21 the input variables. This procedure gives the neuro-fuzzy model an advantage over the
22 ordinary fuzzy logic model because the former does not require the modeller to
23 specify in advance the number of rules.

24

1 The function of each node in the second layer is to multiply the membership grades of
2 all fuzzy sub-sets involved in a specific IF-THEN rule (e.g. u_{AI} and u_{BI}) to obtain the
3 weight for this rule (e.g. w_I) which is normalised in the corresponding node in the
4 third layer. The normalised weight (indicated by a bar \bar{w}_1) is obtained by dividing the
5 weight assigned to that particular IF-THEN rule by the sum of the weights for all
6 rules (e.g. $\bar{w}_1 = w_1 / (w_1 + w_2 + w_3 + w_4)$). The nodes in the fourth layer compute the
7 fractional contribution to the final model output(s) of each IF-THEN rule and this
8 layer represents “the consequent part” of the NFM. This fraction is the product of the
9 normalised weight of the associated IF-THEN rule (e.g. \bar{w}_1) by a value calculated
10 from a function associated with this rule to transform the crisp values of the inputs
11 into a scalar output (e.g. $f_1(x,y)$). In the original NFM formulation, a first order
12 polynomial model, such as a Linear Transfer Function (Box and Jenkins, 1976), was
13 used for this purpose. However, in this study the black-box SLM and conceptual
14 SMAR catchment models are used instead.

15

16 In the fifth layer, each output variable is represented by a neuron. The final output
17 produced by each neuron in the fifth layer is the aggregation of the outputs
18 contributed by all the associated IF-THEN rules. Fig. 3 illustrates the case of a single
19 output Z , from the single neuron in the fifth layer, having the form:

$$20 \quad Z = \bar{w}_1 * f_1(x, y) + \bar{w}_2 * f_2(x, y) + \bar{w}_3 * f_3(x, y) + \bar{w}_4 * f_4(x, y) \quad (1)$$

21

22 **4. Simple Linear Model**

23 The Simple Linear Model (SLM) was introduced by Nash and Foley (1982) as a
24 naïve, benchmark, model against which the performance of more substantive and

1 sophisticated rainfall-runoff models could be compared. The SLM assumes a linear
 2 time invariant relationship between rainfall and discharge, expressed by a convolution
 3 summation relation. Here, an additional term has been added in order to include,
 4 albeit crudely, losses due to evaporation in the modelling, giving the equation:

$$5 \quad q_i = G \sum_{j=1}^m r_{i-j+1} h_j + \alpha \cdot e_i + \varepsilon_i \quad (2)$$

6 where q_i , r_i , and e_i are the discharge, rainfall and evaporation respectively at the i^{th}
 7 time step, h_j is the j^{th} ordinate of the discrete pulse response function, m is the memory
 8 length of the system, G is the gain factor, α is the coefficient of the evaporation term
 9 (this can be set to zero if evaporation is to be ignored) and ε_i is the error term.

10 Usually, the sum of the h_j terms is unity.

11

12 This is a multiple linear regression of the observed discharge on the m previous
 13 observed rainfall values and the current evaporation value. For the pulse response
 14 terms, h_j , either a parametric or non-parametric form can be used, and the two-
 15 parameters Nash cascade model (Nash, 1957) is used here. The discrete h_j terms are
 16 calculated from its impulse response function $h(t)$ which has the following form:

$$17 \quad h(t) = (1/k\Gamma(n))(t/k)^{n-1} \exp^{-t/k} \quad (3)$$

18 where $\Gamma(n)$ is the gamma function.

19

20 Thus the SLM, with the pulse response function in parametric form, has four
 21 parameters, G , n , k , and α .

22

23 **5. Soil Moisture Accounting and Routing (SMAR) model**

1 O'Connell et al. (1970) developed a quasi-physical rainfall-runoff model known as
2 the layers model but later on renamed the Soil Moisture Accounting and Routing
3 (SMAR) model. This model consists of two complementary components. The first
4 implements a water balance (the soil moisture accounting procedure) between rainfall,
5 evaporation, runoff, and simulated soil storage for each time step. The second routes
6 the calculated runoff to the catchment outlet, taking account of attenuation and wave
7 diffusive effects. A number of modifications to the original structure of the model
8 have been introduced (Khan, 1986; Liang, 1992) and the latest version by Tan and
9 O'Connor (1996) is used here. It has eight parameters in the water balance component
10 and three parameters in the routing component. In addition, the initial condition of the
11 groundwater storage is considered as a parameter bringing the total number of
12 parameters to twelve.

13

14 **6. Description of the proposed NFM**

15 Ozelkan and Duckstein (2001) described any catchment model as a system composed
16 of sub-modules to represent the sub-elements of this modelled system coupled
17 together in order to produce a synergic effect reflected at the output of the system.
18 The representation of the catchment model in this modal structure is equivalent to the
19 branching structure in an algorithm flow diagram resulting from 'IF-THEN' fuzzy
20 rules (Gupta and Sorooshian, 1983). In the present work, the aim is not to utilise the
21 'IF-THEN' fuzzy rules as the model core but rather to improve the performance of
22 deterministic catchment models by using a number of 'IF-THEN' fuzzy rules to create
23 specific localised versions of these models which are better able to respond to local
24 variations in the pattern of temporal and spatial data. The approach is similar to that of
25 Jaquin and Shamseldin (2006) who investigated the combination of different

1 empirical sub-models, using a fuzzy logic model, to account separately for variation
2 in catchment wetness and for catchment seasonality.

3

4 In this study, temporal variations are accounted for in a separate modelling scenario,
5 called NFM_T, and the spatial variations in another one, called NFM_S. The NFM
6 structure for both scenarios is similar to the one illustrated in Fig. 1. All NFMs used in
7 this study employ the Gaussian function to represent the membership function of all
8 temporal input variables to the models. This function has the following analytical
9 expression:

$$10 \quad u(x) = \exp^{-(x-c)^2 / 2\sigma^2} \quad (4)$$

11 where $u(x)$ is membership value of a variable x to certain fuzzy sub-set, and
12 parameters c and σ specify the location and spread of the function and require
13 calibration.

14

15 As mentioned earlier, the two models, SLM and SMAR, are used in the consequent
16 part of the NFM in both modelling scenarios. It is worthwhile stressing at this point
17 that the resulting consequent part of the NFM for each scenario can be visualised as a
18 collection of either SLM or SMAR sub- or local-models determined according to the
19 IF-THEN rules acting in parallel. Indeed it is the generation of such a configuration,
20 as an alternative method of involving the temporal and spatial pattern variations of the
21 variables in modelling the rainfall-runoff relationship, that is sought in this study.

22

23 **6.1. NFM_T modelling scenario**

24 In the NFM_T scenario there are two inputs, rainfall and evaporation, and the output,
25 discharge, is calculated using one or other of the catchment models. To distinguish

1 between the NFM_T variant which uses SLM and the other which uses SMAR in the
 2 consequent part they are called NFM_T_SLM and NFM_T_SMAR respectively. For
 3 each model a total of ten possible rainfall and evaporation fuzzy sub-set combinations
 4 are formulated as indicated in Table 1. The performances of all ten cases are
 5 evaluated separately for eleven catchments from different parts in the world. Details
 6 of these eleven catchments are given in Table 2.

7

8 The total number of parameters ($npar$) requiring calibration is determined from

- 9 (i) number of fuzzy subsets for the rainfall (nr_{fsub}) and the evaporation (ne_{fsub});
- 10 (ii) number of the IF-THEN rules (this is equal to $nr_{fsub} * ne_{fsub}$); and
- 11 (iii) number of the model parameters (P) (4 for SLM and 12 for SMAR).

12 The relation used to calculate $npar$ is as follow:

$$13 \quad npar = 2 * (nr_{fsub} + ne_{fsub}) + (nr_{fsub} * ne_{fsub}) * P \quad (5)$$

14 The first term in the above equation gives the total number of the Gaussian function
 15 parameters for all fuzzy sub-sets while the second term gives the total number of the
 16 SLM or SMAR model parameters. Thus there are two sets of parameters that need to
 17 be determined by the calibration process. The first set is the parameters of the
 18 Gaussian membership functions of the rainfall and evaporation. The second set is the
 19 parameters of the models (SLM and SMAR) which are used to relate the rainfall and
 20 evaporation (input variables) with the discharge (output variable). The overall
 21 optimisation problem is non-linear and it has been found that if the two sets of
 22 parameters are determined simultaneously the calibration is often poor. Hence the
 23 calibration is performed in a sequential iterative procedure as follows; (i) Initial
 24 values are given to the parameters of the SLM and SMAR models, (ii) Holding the
 25 SLM and SMAR model parameters constant, the parameters of the Gaussian function

1 sub-sets of the rainfall and evaporation are determined by using the Genetic algorithm
2 (Holland, 1975). (iii) The Gaussian function parameters are then held constant and the
3 parameters of SLM and SMAR models are recalibrated in a second optimisation step.
4 The least squares method is used for the linear optimisation problem required by the
5 NFM_T_SLM whereas the Genetic algorithm is used for the non-linear one in the
6 NFM_T_SMAR. (iv) If the resulting objective function is less than a specified
7 tolerance the calibration stops otherwise step (ii) to (iii) are repeated. Note that the
8 initial values of the parameters of SLM and SMAR models in this case are the ones
9 obtained from the calibration in step (iii).

10
11 A split sampling approach was used for model testing, in which the available data for
12 each catchment was split into two parts. The first part (67% of the data) was used in
13 the model calibration while the second (33% of the data) was used in verifying the
14 calibrated models. Two criteria are used in calibration and validation, (i) the Nash-
15 Sutcliffe index (R^2) (Nash and Sutcliffe, 1970) and (ii) the average relative errors
16 (*ARE*) of the estimated discharge peaks over a threshold, conservatively set here as
17 the mean discharge. In addition to these numerical criteria, the observed and the
18 simulated hydrographs for some catchments, for each calendar year, have been plotted
19 to illustrate the fit of the hydrograph shapes.

20
21 **6.2. NFM_S modelling scenario**

22 Here the performance of the NFM_S model with the SLM and the SMAR sub-models
23 is assessed. The first case is called NFM_S_SLM while the latter is called
24 NFM_S_SMAR. However, unlike the NFM_T scenario the modelled catchment in the
25 NFM_S scenario is divided spatially into a number of Homogenous Hydrologic

1 Characteristics Units (HHCUs). Although, analogous to Hydrologic Response Units
2 (HRUs) (e.g. Quiroga et al., 1996), HHCUs are defined and determined in a
3 somewhat different way. The inputs to each HHCU are the catchment averages of
4 rainfall and evaporation.

5
6 If the rainfall and evaporation for each HHCU are used as fuzzy variables then their
7 fuzzy sub-sets can be used to determine the number of IF-THEN rules in the
8 consequent part of each sub-NFM model for each HHCU. However, as only one
9 fuzzy sub-set is used for rainfall and likewise only one for evaporation the resulting
10 combined sub-NFM models is essentially a model describing different homogenous
11 spatial units, i.e. each IF-THEN rule represents a sub-model describing the rainfall-
12 runoff relationship for a given HHCU and the final estimated runoff value is the
13 weighted sum of the contribution from all the HHCUs. This is a type of semi-
14 distributed modelling that can be easily implemented either within or in conjunction
15 with a GIS by overlaying three map layers, the catchment boundary, land use map,
16 and soil map. The number of the HHCUs obtained with this GIS procedure is based
17 only on elevation, land use and soil type and here they are determined with an
18 innovative approach based on the subtractive clustering algorithm (Vernieuwe et al.,
19 2005).

20
21 **6.2.1. Determination of the HHCUs for the Brosna catchment**

22 Each HHCU is expected to have a unique rainfall-runoff relation used to estimate its
23 contribution to the catchment outflow. A large number of spatially-related parameters
24 such as elevation, soil permeability, soil roughness, bedrock transmissivity, etc. could
25 influence the rainfall-runoff response and could be used to characterise the HHCU.

1 However, for this study, the number of such variables is limited to elevation, land use,
2 soil type and these were used to test the NFM_S for the Brosna catchment only. From
3 these three basic maps, four spatial variables are calculated by the GIS (i) elevation,
4 (ii) slope, (iv) land use, and (iv) soil type. Although the original land use map had
5 nineteen different categories, here land use has been aggregated into four main types,
6 agriculture, urban, forest, and wetland. Similarly the slopes obtained directly from the
7 DEM have been assigned to one of three groups: (i) for slopes between 0 % and 8 % a
8 slope index is taken as 4 %; (ii) for slopes between 8 % and 15 % a slope index is
9 taken as 12 %; and (iv) for slope greater than 15 % a slope index is taken as 20 %.

10 The original categories of soil types and elevation bands are used without any changes
11 since they are primary governing parameters in characterising the response to the
12 rainfall.

13

14 Various combination alternatives, summarised in Table 3, of the four input spatial
15 variables are passed on to the subtractive clustering algorithm in order to obtain
16 different number of HHCUs. The resolution of the resulting clusters in each
17 combination alternative can be adjusted by changing the parameters in the subtractive
18 clustering algorithm. In this study, the reject ratio (RR) (c.f. Vernieuwe et al., 2005)
19 had the most influence on the cluster resolution. The RR is used by the subtractive
20 clustering algorithm as a stopping criterion to halt any further attempts to determine
21 new clusters. For each combination alternative the RR was varied from 0.1 to 0.5 in
22 increments of 0.1 and from 0.5 to 1 in increments of 0.05. The calculated numbers of
23 clusters are plotted against reject ratio in Fig. 4. It is clear that for all combination
24 alternatives changing the RR value between 0.1 and 0.65 did not change the number
25 of the resulting clusters. Then there is a gradual drop in the number of clusters

1 corresponding to an increase in RR up to 0.8 which is followed again by a constant
2 number of clusters until RR reaches the value of 0.95. The RR value of 1 corresponds
3 to one cluster and this is consistent with a lumped catchment. Note that for the
4 combination alternatives 3A and 4 the number of clusters corresponding to RR values
5 less than and equal to 0.75 is significantly higher than the corresponding values for
6 the other cases.

7

8 **6.2.2. NFM_S_SLM and NFM_S_SMAR modelling cases**

9 For each combination of spatial variables an upper limit of 40 clusters (shown by
10 section ϕ - ϕ in Fig. 4) is applied to select cases to be considered in the NFM_s tested
11 here. The choice of 40 is aimed to avoid an excessive number of parameters in the
12 NFM_s. As the number of clusters remains constant for a range of RR values, the
13 number of cases tested for the NFM_S_SLM and NFM_S_SMAR models in the
14 Brosna catchment, varies from one combination alternative to another (Table 4).

15

16 Generally when multiple fuzzy sub-sets are used for banding the rainfall and
17 evaporation then the number of parameters to be calibrated for each case in the
18 NFM_S scenario is obtained by multiplying the number of parameters for the NFM_T
19 scenario, given by Eqn. 5, by the number of clusters or HHCUs involved. However,
20 as one fuzzy sub-set is used for both the rainfall and evaporation in the NFM_S
21 scenario only the parameters of the models (SLM and SMAR) must be calibrated.
22 Therefore there is no need for the sequential iterative procedure used in the NFM_T
23 scenario and instead only the least squares method is used for the linear optimisation
24 problem in the NFM_T_SLM whereas the Genetic algorithm is used for the non-
25 linear one in the NFM_T_SMAR.

1

2 **7. Results**

3 The key issue is to determine whether the introduction of combined sub-models to
4 account for temporal or spatial pattern variations improves the simulation compared
5 to that of a single lumped catchment model. First, the results corresponding to the
6 lumped case (case 1 in Table 1 for NFM_T, and cases 1 of all combination
7 alternatives in Table 4 for NFM_S) are calculated. These provide a baseline to be
8 used in assessing the second set of results corresponding to the best combined case. In
9 each scenario, the best combined case can be described as the one with the highest R^2
10 during the calibration period compared to the others in the same group. The best
11 combined case is an improvement over the lumped case if it scores a higher value for
12 the R^2 criterion and a smaller value of the *ARE* criterion. In addition to these two
13 numerical criteria, a graphical comparison of the simulated and the observed
14 hydrographs allowed a visual assessment of model fit.

15

16 In addition, the suitability of using a linear model, such as SLM, or a non-linear
17 model, such as SMAR, in the fuzzy model is also addressed in the discussion.

18

19 **7.1. Results of the NFM_T scenario**

20 **7.1.1. Lumped case vs. the best combined case**

21 For the NFM_T_SLM and NFM_T_SMAR models, the R^2 and *ARE* values for the
22 calibration and validation periods are summarised for the eleven test catchments in
23 Table 5. There is an improvement in the R^2 values during calibration for the best
24 combined case over the lumped case. However, the best combined case improved the
25 R^2 values for validation in nine catchments, the exceptions being Halda and Sg.

1 Bernam, for the NFM_T_SLM and in seven catchments, the exceptions being Halda,
2 Kelantan, Sg. Bernam, and Shiquan-3, for the NFM_T_SMAR model. Only in one of
3 these catchments, Sg. Bernam, the R^2 values during validation of the best combined
4 case were markedly lower than the corresponding values for the lumped case in both
5 the NFM_T_SLM and NFM_T_SMAR models, the differences being insignificant in
6 the rest of the catchments.

7

8 For the *ARE* criterion during calibration, the best combined model case was better
9 than the lumped case for the NFM_T_SLM in all but three catchments (Bird Creek,
10 Kelantan, and Sg. Bernam). During validation the combined models of the
11 NFM_T_SLM gave better *ARE* values than the lumped case in five catchments but
12 was worse in six catchments (Blue Nile, Halda, Kelantan, Nan, Sg. Bernam, and
13 Wolombi Brook). The values of *ARE* for calibration of the NFM_T_SMAR model
14 exhibited a consistent improvement of the best combined case over the lumped case
15 whereas the values of the corresponding validation were worse in six catchments
16 (Bird Creek, Halda, Nan, Sg. Bernam, Sunkosi-1 and Wolombi Brook).

17

18 The best combined case was not consistent for the NFM_T_SLM and
19 NFM_T_SMAR models. For the former model each of cases 8 and 9 was the best in
20 four catchments while case 10 was the best in three catchments. Different trends was
21 obtained in the latter model as each of cases 4, 5, and 10 was the best in three
22 catchments and case 7 was the best in two catchments.

23

24 **7.1.2. NFM_T_SLM vs. NFM_T_SMAR**

1 The values of R^2 and ARE criteria shown in Table 5 for the eleven catchments and for
2 both the lumped case and the best case did not show which of NFM_T_SLM or
3 NFM_T_SMAR is the overall best model. For the lumped case, the R^2 values for the
4 calibration of NFM_T_SMAR were higher than the values of NFM_T_SLM in all
5 catchments. The same occurred in validation except in two catchments, Sg. Bernam
6 and Wolombi Brook. For the best case, only in Shiquan-3 catchment was the value of
7 R^2 for calibration of NFM_T_SMAR lower than for NFM_T_SLM and in validation
8 the same was true for three catchments, Sg. Bernam, Shiquan-3, and Sunkosi-1.

9

10 The ARE values showed even more mixed results as NFM_T_SMAR did not
11 outperform NFM_T_SLM in terms of ARE for the lumped case at two catchments
12 (Chu and Shiquan-3) for calibration and at five catchments (Blue Nile, Chu, Sg.
13 Bernam, Shiquan-3, and Wolombi Brook) for validation. Similar results hold for the
14 best combined case in calibration. It holds also in validation but with the addition of
15 two more catchments (Halda and Sunkosi-1).

16

17 **7.1.3. Hydrographs matching in the NFM_T scenario**

18 The observed and simulated hydrographs of the best combined cases of
19 NFM_T_SLM and NFM_T_SMAR for four catchments, Blue Nile, Brosna, Chu, and
20 Wolombi Brook, are plotted in Figs. 5 to 8. Each of the four catchments exhibits
21 different hydrological behaviour and this is reflected in the shape of its hydrograph. In
22 addition, the period of each hydrograph is chosen to be within the validation period
23 for two reasons: (i) to verify the model parameters; and (ii) to ensure minimal
24 influence of the initial conditions on the models comparison.

25

1 The four graphs demonstrate the ability of the NFM_T_SMAR to capture most of the
2 hydrograph features. This model showed an outstanding performance in reproducing
3 the observed hydrograph in the Chu catchment, (Fig. 7), and to some extent the one in
4 the Blue Nile catchment, (Fig. 5). However, in the Brosna and Wolombi Brook (Figs.
5 6 and 8 respectively), features such as rising limb, recession, and base flow were
6 better generated by this model than the individual peak values.

7

8 The NFM_T_SLM was able to match the non-linearity in the two flashy catchments,
9 Chu and Wolombi Brook, Figs. 7 and 8 respectively. In contrast, this model, with its
10 linear component, was particularly bad for the Brosna, which has a large base flow
11 component, and for the Blue Nile, which has a strong seasonal pattern.

12

13 **7.2. Results of the NFM_S scenario in the Brosna catchment**

14 **7.2.1. Lumped case vs. the best combined case**

15 Table 6 shows the values of the R^2 and ARE model efficiency criteria for the
16 NFM_S_SLM and the NFM_S_SMAR model for both calibration and validation and
17 for the lumped case and the best case for all spatial combination alternatives in the
18 Brosna catchment. The results for the lumped case (treating the catchment as a single
19 unit) of each of the NFM_S_SLM and the NFM_S_SMAR models were identical for
20 all combination alternatives since this involved a single HHCU and only one fuzzy
21 sub-set for rainfall and evaporation.

22

23 The R^2 results for calibration and validation for the NFM_S_SLM do not differ
24 significantly from each other. In contrast, for the NFM_S_SMAR the R^2 values for
25 calibration for the best cases were significantly higher than for the lumped case. In

1 validation, a significant improvement in R^2 was obtained by the best case of the
2 combination alternatives 2A, 2B, and 4 for which the *ARE* values were amongst the
3 lowest. There were no significant differences among the *ARE* values for the
4 NFM_S_SLM and likewise among those of the NFM_S_SMAR models. However,
5 the *ARE* values of the NFM_S_SLM were all much greater than those of the
6 NFM_S_SMAR. The results in Table 6 suggest that, while the 2A combination
7 alternative performs significantly better than the lumped case, the improvement is not
8 as impressive as that obtained for the NFM_T scenario.

9

10 **7.2.2. NFM_S_SLM vs. NFM_S_SMAR**

11 The superiority of SMAR over SLM can be easily seen from the R^2 and the *ARE*
12 values. The introduction of non-linearity in the SLM through the combination of its
13 sub-models did not produce any significant improvement. This is not surprising
14 because the use of HHCUs in this context has no effect on the SLM itself but it rather
15 assigns weights to similar sub-models with the same characteristics as the lumped
16 model. In contrast, in the SMAR model each sub-model adds to the non-linearity of
17 the combined model and this in turn provides the greater flexibility required in
18 modelling the rainfall-runoff relationship.

19

20 For both NFM_S_SLM and NFM_S_SMAR models, using large number of HHCUs,
21 i.e. sub-models, did not improve the results significantly and this means there is an
22 upper limit for the number of HHCUs above which no significant improvement can
23 be expected. Thus using an excessive number of HHCUs might result in including
24 some redundant HHCUs which add little to the model's performance. Again this
25 behaviour is not surprising because the spatial parameters of the HHCUs have no

1 influence on the models. Different responses would be expected if some inputs to the
2 sub-models depended on the characteristics of HHCUs.

3

4 **7.2.3. NFM_T vs. NFM_S**

5 The important question arising out of the results for the two combination scenarios is
6 which combination NFM scenario performs best. To answer this requires a
7 comparison between the best models of the two scenarios. For illustration only, we do
8 this here for the Brosna catchment as the NFM_S was applied for that catchment only.
9 From Table 5 and 6 it is possible to identify the NFM_T_SMAR_4 (case 4) and
10 NFM_S_SMAR_4_14 (combination alternative 4 and HHCUs = 14) as the best
11 models for the two scenarios respectively in the Brosna catchment. The R^2 and *ARE*
12 results for these two models are not substantially different from each other. The fit
13 between the observed hydrograph and the simulated hydrographs for each model are
14 shown in Fig. 9 and they represent the same period used in Fig. 4. The visual
15 comparison between the observed hydrograph and the two models does not show any
16 major differences between models to the extent that one can be declared consistently
17 superior to the other. Thus the use of the NFM_S scenario, which requires more data
18 than the NFM_T scenario, is not justified if the intention of the modelling is to
19 produce outputs only for the outlet of a catchment.

20

21 **8. Conclusions**

22 In this study, the NFM has been proposed to account for spatial and temporal
23 variations in modeling the rainfall-runoff relationship. The proposed procedure was
24 implemented with two simple lumped models, SLM and SMAR. For each model two
25 scenarios (NFM_T and NFM_S) were used to construct sub-models to address the

1 temporal and spatial pattern variations respectively. In the NFM_T scenario, the two
2 models NFM_T_SLM and NFM_T_SMAR, were applied to eleven catchments from
3 around the world. A split sample technique was used and in most cases the neuro-
4 fuzzy combined sub-models were better than the lumped model. The NFM_T_SMAR
5 model was, in general, better than the NFM_T_SLM.

6

7 To address spatial variation in response, a subtractive clustering algorithm was used
8 in the NFM_S scenario to derive a number of HHCUs which exhibit homogenous
9 hydrologic responses. Three spatial layers representing DEM, land use and soil maps
10 of the Brosna catchment (Ireland) have been processed by a GIS software to prepare
11 data of four variables (elevation, slope index, generalised land use types, and soil
12 types) used in the clustering algorithm. For all possible combination alternatives
13 between the four variables the relation between the reject ratio parameter (RR) of the
14 subtractive clustering algorithm and the resulting number of HHCUs was
15 investigated. A remarkable improvement was achieved by the best case of the sub-
16 models of NFM_S_SMAR compared to the lumped model. The NFM_S_SMAR
17 model significantly outperformed the NFM_S_SLM and this is probably due to its
18 inclusion of non-linearity. Only a small number of HHCUs were required to obtain
19 improved results and using a larger number of HHCUs did not improve the results of
20 the NFM_S_SMAR model.

21

22 **9. Further work**

23 This work has shown that combinations of relative simple models can extend their
24 ability to model a range of catchment behaviour without requiring fully distributed
25 time-varying, physically-based models. While the combination approach has proved

1 useful in our Brosna catchment, it should be applied to other catchments with a wider
2 range of climatic variation and conditions to test its generality. In addition, it should
3 be possible to extend the approach to modelling other types of data, particularly water
4 quality time series, where the output at a single point is all that is required. In such
5 cases, the effort to generate and calibrate a physically-based distributed model may
6 not be justified and a calibrated combination of simple models may suffice. The
7 method can be used in all investigations that compare time-series or model the
8 relationship between two time-series, such as investigating tele-connections between
9 climate variables at different locations.

10

11 **Acknowledgement**

12 This study was carried out at the Centre for Water Resources Research at UCD, as
13 part of an Irish NDP/ERTDI project funded through the Environmental Research
14 Centre of the Environmental Protection Agency in Ireland.

15

16 **References**

- 17 Abrahart, R.J., See, L., 2002. Multi-model data fusion for river flow forecasting; an
18 evaluation of six alternative methods based on two contrasting catchments. *Hydrology
19 and Earth System Sciences* 6(4), 655-670.
- 20 Ahsan, M., O'Connor, K.M., 1994. A simple non-linear rainfall-runoff model with a
21 variable gain factor. *Journal of Hydrology* 155, 151-183.
- 22 Ajami, N.K., Gupta, H., Wagener, T., Sorooshian, S., 2004. Calibration of a semi-
23 distributed hydrologic model for streamflow estimation along a river system. *Journal
24 of Hydrology* 298,112-135.

1 Beven, K., 2000. Uniqueness of place and process representations in hydrological
2 modeling. *Hydrology and Earth System Sciences* 4(2), 203-213.

3 Beven, K., 1993. Prophecy, reality and uncertainty in distributed hydrological
4 modelling. *Advances in Water Resources* 16(1), 41-51.

5 Box, G.E.P., Jenkins, G.M., 1976. *Time Series Analysis: Forecasting and Control*.
6 Holden-Day, San Francisco, California, USA.

7 Bruen, M., Dooge, J.C.I., 1992, Unit hydrograph estimation with multiple events and
8 prior information: I Theory and a computer program. *Hydrological Sciences Journal*
9 37(5), 429-443.

10 Bruen, M., 1985. Black box modelling of catchment behaviour. Ph.D. thesis, National
11 University of Ireland, Dublin.

12 Chen, J., Adams, B.J., 2006. Integration of artificial neural networks with conceptual
13 models in rainfall-runoff modelling. *Journal of Hydrology* 318, 232-249.

14 Chen, S.H., Lin, Y.H., Chang, L.C., Chang, F.J., 2006. The strategy of building a
15 flood forecast model by neuro-fuzzy network. *Hydrological Processes* 20, 1525-1540.

16 Coulibaly, P., Hache, M., Fortin, V., Bobée, B., 2005. Improving Daily reservoir
17 inflow forecasts with model combination. *ASCE Journal of Hydrologic Engineering*
18 10(2), 91-99.

19 Fenicia, F., Solomatine, D.P., Savenije, H.H.G., Matgen, P., 2007. Soft combination
20 of local models in a multi-objective framework. *Hydrology Earth System Science*
21 *Discussions* 4, 91-123.

22 Ganoulis, J.G., 1994. *Engineering risk analysis of water pollution: Probabilities and*
23 *fuzzy sets*. VCH Verlagsgesellschaft mbH, Weinheim (Germany) and VCH
24 Publishers Inc. (USA).

1 Gupta, V.J., Sorooshian, S., 1983. Uniqueness and observability of conceptual
2 rainfall-runoff model parameters: the percolation process examined. *Water Resources*
3 *Research* 19(1), 269-276.

4 Holland, J.H., 1975. *Adaptation in natural and artificial systems*. University Michigan
5 Press, Ann Arbor.

6 Jacquin, A.P., Shamseldin, A.Y., 2006. Development of rainfall-runoff models using
7 Takagi-Sugeno fuzzy inference systems. *Journal of Hydrology* 329, 154-173.

8 Kachroo, R.K., Natale, L., 1992. Non-linear modeling of the rainfall-runoff
9 transformation. *Journal of Hydrology* 133, 341-369.

10 Khan, H., 1986. *Conceptual modelling of rainfall-runoff systems*, Unpublished Master
11 thesis submitted to the Department of Engineering Hydrology, National University of
12 Ireland, Galway.

13 Liang, G.C., 1992. A note on the revised SMAR model. Memorandum to the River
14 Flow Forecasting Workshop Group, Department of Engineering Hydrology, National
15 University of Ireland, Galway. (Unpublished).

16 Marechal, D., Holman, I.P., 2005. Development and application of a soil
17 classification-based conceptual catchment-scale hydrological model. *Journal of*
18 *Hydrology* 312, 277-293.

19 Nash, J.E., Foley, J.J., 1982. Linear models of rainfall-runoff systems. In: *Rainfall*
20 *runoff relationship*, Proceedings of the International Symposium on Rainfall-Runoff
21 modelling, Ed. V.P. Singh, Mississippi State University, Water Resources
22 Publications, USA, pp. 51-66.

23 Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models.
24 Part 1. A discussion of principles. *Journal Hydrology* 10, 282-290.

1 Nash, J.E., 1957. The form of the instantaneous unit hydrograph. IAHS Publication
2 45, 114-127.

3 O'Connell, P.E., Nash, J.E., Farrell, J.P., 1970. River flow forecasting through
4 conceptual models. Part II: The Brosna catchment at Ferbane. Journal of Hydrology
5 10, 317-329.

6 Ozelkan, E. C., Duckstein, L., 2001. Fuzzy conceptual rainfall-runoff model. Journal
7 of Hydrology 253, 41-68.

8 Quiroga, C.A., Singh, V.P., Lam, N., 1996. Land-use hydrology. In: Geographical
9 Information Systems in Hydrology, Eds. V.P. Singh, and M. Fiorentino, Kluwer
10 Academic Publishers, The Netherlands, pp. 389-414.

11 See, L., Openshaw, S., 2000. A hybrid multi-model approach to river level
12 forecasting. Hydrological Sciences Journal 45(4), 523-536.

13 See, L., Openshaw, S., 1999. Applying soft computing approach to river level
14 forecasting. Hydrological Sciences Journal 44(5), 763-778.

15 Shamseldin, A.Y., O'Connor, K.M., Liang, G.C., 1997. Methods for Combining the
16 Outputs of Different Rainfall-Runoff Models. Journal of Hydrology 197, 203-229.

17 Shamseldin, A.Y., O'Connor, K.M., 1996. A nearest neighbour linear perturbation
18 model for river flow forecasting. Journal of Hydrology 179, 353-375.

19 Sugawara, M., 1995. Tank model. In: Singh, V.P. (Ed.), in Computer Models of
20 Watershed Hydrology. Water Resources Publications, Littleton, CO.

21 Takagi, T., Sugeno, M., 1985. Fuzzy identification systems and its application to
22 modelling and control. IEEE Trans. Sys. Man. Cybern. 15, 116-132.

23 Tan, B.Q., O'Connor, K.M., 1996. Application of an empirical infiltration equation in
24 the SMAR conceptual model. Journal of Hydrology 185, 275-295.

1 Todini, E., Wallis, J.R., 1977. Using CLS for daily or longer period rainfall-runoff
2 modelling. In Ciriani, T.A., Maione, U., and Wallis, J.R. (Eds.), Mathematical models
3 for surface water hydrology. (Proc. of the 1974 Workshop held at the IBM Centre,
4 Pisa, Italy.) John Wiley & Sons, London , pp 149-168.

5 Vernieuwe, H., Georgieva, O., De Baets, B., Pauwels, V.R.N., Verhoest, N.E.C., De
6 Troch, F.P., 2005. Comparison of data-driven Takagi-Sugeno models of rainfall-
7 discharge dynamics. Journal of Hydrology 302,173-186.

8 Xiong, L., Shamseldin, A.Y., O'Connor, K.M., 2001. A non-linear combination of the
9 forecasts of rainfall-runoff models by the first-order Takagi-Sugeno fuzzy system.
10 Journal of Hydrol. 245, 196-217.

11 Zhao, R.J., and Liu, X.R., 1995. The Xinanjiang model. In: Singh, V.P. (Ed.),
12 Computer Models of Watershed Hydrology. Water Resources Publications, Littleton,
13 CO.

14
15
16
17
18
19
20
21
22
23
24
25

1 **Figure Captions**

2 **Figure 1. Structure of quasi-linear model proposed by Bruen (1985)**

3

4 **Figure 2. Sub-models combination using fuzzy logic principle of membership**
5 **function**

6

7 **Figure 3. NFM architecture**

8

9 **Figure 4. Number of clusters vs. reject ratio (RR) for all combination**
10 **alternatives used in the subtractive clustering algorithm**

11

12 **Figure 5. Simulated and observed hydrographs of the Blue Nile catchment**

13

14 **Figure 6. Simulated and observed hydrographs of the Brosna catchment**

15

16 **Figure 7. Simulated and observed hydrographs of the Chu catchment**

17

18 **Figure 8. Simulated and observed hydrographs of the Wolombi Brook catchment**

19

20 **Figure 9. Comparison between NFM_T and NFM_S best models in the Brosna**
21 **catchment**

22

23

24

25

1 **Tables**

2 **Table 1. Cases representing the rainfall and evaporation fuzzy sub-sets**

3 **combination for the NFM_T_SLM and NFM_T_SMAR**

Model	Case*	No of fuzzy sub-sets	
		Rainfall	evaporation
NFM_T_SLM_*, NFM_T_SMAR_*	1	1	1
	2	1	2
	3	2	1
	4	2	2
	5	3	1
	6	3	2
	7	3	3
	8	4	4
	9	5	5
	10	6	6

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

1 **Table 2. Details of the test catchments**

Catchment name	Country	Area (km²)	Starting date of data	No. of data points	Memory length (day)
Bird Creek	USA	2344	1-Jan.-1955	2922	15
Blue Nile	Sudan	175125	1-Jan.-1992	1461	15
Brosna	Ireland	1207	1-Jan.-1969	3652	30
Chu	Vietnam	2370	1-Jan.-1965	3652	15
Halda	Bangladesh	779	1-Jan.-1980	2556	15
Kelantan	Malaysia	12867	1-Jan.-1975	2922	20
Nan	Thailand	4609	1-Jan.-1978	3287	20
Sg. Bernam	Malaysia	1090	1-Jan.-1977	2556	25
Shiquan-3	China	3092	1-Jan.-1973	2922	15
Sunkosi-1	Nepal	18000	1-Jan.-1975	2922	30
Wolombi Brook	Australia	1580	1-Jan.-1963	1826	15

2
3
4
5
6
7
8
9
10
11
12
13
14
15

1 **Table 3. Combination alternatives of the four spatial variables used in the**
 2 **subtractive clustering algorithm**

id.	No of variables	Variables
2A	2	Elevation + Land use
2B	2	Elevation + Soil
2C	2	Slope index + Land use
2D	2	Slope index + Soil
2E	2	Land use + Soil
3A	3	Elevation + Land use + Soil
3B	3	Slope index + Land use + Soil
4	4	Elevation + Slope index + Land use + Soil

3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21

- 1 **Table 4. Description of the NFM_SPT_SLM and NFM_SPT_SMAR cases tested**
- 2 **in the Brosna catchment for each combination alternative**

Model	Case*	No of HHCUs
NFM_S_SLM_2A_*, NFM_S_SMAR_2A_*	1	1
	2	9
	3	12
	4	29
	5	36
	6	37
NFM_S_SLM_2B_*, NFM_S_SMAR_2B_*	1	1
	2	32
NFM_S_SLM_2C_*, NFM_S_SMAR_2C_*	1	1
	2	10
NFM_S_SLM_2D_*, NFM_S_SMAR_2D_*	1	1
	2	2
	3	3
	4	4
	5	15
NFM_S_SLM_2E_*, NFM_S_SMAR_2E_*	1	1
	2	4
	3	5
	4	7
	5	20
NFM_S_SLM_3A_*, NFM_S_SMAR_3A_*	1	1
	2	23
NFM_S_SLM_3B_*, NFM_S_SMAR_3B_*	1	1
	2	5
	3	6
	4	8
	5	37
NFM_S_SLM_4_*, NFM_S_SMAR_4_*	1	1
	2	14

3

Table 5. R^2 and ARE results for the lumped case and the best combined case of the NFM_T_SLM and NFM_T_SMAR models in the eleven catchments

Model	Catchment	Best case	R^2				ARE			
			Lumped case		Best combined case		Lumped case		Best combined case	
			Calib.	Valid.	Calib.	Valid.	Calib.	Valid.	Calib.	Valid.
NFM_T_SLM	Bird Creek	9	15.86	23.34	24.37	42.54	71.99	81.57	72.99	79.31
	Blue Nile	10	71.69	71.38	87.44	77.82	28.94	22.45	25.43	24.08
	Brosna	8	49.36	32.36	60.24	40.53	29.92	34.37	28.35	33.27
	Chu	9	15.23	29.13	39.27	56.95	57.64	55.73	55.01	52.72
	Halda	8	53.43	69.84	67.63	67.78	38.95	49.57	38.60	52.75
	Kelantan	10	28.88	22.78	52.39	34.78	33.16	29.25	26.79	30.45
	Nan	10	65.29	68.94	69.57	69.54	39.05	33.69	39.98	44.46
	Sg. Bernam	8	60.35	52.14	62.73	47.90	24.88	26.77	27.00	31.47
	Shiquan-3	8	13.45	6.32	28.33	24.40	54.16	49.95	51.76	49.80
	Sunkosi-1	9	77.80	78.78	80.73	82.10	27.86	25.95	27.35	23.55
	Wolombi	9	10.27	-17.03	30.03	17.31	80.33	71.88	71.60	92.62
NFM_T_SMAR	Bird Creek	7	85.85	66.58	89.72	75.27	67.70	60.50	67.45	63.73
	Blue Nile	4	93.26	83.00	94.57	86.53	17.37	29.65	16.22	24.28
	Brosna	4	87.93	83.86	89.81	86.18	15.66	19.17	15.37	18.01
	Chu	10	35.30	43.20	81.46	64.72	69.68	70.52	59.56	65.51
	Halda	10	62.42	69.56	83.76	68.82	34.45	43.35	33.26	54.25
	Kelantan	5	84.67	47.70	87.26	46.81	20.06	27.70	19.70	27.63
	Nan	7	76.36	80.48	83.88	80.70	34.71	26.29	33.29	27.30
	Sg. Bernam	4	73.51	21.49	76.40	5.93	23.38	43.05	23.22	46.19
	Shiquan-3	5	19.69	17.96	23.32	17.24	78.37	83.93	75.46	79.82
	Sunkosi-1	5	80.49	79.90	82.78	80.28	26.57	24.93	25.64	25.56
	Wolombi	10	34.74	-33.82	89.15	58.39	70.41	108.54	59.11	112.46

Table 6. R^2 and ARE results for the lumped case and the best combined case for all the combination alternatives of the NFM_S_SLM and NFM_S_SMAR models in the Brosna catchment

Model	id.	case	No of HHCUs	R^2		ARE	
				Calib.	Valid.	Calib.	Valid.
NFM_S_SLM	2A,2B,...,4	1	1 (lumped model)	49.36	32.36	29.92	34.37
	2A	2	9	50.18	32.68	29.47	33.98
	2B	2	32	48.87	31.04	30.26	34.71
	2C	2	10	50.07	32.65	29.54	33.99
	2D	3	3	50.44	32.93	29.15	33.76
	2E	3	5	50.42	32.92	29.29	33.90
	3A	2	23	49.26	31.75	30.29	34.75
	3B	2	5	50.22	32.95	29.34	33.90
	4	2	14	49.94	32.35	29.81	34.32
NFM_S_SMAR	2A,2B,...,4	1	1 (lumped model)	87.96	84.18	15.44	18.60
	2A	6	37	91.17	87.91	13.68	16.38
	2B	2	32	90.25	86.50	14.08	16.67
	2C	2	10	90.31	82.90	14.19	19.64
	2D	4	4	91.28	85.82	13.63	17.86
	2E	4	7	91.16	84.53	13.86	19.23
	3A	2	23	90.67	85.68	14.44	18.03
	3B	3	6	91.23	85.59	13.99	17.98
	4	2	14	91.42	86.00	13.47	17.57