Meng-Jung, T. and Abrahart, R.J. and Mount, Nick J. and Chang, F.-J. (2014) Including spatial distribution in a data-driven rainfall-runoff model to improve reservoir inflow forecasting in Taiwan. Hydrological Processes, 28 (3). pp. 1055-1070. ISSN 1099-1085

Including spatial distribution in a data-driven rainfall-runoff model to improve reservoir inflow forecasting in Taiwan

Journal:	*Hydrological Processes*
Manuscript ID:	HYP-12-0116.R2
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Tsai, Meng-Jung; National Taiwan University, Department of Bioenvironmental Systems Engineering; University of Nottingham, School of Geography Abrahart, Robert; University of Nottingham, School of Geography Mount, Nick; University of Nottingham, Geography Chang, Fi-John; National Taiwan University, Bioenvironmental systems engineering
Keywords:	semi-distributed model, rainfall-runoff model, data-driven model, reservoir inflow, radar rainfall, ANFIS

Including spatial distribution in a data-driven rainfall-runoff model to improve reservoir inflow forecasting in Taiwan

Meng-Jung Tsai[1,2], Robert J. Abrahart[2], Nick J. Mount[2], Fi-John Chang[1,*]

[1]*Department of Bioenvironmental Systems Engineering, National Taiwan University, Taipei 10617, Taiwan, R.O.C.*

[2]*School of Geography, University of Nottingham, Nottingham, NG7 2RD, UK*

** corresponding author: Fi-John Chang (E-mail: changfj@ntu.edu.tw)*

Abstract

Multi-step ahead inflow forecasting has a critical role to play in reservoir operation and management in Taiwan during typhoons as statutory legislation requires a minimum of 3-hours warning to be issued before any reservoir releases are made. However, the complex spatial and temporal heterogeneity of typhoon rainfall, coupled with a remote and mountainous physiographic context makes the development of real-time rainfall-runoff models that can accurately predict reservoir inflow several hours ahead of time challenging. Consequently, there is an urgent, operational requirement for models that can enhance reservoir inflow prediction at forecast horizons of more than 3-hours. In this paper we develop a novel semi-distributed, data-driven, rainfall-runoff model for the Shihmen catchment, north Taiwan. A suite of Adaptive Network-based Fuzzy Inference System solutions is created using various combinations of auto-regressive, spatially-lumped radar and point-based rain gauge predictors. Different levels of spatially-aggregated radar-derived rainfall data are used to generate 4, 8 and 12 sub-catchment input drivers. In general, the semi-distributed radar rainfall models outperform their less complex counterparts in predictions of reservoir inflow at lead-times greater than 3-hours. Performance is found to be optimal when spatial aggregation is restricted to 4 sub-catchments, with up to 30% improvements in the performance over lumped and point-based models being evident at 5-hour lead times. The potential benefits of applying semi-distributed, data-driven models in reservoir inflow modelling specifically, and hydrological modelling more generally, is thus demonstrated.

Keywords: semi-distributed model, rainfall-runoff model, data-driven model, reservoir inflow, radar rainfall, ANFIS

1. INTRODUCTION

Data-driven modelling (Solomatine, 2005; Solomatine and Ostfeld, 2008; Elshorbagy et al., 2010a,b) is a major component of hydroinformatics (Abbott, 1991; 1999; See et al., 2007; Abrahart et al., 2008; Holz et al., 2011), in which emerging technological products, primarily related to developments in computational intelligence and machine learning algorithms, are applied to complex hydrological problems. In data-driven modelling, the responsibility for identifying model structure is largely passed to computer algorithms, that are not constrained by a need for their solutions to conform to fundamental concepts in hydrology (Mount and Abrahart, 2011a). However, stronger calls for greater incorporation of scientific knowledge and understanding in the development of data-driven hydrological models are now starting to be published (e.g. Abrahart et al., 2011) in which it is argued that better representation of catchment processes should result in improved data-driven modelling products that offer more than optimized curve fitting solutions (Mount and Abrahart, 2012).

One of the most popular uses of data-driven models is the prediction of runoff from rainfall e.g. by means of developing a neural network (De Vos, 2012), fuzzy logic (Wang and Altunkaynak, 2012) or genetic programming (Rodríguez-Vázquez et al., 2012) solution that will effectively convert observed input into required output. In this application domain, the importance of capturing the spatial variability of rainfall-runoff processes via distributed and semi-distributed hydrological modelling frameworks is well-known (Beven and O'Connell, 1982; Tetzlaff and Uhlenbrook, 2005; Segond et al., 2007; Younger et al., 2009). The need to account for spatial distribution becomes particularly acute if convective rainfall, with a high degree of spatio-temporal heterogeneity and uncertainty, is the key driver of runoff (e.g. typhoon events). In this context, instrumentational deficiencies (e.g. Molini et al., 2005) and the number, positioning and overall distribution of monitoring stations (e.g. Cheng et al., 2008) present significant challenges for modellers. Indeed, hydrologists have long claimed that the main factor limiting predictive performance of distributed rainfall-runoff models was interpolated input derived from point-based rain gauge data (Berne & Krajewski, 2012). Whilst traditional, Thiessien Polygon (Thiessen, 1911) approaches remain popular (e.g. Rajurkar, et al., 2002; Wu and Chau, 2011), the errors that result are well known (American Society of Civil Engineers, 1996: 50). Thus, better approaches to interpolation of point-based, rain gauge data continue to receive a modest amount of interest from the scientific community e.g. Schiemann et al. (2010), Verworn and Haberlandt (2011), Ly et al. (2011), Wagner et al. (in press). In recent years the availability of

spatially-continuous radar rainfall data has led to its widespread utilisation by hydrologists and inclusion in particular sorts of model (e.g. Schell et al., 1992; James et al., 1993; Georgakakos et al., 1996; Bell and Moore, 1998; Vieux and Bedient, 1998; Winchell et al., 1998; Sempere-Torres et al., 1999; Borga et al., 2000; Ogden et al., 2000). The use of radar is especially useful in catchments containing coarse rain gauge networks, producing process-realistic distributed runoff simulations (e.g. Michaud and Sorooshian, 1994; Lange et al., 1999; Woods et al., 2000). Radar rainfall estimation is, nevertheless, also subject to a range of errors caused by factors that include instrumentation issues (e.g. calibration, measurement noise) and complexity and variability in the relationship occurring between recorded measurements and precipitation parameters (Austin, 1987; Joss and Lee, 1995; Andrieu et al., 1997; Borga et al., 2002). These different sources of error act to compound the radar rainfall uncertainty and can have a significant impact on the accuracy of rainfall–runoff forecasting (Borga, 2002).

The argument that data-driven modellers should take greater account of hydrological processes representation, has resulted in a number of recent studies that have attempted to adapt standard data-driven modelling approaches, so that a degree of relevant knowledge about hydrological processes and data uncertainty is better represented in the model structure (e.g. Chen and Adams, 2006; Corzo and Solomatine, 2007; Corzo et al., 2009; Song et al., in press). Nonetheless, the majority of data-driven rainfall-runoff models continue to emphasise temporal variation in hydrological processes, and largely ignore the impact of spatial variation in the model inputs. The result is a dominance of spatially-lumped data-driven studies, and this is especially true in the case of rainfall-runoff models (e.g. Nayak et al., 2005; Chiang and Chang, 2009; Wu and Chau, 2011; De Vos, 2012). Despite its obvious potential as a means by which spatial variation can be captured and incorporated, few data-driven modelling studies have attempted to use raster-based radar rainfall (e.g. Teschl and Randeu, 2006; Teschl et al., 2009; Chaipimonplin et al., 2010) or satellite rainfall (e.g. Akhtar et al., 2009) inputs. One reason for this is that each grid cell ultimately represents a separate potential input for a data-driven model, such that utilising raster data in its raw form is impractical and inefficient. Instead, a trade-off is required in which an optimised spatio-temporal lumping strategy is applied to any continuous rainfall data; resulting in a semi-distributed data-driven modelling framework.

The aim of this paper is to enhance our understanding of the extent to which greater incorporation of fundamental hydrological process knowledge, can deliver superior forecast performance for complex hydrological phenomenon. To this end, our core

objective is to exemplify how different levels of semi-distribution, applied to continuous radar rainfall data inputs, operating in a data-driven rainfall-runoff modelling framework, affect the performance of multi-step-ahead reservoir inflow forecasts in Taiwan. Taiwan represents an excellent study site because it experiences conditions of extreme spatial and temporal rainfall heterogeneity associated with typhoon events. In this paper the Adaptive Network-based Fuzzy Inference System (ANFIS: Jang, 1993; Jang et al., 1997) is used to forecast reservoir inflow for a range of different forecasting horizons. We develop a set of models based on different levels of radar rainfall spatial disaggregation, from which a model with the preferred level of input distribution is identified. Results are compared against counterpart solutions developed on point-based rainfall inputs from traditional rain gauges. Justification of the application of a complex, non-linear data-driven modelling algorithm is made by means of multiple linear regression benchmarks.

2. RESERVOIR INFLOW FORECASTING IN TAIWAN

Taiwan is situated within the main track of western North Pacific typhoons. In an average year, Taiwan experiences between four and five typhoons that occur between June and November. There is little consistency in the direction of the typhoon paths (Figure 1: Taiwan Central Weather Bureau, pers. comm.), primarily due to Taiwan being located at a turning point on the track for most typhoons occurring in the Western North Pacific–East Asian region (Camargo et al., 2007; Chu et al., 2012); but also, potentially, on account of long term trends or low-frequency, large scale atmospheric shifts relating to climate change (Kao et al., 2012; Lee et al., in press). Significant variation exists in the individual strength of a particular typhoon and the speed and direction of motion of its track. These factors control the intensity, spatial distribution and total volume of rainfall it delivers (Lee at al., 2006; Pan et al., 2012). Thus typhoons are characterised by substantial spatio-temporal heterogeneity, which means that the spatial and temporal distribution of typhoon-associated heavy rainfall will be highly complex and differ from event to event. The impact of such rainfall, moreover, will in all likelihood be exacerbated as a result of climate change which is expected to deliver increased typhoon frequencies (Lee and You, 2011).

Taiwan's topography is characterized by a mountainous north-south trending central belt, with steep-slopes and short, fast flowing rivers. Several upland rivers have been dammed to form reservoirs, which can be inundated within a few hours as a

result of extreme runoff events linked to the passage of typhoons (Chang et al., 2002). The inflow to these reservoirs is mainly the result of localised rainfall (Yu et al., 2006), with only limited contributions from groundwater; thus rainfall-runoff models represent an important means by which streamflow is predicted (e.g. Vieux et al., 2003; Yu et al., 2004; Wu et al., 2007).

During a typhoon, important decisions about the timing and amount of any required reservoir releases must be made within a matter of hours by the controlling agency. Most reservoir operations are guided by simultaneously balancing flood control and water supply. In general, reservoir operations for flood control can be separated into three different stages (Hsu and Wei, 2007): (1) stage prior to flood arrival, in which water releases are designed to reserve sufficient reservoir capacity for the upcoming flood; (2) stage preceding peak inflow, in which floodwater releases are applied for disaster mitigation; and (3) stage after peak inflow, in which reservoir releases are used to regulate the storage at the end of each flood, for future water purposes. The correct operation of these stages involves the use of look-up tables, which provide rules for the standardised release of water during typhoon periods. These tables are graded by total forecasted rainfall, observed storage level, and reservoir inflow during flood periods.

The implication of current practice in reservoir decision-making procedures is that for operational management purposes an advanced knowledge of flood peak magnitude and timing is required, empowering the controller to select an appropriate course of action. Moreover, armed with such information, it becomes possible for the controlling agency to deliver an appropriate set of statutory warnings to mass/local media, pertinent institutions and downstream residents several hours in advance of any proposed water releases. Thus, in balancing the reservoir, timing is a critical factor and a rainfall-runoff model that can deliver accurate, real-time predictions for forecasting horizons that exceed three or more hours ahead of present is called for. This requires the best possible short-term multi-step-ahead reservoir inflow forecasting model; one that will eventually form an integrated and trustworthy component of the reservoir management and operational decision-making process.

2.1. Short-term, multi-step ahead reservoir inflow forecasting

Generating short period ahead forecasts of reservoir inflow by means of standard, spatially-distributed physical models, applied in real time (e.g. Wu et al., 2007; Wu et al., 2008), is problematic because the generation of ahead-of-time forecasts

necessitates the use of uncertain, forecasted inputs in a modelling framework that is designed to be instantaneous and continuously-updating. Moreover, the complex data needs and parameterisation requirements of physically-based models make their application difficult in many of Taiwan's reservoirs that are fed from remote, mountainous, catchments (Wu et al., 2008). An alternative approach is multi-step-ahead forecasting in which the general relationship between lagged and instantaneous inputs, and an inflow record shifted progressively forward in time, is quantified and reapplied (e.g. Chang et al., 2007; Toth and Brath, 2007; Yonaba et al., 2010). In this context, a forecast is generated for multiple periods (or steps) ahead using a mix of past records and real-time measurements. Data driven models are particularly good at multi-step-ahead forecasting due to their ability to determine the optimal relationships that relate inputs to outputs; albeit with a reduction in their predictive capabilities over longer forecast horizons (Campolo et al., 1999; Babovic and Keijzer, 2002; Nayak et al., 2005; Xu and Li, 2002; Dawson et al., 2006). They are also flexible enough to enable the development of models that accommodate spatial and temporal heterogeneity in the model inputs (Lorrai and Sechi, 1995; Rajurkar et al., 2002).

From the range of different data-driven rainfall-runoff modelling and streamflow forecasting investigations that have been published, ANFIS has emerged as a particularly promising method of dealing with complex time series data sets, following its first reported application to hydrological modelling problems by Gautam and Holz (2001) e.g. Keskin et al. (2006); Bae et al. (2007); Firat and Güngor (2008); Pramanik and Panda (2009); Talei et al. (2010); Nguyen and Chau (2012); Ghalkhani et al. (in press). Of particular importance for this study is recently-reported success in its ability to discover the optimal relationship between current and antecedent rainfall and reservoir inflow inputs and multi-step-ahead reservoir inflow outputs (El-Shafie, 2007; Jothiprakash and Magar, 2012). However, any implementation of a distributed, ANFIS modelling framework is potentially problematic, because the use of large numbers of distributed inputs is likely to result in a grossly inefficient solution. In addition, local noise (both spatial and temporal) in the rainfall data which is used to drive such models risks masking the broader hydrological signal that one wishes to capture (Lin and Chen, 2005; Dark and Bram, 2007). Therefore, in developing a semi-distributed multi-step-ahead ANFIS rainfall-runoff model, it is essential that some preferred level of input spatial aggregation is identified which both maximises the strength of the hydrological signal that can be modelled, and minimises noise due to input uncertainty and model inefficiency.

3. STUDY AREA AND DATASETS

The Shihmen Reservoir (Figure 2) was built in 1964 and is one of the largest and most important reservoirs in Taiwan. It has been the focus of a number of previous studies that have modelled the catchment hydrology for the purpose of inflow prediction. Recent examples include the development and application a physically-based distributed parameter model (Wu et al., 2008) and two different sorts of neural network model (Chen and Chang, 2009; Lin et al., 2009a). Located on the upper reaches of the Tahan River, it has an upstream contributing area of 763.4 km^2, ranging in elevation from 157 m to 3514 m, with average slope angles of about 30 degrees. It has an effective capacity of 219 million cubic metres and is designed for multiple purposes including water supply for irrigation, industrial and domestic uses, flood control, and hydropower generation. The reservoir is currently managed by the Water Resources Agency which stipulates rules for operational flood control (Water Resources Agency, 1984).

The rainfall and inflow datasets used in this study were provided by the Water Resources Agency and Central Weather Bureau of Taiwan and used as supplied. In common with standard practice in data-driven modelling studies, additional preprocessing was not performed (Abrahart et al., 2010). The use of third-party data inhibited meaningful adjustment or correction of observed records for original measurement error, and detailed information on data collection practices or data quality control procedures were not made available. Similarly, rain gauge point-based rainfall records and radar grid-based rainfall records were used as supplied, without adjustments for elevation. The reservoir inflow series comprised 445 hourly observations for eight typhoon events, occurring between 2007 and 2009, in which peak inflow per event across the series ranged from a maximum of 5300 $m^3 s^{-1}$ for KROSA (Event 2: SSHS Category 4 Super Typhoon) to a minimum of 203 $m^3 s^{-1}$ for KALMAEGI (Event 3: SSHS Category 2 Moderate Typhoon that delivered very little rainfall in our catchment) (Table 1). Figure 3 shows the reservoir inflow records depicted as a continuous series for the eight typhoons, highlighting substantial differences in the duration and magnitude of each individual storm event. Two different sorts of rainfall data were available for the eight typhoon events: hourly rainfall data for 12 gauges distributed across the catchment (Figure 2); and a corresponding radar rainfall data set produced from QPESUMS (Quantitative Precipitation Estimation and Segregation Using Multiple Sensors: http://www.nssl.noaa.gov/projects/qpesums/). Full particulars are provided in

Table 2. The radar rainfall data had a temporal resolution of 10 minutes and a spatial resolution of 1.25 km. This raster data set had already been calibrated, using ground observations, by the Central Weather Bureau, Taiwan. The 10-minute precipitation map was temporally accumulated into an hourly sequence, corresponding to our hourly rain gauge data and hourly reservoir inflow series. The study also utilised a 40 m resolution digital elevation model (DEM) provided by the National Land Surveying and Mapping Center, Taiwan for sub-catchment segmentation analysis. The DEM was constructed from stereo-pair imagery in 1995 by the Aerial Survey Office, Taiwan and has a vertical accuracy of between 2.5 m and 5.0 m.

4. METHODOLOGY

Within the context of the existing rules that govern reservoir releases in Taiwan, we develop modelling procedures that are capable of generating reliable estimates of peak reservoir inflow magnitude and timing over intervals that exceed the statutory minimum requirement. It is important to note that we do not exemplify the application of the operational procedures or seek to improve or change them. Instead, we focus on the development of a model that can deliver improved forecast information, which can subsequently be used to support better application of the existing operational rule sets. The purpose of the modelling exercise is thus to generate the best performing real-time, instantaneous multi-hour step-ahead forecast of reservoir inflow for the reservoir in question. From this, the predicted water accumulation within the reservoir can be continuously updated, and decisions about the timing of any required reservoir release(s) can be made in advance of their occurrence. To this end, the model predictions are not constrained to peak inflows, but extend across the entire hydrologic response of the catchment.

Three stages of model development (Stages 1-3), which are the main focus of this paper, are required before a real-time multi-step ahead model can be applied in an operational capacity (Stage 4):

1. The collation and pre-processing of available rainfall and discharge records for the catchment (see Section 3.1);
2. The construction of numerous multi-step-ahead data-driven models (listed below) using a range of different inputs and different levels of spatial aggregation (see Section 3.2);
3. The identification of a model structure that is best able to deliver step-ahead inflow forecasts over increasingly large time steps (according to both level of

spatial aggregation employed and other inputs) (see Section 3.3);

4. The re-application of our preferred model structure (identified at Stage 3).

Inflow into the Shihmen reservoir was modelled for forecast horizons of Q_{t+1},Q_{t+2}...Q_{t+5} hours ahead using six modelling approaches of increasing complexity (Table 3):

Model A: A simple two-input model based solely on the reservoir inflow record. Inputs comprised inflow at time t, and the change (difference) in inflow between Q_t and Q_{t-1}, hereafter termed ΔQ. This model represents the minimum complexity solution against which models using either rainfall or radar as additional inputs can be compared.

Model B: A rainfall-runoff model with inputs comprising 12 lagged, point-based rain gauge records, Q_t and ΔQ. This model represents a standard configuration for including rain gauge inputs in data-driven rainfall-runoff models (e.g. Deo and Thirumalaiah, 2000).

Model C: A lumped rainfall-runoff model with inputs comprising lagged total rainfall, derived from radar data and spatially averaged across the entire catchment, Q_t and ΔQ. This model represents the simplest application of spatial lumping, and provides a baseline against which the additional performance of semi-distributed modelling configurations can be assessed.

Models D-F: A suite of semi-distributed rainfall-runoff models (ranging from 4 to 12 sub catchments), with inputs comprising lagged total rainfall, derived from radar data and spatially averaged across each sub catchment, Q_t and ΔQ. This sequential partitioning provides specific insight into the relationship between different levels of spatial aggregation and data-driven model prediction accuracy.

In accordance with past recommendations (Abrahart and See, 2007; Mount and Abrahart, 2011b), each ANFIS instantiation of the four approaches was also benchmarked against a set of counterpart linear regression models. In this way the additional benefit of employing complex, ANFIS-based modelling could be determined. In contrast to several earlier ANFIS papers, we do not provide additional neural network model benchmarks; for example, by means of developing and reporting a standard counterpart backpropagation-of-error trained feedforward

neural network (FFNN: Abrahart and See, 2000). Following Chang and Chang (2006), who also only applied an ANFIS solution to their successful data-driven modelling of reservoir water level at this location, we direct our focus away from algorithm comparison for reservoir inflow forecasting as many other papers have already addressed such issues, (e.g. Chang et al., 2007; Lin et al., 2009b; Karimi-Googhari and Lee, 2011). Instead we focus on examining the benefits that result from undertaking data-driven rainfall-runoff modelling in a spatially-disaggregated manner. Two further justifications for our decision to omit the additional reporting of simpler-structured FFNNs should also be documented. First, where ANFIS and FFNN model counterparts have previously been compared in the context of step-ahead hydrological forecasting studies, the performance of ANFIS solutions is consistently equal to, or slightly in excess of its simpler FFNN model counterparts (e.g. Chau et al., 2005; Nayak et al., 2005; Chen et al., 2006; Mukerji et al., 2009; Lohani et al., 2012). There is no reason to presume that the findings of our current study would be significantly different from that of previous work. Second, delivering routine comparisons, in which alternative categories of neural algorithm are directly matched one against another, can at best deliver only 'incremental refinement' of existing expertise and/or scientific knowledge (Abrahart et al., 2012); potentially offering only marginal improvement in accuracy, low intellectual reward and no step-change in hydrological modelling understanding or application development.

4.1. Derivation of model inputs

4.1.1. Inflow Inputs

Numerous modelling studies have shown that, over short forecast horizons, simple one-step-ahead autoregressive models will provide good predictions of inflow over a broad range of different hydrological settings (Niedzielski, 2007). However, one-step-ahead models can suffer from local preferencing; where the predictive power of all lagged inputs is minimised in favour of the last observed record, if the latter is included as an input (Abrahart et al., 2007). The upshot of this for data-driven modelling is that one-step-ahead solutions can easily become trapped into producing a minimally-modified autoregressive single-input single-output model. To reduce the likelihood of this occurring, the lagged inflow input in each of our models is modified into a standardised rate of change: ΔQ. This increases the dimensionality of our drivers and, simultaneously, ensures that both positive and negative values are provided. This deviation from standard data-driven modelling practice is intended to reduce the marginalisation of lagged data during model

training and offer greater potential for non-current inputs to influence the model output. Moreover, if the dominant impact of inflow is countered, the significance of rainfall records should be accentuated and might logically be expected to deliver a reduction in timing error.

4.1.2. Rain Gauge Inputs

In contrast to the simple inflow model, the use of rain gauge inputs requires consideration of the spatial distribution of measurement records, and of variable travel times occurring between the rainfall recorded at the gauge and the inflow response at the point of reservoir inflow forecast. Therefore, the challenge is to identify the most representative travel time for each gauge. In this study, we examine all travel times from 0 to 10 hours by means of correlation analysis: standard practice in data-driven hydrological modelling (Maier and Dandy, 2000; Maier et al., 2010). However, variability in the typhoon tracks presents additional complexity which will result in inconsistent lag response times between each gauge and the reservoir inflow during different events. Thus, a three-stage combination and selection process was adopted:

1. For each of the 12 gauges, compute a correlation coefficient for each of 8 individual typhoon events at each of the 11 reservoir inflow travel times (i.e. calculate 1056 individual correlation coefficients);
2. For each of the 11 travel times at each of the 12 gauges, compute a mean correlation coefficient that spans 8 typhoon events (i.e. convert 1056 individual correlation coefficients into 132 mean correlation coefficients).
3. For each gauge, select the travel time to be used based on the maximum mean correlation coefficient.

The results of Steps 2 and 3 are presented in Table 4.

4.1.3. Radar Inputs

The use of radar data allows us to replace spatially-discrete rain gauge modelling inputs with spatially-continuous representations. In total, 434 grid cells comprise the radar data set for this catchment. In order to prevent the development of an excessively complex solution, spatial lumping was performed by aggregating radar values to increasing numbers of hydrological sub-units. The original catchment was partitioned into sub-catchment polygons, based on different levels of stream segmentation, according to the DEM method of Jensen and Domingue (1988) - as

implemented in Arc Hydro (Maidment, 2002). Flow accumulation thresholds of T=100; T=30 and T=25 resulted in 4, 8 and 12 sub-catchment discretisations; with 12 being equal to the number of rain gauges in Model B. Radar data values were subsequently assigned to the full catchment and each of its sub-catchment polygons according to whether or not the centroid of a particular radar cell fell inside its boundary. The total hourly rainfall for each polygon was thereafter calculated by means of summation, resulting in a set of instantaneous rainfall inputs at four different levels of spatial lumping [1 (Model C), 4 (Model D), 8 (Model E) and 12 (Model F)]. Table 5 contains descriptive statistics of the aggregated radar rainfall data for each individual polygon.

For hourly radar rainfall, correlation analysis was performed on the aggregated radar rainfall data for each polygon in a manner identical to that for gauged rainfall, by again examining travel times that ranged from 0 to 10 hours. Mean correlation coefficients were calculated across the 8 typhoon events, with the representative travel time for each polygon selected on the basis of the maximum mean coefficient. The travel times assigned to each polygon are presented in Figure 4. At 4 and 8 sub catchments, the spatial assignment of travel times appears rational, with travel time increasing with distance from the reservoir inflow. At 12 sub catchments, the spatial pattern is less rational, with instances of upstream sub catchments being assigned quicker travel times than some of their downstream neighbours.

4.2. Data-driven Modelling

ANFIS models were developed in MATLAB for our five forecasting horizons (Q_{t+1},Q_{t+2}...Q_{t+5}) using each of the six approaches (A-F), resulting in 30 final models. The Fuzzy Logic Toolbox *genfis3* function was used to determine the structure of each fuzzy information system, by the application of fuzzy c-means. The number of clusters required by the c-means algorithm was provided by the user and, in turn, this parameter set the number of membership functions per input and number of output rules per model so that in each case they equalled the number of clusters used. For each of these 30 models, a multiple linear regression (MLR) counterpart was also developed.

ANFIS is a five-layer feedforward network, applying a neural network learning algorithm and fuzzy reasoning to map input predictor variables onto an output predictand space. The basic architecture is described in detail in numerous other hydrological modelling papers and, as such, need not be repeated in our paper (e.g. Chang et al., 2005). The optimal structure for each ANFIS model used in our

experiments was determined by heuristic search; with the best-performing configuration selected according to its ability to minimise RMSE. In common with other default settings of the *genfis3* function, a Gaussian fuzzy membership input function was applied in all cases, necessitating subsequent optimisation of mean and standard deviation parameters. The output membership function was by default linear. The number of clusters (number of membership functions per input/ number of output rules per model) searched was varied from 2 (minimum) to 14 via manipulation of the *cluster_n* parameter in *genfis3*. The modelling records were divided into three sub-sets: Events 1-5 formed the training dataset; Events 6 and 7 formed the cross-validation dataset; Event 8 formed the testing dataset. The largest and smallest events were included in the training dataset, such that the need for extrapolation beyond the range of the training dataset was avoided. Early stopping (Coullibaly et al., 2000; Giustolisi and Laucelli, 2005) was applied to prevent over-fitting. The ANFIS default hybrid learning algorithm was employed to identify model parameters: delivering a powerful combination of least-squares fitting and backpropagation gradient descent methods. The optimal configuration and stopping point for each of our 30 preferred models is shown in Table 6.

4.3. Evaluation metrics

The performance of each model was evaluated and compared using five different metrics: correlation coefficient (CC: Equation 1), root mean square error (RMSE: Equation 2), mean square error (MSE: Equation 3), coefficient of efficiency (CE: Equation 4) and skill score (SS: Equation 5). Further particulars on the first four metrics can be found in Dawson et al. (2007). SS is a measure of improvement in RMSE: using a reference model as benchmark, in which a positive score indicates superior performance of a model over that benchmark model. These indices served as criteria for model selection: to identify the preferred input combination and, consequently, an optimal level of spatio-temporal aggregation for the radar rainfall modelling scenarios under test.

$$CC = \frac{\sum_{t=1}^{N}\left[Q_{pre}(t) - \overline{Q}_{pre}\right]\left[Q_{obs}(t) - \overline{Q}_{obs}\right]}{\sqrt{\sum_{t=1}^{N}\left[Q_{pre}(t) - \overline{Q}_{pre}\right]^2}\sqrt{\sum_{t=1}^{N}\left[Q_{obs}(t) - \overline{Q}_{obs}\right]^2}} \qquad \text{Equation 1}$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{t=1}^{N}\left[Q_{pre}(t) - Q_{obs}(t)\right]^2}{N}} \qquad \text{Equation 2}$$

$$\mathrm{MAE} = \frac{\sum_{i=1}^{N} | (Q_{pre}(t) - Q_{obs}(t)) |}{N} \qquad \text{Equation 3}$$

$$\mathrm{CE} = 1 - \frac{\sum_{t=1}^{N} \left[Q_{obs}(t) - \mathrm{Q}_{pre}(t) \right]^2}{\sum_{t=1}^{N} \left[Q_{obs}(t) - \overline{Q}_{obs} \right]^2} \qquad \text{Equation 4}$$

$$\mathrm{SS} = \frac{E_m - E_n}{E_m} \times 100\% \qquad \text{Equation 5}$$

where *N* is the number of observations, $Q_{pre}(t)$ is the predicted inflow at time t, $Q_{obs}(t)$ is the observed inflow at time t, $\overline{Q}_{pre}$ and $\overline{Q}_{obs}$ are the mean value of predicted and observed inflow, respectively. E_m is the RMSE of a reference model, in this study Model A. E_n is the RMSE of the model compared.

5. RESULTS

Test data set results for all 30 ANFIS models are presented in Figures 5 and 6. For the one and two-hour ahead forecasting horizons (Q_{t+1},Q_{t+2}), the metrics indicate similar, high levels of performance for all models. This clearly reflects the limited challenge involved in very short-term reservoir inflow forecasting for the Shihmen catchment. However, as the forecasting horizon is increased, clear differences in the performance of individual models become apparent.

For forecasting horizons Q_{t+3} to Q_{t+5}, Models A and B deliver similarly poor performance. This implies that the addition of spatially-distributed, point-based data provides little advantage over a simple lagged inflow model. Presumably, this is because the degree of useful spatial information that is encoded within the model inputs is highly limited. Indeed, the inclusion of continuous spatial data, even as a wholly lumped input (Model C) is shown to result in improved performance at forecasting horizons greater than Q_{t+2}. This suggests that, even without additional spatial discretisation, continuous rainfall data should be used in preference to

point-based, gauged inputs.

Further improvements in model performance are observed when the catchment is discretised into four separate sub-catchments (Model D). Moreover, the performance advantage over all other models becomes consistently greater as the forecasting horizon is increased towards Q_{t+5}. Of interest is the fact that this pattern of improvement is not carried through to models E and F. Indeed, the use of 8 and 12 sub-catchments results in models that, in most cases, perform worse than their simple lumped catchment counterpart (Model C). It demonstrates that there are clear limits to the amount of discretisation that should be applied in this case. This arises from the need to generalise highly complex, spatio-temporal patterns of typhoon events and their relationship to inflow, something which would require the use of a well-parameterised, spatially and temporally-distributed, physically-based model to properly forecast. Instead, data-driven modellers must adopt a more pragmatic position in which the relationships between data sets must be simplified through both temporal and spatial lumping: both of which are applied in this study.

Figure 7 compares observed and predicted inflow series for Model B and Model D, calculated on the testing data (typhoon Jangmi) across different lead times. The plots confirm that predicted values for Model D are much closer to the observed values. It is also worth noting that for Q_{t+5}, the substantial timing-error displacement of predicted peak inflow decreased from 6 hours for Model B to 3 hours for Model D. This indicates that the spatio-temporal optimisation of radar rainfall is not only capable of increasing the overall performance of models across the five metrics that were applied but also reduces the frequently overlooked problem of a spurious lagged time shift component appearing in data-driven model outputs (De Vos and Rientjes, 2005; Abrahart et al., 2007).

Figures 8 and 9 compare non-linear (ANFIS) and linear (MLR) rainfall-runoff reservoir inflow forecasting counterparts. RMSE (Figure 8) and CE (Figure 9) performance indices showed that ANFIS models were almost invariably superior to MLR models, and always so over longer forecast horizons. These results confirm the near-linear nature of one-step-ahead forecasting and that increased non-linearity occurs over longer forecasting horizons. It also confirms the need to develop non-linear modelling solutions for providing multi-step-ahead forecast reservoir inflow related to typhoon events in Taiwan.

6. DISCUSSION AND WIDER IMPLICATIONS

The novel contribution of this paper is its exploration of how the inclusion of spatial-distribution in a data-driven rainfall-runoff model enhances its predictive performance by better capturing the spatial and temporal heterogeneity of typhoon rainfall events. Our results reveal clear improvements associated with the adoption of a semi-distributed data-driven modelling framework; offering potential benefits for similar studies conducted under differing conditions, since catchment sub-partitioning is easily reproduced. This finding conforms to the widely-held viewpoint that errors associated with the estimation of rainfall intensity from lumped models are very likely to limit a model's ability to predict runoff accurately, and that this will be a particular problem where high-intensity, convective rainfall is known to be the key driver of runoff (Dawdy and Bergman, 1969; Wilson et al., 1979). Such results also support the notion that there may be an optimal level of discretisation beyond which anticipated performance benefit decreases due to local noise and uncertainty in the rainfall data masking the broader hydrological signal for a catchment (Lin and Chen, 2005; Dark and Bram, 2007). Thus, it is important to caution against the assumption that data-driven rainfall-runoff model accuracy will necessarily be improved by simply increasing the degree of spatial distribution employed. It is interesting to note that recommendations of the Distributed Model Intercomparison Project also caution against such an assumption (Reed et al., 2004). That project utilised gridded radar rainfall data provided as hourly accumulations, and showed that lumped models can still outperform distributed models. Further, it revealed that individual catchment characteristics are central to the performance advantages of lumped or distributed models. Indeed, in situations where improved results obtained from distributed models driven by radar rainfall input are reported, they are often associated with isolated case studies rather than routine operational predictions. Consequently, the general benefits of incorporating spatial distribution, and the importance of the different factors that affect it, remain a debated topic amongst physically-based modellers (Tetzlaff and Uhlenbrook, 2005; Berne & Krajewski, 2012).

Our paper also makes a significant contribution by providing some beneficial insights regarding the underutilised advantages of adopting radar data as an input into data-driven rainfall-runoff models, rather than the more established use of point-based rain gauge data. Whilst a large number of studies have compared the use of rain gauge and radar data (e.g. Briggs and Atkinson, 2011) and the impact of radar rainfall error and uncertainty on water resources modelling (e.g. Hossain et al., 2005; Habib et al., 2008; Gourley et al., 2011; He et al., 2011; Schrötera et al., 2011), none have considered the issues from a specifically data-driven modelling perspective. The reported method represents a significant advance over the

previous use of lumped mean areal inputs (Lorrai and Sechi, 1995), or distributed point-based rainfall samples (Campolo et al., 1999; Dawson et al., 2006), given that such past approaches possess no explicit physical or operational underpinnings. It is also a physical and structural enrichment of previous data-driven radar rainfall modelling procedures applied by Teschl and Randeu (2006), and of the spatial clustering arguments of Rajurkar et al. (2002) and Lauzon et al. (2006), which are modified by the adoption of physically-meaningful sub-catchment boundaries, developed using drainage network analysis of a digital elevation model. To this extent, a basic level of meaningful, hydrological process knowledge is incorporated into the modelling framework.

Strong parallels exist between the ideas developed in this work and those associated with hybrid modelling solutions (Perez, 2009). Hybridisation in most cases involves the development of a mixed combination of two or more different types of model in which each individual model fulfils a particular role in some larger scheme. Data-driven solutions, for example, can be externally coupled to a conventional model (standalone solution) or embedded within it (modular component). In a similar manner it is possible to exploit the predictive capabilities of data-driven models within a semi-distributed model structure, drawing upon the physical rationality of such approaches, which better reflects the nature of the hydrological phenomena that is required to be modelled. Hybrid solutions are usually characterised by relatively complex, modular solutions in which different data-driven models are developed for one or more individual components of the catchment being modelled (e.g. a different model for each sub-catchment, or for each hydrological process operating within the catchment), and subsequently combined (e.g. Corzo et al., 2009; Huo et al., 2012). By contrast, our study adopts a simpler approach involving the use of semi-distributed model *inputs* as opposed to enforcing spatial distribution in the model structure itself. In this way the predictive benefits that result from the use of different levels of semi-distribution in the model inputs are tested: an important requirement when one or more model inputs are derived from a spatially-continuous data set. The resultant model is, arguably, a simplified hybrid solution that achieves an optimal degree of generalisation of the spatio-temporal variability in typhoon rainfall-runoff processes, whilst avoiding an overly-complex, fully-distributed model structure that would require inputs and parameters that are difficult to obtain.

As with many scientific disciplines, a range of methodological approaches and associated techniques for tackling hydrological problems have emerged, that are founded on different conceptual and philosophical schools of thought about how

hydrological processes should be represented and captured within a set of relevant models (Wainwright and Mulligan, 2004). Modelling ranges from highly inductive, physically-based, to highly deductive, empirical and data-driven, and clearly different sorts of model are needed to serve a number of different purposes. Many of the approaches associated with particular schools have been developed exclusively, with relatively limited transfer and incorporation of specific techniques and ideas occurring across the wider disciplinary and scientific community. This paper exemplifies how the incorporation of the most basic physical concepts in hydrology may be integrated into a data-driven methodology so that the physical rationality of the data-driven product is a core element of the resultant model. In so doing, it supports the notion that hydrologists, irrespective of their conceptual background or philosophical stance, should where possible, seek to incorporate the ideas and knowledge that is best suited to the nature of the problem that they are trying to solve. These may derive from other parts of the discipline or beyond, and may result in models that are substantially different from those that are accepted practice within a particular school of thought. We accordingly encourage data-driven modellers to engage more fully with physical concepts in hydrology, and physical modellers to consider how data-driven techniques may be of benefit to them.

7. CONCLUSIONS

Four key points emerge from this study:

1. Continuous rainfall data appears to offer performance advantages over discrete, point-based spatial data for reservoir inflow forecasting in Taiwan.
2. Further performance advantages can be achieved by using a semi-distributed modelling framework, but there are limits to the number of catchment sub-units that should be used.
3. The spatio-temporal complexity of typhoon rainfall requires a substantial amount of spatial and temporal generalisation in order to build an effective data-driven rainfall-runoff model.
4. The operational requirement for a minimum 3-hour warning of reservoir release requires the availability of a model that performs well over lead times in excess of three hours. This study indicates that data-driven models are of use in this regard, and that their value is maximised when appropriately-distributed, continuous radar rainfall data is used.

This study has highlighted the importance of the spatial dimension in data-driven, rainfall-runoff modelling. To date, this factor has received little attention by researchers in the field. There is, therefore, a clear need for additional research into the effects of spatio-temporal generalisation on data-driven models, applied in different hydrologic and physiographic contexts. This study has demonstrated the specific capability of ANFIS, selected as a typical data-driven modelling tool. Whilst further potential improvements in performance accuracy could probably be obtained from the application of other data-driven algorithms, it is unlikely that such gains would be anything other than marginal, given that previously-published hydrological comparisons of ANFIS and other data-driven algorithms demonstrate broadly similar outcomes (e.g. Chen et al., 2006).

Acknowledgements

This project was completed in the School of Geography at the University of Nottingham. Meng-Jung Tsai was a visiting scholar from the National Taiwan University for the period 1st March - 31st August 2011. Partial funding was provided by the Sinotech Foundation for Research & Development of Engineering Sciences & Technologies, Taipei, Taiwan and by the Water Resources Agency, Taiwan.

References

Abbott MB. 1991. *Hydroinformatics: Information technology and the aquatic environment*. Avebury Technical, Aldershot, UK.

Abbott MB. 1999. Introducing Hydroinformatics. *Journal of Hydroinformatics* **1**: 2-19.

Abrahart RJ, Antcil F, Coulibaly P, Dawson CW, Elshorbagy A, Mount NJ, See LM, Shamseldin AY, Solomatine DP, Toth E, Wilby RL. 2012. Twenty years of anarchy? Emerging themes and outstanding challenges for neural network modelling of surface hydrology. *Progress in Physical Geography* **36**: 480-513.

Abrahart RJ, Dawson CW, See LM, Mount NJ, Shamseldin AY. 2010. Discussion of "Evapotranspiration modelling using support vector machines". *Hydrological Sciences Journal* **55**: 1442-1450.

Abrahart RJ, Heppenstall AJ, See, LM. 2007. Timing error correction procedure applied to neural network rainfall-runoff modelling. *Hydrological Science Journal* **52**: 414-431.

Abrahart RJ, Mount NJ, Ab Ghani N, Clifford NJ, Dawson CW. 2011. DAMP: A protocol for contextualising goodness-of-fit statistics in sediment-discharge data-driven modelling. *Journal of Hydrology* **409**: 596-611.

Abrahart RJ, See LM. 2000. Comparing neural network and autoregressive moving average techniques for the provision of continuous river flow forecasts in two contrasting catchments. *Hydrological Processes* **14**: 2157-2172.

Abrahart RJ, See LM. 2007. Neural network modelling of non-linear hydrological relationships. *Hydrology and Earth System Sciences* **11**: 1563-1579.

Abrahart RJ, See LM, Solomatine DP. (Eds.) 2008. *Hydroinformatics: Computational Intelligence and Technological Developments in Water Applications*. Springer-Verlag: Berlin and Heidelberg.

Akhtar MK, Corzo GA, Andel SJV, Jonoski A. 2009. River flow forecasting with artificial neural networks using satellite observed precipitation pre-processed with flow length and travel time information: case study of the Ganges river basin. *Hydrology and Earth System Sciences* **13**: 1607-1618.

American Society of Civil Engineers. 1996. *Hydrology handbook* 2nd Ed. American Society of Civil Engineers: New York.

Andrieu H, Creutin JD, Delrieu G, Faure D. 1997. Use of weather radar for the hydrology of a mountainous area. Part I: Radar measurement interpretation. *Journal of Hydrology* **34**: 225–239.

Austin PM. 1987. Relation between measured radar reflectivity and surface rainfall. *Monthly Weather Review* **115**: 1053–1070.

Babovic V, Keijzer M. 2002. Rainfall Runoff Modelling Based on Genetic Programming. *Nordic Hydrology* **33**: 331-346.

Bae D-H, Jeong DM, Kim G. 2007. Monthly dam inflow forecasts using weather forecasting information and neuro–fuzzy technique. *Hydrological Sciences Journal* **52**: 99–113.

Bell VA, Moore RJ. 1998. A grid-based distributed flood forecasting model for use with weather radar data. 2. Case studies. *Hydrology and Earth System Sciences* **2**: 278–283.

Berne A, Krajewski WF. 2012. Radar for hydrology: Unfulfilled promise or unrecognized potential? *Advances in Water Resources* [published online 1 Jun 2012]. doi:10.1016/j.advwatres.2012.05.005

Beven K J, O'Connell P E. 1982. *On the role of physically-based distributed modelling in hydrology*. Wallingford, Institute of Hydrology. Institute of Hydrology Report No.81. 36pp.

Biggs EM, Atkinson PM. 2011. A comparison of gauge and radar precipitation data for simulating an extreme hydrological event in the Severn Uplands, UK.

Hydrological Processes **25**: 795–810.

Borga M. 2002. Accuracy of radar rainfall estimates for streamflow simulation. *Journal of Hydrology* **267**: 26-39.

Borga M, Anagnostou EN, Frank E. 2000. On the use of real-time radar rainfall estimates for flood prediction in mountainous basins. *Journal of Geophysical Research* **105**: 2269–2280.

Borga M, Tonelli F, Moore RJ, Andrieu H. 2002. Long term assessment of bias adjustment in radar rainfall estimation. *Water Resources Research* **38**: 1226.

Campolo M, Andreussi P, Soldati, A. 1999. River flood forecasting with a neural network model. *Water Resources Research* **35**: 1191-1197.

Camargo S J, Robertson AW, Gaffney SJ, Smyth P, Ghil M. 2007. Cluster analysis of typhoon tracks. Part I: General properties. *Journal of Climate* **20**: 3635–3653.

Chaipimonplin T, See LM, Kneale PE. 2010. Using radar data to extend the lead time of neural network forecasting on the River Ping. *Disaster Advances* **3**: 35-43.

Chang F-J, Chiang Y-M, Chang L-C. 2007. Multi-step-ahead neural networks for flood forecasting. *Hydrological Sciences Journal* **52**: 114-130.

Chang F-J, Chang L-C, Huang, H-L. 2002. Real-time recurrent learning neural network for stream-flow forecasting. *Hydrological Processes* **16**: 2577-2588.

Chang F-J, Chang L-C, Wang Y-S. 2007. Enforced self-organizing map neural networks for river flood forecasting. *Hydrological Processes* **21**: 741-749.

Chang F-J, Chang Y-T. 2006. Adaptive neuro-fuzzy inference system for prediction of water level in reservoir. *Advances in Water Resources* **29**: 1-10.

Chang YT, Chang LC, Chang FJ. 2005. Intelligent control for modeling of real-time reservoir operation, part II: artificial neural network with operating rule curves. *Hydrological Processes* **19**: 1431-1444.

Chau KW, Wu CL, Li YS. 2005. Comparison of several flood forecasting models in Yangtze River. *Journal of Hydrologic Engineering* **10**: 485-491.

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

Chen SH, Lin YH, Chang LC, Chang FJ. 2006. The strategy of building a flood forecast model by neuro-fuzzy network. *Hydrological Processes* **20**: 1525-1540.

Chen Y-H, Chang F-J. 2009. Evolutionary artificial neural networks for hydrological systems forecasting. *Journal of Hydrology* **367**:125–137.

Cheng K-S, Lin Y-C, Liou J-J. 2008. Rain-gauge network evaluation and augmentation using geostatistics. *Hydrological Processes* **22**: 2554–2564.

Chiang Y-M, Chang F-J. 2009. Integrating hydrometeorological information for rainfall-runoff modelling by artificial neural networks. *Hydrological Processes* **23**: 1650–1659.

Chu H-J, Liau C-J, Lin C-H, Su B-S. 2012. Integration of fuzzy cluster analysis and kernel density estimation for tracking typhoon trajectories in the Taiwan region. *Expert Systems with Applications* **39**: 9451–9457.

Corzo G, Solomatine, D. 2007. Baseflow separation techniques for modular artificial neural network modelling in flow forecasting. *Hydrological Sciences Journal* **52**: 491-507.

Corzo GA, Solomatine DP, Hidayat, De Wit M, Werner M, Uhlenbrook S, Price RK. 2009. Combining semi-distributed process-based and data-driven models in flow simulation: a case study of the Meuse river basin. *Hydrology and Earth System Sciences* **13**: 1619–1634.

Coulibaly P, Anctil F, Bobee B. 2000. Daily reservoir inflow forecasting using artificial neural networks with stopped training approach. *Journal of Hydrology* **230**: 244-257.

Dark SJ, Bram D. 2007. The modifiable areal unit problem (MAUP) in physical geography. *Progress in Physical Geography* **31**: 471-479.

Dawdy DR, Bergman J M. 1969. Effect of rainfall variability on streamflow simulation. *Water Resources Research* **5**: 958–966.

Dawson CW, See LM, Abrahart RJ, Heppenstall AJ. 2006. Symbiotic adaptive neuro-evolution applied to rainfall-runoff modelling in northern England. *Neural Networks* **19**: 236-247.

Dawson CW, Abrahart RJ, See LM. 2007. HydroTest: a web-based toolbox of evaluation metrics for the standardised assessment of hydrological forecasts. *Environmental Modelling & Software* **22**: 1034-1052.

Deo MC, Thirumalaiah, K. 2000. Real time forecasting using neural networks. In *Artificial Neural Networks in Hydrology*, Govindaraju, Ramachandra Rao A. (eds). Kluwer Academic Publishers: Dordrecht; 53-71.

De Vos NJ. 2012. Reservoir computing as an alternative to traditional artificial neural networks in rainfall-runoff modelling. *Hydrology and Earth System Sciences Discussions* **9:** 6101-6134.

De Vos NJ, Rientjes THM. 2005. Constraints of artificial neural networks for rainfall–runoff modelling: trade-offs in hydrological state representation and model evaluation. *Hydrology and Earth System Sciences* **9**: 111–126.

El-Shafie A, Taha MR, Noureldin A. 2007. A neuro-fuzzy model for inflow forecasting of the Nile river at Aswan high dam. *Water Resources Management* **21**: 533-556.

Elshorbagy A, Corzo G, Srinivasulu S, Solomatine D P. 2010a. Experimental investigation of the predictive capabilities of data driven modeling techniques in hydrology - Part 1: Concepts and methodology. *Hydrology and Earth System Sciences* **14**: 1931-1941.

Elshorbagy A, Corzo G, Srinivasulu S, Solomatine DP. 2010b. Experimental investigation of the predictive capabilities of data driven modeling techniques in hydrology - Part 2: Application. *Hydrology and Earth System Sciences* **14**: 1943-1961.

Firat M, Güngor M. 2008. Hydrological time-series modelling using an adaptive neuro-fuzzy inference system. *Hydrological Processes* **22**: 2122–2132.

Gautam DK, Holz KP. 2001. Rainfall-runoff modelling using adaptive neuro-fuzzy systems. *Journal of Hydroinformatics* **3**: 3-10.

Georgakakos KP, Sperfslage JA, Guetter AK. 1996. Operational GIS based models for NEXRAD radar data in the U.S. In *Proceedings of the International Conference on Water Resources and Environmental Research, 29–31 October, 1996, Water Resources and Environmental Research Center, Kyoto University, Kyoto, Japan*; 603–609.

Ghalkhani H, Golian S, Saghafian B, Farokhnia A, Shamseldin A. (in press). Application of surrogate artificial intelligent models for real-time flood routing. *Water and Environment Journal* [published online: 16 August 2012] doi: 10.1111/j.1747-6593.2012.00344.x

Giustolisi O, Laucelli D. 2005. Improving generalization of artificial neural networks in rainfall–runoff modelling. *Hydrological Sciences Journal* **50**: 439–457.

Gourley, JJ., Hong Y, Flamig ZL, Wang J, Vergara H, Anagnostou EN. 2011. Hydrologic Evaluation of Rainfall Estimates from Radar, Satellite, Gauge, and Combinations on Ft. Cobb Basin, Oklahoma. *Journal of Hydrometeorology* **12**: 973–988.

Habib E, Aduvala AV, Meselhe EA. 2008. Analysis of radar-rainfall error characteristics and implications for streamflow simulation uncertainty. *Hydrological Sciences Journal* **53**: 568-587.

He X, Refsgaard J C, Sonnenborg TO, Vejen F, Jensen KH. 2011. Statistical analysis of the impact of radar rainfall uncertainties on water resources modelling. *Water Resources Research* **47**: W09526. doi:10.1029/2011WR010670.

Holz KP, Cunge JA, Lehfeldt, R, Savic D. 2011. *Hydroinformatics Vision 2011.* Synoptic Report of the IAHR/IWA/IAHS Hydroinformatics Joint Committee Working Group: http://www.iahr.net/site/about/organisation/technical/div_2/Committee_on_Hydroinformatics/Final_Synoptic_Report.pdf

Hossain F, Anagnostou EN, Dinku T, Borga M. 2005. Hydrological model sensitivity to parameter and radar rainfall estimation uncertainty. *Hydrological Processes* **18**: 3277-3291.

Hsu N-S, Wei C-C. 2007. A multipurpose reservoir real-time operation model for flood control during typhoon invasion. *Journal of Hydrology* **336**: 282-293.

Huo Z, Feng S, Kang S, Huang G, Wang F, Guo P. 2012. Integrated neural networks for monthly river flow estimation in arid inland basin of Northwest China. Journal of Hydrology **420–421**: 159–170

Jang J-SR. 1993 ANFIS: Adaptive-Network-Based Fuzzy Inference System. *IEEE Transactions on Systems, Man, and Cybernetics* **23**: 665-685.

Jang J-SR, Sun C-T, Mizutani E. 1997. *Neuro-Fuzzy and Soft Computing: a computational approach to learning and machine intelligence*, Prentice Hall, Upper Saddle River, NJ, USA.

James WP, Robinson CG, Bell JF. 1993. Radar-assisted real-time flood forecasting. *Journal of Water Resources Planning and Management* **119**: 32–44.

Jensen SK, Domingue JO. 1988. Extracting topographic structure from digital elevation data for geographic information system analysis. *Photogrammetric Engineering and Remote Sensing* **54**: 1593-1600.

Joss J, Lee R. 1995. The application of radar–gage comparisons to operational precipitation profile corrections. *Journal of Applied Meteorology* **34**: 2612–2630.

Jothiprakash V, Magar, RB. 2012. Multi-time-step ahead daily and hourly intermittent reservoir inflow prediction by artificial intelligent techniques using lumped and distributed data. *Journal of Hydrology* **450-451**: 293-307.

Kao S-C, Kume T, Komatsu H, Liang W-L. 2012. Spatial and temporal variations in rainfall characteristics in mountainous and lowland areas in Taiwan. *Hydrological Processes* [published online 8 Jun 2012] doi: 10.1002/hyp.9416

Karimi-Googhari SH, Lee TS. 2011. Applicability of Adaptive Neuro-Fuzzy Inference Systems in Daily Reservoir Inflow Forecasting. *International Journal of Soft Computing* **6**: 75-84.

Keskin ME, Taylan D, Terzi O. 2006. Adaptive neural-based fuzzy inference system (ANFIS) approach for modelling hydrological time series. *Hydrological Sciences Journal* **51**: 588–598.

Lange J, Leibundgut C, Greenbaum N, Schick A P. 1999. A non-calibrated rainfall-runoff model for large, arid catchments. *Water Resources Research* **35**: 2161–2172.

Lauzon N, Anctil F, Baxter CW. 2006. Clustering of heterogeneous precipitation fields for the assessment and possible improvement of lumped neural network models for streamflow forecasts. *Hydrology and Earth System Sciences* **10**: 485-494.

Lee C-S, Huang L-R, Shen H-S, Wang S-T. 2006. A Climatology Model for Forecasting Typhoon Rainfall in Taiwan. *Natural Hazards* **37**: 87-105.

Lee M-H, Ho C-H, Kim J-H, Song H-J. (in press). Low-frequency variability of tropical cyclone-induced heavy rainfall over East Asia associated with tropical and North Pacific sea surface temperatures. *Journal of Geophysical Research.* doi:10.1029/2012JD017565.

Lee PS, You GJ-Y. 2011. *The risk analysis of long term impact to reservoir under extreme hydrologic events –Shihmen Reservoir a Case Study*. Paper presented at the World Environmental and Water Resources Congress 2011.

Lin G-F, Chen L-H. 2005. Application of an artificial neural network to typhoon rainfall forecasting. *Hydrological Processes* **19**: 1825-1837.

Lin G-F, Chen G-R, Huang P-Y, Chou Y-C. 2009b. Support vector machine-based models for hourly reservoir inflow forecasting during typhoon-warning periods, *Journal of Hydrology* **372**: 17–29.

Lin G-F, Wu M-C, Chen G-R, Tsai F-Y. 2009a. An RBF-based model with an information processor for forecasting hourly reservoir inflow during typhoons. *Hydrological Processes* **23**: 3598-3609.

Lohani AK, Kumar R, Singh RD. 2012. Hydrological time series modelling: a comparison between adaptive neuro-fuzzy, neural network and autoregressive techniques. *Journal of Hydrology* **442-443**: 23-35.

Lorrai M, Sechi GM. 1995. Neural nets for modelling rainfall-runoff transformations. *Water Resources Management* **9**: 299-313.

Ly S, Charles C, Degré A. 2011. Geostatistical interpolation of daily rainfall at catchment scale: the use of several variogram models in the Ourthe and Ambleve catchments, Belgium. *Hydrology and Earth System Sciences* **15**: 2259–2274.

Maidment DR. 2002. *Arc Hydro: GIS for Water Resources*. Redlands, CA: ESRI Press. 203 pp.

Maier HR, Dandy GC. 2000. Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. *Environmental Modelling & Software* **15**: 101-124.

Maier HR, Jain A, Dandy GC, Sudheer KP. 2010. Methods used for the development of neural networks for the prediction of water resource variables in river systems: Current status and future directions. *Environmental Modelling &*

Software **25**: 891-909.

Michaud JD, Sorooshian S. 1994. Effects of rainfall-sampling errors on simulations of desert flash floods. *Water Resources Research* **30**: 2765–2775.

Molini A, Lanza LG, La Barbera P. 2005. The impact of tipping-bucket raingauge measurement errors on design rainfall for urban-scale applications. *Hydrological Processes* **19**: 1073–1088.

Mount NJ, Abrahart RJ. 2011a. Load or concentration, logged or unlogged? Addressing ten years of uncertainty in neural network suspended sediment prediction. *Hydrological Processes* **25**: 3144-3157.

Mount NJ, Abrahart RJ. 2011b. Discussion of "River flow estimation from upstream flow records by artificial intelligence methods", by M.E. Turan and M.A. Yurdusev [J. Hydrol. 369 (2009) 71-77]. *Journal of Hydrology* **396**: 193-196.

Mount NJ, Abrahart RJ. 2012. The need for operational reasoning in data-driven rating curve prediction of suspended sediment. *Hydrological Processes*: doi: 10.1002/hyp.8439.

Mukerji A, Chatterjee C, Raghuwanshi NS. 2009. Flood forecasting using ANN, neuro-fuzzy and neuro-GA models. *ASCE Journal of Hydrologic Engineering* **14**: 647-653.

Nayak PC, Sudheer KP, Rangan DM, Ramasastri KS. 2005. Short-term flood forecasting with a neurofuzzy model. *Water Resources Research* **41**: W04004.

Nguyen PK-T, Chau LH-C. 2012. The data-driven approach as an operational real-time flood forecasting model. *Hydrological Processes* **26**: 2878-2893.

Niedzielski T. 2007. A data-based regional scale autoregressive rainfall-runoff model: a study from the Odra River. *Stochastic Environmental Research and Risk Assessment* **21**: 649-664.

Ogden FL, Sharif HO, Senarath SUS, Smith JA, Beck ML, Richardson JR. 2000. Hydrologic analysis of the Fort Collins, Colorado, flash flood of 1997. *Journal of Hydrology* **228**: 82–100.

Pan T-Y, Chang L-Y, Lai J-S, Chang H-K, Lee C-S, Tan Y-C. 2012. Coupling typhoon rainfall forecasting with overland-flow modeling for early warning of inundation. *Natural Hazards* [published online 5 January 2012]. doi: 10.1007/s11069-011-0061-9

Perez, GAC. 2009. Hybrid models for hydrological forecasting: integration of data-driven and conceptual modelling techniques. *PhD Thesis.* UNESCO-IHE Institute for Water Education: Delft, The Netherlands.

Pramanik N, Panda RK. 2009. Application of neural network and adaptive

neuro-fuzzy inference systems for river flow prediction. *Hydrological Sciences Journal* **54**: 247–260.

Rajurkar MP, Kothyari UC, Chaube UC. 2002. Artificial neural networks for daily rainfall-runoff modelling. *Hydrological Sciences Journal* **47**: 865-877.

Reed S, Koren V, Smith M, Zhang Z, Moreda F, Seo D-J, et al. 2004. Overall distributed model intercomparison project results. *Journal of Hydrology* **298**: 27–60.

Rodríguez-Vázquez K, Arganis-Juárez ML, Cruickshank-Villanueva C, Domínguez-Mora R. 2012. Rainfall–runoff modelling using genetic programming. *Journal of Hydroinformatics* **14**: 108–121.

Schell GS, Madramootoo CA, Austin GL, Broughton RS. 1992. Use of radar measured rainfall for hydrologic modelling. *Canadian Agricultural Engineering* **34**: 41–48.

Schiemann R, Liniger M A, Frei C. 2010. Reduced space optimal interpolation of daily rain gauge precipitation in Switzerland. *Journal of Geophysical Research* **115**: D14109.

Schröter K, Llort X, Velasco-Forero C, Ostrowski M, Sempere-Torres D. 2011. Implications of radar rainfall estimates uncertainty on distributed hydrological model predictions. *Atmospheric Research* **100**: 237–245.

See LM, Solomatine DP, Abrahart RJ, Toth E. 2007. Hydroinformatics: computational intelligence and technological developments in water science applications—Editorial. *Hydrological Sciences Journal* **52**: 391-396.

Segond, M-L, Wheater HS, Onof C. 2007. The significance of spatial rainfall representation for flood runoff estimation: A numerical evaluation based on the Lee catchment, UK. *Journal of Hydrology* **347**: 116-131.

Sempere-Torres D, Corral C, Raso J, Malgrat P. 1999. Use of weather radar for combined sewer overflows monitoring and control. *ASCE Journal of Environmental Engineering* **125**: 372–380.

Solomatine, DP. 2005. Data-driven modelling and computational intelligence methods in hydrology. In: *Encyclopedia of Hydrological Sciences* (M.G. Andersen, ed.), vol. 1, part 2, chapter 19. John Wiley & Sons. pp. 293-306.

Solomatine, DP, Ostfeld, A. 2008. Data-driven modelling: some past experiences and new approaches. *Journal of Hydroinformatics* **10**: 3-22.

Song X, Kong F, Zhan C, Han J. (in press). A Hybrid Optimization Rainfall-Runoff Simulation Based on Xinanjiang Model and Artificial Neural Network. *ASCE Journal of Hydrologic Engineering*. doi:10.1061/(ASCE)HE.1943-5584.0000548

Talei A, Chua LHC, Wong T. 2012. Evaluation of rainfall and discharge inputs used by

Adaptive Network-based Fuzzy Inference Systems (ANFIS) in rainfall–runoff modeling. *Journal of Hydrology* **391**: 248–262.

Teschl R, Randeu WL. 2006. A neural network model for short term river flow prediction. *Natural Hazards and Earth System Sciences* **6**: 629-635.

Teschl R, Randeu WL, Teschl F. 2009. Weather radar measurements in data-driven rainfall-runoff models. *Geophysical Research Abstracts* **11**: 12712.

Tetzlaff D, Uhlenbrook S. 2005. Significance of spatial variability in precipitation for process-oriented modelling: results from two nested catchments using radar and ground station data. *Hydrology and Earth System Sciences* **9**: 29–41.

Thiessen AH. 1911. Precipitation averages for large areas. *Monthly Weather Review* **39**: 1082-1084.

Toth E, Brath A. 2007. Multistep ahead streamflow forecasting: Role of calibration data in conceptual and neural network modelling. *Water Resources Research* **43**: W11405, doi:10.1029/2006WR005383.

Verworn A, Haberlandt U. 2011. Spatial interpolation of hourly rainfall – effect of additional information, variogram inference and storm properties. *Hydrology and Earth System Sciences* **15**: 569-584.

Vieux BE, Bedient PB. 1998. Estimation of rainfall for flood prediction from WSR-88D reflectivity: a case study, 17–18 October 1994. *Weather and Forecasting* **13**: 407–415.

Vieux BE, Vieux JE, Chen C, Howard KW. 2003. Operational deployment of a physics-based rainfall-runoff model for flood forecasting in Taiwan. *HS03: International Symposium on Information from Weather Radar and Distributed Hydrological Modeling, Proc. IAHS General Assembly, Sapporo, Japan, July 7-8, 2003.*

Wainwright J, Mulligan M. 2004. *Environmental Modelling: Finding Simplicity in Complexity*. John Wiley & Sons: Chichester

Wagner PD, Fiener P, Wilken F, Kumar S, Schneider K. (in press). Comparison and evaluation of spatial interpolation schemes for daily rainfall in data scarce regions. *Journal of Hydrology*. [published online 24 July 2012]. http://dx.doi.org/10.1016/j.jhydrol.2012.07.026

Wang K, Altunkaynak A. 2012. Comparative Case Study of Rainfall-Runoff Modeling between SWMM and Fuzzy Logic Approach. *ASCE Journal of Hydrologic Engineering* **17**: 283–291.

Water Resources Agency. 1984. *Shihmen reservoir operating rules and regulations*. Taoyuan: Taiwan. (in Chinese)

Wilson CB, Valdes JB, Rodriguez-Iturbe I. 1979. On the influence of spatial distribution of rainfall on storm runoff. *Water Resources Research* **15**:

321–328.

Winchell M, Gupta HV, Sorooshian S. 1998. On the simulation of infiltration- and saturation-excess runoff using radar-based rainfall estimates: effects of algorithm uncertainty and pixel aggregation. *Water Resources Research* **34**: 2655–2670.

Woods R, Grayson R, Western A, Duncan M, Wilson D, Young R, Ibbitt R, Henderson R, McMahon T. 2000. Experimental design and initial results from the Mahurangi River Variability Experiment: MARVEX. In: *Land Surface Hydrology, Meteorology and Climate: Observations and Modeling*, Lakshmi V, Albertson JD, Schaake J, Eds., Water Sciences and Application, vol. 3. American Geophysical Union, pp. 201–213.

Wu CL, Chau, KW. 2011. Rainfall–runoff modeling using artificial neural network coupled with singular spectrum analysis. *Journal of Hydrology* **399**: 394-409.

Wu R-S, Shih D-S, Chen S-W. 2007. Rainfall-runoff model for typhoons making landfall in Taiwan. *Journal of the American Water Resources Association*, **43**: 969-980.

Wu R-S, Shih D-S, Li M-H, Niu C-C. 2008. Coupled surface and groundwater models for investigating hydrological processes. *Hydrological Processes* **22**: 1216-1229.

Xu ZX, Li JY. 2002. Short-term inflow forecasting using an artificial neural network model. *Hydrological Processes* **16**: 2423-2439.

Yonaba H, Anctil F, Fortin V. 2010. Comparing Sigmoid Transfer Functions for Neural Network Multistep Ahead Streamflow Forecasting. *ASCE Journal of Hydrologic Engineering* **15**: 275–283.

Younger PM, Freer JE, Beven KJ. 2009. Detecting the effects of spatial variability of rainfall on hydrological modelling within an uncertainty analysis framework. *Hydrological Processes* **23**: 1988-2003.

Yu P-S, Chen ST, Chang IF. 2006. Support vector regression for real-time flood stage forecasting. *Journal of Hydrology:* **328**: 704-716.

Yu P-S, Wang, Y-C, Kuo, C-C. 2004. Simulation of flow hydrographs at an ungauged site in Taiwan using a distributed rainfall-runoff model. *Proc. International Environmental Modelling and Software Society International Conference, Osnabruck, Germany, 14-17th June 2004*. pp. 6.

Table 1 Typhoon dataset

Event	Name	SSHS category[#]	Period	Peak inflow (m^3 s^{-1})
1	SEPAT	5	2007/08/16～08/19	1844.40
2	KROSA	4	2007/10/04～10/07	5300.39
3	KALMAEGI	2	2008/07/16～07/18	203.13
4	SINLAKU	4	2008/09/11～09/16	3351.24
5	MORAKOT	1	2009/08/05～08/10	1837.54
6	WIPHA	4	2007/09/17～09/19	2788.15
7	FUNG-WONG	2	2008/07/26～07/29	2039.78
8	JANGMI	5	2008/09/26～09/29	3291.99

[#]Hurricanes are separated into five categories based on wind strength. The scale is roughly logarithmic: 1) Very dangerous winds will cause some damage; 2) Extremely dangerous winds will cause extensive damage; 3) Devastating damage will occur; 4) Catastrophic damage will occur; 5) Catastrophic damage will occur.

Table 2 Summary statistics for reservoir inflow (m^3 s^{-1}) and gauged rainfall (mm hr^{-1}) datasets

	Mean	**St.dev**	**Max**	**Min**
Inflow	1108.90	1030.90	5300.40	10.84

Rain Gauge	**Mean**	**St.dev**	**Max**	**Min**
G1	6.4	8.3	53.0	0.0
G2	7.3	9.4	50.0	0.0
G3	7.4	9.6	56.0	0.0
G4	8.7	10.5	56.0	0.0
G5	7.5	8.9	52.5	0.0
G6	7.1	7.8	44.0	0.0
G7	7.8	10.0	55.0	0.0
G8	7.3	8.8	49.5	0.0
G9	7.3	8.1	62.0	0.0
G10	9.1	10.1	56.0	0.0
G11	7.6	9.3	54.0	0.0
G12	9.2	9.8	54.0	0.0

Table 3 Modelling configurations

Model	**Inflow Output**	**Inflow Inputs**		**Rain Gauge Inputs**		**Radar Inputs**	
	$Q_{t+1}...Q_{t+5}$	Q_t	$Q_t - Q_{t-1}$	$G_1, G_2...G_{12}$	Lags (Hours)	(Sub)Catchment Units	Lags (Hours)
A	✓	✓	✓	×	×	×	×
B	✓	✓	✓	✓	*Either* 6 *or* 7	×	×
C	✓	✓	✓	×	×	1	6
D	✓	✓	✓	×	×	4	*Either* 5,6 or 7
E	✓	✓	✓	×	×	8	*Either* 5,6,7 *or* 8
F	✓	✓	✓	×	×	12	*Either* 5,6,7 *or* 8

Table 4 Mean correlation coefficient between gauged rainfall and reservoir inflow at various time lags. Highest coefficient for each gauge in bold.

ID	Name	t	t-1	t-2	t-3	t-4	t-5	t-6	t-7	t-8	t-9	t-10
G1	Jhangsing	0.219	0.333	0.420	0.463	0.514	0.528	**0.534**	0.525	0.524	0.506	0.461
G2	Fousing	0.226	0.340	0.419	0.459	0.509	0.516	**0.518**	0.512	0.506	0.486	0.431
G3	Siayun	0.231	0.346	0.439	0.480	0.532	0.542	**0.550**	0.536	0.515	0.496	0.440
G4	Gaoyi	0.203	0.304	0.396	0.461	0.510	0.538	**0.556**	0.553	0.541	0.528	0.480
G5	Baling	0.186	0.285	0.381	0.458	0.520	0.564	0.588	**0.599**	0.595	0.592	0.558
G6	Saguang	0.227	0.328	0.420	0.497	0.553	0.591	**0.599**	0.597	0.578	0.567	0.524
G7	Galahe	0.211	0.292	0.375	0.449	0.500	0.540	**0.553**	0.550	0.543	0.528	0.485
G8	Yufong	0.312	0.409	0.494	0.559	0.595	0.610	**0.619**	0.585	0.560	0.527	0.475
G9	Siouluan	0.303	0.405	0.489	0.559	0.617	0.649	0.655	**0.659**	0.632	0.611	0.560
G10	Baishin	0.315	0.407	0.491	0.560	0.618	0.669	0.673	**0.685**	0.662	0.622	0.579
G11	Jhensibao	0.247	0.340	0.425	0.493	0.560	0.613	0.639	**0.647**	0.639	0.617	0.580
G12	Siciouaihshan	0.223	0.303	0.386	0.456	0.514	0.564	0.565	**0.574**	0.569	0.547	0.523

Table 5 Summary statistics for total radar rainfall (mm hr^{-1}) per sub-catchment. Note minimum value of zero for all sub-catchments.

(Sub)Catchment Units	1			4			8			12		
	Mean	**St.dev**	**Max**	**Mean**	**St.dev**	**Max**	**Mean**	**St.dev**	**Max**	**Mean**	**St.dev**	**Max**
I	7920	7533	40664	3132	3343	20571	716	901	7357	716	901	7357
II				1318	1432	7232	2467	2651	13483	1269	1382	7151
III				1191	1238	7909	469	493	2780	246	307	1932
IV				2156	1939	10790	665	707	4829	137	143	820
V							700	683	3737	330	356	1960
VI							1399	1515	7542	951	1045	5162
VII							1127	958	5126	316	328	2139
VIII							386	381	2100	349	394	2690
IX										700	683	3737
X										1399	1515	7542
XI										1127	958	5126
XII										386	381	2100

* Note total radar rainfall comprises radar grid values aggregated by sub-catchment.

Table 6 Optimal configuration and stopping point for each preferred model.

	T+1		t+2		t+3		T+4		t+5	
Model	**MF**	**TI**	**MF**	**TI**	**MF**	**TI**	**MF**	**TI**	**MF**	**TI**
A	6	500	6	900	5	500	3	1000	5	1400
B	4	1200	2	600	3	900	3	1500	5	600
C	3	800	3	600	4	700	3	600	3	1000
D	2	1000	3	1200	2	1000	2	1300	2	900
E	2	800	2	1500	3	1500	3	600	3	900
F	2	1400	2	1100	2	700	2	1500	2	900

MF = number of membership functions. TI = number of training iterations.

FIGURE CAPTIONS

Fig. 1 Typology of paths for typhoons crossing or proximal to Taiwan, 1911-2010.

Fig. 2 Shihmen Reservoir catchment and its 12 rain gauges (G1 - G12).

Fig. 3 Reservoir inflow series for eight typhoon events.

Fig. 4 Mean time lag maps (a) radar cell map; (b) 4 sub-catchments; (c) 8 sub-catchments; (d) 12 sub-catchments.

Fig. 5 Test data set statistics for 30 ANFIS models.

Fig. 6 Model skill score on test data set using Model A as the benchmark model.

Fig. 7 Comparison of observed and predicted values for Model B and Model D at lead times (a) t+1; (b) t+2; (c) t+3; (d) t+4; (e) t+5.

Fig. 8 MLR and ANFIS, statistical comparison on RMSE.

Fig. 9 MLR and ANFIS, statistical comparison on CE.

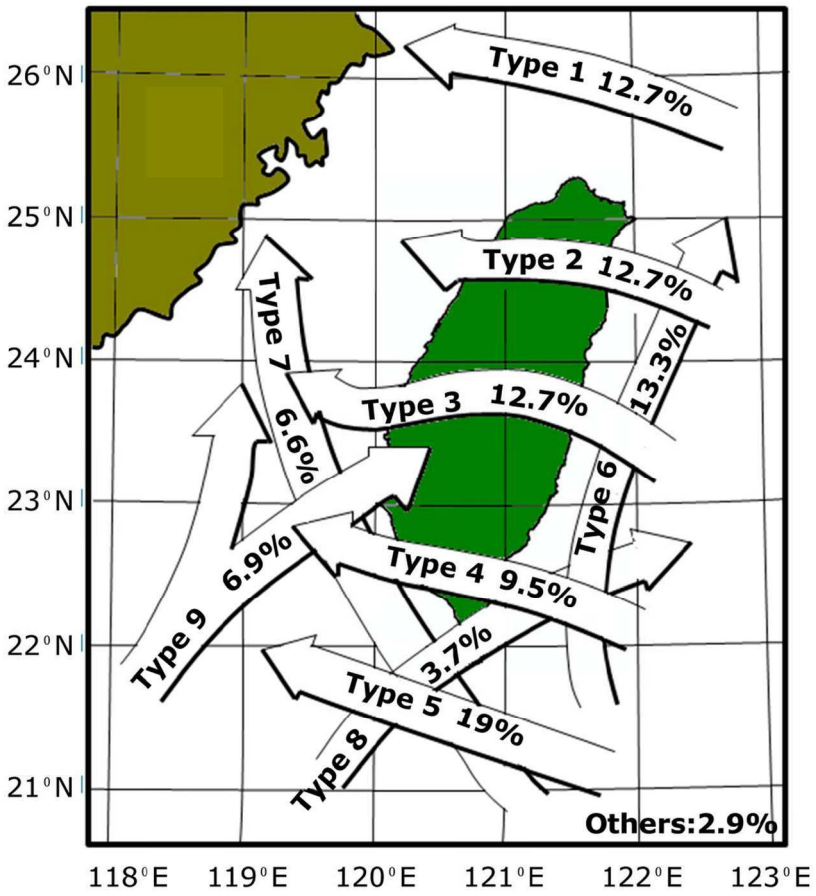

Fig. 1: Typology of paths for typhoons crossing or proximal to Taiwan, 1911-2010. (Taiwan Central Weather Bureau, pers. comm.)
118x168mm (300 x 300 DPI)

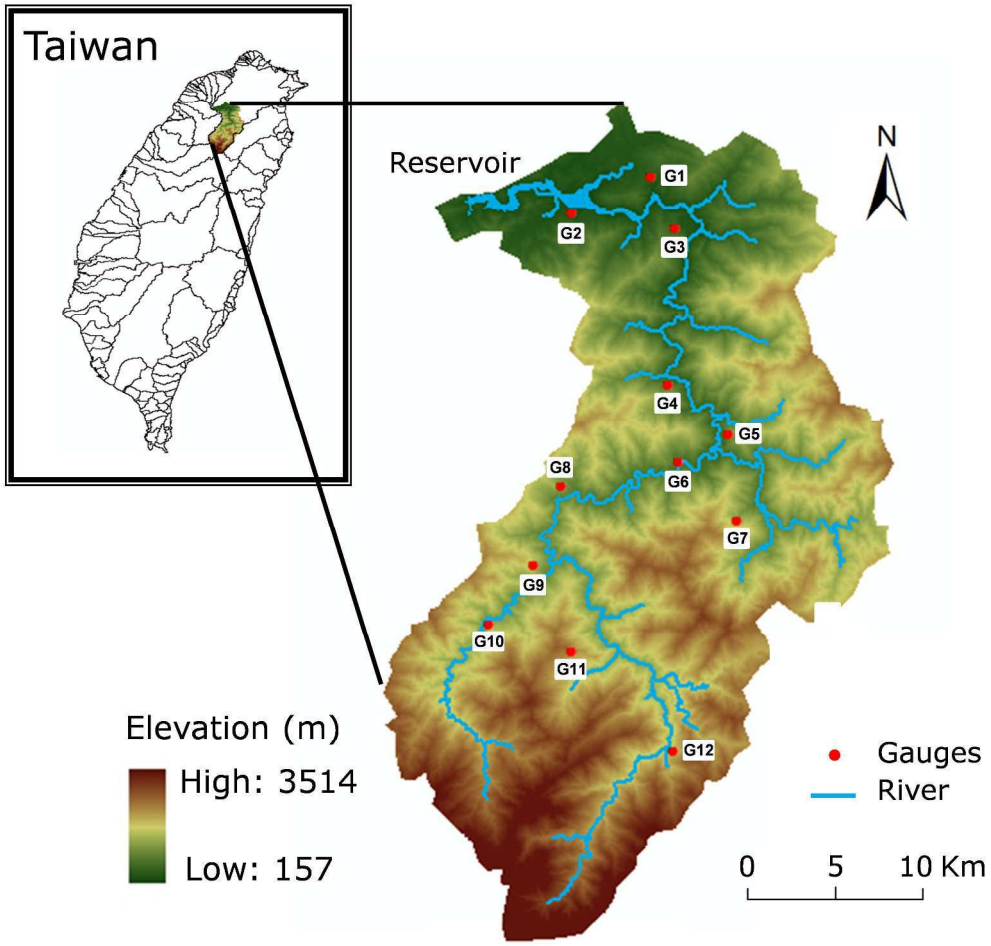

Fig. 2: Shihmen Reservoir catchment and its 12 rain gauges (G1 - G12).
1111x1060mm (96 x 96 DPI)

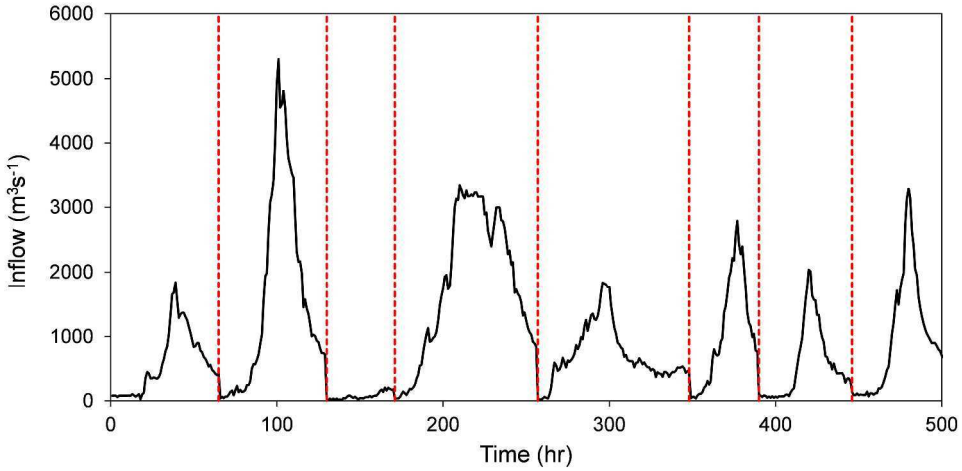

Fig. 3: Reservoir inflow series for eight typhoon events.
1340x654mm (96 x 96 DPI)

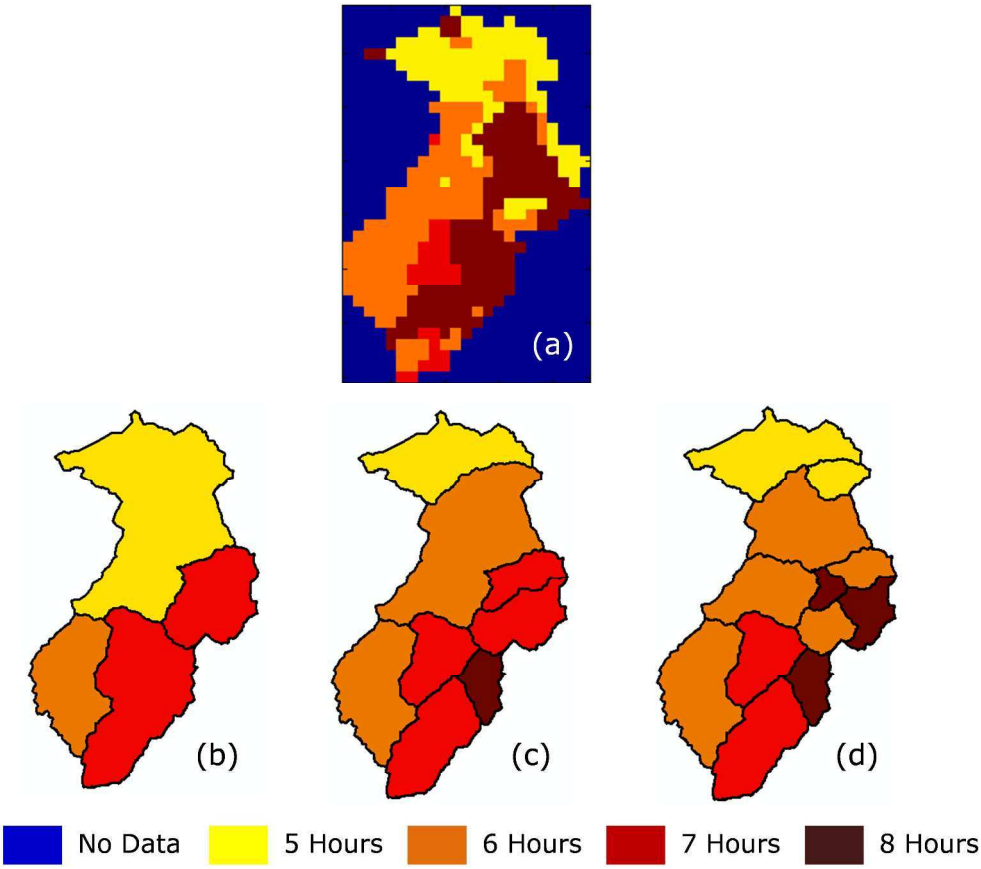

Fig. 4: Mean time lag maps (a) radar cell map; (b) 4 sub-catchments; (c) 8 sub-catchments; (d) 12 sub-catchments.
1217x1069mm (96 x 96 DPI)

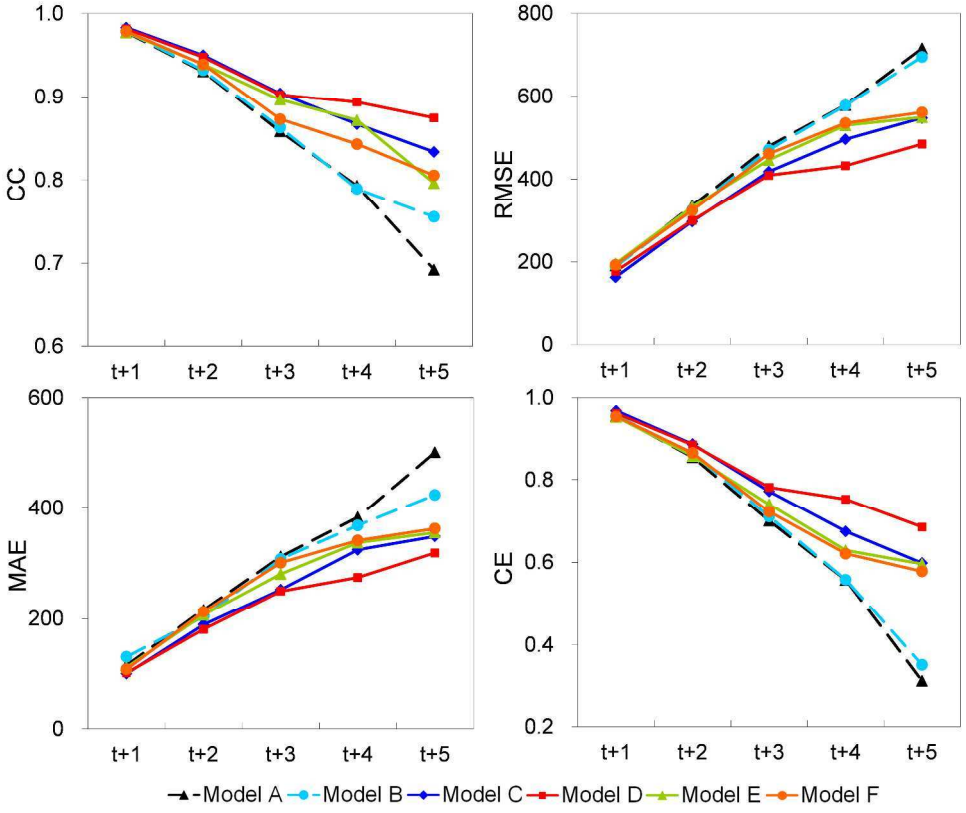

Fig. 5: Test data set statistics for 30 ANFIS models.
1456x1187mm (96 x 96 DPI)

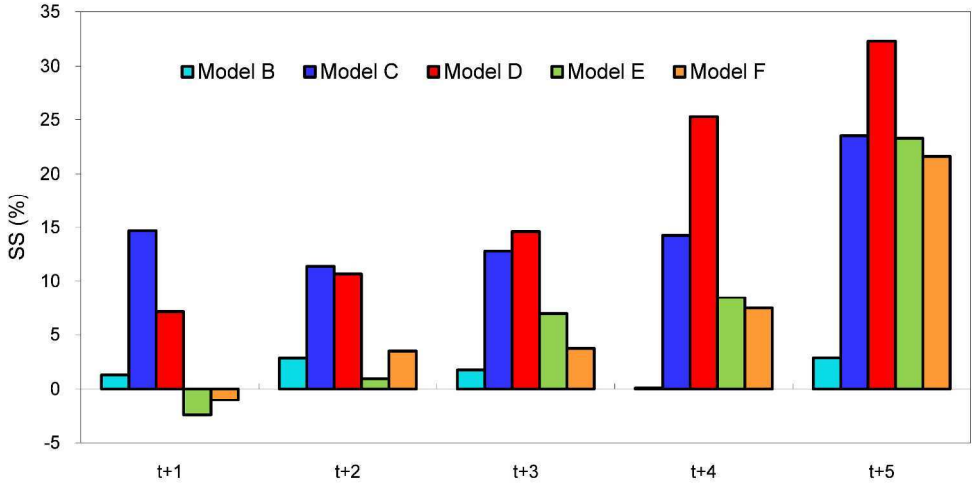

Fig. 6: Model skill score on test data set using Model A as the benchmark model.
1458x743mm (96 x 96 DPI)

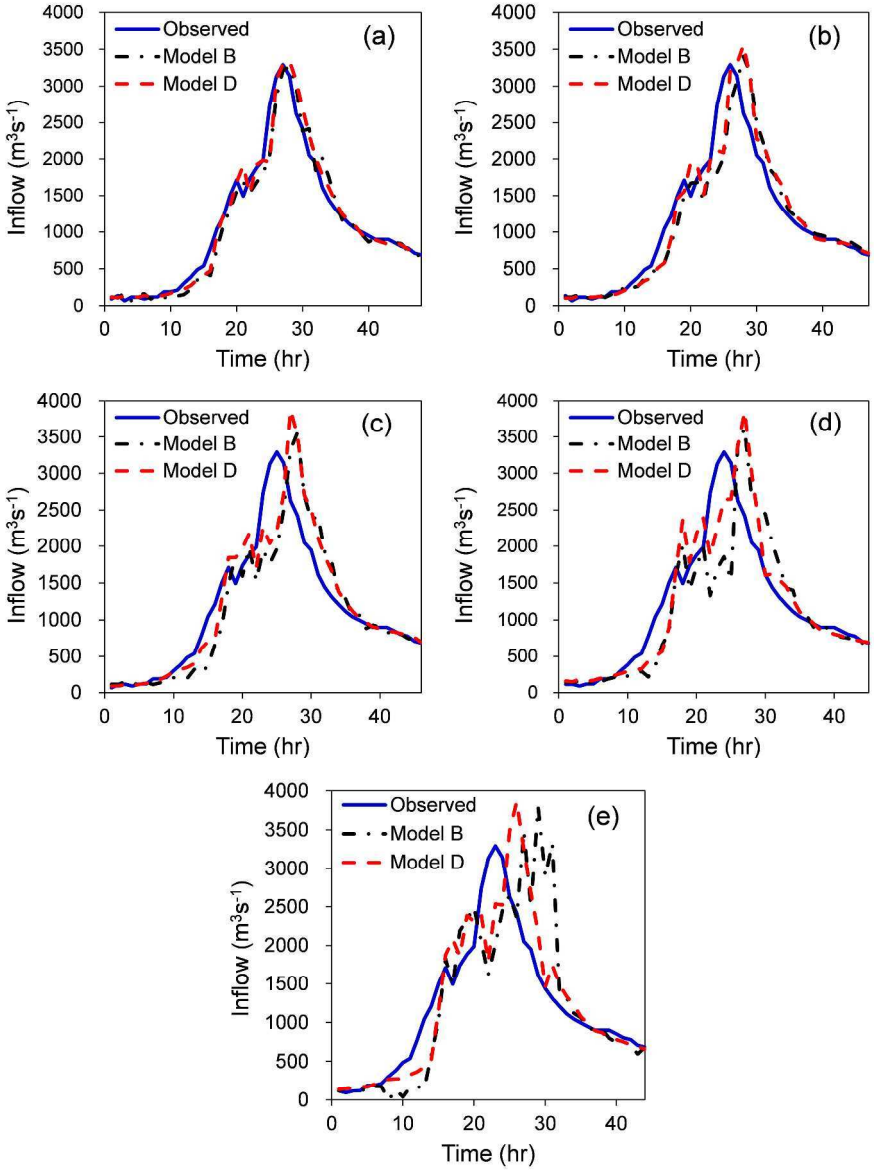

Fig. 7: Comparison of observed and predicted values for Model B and Model D at lead times (a) t+1; (b) t+2; (c) t+3; (d) t+4; (e) t+5.
1190x1587mm (96 x 96 DPI)

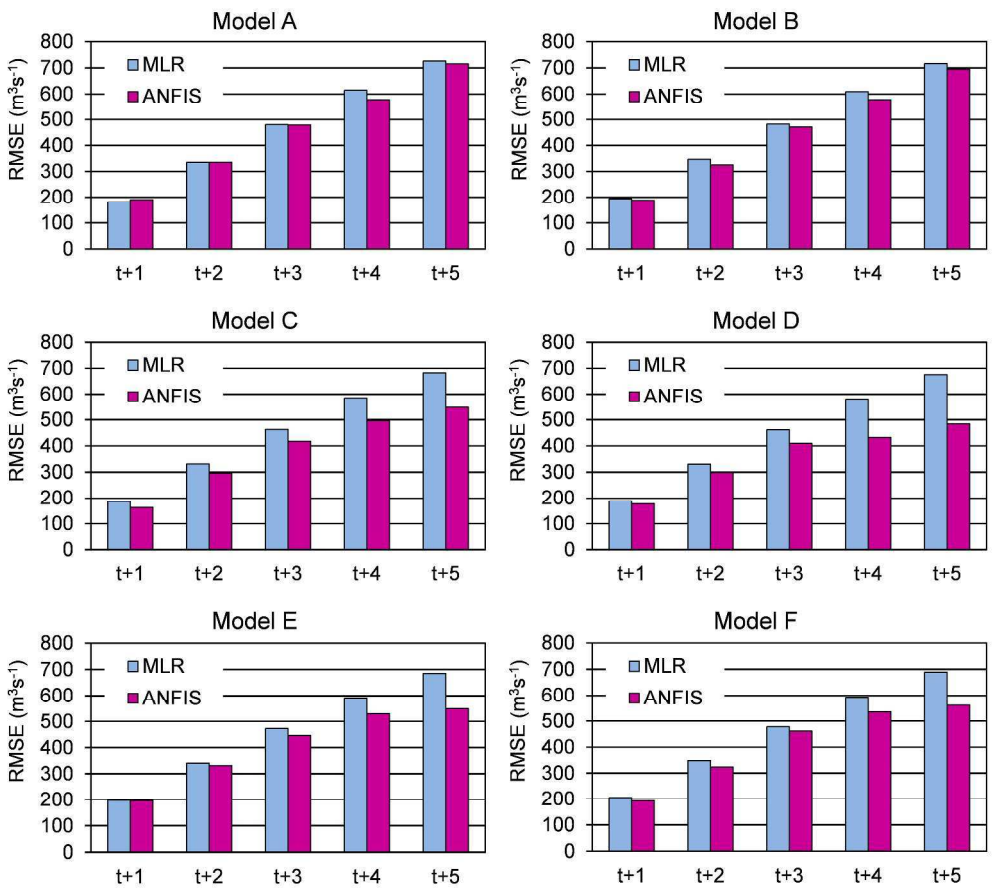

Fig. 8: MLR and ANFIS, statistical comparison on RMSE.
1314x1177mm (96 x 96 DPI)

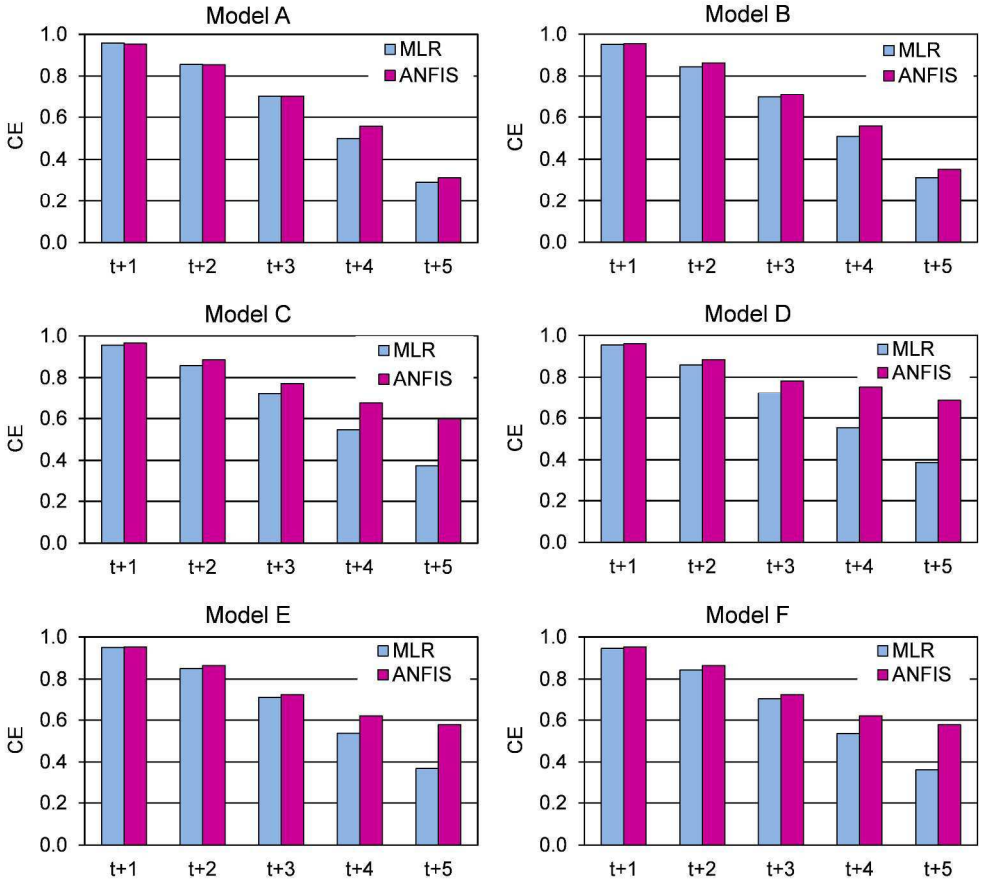

Fig. 9: MLR and ANFIS, statistical comparison on CE.
1318x1172mm (96 x 96 DPI)