



# Implications of uncertainty in inflow forecasting on reservoir operation for irrigation

K. S. Kasiviswanathan<sup>1</sup> · K. P. Sudheer<sup>2,3</sup> · Bankaru-Swamy Soundharajan<sup>4</sup> · Adebayo J. Adeloje<sup>5</sup>

Received: 26 September 2019 / Revised: 10 September 2020 / Accepted: 16 September 2020 / Published online: 6 October 2020  
© The International Society of Paddy and Water Environment Engineering 2020

## Abstract

Accurate and reliable forecasting of reservoir inflows is crucial for efficient reservoir operation to decide the quantity of the water to be released for various purposes. In this paper, an artificial neural network (ANN) model has been developed to forecast the weekly reservoir inflows along with its uncertainty, which was quantified through accounting the model's input and parameter uncertainties. Further, to investigate how the effect of uncertainty is translated in the process of decision making, an integrated simulation–optimization framework that consists of (i) inflow forecasting model; (ii) reservoir operation model; and (iii) crop simulation model was developed to assess the impacts of uncertainty in forecasted inflow on the irrigation scheduling and total crop yield from the irrigation system. A genetic algorithm was used to derive the optimal reservoir releases for irrigation and the area of irrigation. The proposed modeling framework has been demonstrated through a case example, Chittar river basin, India. The upper, lower, and mean of forecasted inflow from the ANN model were used to arrive at the prediction interval of the depth of irrigation, total crop yield, and area of irrigation. From the analysis, the ANN model forecast error of  $\pm 69\%$  to the mean inflow was estimated. However, the error to mean value of simulation for total irrigation, total yield, and area of irrigation was  $\pm 13.3\%$ ,  $\pm 6.5\%$ , and  $\pm 4.6\%$ , respectively. The optimizer mainly contributed to the reduction in the errors (i.e., maximizing the total production with the optimal water releases from the reservoir irrespective of inflow to the reservoir). The results from this study suggested that the information on the uncertainty quantification helps in better understanding the reliability of the systems and for effective decision making.

**Keywords** Artificial neural network · Crop simulation · Reservoir operation · Optimization · Uncertainty

## Introduction

Modeling and simulation are prevalent techniques for the effective management of water and agricultural production in the irrigation systems. The models can be from simple with few parameters to complex with several parameters depending on varying degrees of mathematical relationship. In irrigation modeling, the formulation of objective function has often been to maximize the crop yield with the optimal amount of water being released for irrigation, prevailing all other conditions such as soil nutrients, land suitability, etc., favorably. As there is no single model that can generically be applied everywhere, developing region-specific models are often preferred. However, the interaction among natural processes is complex to understand, and thus, the models developed generally contain a certain degree of uncertainty. These uncertainties mainly arise from the spatial and temporal variability of model inputs, parameters, and model structure (Cristiano et al. 2017). Hence, besides the calibration of

✉ K. S. Kasiviswanathan  
k.kasiviswanathan@wr.iitr.ac.in

<sup>1</sup> Department of Water Resources Development and Management, Indian Institute of Technology Roorkee, Roorkee 247667, India

<sup>2</sup> Department of Civil Engineering, Indian Institute of Technology Madras, Chennai 600036, India

<sup>3</sup> Department of Agricultural and Biological Engineering, Purdue University, West Lafayette, IN 47907, USA

<sup>4</sup> Department of Civil Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore 641112, India

<sup>5</sup> Institute for Infrastructure and Environment, Heriot-Watt University, Edinburgh EH14 4AS, UK

model parameters, quantification of uncertainty has become indispensable in any simulation studies for improving the reliability of model simulations.

In irrigation scheduling, model prediction uncertainty has been quantified through the uncertainty in the rainfall estimation (Burt et al. 1997; Chaubey et al. 1999) and the measurement error in evapotranspiration (Allen et al. 2011; Snyder et al. 2015). For quantifying the uncertainty, application of the Monte Carlo (MC) simulation method has gained significant attention (Prats and Picó 2010; Soundharajan et al. 2016), as it can combine various forms of uncertainty (e.g., input, model, and parameter uncertainties) to estimate the total prediction uncertainty. Specifically, the MC simulation is useful for analyzing the statistical inference from the modeled parameters. However, a prodigious computation, a low rate of model convergence, and the amount of sampling from joint probability distribution are the major constraints that limit the application of MC. It was noted that many attempts so far presented in the literature have focussed on quantifying the effects of uncertainty in the climate variables to the crop yield (Wang et al. 2017; Zhao et al. 2019). However, the impact of reservoir inflow forecast uncertainty on the irrigation water supply, as well as the total estimate of crop yield, should not be ignored as the flow dynamics vary due to high variability in the rainfall and catchment characteristics. In the process, the quantified uncertainty in the reservoir inflow can effectively be accounted into reservoir releases toward meeting the irrigation demands. These releases can further be linked with area irrigated, crop choice, and total crop production for quantifying the possible variations that are otherwise estimated as point values.

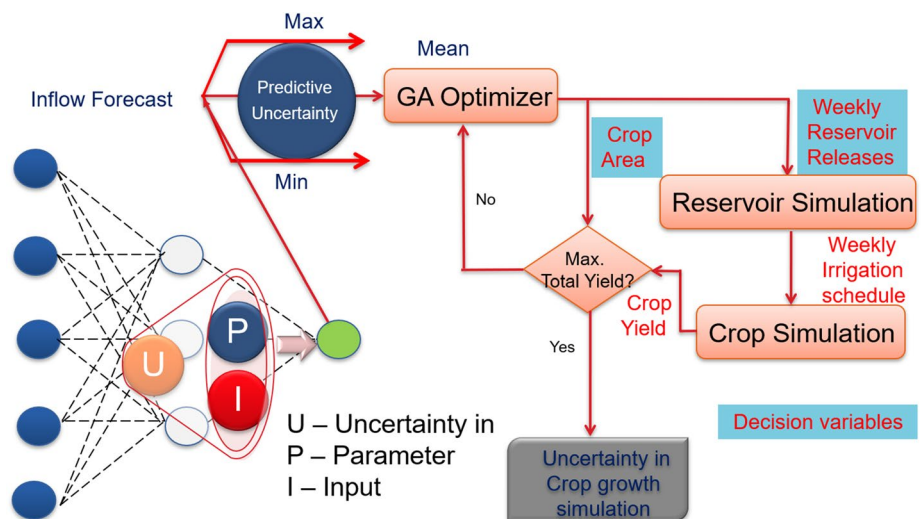
This paper proposed ANN models for forecasting the reservoir inflows along with the quantification of uncertainty in the forecasted values. Several methods, such as the delta technique (Chryssolouris et al. 1996), Bayesian methods

(Zhang et al. 2009, 2011), the Hessian matrix of the cost function for construction of prediction intervals (Papadopoulos et al. 2001), mean–variance estimation-based method (Nix and Weigend 1994), Bootstrap sampling-based methods (Tiwari and Chatterjee 2010), ensemble-based methods (Boucher et al. 2010; Araghinejad et al. 2011), heuristic-based methods (Han et al. 2007), fuzzy-based method (Alvisi and Franchini 2011) have been proposed/applied in ANN model for the quantification of uncertainty. However, all these methods are often limited to quantify only the parameter uncertainty of ANN models without further investigating how the effect of uncertainty is translated in the process of decision making (i.e., integrating them with the management models as presented in this paper). Hence, this paper presents a two-stage ensemble of simulation methods to account for the input and model parameter uncertainty for quantifying the uncertainty in the inflow forecast. Further, a simulation–optimization framework that combines the uncertainty interval of inflow forecast of the ANN models with management models (reservoir simulation and crop simulation) has been proposed to demonstrate the effect of uncertainty in the inflow forecast on the reservoir releases toward irrigation and the crop yield.

## Methodology

A block diagram of the proposed simulation–optimization framework is presented in Fig. 1. An objective function was formulated to maximize the total crop yield with an optimal supply of water from the reservoir storage. The decision variables were crop area and weekly reservoir releases (highlighted in blue color in Fig. 1). The modeling framework has four main components (see Fig. 1): (i) ANN weekly inflow forecast model along

**Fig. 1** Block diagram of simulation–optimization framework

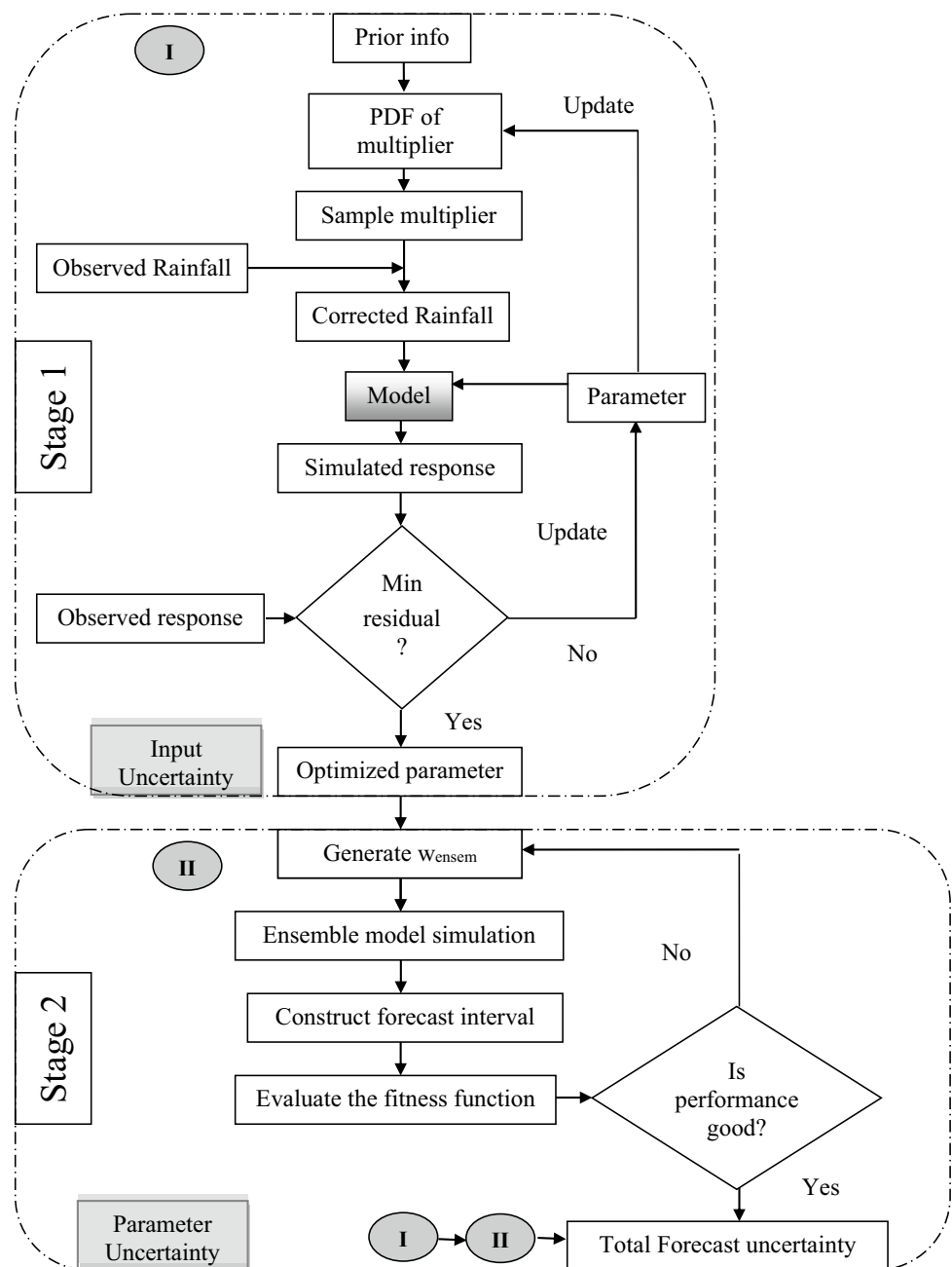


with quantification of forecast uncertainty (in the form of upper, lower, and average values) from the ensembles of simulation; (ii) genetic algorithm (GA) optimizer, to optimize the weekly reservoir releases and irrigated area; (iii) simulation of reservoir operation; and (iv) crop simulation model, which determines the crop yield for the given irrigation schedule (i.e., time of application and net depth of irrigation). The detailed explanation of each component is presented in the following subsections.

### Quantification of inflow forecast uncertainty in ANN models

A two-stage modeling framework, for accounting input and parameter uncertainty in stage 1 and 2, respectively, has been proposed to quantify the uncertainty in the forecasted inflow by the ANN models (see Fig. 2). It has been demonstrated that input uncertainty may change the statistical characteristics of hydrologic model parameters and therefore has significant implications on testing the reliability of the model (Vrugt et al. 2008). In this paper, rainfall has been chosen for the analysis as the accurate measurement/

**Fig. 2** Flowchart describing the quantification of total forecast uncertainty of ANN model



estimation of rainfall is often challenging. In stage 1, quantification of model input uncertainty was performed through randomly sampling the possible error values from the probability distribution. Note that the error in the rainfall could be attributed to various sources such as interpretation error, measurement error, and sampling error. Since the direct estimation of the true values of these errors is always challenging, indirectly forcing the error/noise is a standard practice in modeling (Kasiviswanathan et al. 2017). In principle, each rainfall estimate is associated with an independent error. Therefore, error in each rainfall value must be estimated independently during the model calibration. However, in such an approach, the dimensionality of the calibration procedure grows multi-fold. Consequently, the predictive capability of the model might deteriorate gradually due to over parameterization. Hence, it is necessary to develop an efficient modeling framework that quantifies input uncertainty, also ensuring the model of parsimony. In this paper, this was performed by forcing an error in the form of rainfall multiplier ( $\varphi_t$ ) mainly sampled from probability distribution function (pdf), which perturbs the measured inputs. Statistically, it is assumed that the error follows the normal distribution with zero mean and with certain values of standard deviation depending on the magnitude of error distribution (Kasiviswanathan et al. 2017). In the analysis, the forcing error was assumed to follow lognormal distribution as the multiplier has to be a positive fraction. However, the normal distribution can also be used, converting the negative values into positive ones. The statistical characteristics of the lognormal distribution [mean ( $\mu$ ) and standard deviation ( $\sigma$ ) of the pdf] were estimated within the calibration of the model itself. The ' $\varphi_t$ ' has to be sampled from the pdf for each rainfall values independently. As mentioned above, the mean value of the multiplier was fixed as 'zero,' and the standard deviation of the multiplier was identified through calibration. More details on the input uncertainty quantification that uses a similar approach in ANN rainfall-runoff models have been reported in Kasiviswanathan et al. (2017).

In stage 2, accounting for the parameter uncertainty of the ANN model was performed using an ensemble simulation approach (Kasiviswanathan et al. 2013). The ensemble of model outputs was simulated from perturbing the ANN model parameters obtained in stage 1. As the perturbation of parameters cannot be performed randomly, an optimization function was formulated to obtain the ensemble of model outputs. Thus, the overall objective at this stage was to obtain the narrow prediction interval, which also contains a maximum number of observations within the prediction interval. This was achieved through formulating a multi-objective optimization problem to generate ensembles of simulation, such as (i) minimize the mean square error (MSE) between the ensemble mean of ANN and observation (Eq. 1) and (ii) minimize the average width (AW) of

prediction interval (Eq. 1), and (ii) maximize the percentage of coverage (POC) that contains more number of observations in the forecasted interval (Eq. 2).

$$\text{Minimize} \begin{cases} \text{MSE} = \frac{1}{n} \sum_{t=1}^n \left[ y_t - \frac{1}{K} \sum_{k=1}^K \hat{y}_t^k \right]^2 \\ \text{AW} = \frac{1}{n} \sum_{t=1}^n \left[ \hat{y}_t^U - \hat{y}_t^L \right] \end{cases} \quad (1)$$

$$\text{Maximize} \left\{ \text{POC} = \left( \frac{1}{n} \sum_{t=1}^n c_t \right) * 100 \right. \quad (2)$$

where  $K$  is a total number of the ensemble of networks derived, and for each individual network, the  $t$ th pattern has a predicted value  $\hat{y}_t^k$  obtained from the  $k$ th network ( $k = 1, 2, \dots, K$ ).  $n$  is the total number of patterns used for constructing the prediction interval, and  $\hat{y}_t^U, \hat{y}_t^L$  are the upper and lower bound estimation of the  $t$ th pattern;  $c_t = 1$  if the observed values of target fall within the prediction band  $[\hat{y}_t^U, \hat{y}_t^L]$ , otherwise  $c_t = 0$ . More details on the parameter uncertainty quantification of ANN using an ensemble of simulation has been presented in Kasiviswanathan et al. (2013).

### ORYZA2000–Rice crop simulation model

Crop simulation models play a major role in irrigation water management practices (Kisekka et al. 2017). However, the effectiveness of such models depends on the accurate estimation of irrigation water requirement that decides the crop yield. ORYZA2000 is an eco-physiological crop simulation model developed for simulating the growth, phenological development, and yield of rice crop under potential and limited water and nitrogen conditions on a daily scale (Bouman et al. 2001). The model simulates the phenological development of the crop, such as biomass production and partitioning using average daily temperature and photoperiod on a daily timescale. The simulated biomass is partitioned by the model among the various parts of the crop (roots, leaves, stems, and panicles) using partitioning factors, which are determined through calibration. Please refer to Bouman et al. (2001) for the detailed description of the model along with the program source code.

The model requires inputs of soil, crop, weather, and management parameters (Bouman et al. 2001; Soundharajan and Sudheer 2013). The management inputs are (i) crop variety, (ii) spacing or plant population, (iii) transplanting depth, (iv) nursery duration, and (v) fertilizer and irrigation application. Soil properties include: (i) volumetric soil water content at saturation, (ii) field capacity and wilting point, (iii) depth of puddled soil, and (iv) saturated hydraulic conductivity. Rainfall and temperature are the weather

inputs during the growing season. The crop parameters are (i) relative leaf growth rate, (ii) leaf death rate, (iii) specific leaf area, (iv) crop development rates, and (v) biomass partitioning factors. Please note that all these parameters are required to be calibrated. However, as the motivation of this paper was to show the impact of inflow forecast uncertainty in the crop production, the calibration procedure and calibrated parameter values of the ORYZA2000 model are not presented in this paper (please refer Soundharajan and Sudheer 2013 for more information).

### Optimization -Simulation model formulation

The calibrated ORYZA2000 model was coupled with an optimization framework and a reservoir water balance simulation model (see Fig. 1) to optimize the reservoir releases for irrigation under water deficit situations. The objective of the optimization scheme was to maximize the total crop yield from the reservoir based on the irrigated command area. The optimal release from the reservoir throughout the cropping season was forced for the maximum possible cropping area. The fitness function formulated is as follows:

$$\text{Maximize, } Z = \left( \sum_{s=1}^n A_s Y_s \right) \tag{3}$$

where  $Z$  refers to the total crop yield from the irrigated command area in tons;  $A_s$  is the irrigated area in hectare;  $Y_s$  is the crop yield in tons per hectare. The cropping seasons are Kharif [June–October] and Rabi [October–February]. Decision variables were (i) weekly irrigation depths and (ii) area that can be irrigated. Crop yield ( $Y_s$ ) for a specific irrigation schedule was simulated by the ORYZA2000 model to evaluate the objective function. The candidate solutions of reservoir releases generated by the GA were subjected to reservoir water balance (Eq. 4), proposed by Loucks et al. (1981).

$$\left. \begin{aligned} (1 + a_t)S_{t+1} &= (1 - a_t)S_t + Q_t - R_t - O_t - A_o e_t \\ a_t &= \frac{A_s e_t}{2} \end{aligned} \right\} \text{ for } t = 1, 2, \dots, T \tag{4}$$

$$O_t = \begin{cases} S_{t+1} - S_{\max}; & \text{if } S_{t+1} > S_{\max} \\ 0; & \text{otherwise} \end{cases} \text{ for } t = 1, 2, \dots, T \tag{5}$$

$$\left. \begin{aligned} S_t &\leq S_{\max} \\ S_t &\geq S_o \end{aligned} \right\} \text{ for } t = 1, 2, \dots, T \tag{6}$$

where  $S_t$  and  $S_{t+1}$  are reservoir storage at the beginning and end of periods  $t$  and  $t + 1$ , respectively,  $m^3$ ;  $R_t$  is the gross reservoir release during  $t$ ,  $m^3$ ;  $Q_t$  is the reservoir inflow during  $t$ ,  $m^3$ ;  $O_t$  is the reservoir spill during  $t$ ,  $m^3$ ;  $A_o$  is the water spread area per unit volume of live storage,  $m^2/m^3$ ;  $A_o$  is the

water spread area at the dead storage level,  $m^2$ ;  $e_t$  is the evaporation loss from the reservoir,  $m$ ;  $T$  is the total number of periods in the simulation;  $S_{\max}$  is the reservoir gross storage capacity,  $m^3$ ;  $S_o$  is the dead storage capacity,  $m^3$ . The reservoir behavior simulation was performed on a daily time scale.

The net depth of irrigation has been computed as:

$$\left. \begin{aligned} NR_t &= \eta R_t \\ I_t &= \frac{NR_t}{10A_s} \end{aligned} \right\} \tag{7}$$

where  $\eta$  is the overall irrigation efficiency (60%), fraction;  $NR_t$  is the net reservoir release during ‘ $t$ ,’  $m^3$ ;  $A_s$  is the irrigated area during the season ‘ $s$ ,’ ha;  $I_t$  is the net irrigation during ‘ $t$ ,’ mm. For the complete formulation of an optimization scheme, readers can refer to Soundharajan and Sudheer (2009).

### Outline of the algorithm

The initial decision variables (weekly net irrigation,  $I_t$ ) were generated by the GA optimizer for the entire crop season along with irrigated area ( $A_s$ ).  $I_t$  was converted into reservoir releases ( $R_t$ ) using Eq. 7. The reservoir releases were evaluated using reservoir water balance for the feasibility of generated solutions. If any of the releases within the solution set violates the reservoir water balance (i.e., reservoir releases > reservoir storage), a constrain was forced to maintain the reservoir water balance (i.e., reservoir releases ≤ reservoir storage). Again, the adjusted  $R_t$  was converted back to  $I_t$  using Eq. 7. The final  $I_t$  schedules were given to the ORYZA2000 to simulate the crop growth and yield ( $Y_s$ ). The fitness function of the GA for each solution set was evaluated using Eq. 3. GA generated the population of next-generation solution sets based on the previous generation’s fitness function values through its operations (i.e., selection, crossover, and mutation). This iterative computation was continued until the specified number of generations reached. The final results were the solution set having the highest fitness function value (i.e., the combination of optimal reservoir releases and irrigated area that maximizes objective function).

### Study area and data

The proposed method has been demonstrated through data collected from tributaries of the Chittar river, Tamil Nadu, India. The Chittar river basin (Fig. 3) has a total area of about 1677  $km^2$  and is situated between 77° to 78° E latitude and 8.75° N to 9.25° N longitude.

The Chittar river is the largest tributary of the Tambaraparani river, which originates in the Western Ghats and flows eastwards about 80 km and joins into the

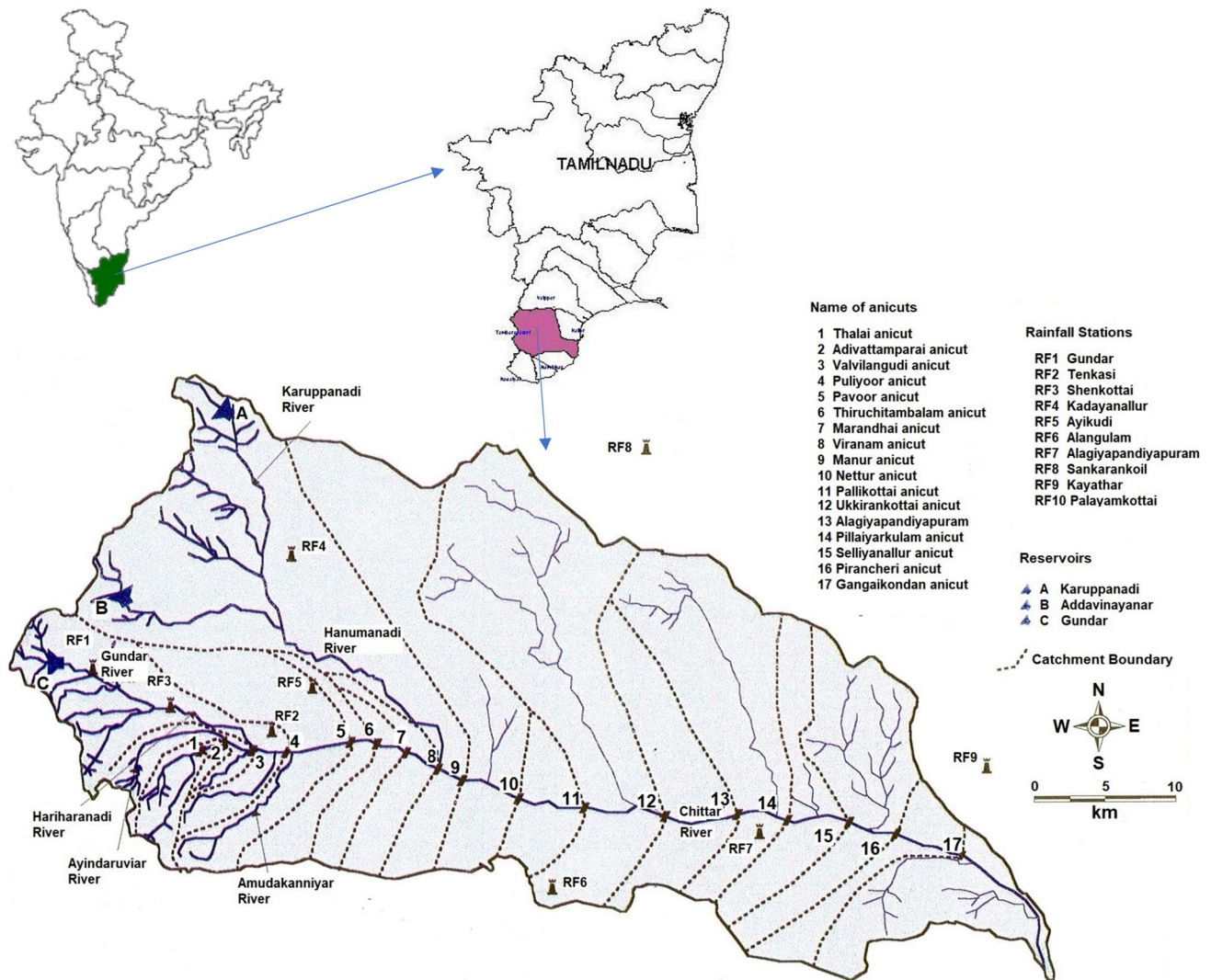


Fig. 3 Study area map ( Source: CWR 2001)

Tambaraparani river. The Chittar river has five tributaries, namely Ayndaruviar, Hariharanadhi, Aludhakanniar, Hanumanadhi, and Uppodai. The reservoir inflow data used in the modeling was collected from Hanumanadhi catchment. The catchment area of Hanumanadhi is 168.69 km<sup>2</sup>.

The Adavinainar kovil reservoir was constructed across the river Hanumanadhi with a capacity of 4.95 Mm<sup>3</sup>. The total dependable annual water yield at the dam site is 17 Mm<sup>3</sup>. The reservoir is meant to impound and utilize 9.85 Mm<sup>3</sup> in two fillings to augment irrigation facilities to the extent of 2250 ha under 16 anicuts across the Hanumanadhi river and supply water to 62 irrigation tanks. The flash floods are stored in all the tanks through respective anicuts (Soundharajan 2011). Only during the flash floods, water reaches the lower anicuts and then to the tanks under them. Since the canals are aligned along the contour, the irrigation return flow from the command areas of the upper anicuts

drains into the lower anicuts or main river. The seepage from the upper command areas also reaches the lower channels, making the water flow available in the channel even when the upstream head sluices are closed. This facilitates some of the farmers in the downstream locations to get water even during non-supply periods (Soundharajan 2011).

The opening of the reservoir is usually carried out either on June 1st of every year or when the storage in the reservoir reaches 1.835 Mm<sup>3</sup>. Any quantity of inflow to the reservoir after 1st of every month, which raises the storage of the reservoir over and above the specified level for that particular month, will be let out in the river through river sluice and shared by the water user associations of the anicuts system. The details of the reservoir, along with catchment information, are presented in Table 1.

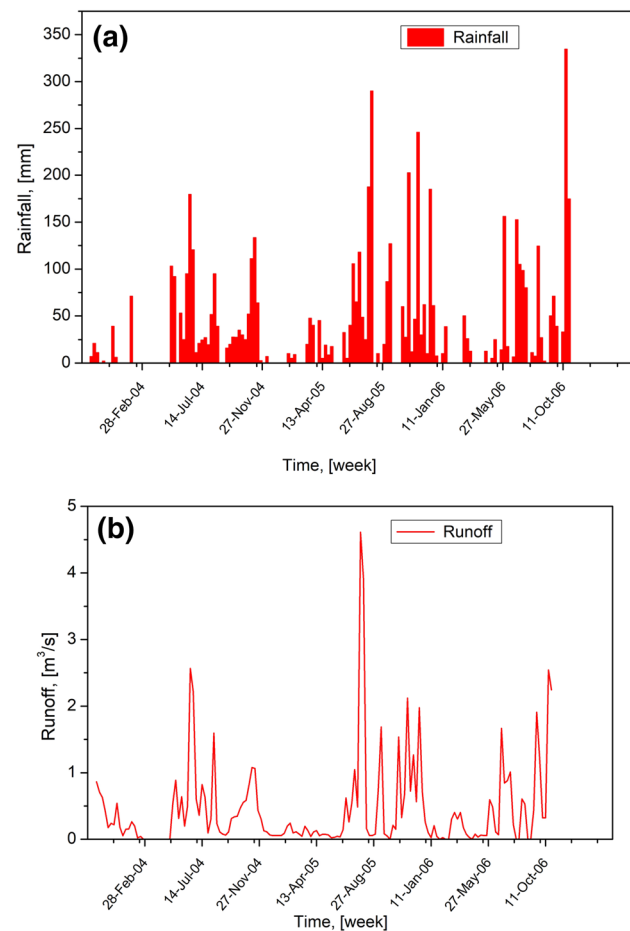
The rainfall and weather (i.e., temperature, humidity, wind speed, sunshine hours, and evaporation) data were

**Table 1** Details of the Adavinainar kovil reservoir irrigation system in the Chittar river basin

River	Hanumanadhi
Reservoir capacity (Mm <sup>3</sup> )	4.95
Direct and indirect command area, ha	2250
No. of anicuts downstream of reservoir	16
No. of tanks	62

collected from Palayamkottai meteorological observatory (maintained by India Meteorological Department).

The data (rainfall, runoff) collected from November 05, 2003, to October 26, 2006, were used to develop ANN inflow forecast models. Please note that the values were measured on a weekly basis. The rainfall and flow data are plotted in Fig. 4 to show their temporal variation. The statistics for the weekly values of rainfall and runoff are reported in Table 2.

**Fig. 4** Data (for the period November 5, 2003, to October 26, 2006) used for model development: **a** weekly rainfall; **b** weekly runoff**Table 2** Summary statistics of data used for model development

Statistical value	Rainfall [mm]	Runoff [m <sup>3</sup> /s]
Mean	38.44	0.44
Standard deviation	58.05	0.69
Skewness	2.50	3.13

## Results and discussion

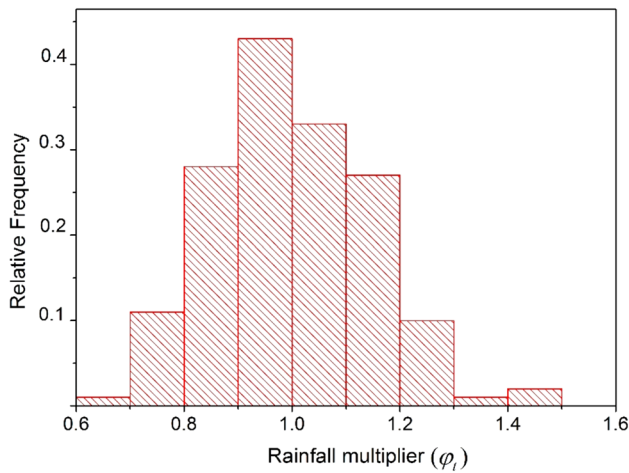
Statistical methods such as cross- and autocorrelation were used to determine the inputs of the ANN model (Sudheer et al. 2002). The identified inputs were:  $[R(t-1), R(t), Q(t-1)]$ , where  $R(t)$  represents the rainfall,  $Q(t)$  represents the runoff at any time ' $t$ ,' and the antecedent values were represented as  $t-1$ . The output was one week ahead forecast of runoff (i.e., inflow to the reservoir). Together, the combination of input and output forms the pattern to develop an ANN model. Out of the total number of input–output patterns comprised of weekly inflow and rainfall values, 110 patterns were used for model calibration, and 46 patterns were used for model validation. The next step was to identify the number of hidden neurons; for that, a trial and error method was used. It was found that two hidden neurons were sufficient to achieve desirable model performance compared to their counterparts, which had more than two hidden neurons. Sigmoidal and linear activation functions were used in hidden and output layers, respectively.

### Model input uncertainty

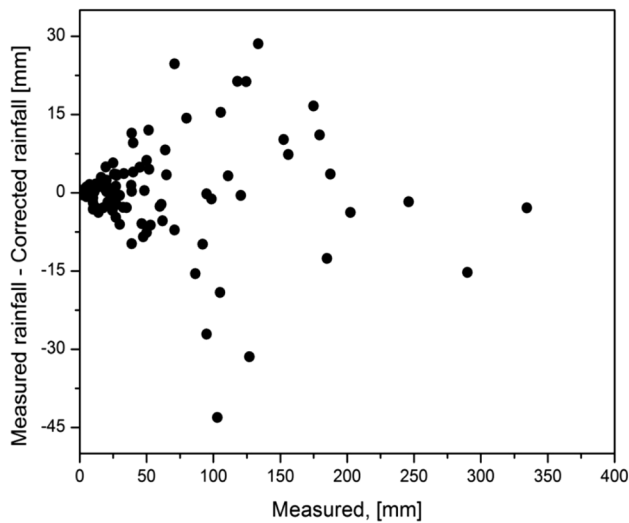
The forced error for correcting the rainfall was sampled from lognormal distribution for which the parameters, mean was kept as zero, and the standard deviation value was optimized along with the model calibration. Through this analysis, a standard deviation value of 0.15 was found to be optimal (Fig. 5).

The presented scatter plot (Fig. 6) shows the variation of bias along with the measured values of rainfall. The results suggested that the bias is consistent and varies linearly when the magnitude of rainfall increases. However, in the case of extreme values of rainfall, low bias was observed. The obvious reason could be the measurement of the extreme value with relatively less error. Overall, the average bias of  $\pm 10\%$  was estimated to the measured weekly rainfall values.

The model performance, including input uncertainty, is presented in Table 3. The performance of the model was assessed with different statistical measures such as coefficient of correlation (CC), Nash Sutcliffe efficiency (NSE), root mean square error (RMSE), and mean biased error (MBE). It is evident from Table 3 that the model performance was good in both calibration and validation with CC values of 0.89 and 0.85, respectively. The NSE



**Fig. 5** Histogram of rainfall multiplier sampled from lognormal distribution



**Fig. 6** Two-dimensional scatter plots of observed rainfall against corrected rainfall

statistics during calibration (78.15%) was slightly better than validation (64.19%). The positive values of MBE during calibration and validation indicated the model of under prediction. The RMSE obtained during calibration and validation were comparable. Thus, the model was consistent across these datasets.

**Table 3** Performance indices evaluated during Stage 1

Calibration				Validation			
CC	NSE (%)	RMSE (m <sup>3</sup> /s)	MBE (m <sup>3</sup> /s)	CC	NSE (%)	RMSE (m <sup>3</sup> /s)	MBE (m <sup>3</sup> /s)
0.89	78.15	0.34	0.08	0.85	64.19	0.36	0.11

### Model parameter Uncertainty

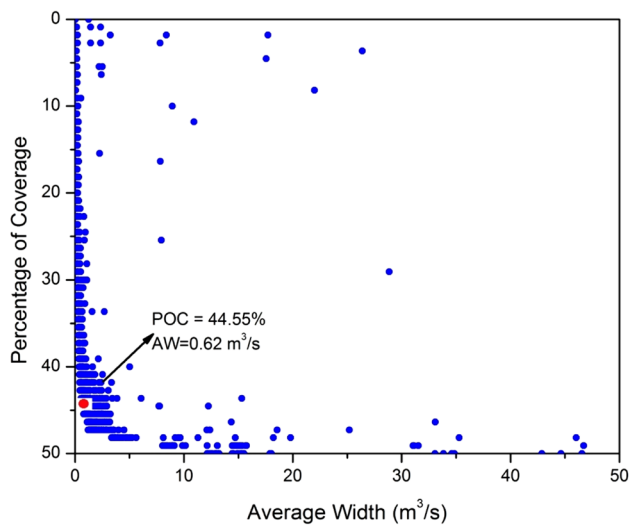
For creating an ensemble of simulation from the ANN model, the initial perturbation level of  $\pm 20\%$  to the parameters (weights and biases of ANN) was initiated in the GA. Figure 7 shows the convergence of ensembles within the multi-objective framework. It may be noted that each point represents the ensemble of simulation through which the forecast interval of the model output can be constructed. However, the forecast interval should be a trade-off solution having maximum POC with less AW. Further, it is expected that the ensemble mean should closely match the observed flow values. In such circumstances, it is a modeler choice to select the point for further analysis. In this study, a red point (see Fig. 7) indicated in Pareto-front was selected for constructing the prediction interval of model output.

The performance of the model and the forecast uncertainty is presented in Table 4. The model performance was consistent during the calibration and validation period with CC values of 0.88 and 0.87, respectively. The NSE value of 77.86% during calibration against the value of 68.82% during validation shows a slightly better estimate of the model during the calibration period. The positive MBE values of 0.04 m<sup>3</sup>/s and 0.08 m<sup>3</sup>/s during calibration and validation, respectively, show the underestimation of the model. The ensemble model has an acceptable level of AW of 0.61 m<sup>3</sup>/s during calibration and 0.53 m<sup>3</sup>/s during validation. The POC of 44.55% was observed during calibration, and 31.11% was observed during validation. It is preferred that the uncertainty interval of forecasted values has more POC with less AW (Kasiviswanathan et al. 2013; Ye et al. 2014). However, as illustrated in Fig. 7, the Pareto front could not converge more than the POC of 50%. The possible reason could be the minimum number of training data used in the modeling. Further, unlike hourly/daily flow values, the weekly flow has large variation, and that could be the potential reason for not being able to learn the process modeled with high POC.

The forecast interval of the hydrograph corresponding to selected points along with the observed values is presented in Fig. 8 for the calibration and validation periods. It may be noted from Fig. 8 that most of the low flow values are not captured in the forecast interval. However, the ensemble of model simulations was slightly above zero values.

It is also to be noted that the width of the forecast interval was wider during the rising and falling limb of the hydrograph, which could be attributed to higher variability among the ensembles. However, the forecast interval derived from the





**Fig. 7** Pareto-optimal front of optimization during ensemble creation

**Table 4** Performance indices for the mean of the ensemble of the selected points during stage 2

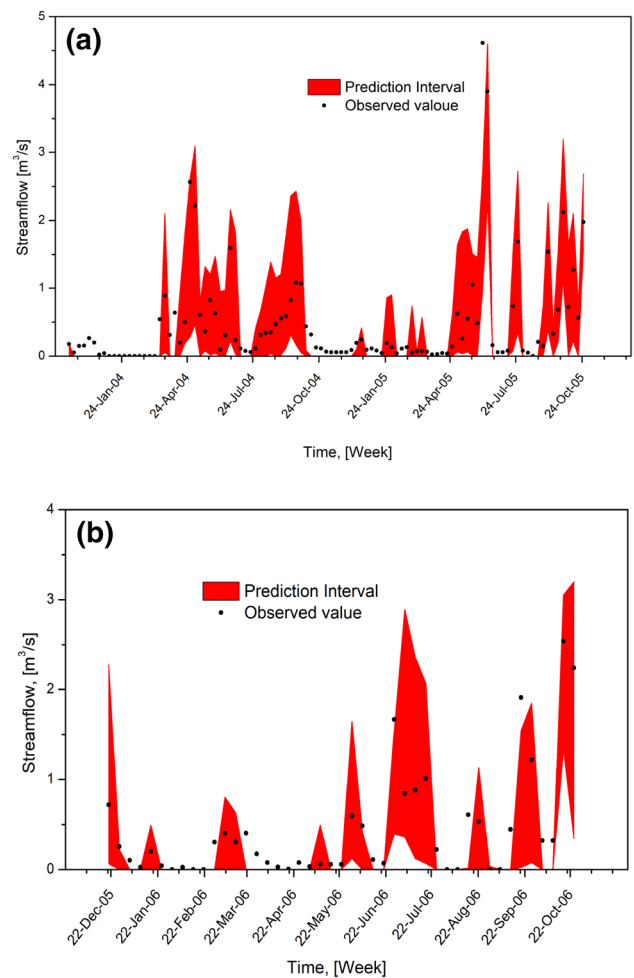
	Calibration	Validation
CC	0.88	0.87
NSE (%)	77.86	68.82
RMSE (m <sup>3</sup> /s)	0.34	0.34
MBE (m <sup>3</sup> /s)	0.04	0.08
POC (%)	44.55	31.11
AW (m <sup>3</sup> /s)	0.61	0.53

ensemble indicates an acceptable level of plausible error due to uncertainty in the model parameters. Please note that it is not based only on POC value. From the Pareto front (Fig. 7), it is clear that different point (solution) has different POC and AW values. As POC and AW indices conflict with each other, an acceptable level of plausible error must be an unbiased selection of point from the Pareto front (Fig. 7) having maximum POC and minimum AW.

**Uncertainty in optimal irrigation schedules and simulated total crop yield**

The total forecast uncertainty in the inflow from the ANN model was propagated into the simulation–optimization framework for making the decision on the irrigation schedules. From the real-time simulation–optimization framework, reservoir releases were estimated on a weekly basis (Fig. 9).

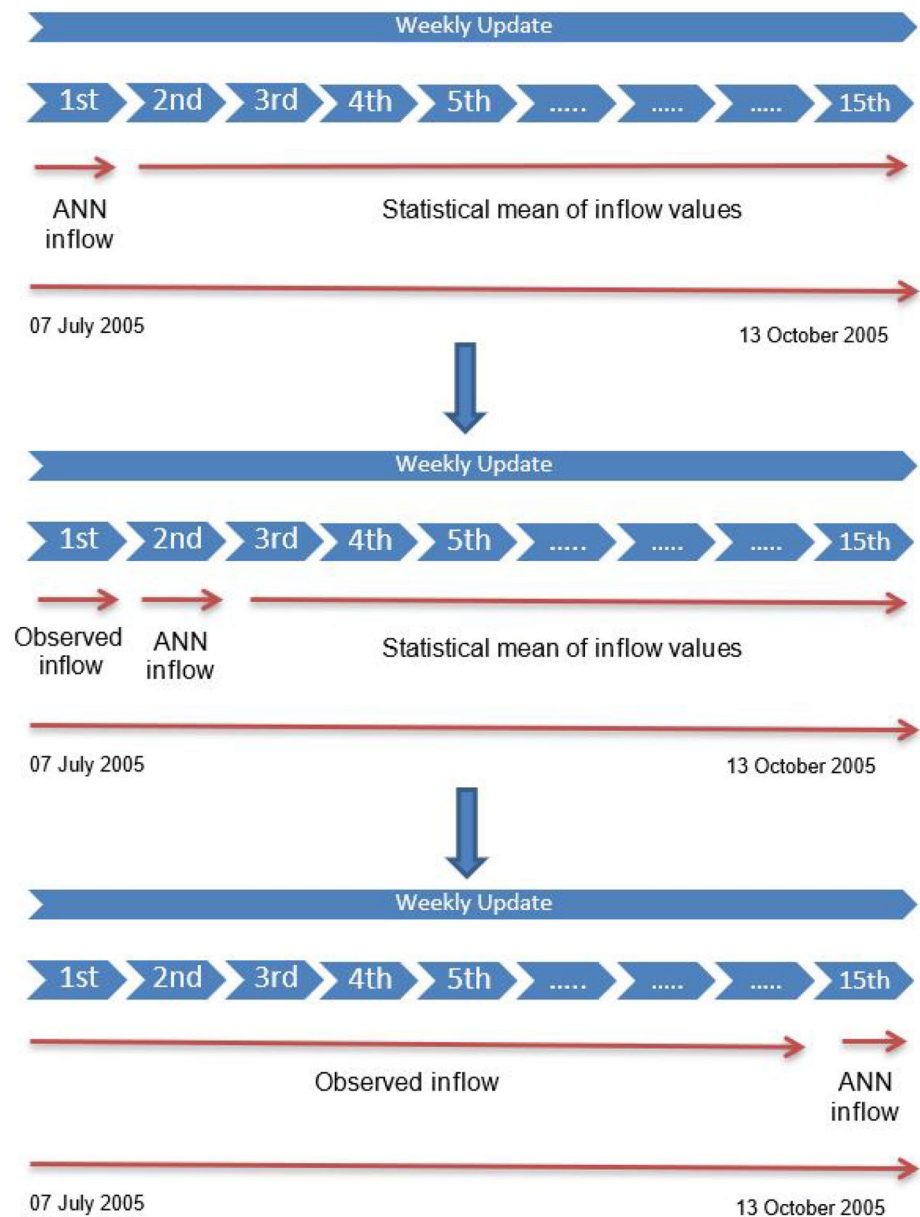
The number of decision variables was 16, which included 15 weekly water releases from the reservoir and area to be irrigated (as a fraction of total command area). The forecast lead time was fixed as 1-week; accordingly, the inflow



**Fig. 8** Prediction interval corresponding to the selected ensemble along with observed flows: **a** calibration (November 12, 2003, to December 15, 2005); **b** validation (December 22, 2005, to October 26, 2006)

was estimated from the calibrated ANN model along with associated forecast uncertainty. However, the crop simulation model requires the irrigation schedules of the complete season for estimating the total yield. Thus, statistical historical mean inflow information was used for the rest of the future period of the season since the higher lead time forecast of ANN was found to be poor. Each week, the optimizer updates its future decision variable (reservoir releases) based on ANN inflow forecast, and previous reservoir releases (if any) were considered as implemented (i.e., the decisions were made already). As the time progresses, actual inflows were considered for the previous weeks (i.e., weeks for which the release decision was already implemented), and inflow forecasts were made only for the rest of the future weeks to use in the simulation as shown in Fig. 9. At the end of each week, the inflow was forecasted from the ANN model for the next one-week lead time, and accordingly, the reservoir releases were optimized. Thus, the

**Fig. 9** Illustration of real-time reservoir operation [The numbers in the arrow refer to the number of weeks from transplanting to maturity (15 weeks), and for this duration weekly optimal irrigation schedules have to be developed. ANN model yields the forecasted inflow for one week ahead. The remaining values are the actual observed inflow and statistical mean of historical inflow for the past and future periods, respectively]



whole framework was updated on a weekly basis. The whole analysis was carried out for three different scenarios using (i) the mean inflow forecast of ANN, (ii) the upper inflow forecast of ANN, and (iii) the lower inflow forecast of ANN. Based on this, the GA finds the optimized values of reservoir releases for ensuring maximum crop yield.

The annual mean and standard deviation of inflow values were calculated and found that there is no appreciable difference among the years from November 2003 to October 2006 in the flow dynamics (Fig. 4). Thus, the model performance was limited to test the data of Kharif 2005 for estimating the optimal irrigation scheduling and cropping area under rice crop. It was evident from the operational framework that the estimated mean prediction of total yield (6154 tons), the total amount of irrigation (24.33 cm), and area of

cultivation (2137 ha) is comparable to the planning framework modeled with actual inflow values (total yield: 6298 tons; total irrigation: 35.05 cm; area cultivated: 1984 ha) as shown in Table 5. This indicates a reasonably accurate forecast of ANN models, which in turn closely matches the total yield based on the irrigation. It may be noted that the upper and lower values of inflow produced higher/lower yield, respectively, as one can expect (see Table 5). Based on the upper and lower value inflow forecast by ANN models, the error to the mean inflow simulation was estimated as  $\pm 69\%$ . This larger bias could be due to the smaller number of training data available for the ANN model. It may be noted that this error was an average value of complete flow series across calibration and validation data. However, the same level of error was not reflected in the total

**Table 5** The summary statistics of estimated uncertainties in irrigation and simulated crop yield

Statistical measures	Operational framework			Planning framework	
	Mean simulated value	Upper simulated value	Lower simulated value	Historical mean value	Observed value
Total yield (tons)	6154	6333	5529	5590	6298
Suggested irrigation (cm)	24.33	32.66	26.19	46.23	35.05
Suggested cultivable area (hectare)	2137	2175	1980	2219	1984

crop yield simulation. Consequently, the error to the mean value of simulated total yield, total irrigation, and area of irrigation were  $\pm 6.5\%$ ,  $\pm 13.3\%$ , and  $\pm 4.6\%$ , respectively. The reduction in error may be attributed to the optimizer that can still try to maximize production with the reservoir releases irrespective of inflow to the reservoir. In addition, the reason for such reduction in error might be, the uncertainty was considered only in inflow forecast, and all other variables were assumed to be a deterministic value. In the case of lower simulated values, the lesser the magnitude in simulated crop yield, the amount of irrigation may be attributed to lower inflow to the reservoir. Consequently, the reduction in reservoir releases resulted in less production. In the case of the planning framework, simulations were performed using observed inflow values of Kharif, 2005, and historical mean inflow values (average values computed using the complete dataset of the year 2003–2006). The total yield with less amount of irrigation, i.e., 35.05 cm, resulted by the simulations performed with observed value (i.e., 6298 tons) is higher than the historical mean value simulation (5590 tons). These results suggested that the excess supply of water (11 cm) in historical mean value simulation might not have been used during the water-sensitive growth stages of the crop, thus reducing overall total yield even with a higher cultivated area.

The simulated mean values from the operational framework are comparable with the planning framework with observed values. However, in the operational framework, less amount of irrigation (24.33 cm) was applied compared to the planning framework (35.05 cm). The possible reason might be that the optimizer updates the amount of release on a weekly basis rather including the complete observed flow series. Thus, the operational framework, along with uncertainty information, can better be relied on in the case of decision making.

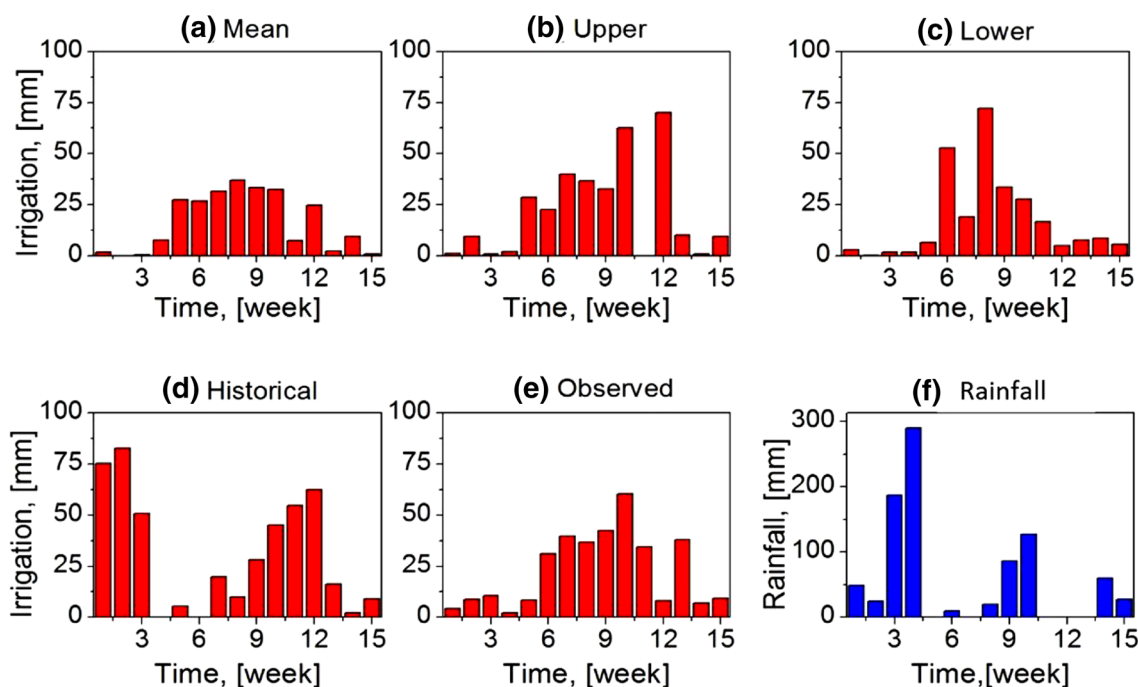
The mean, upper, and lower simulated irrigation releases are presented in Fig. 10 along with the planning framework observed values and historical mean value simulation during the Kharif season, 2005. It is evident from Fig. 10 that immediate weeks after transplantation, and before harvesting stages, the amount of irrigation required was less compared to mid-growth periods as expected. Due to sufficient rainfall during the first four weeks (see Fig. 10f), the

operational framework suggested no or very little irrigation (see Fig. 10a–c). The ‘historical mean’ based irrigation schedule applied a significant amount of irrigation (see Fig. 10d), leading to higher total irrigation (46.23 cm) with lower total yield (5590 tons) since the crop does not need that much water during the early stages. In the case of upper simulation (Fig. 10b), higher total yield and area of irrigation could be attributed due to the higher irrigation at 7th and 12th weeks, coinciding with panicle initiation and grain filling stages, respectively.

In the case of the lower simulated values, the less yield might be due to the non-availability of sufficient irrigation during the panicle initiation (7th week) as well as grain filling (12th week) stages. The amount of irrigation follows similar patterns in the case of operational framework, mean simulation and planning framework-observed value simulation. This indicates the comparably better forecast of inflow values by ANN that closely matches with observed values. The quantity of irrigation water triggered by the proposed model is based on the sensitivity of the crop growth stage to the final yield as well as the water available in the reservoir (Soundharajan 2011).

## Summary and conclusions

An ANN model was developed to forecast a one-week ahead reservoir inflow. A two-stage ensemble simulation approach was used in ANN to account for input and parameter uncertainty for quantifying the uncertainty in the forecasted inflow values. A simulation–optimization framework was proposed to demonstrate the implications of the forecast uncertainty resulted from ANN models in the decision making (i.e., irrigation water management in this case). Thus, this framework integrated an ANN inflow forecast with reservoir simulation and crop simulation models for deriving the optimal irrigation schedule and area to be irrigated (i.e., operational framework) under water deficit condition. Similarly, observed reservoir inflow data were used (i.e., planning framework) to derive the above decisions for comparison. The data from the Chittar river basin, India, were used to demonstrate the proposed modeling framework. It was found that the performance of the operational framework is better



**Fig. 10** Uncertainties in estimated irrigation after transplanting during Kharif season between July 07, 2005, to October 13, 2005 **a** mean inflow simulation, **b** upper inflow simulation, **c** lower inflow simula-

tion, **d** historical mean inflow simulation, **e** observed inflow simulation, **f** observed rainfall (882 mm) for the corresponding period

than planning frameworks that support the effectiveness of the proposed framework. ANN model resulted in a forecast error of  $\pm 69\%$  to the mean inflow. However, comparatively a less error of  $\pm 6.5\%$ ,  $\pm 13.3\%$ , and  $\pm 4.6\%$  was found in the total yield, total irrigation, and area of irrigation, respectively. The reduction in error was mainly contributed by the optimizer that tries to maximize the production with the reservoir releases irrespective of inflow to the reservoir. The reliability of the proposed approach can be evaluated with upper and lower bounds of simulated values and based on which the decisions can be made.

**Acknowledgements** Part of the work reported in this paper has been funded by the Science and Engineering Research Board (SERB) of India under Impacting Research innovation and Technology (IMPRINT-2) initiative (IMP/2018/002012). The authors would like to thank two anonymous reviewers and the Editor for their insightful comments which helped significantly improving the quality of the paper.

## References

- Allen RG, Pereira LS, Howell TA, Jensen ME (2011) Evapotranspiration information reporting: i. Factors governing measurement accuracy. *Agric Water Manag* 98:899–920
- Alvisi S, Franchini M (2011) Fuzzy neural networks for water level and discharge forecasting with uncertainty. *Environ Model Softw* 26:523–537. <https://doi.org/10.1016/j.envsoft.2010.10.016>
- Araghinejad S, Azmi M, Kholghi M (2011) Application of artificial neural network ensembles in probabilistic hydrological forecasting. *J Hydrol* 407:94–104. <https://doi.org/10.1016/j.jhydrol.2011.07.011>
- Boucher MA, Laliberte JP, Anctil F (2010) An experiment on the evolution of an ensemble of neural networks for streamflow forecasting. *Hydrol Earth Syst Sci* 14:603–612
- Bouman BAM, Kropff MJ, Tuong TP, Wopereis MCS, ten Berge HFM, van Laar HH (2001) ORYZA2000. Modeling lowland rice. International Rice Research Institute, Los Banos
- Burt C, Clemmens A, Strelkoff T, Solomon K, Bliesner R, Hardy L, Howell T, Eisenhauer D (1997) Irrigation performance measures efficiency and uniformity. *J Irrig Drain Eng* 123:423–442
- Chaubey I, Haan CT, Grunwald S, Salisbury JM (1999) Uncertainty in the model parameters due to spatial variability of rainfall. *J Hydrol* 220:48–61. [https://doi.org/10.1016/S0022-1694\(99\)00063-3](https://doi.org/10.1016/S0022-1694(99)00063-3)
- Chryssolouris G, Lee M, Ramsey A (1996) Confidence interval prediction for neural network models. *IEEE Trans Neural Netw* 7:229–232
- Cristiano E, ten Veldhuis M-C, de Giesen N-V (2017) Spatial and temporal variability of rainfall and their effects on hydrological response in urban areas—a review. *Hydrol Earth Syst Sci* 21:3859–3878. <https://doi.org/10.5194/hess-21-3859-2017>
- CWR (2001) Study on changes in cropping pattern in drought prone Chittar basin of Tambaraparani river in Tamilnadu: Research Report. Centre for Water Resources, Anna University, Chennai, India
- Han D, Kwong T, Li S (2007) Uncertainties in real-time flood forecasting with neural networks. *Hydrol Process* 21:223–228. <https://doi.org/10.1002/hyp.6184>
- Kasiviswanathan KS, Cibin R, Sudheer KP, Chaubey I (2013) Constructing prediction interval for artificial neural network

- rainfall runoff models based on ensemble simulations. *J Hydrol* 499:275–288. <https://doi.org/10.1016/j.jhydrol.2013.06.043>
- Kasiviswanathan KS, Sudheer KP, He J (2017) Probabilistic and ensemble simulation approaches for input uncertainty quantification of artificial neural network hydrologic models. *Hydrol Sci J* 63:101–113. <https://doi.org/10.1080/02626667.2017.1393686>
- Kisekka I, DeJonge KC, Ma L, Paz J, Douglas-Mankin K (2017) Crop modelling applications in agricultural water management. *Trans ASABE* 60:1959–1964
- Loucks DP, Stedinger JR, Haith DA (1981) *Water resources systems planning and analysis*. Prentice-Hall, Englewood Cliffs
- Nix DA, Weigend AS (1994) Estimating the mean and variance of the target probability distribution. In: *Proceedings of 1994 IEEE international conference on neural networks (ICNN'94)*. Orlando, FL, USA, vol 1, pp 55–60. <https://doi.org/10.1109/ICNN.1994.374138>
- Papadopoulos G, Edwards P, Murray A (2001) Confidence estimation methods for neural networks: a practical comparison. *IEEE Trans Neural Netw* 12:1278–1287
- Prats AG, Picó SG (2010) Performance evaluation and uncertainty measurement in irrigation scheduling soil–water balance approach. *J Irrig Drain Eng* 136:732–743. [https://doi.org/10.1061/\(ASCE\)IR.1943-4774.0000245](https://doi.org/10.1061/(ASCE)IR.1943-4774.0000245)
- Snyder R, Pedras C, Montazar A, Henry J, Ackley D (2015) Advances in ET-based landscape irrigation management. *Agric Water Manag* 147:187–197. <https://doi.org/10.1016/j.agwat.2014.07.024>
- Soundharajan B (2011) *Deficit irrigation scheduling under water deficit conditions using crop simulation models*. PhD Thesis, Indian Institute of Technology Madras, India
- Soundharajan B, Sudheer KP (2009) Deficit irrigation management for rice using crop growth simulation model in an optimization framework. *Paddy Water Environ* 7:135–149. <https://doi.org/10.1007/s10333-009-0156-z>
- Soundharajan B, Sudheer KP (2013) Sensitivity analysis and auto-calibration of ORYZA2000 using simulation–optimization framework. *Paddy Water Environ* 11:59–71. <https://doi.org/10.1007/s10333-011-0293-z>
- Soundharajan B, Adeloye AJ, Remesan R (2016) Evaluating the variability in surface water reservoir planning characteristics during climate change impacts assessment. *J Hydrol* 538:625–639. <https://doi.org/10.1016/j.jhydrol.2016.04.051>
- Sudheer KP, Gosain AK, Ramasastri KS (2002) A data-driven algorithm for constructing artificial neural network rainfall–runoff models. *Hydrol Process* 16:1325–1330. <https://doi.org/10.1002/hyp.554>
- Tiwari MK, Chatterjee C (2010) Development of an accurate and reliable hourly flood forecasting model using wavelet-bootstrap-ANN (WBANN) hybrid approach. *J Hydrol* 394:458–470. <https://doi.org/10.1016/j.jhydrol.2010.10.001>
- Vrugt JA, ter Braak CJF, Clark MP, Hyman JM, Robinson BA (2008) Treatment of input uncertainty in hydrologic modelling. Doing hydrology backward with Markov chain Monte Carlo simulation. *Water Resour. Res* 44:W00B09. <https://doi.org/10.1029/2007WR006720>
- Wang E, Martre P, Zhao Z, Ewert F, Maiorano A, Rötter R, Kimball B, Ottman M, Wall G, White J, Reynolds M, Alderman P, Aggarwal PK, Anothai J, Basso B, Biernath C, Cammarano D, Challinor A, De Sanctis G, Asseng S (2017) The uncertainty of crop yield projections is reduced by improved temperature response functions. *Nat Plants*. <https://doi.org/10.1038/nplants.2017.102>
- Ye L, Zhou J, Zeng X, Guo J, Zhang X (2014) Multi-objective optimization for construction of prediction interval of hydrological models based on ensemble simulations. *J Hydrol* 519:925–933. <https://doi.org/10.1016/j.jhydrol.2014.08.026>
- Zhang X, Liang F, Srinivasan R, Van Liew M (2009) Estimating uncertainty of streamflow simulation using Bayesian neural networks. *Water Resour Res* 45:1–16. <https://doi.org/10.1029/2008WR007030>
- Zhang X, Liang F, Yu B, Zong Z (2011) Explicitly integrating parameter, input, and structure uncertainties into Bayesian neural networks for probabilistic hydrologic forecasting. *J Hydrol* 409:696–709. <https://doi.org/10.1016/j.jhydrol.2011.09.002>
- Zhao Y, Wang C, Zhang Y (2019) Uncertainties in the effects of climate change on maize yield simulation in Jilin Province: a case study. *J Meteorol Res* 33:777–783. <https://doi.org/10.1007/s13351-019-8143-9>