Stoch Environ Res Risk Assess (2017) 31:1659–1670
DOI 10.1007/s00477-016-1369-5

REVIEW

Methods used for quantifying the prediction uncertainty of artificial neural network based hydrologic models

K. S. Kasiviswanathan[1] · K. P. Sudheer[2,3]

Published online: 8 December 2016

Abstract Application of artificial neural network (ANN) models has been reported to solve variety of water resources and environmental related problems including prediction, forecasting and classification, over the last two decades. Though numerous research studies have witnessed the improved estimate of ANN models, the practical applications are sometimes limited. The black box nature of ANN models and their parameters hardly convey the physical meaning of catchment characteristics, which result in lack of transparency. In addition, it is perceived that the point prediction provided by ANN models does not explain any information about the prediction uncertainty, which reduce the reliability. Thus, there is an increasing consensus among researchers for developing methods to quantify the uncertainty of ANN models, and a comprehensive evaluation of uncertainty methods applied in ANN models is an emerging field that calls for further improvements. In this paper, methods used for quantifying the prediction uncertainty of ANN based hydrologic models are reviewed based on the research articles published from the year 2002 to 2015, which focused on modeling streamflow forecast/prediction. While the flood forecasting along with uncertainty quantification has been frequently reported in applications other than ANN in the literature, the uncertainty quantification in ANN model is a recent progress in the field, emerged from the year 2002. Based on the review, it is found that methods for best way of incorporating various aspects of uncertainty in ANN modeling require further investigation. Though model inputs, parameters and structure uncertainty are mainly considered as the source of uncertainty, information of their mutual interaction is still lacking while estimating the total prediction uncertainty. The network topology including number of layers, nodes, activation function and training algorithm has often been optimized for the model accuracy, however not in terms of model uncertainty. Finally, the effective use of various uncertainty evaluation indices should be encouraged for the meaningful quantification of uncertainty. This review article also discusses the effectiveness and drawbacks of each method and suggests recommendations for further improvement.

Keywords Artificial neural network · Forecasting · Hydrologic models · Prediction · Uncertainty

1 Introduction

Modeling hydrologic processes help in effective decision making for the management, planning and operational aspects of water resources. The mathematical modeling with improved accuracy considering various components of watershed is not a simple task as it contains various levels of variability in the space and time, consequently which leads the hydrologic processes to be highly nonlinear and complex in nature. Till date, a plethora of hydrologic models have been developed and reported in the literature, that include the complex physically based distributed models (Jayakrishnan et al. 2005), conceptual models (Kitanidis and Bras 1980), simple linear regression

✉ K. S. Kasiviswanathan
vishwaiitr@gmail.com

[1] Department of Civil Engineering, Schulich School of Engineering, University of Calgary, 2500 University Drive NW, Calgary T2N 1N4, Canada

[2] Department of Civil Engineering, Indian Institute of Technology Madras, Chennai 600 036, India

[3] Department of Agricultural and Biological Engineering, Purdue University, West Lafayette, IN 47906, USA

(Schreiber and Kincaid 1967) and non-linear data driven models (Cobaner 2011). Despite the physics and conceptual hydrologic models has an advantage of capturing the actual physics of the system entirely or partly, the scarcity of data in most of the regions is a major issue that limits the application of these models. In contrast, the regression based models can be developed with limited available information either using simple liner regression or non-linear data driven models based on the complexity of the process. Several types of data driven models have been proposed, in which the artificial neural network (ANN) models have gained significant attention owing to their simplicity in the model development and it produces acceptable level of accuracy in the model prediction (Citakoglu 2015). The ANN models, in general are characterized as universal approximation functions mainly used for deriving the unknown relationship between the variables of interest. There is a progressive development towards improving the performance of ANN models. Consequently, various research studies have been reported for developing different approaches for determining model inputs, architecture, sophisticated training algorithms, criteria for the model training, and so on across diverse fields. All these efforts have made ANN as an alternate promising modeling tool in water resources and environmental studies. Hence, the application of ANN for predicting and forecasting of water resources variables have been well established in last two decades.

Despite the improved prediction, the major criticism that ANN models experience are (a) lack of transparency (Abrahart et al. 2010), (b) the model parameters hardly explain the physical mechanism of the underline processes, (c) the stochastic nature of ANN model reproduce no identical results unless carefully designed (Elshorbagy et al. 2010a, b). It is also reported that the point prediction of ANN models has limited value owing to the variability/uncertainty in the system being modeled (Kasiviswanathan et al. 2013), which necessitates the quantification of uncertainty. However, the quantification of uncertainty in ANN is still a challenging task owing to its parallel computing architecture, which consequently limits the application in solving the real world problems.

It is noted that number of review articles have addressed the potential of ANN in hydrologic modeling (Maier and Dandy 2000; ASCE 2000a, b). The major focus of all these review articles was pertained to the application of ANN models for solving variety of problems discussing the advantages and shortcomings. The review article by Maier et al. (2010) reported different methods used for the development of ANN models (i.e. input, model architecture, structure selection, data division and algorithms for model training) that focused on prediction of water resources variables in the river system. One of the major recommendations by them was to continue developing new methods considering the best way of incorporating the uncertainty into ANN models. Hence, the main focus of this review paper is on the different methods applied to quantify the uncertainty of ANN models, their complexity, advantages, shortcoming and to suggest further investigations required, thereby increasing the reliability of these models. During the initial screening of the articles, it was found that majority of the study has demonstrated the uncertainty analysis in ANN models through flood forecast/rainfall-runoff modeling. Therefore, the scope of this paper is limited to uncertainty methods applied in ANN based rainfall- runoff and flood forecasting models. The objective of this review paper is to analyze the computational efficiency, difficulty of implementation and fulfillment of statistical assumptions of various methods and to address the key issues, which require further improvements. A total of 36 research articles were selected from well-known international journals spread over the period between 2002 and 2015. While uncertainty analysis of other models than ANN was in place for long time, the ANN applications received attention on uncertainty analysis only from the year 2002 onwards. It may be noted that this article does not investigate the general aspects of ANN models, such as analyzing the model performance, selection of model inputs, architecture etc. For more details on the basics of ANN model, the readers can refer articles and textbooks published (Flood and Kartam 1994; Bishop 2004). However it is important to investigate the influence of varying the overall ANN architecture in terms of their connection, activation function, training algorithm on the quantification of model prediction uncertainty, hence it is discussed in this review paper.

The remainder of this paper is organized as follows. The selection of research article for the review based on different combination of search in the database is presented in Sect. 2. In Sect. 3, a detailed description of different sources of uncertainty is described. In Sect. 4, the types of ANN models and their influence on model uncertainty level are presented. In Sect. 5, different uncertainty estimation methods are assessed based on the results of the selected 36 papers. In Sect. 6, different evaluation criteria used for assessing the magnitude of uncertainty is presented. The summary and recommendations for future research is outlined in Sect. 7.

2 Overview

The articles reviewed in this paper were selected from various peer reviewed journals and are presented in Table 1. It may be noted that the listed journals in Table 1 publish range of hydrology and environmental related

Table 1 List of journals selected for the review

S. no	Name of journal	2014 impact factor	No of papers published
1	*Journal of Hydrology*	3.053	7
2	*Hydrology and Earth System Sciences*	3.535	3
3	*Hydrological Processes*	2.677	4
4	*Water Resources Research*	3.549	3
5	*Environmental Modelling and Software*	4.420	1
6	*Stochastic Environment Research and Risk Assessment*	2.086	3
7	*Journal of Hydroinformatics*	1.388	3
8	*Journal of the American Water Resources Association*	1.348	3
9	*Journal of Hydrometeorology*	3.645	1
10	*Water Resources Management*	2.600	2
11	*Engineering Applications of Artificial Intelligence*	2.503	1
12	*Natural Hazards Earth System Sciences*	1.826	1
13	*Journal of Hydrologic Engineering*	1.624	1
14	*Neural Computing and Applications*	1.569	1
15	*Physics and Chemistry of the earth*	1.477	1
16	*Journal of Flood Risk Management*	1.119	1

research papers that are recognized and cited by researchers across the world. The papers for the review were collected from various data base that include Scopus, Thomson Reuters and Google scholar. Different combinations of keywords were used such as “neural network”, ‘ensemble’. “uncertainty” and “prediction/forecast interval”. Note that there was no mention on the specific year during the search. However, the applications were limited to water resources, environmental science and the document type was article which resulted in 36 papers which focused on ANN models applied for flood forecast or rainfall-runoff modeling.

The selected papers are listed in Table 2 including the year of publication, details of study area and the method used to quantify the uncertainty. Figure 1 shows the distribution of papers published from year 2002 to 2015. It is clear that the papers published over the year is not evenly distributed and majority of the papers have been published in the year 2009 and 2015, that has 9 and 8 articles respectively.

3 Sources of uncertainty

The uncertainty in hydrologic models is generally distinguished into modeling and prediction uncertainty. The difference between these two uncertainty is that the modeling uncertainty comes from imperfect fit of the model to the observed values of past, whereas the prediction uncertainty arise from extrapolation error for future variable that may not follow the modeling uncertainty (Morgan et al. 1990; Krupnick et al. 2006). The modeling uncertainty can be further classified into different forms based on the variability arises from model inputs, parameters and structures which combine together contributes in producing the prediction uncertainty. The input (measured/forecasted precipitation in the case of hydrologic models) uncertainty is mainly due to instrument, measurement and sampling error. The parameter uncertainty lies in inability to identify a unique set of best performing parameters. The simplification, inadequacy and ambiguity in the description of real world process through mathematical equation leads to model structure uncertainty (Shrestha and Solomatine 2008). It is noted that different sources of uncertainty may produce different magnitude of error and hence suitable approach must be developed for better characterizing and quantifying the uncertainty. Any misleading assumption or ignoring any form of uncertainty might result in over and/or under estimation of uncertainty at the model output.

The progress towards the improved performance of model ensuring acceptable level of modeling uncertainty would enhance the predictability of the model. In the context of uncertainty in ANN based hydrologic models, uncertainty quantification of input, parameter and model structure combine together or individually is of primary interest, which in turn will have significant impact on model prediction uncertainty. Different forms of uncertainty investigated by various research articles selected for this review is illustrated in Fig. 2. It is evident from the Fig. 2 that the quantification of parameter uncertainty has

Table 2 Details of paper reviewed

S. no.	Author(s) (year)	Study area	Method used to quantify uncertainty
1.	Cannon and Whitfield (2002)	British Columbia, Canada	Ensemble ANN
2.	Jeong and Kim (2005)	Geum river basin, Korea	Ensemble ANN
3.	Khan and Coulibaly (2006)	Serpent River and CDD sub basin in SLSJ watershed, Canada	Parameter
4.	Han et al. (2007)	Bird creek drainage basin, USA	Hueristic based method
5.	Srivastav et al. (2007)	Kolar basin, India	Bootstrap
6.	Kim and Kim (2008)	Wi-Stream Catchment, Korea	Sensitivity analysis
7.	Asefa (2009)	West-central coast of Florida, USA	GLUE
8.	Yang and Chen (2009)	Wu-Shi basin, Taiwan	Self-organizing map and linear transfer function
9.	Boucher et al. (2009a)	Leaf river, USA; Serein and Le Gola river France	Bootstrap
10.	Boucher et al. (2009b)	Leaf river, USA; Serein and Volpajola river France	Monte Carlo simulation
11.	Cullmann et al. (2009)	Freiberger Mulde catchment, Germany	ANN and process based models (HEC-RAS)
12.	Shrestha and Nestmann (2009)	Rhine and Neckar Rivers, Germany	Fuzzy
13.	Shrestha et al. (2009)	Brue catchment, UK	Monte Carlo simulation
14.	Zhang et al. (2009)	Reynold creek, little creek, USA	Bayesian
15.	Sharma and Tiwari (2009)	Upper damodar valley catchment, India	Bootstrap
16.	Khan and Coulibaly (2010)	Serpent River and CDD sub basin in SLSJ watershed, Canada	Bayesian
17.	Tiwari and Chatterjee (2010a)	Mahanadi River basin, India	Bootstrap
18.	Tiwari and Chatterjee (2010b)	Mahanadi River basin, India	Bootstrap
19.	Araghinejad et al. (2011)	Red River, Canada; Zayandeh-rud River, Iran	Ensemble based approach
20.	Zhang et al. (2011)	Little River Experimental Watershed, USA	Bayesian
21.	Alvisi and Franchini (2011)	Reno and Sieve river, Italy	Fuzzy
22.	Tiwari and Chatterjee (2011)	Mahanadi River basin, India	Bootstrap
23.	Zhang and Zhao (2012)	Little River and Reynolds Creek Watershed, USA	Bayesian
24.	Alvisi and Franchini (2012)	Reno river basin, Italy	Grey NN
25.	Artigue et al. (2012)	Gardond'Anduze Basin, France	Rainfall multiplier
26.	Kasiviswanathan et al. (2013)	Kolar Basin, India	Multi objective optimization
27.	Kasiviswanathan and Sudheer (2013)	Upper white river, USA	First order uncertainty analysis
28.	Guo et al. (2013)	Leaf river watershed, USA	Monte Carlo simulation
29.	Kant et al. (2013)	Mahanadi River basin, India	Multi objective optimization and bootstrap
30.	Fleming et al. (2015)	English man river, Canada	Ensemble ANN
31.	Kan et al. (2015)	Chengcun, Dongwan and Zhidan, China	Ensemble ANN
32.	Kim and Seo (2015)	Nakdong River, South Korea	Ensemble ANN
33.	Kumar et al. (2015)	Damodar catchment, India	Bootstrap
34.	Oliveira et al. (2015)	Ijuí River basin, Brazil	Stochastic generation of climate variable
35.	Yu et al. (2015)	Upper Thames river, UK	Monte Carlo simulation
36.	Taormina and Chau (2015)	Susquehanna and Nehalem rivers, USA	Multi objective optimization

mostly been reported in 20 numbers of papers. Followed by this, eight numbers of papers have been published for the input uncertainty. Very few papers have dealt the combination of different sources of uncertainty while quantifying the prediction uncertainty of ANN. The heuristic approach has been proposed in a few of the papers, which relate the overall prediction uncertainty based on the magnitude of deviation in the input variables and/or model parameters.

More often, the input uncertainty is analyzed through sampling multiple realization of inputs either probabilistic

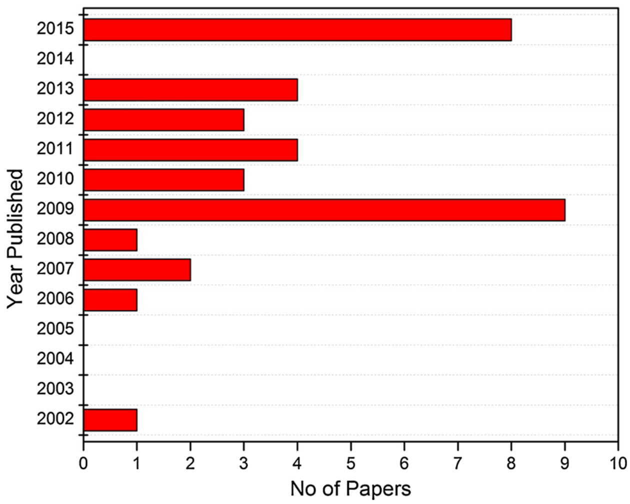

Fig. 1 Distribution of papers by year of publication

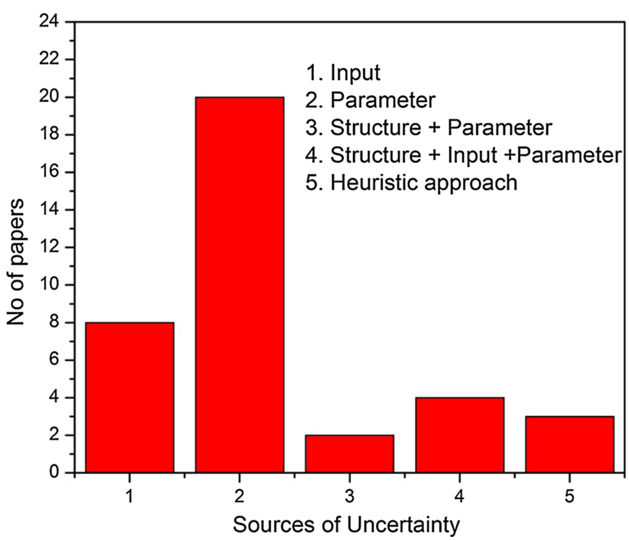

Fig. 2 Number of papers published based on different sources of uncertainty

approach (Oliveira et al. 2015; Yu et al. 2015) or fuzzy approach (Shrestha and Nestmann 2009; Alvisi and Franchini 2011). These realizations would result in different combinations of trained ANN model parameters for obtaining the variability at the model output or prediction uncertainty. The parameter uncertainty can be carried out through different approaches: (a) random initialization of the network weights for training the ANN model (Boucher et al. 2009b), (b) varying the training dataset (Srivastav et al. 2007) and (c) perturbing the weights of the networks (Kasiviswanathan et al. 2013). The structure uncertainty of ANN models was carried out by adding or removing hidden neurons/or hidden layers, which in general requires very high computational effort (Zhang et al. 2011) and that could be the plausible reason, only a few studies have been reported (Fig. 2).

4 Types of ANN and their level of uncertainty

In literature, several variants of neural network models have been reported (Maier et al. 2010), which in general are classified into feed forward and recurrent networks. In feed forward neural network, the information is processed from input to output layer in one direction, whereas in recurrent network the information is processed in both directions. The topology/architecture of the model varies between the types of network in terms of number of hidden layers/nodes, activation function and their connections. It is well established that the selection of a particular architecture is unique for the process to be modeled, which subsequently yields different magnitudes of error/uncertainty while reproducing the observation in terms of model output. Hence it is important to analyze different model architecture for their influence or impact while quantifying the uncertainty. Among different feed forward ANN architecture, the multilayer perceptron (MLP) has been frequently used in ANN based hydrologic models and were reported in 32 papers that are reviewed. Other ANN architecture has been reported only in very few studies. For instance, a fuzzy neural network (Shrestha and Nestmann 2009; Alvisi and Franchini 2011, 2012; Kant et al. 2013) and recurrent neural network (Kim and Kim 2008) has received less attention, hence require further investigation for quantifying the model prediction uncertainty. The input and output nodes are problem dependent, however eliminating less sensitive inputs while training the model reduces the size of the network, which obviously reduce the number of connection weights and the level of model parameter/structure and prediction uncertainty (Kim and Kim 2008).

The number of hidden nodes is responsible for capturing the non-linearity of the process to be modeled. In most of the studies, it has been suggested to initialize with less number of neurons and increase them based on the model accuracy. These nodes are arranged in the form of layer. Most of the researchers suggested using single hidden layer except few (Shrestha and Nestmann 2009; Yu et al. 2015). The complexity of the network, if required, can be enhanced by increasing the number of hidden nodes in a single hidden layer (Neal 1996) without the effect of overfitting, as it would increase the uncertainty. Most of the studies, that are considered in this review article suggested use of trial and error approach to fix the number of hidden layers and nodes. However, Taormina and Chau (2015) have suggested the k-fold cross validation to determine the number of hidden layers based on the complexity of the dataset in an ensemble framework. Though variety of training algorithm has been reported in different studies, the amount of uncertainty related to selection of suitable algorithm still requires further investigation. The activation function responsible for modeling the non-linearity needs to be decided in the hidden and output layers. Most of the studies use sigmoidal and linear function in hidden and output layers respectively, except a few which used non-linear activation function in both the layers (Kasiviswanathan and Sudheer 2013; Oliveira et al. 2015). Overall, it was found that although different model architecture, number of hidden layers and nodes and activation functions are the causes of uncertainty, all these criteria have been utilized to assess the model accuracy for the point prediction and not for the model prediction uncertainty. In addition, the general conclusion reported in the selected articles were that increasing the model complexity might increase the accuracy of the model, however at the expense of uncertainty associated with the model prediction. Hence it is suggested to include these approaches for the quantification of model uncertainty, so that a reliable model with less uncertainty can be identified.

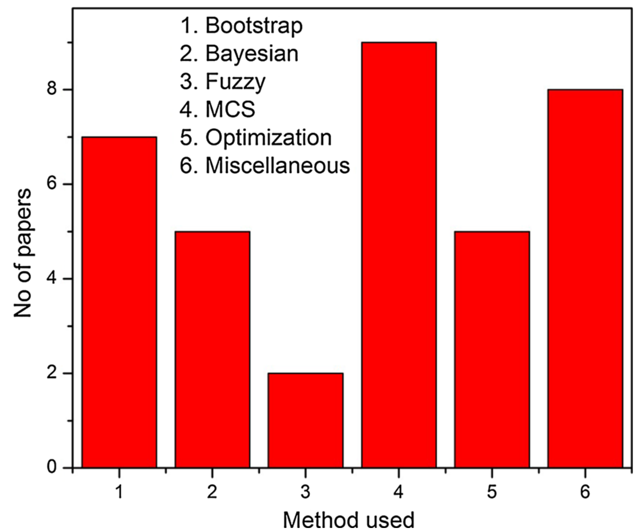

Fig. 3 Number of papers used different methods for estimating uncertainty

5 Methods used for estimating the uncertainty of ANN

The uncertainty methods in general, have been classified into four different approaches (Alvisi and Franchini 2011) such as (a) probabilistic based method (b) analyzing the statistical properties of the errors of the model in reproducing the observed data (c) resampling techniques, generally known as ensemble methods, or the Monte Carlo method and (d) fuzzy based method. In this paper, the same classification has been adopted, however named with specific terms while describing these methods. Figure 3 illustrates different methods that were developed for estimating the uncertainty of neural network hydrologic models. It is evident from Fig. 3 that Monte Carlo Simulation (MCS) and bootstrap based approaches (i.e. 9 and 7 papers respectively) have often been reported due to robust estimation of model prediction uncertainty. All other approaches such as Bayesian and optimization based approaches have been reported in at least five numbers of papers. The fuzzy based approach has been reported only in two papers. The miscellaneous approach such as first order uncertainty analysis, generalized likelihood uncertainty estimation (GLUE), heuristic based approach, sensitivity analysis and self-organizing map are also reviewed. The following subsections describe each of these methods keeping in view of their advantages, shortcomings and general conclusion drawn in different papers.

5.1 Bootstrap method

In the deterministic modeling approach, it is assumed that the model parameters are time invariant. However, such an approach might not always be valid, if the process contains uncertainty or inherent variability. The bootstrap method samples different realization of input–output patterns in order to treat the network parameters (i.e. weights and biases) as a non-deterministic component. Further, the variability created in weights and biases are utilized to quantify the prediction uncertainty. In bootstrap method, it is assumed that the samples follow the statistical characteristics of population data and also they mimic the random component of the process to be modelled. The empirical distribution of model output produced by bootstrap samples determines the confidence interval. It is also reported that training multiple neural network from different random starting points provide a better convergence in parameter space (Boucher et al. 2009a; Kim and Seo 2015). Different research reported the potential use of bootstrap to quantify the uncertainty of ANN (Table 3). Though variety of bootstrap methods is available, the reported studies used a random bootstrap method. In the bootstrap method, determining the number of bootstrap sample is one of the main factors. However, there is no clear guideline mentioned fixing the number of bootstrap sample; consequently different studies used different number of bootstrap (Table 3). It is noted that all these studies report the final model output by estimating the simple average of ensemble of ANN models trained individually with bootstrapped samples. However, the weighted average, probabilistic methods can still be used to improve the model prediction (Araghinejad et al. 2011). Please note that the reported studies considered only the parameter uncertainty of ANN while estimating the prediction uncertainty. The general conclusion drawn from these studies state that compared to deterministic ANN model, the bootstrap based ANN

Table 3 Number of bootstrap sample used while considering the final model

S. no	Author (s) (year)	Number of bootstrap used
1	Srivastav et al. (2007)	300
2	Sharma and Tiwari (2009)	50
3	Tiwari and Chatterjee (2010a)	50
4	Tiwari and Chatterjee (2010b)	50
5	Boucher et al. (2009b)	50
6	Tiwari and Chatterjee (2011)	200
7	Kumar et al. (2015)	100

produced prediction limit around the mean value that improves the reliability of the model.

5.2 Bayesian method

The traditional neural network model optimizes the parameters of ANN model either by maximizing the likelihood function of parameter or minimizing the error function for the best set of parameter. The ANN model trained in this fashion fits the training data well with reasonable accuracy. However the prediction for the new data might not necessarily have less error due to over fitting of the model and this is one of major issue, which discourages the application of ANN. The impact of over fitting can be minimized with the Bayesian approach, in which the neural network parameters are defined as probabilistic distribution function that take care of uncertainty associated. The advantage of Bayesian neural network (BNN) integrates the posterior distribution of network weights to obtain the predictive distribution of model output (Mackay 1992; Neal 1996). The initial studies which used Bayesian method evaluated only a parameter uncertainty in the case of streamflow prediction (Khan and Coulibaly 2006, 2010). The comparison with conventional deterministic ANN approach indicated that the BNN has showed improvement in the model performance. In addition, the prediction interval obtained through BNN model indicated the level of uncertainty, which provides useful information in decision making. Later the combination of structure and parameter uncertainty was evaluated under the framework of Bayesian method (Zhang et al. 2009). Recently, the combined effect of input, parameter and structure uncertainty was analyzed using Bayesian method to show the overall influence of these uncertainties at the model output (Zhang et al. 2011; Zhang and Zhao 2012). The general conclusion from these studies suggested that the interactions between different uncertainty sources make it difficult to identify the contribution of individual uncertainty sources. Though BNN approach produced promising results along with uncertainty quantification, it suffers from huge sampling and computational burden and hence limits the practical application of these methods. Nevertheless, BNNs are the most appropriate tools to achieve reliable ensemble and probabilistic hydrologic forecasting (Khan and Coulibaly 2006).

5.3 Fuzzy method

In the fuzzy method, the uncertainty of ANN model can be analyzed for the model parameters and/or inputs. In such analysis, the model inputs and/or parameters are represented as a fuzzy number for defining the variability and to quantify the prediction uncertainty. The fuzzy method has an advantage over the probability based methods, since it does not require any assumption of underlined distribution. Note that integration of fuzzy with ANN (ANFIS) models were reported (Chang and Chen 2001; Nayak et al. 2005) to demonstrate the newer version of ANN model in terms of improved prediction. However, all these methodology focused on developing automatic if–then fuzzy rule in ANN models, which typically provide the point prediction without any information of model uncertainty. While analyzing the uncertainty in ANN models, Shrestha and Nestmann (2009) used fuzzy method to model the stage-discharge relationship representing the stage as a fuzzy number. In such procedure, the generated series of fuzzy inputs were modeled using ANN and then the uncertainty was estimated. Alvisi and Franchini (2011) quantified the parameter uncertainty of ANN by representing the neural parameters as fuzzy number. In this context, it can be summarized from the reported works that the lower and upper bounds defined by membership function leads to establish a confidence/prediction interval of output which in turn reflects the model prediction uncertainty.

5.4 Monte Carlo simulation method (MCS)

The Monte Carlo simulation method is a probabilistic based approach which samples different realization of model inputs or parameters by assigning the ranges and probability distribution of each variable. The sampling could be random (Tung and Yen 2005) or stratified (for example latin hyper cube (LHS)) (Mckay et al. 1979). The advantage of LHS provides more uniform space—filling coverage of parameter space which resulted in faster convergence. These samples will then be further used in running the hydrologic models to estimate the uncertainty of model outcomes. In the case of ANN models, the randomness may be created in model inputs (i.e. typically in rainfall) or model parameters through samples (i.e. named as *rainfall multipliers*) drawn from the probability distribution functions. The rainfall multipliers are usually sampled from normal random variate, which tend to modify the measured inputs during model calibration so as to evaluate the input uncertainty. For instance, Artigue et al. (2012) sampled 20 different combinations of rainfall multiplier randomly with a range between 0.8 and 1.2. The altered rainfall information was used to train 20 different ANN models and to quantify the uncertainty of the model output.

In addition, some of the studies reported the ANN as a surrogate model or uncertainty processing tool (Shrestha et al. 2009; Cullmann et al. 2009; Yu et al. 2015). In such analysis, the parameter uncertainty of conceptual or physically based model that used MCS was mapped through ANN model. The prediction limits obtained from the ANN model is based on the functional relationship between the hydro meteorological variables and the characteristics of the model output that follow probability distribution. The recent developments in the optimization algorithms (shuffled complex evaluation metropolis—SCEM) include the component of uncertainty while calibrating the model parameters. The algorithm used Marko Chain Monte Carlo (MCMC) sampler that can effectively locate the high probability density region of the feasible neural network parameter space (Guo et al. 2013).

5.5 Optimization based method

Recently, the application of optimization algorithm in estimating the uncertainty of ANN models has gained increasing attention. Despite several sophisticated optimization based search algorithm available, the genetic algorithm (GA) based optimization technique has been frequently used. For instance, the GA was integrated with ANN for identifying the optimal architecture and also to quantify the prediction uncertainty by combining several neural networks (Kant et al. 2013). Araghinejad et al. (2011) used specific function with the consideration of improving the estimate of the peak and low flow through ensemble of ANN models. Recently, the ANN model parameter uncertainty was carried out with the cost function aiming to reduce/optimize the uncertainty interval at the model output using the optimization framework. In contrast to the traditional objective function used for developing ANN models which provide the point prediction through deterministically optimizing the parameter, the newly developed objective function optimizes ensemble of models which provides the indication of uncertainty at the model output (Alvisi and Franchini 2012; Kasiviswanathan et al. 2013; Taormina and Chau 2015). The major difference between the traditional and optimization based method is that it differs in identifying the ANN parameters with acceptable level of uncertainty rather train the model and then quantify the uncertainty. However, these methods consider only parameter uncertainty of ANN and hence methods for including other sources of uncertainty should be developed.

5.6 Miscellaneous method

In addition to the methods described in this paper, several other methods have also been reported in the context of evaluating the uncertainty in ANN based hydrologic models and listed as follows:

- Any variation created in the initial seeds of parameters lead to converge in different combination of parameters. For instance, Boucher et al. (2009b) used an approach which initialized various combinations of random seeds of model parameters for obtaining the variation in final model parameters and to quantify the prediction uncertainty.
- Han and Kwong (2007) proposed a heuristic based method to understand the uncertainty in ANN hydrologic models. The method estimates the distance between the input vector at prediction and all the training data, which provides a valuable indication on how well the prediction would be, relating to the uncertainty in the model input, parameters and structure.
- The sensitivity analysis was reported in some of the research studies (Kim and Kim 2008). It was carried out for the purpose of eliminating the uncertainty that originates from irrelevant input information in ANN models and thereby enhances the model prediction with reduced level of uncertainty.
- Yang and Chen (2009) applied a self-organizing map (SOM) and linear transfer function (LTF) to efficiently determine the intervals of weights and biases. The ranges of parameters obtained through this approach were then used to establish the prediction interval of the model output. The results indicated that compared to the conventional ANN model, the proposed approach produced a narrow interval of model output around the observed values which better characterize the underlined uncertainty.
- The first order uncertainty analysis (FOUA) has been reported to be a promising approach while evaluating meaningful estimate of model prediction uncertainty by deriving the first derivative of model functional form with respect to uncertain variables (i.e. model parameter and/or inputs). However, the application of FOUA remains a challenging task in non-linear models due to the complex derivatives involved. In the case of ANN model, it was explored by Kasiviswanathan and Sudheer (2013) and their results suggested that the

quantified level of uncertainty was found to be varying with magnitude of flow as anticipated.

- Several research studies used ensemble of climate model output while developing ensemble of ANN models for the purpose of flood forecasting (Cannon and Whitfield 2002; Fleming et al. 2015; Oliveira et al. 2015). Consequently each ANN model used different set of predictor variates, so that the overall system provides a collective forecast error which combines different sources of uncertainty.
- Asefa (2009) proposed GLUE approach in ANN models which accounts for uncertainties in model structure and inputs. The method has an advantage of systematically characterizing the uncertainty in the form of non-parametric without explicit consideration of assuming the distribution of residuals. The likelihood measure decides the selection of model output and corresponding model parameters sampled from probability distribution functions. Further, it suggested developing methods for updating the likelihood to improve the model prediction without training the models. However, the complex computation and assumption about parameter in the form of probability distribution functions limit the practical application of these methods.

6 Evaluation criteria used for quantifying the uncertainty of ANN

The accuracy of uncertainty evaluation not only depends on the methods used, however it is also important to convey the information meaningfully in appropriate way. In general, the magnitude of uncertainty can be evaluated using graphical representation and/or statistical measures. The graphical representation such as rank histogram and reliability diagram (Boucher et al. 2009a) provides information of model uncertainty. The visual inspection of model output variability with a set of upside and downside prediction (i.e. optimistic and pessimistic values) around the mean simulation provides overall information of quality of the model prediction. Despite, the graphical representation which provides useful insight about the model uncertainty assessment, the inter comparison of model response under different conditions is a difficult task. Hence, the statistical measures, which are objective and quantitative in nature, are often employed for ranking different alternatives. Figure 4 depicts the number papers which used different uncertainty measure indices while evaluating the model output uncertainty.

Several uncertainty evaluation measures have been reported in literature that include (a) average relative

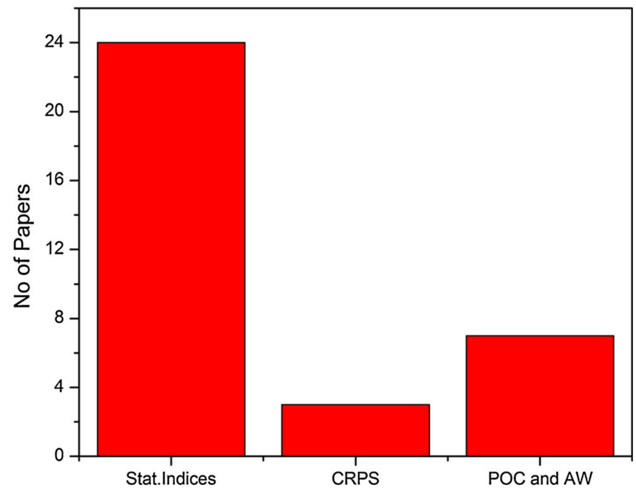

Fig. 4 Number of papers used different uncertainty indices

length (ARL) (Jin et al. 2010), (b) average asymmetry degree (AAD), (c) average deviation amplitude (ADA) (Xiong et al. 2009), (d) percentage of coverage (POC) and (e) average width (AW) (Kasiviswanthan and Sudheer 2013). The first three indices are mainly used to evaluate the asymmetry degree of prediction interval and deviation from observed values. The remaining two indices measure the magnitude of prediction interval width and percentage of observed values that fall within prediction interval. Although, all these indices measure the quality of the model prediction uncertainty in terms of quantitative estimate (Alvisi and Franchini 2011), the POC and AW were reported in seven papers (Fig. 4) as these two indices reflect the accuracy of the model prediction. In addition, three papers have employed the continuous rank probability score (CRPS) (Matheson and Winkler 1976; Hersbach 2000) to estimate the quality of the prediction interval. It measures the distance between the observed and predicted distribution. Another index called logarithmic score (Roulston and Smith 2002) is the logarithm of probability density, which corresponds to observed and predicted values. The major drawback with this index is that it does not consider the lower bound and the probability density becomes zero when the observation falls outside, which results an infinite value. This could be the reason that most of the study did not consider this index while estimating the uncertainty, except few (Boucher et al. 2009b).

It is noted from the reported studies that 24 numbers of papers did not use any indices to quantify uncertainty, but rather the uncertainty was described through statistical indices (i.e. standard deviation, minimum, maximum and inter quartile range of model output). This suggests further improvement/research while reporting the magnitude of uncertainty with proper indices.

7 Summary and recommendations for future research

The absence of uncertainty quantification in the hydrologic models might produce misleading information, which limits their practical application especially in planning and management of water resources. Thus, the meaningful quantification of uncertainty is necessary for understanding different mechanism, so as to find the best possible way of reducing it. In the case of ANN based flood forecast models, the uncertainty quantification using standard procedures as applied in other hydrologic models are computationally challenging owing to the parallel computing architecture, large number of degrees of freedom in their model development, subsequently suitable methods for carrying out uncertainty analysis is still lacking. This motivates researchers to develop appropriate procedures with suitable modifications, assumptions, in order to quantify the uncertainty of ANN. If sufficient information of uncertainty is available, the credibility/reliability of forecasting models would reduce the risk in the decision making (i.e. the design or operation of hydraulic structures). It is noted that many review articles reported the advantages of using ANN and addressed the methodological issues, however not in terms of uncertainty. Therefore, this review paper would be useful in understanding the current state of research, research gaps where further progress is required. Based on the review of 36 research articles, the following recommendations for future research are made:

1. The clear guidelines should be developed to determine the number of bootstraps, since different study reported different number of ensembles. While estimating the model output from the ensemble of neural network formed by individual ANN models, more sophisticated methods should be developed to improve the estimate of prediction rather focusing on simple arithmetic averaging. The potential of other bootstrap methods such as moving block bootstrap, circular block bootstrap, and stationary bootstrap can also be explored. The study reported bootstrap method considered only a parameter uncertainty; however approaches should be developed to include other sources of uncertainty.
2. The application of Bayesian method could be further explored to better characterize the influence of each sources of uncertainty.
3. Since very limited studies focused on employing fuzzy method, further developments in this area can be explored integrating fuzzy with ANN models for the uncertainty quantification. It is also noted that studies based on fuzzy assume a constant membership function while generating the lower and upper values of input or model parameters and hence future study can focus variable membership level and also methods for including the structure uncertainty.
4. The application of MCS based method can be further extended to quantify the uncertainty of ANN, since it is very robust while analyzing the uncertainty.
5. The FOUA can still be used to demonstrate the influence of each parameter while estimating the total prediction uncertainty. It is also suggested that proper methods should be devised for the input identification of ANN model through FOUA.
6. The current studies limit quantifying the parameter uncertainty of ANN using optimization based methods; hence further methods can be extended to combine different sources of uncertainty.
7. The uncertainty is often quantified using general statistical indices such as mean, standard deviation, minimum and maximum values of model output, thus the uncertainty measuring indices listed in (Xiong et al. 2009) could be employed and the new indices could be developed.
8. The selection of network architecture/topology in terms number of hidden layers/nodes, activation function have been reported to finalize the initial model based on the accurate estimate of point prediction/forecast of the model. After fixing the architecture, the quantification of uncertainty was carried out in majority of the studies. Hence further investigation is required focusing the level of uncertainty in terms of reducing the magnitude while developing the models. In addition, most of the studies in uncertainty quantification of ANN reported the MLP architecture; however other types such as radial basis functions, support vector machines can be explored.

Acknowledgements The authors would like to thank two anonymous reviewers for their insightful comments, which helped improving the quality of the paper.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This paper does not contain any studies with human participants or animals performed by any of the authors.

Informed consent For this type of study, formal consent is not required.

References

Abrahart RJ, See LM, Dawson CW, Shamseldin AY, Wilby RL (2010) Nearly two decades of neural network hydrologic

modeling. In: Sivakumar B, Berndtsson R (eds) Advances in data-based approaches for hydrologic modeling and forecasting. World Scientific Publishing, Hackensack, NJ, pp 267–346
Alvisi S, Franchini M (2011) Fuzzy neural networks for water level and discharge forecasting with uncertainty. Environ Model Softw 26:523–537. doi:10.1016/j.envsoft.2010.10.016
Alvisi S, Franchini M (2012) Grey neural networks for river stage forecasting with uncertainty. Phys Chem Earth Parts A/B/C 42–44:108–118. doi:10.1016/j.pce.2011.04.002
Araghinejad S, Azmi M, Kholghi M (2011) Application of artificial neural network ensembles in probabilistic hydrological forecasting. J Hydrol 407:94–104. doi:10.1016/j.jhydrol.2011.07.011
Artigue G, Johannet A, Borrell V, Pistre S (2012) Flash flood forecasting in poorly gauged basins using neural networks: case study of the Gardon de Mialet basin (southern France). Nat Hazards Earth Syst Sci 12:3307–3324. doi:10.5194/nhess-12-3307-2012
ASCE Task Committee on Application of Artificial Neural Networks in Hydrology (2000a) Artificial neural networks in hydrology. I: preliminary concepts. J Hydrol Eng 5(2):115–123
ASCE Task Committee on Application of Artificial Neural Networks in Hydrology (2000b) Artificial neural networks in hydrology. II: hydrologic applications. J Hydrol Eng 5(2):124–137
Asefa T (2009) Ensemble streamflow forecast: a glue-based neural network approach. J Am Water Resour Assoc 45:1155–1163. doi:10.1111/j.1752-1688.2009.00351.x
Bishop CM (2004) Neural networks for pattern recognition. Oxford University Press, Oxford
Boucher MA, Laliberté J-P, Anctil F (2009a) An experiment on the evolution of an ensemble of neural networks for streamflow forecasting. Hydrol Earth Syst Sci Discuss 6:6265–6291
Boucher M-A, Perreault L, Anctil F (2009b) Tools for the assessment of hydrological ensemble forecasts obtained by neural networks. J Hydroinform 11:297. doi:10.2166/hydro.2009.037
Cannon AJ, Whitfield PH (2002) Downscaling recent streamflow conditions in British Columbia, Canada using ensemble neural network models. J Hydrol 259:136–151. doi:10.1016/S0022-1694(01)00581-9
Chang FJ, Chen YC (2001) A counterpropagation fuzzy-neural network modelling approach to real time streamflow prediction. J Hydrol 245:153–164
Citakoglu H (2015) Comparison of artificial intelligence techniques via empirical equations for prediction of solar radiation. Comput Electron Agric 118:28–37
Cobaner M (2011) Evapotranspiration estimation by two different neuro-fuzzy inference systems. J Hydrol 398(3–4):292–302
Cullmann J, Krausse T, Philipp A (2009) Communicating flood forecast uncertainty under operational circumstances. J Flood Risk Manag 2:306–314. doi:10.1111/j.1753-318X.2009.01048.x
Elshorbagy A, Corzo G, Srinivasulu S, Solomatine DP (2010a) Experimental investigation of the predictive capabilities of data driven modeling techniques in hydrology—part 1: concepts and methodology. Hydrol Earth Syst Sci 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. Hydrol Earth Syst Sci 14:1943–1961
Fleming SW, Bourdin DR, Campbell D et al (2015) Development and operational testing of a super-ensemble artificial intelligence flood-forecast model for a Pacific northwest river. JAWRA J Am Water Resour Assoc 51:502–512. doi:10.1111/jawr.12259
Flood I, Kartam N (1994) Neural networks in civil engineering. I: principles and understanding. J Comput Civil Eng 8(2):131–148
Guo J, Zhou J, Song L et al (2013) Uncertainty assessment and optimization of hydrological model with the shuffled complex evolution metropolis algorithm: an application to artificial neural network rainfall-runoff model. Stoch Environ Res Risk Assess 27:985–1004. doi:10.1007/s00477-012-0639-0
Han DT, Kwong Li S (2007) Uncertainties in real-time flood forecasting with neural networks. Hydrol Process 21(2):223–228. doi:10.1002/hyp.6184
Hersbach H (2000) Decomposition of the continuous ranked probability score for ensemble prediction systems. Weather Forecast 15:559–570
Jayakrishnan R, Srinivasan R, Santhi C, Arnold JG (2005) Advances in the application of the SWAT model for water resources management. Hydrol Process 19:749–762
Jeong D, Kim Y (2005) Rainfall-runoff models using artificial neural networks for ensemble streamflow prediction. Hydrol Process 19:3819–3835. doi:10.1002/hyp.5983
Jin X, Xu CY, Zhang Q et al (2010) Parameter and modeling uncertainty simulated by GLUE and a formal Bayesian method for a conceptual hydrological model. J Hydrol 383(3–4):147–155
Kan G, Yao C, Li Q et al (2015) Improving event-based rainfall-runoff simulation using an ensemble artificial neural network based hybrid data-driven model. Stoch Environ Res Risk Assess 29:1345–1370. doi:10.1007/s00477-015-1040-6
Kant A, Suman PK, Giri BK et al (2013) Comparison of multi-objective evolutionary neural network, adaptive neuro-fuzzy inference system and bootstrap-based neural network for flood forecasting. Neural Comput Appl 23:231–246. doi:10.1007/s00521-013-1344-8
Kasiviswanathan KS, Sudheer KP (2013) Quantification of the predictive uncertainty of artificial neural network based river flow forecast models. Stoch Environ Res Risk Assess 27:137–146. doi:10.1007/s00477-012-0600-2
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. doi:10.1016/j.jhydrol.2013.06.043
Khan MS, Coulibaly P (2006) Bayesian neural network for rainfall-runoff modeling. Water Resour Res 42:1–18. doi:10.1029/2005WR003971
Khan MS, Coulibaly P (2010) Assessing hydrologic impact of climate change with uncertainty estimates: Bayesian neural network approach. J Hydrometeorol 11:482–495. doi:10.1175/2009JHM1160.1
Kim S, Kim HS (2008) Uncertainty reduction of the flood stage forecasting using neural networks model 1. J Am Water Resour Assoc 44:148–165
Kim SE, Seo IW (2015) Artificial neural network ensemble modeling with conjunctive data clustering for water quality prediction in rivers. J Hydro-Environ Res. doi:10.1016/j.jher.2014.09.006
Kitanidis PK, Bras RL (1980) Real-time forecasting with a conceptual hydrologic model 2. Applications and results. Water Resour Res 16(6):1034–1044
Krupnick AR, Morgenstern M, Batz P et al (2006) Not a sure thing: making regulatory choices under uncertainty. Resources for the Future, Washington, DC
Kumar S, Tiwari MK, Chatterjee C, Mishra A (2015) Reservoir inflow forecasting using ensemble models based on neural networks, wavelet analysis and bootstrap method. Water Resour Manag 29:4863–4883. doi:10.1007/s11269-015-1095-7
MacKay DJC (1992) Bayesian interpolation. Neural Comput 4:720–736
Maier HR, Dandy GC (2000) Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. Environ Model Softw 15(1):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. Environ Model Softw 25:891–909. doi:10.1016/j.envsoft.2010.02.003

Matheson JE, Winkler RL (1976) Scoring rules for continuous probability distributions. Manag Sci 22:1087–1095

McKay M, Beckman R, Conover W (1979) A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics 21(2):239–245

Morgan M, Henrion GM, Small M (1990) Uncertainty: a guide to dealing with uncertainty in quantitative risk and policy analysis. Cambridge University Press, New York

Nayak PC, Sudheer KP, Rangan DM, Ramasastri KS (2005) Short-term flood forecasting with a neurofuzzy model. Water Resour Res 41:1–16. doi:10.1029/2004WR003562

Neal RM (1996) Bayesian learning for neural networks. Springer, New York

Oliveira GG, Pedrollo OC, Castro NMR (2015) Stochastic approach to analyzing the uncertainties and possible changes in the availability of water in the future based on scenarios of climate change. Hydrol Earth Syst Sci 19:3585–3604. doi:10.5194/hess-19-3585-2015

Roulston MS, Smith LA (2002) Evaluating probabilistic fore- casts using information theory. Mon Weather Rev 130:1653–1660

Schreiber HA, Kincaid DR (1967) Regression models for predicting on-site runoff from short-duration convective storms. Water Resour Res 3(2):389–395

Sharma SK, Tiwari KN (2009) Bootstrap based artificial neural network (BANN) analysis for hierarchical prediction of monthly runoff in Upper Damodar Valley Catchment. J Hydrol 374:209–222. doi:10.1016/j.jhydrol.2009.06.003

Shrestha RR, Nestmann F (2009) Physically based and data-driven models and propagation of input uncertainties in river flood prediction. J Hydrol Eng 14:1309–1319. doi:10.1061/(ASCE)HE.1943-5584.0000123

Shrestha DL, Solomatine DP (2008) Data-driven approaches for estimating uncertainty in rainfall-runoff modelling. Int J River Basin Manag 6:109–122. doi:10.1080/15715124.2008.9635341

Shrestha DL, Kayastha N, Solomatine DP (2009) A novel approach to parameter uncertainty analysis of hydrological models using neural networks. Hydrol Earth Syst Sci 13:1235–1248. doi:10.5194/hess-13-1235-2009

Srivastav RK, Sudheer KP, Chaubey I (2007) A simplified approach to quantifying predictive and parametric uncertainty in artificial neural network hydrologic models. Water Resour Res 43. doi:10.1029/2006WR005352

Taormina R, Chau K-W (2015) ANN-based interval forecasting of streamflow discharges using the LUBE method and MOFIPS. Eng Appl Artif Intell 45:429–440. doi:10.1016/j.engappai.2015.07.019

Tiwari MK, Chatterjee C (2010a) Development of an accurate and reliable hourly flood forecasting model using wavelet-bootstrap-ANN (WBANN) hybrid approach. J Hydrol 394:458–470. doi:10.1016/j.jhydrol.2010.10.001

Tiwari MK, Chatterjee C (2010b) Uncertainty assessment and ensemble flood forecasting using bootstrap based artificial neural networks (BANNs). J Hydrol 382:20–33. doi:10.1016/j.jhydrol.2009.12.013

Tiwari MK, Chatterjee C (2011) A new wavelet–bootstrap–ANN hybrid model for daily discharge forecasting. J Hydroinf 13:500. doi:10.2166/hydro.2010.142

Tung YK, Yen BC (2005) Hydrosystems engineering uncertainty analysis. McGraw Hill, New York

Xiong L, Wan M, Wei X, O'Connor KM (2009) Indices for assessing the prediction bounds of hydrological models and application by generalised likelihood uncertainty estimation/Indices pour évaluer les bornes de prévision de modèleshydrologiques et miseenœuvre pour une estimation d'incertitude. Hydrol Sci J 54:852–871. doi:10.1623/hysj.54.5.852

Yang CC, Chen CS (2009) Application of integrated back-propagation network and self organizing map for flood forecasting. Hydrol Process 23:1313–1323. doi:10.1002/hyp.7248

Yu JJ, Qin XS, Larsen O (2015) Uncertainty analysis of flood inundation modelling using GLUE with surrogate models in stochastic sampling. Hydrol Process 29:1267–1279. doi:10.1002/hyp.10249

Zhang X, Zhao K (2012) Bayesian neural networks for uncertainty analysis of hydrologic modeling: a comparison of two schemes. Water Resour Manag 26:2365–2382. doi:10.1007/s11269-012-0021-5

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. doi: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. doi:10.1016/j.jhydrol.2011.09.002