

Review

Prediction Interval Estimation Methods for Artificial Neural Network (ANN)-Based Modeling of the Hydro-Climatic Processes, a Review

Vahid Nourani ^{1,2,*}, Nardin Jabbarian Paknezhad ¹  and Hitoshi Tanaka ³ 

¹ Center of Excellence in Hydroinformatics and Faculty of Civil Engineering, University of Tabriz, Tabriz 51368, Iran; n.jabbarian@tabrizu.ac.ir

² Faculty of Civil and Environmental Engineering, Near East University, N. Cyprus, via Mersin 10, Nicosia 99138, Turkey

³ Department of Civil Engineering, Tohoku University, 6-6-06 Aoba, Sendai 980-8579, Japan; hitoshi.tanaka.b7@tohoku.ac.jp

* Correspondence: vnourani@yahoo.com or nourani@tabrizu.ac.ir; Tel.: +98-914-403-0332

Abstract: Despite the wide applications of artificial neural networks (ANNs) in modeling hydro-climatic processes, quantification of the ANNs' performance is a significant matter. Sustainable management of water resources requires information about the amount of uncertainty involved in the modeling results, which is a guide for proper decision making. Therefore, in recent years, uncertainty analysis of ANN modeling has attracted noticeable attention. Prediction intervals (PIs) are one of the prevalent tools for uncertainty quantification. This review paper has focused on the different techniques of PI development in the field of hydrology and climatology modeling. The implementation of each method was discussed, and their pros and cons were investigated. In addition, some suggestions are provided for future studies. This review paper was prepared via PRISMA (preferred reporting items for systematic reviews and meta-analyses) methodology.

Keywords: artificial neural network; uncertainty; sustainability; prediction intervals



Citation: Nourani, V.; Paknezhad, N.J.; Tanaka, H. Prediction Interval Estimation Methods for Artificial Neural Network (ANN)-Based Modeling of the Hydro-Climatic Processes, a Review. *Sustainability* **2021**, *13*, 1633. <https://doi.org/10.3390/su13041633>

Academic Editor: Saeed Chehreh Chelgani

Received: 9 January 2021

Accepted: 1 February 2021

Published: 3 February 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Sustainable water resources management includes designing and managing various aspects such as ecology, environment and hydrology integrity in the present and future [1]. Sustainable management requires adequate information about the state of water resources. Thus, appropriate modeling to investigate the situation in present and future times is necessary. On the other hand, reliability of the modeling is an important issue that directly influences the management and decision-making of the problems. Therefore, the uncertainty involved in modeling should be carefully considered so as to achieve more realistic decisions.

Recently, the artificial neural network (ANN) as a prevalent modeling method has been used for identification of the complicated non-linear relationship of inputs and output (e.g., see, [2–8]). The relationship between the hydrological phenomena is a complicated issue and hot topic in hydrological studies, due to spatial and temporal changes of factors that influence the process. Therefore, many hydrological models with various degrees of complexity have been used for the simulation of such a stochastic process [9]. In spite of the numerous applications of the ANN, it has been indicated in many previous studies that ANN models are inherently stochastic, as identical results would be difficult to be reproduced on different occasions [10]. Classic applications of ANN include some imperfections; e.g., the weights of the ANN are randomly assigned, which leads to a long training time; ANN behavior is unexpected and there is not a specified way for determination of the best structure. These features are a deficiency of ANNs, which can have a negative effect on the reliability of the modeling.

Generally, point prediction of the ANN has been considered in most of the studies, but the reliability of the point prediction decreases when the level of uncertainty is high. In the point prediction method, only a point is directly predicted, which is the unknown true targeted value, so its application is questionable. The point prediction is not able to give information about the uncertainty of the modeling, and it is not able to describe the prediction accuracy [11]. Furthermore, via the point prediction, only a prediction error can be obtained, and the probability for correct predictions remains unknown, which can make decision making more difficult. Most of the models associated with the water management and modeling are in the form of point prediction, so various sources of uncertainties have not been investigated, which may affect the modeling outcome. Uncertainty includes model uncertainty, input uncertainty and parameter uncertainty. Prediction uncertainty sources can be errors in measuring, lack of knowledge of constants, sparse and noisy input data, and model approximation errors (e.g., due to imperfections in the model formulation) [12,13], or the target values are affected by some probabilistic events [14]. The importance of determining the total model uncertainty is the same as model output and it has efficient impact on decision making. Without investigating the resources of the uncertainty, the assessment of the ANN modeling quality is impossible.

Prediction interval (PI) is a powerful measure of uncertainties associated with prediction to inform decision makers [15]. PIs lead to the proper arrangement of the future trends and plans, appropriate risk management and an increment of the benefits of modeling. PI presents a bound that captures the observed values by measuring the indication of accuracy called the confidence level $((1-\alpha)\%)$ [16–18]. As PIs contain more sources of uncertainties compared to similar tools such as the confidence interval (CI), they are superior and more practical in order to help decision makers distinguish the best and the worst of the modeling scenarios. Wider PIs present more uncertainty, so more awareness is needed in decisions, and narrower PIs show more confidence in decisions. There are some methods to quantify uncertainty, such as ensemble of the ANN, sensitivity analysis and the self-organizing map. Moreover, some techniques are applied to compute the PIs, such as delta, Bayesian, Monte Carlo, bootstrap and lower upper bound estimation (LUBE). The delta method basic concept is about analyzing the ANNs regression models and application of Taylor's series [19,20]. The procedure of this technique is based on the homogeneity of the noise and its normal distribution. As most of the natural phenomena are heterogeneous, this method's reliability may be questionable. The Bayesian method is based on the consideration of a pre-defined probability distribution of the ANN's parameters, instead of a single value, so the output will also have distributions conditional on the observed training set [21]. The Bayesian technique includes great computations; moreover, to construct the PIs, calculating the Hessian matrix is needed. The Monte Carlo method constructs the PIs based on allocation of the ranges and probability distribution of each variable. These classic methods contain some assumptions about the data distribution [22]. The bootstrap method is a simple and frequently used technique to calculate the PIs [23,24]. This method is based on resampling and training different ANNs. It does not need any assumption about the data distribution but consists of high computational costs for large datasets. Implementation of this method is easy, and it is independent of massive calculation. The main disadvantage of this method is its computational cost for large datasets. The LUBE technique [25] is a non-parametric method, independent of information about the data or error distributions. Kasiviswanathan and Sudheer [22] have reviewed some techniques to quantify the uncertainty of the ANN models in hydrology. Thirty-six research articles associated with uncertainty analysis of the ANN-based stream flow and flood prediction from the years 2002 to 2015 were reviewed, and it is concluded that in order to distinguish the best procedure to encompass different sources of uncertainty, more investigations are needed. The applied methodologies in papers from well-known international journals about uncertainty analysis in hydrology from the year 2002 to quantify uncertainty are tabulated in Table 1.

Table 1. Papers from well-known international journals about uncertainty analysis in hydrology.

Method	Author	Name of the Journal
Ensemble ANN	Cannon and Whitfield [26]	<i>Journal of Hydrology</i>
	Jeong and Kim [27]	<i>Hydrological Processes</i>
	Fleming, Bourdin, Campbell, Stull and Gardner [28]	<i>Water Resources Research</i>
	Kan, Yao, Li, Li, Yu, Liu, Ding, He and Liang [29]	<i>Stochastic Environment Research and Risk Assessment</i>
	Kim and Seo [30]	<i>Journal of Hydro-environment Research</i>
Sensitivity analysis	Kim and Kim [31]	<i>The Journal of the American Water Resources Association</i>
Self-organizing map	Yang and Chen [32]	<i>Hydrological Processes</i>
Bootstrap	Srivastav, Sudheer and Chaubey [33]	<i>Water Resources Research</i>
	Boucher, Perreault and Anctil [34]	<i>Hydrology and Earth System Sciences</i>
	Sharma and Tiwari [35]	<i>Journal of Hydrology</i>
	Kant, Suman, Giri, Tiwari, Chatterjee, Nayak and Kumar [36]	<i>Neural Computing and Applications</i>
Bayesian	Zhang, Liang, Srinivasan and Van Liew [37]	<i>Water Resources Research</i>
	Khan and Coulibaly [38]	<i>Journal of Hydrometeorology</i>
	Zhang, Liang, Yu and Zong [39]	<i>Journal of Hydrology</i>
	Zhang and Zhao [40]	<i>Journal of Hydrology</i>
	Humphrey, Gibbs, Dandy and Maier [41]	<i>Journal of Hydrology</i>
Markov Chain Monte Carlo	Shen, Zeng, Liang, Li, Tan, Li and Li [42]	<i>Water Resources Research</i>
GLUBE	Tongal and Booi [43]	<i>Stochastic Environmental Research and Risk Assessment</i>

The application of PIs to quantify uncertainty has increased and attracted significant attention in the last decade. Therefore, it is worthy to evaluate different PIs development methods in order to identify the best performance of the methods and assess each method suitability.

The lack of review papers investigating PIs development methods in the field of hydrology caused the preparation of the current review paper. The number of published papers regarding PIs construction methods is depicted in Figure 1. The major objective of this review paper is to categorize and enumerate the PIs construction methods and their applications in hydro-climatic studies. Moreover, some suggestions for future works in order to develop and improve the PIs applications are presented. The reviewed sources are mostly included by the Scopus abstract and citation database (www.scopus.com (accessed on 26 January 2021)). Elsevier's Scopus is the most frequently used research engine, and it is updated earlier than the Web of Science on which the papers may be updated lately. In addition, as authors can load any paper onto Google Scholar, some information may not be reliable. The search terms were ("ANN"; "uncertainty") and ("ANN"; "Prediction interval"), respectively, for the uncertainty analysis and PIs. The search operator was "and". Then, the appropriate papers in the fields of hydrology and climatology were selected by abstract reviewing. Moreover, only journal articles published in English were considered, as most of the papers in Scopus research engine are in English. The initial number of obtained papers about the uncertainty assessment of hydrological and hydro-climatological studies was 36 papers; 18 papers from well-known international journals from the year 2002 are tabulated in Table 1. In addition, 69 papers associated with PIs construction of the ANNs were investigated, then 17 papers were selected according to an abstract review and relation to hydrology and climatology. Papers from the years 2002 to 2020 were presented as selected papers.

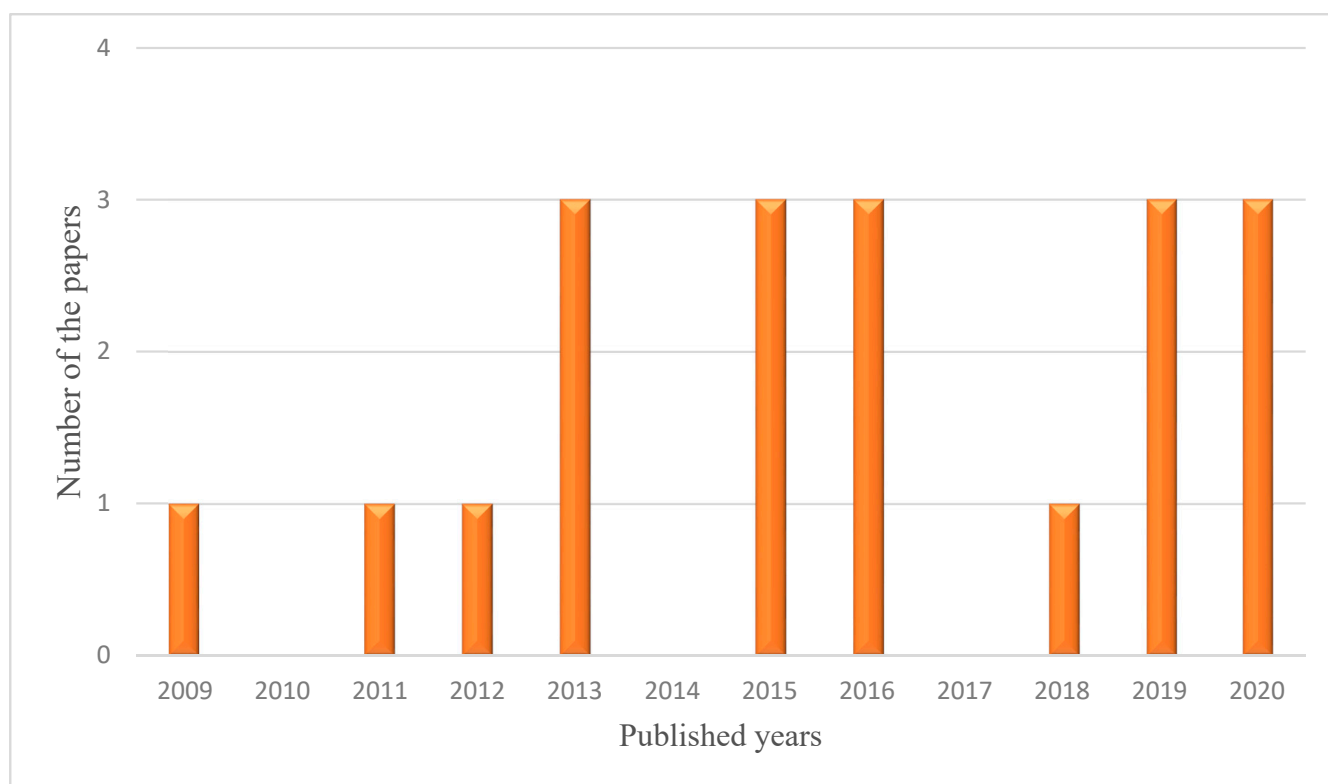


Figure 1. Number of published papers regarding prediction intervals (PIs) construction (indexed in Scopus) with respect to year of publication.

Some methodologies concerning standards of literature reviews and the way of reporting and structuring of them are RAMESES (realist and meta-narrative evidence syntheses: evolving standards), PRISMA (preferred reporting items for systematic reviews and meta-analyses) and PSALSAR (research protocol, appraisal, synthesis and analysis, reporting results). RAMESES could be an appropriate choice for systematic narrative reviews. PRISMA was developed for systematic literature reviews and meta-analyses [44]. PRISMA consists of a 27-item evaluation checklist and a specific flowchart to follow [45]. Moreover, PRISMA protocols (PRISMA-P) checklist and the Explanation and Elaboration [46] document could lead to the improvement of a more complete and reliable review report [47]. The PSALSAR method consists of six basic steps [48], while the common systematic literature review methods include four steps, which are search, appraisal, synthesis and analysis (SALSA). The PSALSAR method contains research protocol and reporting results at the first and last steps. This review paper attempts to follow most of the checklist's items of the PRISMA method, since this method is used as a basis for reporting systematic reviews for most research types. Some examples of systematic literature reviews are [49,50]. The systematic research was conducted on September 1, 2020, and it was updated on December 20, 2020 and also on January, 20, 2021 for preparing a revision. The list of reviewed papers is tabulated in Table 2.

In the following, Section 2 presents the base concepts of PIs and measurement criteria of PIs, Section 3 describes the different PIs construction methods, Section 4 compares different pros and cons of the methods and finally Section 5 recommends some suggestions for future studies.

Table 2. Number of the papers that applied PI to quantify the uncertainty of the artificial neural networks (ANN).

Keywords	Number of Cites	Author	PI Construction Method	Number of Papers
ANN; Bayesian; Bootstrap; PI; Uncertainty	27	Kasiviswanathan and Sudheer [51]	Bayesian	1
Uncertainty analysis; PIs; Lake level; ANN; ANFIS; Bootstrapping	129	Talebizadeh and Moridnejad [52]		
ANN; Bootstrap technique; Hydrological processes; Non-linear function; Taylor series	72	Kasiviswanathan and Sudheer [53]		
Uncertainty; Flood forecasting; Bootstrap; ANNs; Ensemble	60	Kumar, Tiwari, Chatterjee and Mishra [54]		
Water quality forecasting; Wavelet neural network; Bootstrap; Uncertainty; Data missing; Data filling; Songhua River	33	Wang, Zheng, Zhao, Jiang, Wang, Guo and Wang [55]	Bootstrap	7
ANN; Bayesian; Bootstrap; PI; Uncertainty	27	Kasiviswanathan and Sudheer [51]		
ANNs; ensemble simulation; input uncertainty; prediction uncertainty; rainfall–runoff modeling	82	Kasiviswanathan, He, Sudheer and Tay [56]		
General circulation models; Downscaling; PIs; ANN	9	Nourani, Paknezhad, Sharghi and Khosravi [57]		
	111	Shrestha, Kayastha and Solomatine [58]		
groundwater; artificial intelligence; hydrologic model; groundwater level prediction; machine learning; artificial neural network	7	Seifi, Ehteram, Singh and Mosavi [59]	Monte Carlo	3
ANNs; Bayesian uncertainty; fuzzy logic; kriging; uncertainty analysis	5	Tapoglou, Varouchakis, Trichakis and Karatzas [60]		
Ensemble Optimization; PI; Rainfall runoff models	68	Kasiviswanathan, Cebin, Sudheer and Chaubey [61]		
MOFIPS; PSO; Prediction interval; LUBE; Neural networks; Streamflow prediction	63	Taormina and Chau [62]		
PI; Symmetry; ANN; Uncertainty; Flood forecasting; Shuffled complex evolution	16	Zhang, Zhou, Ye, Zeng and Chen [63]		
ANN; Bayesian; Bootstrap; PI; Uncertainty	27	Kasiviswanathan and Sudheer [51]		
ANN; ensemble simulation; input uncertainty; prediction uncertainty; rainfall–runoff modeling	8	Kasiviswanathan, Sudheer and He [64]	LUBE	9
General circulation models; Downscaling; PI; ANN	9	Nourani, Paknezhad, Sharghi and Khosravi [57]		
Evaporation; Neural network; Prediction interval; Uncertainty quantifying; Wavelet de-noising; Jittered data	1	Nourani, Sayyah-Fard, Alami and Sharghi [65]		
Uncertainty analysis; Hybrid double feedforward neural network; Sediment load estimation; Lower upper bound estimation	27	Chen and Chau [66]		
ANN; Crop simulation; Reservoir operation; Optimization; Uncertainty	0	Kasiviswanathan, Sudheer, Soundharajan and Adeboye [67]		

2. PIs Concepts

An interval includes the upper and lower limits that capture indeterminate future value with a prescribed probability [68]. This limit and interval respectively are known as prediction limit and PI (see Figure 2).

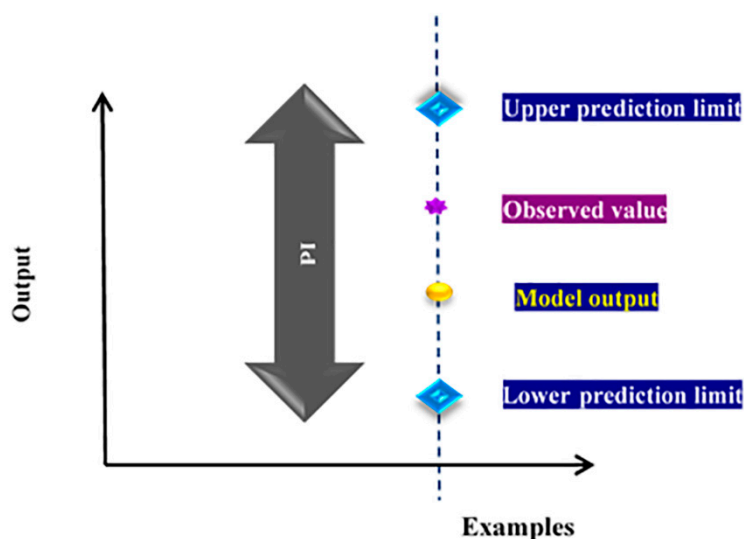


Figure 2. PI description [68].

PI and CI are different measures, and it is important to distinguish them. The CI corresponds to the accuracy of the estimation of the true regression, while PI corresponds to the accuracy of the estimation concerning the observed target value. Actually, PI is more applicable than CI, since it is associated with the accuracy of observed target prediction, whereas CI presents the accuracy of true regression estimation. In order to investigate the differences between the two measures, the following description should be considered. For the estimation of the unknown function $f(x_i; \theta)$ presenting the true underlying model, where θ is actual parameter set, for N data samples $\{(x_i, t_i)\}_{i=1}^N$ we have:

$$t_i = f(x_i; \theta) + e_i \quad (1)$$

where x_i and t_i are, respectively, input, observed target and model error. So, the objective is to estimate the true model $f(x_i; \theta)$. The approximate model $f(x_i; \hat{\theta})$ is the mean of the distribution of the targets, where the estimated parameters $\hat{\theta}$ are determined using machine learning methods. To quantify the uncertainty, two aspects should be considered. One of the aspects is that CI that indicates the accuracy of the estimation of the true model. CI is measured by the distribution of the quantity $f(x_i; \theta) - f(x_i; \hat{\theta})$. The second aspect is that PI indicates the accuracy of the prediction of the target. Therefore, Equation (2) can present the relation between PI and CI as:

$$t_i - f(x_i; \hat{\theta}) = [f(x_i; \theta) - f(x_i; \hat{\theta})] + e_i \quad (2)$$

It can be concluded from Equation (2) that PIs are wider than CIs, where PI contains CI and covers more sources of uncertainty [54].

3. PIs Assessment Measures

The most commonly used measures for quantifying PIs construction are PI coverage probability (PICP) and mean PI width (MPIW). The coverage measure corresponds to the

encompassments of the obtained bounds. The wider PIs increase the PICP. The PICP is calculated as [24]:

$$\text{PICP} = \frac{1}{N} \sum_{i=1}^N c_i c_i = \begin{cases} 1, & x_i \in [L_i, U_i] \\ 0, & x_i \notin [L_i, U_i] \end{cases} \quad (3)$$

where N is the number of samples, and L_i and U_i are the lower and upper bounds of the i th PI, respectively. The second measure is used to evaluate the width of PIs. Normalized MPIW (NMPIW) shows the normalized width as [24]:

$$\text{NMPIW} = \frac{1}{NR} \sum_{i=1}^N L(X_i) - U(X_i) \quad (4)$$

where R is the range of the observed values. NMPIW, as a dimensionless criterion, indicates the mean width of PIs. Each of these criteria separately cannot lead to a clear judgment due to their inverse relationship. Therefore, the combinational coverage width-based criterion (CWC) containing both criteria can be used to evaluate the estimated PIs as [24]:

$$\text{CWC} = \text{NMPIW} (1 + \gamma (\text{PICP}) e^{-\eta (\text{PICP} - \mu)}) \gamma = \begin{cases} 0, & \text{PICP} \geq \mu \\ 1, & \text{PICP} < \mu \end{cases} \quad (5)$$

η and μ are fixed parameters, which determine the PIs with the lower value of the PICP. μ represents the confidence level of the PIs. η magnifies variation of the PICP and μ . Different η values should be examined to determine the most appropriate η value via a trial–error process. The coverage and width criteria are the most commonly used measures to evaluate the PIs' quality, however some other statistical measures are also applied to evaluate the calculated PIs. For example, in order to evaluate the constructed PIs via the Monte Carlo method, mean and standard deviation of the Nash–Sutcliffe model efficiency for each run are calculated to measure the PIs' coverage [58].

4. PIs Construction Methods

There are some methods to calculate the PIs. The Bayesian and Monte Carlo are the traditional methods. The other most frequently used method is the bootstrap technique; besides LUBE, it is one of the reliable methods. In addition, there are some other methods, such as mean variance estimation [11] and first order uncertainty analysis (FOUA) [53], but as they are not so prevalent, in the following sub sections the most common methods and their applications in hydro-climatic studies are described.

4.1. Bayesian Method

The Bayesian method was introduced by MacKay [21]. Training of the classic ANN is based on minimizing the error function, which leads to obtaining the optimum weights. Whereas, the Bayesian method attempts to train the ANN for the posterior probability distribution of weights from assumed prior probability distribution using Bayes' theorem. In the Bayesian method, ANN training is performed based on the regularized cost function as:

$$E(\omega) = \rho E_{\omega} + \beta E_D \quad (6)$$

where E_D is the sum of squared error and E_{ω} is the sum of squares of the network weights. ρ and β are used to determine training goals. The concept of this method is based on consideration of the set of ANN parameters, ω as a random set of variables with presumed distributions.

The Bayes' rule is applied to update the density function of the weights as [14]:

$$P(\omega|D, \rho, \beta, M) = \frac{P(D|\omega, \beta, M) P(\omega|\rho, M)}{P(D|\rho, \beta, M)} \quad (7)$$

where M and D are the NN model and the training dataset. $P(D|\omega, \beta, M)$ and $P(\omega|\rho, M)$ are the likelihood function of data occurrence and the prior density of parameters, respectively. Representing our knowledge, $P(D|\rho, \beta, M)$ is a normalization factor enforcing that the total probability is 1. Assuming that noises are normally distributed and $P(D|\omega, \beta, M)$ and $P(\omega|\rho, M)$ have normal distributions, it can be concluded that [14]:

$$P(D|\omega, \beta, M) = \frac{1}{Z_D(\beta)} e^{-\beta E_D} \quad (8)$$

$$P(\omega|\rho, M) = \frac{1}{Z_\omega(\rho)} e^{-\rho E_\omega} \quad (9)$$

where $Z_D(\beta) = (\frac{\pi}{\beta})^{n/2}$ and $Z_\omega(\rho) = (\frac{\pi}{\rho})^{p/2}$. n and p are the number of training samples and NN parameters, respectively. By substituting Equations (8) and (9) into (7), Equation (10) is obtained as [14]:

$$P(\omega|D, \rho, \beta, M) = \frac{1}{Z_F(\beta, \rho)} e^{-(\rho E_\omega + \beta E_D)} \quad (10)$$

The ANN is trained via maximization of the posterior probability $P(\omega|D, \rho, \beta, M)$, which is based on the minimizing Equation (6). By taking derivatives with respect to the logarithm of (10) and setting it equals to zero, the optimal values for β and ρ are obtained [14]:

$$\beta^{MP} = \frac{\gamma}{E_D(\omega^{MP})} \quad (11)$$

$$\rho^{MP} = \frac{n - \gamma}{E_\omega(\omega^{MP})} \quad (12)$$

where $\gamma = p - 2\rho^{MP} \text{tr}(H^{MP})^{-1}$ is the so-called effective number of ANN parameters, and p is the total number of ANN model parameters. ω^{MP} are the most probable values of the ANN parameters. H^{MP} (Equation (13)) is the Hessian matrix of $E(\omega)$ as [14]:

$$H^{MP} = \rho \nabla^2 E_\omega + \beta \nabla^2 E_D \quad (13)$$

The approximation of the Hessian matrix is generally performed using the Levenberg–Marquardt optimization algorithm. Application of this technique for the training process results in ANNs having the variance as [14]:

$$\sigma_i^2 = \sigma_D^2 + \sigma_{\omega^{MP}}^2 \frac{1}{\beta} + \nabla_{\omega^{MP}}^T \hat{y}_i (H^{MP})^{-1} \nabla_{\omega^{MP}} \hat{y}_i \quad (14)$$

The uncertainties corresponding to the data and parameters, respectively, are quantified via the term in the right and left sides of Equation (15). Finally, PI can be calculated considering the total variance of the i th future sample as [14]:

$$PI = \hat{y}_i \pm z^{1-\frac{\alpha}{2}} \left(\frac{1}{\beta} + \nabla_{\omega^{MP}}^T \hat{y}_i (H^{MP})^{-1} \nabla_{\omega^{MP}} \hat{y}_i \right)^{\frac{1}{2}} \quad (15)$$

where $z^{1-\frac{\alpha}{2}}$ is the $1-(\alpha/2)$ quantile of the normal distribution function with zero mean and unit variance. Additionally, $\nabla_{\omega^{MP}}^T \hat{y}_i$ is the gradient of the ANN output with respect to its parameters' set of ω^{MP} . The Bayesian method for PI construction has a strong mathematical foundation. This method requires calculation of the Hessian matrix, which needs a significant amount of time, but it should be considered that the computational load is lower in the process of constructing PIs because of only calculating the gradient of Neural Network (NN) output.

Some studies applied the Bayesian method in order to construct the PIs. Kasiviswanathan and Sudheer [51] used the bootstrap and Bayesian techniques to assess the uncertainty of the flood forecasting models of the ANN. The method application was based on the assumption that the model structure is deterministic, therefore, only the parameter of uncertainty was assessed in this study. It was concluded that model implementation is acceptable when the ensemble mean is considered. It was concluded that the bootstrap method is simple and easy in the case of the implementation as compared to the Bayesian method, but comparison of the obtained results showed that the Bayesian method led to narrower PIs and lower variance in parameter convergence.

4.2. Monte Carlo Method

The performance of the Monte Carlo method is based on alteration of the model inputs, parameters or structure of their ensemble. The number of iterations depends on the required level of reliability and is a problem-dependent task. More repetition leads to more reliable results, but higher computational cost should also be examined. If the model structure and the input data are assumed to be certain, Equation (16) can be presented as [58]:

$$\hat{y}_{t,i} = M(x, \theta_i); t = 1, 2, \dots, n; i = 1, 2, \dots, s \quad (16)$$

where θ_i is the set of parameters sampled for the i th run of the Monte Carlo simulation, $\hat{y}_{t,i}$ is the model output of the i th time step for the i th run, n is the number of time steps and s is the number of simulations. The statistical properties (such as moments and quantiles) of the model output for each time step t are estimated from the realizations $\hat{y}_{t,i}$. In order to quantile the uncertainty, the following equation can be expressed [58]:

$$P(\hat{y}_t < \hat{Q}(p)) = \sum_{i=1}^s w_i |\hat{y}_{t,i} < \hat{Q}(p)| \quad (17)$$

where, \hat{y}_t is the model output at time step t , $\hat{y}_{t,i}$ is the value of model output at time t simulated by the model $M(x, \theta_i)$ in i th simulation, $\hat{Q}(p)$ is p th [0,1] quantile, w_i is the weight given to the model output in i th at simulation. Quantiles obtained in this way are conditioned on the inputs to the model, the model structure and the weight vector w_i . The computation of the model's PI with confidence level α ($0 < \alpha < 1$) is achieved through estimation of the $\frac{1-\alpha}{2} * 100\%$ and $\frac{1+\alpha}{2} * 100\%$ via the $\hat{y}_{t,i}$. The lower prediction limit PL^L and the upper prediction limit PL^U are calculated as [58]:

$$\hat{Q}(p) = PL^L \text{ where } p = (1 - \alpha)/2 \quad (18)$$

$$\hat{Q}(p) = PL^U \text{ where } p = (1 + \alpha)/2 \quad (19)$$

Then, the PI is derived considering the output of the calibrated (optimal) model (\bar{y}) as:

$$PI^L = \bar{y} - PL^L, PI^U = PL^U - \bar{y} \quad (20)$$

where PI^L and PI^U are the interval of the obtained results as lower and upper bounds, respectively. It should define the purpose of the work and its significance.

Shrestha, Kayastha and Solomatine [58] applied the Monte Carlo method to assess the parametric uncertainty of the analysis of hydrological models of rainfall runoff using ANN. It was shown that the Monte Carlo method could be used to determine the other sources of uncertainty, such as input, structure or their combination. Tapoglou, Varouchakis, Trichakis and Karatzas [60] applied the Monte Carlo technique to investigate the uncertainty associated with modeling of the hydraulic head in an aquifer via the ANN. The model was performed 300 times by various training sets, and initial random values and the training results constituted a sensitivity analysis of the ANN training to the kriging part of the algorithm. This study concluded that error intervals for the train and test data of the ANN and kriging PIs were narrow, considering the complexity of the study area. Application of

the Bayesian kriging methodology was assessed, and it was concluded that the difference between the predicted values and the results of simulation of the actual data was low. This method led to consistent and reliable performance in different conditions. It was assessed that this method is appropriate to simulate groundwater-level, chiefly in cases with complicated behavior and unknown geological data.

Seifi, Ehteram, Singh and Mosavi [59] attempted to evaluate the uncertainty of groundwater level modeling via the hybrid ANN modeling and some other black box models and six meta-heuristic optimization methods, such as the grasshopper algorithm, cat swarm, weed algorithm, genetic algorithm, krill algorithm and particle swarm optimization, and it was mentioned that hybrid methods led to better performance and accuracy than sole methods.

4.3. Bootstrap Method

In this technique, several ANNs (B ANNs) are trained with randomly selected sub-sets [22] (see Figure 3). The randomly selected samples from total data are used for training each of networks. This technique is based on ensembling some ANNs, which could lead to lower estimation errors with regard to a single ANN. This method is independent of any complicated calculation using the non-linear operator or function. A model as $f_{ANN}(x)$ is fitted to each of the generated bootstrap sub-sets, and the bootstrapping estimate is calculated as the average and variance of each model as:

$$\hat{y}_{boot}(x) = \frac{1}{B} \sum_{b=1}^B f_{ANN}^b(x) \quad (21)$$

$$\hat{\sigma}_{boot}^2(x) = \frac{1}{B-1} \sum_{b=1}^B (f_{ANN}^b(x) - \hat{y}_{boot}(x))^2 \quad (22)$$

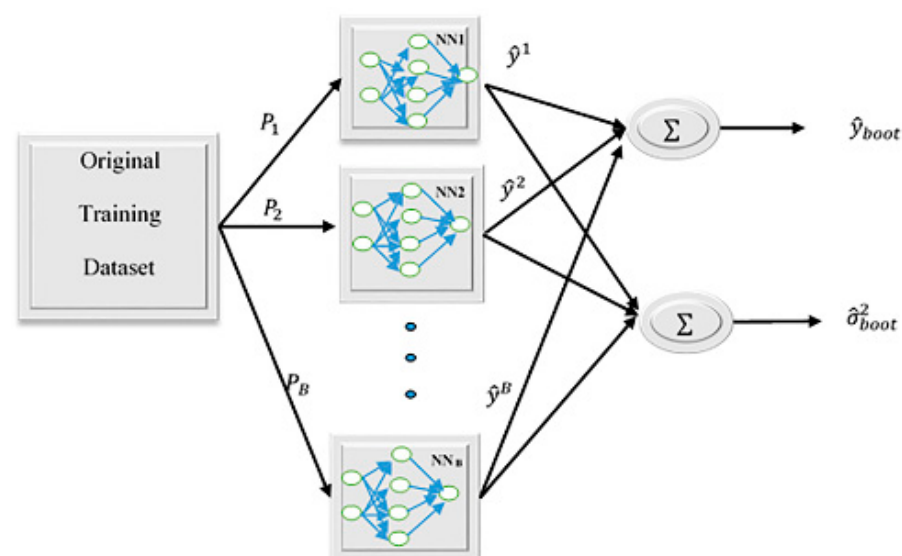


Figure 3. Schematic of the bootstrap method [14]. P_B stands for training data sub-sets.

For constructing the PIs [l,u], values of the observed $X = (x_1, x_2, \dots, x_3)$ with normal distribution probability of P and according to Equations (21) and (22), $P(l < X < u)$ are as:

$$P(l < X < u) = P\left(\frac{l - \hat{y}_{boot}}{\hat{\sigma}_{boot}} < \frac{X - \hat{y}_{boot}}{\hat{\sigma}_{boot}} < \frac{u - \hat{y}_{boot}}{\hat{\sigma}_{boot}}\right) \quad (23)$$

where $Z = \frac{X - \hat{y}_{boot}}{\hat{\sigma}_{boot}}$, the standard score of X , is distributed as standard normal [69], hence:

$$\frac{1 - \hat{y}_{boot}}{\hat{\sigma}_{boot}} = -z ; \quad (24)$$

or $l = \hat{y}_{boot} - z\hat{\sigma}_{boot}$; $u = \hat{y}_{boot} + z\hat{\sigma}_{boot}$.

Table 3 presents the corresponding values of Z and PIs.

Table 3. Values of Z for different PIs [70].

PI	z
75%	1.15
90%	1.64
95%	1.96
99%	2.58

Talebizadeh and Moridnejad [52] applied the bootstrap method to assess the uncertainty arising from measurement error and also the uncertainty of the ANNs' output for forecasting the lake level fluctuations. It was stressed that PIs' estimation provided beneficial information for decision making and designing. Kasiviswanathan and Sudheer [53] combined the bootstrap and the FOUA method to investigate the parametric and predictive uncertainty of the rainfall-runoff ANN-based modeling to forecast river flow. It was concluded that the FOUA method could compute the sensitivity coefficients that are the first order partial derivative of the model output and parameters of modeling. In this method, the computational burden and time of simulation for uncertainty analysis are reduced due to the usage of the statistical parameters such as mean and variance of the ANN weight vectors and biases. The parameter variability was determined via the bootstrap method. The obtained results for uncertainty analysis were quantified via the coverage and width criteria. It was concluded that the results for training and verifying data sets matched each other. Moreover, uncertainty associated with various domains of flow (low, medium and high) was assessed to identify the effect of the magnitude of flow on uncertainty; the results indicated that the uncertainty level changed with different flow regimes, proportionally. The overall results, considering both width and coverage criteria, show that the FOUA method led to a better quantification of the prediction uncertainty compared to the bootstrap method. Wang, Zheng, Zhao, Jiang, Wang, Guo and Wang [55] used the bootstrap method to calculate the PIs of water quality modeling via the wavelet-ANN approach. The uncertainty of the model structure and data noise was investigated, and it was shown that the application of the wavelet data pre-processing could lead to more accurate results. Kumar, Tiwari, Chatterjee and Mishra [39] used the combination of the bootstrap method and wavelet-ANN to quantify the uncertainty associated with the reservoir inflow forecasting. Moreover, multiple linear regression model implementation was compared to the bootstrap method and it was concluded that the performance of the bootstrap method is more reliable. It was also mentioned that PIs' estimation could provide more useful information in operational inflow forecasting. Kasiviswanathan, He, Sudheer and Tay [56] used the bootstrap technique for the quantification of the uncertainty associated with modeling streamflow and flood management. The coverage and width criteria were applied for quantification of the model's performance uncertainty. Results indicated that the bootstrap method is the proper method for streamflow forecasting and flood management. Moreover, it was mentioned that there are some limitations associated with forecasting high flow due to the lower samples and dependence of the ANN on the number of samples. Thus, it was suggested that one use hybrid modeling and integrate the data-driven models with physically-based/conceptual models and/or empirical relationships between high flows and influencing factors for the enhancement of the accuracy of the model for modeling high flow regimes.

4.4. LUBE Method

The LUBE method is based on training an ANN with two outputs for developing the PIs in one level (see Figure 4). Two outputs present the upper and lower bounds of the PIs. The proposed ANN is trained based on minimizing the defined cost function. Cost function contains both coverage and width criteria. This method is independent of special information about the PI bound. Previous studies assumed that PIs developed via this method are superior to the other PIs construction methods. In addition, its computational expense is insignificant [24]. Unlike traditional methods, LUBE method performance is independent of point prediction and it is a non-parametric technique. LUBE method application does not depend on parametric distribution of data. It is fast and simple [71].

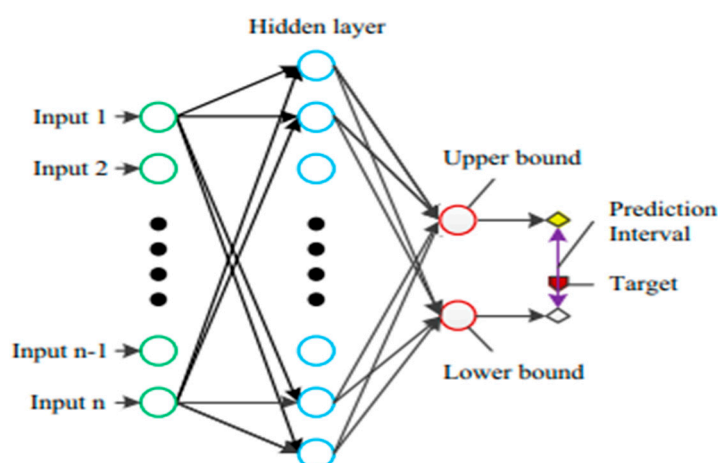


Figure 4. The structure of the ANN in the lower upper bound estimation (LUBE) method.

Kasiviswanathan, Cibin, Sudheer and Chaubey [61] attempted to develop the PIs of the ANN-based rainfall-runoff modeling by generation of the ensemble predictions (similar to the LUBE method). PIs were developed at two levels. At the first level, optimum ANN parameters were obtained via the genetic algorithm. In the second step, the ensemble of the models was created by optimization of the verity of the ANN parameters. PIs were calculated by minimizing the residual variance of the ensemble mean and maximization of the covering targets. Moreover, at the same time, the minimization of the PIs' width was taken into account. It was stated that by consideration of the ensemble mean value as the output of the model, the peak flow could be predicted more precisely compared to the classic point prediction of the ANN. Taormina and Chau [62] examined the LUBE method for construction of the PIs at different confidence levels for the 6 hours ahead streamflow discharges forecasting. Particle swarm optimization was used to minimize the CWC cost function. It was shown that the obtained results depend on the used particle swarm optimization paradigm. They claimed that the multi-objective framework led to more appropriate results than single-objective swarm optimization. It was also concluded that the applied algorithm to develop the model could have a remarkable effect on the PIs quality. Zhang, Zhou, Ye, Zeng and Chen [63] applied the LUBE method to construct the PIs of flood forecasting. This study proposed a PI symmetry index and objective function to evaluate the coverage, width and symmetry of PIs. To optimize the proposed objective function, the shuffled complex evolution algorithm was applied. The mean of the bounds was used to present deterministic forecasting. Kasiviswanathan and Sudheer [51] applied the PI method (similar to the LUBE method) for quantification of the ANN modeling uncertainty. The coverage and width criteria were used to quantify the PIs. This paper compared other PI construction methods results, such as the Bayesian and bootstrap methods, and concluded that the PI method was more reliable. In addition, it was presented that the PI method could successfully capture the peak points. Kasiviswanathan, Sudheer and He [64] assessed the uncertainty associated with input and parameters for

the ANN-based rainfall-runoff modeling. A two-stage optimization method was applied to estimate the PIs. Chen and Chau [66] used the LUBE method to construct the PIs of sediment load modeling via the hybrid double feedforward NN. It was demonstrated that PIs led to appropriate results in the 90% and 95% CLs. The proposed method could generate reliable PIs. It was discussed that application of the hybrid double feedforward NN could improve performance of the PIs construction in the classification of the low, medium and high sediment loads, and coverage probability about 100% for low and medium sediment loads, but its performance was weak for modeling high sediment loads. Moreover, it was concluded that the LUBE method was efficient in quantifying the uncertainty of data-driven models. Nourani, Paknezhad, Sharghi and Khosravi [57] used the LUBE method to construct the PIs associated with the ANN-based downscaling of the general circulation models. In this study, the LUBE method was applied by generating multiple sets of weights to develop narrow PIs with high coverage probability. It was indicated that the LUBE method could be successfully used to compute the PIs of ANN-based downscaling with reliable performance. Nourani, Sayyah-Fard, Alami and Sharghi [65] quantified the uncertainty of the ANN-based evaporation modeling via the LUBE method. It was claimed that the LUBE method could construct PIs with an appropriate level of reliability; however, data pre-processing methods could affect the uncertainty. This study applied simulated annealing optimization algorithms to construct PIs with higher coverage and lower width. It was mentioned that this method could overcome the problem of trapping in local minima. Kasiviswanathan, Sudheer, Soundharajan and Adeloye [67] applied upper lower bound and mean of forecasting to evaluate uncertainty in inflow modeling via the ANN for optimizing the reservoir operation and decision making. An integrated simulation–optimization was applied, which led to minimizing the error. In Figure 5, the procedure of PIs construction is depicted.

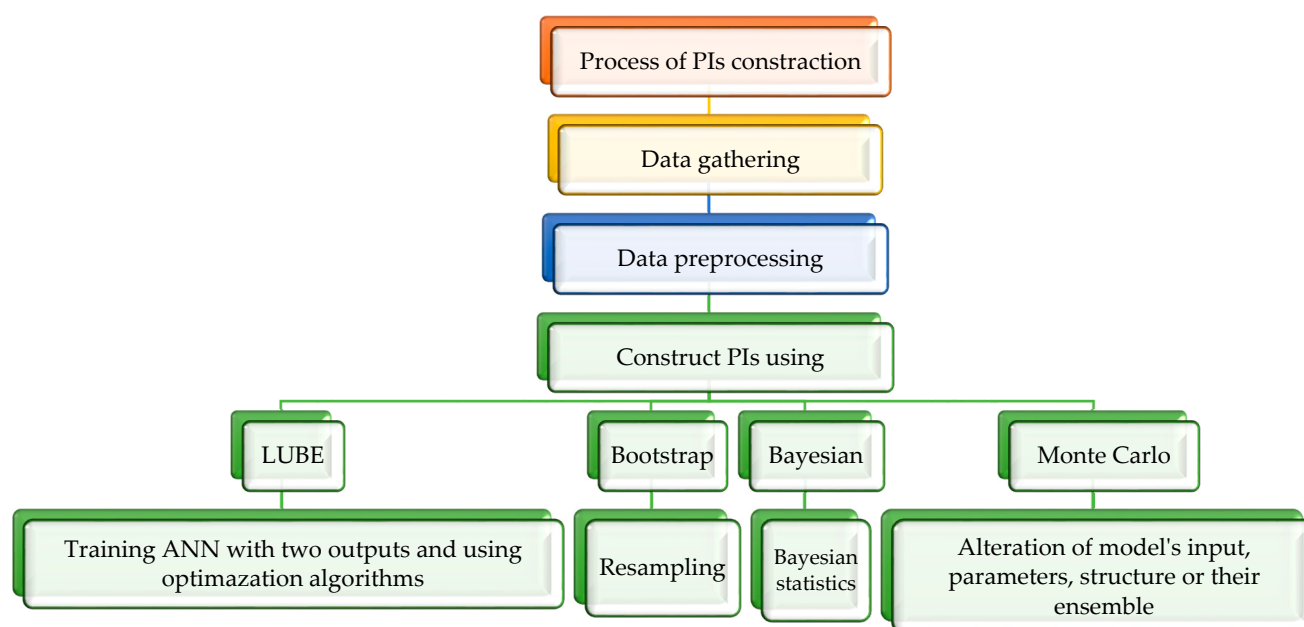


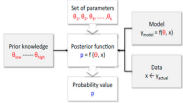
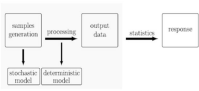
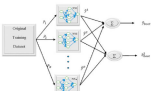
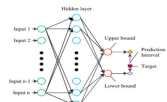
Figure 5. The PIs construction procedure.

5. Comparison of the PIs Construction Methods

Various methods of PIs construction with different levels of complexity, computational burden, difficulty in implementation, reliability and required times have been developed. Undoubtedly, it is impossible to claim that a particular technique is superior to the others, but each method has its own advantages. Therefore, in this section, the advantage and

disadvantages of the prevalent methods are investigated. A brief comparison of each method is tabulated in Table 4.

Table 4. The brief comparison of the proposed methods.

Method		Run Time	Number of Required Networks	Computational Burden	Reliability	Reference
Bayesian	 <p>Conceptual representation of the Bayesian model for inference of parameters.</p>	high	One	high	high	[21]
Monte Carlo	 <p>General overview of the Monte Carlo algorithm.</p>	high	More than 1	medium	medium	[58]
Bootstrap	 <p>Schematic of the bootstrap method, PB stands for training data sub-sets.</p>	high	More than 1	low	medium	[14]
LUBE	 <p>The structure of the ANN in the LUBE method</p>	low	One	low	high	[24]

Khosravi, Nahavandi, Creighton and Atiya [14] compared the different methods of PI construction for data from various domains. It was assumed that the delta method obtains the highest quality of the PIs, but its repeatability (performance of the method in worst cases) is not acceptable. Moreover, the constructed PIs via the delta technique consist of fixed PI width. The Bayesian method is the most acceptable method in the case of reproducibility (various iterations of the method lead to similar results). Moreover, the stability of developed PIs is the other advantage of this method. The bootstrap method is favorable in the case of variability (the response of PIs to the level of uncertainty associated with data). However, the obtained PIs via the bootstrap method may lead to low quality in comparison to the delta and Bayesian methods. The variance of the outputs may be overestimated, which may cause wider PIs. In addition, it has been demonstrated that by increasing the number of training iterations, the bootstrap method might not definitely improve the results. It has also been indicated that each method has its pros and cons, and implementation of different methods by consideration of various criteria may lead to different outcomes. Kasiviswanathan and Sudheer [51] compared different PIs construction methods and showed that parameters coverage and peak flow prediction are high for the PI method and low for the bootstrap method. Fulfillment of statistical and probabilistic assumptions is low for the bootstrap method, but it is high for the Bayesian method.

The difficulty of the implementation of the bootstrap, Bayesian and LUBE methods are, respectively, medium, high and low.

6. Gaps and Suggestions for Future Studies

The following issues are suggested for future studies to fill the existing gaps in already performed studies.

- i. Most of the PIs construction methods applied coverage and width criteria and their combination, but there is not any criterion in order to present information about the probability and reliability of the lower and upper bounds of the constructed PIs. Therefore, future studies may develop and present some criteria about this issue, for example, proximity of the target to the upper or lower bounds.
- ii. The reviewing of multiple studies showed that there is not any study that uses the delta method for PIs construction. Although it may have high level of computational cost, according to its reliable performance in other fields as expressed in Khosravi, Nahavandi, Creighton and Atiya [14], it is proposed that one apply this method in hydrological studies as well.
- iii. The performance of the LUBE method, as the most robust method of PI construction, can be improved in different aspects in order to obtain more reliable PIs. The implementation could be augmented by combining with the ANN structure selection techniques.
- iv. Adaptation between point prediction and PIs has not been examined, yet significantly presented. Therefore, it can be recommended that future works analyze the adaptation and correlation between point prediction and PIs construction. Moreover, some criteria can be defined to capture simultaneously.
- v. It is recommended that one apply the presented methods to assess the uncertainty associated with the improved version of the ANN, such as emotional ANN [72] and to investigate the effects of hormonal parameters in the reliability of the models.
- vi. The LUBE method performance is based on the cost functions. Some studies used multi-objective optimization cost function in which the coverage and width criteria were considered as cost function simultaneously. In contrast, some other studies used coverage and width combination criteria as the cost function, in which some parameters should be determined by trial and error. Therefore, future studies can compare the implementation of multi-objective and single-objective optimization methods. Moreover, future studies can propose the appropriate value of CWC parameters or propose a method for better and faster determination of the parameters' values.
- vii. As there are few studies attempting to investigate other artificial intelligence methods such as ANFIS, it is recommended that one construct the PIs of those methods too.

7. Conclusions

This study investigated multiple PIs construction methods applied in hydro-climatic studies. It concluded that the bootstrap method has been used in the majority of the studies as it is simple and can be applied easily. Moreover, the LUBE method has gained noticeable attention recently in hydrological studies due to its superiority in implementation and reliability compared to other methods. Nevertheless, there are few applications of the Bayesian or delta method in the development of PIs in the hydrological issues.

Author Contributions: All authors have contributed equally to all sections of this manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data will be available upon request.

Acknowledgments: This work has been supported by University of Tabriz, International and Academic Cooperation Directorate, in the framework of TabrizU-300 program.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Loucks, D.P. Sustainable Water Resources Management. *Water Int.* **2000**, *25*, 3–10. [\[CrossRef\]](#)
2. Nourani, V.; Kalantari, O. Integrated Artificial Neural Network for Spatiotemporal Modeling of Rainfall–Runoff–Sediment Processes. *Environ. Eng. Sci.* **2010**, *27*, 411–422. [\[CrossRef\]](#)
3. Nourani, V.; Kisi, O.; Komasi, M. Two hybrid Artificial Intelligence approaches for modeling rainfall–runoff process. *J. Hydrol.* **2011**, *402*, 41–59. [\[CrossRef\]](#)
4. Sharghi, E.; Nourani, V.; Najafi, H.; Molajou, A. Emotional ANN (EANN) and Wavelet-ANN (WANN) Approaches for Markovian and Seasonal Based Modeling of Rainfall-Runoff Process. *Water Resour. Manag.* **2018**, *32*, 3441–3456. [\[CrossRef\]](#)
5. Sharifi, S.S.; Delirhasannia, R.; Nourani, V.; Sadraddini, A.A.; Ghorbani, A. Using artificial neural networks (ANNs) and adaptive neuro-fuzzy inference system (ANFIS) for modeling and sensitivity analysis of effective rainfall. *Recent Adv. Contin. Mech. Hydrol. Ecol.* **2013**, *3*, 133–139.
6. Latifoglu, L. Evaluating Stream Flow Forecasting Performance Using Adaptive Network Based Fuzzy Logic Inference System, Artificial Neural Networks with Feature Selection. *EPSTEM* **2020**, *11*, 125–130.
7. Klomjit, J.; Ngaopitakkul, A. Comparison of Artificial Intelligence Methods for Fault Classification of the 115-kV Hybrid Transmission System. *Appl. Sci.* **2020**, *10*, 3967. [\[CrossRef\]](#)
8. Niu, W.-J.; Feng, Z.-K.; Feng, B.-F.; Xu, Y.-S.; Min, Y.-W. Parallel computing and swarm intelligence based artificial intelligence model for multi-step-ahead hydrological time series prediction. *Sustain. Cities Soc.* **2021**, *66*, 102686. [\[CrossRef\]](#)
9. Nourani, V.; Komasi, M.; Mano, A. A Multivariate ANN-Wavelet Approach for Rainfall–Runoff Modeling. *Water Resour. Manag.* **2009**, *23*, 2877–2894. [\[CrossRef\]](#)
10. Elshorbagy, A.; Corzo, G.; Srinivasulu, S.; Solomatine, D.P. Experimental investigation of the predictive capabilities of data driven modeling techniques in hydrology—Part 1: Concepts and methodology. *Hydrol. Earth Syst. Sci.* **2010**, *14*, 1931–1941. [\[CrossRef\]](#)
11. Huang, F.; Huang, J.; Jiang, S.; Zhou, C.-B. Landslide displacement prediction based on multivariate chaotic model and extreme learning machine. *Eng. Geol.* **2017**, *218*, 173–186. [\[CrossRef\]](#)
12. Refsgaard, J.C.; Van Der Sluijs, J.P.; Brown, J.; Van Der Keur, P. A framework for dealing with uncertainty due to model structure error. *Adv. Water Resour.* **2006**, *29*, 1586–1597. [\[CrossRef\]](#)
13. Zio, E.; Aven, T. Uncertainties in smart grids behavior and modeling: What are the risks and vulnerabilities? How to analyze them? *Energy Policy* **2011**, *39*, 6308–6320. [\[CrossRef\]](#)
14. Khosravi, A.; Nahavandi, S.; Creighton, D.; Atiya, A.F. Comprehensive Review of Neural Network-Based Prediction Intervals and New Advances. *IEEE Trans. Neural Netw.* **2011**, *22*, 1341–1356. [\[CrossRef\]](#)
15. Quan, H.; Srinivasan, D.; Khosravi, A. Incorporating Wind Power Forecast Uncertainties into Stochastic Unit Commitment Using Neural Network-Based Prediction Intervals. *IEEE Trans. Neural Netw. Learn. Syst.* **2015**, *26*, 2123–2135. [\[CrossRef\]](#)
16. Chatfield, C. Calculating interval forecasts. *J. Bus. Econ. Stat.* **1993**, *11*, 121–135.
17. Ma, J.; Niu, X.; Tang, H.; Wang, Y.; Wen, T.; Zhang, J. Displacement Prediction of a Complex Landslide in the Three Gorges Reservoir Area (China) Using a Hybrid Computational Intelligence Approach. *Complexity* **2020**, *2020*, 1–15. [\[CrossRef\]](#)
18. Shuai, Y.; Notton, G.; Duchaud, J.-L.; Almorox, J.; Yaseen, Z.M. Solar irradiation prediction intervals based on Box–Cox transformation and univariate representation of periodic autoregressive model. *Renew. Energy Focus* **2020**, *33*, 43–53. [\[CrossRef\]](#)
19. Chrysosolouris, G.; Lee, M.; Ramsey, A. Confidence interval prediction for neural network models. *IEEE Trans. Neural Netw.* **1996**, *7*, 229–232. [\[CrossRef\]](#)
20. Momotaz, B.; Dohi, T. Prediction Interval of Cumulative Number of Software Faults Using Multilayer Perceptron. *Flex. Gen. Uncertain. Optim.* **2015**, *619*, 43–58. [\[CrossRef\]](#)
21. Mackay, D.J.C. A Practical Bayesian Framework for Backpropagation Networks. *Neural Comput.* **1992**, *4*, 448–472. [\[CrossRef\]](#)
22. Kasiviswanathan, K.S.; Sudheer, K. Methods used for quantifying the prediction uncertainty of artificial neural network based hydrologic models. *Stoch. Environ. Res. Risk Assess.* **2017**, *31*, 1659–1670. [\[CrossRef\]](#)
23. Efron, B.; Tibshirani, R.J. *An Introduction to the Bootstrap*; CRC Press: Boca Raton, FL, USA, 1994.
24. Lu, J.; Ding, J.; Dai, X.; Chai, T. Ensemble Stochastic Configuration Networks for Estimating Prediction Intervals: A Simultaneous Robust Training Algorithm and Its Application. *IEEE Trans. Neural Netw. Learn. Syst.* **2020**, *31*, 5426–5440. [\[CrossRef\]](#)
25. Khosravi, A.; Nahavandi, S.; Creighton, D.; Atiya, A.F. Lower Upper Bound Estimation Method for Construction of Neural Network-Based Prediction Intervals. *IEEE Trans. Neural Netw.* **2010**, *22*, 337–346. [\[CrossRef\]](#)
26. Cannon, A.J.; Whitfield, P.H. Downscaling recent streamflow conditions in British Columbia, Canada using ensemble neural network models. *J. Hydrol.* **2002**, *259*, 136–151. [\[CrossRef\]](#)
27. Jeong, D.-I.; Kim, Y. Rainfall-runoff models using artificial neural networks for ensemble streamflow prediction. *Hydrol. Process.* **2005**, *19*, 3819–3835. [\[CrossRef\]](#)
28. Fleming, S.W.; Bourdin, D.R.; Campbell, D.; Stull, R.B.; Gardner, T. Development and Operational Testing of a Super-Ensemble Artificial Intelligence Flood-Forecast Model for a Pacific Northwest River. *JAWRA J. Am. Water Resour. Assoc.* **2014**, *51*, 502–512. [\[CrossRef\]](#)

29. Kan, G.; Yao, C.; Li, Q.; Li, Z.; Yu, Z.; Liu, C.; Ding, L.; He, X.; Liang, K. Improving event-based rainfall-runoff simulation using an ensemble artificial neural network based hybrid data-driven model. *Stoch. Environ. Res. Risk Assess.* **2015**, *29*, 1345–1370. [\[CrossRef\]](#)
30. Kim, S.E.; Seo, I.W. Artificial Neural Network ensemble modeling with conjunctive data clustering for water quality prediction in rivers. *HydroResearch* **2015**, *9*, 325–339. [\[CrossRef\]](#)
31. Kim, S.; Kim, H.S. Uncertainty Reduction of the Flood Stage Forecasting Using Neural Networks Model1. *JAWRA J. Am. Water Resour. Assoc.* **2008**, *44*, 148–165. [\[CrossRef\]](#)
32. Yang, C.-C.; Chen, C.-S. Application of integrated back-propagation network and self organizing map for flood forecasting. *Hydrol. Process.* **2009**, *23*, 1313–1323. [\[CrossRef\]](#)
33. Srivastav, R.; Sudheer, K.; Chaubey, I. A simplified approach to quantifying predictive and parametric uncertainty in artificial neural network hydrologic models. *Water Resour. Res.* **2007**, *43*. [\[CrossRef\]](#)
34. Boucher, M.; Perreault, L.; Anctil, F. Tools for the assessment of hydrological ensemble forecasts obtained by neural networks. *J. Hydroinform.* **2009**, *11*, 297–307. [\[CrossRef\]](#)
35. Sharma, S.K.; Tiwari, K. Bootstrap based artificial neural network (BANN) analysis for hierarchical prediction of monthly runoff in Upper Damodar Valley Catchment. *J. Hydrol.* **2009**, *374*, 209–222. [\[CrossRef\]](#)
36. Kant, A.; Suman, P.K.; Giri, B.K.; Tiwari, M.K.; Chatterjee, C.; Nayak, P.C.; Kumar, S. Comparison of multi-objective evolutionary neural network, adaptive neuro-fuzzy inference system and bootstrap-based neural network for flood forecasting. *Neural Comput. Appl.* **2013**, *23*, 231–246. [\[CrossRef\]](#)
37. Zhang, X.; Liang, F.; Srinivasan, R.; Van Liew, M. Estimating uncertainty of streamflow simulation using Bayesian neural networks. *Water Resour. Res.* **2009**, *45*. [\[CrossRef\]](#)
38. Khan, M.S.; Coulibaly, P. Assessing Hydrologic Impact of Climate Change with Uncertainty Estimates: Bayesian Neural Network Approach. *J. Hydrometeorol.* **2010**, *11*, 482–495. [\[CrossRef\]](#)
39. Zhang, X.; Liang, F.; Yu, B.; Zong, Z. Explicitly integrating parameter, input, and structure uncertainties into Bayesian Neural Networks for probabilistic hydrologic forecasting. *J. Hydrol.* **2011**, *409*, 696–709. [\[CrossRef\]](#)
40. Zhang, X.; Zhao, K. Bayesian Neural Networks for Uncertainty Analysis of Hydrologic Modeling: A Comparison of Two Schemes. *Water Resour. Manag.* **2012**, *26*, 2365–2382. [\[CrossRef\]](#)
41. Humphrey, G.B.; Gibbs, M.S.; Dandy, G.; Maier, H.R. A hybrid approach to monthly streamflow forecasting: Integrating hydrological model outputs into a Bayesian artificial neural network. *J. Hydrol.* **2016**, *540*, 623–640. [\[CrossRef\]](#)
42. Shen, S.; Zeng, G.; Liang, J.; Li, X.; Tan, Y.; Li, Z.; Li, J. Markov Chain Monte Carlo Approach for Parameter Uncertainty Quantification and Its Impact on Groundwater Mass Transport Modeling: Influence of Prior Distribution. *Environ. Eng. Sci.* **2014**, *31*, 487–495. [\[CrossRef\]](#)
43. Tongal, H.; Booi, M.J. Quantification of parametric uncertainty of ANN models with GLUE method for different streamflow dynamics. *Stoch. Environ. Res. Risk Assess.* **2017**, *31*, 993–1010. [\[CrossRef\]](#)
44. Snyder, H. Literature review as a research methodology: An overview and guidelines. *J. Bus. Res.* **2019**, *104*, 333–339. [\[CrossRef\]](#)
45. Torres-Carrión, P.V.; González-González, C.S.; Aciar, S.; Rodríguez-Morales, G. Methodology for Systematic Literature Review Applied to Engineering and Education. In Proceedings of the 2018 IEEE Global Engineering Education Conference (EDUCON), Tenerife, Spain, 17–20 April 2018; pp. 1364–1373.
46. Shamseer, L.; Moher, D.; Clarke, M.; Ghersi, D.; Liberati, A.; Petticrew, M.; Shekelle, P.; Stewart, L.A. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: Elaboration and explanation. *BMJ* **2015**, *349*, 1–25. [\[CrossRef\]](#)
47. Moher, D.; Shamseer, L.; Clarke, M.; Ghersi, D.; Liberati, A.; Petticrew, M.; Shekelle, P.; Stewart, L.A. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst. Rev.* **2015**, *4*, 1. [\[CrossRef\]](#)
48. Mengist, W.; Soromessa, T.; Legese, G. Method for conducting systematic literature review and meta-analysis for environmental science research. *MethodsX* **2020**, *7*, 100777. [\[CrossRef\]](#)
49. Hasan, H.H.; Razali, S.F.M.; Muhammad, N.; Ahmad, A. Research Trends of Hydrological Drought: A Systematic Review. *Water* **2019**, *11*, 2252. [\[CrossRef\]](#)
50. Harrison-Atlas, D.; Theobald, D.M.; Goldstein, J.H. A systematic review of approaches to quantify hydrologic ecosystem services to inform decision-making. *Int. J. Biodivers. Sci. Ecosyst. Serv. Manag.* **2016**, *12*, 160–171. [\[CrossRef\]](#)
51. Kasiviswanathan, K.S.; Sudheer, K.P. Comparison of methods used for quantifying prediction interval in artificial neural network hydrologic models. *Model. Earth Syst. Environ.* **2016**, *2*, 1–11. [\[CrossRef\]](#)
52. Talebizadeh, M.; Moridnejad, A. Uncertainty analysis for the forecast of lake level fluctuations using ensembles of ANN and ANFIS models. *Expert Syst. Appl.* **2011**, *38*, 4126–4135. [\[CrossRef\]](#)
53. Kasiviswanathan, K.S.; Sudheer, K. Quantification of the predictive uncertainty of artificial neural network based river flow forecast models. *Stoch. Environ. Res. Risk Assess.* **2012**, *27*, 137–146. [\[CrossRef\]](#)
54. Kumar, S.; Tiwari, M.K.; Chatterjee, C.; Mishra, A. Reservoir Inflow Forecasting Using Ensemble Models Based on Neural Networks, Wavelet Analysis and Bootstrap Method. *Water Resour. Manag.* **2015**, *29*, 4863–4883. [\[CrossRef\]](#)
55. Wang, Y.; Zheng, T.; Zhao, Y.; Jiang, J.; Wang, Y.; Guo, L.; Wang, P. Monthly water quality forecasting and uncertainty assessment via bootstrapped wavelet neural networks under missing data for Harbin, China. *Environ. Sci. Pollut. Res.* **2013**, *20*, 8909–8923. [\[CrossRef\]](#)
56. Kasiviswanathan, K.; He, J.; Sudheer, K.; Tay, J.-H. Potential application of wavelet neural network ensemble to forecast streamflow for flood management. *J. Hydrol.* **2016**, *536*, 161–173. [\[CrossRef\]](#)

57. Nourani, V.; Paknezhad, N.J.; Sharghi, E.; Khosravi, A. Estimation of prediction interval in ANN-based multi-GCMs downscaling of hydro-climatologic parameters. *J. Hydrol.* **2019**, *579*, 124226. [[CrossRef](#)]
58. Shrestha, D.L.; Kayastha, N.; Solomatine, D.P. A novel approach to parameter uncertainty analysis of hydrological models using neural networks. *Hydrol. Earth Syst. Sci.* **2009**, *13*, 1235–1248. [[CrossRef](#)]
59. Seifi, A.; Ehteram, M.; Singh, V.P.; Mosavi, A. Modeling and Uncertainty Analysis of Groundwater Level Using Six Evolutionary Optimization Algorithms Hybridized with ANFIS, SVM, and ANN. *Sustainability* **2020**, *12*, 4023. [[CrossRef](#)]
60. Tapoglou, E.; Varouchakis, E.A.; Trichakis, I.; Karatzas, G.P. Hydraulic head uncertainty estimations of a complex artificial intelligence model using multiple methodologies. *J. Hydroinform.* **2020**, *22*, 205–218. [[CrossRef](#)]
61. Kasiviswanathan, K.; Cibin, R.; Sudheer, K.; Chaubey, I. Constructing prediction interval for artificial neural network rainfall runoff models based on ensemble simulations. *J. Hydrol.* **2013**, *499*, 275–288. [[CrossRef](#)]
62. Taormina, R.; Chau, K.-W. ANN-based interval forecasting of streamflow discharges using the LUBE method and MOFIPS. *Eng. Appl. Artif. Intell.* **2015**, *45*, 429–440. [[CrossRef](#)]
63. Zhang, H.; Zhou, J.; Ye, L.; Zeng, X.; Chen, Y. Lower Upper Bound Estimation Method Considering Symmetry for Construction of Prediction Intervals in Flood Forecasting. *Water Resour. Manag.* **2015**, *29*, 5505–5519. [[CrossRef](#)]
64. Kasiviswanathan, K.; Sudheer, K.; He, J. Probabilistic and ensemble simulation approaches for input uncertainty quantification of artificial neural network hydrological models. *Hydrol. Sci. J.* **2018**, *63*, 101–113. [[CrossRef](#)]
65. Nourani, V.; Sayyah-Fard, M.; Aalami, M.T.; Sharghi, E. Data pre-processing effect on ANN-based prediction intervals construction of the evaporation process at different climate regions in Iran. *J. Hydrol.* **2020**, *588*, 125078. [[CrossRef](#)]
66. Chen, X.-Y.; Chau, K.-W. Uncertainty Analysis on Hybrid Double Feedforward Neural Network Model for Sediment Load Estimation with LUBE Method. *Water Resour. Manag.* **2019**, *33*, 3563–3577. [[CrossRef](#)]
67. Kasiviswanathan, K.S.; Sudheer, K.P.; Soundharajan, B.-S.; Adeloye, A.J. Implications of uncertainty in inflow forecasting on reservoir operation for irrigation. *Paddy Water Environ.* **2020**, 1–13. [[CrossRef](#)]
68. Shrestha, D.L.; Solomatine, D.P. Machine learning approaches for estimation of prediction interval for the model output. *Neural Netw.* **2006**, *19*, 225–235. [[CrossRef](#)]
69. Shrivastava, N.A.; Khosravi, A.; Panigrahi, B. Prediction Interval Estimation of Electricity Prices Using PSO-Tuned Support Vector Machines. *IEEE Trans. Ind. Inform.* **2015**, *11*, 322–331. [[CrossRef](#)]
70. Grant, E.L.; Leavenworth, R.S. *Statistical Quality and Control*; McGraw-Hill: New York, NY, USA, 1972.
71. Quan, H.; Srinivasan, D.; Khosravi, A. Particle swarm optimization for construction of neural network-based prediction intervals. *Neurocomputing* **2014**, *127*, 172–180. [[CrossRef](#)]
72. Nourani, V. An Emotional ANN (EANN) approach to modeling rainfall-runoff process. *J. Hydrol.* **2017**, *544*, 267–277. [[CrossRef](#)]