

Article Deep Learning-Based Predictive Framework for Groundwater Level Forecast in Arid Irrigated Areas

Wei Liu¹, Haijiao Yu^{2,*}, Linshan Yang¹, Zhenliang Yin¹, Meng Zhu¹, and Xiaohu Wen^{1,*}

- Key Laboratory of Ecohydrology of Inland River Basin, Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou 730000, China; weiliu@lzb.ac.cn (W.L.); yanglsh08@lzb.ac.cn (L.Y.); yinzhenliang@lzb.ac.cn (Z.Y.); zhumeng@lzb.ac.cn (M.Z.)
- Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China

Correspondence: yuhaijiao2011@163.com (H.Y.); xhwen@lzb.ac.cn (X.W.)

Abstract: An accurate groundwater level (GWL) forecast at multi timescales is vital for agricultural management and water resource scheduling in arid irrigated areas such as the Hexi Corridor, China. However, the forecast of GWL in these areas remains a challenging task owing to the deficient hydrogeological data and the highly nonlinear, non-stationary and complex groundwater system. The development of reliable groundwater level simulation models is necessary and profound. In this study, a novel ensemble deep learning GWL predictive framework integrating data pro-processing, feature selection, deep learning and uncertainty analysis was constructed. Under this framework, a hybrid model equipped with currently the most effective algorithms, including the complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) for data decomposition, the genetic algorithm (GA) for feature selection, the deep belief network (DBN) model, and the quantile regression (QR) for uncertainty evaluation, denoted as CEEMDAN-GA-DBN, was proposed for the 1-, 2-, and 3-month ahead GWL forecast at three GWL observation wells in the Jiuquan basin, northwest China. The capability of the CEEMDAN-GA-DBN model was compared with the hybrid CEEMDAN-DBN and the standalone DBN model in terms of the performance metrics including R, MAE, RMSE, NSE, RSR, AIC and the Legates and McCabe's Index as well as the uncertainty criterion including MPI and PICP. The results demonstrated the higher degree of accuracy and better performance of the objective CEEMDAN-GA-DBN model than the CEEMDAN-DBN and DBN models at all lead times and all the wells. Overall, the CEEMDAN-GA-DBN reduced the RMSE of the CEEMDAN-DBN and DBN models in the testing period by about 9.16 and 17.63%, while it improved their NSE by about 6.38 and 15.32%, respectively. The uncertainty analysis results also affirmed the slightly better reliability of the CEEMDAN-GA-DBN method than the CEEMDAN-DBN and DBN models at the 1-, 2- and 3-month forecast horizons. The derived results proved the ability of the proposed ensemble deep learning model in multi time steps ahead of GWL forecasting, and thus, can be used as an effective tool for GWL forecasting in arid irrigated areas.

Keywords: groundwater level forecasting; complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN); genetic algorithm (GA); feature selection (FS); deep belief network (DBN); quantile regression (QR); uncertainty evaluation

1. Introduction

Groundwater is an indispensable natural resource to provide the necessary water supply for the irrigation system and industry, as well as to support the health and survival of the natural ecosystem in arid and semi-arid regions where surface water is scarce [1–3]. However, with groundwater being over-exploited and potentially facing a risk of depletion, the sustainable management of the groundwater resources becomes a crucial task in these regions [4–6]. Moreover, in areas where groundwater is used for irrigation supply, water allocation is usually scheduled ahead of time according to the projected water



Citation: Liu, W.; Yu, H.; Yang, L.; Yin, Z.; Zhu, M.; Wen, X. Deep Learning-Based Predictive Framework for Groundwater Level Forecast in Arid Irrigated Areas. *Water* 2021, *13*, 2558. https:// doi.org/10.3390/w13182558

Academic Editor: Adriana Bruggeman

Received: 21 July 2021 Accepted: 16 September 2021 Published: 17 September 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/).



availability [7]. For this purpose, the practical forecast of groundwater resources is quite essential.

The groundwater level (GWL) is a key parameter required to quantify groundwater resources. The accurate and reliable prediction of the GWL is fundamental for evaluating the groundwater availability and effective management of the irrigation schedules [8,9]. Similar to other arid regions globally, the Jiuquan basin has experienced a general groundwater deterioration in both storage and sustainability [10]. Well-known for lying on the Silk Road and being a significant irrigated agricultural area of northwest China, the area urgently needs sustainable groundwater management. To achieve this, the accurate GWL forecast is of critical importance. However, detailed predictions for the region are still lacking.

Generally, the physically based models have traditionally been primary tools for GWL simulation [11–13]. However, the highly nonlinear, non-stationary and relatively complex hydrological system, the need of lots of hydrogeological data as well as the proper initial and boundary conditions bring plenty of difficulties in the characterization of the real hydrological system, and thus threaten the accuracy and the popularization of these models [14,15]. With the rapid growth of the data-based methods (mainly machine learning models), conventional algorithms such as the artificial neural network (ANN), the support vector machine (SVM), the extreme learning machine (ELM), the adaptive neuro-fuzzy inference system (ANFIS) and genetic programming (GP) have become viable techniques for groundwater forecasting owing to the greater simplicity in design and flexibility [16,17]. A comprehensive and explicit review of the machine learning application in GWL modeling can be found in Rajaee et al. [18]. However, the simplified architecture limits their ability to perform deeper feature extraction and they continue to be regarded as 'shallow learning' models as well. The dependable simulating and predicting GWL (including evaluating the underlying uncertainties) remains relatively challenging.

It is not until recently that the deep learning models (as a new class of machine learning models) have attracted great research attention, especially in pattern recognition, signal processing, time series analysis and complex modelling tasks such as image processing [19–21]. In terms of groundwater modeling, Zhang et al. [22] developed a Short-Term Long Memory (LSTM) network for the groundwater table, concluding its robust learning capability to replicate the patterns of the groundwater table. Compared with the other fields (e.g., solar radiation modeling), the application of the deep learning models in groundwater prediction is deficient.

In this study, one specific type of the deep learning model of interest is the deep belief network (DBN), which exhibits an outstanding ability to extract the inherent pairwise input(s)-target features from the lowest to the highest level [23–25]. Considering the complexity and nonlinearity of the GWL fluctuations, the DBN algorithm could be an appropriate choice. However, the utility of the DBN in GWL forecasting is few reported, urging the necessity to explore its potentiality in this field.

In light of the stochastic and non-stationary characteristics of the GWL fluctuations, effective data pre-processing tools such as wavelet transform (WT) are usually advocated. Successful applications of WT have been conducted in hydrological forecasting [26–29]. However, the excessive dependency on the selection of the mother wavelet and the decomposition level make the WT neither prior knowledge-based nor a regular or automated tool [18]. Recently, the Empirical Mode Decomposition (EMD), Ensemble EMD (EEMD) and their improved version (e.g., Complete Ensemble EMD with Adaptive Noise, CEEMDAN) have been proposed as alternatives. Compared with EMD and EEMD, the CEEMDAN algorithm reconstructs the original signal completely and noiselessly without trial-and-error processes. This outstanding advantage results in its feasible application in several hydrological prediction problems, such as the prediction of runoff, soil water, wind speed and others [30–34]. Comparatively speaking, the CEEMDAN model seems to outperform the other two models. Despite the achieved acceptable results of the CEEMDAN algorithm,

no prior study in multi-timescale GWL forecasting has been carried out to the best of the authors' knowledge.

Although the machine learning models' coupling data pre-processing techniques (e.g., EMD, EEMD or CEEMDAN) have achieved more accuracy than the traditional models, limitations still exist in practical usage. For example, the EMD, EEMD or the CEEMDAN algorithm decomposes a complex series into the intrinsic mode functions (IMFs) and residual component (Res), which may lead to the loss of vital information since some of the resulting IMFs become the unsystematic and disorderly parts among the other IMFs, and consequently, are likely to deteriorate the forecasting accuracy.

To solve this problem, solutions either to eliminate the IMFs that contained highly perturbing frequencies signals [35] or to build a model with a two-phase decomposition system [36–39] are commonly accepted. The removal of the irrelevant IMFs and selecting the most influential features are thus considered necessary steps when establishing any EMD, EEMD and CEEMDAN hybrid hydrological forecasting models [40,41]. Normally, the feature selection (FS) process is achieved collectively using the autocorrelation function (ACF) and the partial autocorrelation (PACF) method. Nevertheless, being purely linear, the ACF and PACF procedures cannot capture the nonlinear relationship between the target and the exploratory variables. Regarding the superior capability of machine learning in feature selection, the genetic algorithm (GA) is able to deduce the nonlinear relationships through the heuristic search and optimization techniques [42]. This study, therefore, attempts to integrate for the first time, the CEEMDAN decomposition method and the GA feature selection method into a deep belief network model for GWL forecasting.

In addition, in practical GWL forecasting applications, the machine learning models are often applied without considering the uncertainties. Although a majority of the EMD-, EEMD- and CEEMDAN (or WT-equivalent)-based deep learning models appear to improve the forecasting capability, the issue of uncertainty is often neglected. Moreover, the deterministic prediction may not be sufficient to describe the inherent uncertainties. In this respect, the quantile regression (QR) method, widely used for conditional quantiles estimation, can be a powerful tool. Without any parameters or prior assumptions, the QR method can account for the uncertainty from different sources [43]. This paper adopted the QR method as a data post-processing tool to estimate the uncertainty in modeling GWL variations.

Considering the numerous machine learning applications in GWL forecasting, this study aims to implement an innovative and systematic hybrid modelling framework in GWL forecasting based on the integrated inclusion of data decomposition (data preprocessing), deep learning, feature selection and uncertainty evaluation (data post-processing) (Table 1). The objectives are threefold.

Table 1. Summary of machine learning methods in groundwater level modeling for arid and semi-arid regions in the latest years.

Data Pre-Processing	Feature Selection	Machine Uncertainty Learning/Deep Analysis Learning Method		Hybrid Model	Reference
WT		SVM; Random Forest (RF); Linear Boosting of XGB; Tree Boosting of XGB		WT-SVR; WT-RF; WT-XGBL; WT-XGBT	Rahman et al. (2020) [16]
WT	Pearson correlation analysis	LSTM		WT-multivariate LSTM	Wu et al. (2021) [27]
		Deep Neural Network (DNN); SVM; ANFIS			Mohapatra et al. (2021) [44]

Data Pre-Processing	Feature Selection	Machine Learning/Deep Learning Method	Uncertainty Analysis	Hybrid Model	Reference
WT		Support Vector Regression (SVR); Gaussian Process Regression (GPR)		W-SVR; W-GPR	Band et al. (2021) [45]
	Information Gain (IG);Variance Inflation Factor (VIF)	RF; Logistic Regression; Decision Tree; ANN			Namous et al. (2021) [46]
		Multi-layer Perceptron; Radial Basis Function Network; SVM; DNN			Chen et al. (2021) [47]
		ANFIS; SVM; Kernel-based Nonlinear Arps decline; LSTM; Gated Recurrent Unit			Ao et al. (2021) [48]
		DNN	Modified delayed acceptance Markov Chain Monte Carlo (MCMC)		Lykkegaard et al. (2021) [49]
VMD	Boruta	ELM	Bootstrap	VBELM	Wu et al. (2021) [50]
CEEMDAN	GA	DBN	QR	CEEMDAN-GA- DBN	This study

Table 1. Cont.

(1) To develop a hybrid model (denoted as CEEMDAN-GA-DBN) with the highest configuration in data decomposition (CEEMDAN), feature selection (GA) and deep learning (DBN) and to test its validity in GWL forecasting at multi spatial and temporal scales for the arid Jiuquan Basin.

(2) To evaluate the robustness of the newly proposed GWL prediction scheme (i.e., CEEMDAN-GA-DBN) versus a pre-decomposed deep learning (i.e., CEEMDAN-DBN) model and a standalone deep learning (i.e., DBN) model.

(3) To evaluate all the possible uncertainties of the GWL forecasts by adopting the quantile regression (QR) method and combining the modelling procedure with the deterministic prediction results of the hybrid CEEMDAN-GA-DBN, the hybrid CEEMDAN-DBN and the standalone DBN model.

2. Materials and Methods

2.1. Methodological Framework

To achieve the goal of an accurate GWL forecast, the methodological framework illustrated in Figure 1 is proposed. Firstly, the original input series is partitioned into the training and testing subsets to prevent future data falling into the training process. After that, the pre-processing method is implemented to reconstruct the original input time series and to remove the stochastic and non-stationary characteristics. In this study, the CEEMDAN was applied to decompose the original data series into IMFs and residuals (Res). Then, the testing set was sequentially appended to the training set to escape from the 'hindcasting experiments' [51]. The same number of IMFs/Res was set for the training and testing datasets since both can be integrated to simulate the test data series. Secondly, the feature selection method ought to be applied to identify and remove the

unneeded, irrelevant and redundant attributes from the IMFs and Res decomposed using the CEEMDAN method that do not contribute to the accuracy of the predictive models. In this study, the GA feature selection method was applied to select the most appropriate variable(s) from all the CEEMDAN decomposed subseries. Then, these selected variables were further adopted for modelling. Thirdly, the deep learning model (DBN) uses subset data containing selected variables to forecast each IMF and Res. Afterwards, all extracted IMFs and Res predicted values were aggregated to generate the 1-, 2- and 3-month ahead GWL.



Figure 1. Schematic of the CEEMDAN-GA-DBN model.

Once the results of the GWL forecasting are obtained, the predictive uncertainty of the models is assessed using QR. The QR model is calibrated using the predicted GWL values as inputs and the true GWL values from the training dataset as outputs. Then, the desired quantiles (i.e., 0.05 and 0.95) of the GWL values are estimated with the calibrated QR model based on the predicted GWL in the testing set as inputs.

In this study, the hybrid CEEMDAN-GA-DBN model was proposed based on the abovementioned methodological framework. To benchmark the GWL forecast performance of the hybrid CEEMDAN-GA-DBN model, a hybrid CEEMDAN-DBN model without feature selection process and a standalone DBN model without any decomposition or feature selection were also applied for 1-, 2- and 3-month ahead GWL forecast at three GWL observation wells in the Jiuquan Basin, northwest China.

2.2. Data Decomposition Algorithm (Complete Ensemble Empirical Mode Decomposition with Adaptive Noise)

The complete ensemble empirical mode decomposition with adaptive noise (CEEM-DAN) algorithm is adopted as a data decomposition method to adaptively decompose the complex signals (e.g., GWL) into intrinsic mode functions (IMFs). The produced IMFs with less noise have better stability and regularity than the original data series [52]. Moreover, the CEEMDAN algorithm can alleviate the mode mixing problem by adopting the white Gaussian noise. The decomposition process of the algorithm can be summarized as follows.

(1) The mixed signal $x_i(t)$ is described as follows:

$$x_i(t) = x(t) + \omega_i(t), i = 1, 2, ..., k$$
 (1)

where i is the ensemble number of the trials, and x(t) and $w_i(t)$ denote the original signal and the white noise sequences, respectively.

(2) The CEEMDAN decomposes all the $x_i(t)$ into $c_{i1}(t)$ and $r_i(t)$ by using the EMD. It should be noted that the purpose of this step is to obtain the IMF1 and its residual item. The mean of the $c_{i1}(t)$ is calculated to obtain the $c_1(t)$ for the first IMF.

$$IMF_1 = c_1(t) = \frac{1}{k} \sum_{i=1}^k c_{i1}(t)$$
(2)

Correspondingly, the first residue item is defined.

$$r_1(t) = x(t) - c_1(t)$$
 (3)

(3) The CEEMDAN decomposes the white noise $\omega_i(t)$ using the EMD. The $E_j(\omega_i(t))$ is defined as a vector representing the j-th IMF of $\omega_i(t)$, where IMF2 can be express as follows:

$$IMF_{2} = c_{2}(t) = \frac{1}{k} \sum_{i=1}^{k} (E_{1}(\omega_{i}(t)) + r_{1}(t))$$
(4)

Additionally, its residual mode is as follows:

$$\mathbf{r}_{2}(t) = \mathbf{r}_{1}(t) - \mathbf{c}_{2}(t) \tag{5}$$

(4) Finally, the CEEMDAN algorithm calculates the following j-th residue:

$$r_{j}(t) = x(t) - c_{j}(t)$$
 (6)

(5) By repeating process (3)–(4), the other IMFs are obtained.

$$IMF_{(j+1)} = c_{j+1}(t) = \frac{1}{k} \sum_{i=1}^{k} \left(E_1(r_j(t)) + E_j(\omega_i(t)) \right)$$
(7)

$$r_{j+1}(t) = r_j(t) - c_j(t)$$
 (8)

The original signal x(t) is then expressed as follows:

$$\mathbf{x}(t) = \sum_{j=1}^{q} c_j(t) + \mathbf{r}(t)$$
 (9)

where q is the number of $c_i(t)$, and r(t) is the residual item of x(t).

This study conducted the CEEMDAN algorithm by using the 'Rlibeemd' package of the R software. The ensemble number was set to 200 and the amplitude of the added white noise was 0.2, both of which were acceptable [38,39].

Once the procedure of CEEMDAN was accomplished, the GA method was implemented to select the informative features. GA has long been applied in feature selection [53–55], readers can refer to Yang and Honavar [56] for more details, no longer repeated here. In terms of the parameters of the GA algorithm, the population size was set to 20, the maximum iteration number was 150, the crossover rate was 0.6 and the mutation rate was 0.02.

2.3. Deep Learning Forecast Algorithm (Deep Belief Network)

Integrated with the CEEMDAN process, an advanced form of the deep belief network (DBN) model was developed in this study. The DBN, by incorporating multiple hidden neuronal layers, is typically designed to improve the structure and the learning ability of the traditional ANN [57]. Thus, the model has been extensively applied to image recognition, feature extraction, time series forecasting and other practical fields that encounter relatively complex data [56].

The DBN is a generative model that produces the training data by training the weights between neurons according to the maximum probability in multiple sequentially stacked Restricted Boltzmann Machines (RBMs) (Figure 1). RBM is the primary unit of a DBN and consists of only two layers of neurons, the visible layer for the training input and the hidden layer for feature detection. Notably, there is a connection between the adjacent layers, but there are no relationships among the whole RBM layers. The process of training a DBN model is facilitated by the layer-by-layer approach, for which the hidden layer inferred from the data vector can be treated as that of the next layer (higher layer).

The learning process of a DBN model can be partitioned into two phases, an unsupervised pre-training phase and a supervised fine-tuning phase. In the unsupervised pre-training process, no target variable during the training procedure exists [57]. In this phase, when one of the RBMs has been trained, the learned parameters are kept, whereas the outputs of the hidden layer are seen as the input of the next RBM layer. In this manner, the RBM can be trained in a sequence until all RBM layers are finally trained. The pretraining process in a DBN model is required to acquire the optimum solution of the initial weights, and these values are later used in the subsequent (supervised fine-tuning) phase. The fine-tuning procedure is implemented to improve the model with supervised learning labeled data. Finally, further adjustment and optimization of the parameters of the entire networks are made after training the RBM layers.

In the current study, a 5-layer DBN including 1 input layer, 3 hidden layers and 1 output layer was implemented. Before employing the DBN model, the learning rate of each RBM hidden layer was set to 0.01, the maximum number of iterations was 500. When training the DBN model structure, the hidden layers were set to RBM1, RBM2 and RBM3 layers and the number of hidden neurons in each RBM was set to 10–100, the number of hidden neurons was selected layer by layer through a trial-and-error method, and the prediction effects of different network structures were compared. The DBN was executed by the R software under the 'darch' package.

2.4. Uncertainty Evaluation (Quantile Regression)

In this study, the predictive uncertainty of the proposed models was evaluated using the QR method. QR, developed by Koenker and Bassettto, is a linear method used to estimate the quantiles of a response variable without the need for a prescribed hypothesis [58,59]. Although being proposed more than 50 years, QR's ability in evaluating the robustness of the model remains strong and timeless [60]. It has been found that using QR for estimating forecast errors conditional on the forecasted water levels provides a relatively simple, efficient and robust means for estimation of predictive uncertainty [61].

For each quantile τ , the relationship between the observed (y) and predicted (\hat{y}) data can be expressed as follows:

$$v = \alpha_{\tau} \hat{y} + b_{\tau} \tag{10}$$

where α_{τ} and b_{τ} are the slope and intercept of equation (10), respectively, estimated by minimizing the sum of residuals as follows:

$$\min \sum_{i=1}^{N} \rho_{\tau (y_i - (\alpha_{\tau} \hat{y}_i))}$$
(11)

where y_i and \hat{y}_i are i-th samples from a dataset and ρ_{τ} is the QR function of the τ -th quantile as follows:

$$\rho_{\tau}(\epsilon_{i}) = \begin{cases} (\tau - 1) \cdot \epsilon_{i} \, \epsilon_{i} < 0\\ \tau \cdot \epsilon_{i} \quad \epsilon_{i} \ge 0 \end{cases}$$
(12)

The uncertainties generated using machine learning models were estimated as a set of forecasted quantiles for each of the desired quantile levels (i.e., p = 0.05 and 0.95). This study used the 'quantreg' package in R software to perform this QR analysis.

2.5. Materials

2.5.1. Study Area

The study area is situated inside the Jiuquan basin of Gansu province, northwest China, and refers to the oasis area around the Jiuquan city (Figure 2). The study area has long been an important commodity grain production base and the famous concentrated industrial crops production area in the Hexi Corridor, more than 2/3 commodity grain, almost all cotton, 9/10 sugar beet, more than 2/5 oil, beer barley, melon, fruit and vegetables are provided for the whole province. Moreover, being a significant node in the Gansu section of the Silk Road Economic Belt, the unique advantages in location, transportation, cultural and natural resources make the Jiuquan basin superior to promote multi-level and wide-field exchanges and cooperation. However, limited to the rare precipitation and surface water resources, groundwater-based oasis agriculture is mainly developed.



Therefore, the accurate forecast of groundwater level plays an essential role in agricultural management and groundwater resource scheduling.

Figure 2. Location of the Jiuquan basin, the study area, the groundwater level observation wells and the meteorological station.

In terms of the geology condition, the Jiuquan basin is surrounded by the Jiayuguan fault and the Wenshushan uplift in the west, as well as the Gaotai uplift in the east. Due to the influence of multiple tectonic movements, the Qilian fold belt in the southern Jiuquan Basin and the Gobi plains are separated by the Geermo Fault, Langweishan-Niutoushan Fault and the deep fault at the southern edge of the Hexi Corridor. Moreover, due to the resistance of the Cretaceous–Tertiary strata in the southern basin, the mountainous area and the Gobi plains have become two independent groundwater system units [62]. The large fault in front of the mountain objectively hinders the supply of fissure water in the bedrocks of the mountainous area to the Gobi plains [63]. Therefore, the basin presents a complete hydrogeological unit of the recharge, runoff and discharge processes ranging from the Gobi plains to the fine soil plain region and transitioning from a single thick layer of sand and gravel aquifer to a double-layer or a multi-layer fine-grain aquifer. The aquifer has a thickness of 50–400 m filled with a large volume of unconsolidated Quaternary sediments. The groundwater becomes shallower from the southwest to the northeast end with the groundwater depth of more than 100 m in the southern piedmont belt, 20–40 m in the middle basin and less than 5–10 m at the outlet of the Beidahe River. According to Wan et al. (2017), the unconfined groundwater in this area belongs to a snow and ice melt water-groundwater system, while the confined groundwater was recharged in the late Pleistocene and middle Holocene and is non-renewable [64]. Ice-snow melted water and precipitation are the main sources of the shallow and deep phreatic groundwater [65].

The climate of this area is arid continental with an average annual temperature range of about 5.8–10 °C. The mean annual precipitation is 117.5 mm, whereas the annual average potential evaporation is 2148.8 mm. Harsh natural conditions of rare precipitation and powerful evaporation make the area extremely short of water. Moreover, in recent decades, water demand has increased dramatically, groundwater has been exploited due to urbanization, agriculture and economic development [66]. There is no doubt that the overuse of the surface water and overexploitation of the groundwater would affect the water balance situation, leading to a serious decline in groundwater level and a series of complicated environmental problems (e.g., desertification, frequent sandstorms and ecological degradation). Thus, reliable groundwater level forecasting results are of great significance for the sustainable development of this region.

2.5.2. Dataset

To assess the validity of the hybrid CEEMDAN-GA-DBN model, the monthly GWL records of three observation wells (i.e., Well I, Well II and Well III) in the Jiuquan Basin were considered in this study (Figure 2). Due to data availability, monthly GWL data (m above the sea level) from January 2000 to December 2015 were applied from the Water Authority of the Jiuquan City. Moreover, precipitation and temperature are the most frequently used parameters in GWL prediction problems [18]. Hence, the nearby meteorological data were applied as model inputs, providing a more complete dataset (i.e., GWL; precipitation, Pre; temperature, Tem). The measured daily Pre and Tem were obtained from the Chinese Meteorological Administration with a time range of January 2000–December 2015. Since the purpose of this study is to predict the groundwater at the 1, 2 and 3 monthly time steps, the daily Pre was added to generate monthly equivalent ones, while the daily Tem was averaged to acquire monthly average value beforehand.

To escape from overfitting problem, the cross-validation method was performed in this study. All the data were divided into the following two distinct sets: the training set (January 2000 to December 2010) and the testing set (January 2011 to December 2015). Table 2 describes the hydrological statistics of the monthly GWL (m above the sea level), total precipitation per month (mm) and mean monthly temperature (°C) used in the training, testing and total sets. According to the statistical characteristics, no statistically significant difference existed in the training and testing datasets, indicating that the model features within the two sets were generally similar. The training set captured enough reliable information about the hydrological system, and therefore, can be used to train the predictive model. Moreover, the inputs and target time series were normalized between (0, 1) before model design.

$$X_{\rm norm} = \frac{X - X_{\rm min}}{X_{\rm max} - X_{\rm min}} \tag{13}$$

where X_{norm} is the normalized data; and X_{min} and X_{max} are the minimum and maximum values of the data, respectively.

		Total Precipitation Per Month	Mean Monthly Temperature		Monthly GWL	
		Pre (mm)	Tem (°C)	Well I (m)	Well II (m)	Well III (m)
	All	0	-14.11	1440.03	1436.22	1328.63
Minimum	Training	0	-14.11	1440.12	1436.39	1328.63
	Testing	0	-13.91	1440.03	1436.22	1328.78
	All	79.50	24.38	1442.63	1441.28	1329.93
Maximum	Training	52.40	24.38	1442.36	1441.28	1329.93
	Testing	79.50	23.10	1442.63	1440.78	1329.82
	All	7.75	8.20	1441.36	1438.74	1329.33
Mean	Training	7.77	8.19	1441.30	1438.79	1329.30
	Testing	7.70	8.21	1441.50	1438.64	1329.38
	All	11.05	11.18	0.54	1.08	0.30
Standard	Training	10.01	11.21	0.46	1.07	0.31
Deviation	Testing	13.15	11.21	0.66	1.10	0.28
	All	2.83	-0.24	-0.13	-0.08	-0.22
Skewness	Training	2.19	-0.23	-0.38	-0.04	-0.08
	Testing	3.39	-0.27	-0.28	-0.15	-0.55

Table 2. The statistical parameters of groundwater level (GWL) and related climatic data for three GWL observation wells in the whole dataset (January 2000 to December 2015), the training dataset (January 2000 to December 2010) and the testing dataset (January 2011 to December 2015).

2.5.3. Model Structure

In this study, the monthly GWL, total precipitation and average temperature were taken as potential inputs for the hybrid CEEMDAN-GA-DBN model, the hybrid CEEMDAN-DBN model and the standalone DBN model. Considering the general difficulties in obtaining GWL series in scarce data areas, in the current research, the maximum time delay was set to 3. This means that the GWL, total precipitation and average temperature at current month (t) as well as the antecedent months, which lagged 2 months behind (or t – 2) and 1 month behind (or t – 1), were incorporated as model input for the forecasting of GWL at the t + 1, t + 2 and t + 3 timescales. Consequently, the input structure of the hybrid CEEMDAN-GA-DBN, hybrid CEEMDAN-DBN and the standalone DBN models was set to

$$GWL(t+n) = f\{GWL_{t-2}, GWL_{t-1}, GWL_{t}, Pre_{t-2}, Pre_{t-1}, Pre_{t}, Tem_{t-2}, Tem_{t-1}, Tem_{t}\}$$
(14)

where GWL(t + n) was the forecasted groundwater level and the time-series of GWL_{t-2} , GWL_{t-1} , GWL_t , Pre_{t-2} , Pre_{t-1} , Pre_t , Tem_{t-2} , Tem_{t-1} and Tem_t were the model's input variables.

2.5.4. Model Performance Evaluation

To ascertain the preciseness of the objective model (i.e., the hybrid CEEMDAN-GA-DBN) against the CEEMDAN- DBN and the standalone DBN models, multiple criteria, such as correlation coefficient (R), mean absolute error (MAE), root mean squared error (RMSE), Nash–Sutcliffe efficiency coefficient (NSE), the ratio of the RMSE to the standard deviation of the observed GWL (RSR) and the Legates and McCabe's Index were applied. As a widely used metric, the R value presents the degree of linearity between the forecasted and the observed variables, and the RMSE measures the global fitness of the predictive models. In contrast, the MAE provides a more balanced perspective of the goodness-of-fit based on the estimation errors. The NSE adopts a range of $-\infty$ and 1, indicating the error between the simulated and the mean observation values. The optimal value of RSR is 0, indicating the RMSE or residual variation of the model is 0. The Legates and McCabe's Index (I), ranging from 0 to 1, is strongly recommended as supplementary model evaluation tool [67]. The following equations were used to evaluate these metrics:

$$R = \frac{\sum_{i=1}^{n} (GWL_{i}^{o} - \overline{GWL^{o}})(GWL_{i}^{p} - \overline{GWL^{p}})}{\sqrt{\sum_{i=1}^{n} (GWL_{i}^{o} - \overline{GWL^{o}})^{2} \sum_{i=1}^{n} (GWL_{i}^{p} - \overline{GWL^{p}})^{2}}}$$
(15)

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |GWL_i^p - GWL_i^o|$$
(16)

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(GWL_{i}^{p} - GWL_{i}^{o} \right)^{2}}$$
(17)

$$NSE = 1 - \frac{\sum_{i=1}^{n} (GWL_{i}^{o} - GWL_{i}^{p})^{2}}{\sum_{i=1}^{n} (GWL_{i}^{o} - \overline{GWL^{o}})^{2}}$$
(18)

$$RSR = \frac{RMSE}{\sigma_0}$$
(19)

$$I = 1.0 - \frac{\sum_{i=1}^{n} \left| GWL_{i}^{o} - GWL_{i}^{p} \right|}{\sum_{i=1}^{n} \left| GWL_{i}^{o} - \overline{GWL^{o'}} \right|}$$
(20)

where n = the number of input samples; GWL_i^o and GWL_i^p = the observed and simulated GWL at time t, respectively; and $\overline{GWL^o}$ and $\overline{GWL^p}$ = the average of the observed and forecasted GWL, respectively.

For accurately simulated GWL, $R \approx 1$, MAE/RMSE/RSR ≈ 0 and NSE ≈ 1 . In accordance with the relative literature, a model is normally considered very good if NSE

exceeds 0.75 and RSR exceeds 0.50, good if NSE is greater than 0.65 and RSR is less than 0.6 and satisfactory if NSE is greater than 0.5 and RSR is less than 0.70 [68].

For uncertainty evaluation, this study employed the mean prediction interval (MPI) and prediction interval coverage probability (PICP), which are both the typical metrics frequently used to evaluate the uncertainty of the data driven models [69]. The PICP, a measure of forecast reliability, is the probability that the observed values of an input pattern lie within the prediction limits; while the MPI, a measure of forecast resolution, monitors the ability to enclose observed values inside the prediction intervals [70]. Practically, a model with lower MPI and higher PICP (closer to $(1 - \alpha)$ %) is regarded as a better model [71]. Mathematically, the MPI and PICP are described as follows:

$$MPI = \frac{1}{n} \sum_{t=1}^{n} \left(PL_t^u - PL_t^l \right)$$
(21)

$$PICP = \frac{1}{n}count t$$
 (22)

where

$$t: PL_t^1 \le y_i \le PL_t^u \tag{23}$$

where y_i is the observed value; and PL_t^l and PL_t^u are the lower and upper limits, respectively.

When performing data driven models in GWL forecasting, it is important to compare the models not only in accuracy, but also in complexity since a complex model is not practical. The Akaike information criterion (AIC), a standard to measure the goodness of fit of a statistical model, is able to weigh the complexity of the estimated models [72]. The lower the AIC index value is, the better model would be.

$$AIC = \frac{2kn + (nln(\sigma_{\varepsilon}^2)(n-k-1))}{n-k-1}$$
(24)

where *k* is the number of parameters, *n* is the number of samples, and σ_{ε}^2 is the standard deviation of the residual.

3. Results and Discussion

In this section, a comprehensive evaluation among the objective model (i.e., CEEMDAN-GA-DBN), the hybrid CEEMDAN-DBN and the standalone DBN models for 1-, 2- and 3-month ahead GWL forecasting is analyzed and discussed. The performance metrics of the CEEMDAN-GA-DBN, the CEEMDAN-DBN and the DBN models of Wells I, II and III (Figure 2) in the training and the testing periods are enumerated in Tables 3–5. Considering the GWL forecasting purpose in this study, the performance of the proposed models in the testing period is mainly focused.

Table 3. Performance metrics of the hybridized CEEMDAN-GA-DBN, CEEMDAN-DBN and DBN models in the training and testing phases for 1-, 2- and 3-month ahead GWL forecasting of Well I.

		Training Phase					Testing Phase			
	R	MAE (m)	RMSE (m)	NSE	RSR	R	MAE (m)	RMSE (m)	NSE	RSR
			C	EEMDAN	J-GA-DBI	N				
1-month ahead	0.967	0.095	0.117	0.935	0.254	0.930	0.210	0.277	0.819	0.421
2-month ahead	0.940	0.123	0.157	0.883	0.341	0.904	0.225	0.315	0.759	0.486
3-month ahead	0.883	0.204	0.253	0.700	0.545	0.865	0.247	0.333	0.732	0.513
				CEEMDA	AN-DBN					
1-month ahead	0.966	0.096	0.119	0.932	0.259	0.923	0.241	0.313	0.770	0.476
2-month ahead	0.941	0.118	0.155	0.886	0.337	0.891	0.248	0.333	0.732	0.513
3-month ahead	0.917	0.131	0.169	0.867	0.364	0.853	0.261	0.363	0.680	0.560
				DE	BN					
1-month ahead	0.920	0.147	0.180	0.845	0.392	0.881	0.259	0.341	0.727	0.518
2-month ahead	0.897	0.163	0.203	0.804	0.441	0.815	0.290	0.394	0.624	0.607
3-month ahead	0.874	0.180	0.225	0.764	0.484	0.839	0.304	0.396	0.621	0.610

Note: R = correlation coefficient between observed and forecasted GWL, MAE = mean absolute error, RMSE = root mean square error, NSE = Nash Sutcliffe Coefficient and RSR = ratio of RMSE to the standard deviation.

		Training Phase					Testing Phase			
	R	MAE (m)	RMSE (m)	NSE	RSR	R	MAE (m)	RMSE (m)	NSE	RSR
			С	EEMDAN	J-GA-DBI	V				
1-month ahead	0.951	0.238	0.291	0.920	0.283	0.968	0.254	0.291	0.929	0.264
2-month ahead	0.942	0.268	0.338	0.887	0.334	0.941	0.375	0.443	0.840	0.397
3-month ahead	0.942	0.269	0.337	0.887	0.334	0.906	0.412	0.495	0.801	0.442
				CEEMDA	AN-DBN					
1-month ahead	0.957	0.240	0.299	0.915	0.291	0.954	0.295	0.365	0.889	0.330
2-month ahead	0.941	0.267	0.343	0.884	0.339	0.942	0.413	0.509	0.789	0.456
3-month ahead	0.944	0.258	0.331	0.891	0.328	0.895	0.428	0.515	0.786	0.459
				DE	BN					
1-month ahead	0.915	0.336	0.415	0.836	0.404	0.932	0.345	0.414	0.857	0.375
2-month ahead	0.856	0.416	0.521	0.733	0.515	0.876	0.458	0.575	0.730	0.515
3-month ahead	0.833	0.458	0.556	0.694	0.551	0.835	0.518	0.656	0.652	0.585

Table 4. Performance metrics of the hybridized CEEMDAN-GA-DBN, CEEMDAN-DBN and DBN models in the training and testing phases for 1-, 2- and 3-month ahead GWL forecasting of Well II.

Note: *R* = correlation coefficient between observed and forecasted GWL, *MAE* = mean absolute error, *RMSE* = root mean square error, *NSE* = Nash Sutcliffe Coefficient and *RSR* = ratio of *RMSE* to the standard deviation.

Table 5. Performance metrics of the hybridized CEEMDAN-GA-DBN, CEEMDAN-DBN and DBN models in the training and testing phases for 1-, 2- and 3-month ahead GWL forecasting of Well III.

		Training Phase				Testing Phase				
	R	MAE (m)	RMSE (m)	NSE	RSR	R	MAE (m)	RMSE (m)	NSE	RSR
			С	EEMDAN	J-GA-DBN	J				
1-month ahead	0.909	0.098	0.131	0.827	0.415	0.843	0.110	0.157	0.688	0.554
2-month ahead	0.904	0.105	0.135	0.816	0.428	0.810	0.129	0.169	0.630	0.602
3-month ahead	0.872	0.122	0.153	0.757	0.491	0.760	0.134	0.178	0.570	0.650
				CEEMDA	AN-DBN					
1-month ahead	0.911	0.096	0.130	0.831	0.410	0.808	0.123	0.175	0.614	0.616
2-month ahead	0.910	0.101	0.130	0.828	0.414	0.798	0.138	0.180	0.583	0.640
3-month ahead	0.897	0.107	0.139	0.799	0.447	0.749	0.149	0.185	0.534	0.677
				DE	BN					
1-month ahead	0.880	0.111	0.150	0.772	0.475	0.778	0.133	0.180	0.588	0.636
2-month ahead	0.908	0.096	0.132	0.823	0.419	0.808	0.144	0.182	0.573	0.648
3-month ahead	0.893	0.104	0.139	0.798	0.448	0.820	0.152	0.191	0.507	0.696

Note: *R* = correlation coefficient between observed and forecasted GWL, *MAE* = mean absolute error, *RMSE* = root mean square error, *NSE* = Nash Sutcliffe Coefficient and *RSR* = ratio of *RMSE* to the standard deviation.

3.1. Performance of the Hybrid CEEMDAN-GA-DBN Model

For 1-month ahead GWL forecasting, the hybrid CEEMDAN-GA-DBN model obtained very good performances at Well I, Well II and a good performance at Well III following the NSE and RSR values. Evidently, the hybrid CEEMDAN-GA-DBN model attained the R, MAE, RMSE, NSE and RSR values of 0.930, 0.210 m, 0.277 m, 0.819 and 0.421 for Well I, respectively, and 0.968, 0.254 m, 0.291 m, 0.929 and 0.264 for Well II, respectively. For Well III, the R, MAE, RMSE, NSE and RSR values for all the three GWL observation wells were significantly higher than 0.80 for the 1-month ahead forecasting, while RMSE and MAE were relatively small. That is to say, the hybrid CEEMDAN-GA-DBN model developed in the present study was able to provide very accurate and reliable forecasts for 1-month ahead GWL forecasting.

For the prediction of the 2- and 3-month ahead, the evaluation metrics in the testing phase revealed that the performance of the hybrid CEEMDAN-GA-DBN model was worse at the 1-month ahead timescale. Hence, there was a notable reduction in the magnitudes of R and NSE, and a consequent increase in the magnitudes of errors (i.e., RSR, MAE and RMSE) for Well I, Well II and Well III. The result, demonstrating an inferior performance of the proposed model at the 2- and 3-month ahead timescales, was consistent with other studies [38,39]. However, the magnitude of the performance metrics remained within a high predictive accuracy range. Specifically, taking Well I for example, the hybrid model (CEEMDAN-GA-DBN) registered a value at the 2-month ahead forecast of R, MAE, RMSE, NSE and RSR equivalent to 0.904, 0.225 m, 0.315 m, 0.735 and 0.513 for the 3-month ahead the acquired values were 0.865, 0.247 m, 0.333 m, 0.735 and 0.513 for the 3-month ahead

GWL forecast, respectively. Importantly, the prediction results showed that the hybrid CEEMDAN-GA-DBN model obtained relatively good performance for the 2-month ahead GWL prediction and acceptable results for the 3-month ahead GWL prediction for the case of Well I. Meanwhile, the present study also found that the R values in the testing period were greater than 0.90, the MAE and RMSE remained significantly low, the NSE values exceeded 0.80 and the RSR values were less than 0.50 for the case of Well II. For the case of Well III, the simulation results showed that the hybrid CEEMDAN-GA-DBN method obtained acceptable performance for 2- and 3-month ahead GWL prediction with the NSE values being greater than 0.50 and the RSR values being less than 0.70.

Despite the slightly deteriorated performance at the longer lead time, the R values in the testing period were greater than 0.75 for all the wells. The results stated clearly that the predicted GWL values fitted the target data quite closely. Indeed, the low MAE and RMSE values and the high NSE and RSR values demonstrated the high-quality forecast. It should be noted that the performance of the hybrid CEEMDAN-GA-DBN model was in agreement with that of the previous researchers who utilized machine learning models to predict the hydrological systems at multiple timescales [71,73]. In those studies, the predictive capacity of their models also deteriorated with an increase in the forecasting timescale, which may be attributed to the reduction in data patterns at longer time steps [74]. Despite the deteriorated accuracy of 2- and 3- month ahead forecasting, statistical metrics for the hybrid CEEMDAN-GA-DBN model in the testing phase satisfied the requirements of good GWL simulation. Therefore, it is certain that the hybrid CEEMDAN-GA-DBN model achieved good forecasts for long-term GWL simulation.

3.2. Comparison of the CEEMDAN-GA-DBN Model, the CEEMDAN-DBN Model and the DBN Model

When comparing the capability of the hybrid CEEMDAN-GA-DBN, the CEEMDAN-DBN and the DBN models for 1-, 2- and 3- month ahead GWL forecasting, it can be seen that the proposed CEEMDAN-GA-DBN model outperformed the hybrid CEEMDAN-DBN and DBN models not only for the 1-month ahead GWL predicting, but also for the 2- and 3-month forecasting horizons (see Tables 3–5).

Taking Well I for instance, for 1-month ahead forecasting, the hybrid CEEMDAN-DBN and the standalone DBN model yielded R, MAE, RMSE, NSE and RSR values of 0.923, 0.214 m, 0.313 m, 0.770, 0.476 and 0.881, 0.259 m, 0.341 m, 0.727, 0.518, respectively. By contrast, the hybrid CEEMDAN-GA-DBN model attained R and NSE values of 0.930 and 0.819, while the MAE, RMSE and RSR were reduced to 0.210 m, 0.277 m and 0.421, respectively. The CEEMDAN-GA-DBN model reduced 11.50% of RMSE less than the CEEMDAN-DBN model, and 18.77% less than the DBN model; while it improved NSE by 6.36% more than the CEEMDAN-DBN model, and 12.65% more than the DBN model. That is, the hybrid CEEMDAN-GA-DBN predictive model developed in this study provided more accurate and reliable 1-month ahead GWL forecasting compared with the CEEMDAN-DBN and DBN models (see Tables 3–5). Although the accuracy of the 2- and 3-month ahead predicted GWL values was worse than that of the 1-month forecasts, the present results obtained using the hybrid CEEMDAN-GA-DBN model were comparatively better than the other two models. For example, compared with that of the CEEMDAN-DBN (0.363 m) and DBN (0.396 m) models, the RMSE of the CEEMDAN-GA-DBN decreased by 8.26 and 15.91%, respectively, for 3-month ahead forecasting. Overall, the CEEMDAN-GA-DBN reduced the RMSE of the CEEMDAN-DBN and DBN models in the testing period about 9.16 and 17.63%, while it improved their NSE by about 6.38 and 15.32%, respectively, for all the lead times and the three wells.

The hybrid CEEMDAN-GA-DBN model can also be considered good since the NSE value was 0.840 and the RSR value was 0.397 for the 2-month ahead forecasting at Well II. In contrast, the hybrid CEEMDAN-DBN and DBN models were inferior to the CEEMDAN-GA-DBN. The NSE values reduced to 0.789 and 0.730, while the RSR values increased to 0.456 and 0.515 for the 2-month ahead forecasting of Well II. The results demonstrated the better performance that the hybrid CEEMDAN-GA-DBN model achieved. Similarly, the hybrid CEEMDAN-GA-DBN model yielded more accurate results than the hybrid CEEMDAN-DBN and DBN model for the 3 month-ahead GWL simulation. Therefore, the simulation results reveal that the hybrid CEEMDAN-GA-DBN method can significantly improve performance relative to the hybrid CEEMDAN-DBN and the standalone DBN model in terms of the 2- and 3-month ahead GWL estimations.

While the performance metrics can statistically evaluate the ability of the CEEMDAN-GA-DBN, the CEEMDAN-DBN and the DBN models, the hydrograph and scatter plots are capable of assessing and displaying the temporal correspondence of the observed GWL data and the predicted values. Figures 3-8 show the hydrographs and scatter plots of the observed and predicted GWL at the 1-, 2- and 3-month ahead times for the CEEMDAN-GA-DBN, CEEMDAN-DBN and DBN models in the testing period. For interpretation purposes, the least square regression line, y = ax + b was incorporated in each panel. The constant a and intercept b were used to assess the model's overall accuracy. Clearly, the CEEMDAN-GA-DBN model performed far better for 1-month lead GWL forecasting than the 2- and 3-month lead time. The performance for a longer lead time deteriorated gradually (Figures 3–8). Furthermore, it is apparent that the GWL forecasted using the CEEMDAN-GA-DBN model closely matched with the corresponding observed values. Figures 3–8 also show the hydrographs and scatter plots of the observed and estimated GWL values generated using the CEEMDAN-DBN and DBN models in the testing period. Both the CEEMDAN-DBN and DBN models show relatively good accuracy for the 1-, 2and 3-month ahead GWL forecasting. Moreover, it can be seen from the fitted equations that the CEEMDAN-GA-DBN model yielded a and b closer to 1 and 0 when compared with that of the CEEMDAN-DBN and DBN models in most cases. The results demonstrated the significantly better fit the CEEMDAN-GA-DBN model achieved. These figures confirmed that the CEEMDAN-GA-DBN model has a better generalization skill of the predictive data compared with the CEEMDAN-DBN and DBN models considered in the current study. Thus, the best accuracy in GWL forecasting was achieved using the CEEMDAN-GA-DBN model.



Figure 3. Cont.



Figure 3. Observed vs. predicted groundwater level for Well I generated by the CEEMDAN-GA-DBN, CEEMDAN-DBN and DBN models for 1-, 2- and 3-month ahead forecast in the testing period.



Figure 4. Cont.



Figure 4. Scatter plot of the observed and predicted groundwater level for Well I generated by the CEEMDAN-GA-DBN, CEEMDAN-DBN and DBN models for 1-, 2- and 3-month ahead forecast in the testing period.



Figure 5. Observed vs. predicted groundwater level for Well II generated by the CEEMDAN-GA-DBN, CEEMDAN-DBN and DBN models for 1-, 2- and 3-month ahead forecast in the testing period.



Figure 6. Scatter plot of the observed and predicted groundwater level for Well II generated by the CEEMDAN-GA-DBN, CEEMDAN-DBN and DBN models for 1-, 2- and 3-month ahead forecast in the testing period.



Figure 7. Cont.



Figure 7. Observed vs. predicted groundwater level of Well III generated by the CEEMDAN-GA-DBN, CEEMDAN-DBN and DBN models for 1-, 2- and 3-month ahead forecast in the testing period.



Figure 8. Cont.



Figure 8. Observed vs. predicted groundwater level of Well III generated by the CEEMDAN-GA-DBN, CEEMDAN-DBN and DBN models for 1-, 2- and 3-month ahead forecast in the testing period.

When evaluating the feasibility of a predictive model in GWL prediction, the distribution of error can show a robust and reliable consequence of the model's predictability. Figure 9 demonstrates the error indicators for the hybrid CEEMDAN-GA-DBN, CEEMDAN-DBN and DBN models. The boxplot diagram shows the degree of the dispersion and skewness of the error for all the wells. Clearly, the figure illustrates that the CEEMDAN-GA-DBN technique obtained the minimum error compared with the CEEMDAN-DBN and DBN models in most instances. This is especially true for the 1-month ahead GWL prediction and the 2- and 3-month ahead prediction values. Therefore, the current results confirm that the CEEMDAN-GA-DBN model is superior to the CEEMDAN-DBN and DBN models in terms of error distributions.



Figure 9. Boxplots of the predicted error for the 1-, 2- and 3-month ahead groundwater level forecast generated by the CEEMDAN-GA-DBN, CEEMDAN-DBN and DBN models in the testing period.

A visualized comparison of the distribution of the observed and forecasted values for a forecasting model is an effective way to verify and validate a model. Considering that the boxplot can present a clear visualization of the data distribution concerning the quartiles distinctly indicating the outliers. The boxplots diagrammatizing the distribution of the observed and forecasted GWL values from the CEEMDAN-GA-DBN, CEEMDAN-DBN and DBN models are shown in Figure 10. It can be seen that the box areas for all three models are very close to the observed values. Comparatively speaking, the distribution of the CEEMDAN-GA-DBN-forecasted GWL approaches the observed GWL values than the CEEMDAN-DBN and DBN models under most circumstances. Consequently, it is further ascertained that the CEEMDAN-GA-DBN model is expected to generate forecasted GWL values that closely resemble the measured values.



Figure 10. Boxplots of the distribution for the observed and simulated groundwater level forecast for the 1-, 2- and 3-month ahead groundwater level generated by the CEEMDAN-GA-DBN, CEEMDAN-DBN and DBN models in the testing period.

To determine the accuracy of the hybrid CEEMDAN-GA-DBN model, the Legates and McCabe's Index was imported. Figure 11 plots the Legates and McCabe's Index computed for all the simulated and observed GWLs in the testing phase. It is noteworthy that for a perfect model, the index must approach one. There is no doubt that the hybrid deep learning model (i.e., CEEMDAN-GA-DBN) produces significantly better forecasts, evidenced by the larger Legates and McCabe's Index. This is important for all the forecasting horizons, ranging from t + 1 to t + 3, although the accuracy deteriorates as the forecast horizon increases. Nonetheless, the performance of the CEEMDAN-GA-DBN model, where the data decomposition was used together with the feature selection, far exceeds that of the CEEMDAN-DBN and the standalone DBN model.



Figure 11. Cont.



Figure 11. The Legates and McCabe's Index in testing phase of the hybrid CEEMDAN-GA-DBN vs. CEEMDAN-DBN and DBN models for (**a**) Well I, (**b**) Well II and (**c**) Well III.

In terms of model simplicity, the AIC index was used to compare the performance of the proposed methods. Table 6 shows the AIC values of the CEEMDAN-GA-DBN, CEEMDAN-DBN, DBN models for 1-, 2- and 3-month ahead GWL forecasting at three GWL wells. It can be clearly seen that the CEEMDAN-GA-DBN obtained the lowest AIC values, while the DBN model derived the highest AICs. Taking the results of Well I as an example, the DBN model obtained AIC values of -114.275, -98.281 and -97.732, respectively at three forecasting horizons; however, the AIC values of the CEEMDAN-DBN model decreased by 8.12, 18.93 and 9.62%; for the CEEMDAN-GA-DBN model, the AIC values reduced by 19.81, 24.88 and 19.50%. The results indicate that the data decomposition and feature selection processes improved the accuracy of the model rather than increasing model complexity. Therefore, combining the above accuracy and complexity analysis results of the proposed models, it can be found that although the CEEMDAN-GA-DBN model had two more steps (data decomposition and feature selection) than the other models, it is able to achieve a simpler and more accurate model.

	CEEMDAN-GA-DBN	CEEMDAN-DBN	DBN
	Well I		
1-month ahead	-136.917	-123.559	-114.275
2-month ahead	-122.73	-116.889	-98.281
3-month ahead	-116.793	-107.13	-97.732
	Well II		
1-month ahead	-131.507	-106.671	-92.663
2-month ahead	-85.33	-70.112	-56.719
3-month ahead	-73.025	-68.82	-42.152
	Well III		
1-month ahead	-199.541	-187.797	-184.252
2-month ahead	-191.245	-184.587	-183.257
3-month ahead	-185.635	-181.233	-178.13

Table 6. AIC values of the CEEMDAN-GA-DBN, CEEMDAN- DBN and DBN models for the 1-, 2- and 3-month ahead GWL forecasts at Well I, Well II, Well III.

3.3. Uncertainty Analysis

Typically, higher GWL prediction accuracy means less uncertainty and smoother fluctuations, which is significant for the utilization and planning of groundwater resources. Therefore, analyzing the uncertainty of the proposed models in GWL simulation is of great significance. In this study, the predictive uncertainties of the CEEMDAN-GA-DBN, CEEMDAN-DBN and DBN models were assessed using the QR, the uncertainty information of GWL predictions was estimated at 90 and 95% confidence levels.

Tables 7 and 8 present the MPI and PICP values of different wells and different forecasting horizons. Note that perfect reliability is shown as the PICP value equals the corresponding confidence level. If similar PICP scores are derived, then the one with lower MPI is the better. The tables show that the uncertainty analysis results of the CEEMDAN-GA-DBN, CEEMDAN-DBN and DBN models are not identical in terms of the MPI and PICP values either at the confidence level of 90% or the confidence level of 95%. It seems very difficult to derive balanced low MPI as well as high PICP values since other researchers also encountered this problem [65,66]. Nevertheless, the PICP values of the proposed CEEMDAN-GA-DBN model are higher than that of the CEEMDAN-DBN and DBN models in most circumstances. Although higher PICP values remain higher than the CEEMDAN-GA-DBN model. Hence, the proposed CEEMDAN-GA-DBN method exhibits a slightly better forecasting reliability compared with the CEEMDAN-DBN and DBN models on the whole.

Table 7. Uncertainty analysis for the 1-, 2- and 3-month ahead GWL forecasts of the CEEMDAN-GA-DBN, CEEMDAN-
DBN and DBN models at 90% confidence level.

1 01117 6

.

	T 1 m'	CEEMDAN-GA-DBN		CEEMD	AN-DBN	DBN	
Observation well	Lead Time	MPI (m)	PICP (%)	MPI (m)	PICP (%)	MPI (m)	PICP (%)
	1	0.38	56.36	0.37	47.27	0.59	65.45
Well I	2	0.51	69.09	0.53	63.64	0.65	63.64
	3	0.74	76.36	0.51	62.82	0.68	56.36
	1	0.98	92.73	0.97	85.45	1.41	92.73
Well II	2	1.14	80.00	1.07	70.91	1.69	81.82
	3	1.08	70.90	1.19	69.09	1.81	76.36
	1	0.37	81.82	0.37	78.18	0.46	85.45
Well III	2	0.45	89.09	0.37	78.18	0.44	81.82
	3	0.47	83.64	0.47	80.00	0.44	80.00

Note: MPI = mean prediction interval; PICP = prediction interval coverage probability.

	T	CEEMDAN-GA-DBN		CEEMD	AN-DBN	DBN	
Observation well	Lead Time	MPI (m)	PICP (%)	MPI (m)	PICP (%)	MPI (m)	PICP (%)
	1	0.41	56.36	0.43	58.18	0.71	70.91
Well I	2	0.62	72.73	0.67	83.64	0.84	78.20
	3	0.81	81.82	0.69	72.73	1.00	81.82
	1	1.10	96.36	1.17	92.72	1.67	94.55
Well II	2	1.30	83.64	1.46	89.09	1.96	83.64
	3	1.24	74.55	1.31	76.36	2.07	83.64
	1	0.48	90.91	0.46	89.01	0.57	90.91
Well III	2	0.47	89.09	0.52	89.09	0.53	87.27
	3	0.61	90.91	0.56	89.27	0.61	90.91

Table 8. Uncertainty analysis for the 1-, 2- and 3-month ahead GWL forecasts of the CEEMDAN-GA-DBN, CEEMDAN-DBN and DBN models at 95% confidence level.

Note: MPI = mean prediction interval; *PICP* = prediction interval coverage probability.

3.4. Model Interpretation

From the viewpoint of the predictive performances and uncertainty criterion, it can be appropriately concluded that the proposed CEEMDAN-GA-DBN model provides a more robust and stable predictive performance in GWL forecasting for multiple forecasting horizons. The high precision of the hybrid CEEMDAN-GA-DBN model may lie in three aspects. Firstly, the deep belief network algorithm can discover the inherent physical structure and antecedent features powerfully in datasets without prior knowledge. Therefore, the highlevel nonlinear groundwater dynamic characteristics can be effectively captured through a deep learning process. Secondly, the intrinsic mode functions (IMFs) decomposed by the CEEMDAN algorithm express the non-stationary and uncertainty behaviors of GWL more clearly, making the repeating feature catching process more reliable and predictable. That is the reason for the much better prediction performance of the CEEMDAN-DBN model than the DBN model. Thirdly, the irrelevant and redundant attributes of the GWL were identified and removed from the IMFs using the genetic algorithm-based feature selection procedure, thus the accuracy and stability of the model was improved. However, the inconformity of the uncertainty analysis results among the three wells and the three models reveals that uncertainty is inevitable, even for those models with high accuracy. Typically, the uncertainty increases as the noise increases [65]. In this study, reasonable interpretations for this phenomenon may lie in the collective affect from every possible step of the methodological framework, the hydrogeological heterogeneity of the three distinguished GWL observation wells, the observed GWL data (the observation or the data entry processes) and the other sources.

However, when comparing the performance of the proposed models in training and testing phases, it can be seen that the performance in the former phase is much better than that in the testing phase in most cases, meaning there is a phenomenon of overfitting to some extent. Overfitting is an issue within machine learning based applications where a model learns the patterns of the training dataset too well, perfectly explaining the training set but failing to generalize this predictive power to other sets. Actually, in this study, the cross-validation procedure was applied to deal with this problem through dividing all the data into two distinct sets: the training set and the testing set, but still did not avoid it. Through a comprehensive literature review, it can be found that the phenomenon that the performance in the training period is superior to that in the testing period is common in machine learning/deep learning based GWL forecasting studies [75,76]. In this study, according to the derived results, although the MAE values in the testing phase were higher than that in the training phase, the MAE for Well I, Well II and Well III was less than 0.4, 0.7 and 0.2 m, respectively. Compared with the average GWL of 1141.36, 1138.73 and 1329.33 m of the three GWL wells, the error is acceptable. Moreover, the analysis in this study also proved that the proposed model achieved high accuracy since the magnitude of the other performance metrics remained within high predictive accuracy ranges in both the

training and testing periods. Considering this, it can be said that the overfitting problem only demonstrates the inferior generalization ability of the proposed models, rather than the accuracy or the application of the model.

In addition, it is noteworthy that in the irrigated areas of the Jiuquan basin (e.g., this study area), the dynamics of the phreatic groundwater are closely related to agricultural irrigation. The irrigation water infiltration has important significance to the replenishment of the groundwater in this area. That is mainly because the groundwater level drops when pumping during irrigation time, while it rises after irrigation. Overall, the variation of the GWL changes little over the years, thus no obvious upward or downward trend was found [77]. This means that the proposed model and derived results in this study remain useful in multi time steps ahead GWL forecasting since the GWL dynamics were captured using the CEEMDAN-GA-DBN model through the training process.

4. Conclusions

The accurate and reliable GWL prediction is extremely important for the planning and management of groundwater resources in arid irrigated regions. In this study, a novel hybrid forecasting framework (the CEEMDAN-GA-DBN model) was proposed for 1-, 2and 3-month ahead GWL prediction. The performance of the CEEMDAN-GA-DBN model was compared with the hybrid CEEMDAN-DBN model and the standalone DBN model by the performance metrics and uncertainty criterion.

The results show that the hybrid CEEMDAN-GA-DBN model is able to successfully forecast 1-, 2- and 3-month ahead GWL and outperform the CEEMDAN-DBN and DBN model in terms of the performance criteria including R, MAE, RMSE, NSE, RSR and the Legates and McCabe's Index. Moreover, based on the QR method, the proposed CEEMDAN-GA-DBN model is also very effective from uncertainties analysis. Therefore, it is certain that the CEEMDAN-GA-DBN model has a high potential for practical application in GWL prediction in arid irrigated areas. The CEEMDAN-GA-DBN model can be explored as a scientific tool applied for GWL forecast without understanding the intrinsic mechanisms and hydrogeological characteristics or collecting the condition data required for various interacting elements. Therefore, it is especially valuable for regions with limited measurement data to develop a physical based hydrogeological model. Yet, attention should be paid to the uncertainty in the process of model building.

The implementation of the CEEMDAN-GA-DBN proves the conspicuous results of the proposed framework, indicating a forecasting framework that, when implemented in practice in arid irrigated areas, can improve the GWL forecasting accuracy. Considering the extensive range of feature selection methods and deep learning models, in further studies, any other methods can be explored as alternative tools in their place if necessary (Figure 12).

Although the proposed deep learning model has achieved an excellent accuracy for GWL prediction, considerable room remains for further improvement. Firstly, due to a lack of data in recent years, the performance of the model was mainly focused from 2000 to 2015. Considering the fact that in recent years, water demand has increased dramatically, and groundwater has been excessively exploited with an annual groundwater extraction rate of 2.6 \times 10⁸ m³, whether the same high accuracy can be derived using the model remains uncertain and needs to be further studied. Secondly, a comprehensive assessment should be implemented to analyze the possible uncertainties that had influences on the performance of the models. For example, the spatial variation of the hydrogeological parameters should be considered in a future study to improve the model's accuracy. Thirdly, certain anthropic or natural factors in the specific area, such as the excessive groundwater extraction, the change of water users as well as the irrigated areas in this groundwater-supported agricultural region, the intricate and intense interaction between river and groundwater in this inland river basin, may affect the accuracy of the GWL forecast. Therefore, there stands a chance that the model's accuracy will be improved remarkably if these universal interfering factors and the controlling factors of different

wells are taken into consideration. Finally, the performance of the CEEMDAN-GA-DBN model for 1-, 2- and 3-month ahead GWL forecasting was mainly discussed in this study; however, as a designed simulation model, further investigation of its potentiality in both short- and long-term forecasting would be an interesting exploration.



Figure 12. Proposed methodological framework and possible alternative methods.

Author Contributions: Conceptualization, W.L. and X.W.; methodology, X.W. and H.Y.; software, H.Y. and Z.Y.; validation, W.L. and H.Y.; formal analysis, L.Y.; investigation, H.Y.; resources, W.L. and H.Y.; data curation, H.Y.; writing—original draft preparation, W.L. and X.W.; writing—review and editing, W.L. and M.Z.; visualization, H.Y. and X.W.; supervision, W.L.; project administration, W.L. and X.W.; funding acquisition, W.L. and X.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Fund of China (Grant No. 41771252, 42001035, 41801015), the Key Research Program of Frontier Sciences, CAS (QYZDJ-SSW-DQC031), the CAS 'Light of West China' Program, the Forestry and Grassland Science and Technology Innovation Program of Gansu Province (Grant No. GYCX [2020]01), the Key R&D Program of Gansu Province, China (Grant No. 20YF8FA002), the XPCC Science and Technique Foundation (Grant No. 2021AB021).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author on reasonal request.

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

References

- 1. Barua, S.; Cartwright, I.; Dresel, E.P.; Daly, E. Using multiple methods to investigate the effects of land-use changes on groundwater recharge in a semi-arid area. *Hydrol. Earth Syst. Sci.* 2021, 25, 89–104. [CrossRef]
- Khalil, M.M.; Tokunaga, T.; Heggy, E.; Abotalib, A.Z. Groundwater mixing in shallow aquifers stressed by land cover/land use changes under hyper-arid conditions. *J. Hydrol.* 2021, 598, 126245. [CrossRef]
- Herrera, C.; Godfrey, L.; Urrutia, J.; Custodio, E.; Jordan, T.; Jódar, J.; Delgado, K.; Barreneche, F. Recharge and residence times of groundwater in hyper arid areas: The confined aquifer of Calama, Loa River Basin, Atacama Desert, Chile. *Sci. Total. Environ.* 2021, 752, 141847. [CrossRef]

- Li, H.; Lu, Y.; Zheng, C.; Zhang, X.; Zhou, B.; Wu, J. Seasonal and inter-annual variability of groundwater and their responses to climate change and human activities in arid and desert areas: A case study in Yaoba Oasis, northwest China. *Water* 2020, *12*, 303. [CrossRef]
- 5. De Graaf, I.E.M.; Gleeson, T.; Rens van Beek, L.P.H.; Sutanudjaja, E.H.; Bierkens, M.F.P. Environmental flow limits to global groundwater pumping. *Nature* **2019**, *574*, 90–94. [CrossRef]
- Abu-Bakr, H.A.e.-A. Groundwater vulnerability assessment in different types of aquifers. *Agric. Water Manag.* 2020, 240, 106275. [CrossRef]
- 7. Amaranto, A.; Munoz-Arriola, F.; Solomatine, D.P.; Corzo, G. A spatially enhanced data-driven multimodel to improve semiseasonal groundwater forecasts in the high plains aquifer, USA. *Water Resour. Res.* **2019**, *55*, 5941–5961. [CrossRef]
- 8. Ghiglieri, G.; Buttau, C.; Arras, C.; Funedda, A.; Soler, A.; Barbieri, M.; Carrey, R.; Domènech, C.; Torrentó, C.; Otero, N.; et al. Using a multi-disciplinary approach to characterize groundwater systems in arid and semi-arid environments: The case of Biskra and Batna regions (NE Algeria). *Sci. Total Environ.* **2021**, 757, 143797. [CrossRef]
- 9. Fang, C.; Sun, S.; Jia, S.; Li, Y. Groundwater level analysis using regional Kendall Test for trend with spatial autocorrelation. *Groundwater* **2019**, *57*, 320–328. [CrossRef] [PubMed]
- Wang, S.; Liu, H.; Yu, Y.; Zhao, W.; Yang, Q.; Liu, J. Evaluation of groundwater sustainability in the arid Hexi Corridor of northwestern China, using GRACE, GLDAS and measured groundwater data products. *Sci. Total Environ.* 2020, 705, 135829. [CrossRef] [PubMed]
- 11. Hellwig, J.; de Graaf, I.E.M.; Weiler, M.; Stahl, K. Large-scale assessment of delayed groundwater responses to drought. *Water Resour. Res.* 2020, *56*, e2019WR025441. [CrossRef]
- 12. De Graaf, I.; Condon, L.; Maxwell, R. Hyper-resolution continental-scale 3-D aquifer parameterization for groundwater modeling. *Water Resour. Res.* 2020, *56*, 1–14. [CrossRef]
- 13. Deng, C.; Bailey, R.T. Assessing causes and identifying solutions for high groundwater levels in a highly managed irrigated region. *Agric. Water Manag.* **2020**, 240, 106329. [CrossRef]
- 14. Chang, F.J.; Chang, L.C.; Huang, C.W.; Kao, I.F. Prediction of monthly regional groundwater levels through hybrid soft-computing techniques. *J. Hydrol.* **2016**, *541*, 965–976. [CrossRef]
- 15. Marçais, J.; de Dreuzy, J.R. Prospective interest of deep learning for hydrological inference. *Groundwater* **2017**, *55*, 688–692. [CrossRef] [PubMed]
- 16. Rahman, A.T.M.S.; Hosono, T.; Quilty, J.M.; Das, J.; Basak, A. Multiscale groundwater level forecasting: Coupling new machine learning approaches with wavelet transforms. *Adv. Water Resour.* **2020**, *141*, 103595. [CrossRef]
- 17. Wei, Z.L.; Wang, D.F.; Sun, H.Y.; Yan, X. Comparison of a physical model and phenomenological model to forecast groundwater levels in a rainfall-induced deep-seated landslide. *J. Hydrol.* **2020**, *586*, 124894. [CrossRef]
- Rajaee, T.; Ebrahimi, H.; Nourani, V. A review of the artificial intelligence methods in groundwater level modeling. *J. Hydrol.* 2019, 572, 336–351. [CrossRef]
- 19. Mo, S.; Zabaras, N.; Shi, X.; Wu, J. Deep autoregressive neural networks for high-dimensional inverse problems in groundwater contaminant source identification. *Water Resour. Res.* **2019**, *55*, 3856–3881. [CrossRef]
- Shahabi, H.; Shirzadi, A.; Ronoud, S.; Asadi, S.; Pham, B.T.; Mansouripour, F.; Geertsema, M.; Clague, J.J.; Bui, D.T. Flash flood susceptibility mapping using a novel deep learning model based on deep belief network, back propagation and genetic algorithm. *Geosci. Front.* 2021, 12, 101100. [CrossRef]
- 21. Zhong, Z.; Sun, A.Y.; Yang, Q.; Ouyang, Q. A deep learning approach to anomaly detection in geological carbon sequestration sites using pressure measurements. *J. Hydrol.* **2019**, *573*, 885–894. [CrossRef]
- Zhang, J.; Zhu, Y.; Zhang, X.; Ye, M.; Yang, J. Developing a Long Short-Term Memory (LSTM) based model for predicting water table depth in agricultural areas. J. Hydrol. 2018, 561, 918–929. [CrossRef]
- 23. Gochoo, M.; Akhter, I.; Jalal, A.; Kim, K. Stochastic remote sensing event classification over adaptive posture estimation via multifused data and deep belief network. *Remote Sens.* **2021**, *13*, 912. [CrossRef]
- 24. Niu, G.Q.; Yi, X.H.; Chen, C.; Li, X.Y.; Han, D.H.; Yan, B.; Huang, M.Z.; Ying, G.G. A novel effluent quality predicting model based on genetic-deep belief network algorithm for cleaner production in a full-scale paper-making wastewater treatment. *J. Clean. Prod.* 2020, *265*, 121787. [CrossRef]
- 25. Rizk, Y.; Hajj, N.; Mitri, N.; Awad, M. Deep belief networks and cortical algorithms: A comparative study for supervised classification. *Appl. Comput. Inform.* **2019**, *15*, 81–93. [CrossRef]
- 26. Wen, X.; Feng, Q.; Yu, H.; Wu, J.; Si, J.; Chang, Z.; Xi, H. Wavelet and adaptive neuro-fuzzy inference system conjunction model for groundwater level predicting in a coastal aquifer. *Neural Comput. Appl.* **2015**, *26*, 1203–1215. [CrossRef]
- Wu, C.C.; Zhang, X.Q.; Wang, W.J.; Lu, C.P.; Zhang, Y.; Qin, W.; Tick, G.R.; Liu, B.; Shu, L.C. Groundwater level modeling framework by combining the wavelet transform with a long short-term memory data-driven model. *Sci. Total. Environ.* 2021, 783, 146948. [CrossRef]
- Yu, H.; Wen, X.; Feng, Q.; Deo, R.C.; Si, J.; Wu, M. Comparative study of hybrid-wavelet artificial intelligence models for monthly groundwater depth forecasting in extreme arid regions, northwest China. *Water Resour. Manag.* 2018, 32, 301–323. [CrossRef]
- 29. Zhang, J.; Zhang, X.; Niu, J.; Hu, B.X.; Soltanian, M.R.; Qiu, H.; Yang, L. Prediction of groundwater level in seashore reclaimed land using wavelet and artificial neural network-based hybrid model. *J. Hydrol.* **2019**, *577*, 123948. [CrossRef]

- 30. Meng, E.; Huang, S.; Huang, Q.; Fang, W.; Wu, L.; Wang, L. A robust method for non-stationary streamflow prediction based on improved EMD-SVM model. *J. Hydrol.* **2019**, *568*, 462–478. [CrossRef]
- 31. Qu, Z.; Mao, W.; Zhang, K.; Zhang, W.; Li, Z. Multi-step wind speed forecasting based on a hybrid decomposition technique and an improved back-propagation neural network. *Renew. Energy* **2019**, *133*, 919–929. [CrossRef]
- 32. Rezaie-Balf, M.; Kim, S.; Fallah, H.; Alaghmand, S. Daily river flow forecasting using ensemble empirical mode decomposition based heuristic regression models: Application on the perennial rivers in Iran and South Korea. *J. Hydrol.* **2019**, *572*, 470–485. [CrossRef]
- 33. Huang, G.; Li, X.; Zhang, B.; Ren, J. PM 2.5 concentration forecasting at surface monitoring sites using GRU neural network based on empirical mode decomposition. *Sci. Total Environ.* **2021**, *768*, 144516. [CrossRef]
- 34. Liu, H.; Cao, L.; Jia, J.; Gong, H.; Qi, X.; Xu, X. Effects of land use changes on the nonlinear trends of net primary productivity in arid and semiarid areas, China. *Land Degrad. Dev.* **2021**, *32*, 2183–2196. [CrossRef]
- 35. Huang, S.; Chang, J.; Huang, Q.; Chen, Y. Monthly streamflow prediction using modified EMD-based support vector machine. *J. Hydrol.* **2014**, *511*, 764–775. [CrossRef]
- 36. Zhang, Z.J.; Zhang, Y.M.; Lu, J.W.; Gao, F.; Xiao, G. A novel complex manufacturing business process decomposition approach in cloud manufacturing. *Comput. Ind. Eng.* 2020, 144, 106442. [CrossRef]
- Fijani, E.; Barzegar, R.; Deo, R.; Tziritis, E.; Konstantinos, S. Design and implementation of a hybrid model based on two-layer decomposition method coupled with extreme learning machines to support real-time environmental monitoring of water quality parameters. *Sci. Total Environ.* 2019, 648, 839–853. [CrossRef]
- 38. Prasad, R.; Deo, R.C.; Li, Y.; Maraseni, T. Weekly soil moisture forecasting with multivariate sequential, ensemble empirical mode decomposition and Boruta-random forest hybridizer algorithm approach. *Catena* **2019**, *177*, 149–166. [CrossRef]
- Wen, X.; Feng, Q.; Deo, R.C.; Wu, M.; Yin, Z.; Yang, L.; Singh, V.P. Two-phase extreme learning machines integrated with the complete ensemble empirical mode decomposition with adaptive noise algorithm for multi-scale runoff prediction problems. *J. Hydrol.* 2019, 570, 167–184. [CrossRef]
- 40. Yang, Y.; Wang, J. Forecasting wavelet neural hybrid network with financial ensemble empirical mode decomposition and MCID evaluation. *Expert Syst. Appl.* **2021**, *166*, 114097. [CrossRef]
- 41. Ali, M.; Deo, R.C.; Maraseni, T.; Downs, N.J. Improving SPI-derived drought forecasts incorporating synoptic-scale climate indices in multi-phase multivariate empirical mode decomposition model hybridized with simulated annealing and kernel ridge regression algorithms. *J. Hydrol.* **2019**, *576*, 164–184. [CrossRef]
- 42. Das, A.K.; Das, S.; Ghosh, A. Ensemble feature selection using bi-objective genetic algorithm. *Knowl.-Based Syst.* 2017, 123, 116–127. [CrossRef]
- Rahmati, O.; Choubin, B.; Fathabadi, A.; Coulon, F.; Soltani, E.; Shahabi, H.; Mollaefar, E.; Tiefenbacher, J.; Cipullo, S.; Ahmad, B.B.; et al. Predicting uncertainty of machine learning models for modelling nitrate pollution of groundwater using quantile regression and UNEEC methods. *Sci. Total Environ.* 2019, *688*, 855–866. [CrossRef]
- 44. Mohapatra, J.B.; Jha, P.; Jha, M.K.; Biswal, S. Efficacy of machine learning techniques in predicting groundwater fluctuations in agro-ecological zones of India. *Sci. Total Environ.* **2021**, *785*, 147319. [CrossRef]
- 45. Band, S.S.; Heggy, E.; Bateni, S.M.; Karami, H.; Rabiee, M.; Samadianfard, S.; Chau, K.W.; Mosavi, A. Groundwater level prediction in arid areas using wavelet analysis and Gaussian process regression. *Eng. Appl. Comp. Fluid.* **2021**, *15*, 1147–1158. [CrossRef]
- 46. Namous, M.; Hssaisoune, M.; Pradhan, B.; Lee, C.W.; Alamri, A.; Elaloui, A.; Edahbi, M.; Krimissa, S.; Eloudi, H.; Ouayah, M.; et al. Spatial prediction of groundwater potentiality in large semi-arid and karstic mountainous region using machine learning models. *Water* **2021**, *13*, 2273. [CrossRef]
- 47. Chen, C.; He, W.; Zhou, H.; Xue, Y.R.; Zhu, M.D. A comparative study among machine learning and numerical models for simulating groundwater dynamics in the Heihe River Basin, northwestern China. *Sci. Rep.* **2021**, *10*, 3904. [CrossRef] [PubMed]
- 48. Ao, C.; Zeng, W.Z.; Wu, L.F.; Qian, L.; Srivastava, A.K.; Gaiser, T. Time-delayed machine learning models for estimating groundwater depth in the Hetao Irrigation District, China. *Agric. Water Manag.* **2021**, 255, 107032. [CrossRef]
- 49. Lykkegaard, M.B.; Dodwell, T.J.; Moxey, D. Accelerating uncertainty quantification of groundwater flow modelling using a deep neural network proxy. *Comput. Methods Appl. Mech. Eng.* 2021, 383, 113895. [CrossRef]
- Wu, M.; Feng, Q.; Wen, X.H.; Yin, Z.L.; Yang, L.S.; Sheng, D.R. Deterministic analysis and uncertainty analysis of ensemble forecasting model based on variational mode decomposition for estimation of monthly groundwater level. *Water* 2021, *13*, 139. [CrossRef]
- 51. Zuo, G.; Luo, J.; Wang, N.; Lian, Y.; He, X. Decomposition ensemble model based on variational mode decomposition and long short-term memory for streamflow forecasting. *J. Hydrol.* **2020**, *585*, 124776. [CrossRef]
- Huang, N.E.; Shen, Z.; Long, S.R.; Wu, M.C.; Snin, H.H.; Zheng, Q.; Yen, N.C.; Tung, C.C.; Liu, H.H. The empirical mode decomposition and the Hubert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. A Math. Phys. Eng. Sci.* 1998, 454, 903–995. [CrossRef]
- 53. Huang, C.L.; Wang, C.J. A GA-based feature selection and parameters optimization for support vector machines. *Expert Syst. Appl.* **2006**, *31*, 231–240. [CrossRef]
- 54. Ahn, G.; Hur, S. Efficient genetic algorithm for feature selection for early time series classification. *Comput. Ind. Eng.* **2020**, 142, 106345. [CrossRef]

- 55. Divya, R.; Shantha Selva Kumari, R. Genetic algorithm with logistic regression feature selection for Alzheimer's disease classification. *Neural Comput. Appl.* **2021**, *33*, 8435–8444. [CrossRef]
- 56. Yang, J.; Honavar, V. Subset selection using a genetic algorithm. In *Feature Extraction, Construction and Selection*; Liu, H., Motoda, H., Eds.; Longman: Boston, UK, 1998; Chapter 8; p. 117.
- 57. Hinton, G.E.; Osindero, S.; Teh, Y.W. A fast learning algorithm for deep belief nets. Neural Comput. 2006, 18, 1527–1554. [CrossRef]
- 58. Koenker, R.; Bassett, G. Regression Quantiles. *Econometrica* **1978**, *46*, 33–50. [CrossRef]
- 59. Koenker, R.; Hallock, K.F. Quantile regression. J. Econ. Perspect. 2001, 15, 143–156. [CrossRef]
- 60. Kasraei, B.; Heung, B.; Saurette, D.D.; Schmidt, M.G.; Bulmer, C.E.; Bethel, W. Quantile regression as a generic approach for estimating uncertainty of digital soil maps produced from machine-learning. *Environ. Modell. Softw.* **2021**, *144*, 105139. [CrossRef]
- 61. Weerts, A.H.; Winsemius, H.C.; Verkade, J.S. Estimation of predictive hydrological uncertainty using quantile regression: Examples from the national flood forecasting system (England and Wales). *Hydrol. Earth Syst. Sci.* **2011**, *15*, 255–265. [CrossRef]
- 62. The second hydrogeological brigade for the bureau of Gansu geology and mineral resources. In *Manual of Comprehensive Hydrogeological Map of the People's Republic of China-Jiuquan Sheet;* SinoMaps Press: Beijing, China, 1981; p. 6.
- 63. Wei, G.; Wang, G.; Xu, T.; Zhu, F. The evaluation of groundwater system in Jiuquan Basin. J. Arid Land Res. Environ. 2008, 22, 38–44.
- 64. Wan, L.; Dong, Y.; Xu, Z. A synthesis of hydrochemistry with an integrated conceptual model for groundwater in the Hexi Corridor, northwestern China. *J. Asian Earth Sci.* **2017**, *146*, 20–29. [CrossRef]
- 65. Ren, X.; Gao, Z.; An, Y.; Liu, J.; Wu, X.; He, M.; Feng, J. Hydrochemical and isotopic characteristics of groundwater in the Jiuquan East Basin, China. *Arab. J. Geosci.* **2020**, *13*. [CrossRef]
- 66. He, J.; Ma, J.; Zhang, P.; Tian, L.; Zhu, G.; Edmunds, M.W.; Zhang, Q. Groundwater recharge environments and hydrogeochemical evolution in the Jiuquan Basin, Northwest China. *Appl. Geochem.* **2012**, *27*, 866–878. [CrossRef]
- 67. Legates, D.R.; McCabe, G.J. Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. *Water Resour. Res.* **1999**, *35*, 233–241. [CrossRef]
- 68. Moriasi, D.N.; Arnold, J.G.; Van Liew, M.W.; Bingner, R.L.; Harmel, R.D.; Veith, T.L. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans. ASABE* 2007, *50*, 885–900. [CrossRef]
- 69. 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]
- 70. Muthusamy, M.; Godiksen, P.N.; Madsen, H. Comparison of different configurations of quantile regression in estimating predictive hydrological uncertainty. *Proc. Eng.* **2016**, 154, 513–520. [CrossRef]
- 71. Feng, Q.; Wen, X.; Li, J. Wavelet analysis-support vector machine coupled models for monthly rainfall forecasting in arid regions. *Water Resour. Manag.* **2015**, *29*, 1049–1065. [CrossRef]
- 72. Zeynoddin, M.; Bonakdari, H.; Ebtehaj, I.; Esmaeilbeiki, F.; Gharabaghi, B.; Haghi, D.Z. A reliable linear stochastic daily soil temperature forecast model. *Soil Till. Res.* **2019**, *189*, 73–87. [CrossRef]
- Yin, Z.; Feng, Q.; Wen, X.; Deo, R.C.; Yang, L.; Si, J.; He, Z. Design and evaluation of SVR, MARS and M5 Tree models for 1, 2 and 3-day lead time forecasting of river flow data in a semiarid mountainous catchment. *Stoch. Environ. Res. Risk Assess.* 2018, 32, 2457–2476. [CrossRef]
- 74. Wen, X.; Feng, Q.; Deo, R.C.; Wu, M.; Si, J. Wavelet analysis-artificial neural network conjunction models for multi-scale monthly groundwater level predicting in an arid inland river basin, northwestern China. *Hydrol. Res.* **2017**, *48*, 1710–1729. [CrossRef]
- 75. Bahmani, R.; Ouarda, T.B.M.J. Groundwater level modeling with hybrid artificial intelligence techniques. *J. Hydrol.* **2021**, *595*, 125659. [CrossRef]
- 76. Barzegar, R.; Aalami, M.T.; Adamowski, J. Coupling a hybrid CNN-LSTM deep learning model with a Boundary Corrected Maximal Overlap Discrete Wavelet Transform for multiscale Lake water level forecasting. *J. Hydrol.* **2021**, *598*, 126196. [CrossRef]
- 77. Sui, G.S. Research on groundwater dynamic characteristics and influencing factors in Jiuquan west basin. *Groundwater* **2019**, *41*, 41–42; 57. [CrossRef]