water

MDPI

Article

Modeling and Investigating the Mechanisms of Groundwater Level Variation in the Jhuoshui River Basin of Central Taiwan

Tao Bai [1], Wen-Ping Tsai [1,2,*], Yen-Ming Chiang [3], Fi-John Chang [4], Wan-Yu Chang [4], Li-Chiu Chang [5] and Kuang-Chih Chang [6]

1 State Key Laboratory of Eco-Hydraulics in Northwest Arid Region of China, Xi'an University of Technology, Xi'an 710048, China
2 Department of Civil and Environmental Engineering, Pennsylvania State University, University Park, PA 16802, USA
3 Institute of Hydrology and water resources, Zhejiang University, Hangzhou 310058, China
4 Department of Bioenvironmental Systems Engineering, National Taiwan University, Taipei 10617, Taiwan
5 Department of Water Resources and Environmental Engineering, Tamkang University, New Taipei City 25137, Taiwan
6 Water Resources Agency, Ministry of Economic Affairs, Taipei 10651, Taiwan
* Correspondence: wptsai11@gmail.com

Received: 11 May 2019; Accepted: 24 July 2019; Published: 27 July 2019

Abstract: Due to nonuniform rainfall distribution in Taiwan, groundwater is an important water source in certain areas that lack water storage facilities during periods of drought. Therefore, groundwater recharge is an important issue for sustainable water resources management. The mountainous areas and the alluvial fan areas of the Jhuoshui River basin in Central Taiwan are considered abundant groundwater recharge regions. This study aims to investigate the interactive mechanisms between surface water and groundwater through statistical techniques and estimate groundwater level variations by a combination of artificial intelligence techniques and the Gamma test (GT). The Jhuoshui River basin in Central Taiwan is selected as the study area. The results demonstrate that: (1) More days of accumulated rainfall data are required to affect variable groundwater levels in low-permeability wells or deep wells; (2) effective rainfall thresholds can be properly identified by lower bound screening of accumulated rainfall; (3) daily groundwater level variation can be estimated effectively by artificial neural networks (ANNs); and (4) it is difficult to build efficient models for low-permeability wells, and the accuracy and stability of models is worse in the proximal-fan areas than in the mountainous areas.

Keywords: groundwater level; recharge groundwater; Gamma test (GT); accumulated rainfall; artificial neural networks (ANNs)

1. Introduction

Due to industrial and economic development along with rapid population growth in Taiwan in recent years, more water sources are required to satisfy the increasing water demands from the development of human civilization and livelihoods. Due to its easy accessibility, low cost, and stable quality and temperature, groundwater is often considered an important water source. The interactive mechanism of groundwater level variations is rather complex and includes precipitation, tides, atmospheric pressure, and even earthquakes. Previous research has shown that rainfall can be regarded as a leading indicator of groundwater levels by applying time series to analyze the relationship between rainfall and groundwater levels [1–7]. However, the temporal relationship between groundwater and rainfall might vary widely in different regions. In this study area, Taiwan, some research has suggested that time lags between rainfall and groundwater level variation might occur [8]. Chen [9] identified that different durations of rainfall affected groundwater levels and that response time spans range from a few hours to a few days.

Artificial neural networks (ANNs) are recognized as an effective technique for modeling complicated systems (nonlinear issues) and have been widely adopted in various hydro system problems in past decades. The advantages of ANNs include their ability to learn from datasets, their flexible noise tolerance, and their ability to generalize. Additionally, ANNs can extract significant and meaningful features from complex data structures and can learn the implicit relationship between inputs and outputs based on sufficient training data [10,11]. Pradhan and Lee [12] used a back-propagation neural network (BPNN) coupled with a geographic information system (GIS) to analyze the risk of landslides. Karthikeyan et al. [13] predicted ground water levels in India by using two networks, feed forward neural networks (FFNN) and recurrent neural network (RNN), and adjusting the parameters according to the related input factors revealed that the hydro-meteorological parameters affect the models. Sreekanth et al. [14] predicted groundwater levels to deal with the problems of water resource usage through FFNN-LMB applied to groundwater level forecasting and verified that ANN is a promising tool. Li et al. [15] executed the back-propagation (BP)coupled with a sensitivity analysis to determine that the most sensitive anthropogenic factor, coal mining drainage, has a major effect on groundwater levels in the Jinci Spring Basin in northern China. Mohanty et al. [16] used various methods to construct models of groundwater systems and found that ANN models provided better short-term groundwater level prediction than MODFLOW-based numerical models without sufficient parameters, data and boundary conditions. New techniques such as the neuro-fuzzy inference system that integrates ANNs and fuzzy logic methods have also proven powerful. The neuro-fuzzy inference system has the potential to incorporate the strengths of both ANN and fuzzy logic in one framework. The adaptive neuro-fuzzy inference system (ANFIS) was introduced to provide a new approach on the reservoirs' optimal operations [17] and rainfall-runoff simulations [18].

Some studies have focused on the interactive mechanism of surface water and groundwater in mountainous areas. The purpose of this study is to explore the interactive mechanism of surface and subsurface water by analyzing and modeling groundwater level variations using statistical and artificial intelligence techniques. The approaches consist of correlation analysis between precipitation and groundwater level variation and the estimation of groundwater level variations. The Jhuoshui River basin in Central Taiwan is selected as a study area.

2. Methodology

This study investigates the interactive mechanism of groundwater level variations through artificial neural networks, statistical analyses, the Gamma test and sensitivity analysis. The methodologies used in this research are briefly addressed as follows:

2.1. Pearson Correlation Coefficient

The Pearson correlation coefficient [19] is a quantitative method used to measure the linear correlation between two variables. In this study, it is used to determine the temporal relationship between groundwater level variations and precipitation as well as stream flow. The Pearson correlation coefficient ranges between $[-1,1]$, in which 1 denotes a positive exact linear relationship, -1 indicates a negative perfect linear relationship, and 0 indicates that no linear relationship exists between the two variables.

For variables of P (precipitation) and G (groundwater) with n datasets (each dataset can be denoted by p_i and g_i, where $i = 1, 2, \ldots, n$), the correlation coefficient can estimate the Pearson correlation (R) between P and G. The calculation of the correlation coefficient is given as:

$$R_{P,G} = \frac{\sum_{i=1}^{n}\left(p_i - \overline{P}\right)\left(g_i - \overline{G}\right)}{\sqrt{\sum_{i=1}^{n}\left(p - \overline{P}\right)^2}\sqrt{\sum_{i=1}^{n}\left(g_i - \overline{G}\right)^2}} \tag{1}$$

where $\overline{P}$ and $\overline{G}$ are the means of P and G accordingly.

2.2. Gamma Test (GT)

The Gamma test (GT) is designed to estimate noise variance by measuring the smooth relationship between the input-output dataset. The GT was initially proposed by Končar and Aðalbjörn [20,21] to determine the best input combination for a neural network [22,23]. This study implements the GT for the extraction of nonlinearity between groundwater level variations and precipitation, along with the identification of important factors affecting groundwater level variations.

Given an input-output pair denoted by $(X, y) = ((x_1, \ldots, x_m), y)$, where X is the input items and scalar y is the output items, the observations can be described as Equation (2).

$$\{(X_i, y_i), 1 \le i \le M\} \tag{2}$$

where, $X_i \in R^m$ are the m-dimensional input variables with a dataset length of M, which is constrained to a closed bounded set $C \in R^m$. The corresponding outputs $y_i \in R$ are scalars. The underlying relationship of the data set can be denoted:

$$y = f(x_1, \ldots x_m) + r \tag{3}$$

where, f is a smooth function, and r denotes a random variable for noise. The Gamma statistic (Γ) is a variance estimation of noise. Let $X_{i,k}$ denote the kth nearest neighbor to X_i in terms of Euclidean distance. The delta function is defined:

$$\delta_M(k) = \frac{1}{M}\sum_{i=1}^{M}|X_{i,k} \quad X_i| \quad (1 \le k \le p) \tag{4}$$

where, $|\cdots|$ is the Euclidean distance, and p is the number of neighbors. For computing the Γ, the line of least squares regression is built for p points:

$$\gamma = A\delta + \Gamma \tag{5}$$

where, A is the gradient. The steeper the gradient is, the more complex is the model. The Γ is an index to evaluate the model performance of an input combination. A Γ closer to 0 implies a more appropriate input combination.

2.3. Back-Propagation Neural Network (BPNN)

The back-propagation neural network (BPNN), developed by Rumelhart et al. [24], is used to build a model for estimating the groundwater level variations in this study. Figure 1 shows the BPNN structure in this study.

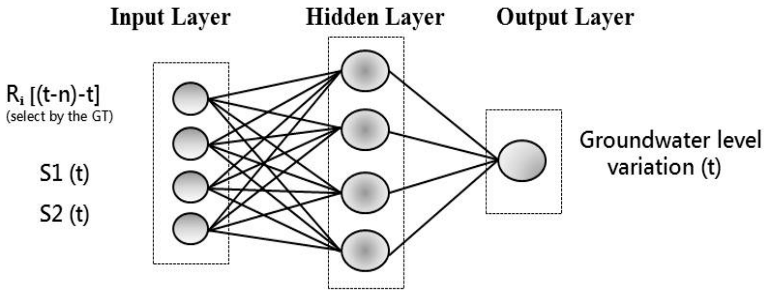


Figure 1. Architecture of the back-propagation neural network (BPNN) model.

The input items comprise the streamflow and accumulated rainfall. The node number of the hidden layer is determined by trial and error. The model target is groundwater level variation. The BPNN applies the gradient steepest descent method for adjusting the neuron weights for minimizing the output error. In the learning process (calculation iterations), the neuron weights can be adjusted by an error convergence technique to approximate the model target based on the given input set. The output layer error may propagate backward to the input layer. Further details of the BPNN algorithm are found in Rumelhart et al. [24].

2.4. Adaptive Neuro-Fuzzy Inference System (ANFIS)

The adaptive neuro-fuzzy inference system (ANFIS) was proposed by Jang [25] who used the fuzzy inference system as an essential core in combination with the artificial neural network. The ANFIS preserves the learning process of ANNs for projecting input features to an output space and retains the fuzzy advantages, including if-then rules (rule layer), for describing the regional behavior of such projection. Then, the results of the inference system are acquired through the reasoning capability of fuzzy logic. The ANFIS was shown to have powerful modeling abilities in comprehensive fields such as motor fault detection and diagnosis [26], power systems dynamic load [27], forecasting systems for the demand of teachers' human resources [28] and real-time reservoir operations [29].

This study applies the Takagi-Sugeno fuzzy model [30] in a fuzzy rule layer. Here, 2 input variables, p (precipitation) and s (streamflow), and 1 output variable, g (groundwater), are taken as an example for describing the rule layer. The rule sets of the rule layer can be expressed as:

Rule 1: If p is A_1 and s is B_1 then $g_1 = k_1{*}p + t_1{*}s + r_1$
Rule 2: If p is A_2 and s is B_2 then $g_2 = k_2{*}p + t_2{*}s + r_2$

where k, t and r are linear parameters in the consequent part (then-part) of the first-order Takagi-Sugeno fuzzy model.

The typical ANFIS includes 5 layers (Figure 2). The input layer—in this layer, each node produces membership degrees that belong to each of the fuzzy sets by using appropriate membership functions. The rule layer—the AND operator is used to obtain one output that represents the result of the antecedent for that rule. The average layer—the ratio of each ith rule's firing strength to the sum of all the rules' firing strength is calculated. The consequent layer—the node function determines the contribution of each ith rule's toward the total output. The output layer—the defuzzification process transforms each rule's fuzzy results into the model output. The details of the ANFIS are found in Chang and Chang [31].

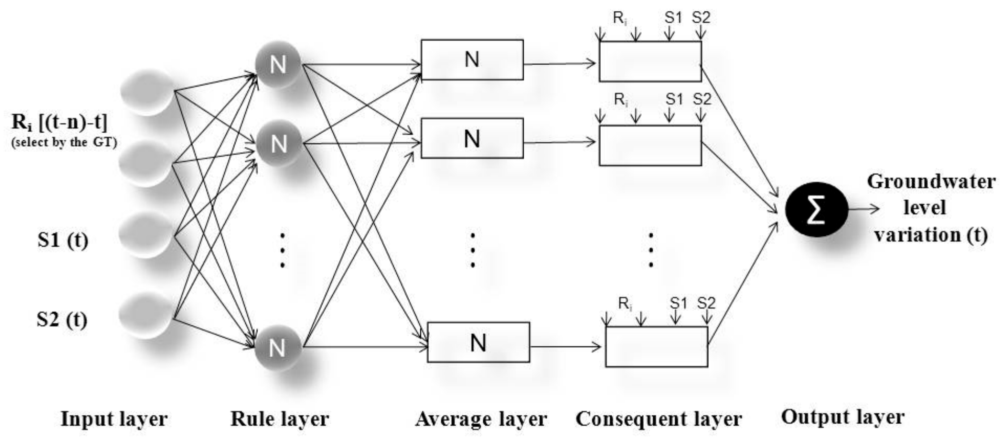
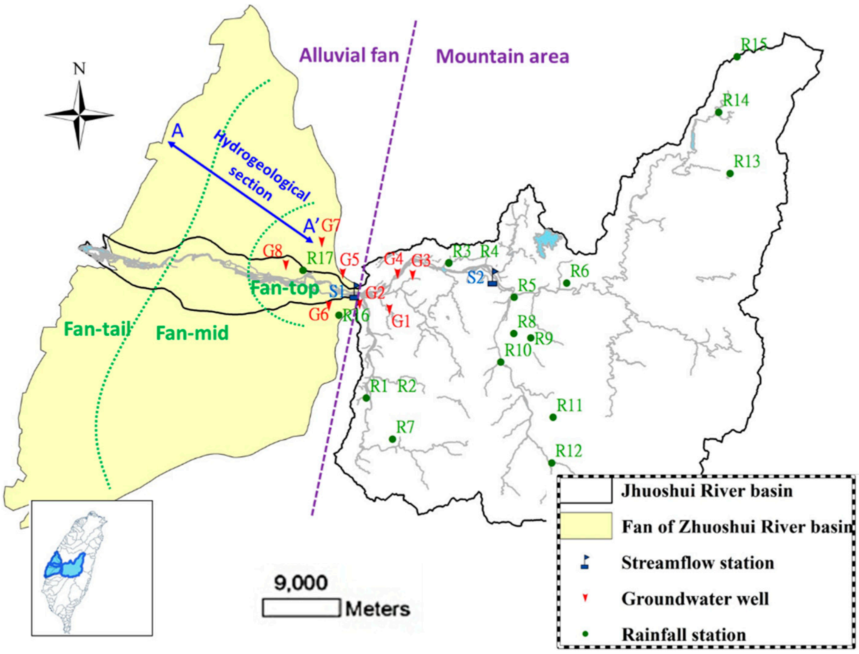


Figure 2. Architecture of the adaptive neuro-fuzzy inference system (ANFIS) model.

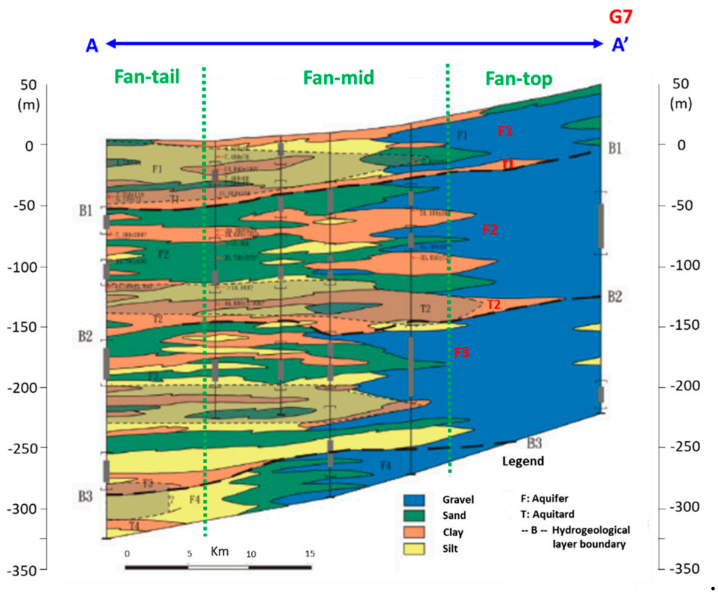
3. Case Study

This study investigates the interactive mechanism of groundwater level variations and precipitation in mountainous areas in the Jhuoshui River basin by using statistical methods. The Jhuoshui River basin is located in Central Taiwan with a maximum elevation of 3200 m. The watershed area of Jhuoshui River is approximately 3157 km^2. Figure 3a shows the distribution of the gauge stations and the hydrogeological information of groundwater monitoring wells in the study area. There are 2 main regions in this study area: The mountain area and the alluvial fan. The alluvial fan can be roughly divided into 3 districts: The fan-top district; the fan-mid district; the fan-tail district (Figure 3a). The G1~G4 monitoring wells belong to the mountain area, and the G5~G8 monitoring wells belong to the fan-top of the alluvial fan. As shown in Figure 3b, the fan-top district is the only district without a significant confined aquifer. In general, the hydrogeological feature of the fan-top district in this study area is recognized as a high potential groundwater recharge area because the main geological components are sandstone, phyllite and slate (the gravel of the geological structure) (Figure 3b). According to investigations of core drilling samples, the soil thickness of this study area (the depth to bedrock, including the soil layer, colluvium layer and saprolite) is deep [32]. The Central Geological Survey (CGS) (Ministry of Economic Affairs (MOEA), R.O.C.) indicated that the hydrogeological structure of the study area can be divided into several strata. Based on the depth from the land surface (Figure 3b), these strata are as follows: Aquifer 1 (F1), Aquitard 1 (T1), Aquifer 2 (F2), Aquitard 2 (T2) and Aquifer 3 (F3). The average thicknesses of Aquifer 1, Aquifer 2 and Aquifer 3 are 42 m, 95 m and 86 m, respectively. The thickness of the gravel and sand strata in the fan-top areas can reach more than 130 m [33]. In this study area, some groundwater monitoring wells include two different depth monitoring records (in a different aquifer), such as G1, G4, G7 and G8 (Figure 3c). However, according the hydrogeological information provided by the CGS (Figure 3b), it is recognized that the shallow groundwater well in G7 belongs to Aquifer 3 (F3). The rainfall is distributed unevenly, which mainly occurs from May to September due to the unique topographical terrain and location. The total rainfall in wet periods comprises 75% of the annual rainfall, which implies rainfall differs significantly between wet and dry periods.

The groundwater level data at twelve groundwater monitoring wells and the flow data at two streamflow gauging stations were collected from the Water Resources Agency in Taiwan during 2001 and 2011, and rainfall data at seventeen rainfall gauging stations were collected from the Water Resources Agency, Central Weather Bureau and Taiwan Power Company in Taiwan in 2001 and 2011 (Tables 1–3). The missing data were infilled by using linear regression techniques based on the data of the surrounding stations. The flowchart of this study is shown in Figure 4. First, this study conducts rainfall duration analyses by investigating the effective duration of accumulated rainfall on groundwater level variations. Second, this study conducts effective rainfall analysis by identifying significant accumulated rainfall amount thresholds affecting groundwater level variations, and therefore the correlation between accumulated rainfall amounts and groundwater level variations can be estimated. Third, this study constructs estimation models of the groundwater level variations by using the BPNN and ANFIS based on rainfall, streamflow and groundwater level variation data, in which the GT is used to determine the proper input variables associated with the rainfall gauging stations as well as the streamflow gauging stations (S1), and rainfall durations for each model at individual groundwater monitoring stations. Finally, the performance of the models is compared.

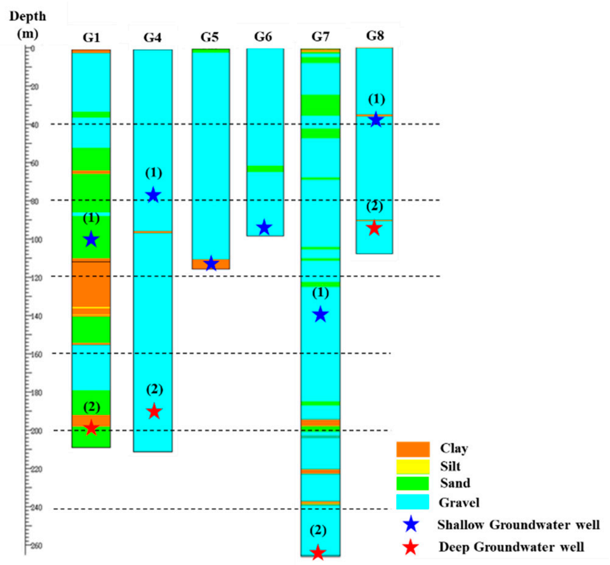


(a) The distribution of the streamflow stations, the groundwater wells and the rainfall stations.



(**b**) The hydrogeological section. [Original sources are provided by the CGS, MOEA, R.O.C].

Figure 3. *Cont.*



(**c**) The vertical profile of the soil texture of the groundwater monitoring wells (information of G2 and G3 are unavailable). [Original sources are provided by the CGS, MOEA, R.O.C].

Figure 3. The distribution of the gauge stations and hydrogeological information of the groundwater monitoring wells.

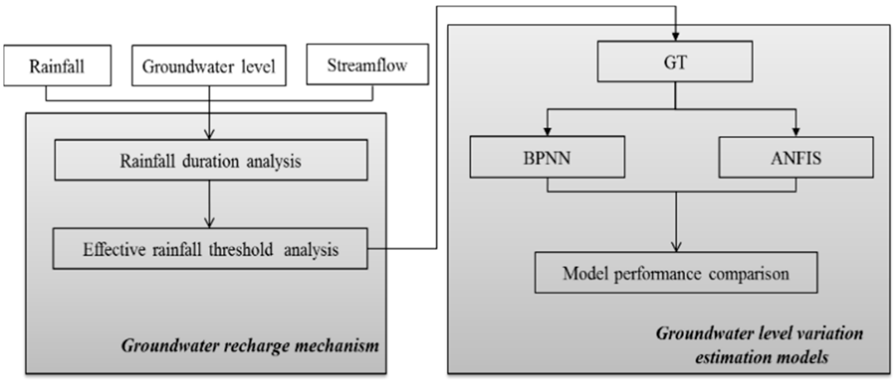


Figure 4. Implementation flowchart of this study.

Table 1. The statistics of seventeen rainfall stations.

Rainfall Station	Elevation (m)	Rainfall (mm)		
		Mean	SD [1]	Annual Rainfall
R1	400	8.18	29.16	2985
R2	231	8.43	30.64	3076
R3	203	6.24	20.77	2278
R4	215	6.35	21.33	2320
R5	296	5.93	22.96	2166
R6	393	5.76	20.99	2103
R7	724	8.14	37.45	2970
R8	322	5.64	21.60	2059
R9	1666	7.31	26.50	2669
R10	485	8.91	29.34	3250
R11	2200	7.84	30.56	2863
R12	1135	6.04	25.60	2203
R13	1200	7.38	26.33	2695
R14	1520	7.01	27.07	2558
R15	2303	5.46	22.40	1991
R16	82	5.15	19.30	1880
R17	110	4.44	18.33	1620

[1] Standard deviation.

Table 2. The basic statistics of twelve groundwater monitoring wells.

Monitoring Well	Well Depth (m)	Elevation (m)	Groundwater Level (m)	
			Mean	SD
G1(1) [shallow]	102.6	151.2	141.7	1.86
G1(2) [deep]	199.3	151.2	142.7	1.46
G2	150	113.3	109.0	4.63
G3	24.1	179.3	169.1	1.03
G4(1) [shallow]	78.2	151.1	137.6	2.17
G4(2) [deep]	193.2	151	135.5	1.45
G5	112.7	82.4	38.75	3.35
G6	96	72.3	38.48	2.84
G7(1) [shallow]	140	49.5	35.97	2.38
G7(2) [deep]	269	49.6	34.02	2.31
G8(1) [shallow]	38.7	46.6	34.73	1.55
G8(2) [deep]	97.5	46.5	34.73	1.52

Table 3. The basic information of two streamflow gauging stations.

Stream Flow Station	Elevation (m)	Discharge (cm s)	
		Mean	SD
S1	107.17	136.51	401.22
S2	279.09	113.59	272.19

The Pearson correlation coefficient and the root mean square error (RMSE) are used to evaluate the performance of the estimation model. The calculation of RMSE is given below:

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - d_i)^2} \tag{6}$$

where, y_i is the model estimation and d_i is the observation, and N is the number of datasets. The RMSE is used to evaluate the accuracy of the estimations of the groundwater level variations. The lower the RMSE value is, the better is the model's performance.

4. Results and Discussion

4.1. Duration of Accumulated Rainfall Analysis

First, this study uses the Pearson correlation coefficient to investigate the relationship between the duration of accumulated rainfall (1–10 days) and the groundwater level variation. A high correlation coefficient indicates a strong relationship between the accumulated rainfall and the groundwater level variation. To generate a comprehensive data set, rainfall levels at each rainfall gauging station can be determined by the average rainfall over the basin obtained from the Thiessen polygon method [34]. Figure 5 shows the time series of groundwater levels at the groundwater monitoring wells and the rainfall data at R2. This study also analyzed the relationship between the groundwater level variation and the rainfall at a 1-day lag (from the current day to the previous day), and further compared the results with a duration analysis. At all groundwater monitoring stations, the results show that the correlation coefficient of groundwater level variations and accumulated rainfall with different durations (2–10 days) is higher than the rainfall at the current day and 1-day lag. In other words, the accumulated rainfall has a stronger positive relationship with the groundwater level variation than the rainfall of the previous day at all wells.

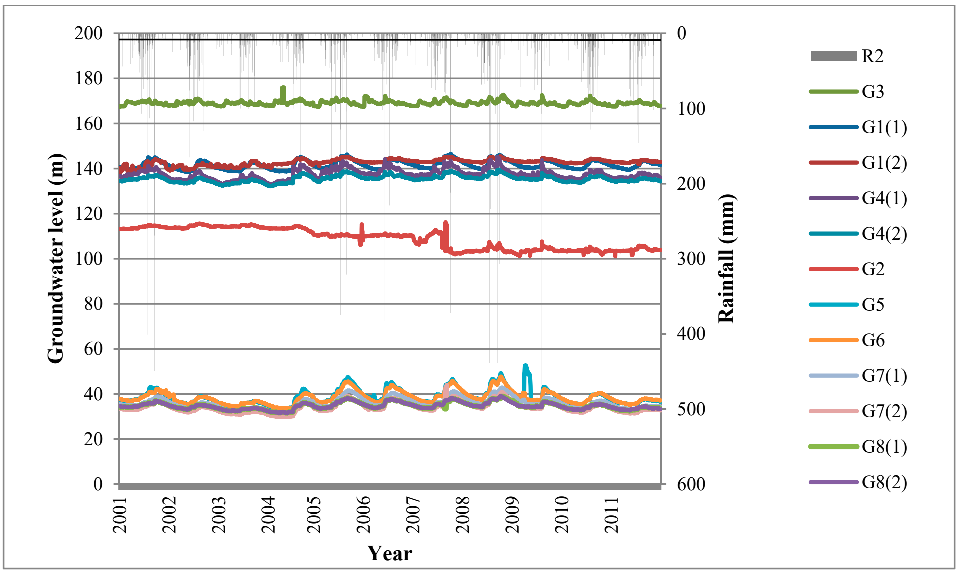


Figure 5. Time series of the groundwater level at the groundwater monitoring wells and the rainfall data at R2.

As shown in Figure 6, the highest correlation between the groundwater level variation and two-day accumulated rainfall occurs at G4(1), G4(2) and G2, which might result from the gravel of the geological structure at G4 and G2. The highest correlation between groundwater level variations and three-day accumulated rainfall occurs at G1(1) and G3, which may be because of the shallow well depth of G1(1) and G3. The highest correlation between the groundwater level variation and a five-day accumulated rainfall occurs at G1(2), which might be caused by the deep well depth of G1(2). Therefore, the accumulated rainfall over different durations is selected as an input variable in the ANN models for estimating the groundwater level variations. In addition, the differences in geological structure and depth at each groundwater monitoring well produces different relationships between the groundwater level variations and the accumulated rainfall of different durations.

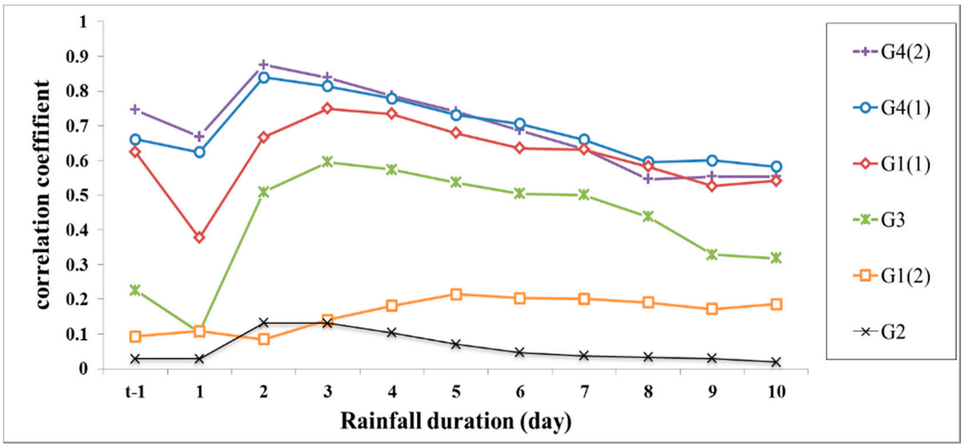


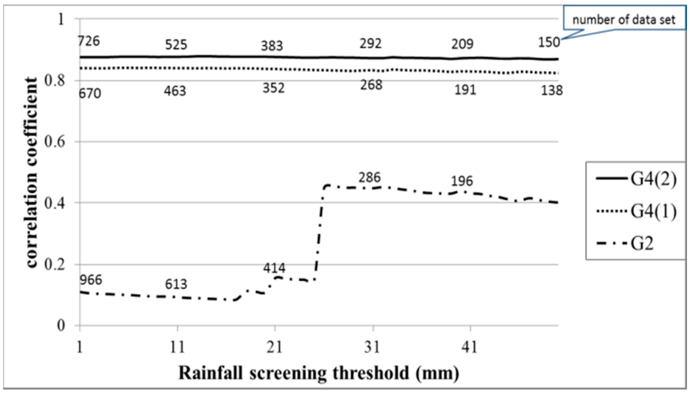
Figure 6. The correlation coefficient between the groundwater level variation and rainfall durations at different groundwater monitoring wells.

4.2. *Effective Rainfall Analysis*

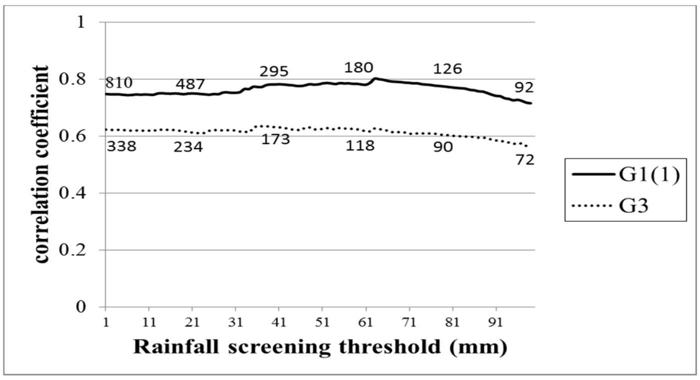
The different quantities of rainfall may have different impacts on the groundwater level variation. This study considers the average rainfalls associated with different thresholds for evaluating the rainfall amount impacts on the groundwater level variations at G1 to G4 by the use of the Pearson correlation coefficient. The duration of the accumulated rainfall adopts the result of Section 4.1 *Duration* *of accumulated rainfall analysis*: G4(1), G4(2) and G2—two days; G1(1) and G3—three days; G1(2)—five days. This study includes two types of screening:

(1) Lower bound screening: Delete the data sets below the rainfall threshold

The rainfall data (a total more than 3986) were screened at thresholds from 1–50 or 1–100 mm (with 1 mm increasing). Figure 7 shows the Pearson correlation between the groundwater level variations and the accumulated rainfall at different lower bound thresholds. In Figure 7a, the correlation patterns of G4(1) and G4(2) change slightly, and the correlation curve rises at the 13 mm threshold of the accumulated rainfall. This might be due to the greater structural permeability of G4 (gravel), as well as the location of G4 close to the Jhuoshui river, so that the groundwater level at G4 is affected by rainfall and the stream flow at the same time. The correlation curve of G2 rises obviously at the accumulated rainfall threshold of 26 mm. It is suspected that this causes saturation when the accumulated rainfall is 26 mm, and furthermore, the excess rainfall causes percolation to recharge the groundwater layer. Therefore, the effective rainfall thresholds affecting the groundwater level variation at G2 can be identified as 26 mm. In Figure 7b, the correlation pattern of G1(1) and G3 are uniform. At G1(1), the correlation increases gradually at the accumulated rainfall threshold of 34 mm, and the highest correlation coefficient is 0.8 at the accumulated rainfall threshold of 63 mm. In Figure 7c, the correlation pattern of G1(2) is similar to G2. The correlation curve of G1(2) rises obviously at the threshold of 48 mm of the accumulated rainfall, which causes effective groundwater recharge activity when the accumulated rainfall rises to 48 mm in five days.

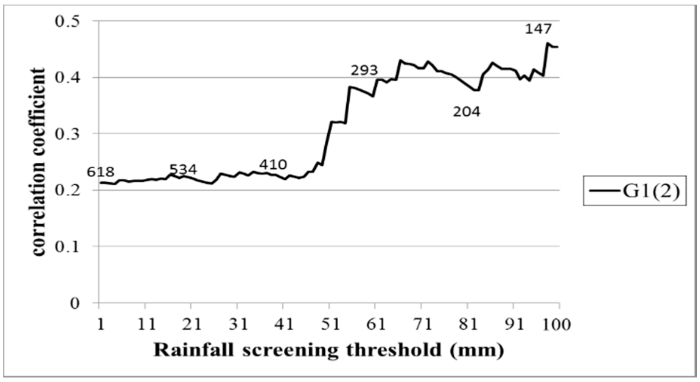


(**a**) G4(1), G4(2), G2—2 days



(**b**) G1(1), G3—2 days

Figure 7. *Cont.*

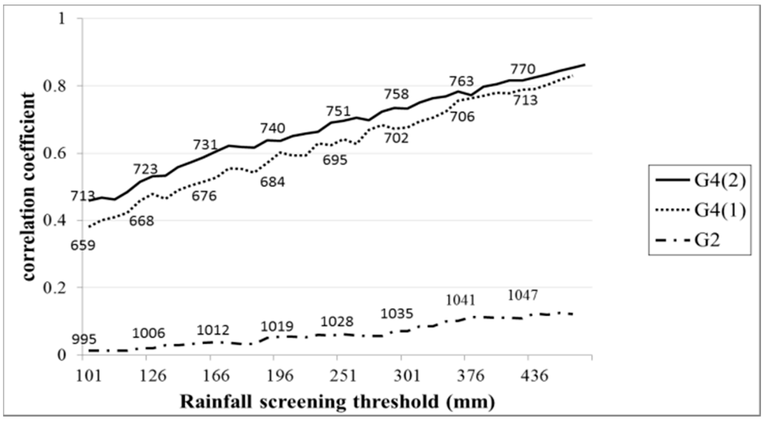


(**c**) G1(2)—5 days

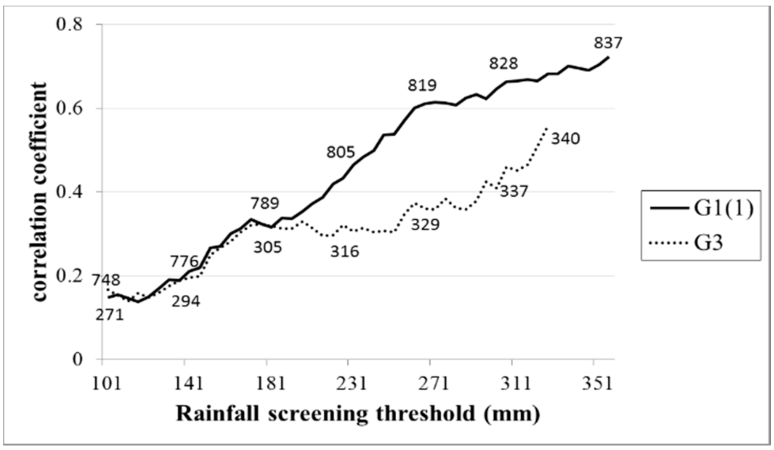
Figure 7. The correlation between the groundwater level variations and the accumulated rainfall amount at different low thresholds (Lower bound screening).

(2) Upper bound screening: deleting the data sets over the rainfall threshold

Figure 8 demonstrates the Pearson correlation between the groundwater level variations and the accumulated rainfall at different upper bound thresholds. All the groundwater monitoring wells show the same correlation patterns well.

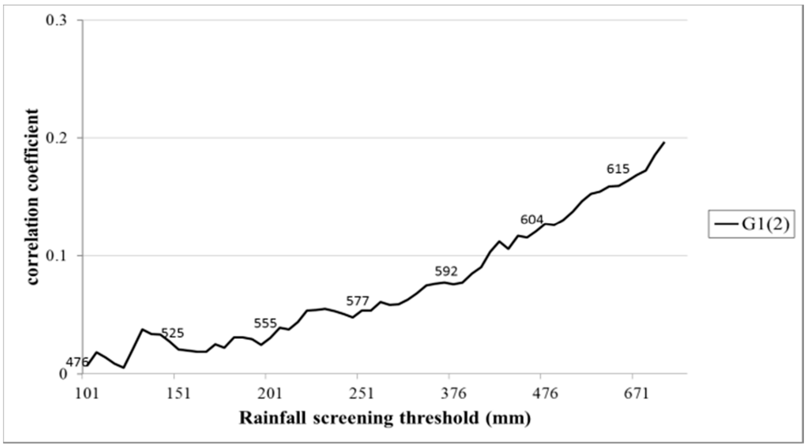


(**a**) G4(1), G4(2), G2—2 days



(**b**) G1(1), G3—3 days

Figure 8. *Cont.*



(**c**) G1(2)—5 days

Figure 8. The correlation between the groundwater level variations and the accumulated rainfall amount at different up thresholds (Upper bound screening).

With the higher upper bound rainfall threshold (which retain more heavy rainfall data), the correlation coefficient increases. This indicates that the groundwater level variations increase with more accumulated rainfall. This study can identify the lower bound threshold of the accumulated rainfall affecting the groundwater level effectively, but the authors cannot identify the upper bound threshold of accumulated rainfall.

4.3. Estimation of Groundwater Level Variations

This study built two estimation models (BPNN and ANFIS) for the groundwater level variations for each groundwater monitoring well. The learning and training efficiency of ANN models decrease if the data from all hydrology information is incorporated into the models. Therefore, the GT was used to determine the critical input factors (i.e., rainfall and streamflow) that correlated strongly with the groundwater level variations for ANN models. According to Chang et al. [35], it is better to use rainfall gauging stations as well as streamflow gauging stations as simultaneous input factors while constructing the groundwater level variation models at G1 to G4. Therefore, the study adopts the GT to evaluate the critical rainfall stations from G1 to G4, and streamflow gauging stations S1 and S2 are used as two other inputs for the ANN models. However, due to uncertainty as to whether or not the streamflow information is the critical factor of the groundwater level in the proximal-fan areas, this study adopts the GT to determine the critical rainfall gauging stations and streamflow gauging stations at G5 to G8. The GT is also used to detect rainfall durations (2–5 days) that have the highest correlations with the groundwater level variations at all wells. Table 4 shows the results of the GT analysis.

Table 4. The results of the Gamma test at the groundwater monitoring wells.

Groundwater Monitoring Well	Input Type: Streamflow	Input Type: Rainfall				Number of Datasets
		Rainfall Duration (days)	Rainfall Gauging Station Selected	Average Rainfall (mm)	Standard Deviation (mm)	
G1(1)	S1; S2	three	R2; R4	130	141	236
G1(2)	S1; S2	four	R3; R13	146	152	169
G2	S1; S2	two	R3; R6	60	79	372
G3	S1; S2	three	R3	102	98	243
G4(1)	S1; S2	two	R4; R7; R9	88	134	312
G4(2)	S1; S2	two	R7; R9; R13	88	135	364
G5	S1	two	R4; R12	64	94	251
G6	S1	two	R3; R12	62	93	330

Table 4. *Cont.*

Groundwater Monitoring Well	Input Type: Streamflow	Input Type: Rainfall				Number of Datasets
		Rainfall Duration (days)	Rainfall Gauging Station Selected	Average Rainfall (mm)	Standard Deviation (mm)	
G7(1)	S1	four	R17	145	120	139
G7(2)		four	R6; R7	171	252	154
G8(1)	S1	four	R17	139	120	142
G8(2)	S1	four	R17	138	119	150

The node numbers of the hidden layer in BPNN and the rule numbers in ANFIS were determined by trial and error. The optimal model structure is determined by the lowest RMSE value and the highest correlation coefficient in the validation phases at each groundwater monitoring station. The observations were divided into the training, validation and testing phases in the ratio of 5:2.5:2.5 or 6:2:2. Table 5 shows the estimation results of BPNN and ANFIS at each groundwater monitoring well (G1 to G8), and Table 6 shows the statistics of the observed data.

Table 5. The estimation performance of the BPNN and the ANFIS at the groundwater monitoring wells.

Well	ANN Model: BPNN [1] ANFIS [2]	RMSE (m)			Correlation Coefficient		
		Training	Validation	Testing	Training	Validation	Testing
G1(1)	BPNN (4-3-1)	0.091	0.108	0.11	0.844	0.696	0.724
	ANFIS (4-2-1)	0.085	0.112	0.133	0.867	0.666	0.578
G1(2)	BPNN (4-4-1)	0.115	0.137	0.149	0.481	0.304	0.221
	ANFIS (4-2-1)	0.114	0.138	0.149	0.491	0.275	0.22
G2	BPNN (4-4-1)	0.055	0.091	0.112	0.342	0.394	0.274
	ANFIS (4-2-1)	0.056	0.092	0.115	0.308	0.312	0.121
G3	BPNN (3-6-1)	0.058	0.083	0.128	0.844	0.846	0.682
	ANFIS (3-2-1)	0.063	0.107	0.129	0.806	0.753	0.667
G4(1)	BPNN (5-9-1)	0.063	0.089	0.109	0.895	0.806	0.893
	ANFIS (5-3-1)	0.058	0.094	0.144	0.912	0.779	0.775
G4(2)	BPNN (5-10-1)	0.048	0.067	0.069	0.928	0.88	0.865
	ANFIS (5-2-1)	0.045	0.064	0.062	0.938	0.887	0.89
G5	BPNN (3-3-1)	0.077	0.084	0.164	0.522	0.407	0.355
	ANFIS (3-2-1)	0.076	0.086	0.168	0.541	0.365	0.289
G6	BPNN (3-3-1)	0.06	0.079	0.126	0.438	0.368	0.167
	ANFIS (3-3-1)	0.049	0.065	0.124	0.677	0.705	0.236
G7(1)	BPNN (2-3-1)	0.12	0.121	0.138	0.752	0.715	0.734
	ANFIS (2-3-1)	0.076	0.126	0.147	0.907	0.703	0.662
G7(2)	BPNN (2-2-1)	0.132	0.186	0.186	0.619	0.625	0.582
	ANFIS (2-2-1)	0.124	0.192	0.213	0.675	0.599	0.504
G8(1)	BPNN (2-3-1)	0.093	0.096	0.146	0.847	0.843	0.841
	ANFIS (2-3-1)	0.055	0.121	0.207	0.949	0.692	0.841
G8(2)	BPNN (2-3-1)	0.103	0.112	0.118	0.874	0.859	0.832
	ANFIS (2-2-1)	0.07	0.122	0.126	0.944	0.841	0.909

[1] (number of input-number of nodes in the hidden layer-number of output); [2] (number of input-number of rules-number of output).

Table 6. The statistics of the observed data at the groundwater monitoring wells.

Well	Statistics (m)			
	Mean	SD	Maximum	Minimum
G1(1)	0.15	0.19	1.19	0.01
G1(2)	0.08	0.11	0.76	0.01
G2	0.16	0.65	7.95	0.01
G3	0.15	0.22	1.67	0.01
G4(1)	0.27	0.45	2.78	0.01
G4(2)	0.11	0.16	1.26	0.01
G5	0.07	0.1	0.86	0.01
G6	0.06	0.07	0.8	0
G7(1)	0.06	0.07	0.37	0.01
G7(2)	0.08	0.13	0.76	0.01
G8(1)	0.08	0.09	0.45	0.01
G8(2)	0.08	0.09	0.45	0.01

The results indicate that the ANN models provide an effective simulation and prediction of the groundwater level variation because the observations and the estimations have high correlations and small RMSE values. When comparing Tables 5 and 6, the RMSEs are smaller than the standard deviation at G1 to G4. When comparing the two types of models, the estimations from BPNN are better than those from the ANFIS model. In particular, BPNN shows more than 23% improvement in terms of the correlation coefficient at G1(1), G2 and G5. Therefore, when comparing the two areas of this study, it is easier to predict the groundwater level variation with adequate reliability and accuracy in mountainous areas than in proximal-fan areas.

5. Conclusions

Water resource deficiency is a global problem, especially under severe climate change. In Taiwan, groundwater is considered an important alternative water resource due to its low cost and convenient accessibility. The improper groundwater development may result in disasters coupled with environmental and economic losses, and strategic development for groundwater conservation is therefore a critical issue. Precipitation is the main source of groundwater recharge, so this study used statistical methods to explore the relationship between precipitation (duration and amount) and the groundwater level variation. The statistical results demonstrate that the connection of precipitation and groundwater can be described by the hydrogeological features. The geological structure (potential high-porosity and high-permeability) of the gravel and the thin soil thickness (shallow monitoring wells) resulted in a high correlation between the groundwater level variation and the short precipitation duration. The results also show that the groundwater level variation in the high-permeability geological structures would have minor effects on accumulated rainfall. In the other words, the high-permeability area can be considered to be an area sensitive to precipitation.

After identifying the relationship between precipitation and the groundwater level variations, this study built an estimation model of the groundwater level variation by artificial intelligence techniques. This study demonstrates that limited data can be used to build a robust estimation model for the groundwater level variation in the potential groundwater recharge area by artificial intelligence techniques. In other words, a model estimating the groundwater level variations with less parameter bias can be built since this model needs only the streamflow and the rainfall variables (which are easier to measure) as input items. This model feature is helpful when applied to complex hydrogeological systems. Furthermore, from the authors' viewpoint, the groundwater level variation is an important issue for water resource management, especially in the potential recharge area. If information can be accurately provided about the variation in the groundwater level to water resources decision makers, they can develop more appropriate water resource allocations or water management policies. Overall, the model shows better performance for high-permeability areas (mountain regions) than the low-permeability area (the proximal-fan region). This study also adopted the Gamma test to deal with the complex correlation between the surface water and the groundwater. The model results show that neural network models can provide accurate performance with suitable data preprocessing. In summary, the results indicate that neural network-based estimation models perform well. These results can also provide valuable information for the prevention and treatment of land subsidence and can serve as an effective reference for water resource management in the alluvial fan and mountain areas. The results also demonstrate that the application of artificial neural networks in the study of groundwater is a promising avenue.

Author Contributions: T.B.: Literature search, study design, figure design, data interpretation and writing; W.-P.T.: Literature search, study design, figure design, data analysis, data interpretation and writing; Y.-M.C.: Literature search, study design and data interpretation; F.-J.C.: Literature search, study design and data interpretation; W.-Y.C.: Literature search, data collection and data analysis; L.-C.C.: Literature research, study design and data interpretation; K.-C.C.: Literature research and data interpretation.

Funding: National Key R & D Program of China (2017YFC0405900); Water Resources Agency, Taiwan, ROC (Grant number: MOEAWRA1010425).

Acknowledgments: The streamflow and groundwater level data provided by the Water Resources Agency, Taiwan, and the rainfall data provided by the Water Resources Agency, Central Weather Bureau and Taiwan Power Company, Taiwan, are very much appreciated. In addition, the authors are indebted to the reviewers for their valuable comments and suggestions.

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

References

1. Changnon, S.A.; Huff, F.A.; Hsu, C.F. Relations between precipitation and shallow groundwater in illinois. *J. Clim.* **1988**, *1*, 1239–1250. [CrossRef]
2. Cheng, J.D.; Hong, H.N.; Chou, L.H. Use Time Series to Analysis Rainfall and Groundwater Level. *J. Chin. Soil Water Conserv.* **2003**, *35*, 47–56. (In Chinese)
3. Chen, Z.; Grasby, S.E.; Osadetz, K.G. Relation between climate variability and groundwater levels in the upper carbonate aquifer, southern Manitoba, Canada. *J. Hydrol.* **2004**, *290*, 43–62. [CrossRef]
4. Lee, L.; Lawrence, D.; Price, M. Analysis of water-level response to rainfall and implications for recharge pathways in the Chalk aquifer, SE England. *J. Hydrol.* **2006**, *330*, 604–620. [CrossRef]
5. Nayak, P.C.; Rao, Y.R.S.; Sudheer, K.P. Groundwater level forecasting in a shallow aquifer using artificial neural network approach. *Water Resour. Manag.* **2006**, *20*, 77–90. [CrossRef]
6. Yoon, H.; Jun, S.C.; Hyun, Y.; Bae, G.O.; Lee, K.K. A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer. *J. Hydrol.* **2011**, *396*, 128–138. [CrossRef]
7. Tsai, W.P.; Chiang, Y.M.; Huang, J.L.; Chang, F.J. Exploring the Mechanism of Surface and Ground Water through Data-Driven Techniques with Sensitivity Analysis for Water Resources Management. *Water Resour. Manag.* **2016**, *30*, 4789–4806. [CrossRef]
8. Chen, T.S.; Jan, C.D.; Chen, S.H.; Tseng, C.M. Correlation between Groundwater Level Variations and Rainfall Depths. *Taiwan Water Conserv.* **2005**, *53*, 1–12. (In Chinese)
9. Chen, T.S. Study on Rainfall-Induced Groundwater Water Level Variation–Case study at the Naba, Liujar, and Donher Well Stations. Ph.D. Thesis, National Cheng Kung University, Tainan, Taiwan, 2006. (In Chinese)
10. Chang, F.J.; Kao, L.S.; Kuo, Y.M.; Liu, C.W. Artificial neural networks for estimating regional arsenic concentrations in a blackfoot disease area in Taiwan. *J. Hydrol.* **2010**, *388*, 65–76. [CrossRef]
11. Chang, L.C.; Chen, P.A.; Chang, F.J. Reinforced two-step-ahead weight adjustment technique for online training of recurrent neural networks. *IEEE Trans. Neural Netw. Learn. Syst.* **2012**, *23*, 1269–1278. [CrossRef]
12. Pradhan, B.; Lee, S. Regional Landslide Susceptibility Analysis Using Back-Propagation Neural Network Model at Cameron Highland, Malaysia. *Landslides* **2010**, *7*, 13–30. [CrossRef]
13. Karthikeyan, L.; Kumar, D.N.; Gaur, S. Prediction of Ground Water Levels in the Uplands of a Tropical Coastal Riparian Wetland Using Artificial Neural Networks. *Water Resour. Manag.* **2013**, *27*, 871–883. [CrossRef]
14. Sreekanth, P.D.; Geethanjali, N.; Sreedevi, P.D.; Ahmed, S.; Kumar, N.R.; Jayanthi, P.D.K. Forecasting Groundwater Level Using Artificial Neural Networks. *Curr. Sci.* **2009**, *96*, 933–939.

15. Li, X.; Shu, L.; Liu, L.; Yin, D.; Wen, J. Sensitivity analysis of groundwater level in Jinci Spring Basin (China) based on artificial neural network modeling. *Hydrogeol. J.* **2012**, *20*, 727–738. [CrossRef]
16. Mohanty, S.; Jha, M.K.; Kumar, A.; Panda, D. Comparative evaluation of numerical model and artificial neural network for simulating groundwater flow in Kathajodi–Surua Inter-basin of Odisha, India. *J. Hydrol.* **2013**, *495*, 38–51. [CrossRef]
17. Chiang, Y.M.; Chang, L.C.; Tsai, M.J.; Wang, Y.F.; Chang, F.J. Auto-control of pumping operations in sewerage systems by rule-based fuzzy neural networks. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 185–196. [CrossRef]
18. Talei, A.; Chua, L.H.C.; Quek, C. A novel application of a neuro-fuzzy computational technique in event-based rainfall–runoff modeling. *Expert Syst. Appl.* **2010**, *37*, 7456–7468. [CrossRef]
19. Pearson, K. Mathematical contributions to the theory of evolution. III.; Regression, Heredity and Panmixia. *Philos. Trans. R. Soc. Lond. Ser.* **1896**, *187*, 253–318. [CrossRef]
20. Stefánsson, A.; Končar, N.; Jones, A.J. A note on the Gamma test. *Neural Comput. Appl.* **1997**, *5*, 131–133. [CrossRef]
21. Končar, N. Optimisation Methodologies for Direct Inverse Neurocontrol. Ph.D. Thesis, University of London, London, UK, 1997.
22. Moghaddamnia, A.; Gousheh, M.G.; Piri, J.; Amin, S.; Han, D. Evaporation estimation using artificial neural networks and adaptive neuro-fuzzy inference system techniques. *Adv. Water Resour.* **2009**, *32*, 88–97. [CrossRef]
23. Noori, R.; Hoshyaripour, G.; Ashrafi, K.; Araabi, B.N. Uncertainty analysis of developed ANN and ANFIS models in prediction of carbon monoxide daily concentration. *Atmos. Environ.* **2010**, *44*, 476–482. [CrossRef]
24. Rumelhart, D.; Hinton, G.; Williams, R. Learning Internal Representations by Error Propagation. *Read. Cogn. Sci.* **1988**, *323*, 399–421.
25. Jang, J.S. ANFIS: Adaptive-network-based fuzzy inference system. *IEEE Trans. Syst. Man Cybern.* **1993**, *23*, 665–685. [CrossRef]
26. Altug, S.; Chen, M.Y.; Trussell, H. Fuzzy inference systems implemented on neural architectures for motor fault detection and diagnosis. *IEEE Trans. Ind. Electron.* **1999**, *46*, 1069–1079. [CrossRef]
27. Oonsivilai, A.; El-Hawary, M.E. Power System Dynamic Load Modeling Using Adaptive-Network-Based Fuzzy Inference System. In Proceedings of the Engineering Solutions for the Next Millennium 1999 IEEE Canadian Conference on Electrical and Computer Engineering (Cat. No. 99TH8411), Edmonton, AB, Canada, 9–12 May 1999; pp. 1217–1222.
28. Liao, H.P.; Su, J.P.; Wu, H.M. An application of ANFIS to modeling of a forecasting system for the demand of teacher human resources. *J. Educat. Psychol.* **2001**, *24*, 1–17.
29. Chang, L.C.; Chang, F.J. Intelligent control for modelling of real-time reservoir operation. *Hydrol. Process.* **2001**, *15*, 1621–1634. [CrossRef]
30. Takagi, T.; Sugeno, M. Fuzzy Identification of Systems and Its Applications to Modeling and Control. *Read. Fuzzy Sets Intell. Syst.* **1993**, *15*, 387–403.
31. Chang, F.-J.; Chang, Y.-T. Adaptive neuro-fuzzy inference system for prediction of water level in reservoir. *Adv. Water Resour.* **2006**, *29*, 1–10. [CrossRef]
32. Chang, W.Y. Investigating the Mechanisms of Groundwater Level Variation and Recharge at Zhuoshui River Basin. Master's Thesis, National Taiwan University, Taipei, Taiwan, 2013. (In Chinese)
33. Tsai, J.P.; Chen, Y.W.; Chang, L.C.; Kuo, Y.M.; Tu, Y.H.; Pan, C.C. High Recharge Areas in the Choushui River Alluvial Fan (Taiwan) Assessed from Recharge Potential Analysis and Average Storage Variation Indexes. *Entropy* **2015**, *17*, 1558–1580. [CrossRef]
34. Thiessen, A.H. Precipitation averages for large areas. *Mon. Weather. Rev.* **1911**, *39*, 1082–1089. [CrossRef]
35. Chang, F.J.; Lin, C.H.; Chang, K.C.; Kao, Y.H.; Chang, L.C. Investigating the Interactive Mechanisms between Surface Water and Groundwater over the Jhuoshuei River Basin in Central Taiwan. *Paddy Water Environ.* **2014**, *12*, 365–377. [CrossRef]