Article

How Much Meteorological Information Is Necessary to Achieve Reliable Accuracy for Rainfall Estimations?

Mohammad Valipour

Young Researchers and Elite Club, Kermanshah Branch, Islamic Azad University, Kermanshah, Iran; vali-pour@hotmail.com

Abstract: This paper reports the study of the effect of the length of the recorded data used for monthly rainfall forecasting. Monthly rainfall data for three periods of 5, 10, and 49 years were collected from Kermanshah, Mashhad, Ahvaz, and Babolsar stations and used for calibration time series models. Then, the accuracy of the forecasting models was investigated by the following year's data. The following was concluded: In temperate and semi-arid climates, 60 observation data is sufficient for the following year's rainfall forecasting. The accuracy of the time series models increased with increasing amounts of observation data of arid and humid climates. Time series models are appropriate tools for forecasting monthly rainfall forecasting in semi-arid climates. Determining the most critical rainfall month in each climate condition for agriculture schedules is a recommended aim for future studies.

Keywords: Iran; weather data; precipitation

1. Introduction

More accurate forecasting of monthly rainfall is significantly important in water resource management and crop pattern design. In this study, time series models were used to forecast monthly rainfall in four different climates. After publishing the paper of Box and Jenkins [1], Box–Jenkins models became a general time series model of hydrological forecasting. These models include the Auto Regressive Integrated Moving Average (ARIMA), the AR Moving Average (ARMA), the Auto Regressive (AR), and the Moving Average (MA). Access to basic information requires integration from the series (for a continuous series) or calculating all differences in the series (for a continuous series). Since the constant of integration in derivation or differences are deleted, the probability of using accurate amount in this process is not possible. Therefore, ARIMA models are non-static and cannot be used to reconstruct the missing data. However, these models are very useful for forecasting changes in hydrological processes considering agricultural water management, evapotranspiration, and water crisis issues [2–17]. Models of time series analysis (Box–Jenkins models) in various fields of hydrology and rainfall forecasting in irrigation schedules are widely applied, some of which will be described in what follows.

Serinaldi and Kilsby [18] presented a modular class of multisite monthly rainfall generators for water resource management and impact studies. The results of the case study show that the model can capture several characteristics of the rainfall series. In particular, it enables the simulation of low and high rainfall scenarios more extreme than those observed as well as the reproduction of the distribution of the annual accumulated rainfall, and of the relationship between the rainfall and circulation indices such as the North Atlantic Oscillation (NAO) and the Sea Surface Temperature (SST), thus proving the framework to be well-suited for sensitivity analysis under alternative climate scenarios and additional

forcing variables. Luc et al. [19] studied an application of artificial neural networks (ANNs) for rainfall forecasting successfully. Wei et al. [20] used weather satellite imagery to forecast rainfall in Taiwan. Andrieu et al. [21] studied the adaptation and application of a quantitative rainfall forecasting model in a mountainous region. This work showed that a limit on forecast lead-time may be related to the response time of the precipitating cloud system. Burlando et al. [22], using ARMA models, forecasted short-term rainfall. The hourly rainfall parameter from two gaging stations in Colorado, USA, and from several stations in Central Italy, was used. Results showed that the event-based estimation approach yields better forecasts. Hu et al. [23] studied rainfall, mosquito density, and the transmission of the Ross River virus using a time-series forecasting model. Their results showed that both rainfall and mosquito density were strong predictors of Ross River virus transmission in simple models. Ramírez et al. [24] used an ANN technique for rainfall forecasting applied to the São Paulo region. The results showed that ANN forecasts were superior to the ones obtained by the linear regression model, thus revealing a great potential for an operational suite. Han et al. [25] successfully forecasted drought based on remote sensing data using ARIMA model. Chattopadhyay and Chattopadhyay [26] compared ARIMA and ARNN models using univariate modeling of summer-monsoon rainfall time series. Anctil et al. [27] studied the impact of the length of observed records on the performance of ANN and of conceptual parsimonious rainfall-runoff forecasting models. The results showed that the best performance about evenly for 3- and 5-year training sets, but multiple-layer perceptrons (MLPs) performed better whenever the training set was dominated by wet weather. The MLPs continued to improve for input vectors of 9 years and more, which was not the case of the conceptual model. Jia and Culver [28] using bootstrapped ANNs suggested that even a small set of periodic instantaneous observations of stage from a staff gauge, which can easily be collected by volunteers, can be a useful data set for effective hydrological modeling. M. Baareh et al. [29] applied the ANN and AR models to the river flow forecasting problem. A comparative study of both ANN and the AR conventional model networks indicated that the ANNs performed better than the AR model. They showed that ANN models can be used to train and forecast the daily flows of the Black Water River near Dendron in Virginia and the Gila River near Clifton in Arizona. Xiong and O'Connor [30] used four different error-forecast updating models—AR, AR-threshold (AR-TS), fuzzy AR-threshold (FU-AR-TS), and an ANN—for real-time river flow forecasting. They found that all of these four updating models are very successful in improving the flow forecast accuracy. Chenoweth et al. [31] estimated the ARMA model parameters using neural networks. Their results showed that the ability of neural networks to accurately identify the order of an ARMA model was much lower than reported by previous researchers, and is especially low for time series with fewer than 100 observations. Using forecasting of hydrologic time series with ridge regression in feature space, Yu and Liong [32] showed that the training speed in the data mining method was much faster than the ARIMA model. See and Abrahart [33] used data fusion for hydrological forecasting. Their results showed that, using data fusion methodologies for ANN, fuzzy logic, and ARMA models, accuracy of forecasting increased. Using hybrid approaches, Srinivas and Srinivasan [34] improved the accuracy of AR model parameters for annual streamflows. Using Fourier coefficients, Ludlow and Enders [35] estimated the ARMA model parameters with relatively good accuracy. Chenoweth et al. [36] estimated the ARMA model parameters using Hilbert coefficients. Their results showed that Hilbert coefficients are a useful tool for estimating ARMA model parameters. Balaguer et al. [37] used the time delay neural network (TDNN) and ARMA model in support centers for crisis management. The obtained correlation results for TDNN model and ARMA were 0.88 and 0.97, respectively. This study confirmed the superiority of the ARMA model to the TDNN. Toth et al. [38] used the ANN and ARMA models to forecast rainfall. The results show the success of both short-term rainfall-forecasting models for forecasting floods in real time. Mohammadi et al. [39] forecasted Karaj reservoir inflow using data of melting snow and ANN and ARMA methods, as well as regression analysis. Sixty percent of inflow in dams occurs between April and June, so forecasting the inflow in this season is very important for a dam's performance. The highest inflows were in the spring due to the snow melt caused by

draining in threshold winter. The results showed that an ANN, compared with other methods, has lower significant errors. Mohammadi et al. [40] in another study estimated the parameters of an ARMA model for river flow forecasting using goal programming. Their results showed that the goal programming is a precise and effective method for estimating ARMA model parameters for forecasting inflow. Other researchers [41–48] estimated the parameters of the ARMA and ARIMA models and compared their ability for inflow forecasting. By comparing root mean square errors of the model, it was determined that the ARIMA model can forecast inflow to the Dez reservoir from 12 months with a lower error than can the ARMA model.

Therefore, considering the above-mentioned research, the efficacy of time series models forecasting field and hydrologic sampling can be determined. The effect of recorded data in ANN forecasting has been previously studied. However, concurrent use of time series models to determine the number of required observation data has not been performed. This study aims to forecast monthly rainfall using time series models and determine appropriate observation data according to different climate conditions.

2. Material and Methods

In this study, to investigate the effect of climate conditions on the amount of appropriate observation data, four synoptic stations with different climate conditions in Iran, including temperate, humid, arid, and semi-arid were selected.

Table 1 displays information on the synoptic stations used in this study. In order to forecast rainfall on a monthly basis, the rainfall data period from 1951 to 2000 was gathered. In fact, the used data involved 2400 data (all stations) that began in January 1951 and ended in December 2000. Figure 1 shows the locations of the synoptic stations. In this study, time series models were used to forecast monthly rainfall. These models include ARIMA, ARMA, AR, and MA. For this purpose, MINITAB software was used to run all time series models. Finally, via three methods, the ability of time series models to forecast monthly rainfall was surveyed. In the first method, 60 data (1995–1999) regarding the amount of rainfall forecasted for the year 2000 was used. In the second method, 10 years of monthly rainfall data (1990–1999) regarding the amount of rainfall forecasted for the year 2000 was used. In the third method, 588 data (1951–1999) regarding the amount of rainfall forecasted for the year 2000 was used.

Table 1. Information of synoptic stations used in this research.

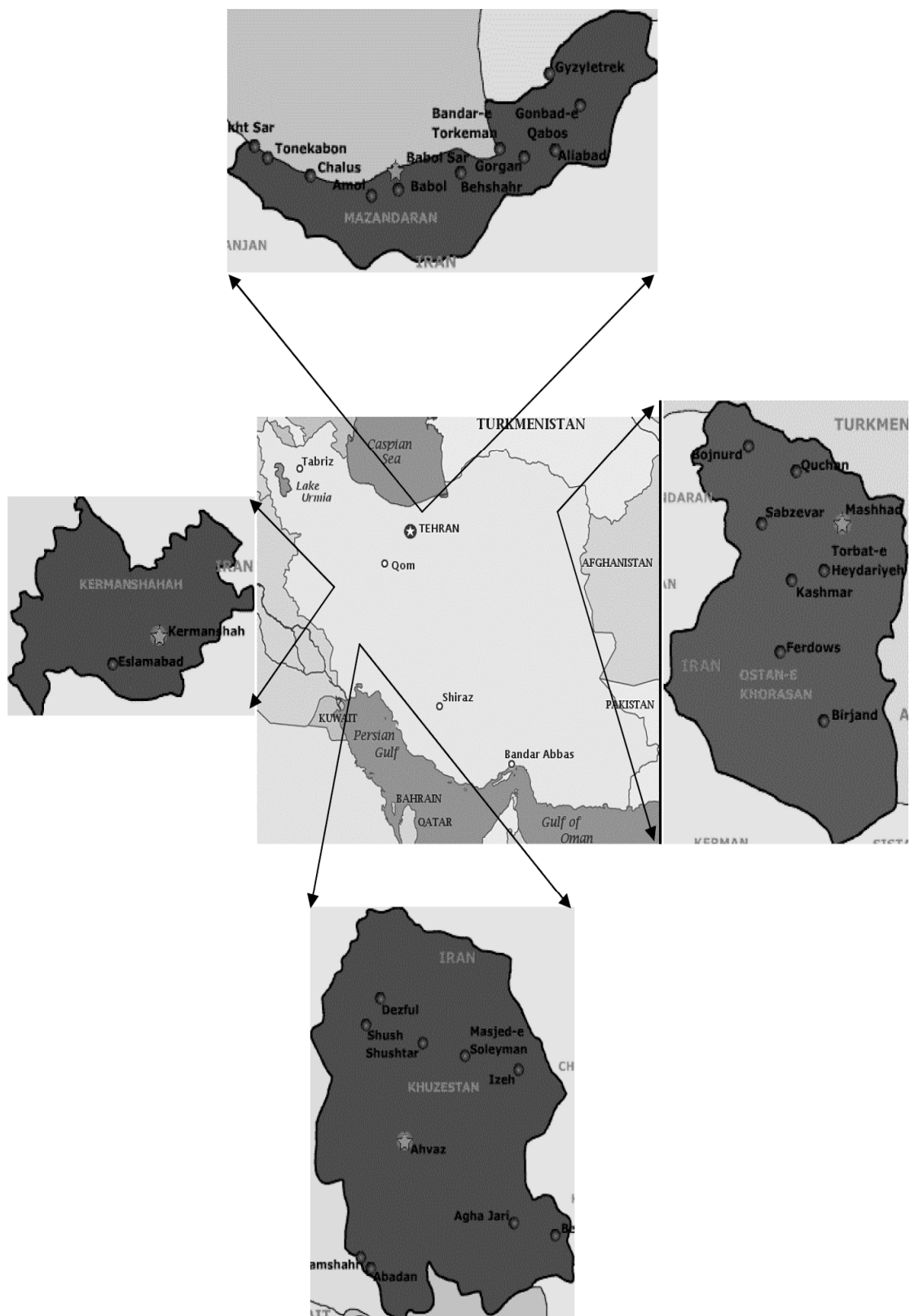| Station	Climate	Elevation Related to Sea Level (m)	Longitude	Altitude
Kermanshah	Temperate	999.2	59°38′ E	36°16′ N
Mashhad	Semi-arid	1322.0	47°7′ E	34°17′ N
Ahvaz	Arid	22.5	48°40′ E	31°20′ N
Babolsar	Humid	−21.0	52°39′ E	36°43′ N

Figure 1. Position of synoptic stations in Iran.

3. Results and Discussion

Table 2 shows the obtained R^2 for four different stations according to the amount of observation data. Figures 2–5 show the amount of forecasted rainfall for Kermanshah, Mashhad, Ahvaz, and Babolsar synoptic stations, respectively.

Table 2. Obtained R^2 for four different stations according to the amount of observation data.

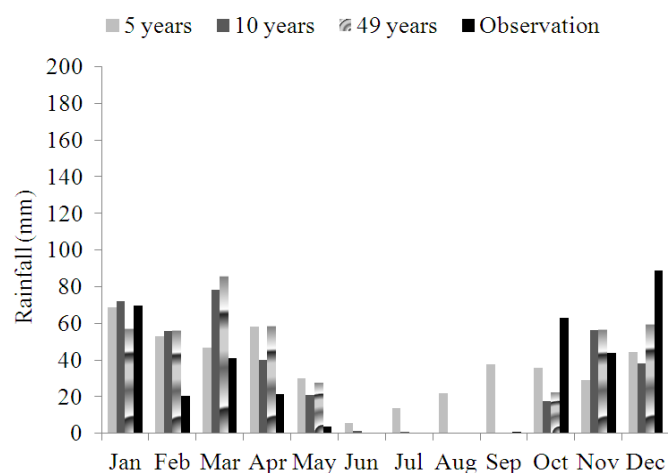| Station	5 Years (60 Data)	10 Years (120 Data)	49 Years (588 Data)
Kermanshah	0.81	0.81	0.81
Mashhad	0.96	0.96	0.96
Ahvaz	0.77	0.83	0.87
Babolsar	0.70	0.73	0.82

Figure 2. Amount of forecasted rainfall for the Kermanshah synoptic station.
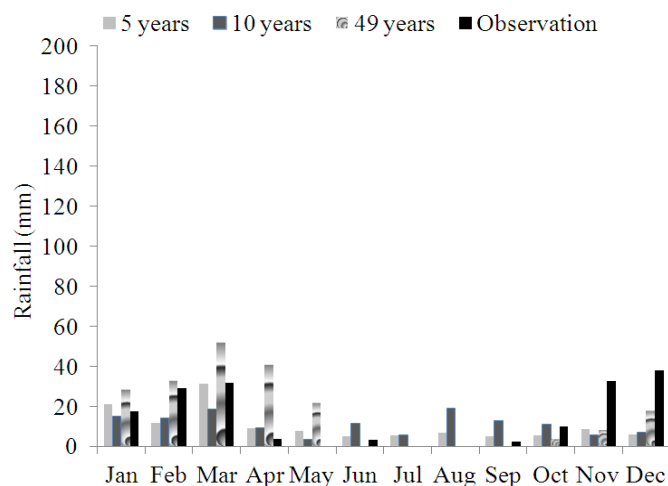

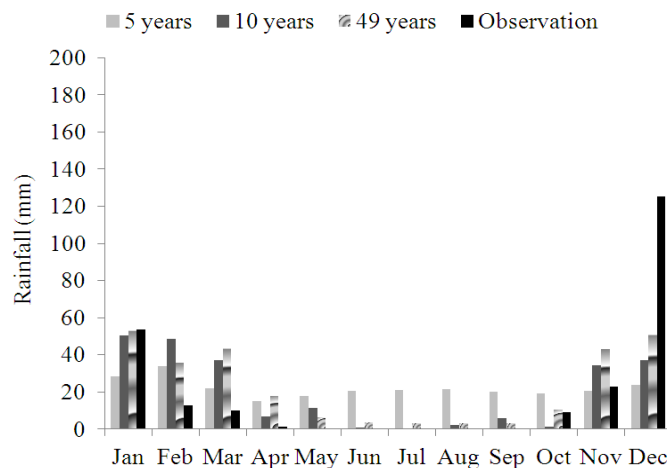**Figure 3.** Amount of forecasted rainfall for the Mashhad synoptic station.

Figure 4. Amount of forecasted rainfall for the Ahvaz synoptic station.
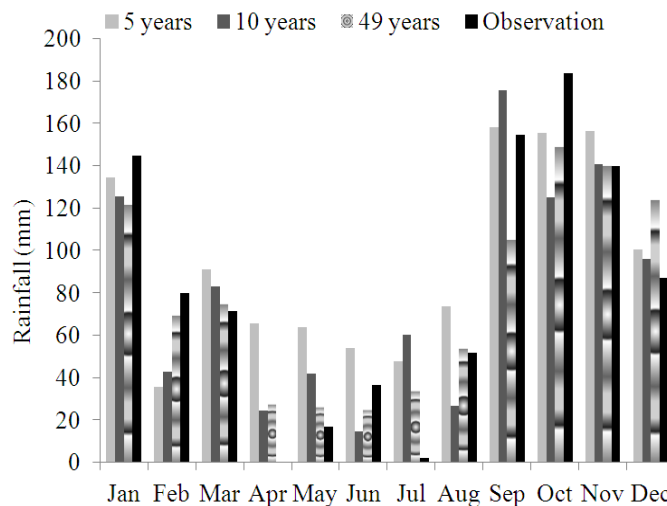

Figure 5. Amount of forecasted rainfall for the Babolsar synoptic station.

According to the obtained results for R^2 in Table 2, in a temperate climate (Kermanshah), due to low changes of monthly rainfall over the long term, 60 observation data was sufficient for forecasting. In this station, the R^2 was 0.81, and an increase in observation data is not recommend.

In a semi-arid climate (Mashhad), due to the low monthly rainfall relative to the Kermanshah and Babolsar stations and the lack of torrent precipitation in arid climates (for example, December in Figure 4), the highest R^2 was obtained.

In the Mashhad station, data from 5 years prior was sufficient for forecasting, and this is a good result for deficit data situations. However, according to Table 2 and Figures 4 and 5, with 60 observation data, the accuracy of forecasted rainfall was not ideal.

The Ahvaz station, with an arid climate due to drought periods and torrent precipitation, needs at least 588 data for accurate forecasting ($R^2 = 0.87$ rather than $R^2 = 0.77$).

The Babolsar station, with a humid climate due to flood periods in some years, depends on an increase in observation data if rainfall forecasting is to improve in accuracy. A difference in R^2 from 49 years of observation data accrued in the Babolsar synoptic station, compared with only 5 years of observation data ($R^2 = 0.82$ rather than $R^2 = 0.70$) was observed.

4. Conclusions

This paper reports the study of the effect of the length of the recorded data for monthly rainfall forecasting. Monthly rainfall data for three periods of 5, 10, and 49 years were collected from Kermanshah, Mashhad, Ahvaz, and Babolsar stations and used for calibration time series models. Then, the accuracy of forecasting models was investigated with the following year's data. The following was concluded:

1　In temperate and semi-arid climates, 60 observation data is sufficient for the following year's rainfall forecasting.
2　The accuracy of the time series models increased with increasing amounts of observation data of arid and humid climates.
3　Time series models are appropriate tools for forecasting monthly rainfall forecasting in semi-arid climates.
4　Determining the most critical rainfall month in each climate condition for agriculture schedules is a recommended aim for future studies.

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

References

1. Box, G.E.P.; Jenkins, G.M. *Series Analysis Forecasting and Control*; Prentice-Hall Inc.: London, UK, 1976.
2. Khoshravesh, M.; Gholami Sefidkouhi, M.A.; Valipour, M. Estimation of reference evapotranspiration using multivariate fractional polynomial, Bayesian regression, and robust regression models in three arid environments. *Appl. Water Sci.* **2015**. [CrossRef]
3. Valipour, M.; Singh, V.P. Global Experiences on Wastewater Irrigation: Challenges and Prospects. In *Balanced Urban Development: Options and Strategies for Liveable Cities*; Maheshwari, B., Singh, V.P., Thoradeniya, B., Eds.; Springer: Bern, Switzerland, 2016; pp. 289–327.
4. Valipour, M.; Gholami Sefidkouhi, M.A.; Raeini-Sarjaz, M. Selecting the best model to estimate potential evapotranspiration with respect to climate change and magnitudes of extreme events. *Agric. Water Manag.* **2017**. [CrossRef]
5. Valipour, M.; Gholami Sefidkouhi, M.A.; Khoshravesh, M. Estimation and trend evaluation of reference evapotranspiration in a humid region. *Ital. J. Agrometeorol.* **2017**, in press.
6. Valipour, M. Future of agricultural water management in Africa. *Arch. Agron. Soil Sci.* **2015**, *61*, 907–927. [CrossRef]
7. Valipour, M. Calibration of mass transfer-based models to predict reference crop evapotranspiration. *Appl. Water Sci.* **2015**. [CrossRef]
8. Valipour, M. Analysis of potential evapotranspiration using limited weather data. *Appl. Water Sci.* **2015**. [CrossRef]
9. Valipour, M. Increasing irrigation efficiency by management strategies: Cutback and surge irrigation. *ARPN J. Agric. Biol. Sci.* **2013**, *8*, 35–43.
10. Valipour, M. Necessity of Irrigated and Rainfed Agriculture in the World. *Irrig. Drain. Syst. Eng.* **2013**, *S9*, e001. [CrossRef]
11. Valipour, M. Evolution of Irrigation-Equipped Areas as Share of Cultivated Areas. *Irrig. Drain. Syst. Eng.* **2013**, *2*, e114. [CrossRef]
12. Valipour, M. Use of surface water supply index to assessing of water resources management in Colorado and Oregon, US. *Adv. Agric. Sci. Eng. Res.* **2013**, *3*, 631–640.
13. Valipour, M. Hydro-Module Determination for Vanaei Village in Eslam Abad Gharb, Iran. *ARPN J. Agric. Biol. Sci.* **2012**, *7*, 968–976.
14. Valipour, M. Ability of Box–Jenkins Models to Estimate of Reference Potential Evapotranspiration (A Case Study: Mehrabad Synoptic Station, Tehran, Iran). *IOSR J. Agric. Vet. Sci.* **2012**, *1*, 1–11. [CrossRef]
15. Valipour, M. A Comparison between Horizontal and Vertical Drainage Systems (Include Pipe Drainage, Open Ditch Drainage, and Pumped Wells) in Anisotropic Soils. *IOSR J. Mech. Civil Eng.* **2012**, *4*, 7–12. [CrossRef]

16. Valipour, M. Application of new mass transfer formulae for computation of evapotranspiration. *J. Appl. Water Eng. Res.* **2014**, *2*, 33–46. [CrossRef]

17. Valipour, M.; Montazar, A.A. An Evaluation of SWDC and WinSRFR Models to Optimize of Infiltration Parameters in Furrow Irrigation. *Am. J. Sci. Res.* **2012**, *69*, 128–142.

18. Serinaldi, F.; Kilsb, C.G. A modular class of multisite monthly rainfall generators for water resource management and impact studies. *J. Hydrol.* **2012**, *464*, 528–540. [CrossRef]

19. Luc, K.C.; Ball, J.E.; Sharma, A. An application of artificial neural networks for rainfall forecasting. *Math. Comput. Model.* **2001**, *33*, 683–693.

20. Ludlow, J.; Enders, W. Estimating non-linear ARMA models using Fourier Coefficients. *Int. J. Forecast.* **2000**, *16*, 333–347. [CrossRef]

21. Baareh, A.K.M.; Sheta, A.F.; Khnaifes, K.A. Forecasting River Flow in the USA: A Comparison between Auto Regressive and Neural Network Non-Parametric Models. *J. Comput. Sci.* **2006**, *2*, 775–780. [CrossRef]

22. Mohammadi, K.; Eslami, H.R.; Dardashti, S.D. Comparison of Regression ARIMA and ANN Models for Reservoir Inflow Forecasting using Snowmelt Equivalent (A Case Study of Karaj). *J. Agric. Sci. Technol.* **2005**, *7*, 17–30.

23. Mohammadi, K.; Eslami, H.R.; Kahawita, R. Parameter estimation of an ARMA model for river flow forecasting using goal programming. *J. Hydrol.* **2006**, *331*, 293–299. [CrossRef]

24. Ramírez, M.C.V.; de Campos Velho, H.F.; Ferreira, N.J. Artificial neural network technique for rainfall forecasting applied to the São Paulo region. *J. Hydrol.* **2005**, *301*, 146–162. [CrossRef]

25. Rezaei, M.; Valipour, M.; Valipur, M. Modelling evapotranspiration to increase the accuracy of the estimations based on the climatic parameters. *Water Conserv. Sci. Eng.* **2016**. [CrossRef]

26. See, L.; Abrahat, R.J. Multi-model data fusion for hydrological forecasting. *Comput. Geosci.* **2001**, *27*, 987–994. [CrossRef]

27. Srinivas, V.V.; Srinivasan, K. Post-blackening approach for modeling dependent annual streamflows. *J. Hydrol.* **2000**, *230*, 86–126. [CrossRef]

28. Toth, E.; Brath, A.; Montanari, A. Comparison of short-term rainfall predication models for real-time flood forecasting. *J. Hydrol.* **2000**, *239*, 132–147. [CrossRef]

29. Valipour, M. Long-term runoff study using SARIMA and ARIMA models in the United States. *Meteorol. Appl.* **2015**, *22*, 592–598. [CrossRef]

30. Valipour, M. Temperature analysis of reference evapotranspiration models. *Meteorol. Appl.* **2015**, *22*, 385–394. [CrossRef]

31. Valipour, M. Use of average data of 181 synoptic stations for estimation of reference crop evapotranspiration by temperature-based methods. *Water Resour. Manag.* **2014**, *28*, 4237–4255. [CrossRef]

32. Valipour, M. Evaluation of radiation methods to study potential evapotranspiration of 31 provinces. *Meteorol. Atmos. Phys.* **2015**, *127*, 289–303. [CrossRef]

33. Valipour, M.; Eslamian, S. Analysis of potential evapotranspiration using 11 modified temperature-based models. *Int. J. Hydrol. Sci. Technol.* **2014**, *4*, 192–207. [CrossRef]

34. Anctil, F.; Perrin, C.; Andréassian, V. Impact of the length of observed records on the performance of ANN and of conceptual parsimonious rainfall-runoff forecasting models. *Environ. Model. Softw.* **2004**, *19*, 357–368. [CrossRef]

35. Andrieu, H.; French, M.N.; Thauvin, V.; Krajewski, W.F. Adaptation and application of a quantitative rainfall forecasting model in a mountainous region. *J. Hydrol.* **1996**, *184*, 243–259. [CrossRef]

36. Balaguer, E.; Palomares, A.; Sorie, E.; Martin-Guerrero, J.D. Predicting service request in support centers based on nonlinear dynamics, ARMA modeling and neural networks. *Expert Syst. Appl.* **2008**, *34*, 665–672. [CrossRef]

37. Burlando, C.; Rosso, R.; Cadavid, L.G.; Salas, J.D. Forecasting of short-term rainfall using ARMA models. *J. Hydrol.* **1993**, *144*, 193–211. [CrossRef]

38. Chattopadhyay, S.; Chattopadhyay, G. Univariate modelling of summer-monsoon rainfall time series: Comparison between ARIMA and ARNN. *C. R. Geosci.* **2010**, *342*, 100–107. [CrossRef]

39. Chenoweth, T.; Dowling, K.; Hubata, R.; Louis, R. Automatic ARMA identification using neural networks and the extended sample autocorrelation function: A reevaluation. *Decis. Support Syst.* **2004**, *29*, 21–30. [CrossRef]

40. Chenoweth, T.; Dowling, K.; Hubata, R.; Louis, R. Distance and prediction error variance constraints for ARMA model portfolios. *Int. J. Forecast.* **2004**, *20*, 41–52. [CrossRef]

41. Han, P.; Wang, P.X.; Zhang, S.Y.; Zhu, D.H. Drought forecasting based on the remote sensing data using ARIMA models. *Math. Comput. Model.* **2010**, *51*, 1398–1403. [CrossRef]

42. Hu, W.; Tong, S.; Mengersen, K.; Oldenburg, B. Rainfall, mosquito density and the transmission of Ross River virus: A time-series forecasting model. *Ecol. Model.* **2006**, *196*, 505–514. [CrossRef]

43. Jia, Y.; Culver, T.B. Bootstrapped artificial neural networks for synthetic flow generation with a small data sample. *J. Hydrol.* **2006**, *331*, 580–590. [CrossRef]

44. Karamouz, M.; Araghinejad, S.H. *Advance Hydrology*; Amirkabir University of Technology Press: Tehran, Iran, 2012.

45. Wei, C.; Hung, W.C.; Cheng, K.S. A multi-spectral spatial convolution approach of rainfall forecasting using weather satellite imagery. *Adv. Space Res.* **2006**, *37*, 747–753. [CrossRef]

46. Xiong, L.; O'Connor, K.M. Comparison of four updating models for real-time river flow forecasting. *Hydrol. Sci. J.* **2002**, *47*, 621–639. [CrossRef]

47. Yannopoulos, S.I.; Lyberatos, G.; Theodossiou, N.; Li, W.; Valipour, M.; Tamburrino, A.; Angelakis, A.N. Evolution of Water Lifting Devices (Pumps) over the Centuries Worldwide. *Water* **2015**, *7*, 5031–5060. [CrossRef]

48. Yu, X.; Liong, S. Forecasting of hydrologic time series with ridge regression in feature space. *J. Hydrol.* **2006**, *332*, 290–302. [CrossRef]