

Research Article

# Generation of Hourly Precipitation Time Series Incorporating Extreme Values Under Data Limitations

Mohammad Reza Khazaei<sup>a</sup>, Vahid Zarezadeh<sup>a</sup>, Bagher Zahabiyoun<sup>b</sup>

<sup>a</sup> Department of Civil Engineering, Payame Noor University, Tehran, Iran.

<sup>b</sup> School of Civil Engineering, Iran University of Science and Technology, Tehran, Iran

## ARTICLE INFO

**Received date:** 22 Jan 2025  
**Accept date:** 12 May 2025  
**Published date:** 18 May 2025

## Keywords:

Data Limitations, Extreme Precipitation, Hourly Precipitation Series, Nsrp, Stochastic Modeling

## Abstract

To simulate floods in small basins, precipitation data with fine temporal resolution is necessary. Especially in assessing the impact of climate change on flooding, it is essential to have a continuous precipitation series in short time steps that includes extreme values. However, precipitation data is often recorded daily, and long-term hourly precipitation series are usually unavailable. In such cases, generating long-term precipitation series with small time steps that incorporate extreme values becomes critical. This study proposes and validates a method for generating long-term 6-hour precipitation series when the available 6-hour series is short, but long-term daily precipitation series is available. The proposed method utilizes an empirical relationship between hourly and daily precipitation statistics over a short-term period to estimate long-term hourly precipitation statistics from long-term daily statistics. Then, using the generated hourly precipitation statistics and the Neyman-Scott Rectangular Pulse (NSRP) stochastic point process model, hourly precipitation series of any desired length can be produced. Using this approach, a long-term 6-hour precipitation series incorporating extreme values was generated and validated with high accuracy based on a 5-year 6-hour precipitation series and a 19-year daily precipitation series.

Homepage: [www.wss.torbath.ac.ir](http://www.wss.torbath.ac.ir)

\*Corresponding Author:

Khazaei, Mohammad Reza

Email: [khazaei@pnu.ac.ir](mailto:khazaei@pnu.ac.ir)



ORCID: 0000-0002-0911-101X



<https://doi.org/10.22048/WSS.2025.501475.1016>

## How to cite this article:

Khazaei, M. R. , Zarezadeh, V. and Zahabiyoun, B. (2025). Generation of Hourly Precipitation Time Series Incorporating Extreme Values Under Data Limitations.. *Journal of Advanced Informatics in Water, Soil, and Structure*, 1(1),164-173.



## Introduction

Access to long-term precipitation data with an acceptable temporal resolution is crucial for simulating the flood regime of basins, particularly small basins. Such data are essential for continuous hydrological modeling of watershed floods. One key aspect of flood simulation using continuous models is the accurate determination of soil moisture conditions before a flood event, which significantly influences the accuracy of simulated floods (Roy et al., 2001; Khazaei, 2025). Moreover, having access to long-term precipitation series with small time steps is vital for assessing the impact of climate change on floods in small basins. The lack of such data poses a significant challenge in evaluating the effects of climate change on flooding in many small basins (Khazaei et al., 2013; IPCC, 2012).

With stochastic rainfall modeling, it is possible to generate synthetic precipitation series of any desired length. Stochastic models are fitted to observed precipitation data and simulate long-term precipitation series that statistically is similar to the observed data (Khazaei, 2024; Khazaei, 2025; Rodriguez et al., 1987). However, for stochastic precipitation modeling in studies of small basin flooding, it is essential to generate precipitation series at short temporal intervals while accurately reproducing extreme precipitation events.

Most stochastic precipitation models are developed for daily time steps, and accurately simulating extreme values remains a significant challenge in stochastic precipitation modeling (Khazaei et al., 2012; Khazaei, 2025). Additionally, for reliable stochastic precipitation simulations, a minimum of 20 years of observed precipitation data is typically required (Semenov et al., 1998). However, precipitation data are often recorded at daily intervals, and in most cases, hourly precipitation series with sufficient statistical length are not available (Khazaei, 2025).

By simulating precipitation using the continuous point process approach, the challenge of temporal scale can be effectively addressed. The NSRP model (Neyman-Scott Rectangular Pulse), introduced by Rodriguez et al. (1987), is one of the continuous stochastic point process models that is fitted to precipitation statistics at a given time step and generates precipitation continuously, independent of the time scale. Consequently, the generated precipitation can be rescaled into any

desired temporal resolution (Burton et al., 2008).

The NSRP (Neyman-Scott Rectangular Pulse) model has significant capabilities for generating long-term synthetic precipitation series with any desired time step while preserving the characteristics of observed precipitation. Some of its notable features include (Khazaei et al., 2021; Kilsby et al., 2007): a) Continuous precipitation Generation: The NSRP model generates precipitation continuously, making it suitable for producing long-term precipitation series at any desired time step. A comprehensive study by Olsson and Burlando (2002) evaluated the capability of the NSRP model to reproduce precipitation characteristics across various time scales. They fitted the NSRP model to 20-minute precipitation data from Italy and demonstrated that the model successfully reproduced precipitation statistics across time scales ranging from 20 minutes to two weeks. b) Calibration Based on Precipitation Statistics: The NSRP model does not require direct precipitation data for calibration; instead, it is fitted to precipitation statistics, allowing it to generate precipitation at any desired time step (including shorter time steps than the observed data). However, using hourly precipitation statistics is more suitable for generating hourly series. For instance, Kilsby et al. (2004) combined daily and hourly precipitation statistics to calibrate the NSRP model and successfully reproduced hourly precipitation series while preserving extreme hourly events. This approach is particularly useful when hourly precipitation data are available for a short time period, as it allows for estimating the relationship between daily and hourly statistics. Using this relationship, the model can be calibrated with recorded daily data to produce long-term hourly series. Moreover, this method is crucial when working with future climate scenarios, where climate model outputs are often only available at a daily resolution (Kilsby et al., 2004). c) Reproduction of Extreme Events: The NSRP model has consistently demonstrated its ability to accurately reproduce extreme precipitation events in various studies (Khazaei et al., 2012; Kilsby et al., 2007; Cowpertwait et al., 2002; Khazaei, 2021; Khazaei, 2025).

In this study, the NSRP model was used to simulate precipitation for the North Tehran synoptic station. Initially, the model was fitted to a

19-year series of 6-hourly precipitation data, and its accuracy in generating a long-term (1,900-year) precipitation series, with correctly reproduced 6-hour extreme precipitation values, was evaluated. Additionally, the combination of precipitation statistics for producing the aforementioned series was optimized and introduced.

Subsequently, a method was proposed and validated for generating long-term 6-hourly precipitation series when the recorded 6-hourly precipitation data is short (5 years). The primary motivation for this research was to provide long-term hourly precipitation data for the current climate of a watershed in North Tehran. Such data would enable continuous rainfall-runoff simulation and the assessment of climate change impacts on floods in the watershed. In this region, only 5 years of hourly precipitation data is available, but a 30-year daily precipitation record was accessible.

## Methodology

### Case study

To simulate short-time-step precipitation series that include extreme precipitation events, the recorded precipitation data from the North Tehran Synoptic Station was utilized. This station was chosen due to the availability of a 19-year (1988–2006) 6-hour precipitation series. The station is located in Tehran Province at a longitude of

51°29'E, latitude of 35°48'N, and an elevation of 1,549 meters above sea level.

Initially, the NSRP rainfall model was fitted to the statistics of the 19-year 6-hourly precipitation series. Using the fitted model, 100 synthetic 19-year series (equivalent to 1,900 years) of 6-hourly precipitation were generated and validated. Subsequently, the feasibility of generating long-term 6-hourly precipitation series was assessed in a scenario where only 5 years of observed 6-hourly data and the remaining 19 years of daily data were available.

In the following, the NSRP model and methods are described in detail.

### NSRP Model

The **NSRP (Neyman-Scott Rectangular Pulse)** model is a stochastic precipitation simulation model that describes the physical structure of precipitation using stochastic methods. This model has demonstrated its ability to preserve the statistics of observed precipitation across a wide range of time scales and extreme events (Khazaei et al., 2012; Khazaei et al., 2020; Khazaei, 2024).

The structure of the NSRP model is described through the following steps and illustrated in Figure (1):

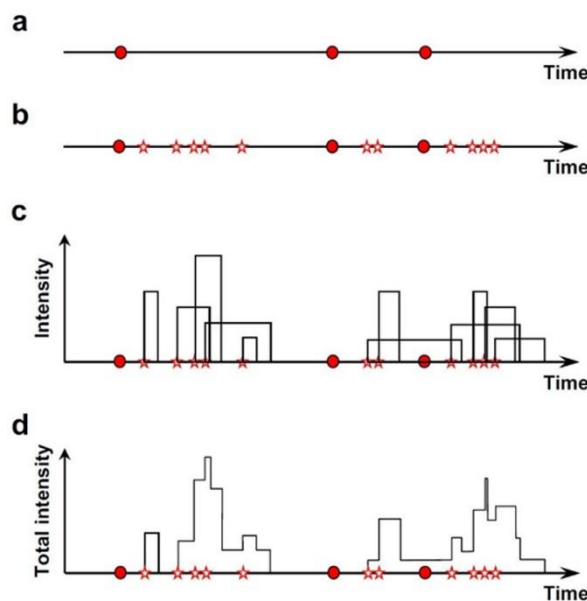


Figure 1. Schematic of the NSRP model (Burton et al., 2008)

- a) Storm origins occur independently following a Poisson process, characterized by an arrival rate denoted by the parameter  $\lambda$ .
- b) Each storm origin produces a Poisson-distributed random number  $C$  of raincells with a mean value of  $v$ . Each raincell follows the storm origin after a time interval that is independently drawn from an exponential distribution with rate parameter  $\beta$ .
- c) Each raincell has an independently and uniformly determined duration and intensity. The duration and intensity of each raincell follow exponential distributions with rate parameters  $\eta$  and  $\xi$ , respectively.
- d) At any given moment, the precipitation intensity is equal to the sum of the intensities of all raincells that are active at that time.

Since the process is continuous in time, a time series is obtained by discretizing it into daily intervals or any other desired time step. The NSRP model effectively captures both the continuous nature of precipitation and the statistical characteristics of observed data, making it a powerful tool for simulating long-term precipitation series.

Thus, the precipitation series can even be generated at smaller time scales than the observed data. The NSRP model has five parameters that are estimated through model calibration. For calibration, at least five statistics are calculated from the observed time series, and their values are equated to the corresponding model statistics, which are functions of the model parameters. This results in a system of equations, which is solved simultaneously to estimate the model parameters.

To preserve the seasonal characteristics of precipitation, these parameters are estimated separately for each month of the year (Burton et al., 2008; Khazaei et al., 2025).

## Model Evaluation

### Calibration of the NSRP Model

The NSRP model was calibrated to the 19-year observed 6-hourly precipitation series (1988–2006) recorded at the North Tehran synoptic station. Various combinations of precipitation statistics, with different weighting coefficients, were tested during the calibration process. As a result, the optimal combination of statistics and their corresponding weights was identified to achieve the best performance in simulating the target precipitation series.

Using the calibrated model, 100 synthetic 19-year 6-hourly precipitation series (matching the length of the observed data) were generated and

validated to ensure the model's accuracy.

### Model Validation

Typically, the performance of stochastic precipitation generation models is assessed by comparing the statistics of synthetic precipitation series with the corresponding observed precipitation values (Semenov et al., 1998; Kim et al., 2012; Chen et al., 2019; Khazaei, 2024; Khazaei, 2025). Due to the uncertainty of sampling variability, 100 synthetic 6-hour precipitation series were generated over 19 years. For each synthetic series, precipitation statistics and the distribution of annual maximum 6-hour precipitation were calculated. The 95 % confidence interval of the simulated statistics and the distribution of maximum precipitation were compared with their corresponding observed values. If the observed statistics fall within a confidence interval of the synthetic statistics, it indicates good model performance at that confidence level (Khazaei et al., 2012; Holman et al., 2009; Khazaei, 2025).

## Precipitation Generation

Since the NSRP model is fitted to precipitation statistics and generates precipitation continuously, it is possible to produce precipitation at any time step, including hourly, using daily precipitation statistics. However, the use of a combination of 6-hour and daily precipitation statistics is aimed at achieving greater accuracy in generating 6-hour precipitation. In fitting the NSRP model to the 19-year series of 6-hour precipitation, the necessary statistics for calibrating the model were determined, which is a combination of 6-hour and daily statistics.

To produce a long-term series of 6-hour precipitation using a 5-year series of 6-hour precipitation and 19 years of daily data, the following steps were taken: 1) Over a 5-year period, regression equations were fitted between 6-hour statistics (required for model calibration) and daily statistics. 2) Using the derived relationship, the 19-year daily statistics were converted into 19-year 6-hour statistics. 3) The NSRP model was calibrated using a combination of the observed daily statistics and the generated 6-hour statistics, resulting in the production of a long-term series of 6-hour precipitation. 4) For method validation, the 6-hour precipitation statistics derived from the long-term simulated precipitation were compared with the observed precipitation statistics over a 19-year period. Additionally, the characteristics of the precipitation generated by the NSRP model were

compared under the two following scenarios: (a) using 19 years of observed 6-hour precipitation data, and (b) using 19 years of daily precipitation data along with 5 years of observed 6-hour precipitation data. 5) To assess the uncertainty of sampling variability and evaluate the reliability of the results from the above method, the 19-year series was divided into 4 periods of 5 years each. Then, the above steps were repeated for four scenarios (assuming that only one of the 5-year periods of 6-hour precipitation is available each time), and the results were compared.

## Results

The NSRP model was calibrated for 19 years of observed 6-hourly precipitation data (1988 to 2006) at the synoptic station in northern Tehran. Through a trial and error process, the observed precipitation statistics were determined with optimal weights for model calibration. These statistics include daily mean (Mean24), 6-hourly and daily variance (Var24 and Var6), daily skewness (Skew24), the probability of a rainy days (Pw24), and the probability of a rainy day following a rainy day (Pww24) (Table 1).

**Table 1.** The assigned weight values for the statistics in the model.

statistics	Mean24	Var6	Var24	Pw24	Pww24	Skew24
weight	20	3	3	3	3	10

To validate the model, 100 series of 6-hour precipitation data spanning 19 years were generated, and the observed precipitation statistics were compared with 95% confidence interval of the generated series statistics. All precipitation statistics presented in Table (1) were well reproduced for all months of the year (the corresponding figures are not presented for brevity). The validation results for the distribution of annual maximum 6-hour precipitation are shown in Figure (2). The observed values fall within the 95% confidence interval of the simulated values, and the slope of the distributions is acceptably reproduced. However, the observed distribution of extreme precipitation is subject to uncertainty due to sampling variability (because of the limited length of the recorded data). For instance, the largest observed data point has a significant impact on the slope of the distribution; however, it is unclear whether the assigned empirical return period for this event is accurate. Additionally, measuring observed extreme events is often challenging and may be associated with errors (Boughton et al., 2006). Although the NSRP model is not directly fitted to extreme data (but rather to the observed precipitation statistics in Table 1), the results indicate a significant consistency between the generated extreme values distribution and the corresponding observed distribution at the 6-hour interval.

Among the statistics used for model calibration for generating 6-hour precipitation series, only the

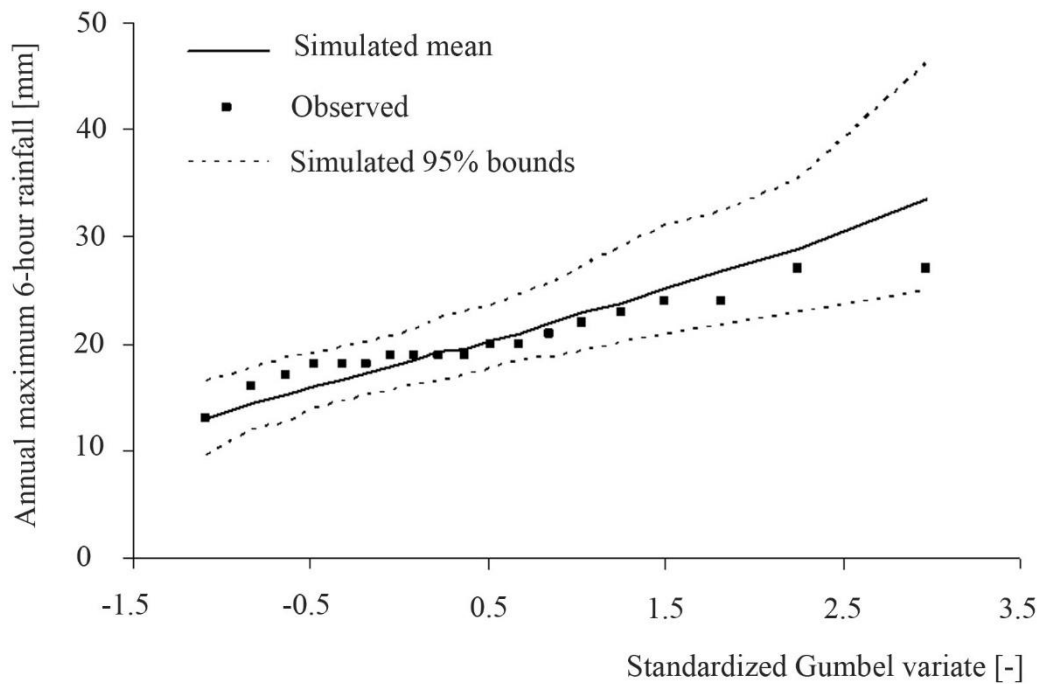
6-hour variance requires a 6-hour series; other statistics are calculated from daily precipitation. Therefore, if the long-term 6-hour variance can be obtained using the short-term 6-hour precipitation series and the long-term daily series, it is possible to simulate a 6-hour precipitation series of any length in situations where the length of the 6-hour data is short but long-term daily precipitation is available. To produce a long-term 6-hour series using 5 years of 6-hour data and 19 years of daily data, regression relationships were fitted between the 6-hour and daily variances. It is necessary for the relationships between the 6-hour and daily statistics in the 5-year period to be stable and generalizable to the 19-year period. To assess the uncertainty of sampling variability, these relationships were established for 4 different 5-year periods to compare the results. In Figure (3), the coefficient of determination ( $R^2$ ) of the linear equations for each month and each 5-year period is shown.

In most cases, the coefficient of determination is greater than 0.8, indicating a strong linear relationship between the 6-hour and daily variances. To assess the reliability of these relationships, the total 6-hour variance over the entire period (19 years) was calculated from the daily variance using the relationships obtained over 5-year periods. In Figure (4), the observed variances are compared with the upper and lower bounds of the estimated variances for the 19-year

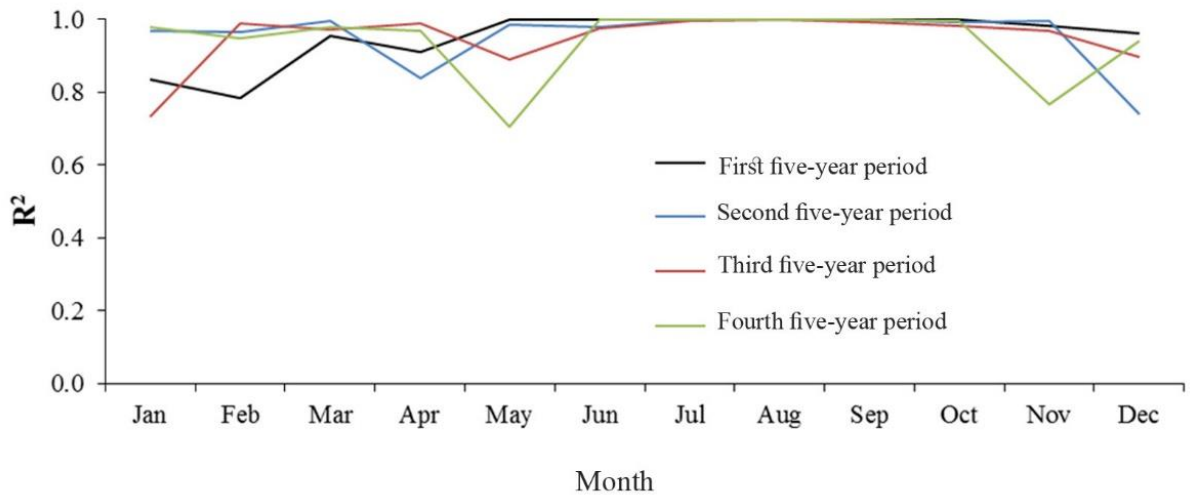
period based on the 5-year series. The results show that the seasonal distribution of the observed variances is well preserved, and the maximum error values are small. Furthermore, the impact of the errors in the estimated variances on the distribution of extreme values was evaluated. For this purpose, in calibrating the NSRP model for the 19-year period, the estimated 6-hour variances mentioned above were used instead of the observed 6-hour variances. In Figure (5), the distribution of maximum 6-hour precipitation is

compared for two scenarios: a) using 19 years of 6-hour precipitation data (Observed); and b) using 5 years of 6-hour data and 19 years of daily data (from the first to the fourth five years' periods).

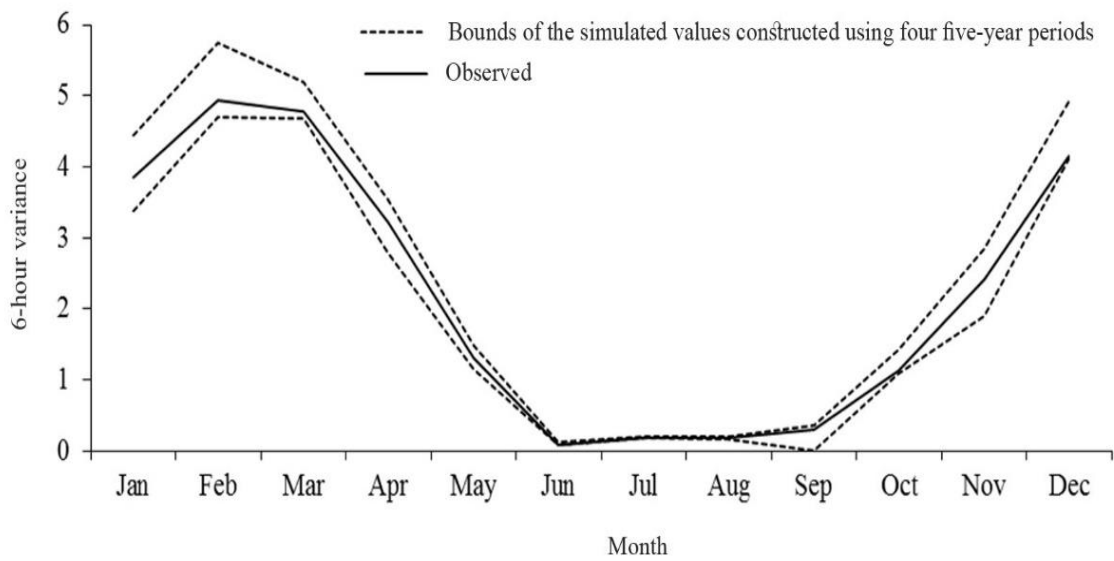
Additionally, the maximum percentage difference between the maximum precipitation values for cases with 6 years of 6-hour data and 19 years of daily data, compared to the case with 19 years of 6-hour precipitation data, for return periods of 2 to 50 years, is presented in Table (2).



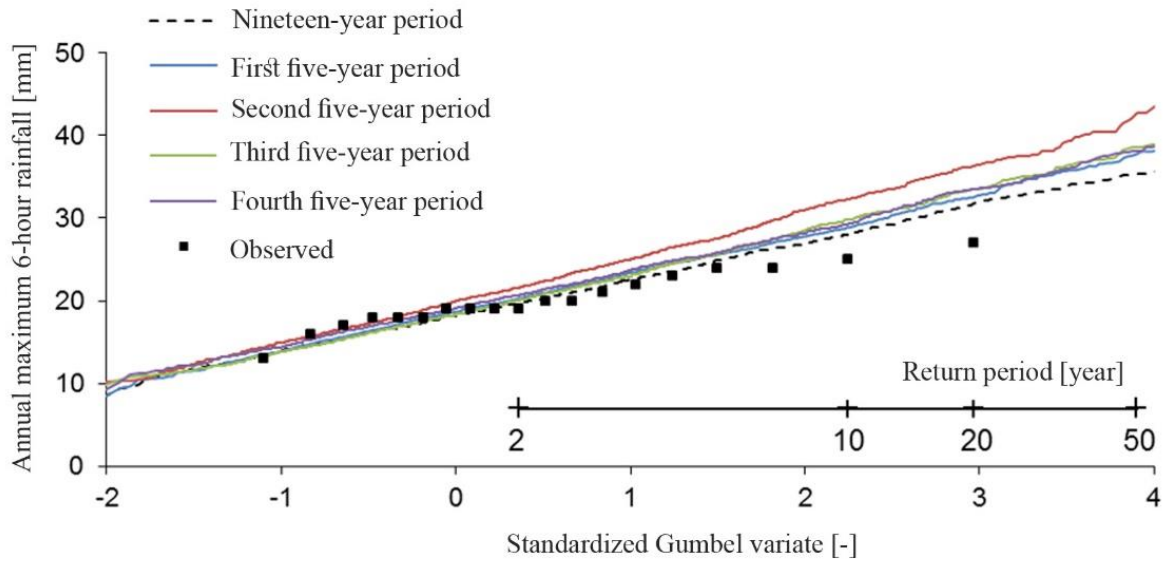
**Figure 2.** Comparison of the observed and simulated annual maximum 6-hour precipitation distributions at the 95% confidence level.



**Figure 3.** Coefficients of determination for the relationship between 6-hour and daily variance across different five-year periods.



**Figure 4.** Bounds of the simulated and observed 6-hour variance.



**Figure 5.** Comparison of the simulated and observed annual extreme 6-hour precipitation distributions.

Based on this, the maximum percentage error for the four cases examined, regarding annual maximum precipitation up to a return period of 25 years, is less than 10%, while for a 50-year return period, it reaches a maximum of 15%. Moreover, except for the 6-hour variance—which is expected to have an error approximately within the range presented in Figure 4—other precipitation statistics have been well reproduced in the synthetic series. The results indicate that the accuracy of the proposed method for extending the 5-year series of 6-hour precipitation is acceptable. Moreover, this study can be valuable for developing methods to reproduce synthetic data with short time intervals under data scarcity conditions.

### Conclusion

The study and analysis of extreme precipitation

events at time steps shorter than daily are particularly important because accurate simulation of flood regimes and assessment of the impact of climate change on flooding in small basins require continuous precipitation data in short time intervals. Recorded precipitation series often have a limited statistical length, which, due to climate variability, introduces uncertainty in the results, especially in the study of extreme values. Furthermore, in most meteorological stations, precipitation data are recorded on a daily time scale, and even if hourly precipitation time series are available, they are often much shorter than daily series, making them insufficient for robust analysis. Therefore, simulating long-term synthetic precipitation series at short time steps, particularly under data-scarce conditions, is highly significant.

**Table 2.** Percentage error in simulating extreme precipitation using a single 5-year period of 6-hour precipitation data and 19 years of daily precipitation data, compared to using a 19-year period of 6-hour precipitation data.

Return Period (years)	Percentage Error Compared to 19-Year 6-Hour Precipitation			
	First 5-Year Period	Second 5-Year Period	Third 5-Year Period	Fourth 5-Year Period
T				
2	7	8	4	7
5	7	8	8	8
10	9	10	8	9
25	9	10	8	9
50	9	15	10	10

In this study, a long-term series of 6-hourly

precipitation from the northern Tehran station was generated using a 19-year precipitation series

through the NSRP stochastic model. The results of model validation indicate the model's capability to produce a long-term series of 6-hourly precipitation, where the statistical characteristics of observed precipitation and the distribution of extreme precipitation events were reproduced at a 95% confidence level. Additionally, the possibility of generating a long-term series of 6-hourly precipitation that represents the 19-year 6-hourly precipitation was examined for a scenario where only 5 years of 6-hourly data and 19 years of daily data are available. For this purpose, the relationship between the 6-hourly and daily statistics was extracted over a 5-year period. Then, the 6-hourly statistics for the 19-year period were calculated from the daily statistics and were used for calibrating the NSRP model and producing a long-term 6-hourly series. To assess the uncertainty of sampling variability, four different 5-year periods were utilized, and the results were compared with the scenario using 19 years of 6-hourly precipitation. The results showed that using the proposed method, the distribution of reproduced extreme 6-hourly precipitation for scenarios utilizing 5 years of 6-hourly data is significantly close to the scenario using 19 years of 6-hourly precipitation. For instance, the error for the four examined scenarios for maximum annual precipitation up to a return period of 25 years is less than 10%, and for a return period of 50 years, it is less than 15%. Consequently, the method presented in this paper can be recommended for generating long-term precipitation series with short time steps, which also encompass extreme precipitation events, even in cases where the length of observed hourly precipitation series is short and data is scarce.

### **Funding sources**

*This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors*

### **Declaration of competing interest**

*All authors have no financial or personal relationships that could potentially bias the content or findings of this research.*

### **Authors contribution statement**

*M.R.K conceived of the presented idea and developed the theory. M.R.K and V.Z performed the computations. M.R.K, V.Z, and B.Z prepared the initial draft and performed the writing.*

### **Acknowledgments**

*The NSRP rainfall model is adopted from the RainSim V3 model (Burton et al., 2008). I am grateful to A. Burton, C.G. Kilsby, H.J. Fowler, P.S.P. Cowpertwait, and P.E. O'Connell for providing the RainSim V3. Moreover, I would like to acknowledge the data providers.*

### **References**

- Boughton, W. (2006). Calibrations of a daily rainfall-runoff model with poor quality data. *Environmental Modelling & Software*, 21(8), 1114–1128.
- Burton, A., Kilsby, C. G., Fowler, H. J., Cowpertwait, P. S. P., & O'Connell, P. E. (2008). RainSim: A spatial-temporal stochastic rainfall modelling system. *Environmental Modelling & Software*, 23(12), 1356–1369.
- Chen, J., Arsenault, R., Brissette, F. P., Côté, P., & Su, T. (2019). Coupling annual, monthly and daily weather generators to simulate multisite and multivariate climate variables with low-frequency variability for hydrological modelling. *Climate Dynamics*, 53, 3841–3860.
- Cowpertwait, P. S. P., Kilsby, C. G., & O'Connell, P. E. (2002). A space-time Neyman–Scott model of rainfall: Empirical analysis of extremes. *Water Resources Research*, 38(8), 6-1–6-14.
- Holman, I., Tascone, D., & Hess, T. (2009). A comparison of stochastic and deterministic downscaling methods for modelling potential groundwater recharge under climate change in East Anglia, UK: Implications for groundwater resource management. *Hydrogeology Journal*,

- 17, 1629–1641.
- Intergovernmental Panel on Climate Change (IPCC). (2012). *Managing the risks of extreme events and disasters to advance climate change adaptation* (C. B. Field et al., Eds.). Cambridge University Press.
- Khazaei, M., Ahmadi, S., Saghafian, B., & Zahabiyoun, B. (2013). A new daily weather generator to preserve extremes and low-frequency variability. *Climatic Change*, 119, 631–645.
- Khazaei, M. R., Zahabiyoun, B., & Saghafian, B. (2012). Assessment of climate change impact on floods using weather generator and continuous rainfall-runoff model. *International Journal of Climatology*, 32(13), 1997–2006.
- Khazaei, M. R., Zahabiyoun, B., & Hasirchian, M. (2020). Comparison of IWG and SDSM weather generators for climate change impact assessment. *Theoretical and Applied Climatology*, 140, 859–870.
- Khazaei, M. R., Hasirchian, M., & Zahabiyoun, B. (2021). An improved daily weather generator for the assessment of regional climate change impacts. *Theoretical and Applied Climatology*, 146, 475–487.
- Khazaei, M. R. (2021). A robust method to develop future rainfall IDF curves under climate change condition in two major basins of Iran. *Theoretical and Applied Climatology*, 144, 179–190.
- Khazaei, M. R. (2024). Performance comparison of four daily weather generators for historical period and downscaling future GCM scenarios. *Journal of Water and Climate Change*, 15(10), 5258–5271.
- Khazaei, M. R. (2025). A proposed method to estimate climate change impacts on dam's spillway design flood. *Journal of Flood Risk Management*, 18(2), e70040.
- Kilsby, C. G., Jones, P. D., Burton, A., Ford, A. C., Fowler, H. J., Harpham, C., ... & O'Connell, P. E. (2007). A daily weather generator for use in climate change studies. *Environmental Modelling & Software*, 22(12), 1705–1719.
- Kilsby, C. G., Moaven-Hashemi, A., & O'Connell, P. E. (2004). Simulation of rainfall extremes: Fitting to observed annual maxima. In *First International Conference on Flood Risk*. University of Bath, UK: Institute of Mathematics and its Applications.
- Kim, D., & Olivera, F. (2012). Relative importance of the different rainfall statistics in the calibration of stochastic rainfall generation models. *Journal of Hydrologic Engineering*, 17(3), 368–376.
- Olsson, J., & Burlando, P. (2002). Reproduction of temporal scaling by a rectangular pulses rainfall model. *Hydrological Processes*, 16(3), 611–630.
- Rodriguez-Iturbe, I., Cox, D. R., & Isham, V. (1987). Some models for rainfall based on stochastic point processes. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 410(1839), 269–288.
- Roy, L., Leconte, R., Brissette, F. P., & Marche, C. (2001). The impact of climate change on seasonal floods of a southern Quebec river basin. *Hydrological Processes*, 15(16), 3167–3179.
- Semenov, M. A., Brooks, R. J., Barrow, E. M., & Richardson, C. W. (1998). Comparison of the WGEN and LARS-WG stochastic weather generators for diverse climates. *Climate Research*, 10(2), 95–107.