American Academic Scientific Research Journal for Engineering, Technology, and Sciences
ISSN (Print) 2313-4410, ISSN (Online) 2313-4402

http://asrjetsjournal.org/

Generation of Synthetic Daily Rainfall Data in Jordan

Ahmad Osama Musleh[a*], Fayez Ahmad Abdulla[b]

[a,b]*Jordan University of Science and Technology, Department of Civil Engineering, Irbid, Jordan*
[a]*Email: aomusleh17@eng.just.edu.jo*

Abstract

This study aims to generate synthetic daily rainfall data for 39 meteorological stations in Jordan by estimating the distributional parameters of daily rainfall occurrence and amounts. Daily rainfall occurrence was modeled by the use of rainfall interarrival times, which were fitted to the one-parameter exponential distribution, except zero values which were represented using the ratio of the number of zero interarrival times to the total number of times, which was called zero ratio. Daily rainfall amounts were fitted to the two-parameter gamma distribution. Goodness-of-fit for one of the stations was tested using chi-square test. This test was performed using Microsoft Office Excel. Distributional parameters were calculated for both occurrence and amounts models, and 100 sequences of synthetic daily rainfall data were then generated, of which every sequence included 1000 non-zero daily rainfall data points (1000 wet days) and 1000 interarrival times (1000 dry spells of which some have a length of zero). These sequences were generated using the embedded random number generators in Python, for one-parameter exponential distribution, two-parameter gamma distribution, and uniform distribution. Percent errors were then calculated and found all to be less than 10%, which was considered acceptable.

Keywords: rainfall; daily rainfall; synthetic data; occurrence model; amounts model; gamma distribution; exponential distribution; Jordan.

1. Introduction

1.1. Overview

Rainfall maintains the hydrological cycle and plays a critical role for sustainable agro-economical, water management and sustenance of human livelihoods [1]. Also, rainfall precipitation is one of the most important weather variables in simulation models [2]. Since the availability of the weather data limits the applicability of the simulation method [3], stochastic rainfall models (SRMs) are used as tools for creating long unlimited rainfall time series data whose statistical properties are close to those of observational records [4]. They are used for augmenting rainfall time series, producing multiple climate realizations for vulnerability assessment, and generating synthetic rainfall records for ungauged stations through interpolating model parameters from adjacent gauged sites [5, 4].

----

* Corresponding author.

Precipitation is the most important variable in the rainfall–runoff models and extreme flood runoff models for analyzing dam safety risk, assessing flood and drought risk, designing infrastructure, managing water resources, and examining reservoir operations and performance [6, 4]. Stochastic rainfall data are also often used as inputs into hydrological models to quantify uncertainty in environmental systems associated with climatic variability, which facilitates making decisions about risk-based design and system operations [7]. Due to the scarcity of historical data, randomly generated synthetic data (stochastic replicates of the historical data) that are based on the statistical characteristics of the historical data are important to overcome the scarcity issue, since the synthetic data can be used for estimating the missing historical data due to their similarity in the statistical characteristics. Also, forecast accuracy of rainfall models depends in the first place on the reliability of the past rainfall data provided to the model [8].

Simulation of rainfall over a region for long time-sequences can be very useful for planning and policymaking, especially when the economy is heavily reliant on rainfall [9]. SRMs are useful in the fields of water resources management, hydrology, ecology, meteorology, and agricultural science and engineering, and help in water resources planning, reservoir and watershed management, flood risk assessment, drought risk assessment, rainfall–runoff models, hydraulic structure design, infrastructure design, erosion prediction, design of landfills, design of facilities for storage and disposal of hazardous wastes, agricultural production and planning, crop-yield models, and climate change impact studies [5, 4, 10, 11, 6, 7, 12, 1]. Thus, proper monitoring and forecasting of rainfall are inevitable [1].

Simulation of rainfall precipitation has two proposed basic approaches: physical and mathematical approaches [10]. Physical approaches are limited in their scope and applicability due to the complexity of the underlying rainfall generation mechanisms [10]. Mathematical approaches are more widely used and called stochastic approaches as they consider rainfall precipitation as a random process [10].

Rainfall precipitation is considered a stochastic process. The basic stochastic processes follow the point process theory [6]. A point process is the process of occurrence of an event continuously along a temporal or spatial dimension. For instance, the arrival of a vehicle to a certain point on a road is a time series event, since the position is fixed, and the events occur during a time interval. This is called a temporal point process. If the time is fixed and a picture is taken to the whole road, the positions of the vehicles on the road form a spatial point process. The point process of interest in this study is rainfall precipitation, which is a temporal point process.

Point processes are Poisson processes, of which the time intervals among successive points follow the exponential probability distribution. When the time is divided into equal-width classes, e.g., 24-hour classes, frequency analysis can be done and histograms result. The shapes observed in the histograms usually follow the gamma probability distribution. In this study, daily rainfall precipitation is modeled. Thus, the width of classes into which the time is divided is equal to one day. The time interval between each two successive rainfall precipitation events is called interarrival time. Interarrival times describe the occurrence model, while the frequency analysis of the rainfall precipitation depths describes the amounts model.

One of the challenges in the simulation of rainfall precipitation is achieving satisfactory representation of the observed process with high comprehensibility, applicability, and computational soundness, without falling into over-parameterization [10].

Different timescales are used in rainfall modeling and range from a year to a few minutes, but the daily timescale has been gaining the highest attention [10].

The simulation scheme of the SRMs is based on two steps of random processes: occurrence model and amounts model [4, 10]. One of approaches for simulating rainfall occurrence is the alternating renewal process. In the alternating renewal process, lengths of consecutive wet and dry spells are considered a random variable in a truncated negative binomial or truncated geometric probability distribution, or a mixture of two geometric distributions [10]. Unfortunately, parameter estimation in the alternating renewal process is problematic. This problem can be overcome by using another approach for simulating rainfall occurrence that is the Markov chain (MC) process, which is the most widely used base of the occurrence models due to its simplicity and effectiveness [4, 13]. Transformation of MC transition probabilities are used to rewrite MC-based models as generalized linear models (GLMs), which can be incorporated in commonly available statistical packages [10]. In an MC, the state of a day is estimated based upon the state(s) of its preceding day(s) [10]. The number of the preceding days on which the state of the day of interest depends is called the order of the MC [11]. Despite the wide use of first-order MCs for daily rainfall simulation, higher-order MCs are getting more popular in sake of improving the dependence structure of wet and dry spells, especially for long dry spells [4]. Low-order MC-based models (i.e., first- and second-order) cannot satisfactorily reproduce observed low-frequency variability; cannot be generalized to represent spatial dependence across multiple point locations; underestimate the interannual variability of wet days, which is governed by the day-to-day and low-frequency variations in the rainfall; and provide poor distribution of number of wet days and rainfall totals at annual time scale [2]. These shortcomings can lead to improper evaluation of the hydrological or agricultural behavior of a region and suboptimal policies for system management [2]. The variance in the rainfall that is unexplained by low-order MC-based models (i.e., overdispersion) is said to be associated with climatic nonstationarity and/or longer time scale variations in the rainfall [2]. In order to incorporate these factors in an SRM, a covariate containing atmospheric signals can be imposed to allow variations in the SRM parameters, which can also be conditionally modified based solely on previous values of aggregated time scale predictors [2].

Amounts models are classified into parametric, semi-parametric and nonparametric models [5]. In parametric methods, the shape of the underlying probability density function (pdf) and correlation structure are presumed, while in nonparametric methods neither is presumed, and the pdf is characterized by the observed time series [10, 14]. In parametric methods, rainfall amounts can be simulated conditionally or unconditionally [10]. In conditional simulation, rainfall amounts have different probability distributions depending on the states of the corresponding wet days, whether they are similar to or different from the day of interest, while unconditional simulation assumes a generic distribution for the rainfall amounts [10]. Probability distributions incorporated in parametric approaches are highly positively skewed and include two-parameter gamma, shifted gamma, one-parameter exponential, three-parameter mixed exponential, kappa, and Weibull distributions [10, 11]. The advantage of the parametric techniques is their capability for extrapolation of non-observed extreme values [5, 4], but the nonparametric techniques are more widely used due to their simplicity and high capability of reproducing observations, although the latter lacks the extrapolation ability for unobserved extremes [5, 14]. Recently, mixed and hybrid probability distributions have been used in parametric techniques, e.g., mixed exponential, double gamma, and GP-Type III distributions [5]. The advantage of these probability distributions is their ability to reproduce the whole range of rainfall amounts, including the most extreme rainfall [4]. Examples of nonparametric techniques include histograms, nearest-neighbor algorithm, and kernel density estimation approach [5, 10, 14]. WGEN, CLIGEN, ClimGen and WeaGETS are examples of recently generated MC-based SRMs with parametric probability distributions [4].

Seasonality (periodicity) in rainfall simulation can be expressed in two approaches; one is imposing a seasonal trend on the important model parameters, such as a polynomial or Fourier function, and the other is handling seasons discretely [10, 11]. Different seasons may have different probability distributions or different dependence characteristics [10]. In Fourier series, every day of the year has its own unique model parameters, while in other models, days are grouped into seasonal groups for which the parameters are estimated [14].

SRMs can represent one point location (single-site models) or multiple point locations (multi-site models), of which the latter is comparatively complex [5, 2]. A simple approach for developing a multi-site model is to extend single-site models by driving them with temporally independent but spatially correlated random numbers [5, 2]. Also, multi-site models sometimes use transformations of the multivariate normal distribution, but when nonparametric methods are adopted, nearest-neighbor resampling is an easy choice [14]. The spatial and temporal intermittence of daily rainfall makes it the most difficult weather variable to simulate [11].

SRMs suffer the limitation of not representing rainfall event characteristics, which include wet spell (rainfall duration), total rainfall amount in one rainfall event (rainfall depth), distribution and dependence structure of daily rainfall amounts in a wet spell (temporal rainfall patterns), and the correlation structure of these three characteristics [4]. Rainfall event characteristics have essential influences on runoff and flood modeling. In multi-day extreme rainfall events, flood volume and duration are highly influenced by rainfall depth and duration, and the resulting surface runoff is significantly affected by their dependence structure [4]. Also, rainfall peak delay in rainfall events increases the severity of the resulting runoff and floods, which is an example of the importance of the effect of temporal rainfall patterns [4]. Consequently, the importance of including rainfall event characteristics in SRMs can be concluded. Thus, event-based rainfall models (rainfall event models), which reproduce rainfall event characteristics, have been developed to overcome the limitations of SRMs [4]. However, the output of rainfall event models can only be used as input to event-based hydrological models, since it is in the image of a sequence of rainfall events [4]. For this reason, a recently published research paper has suggested an MC-based SRM named as SDRM-MCREM (stochastic daily rainfall model with Markov chain rainfall event model) that generates rainfall time series while maintaining rainfall event characteristics [4].

Rainfall event models are classified into profile-based and pulse-based models, with representative Barlett-Lewis and Neyman-Scott models [5].

Rainfall depth and duration can be simulated jointly by copula functions or stochastically using Monte-Carlo method which is preferred to be used for simulating temporal rainfall patterns as well [4].

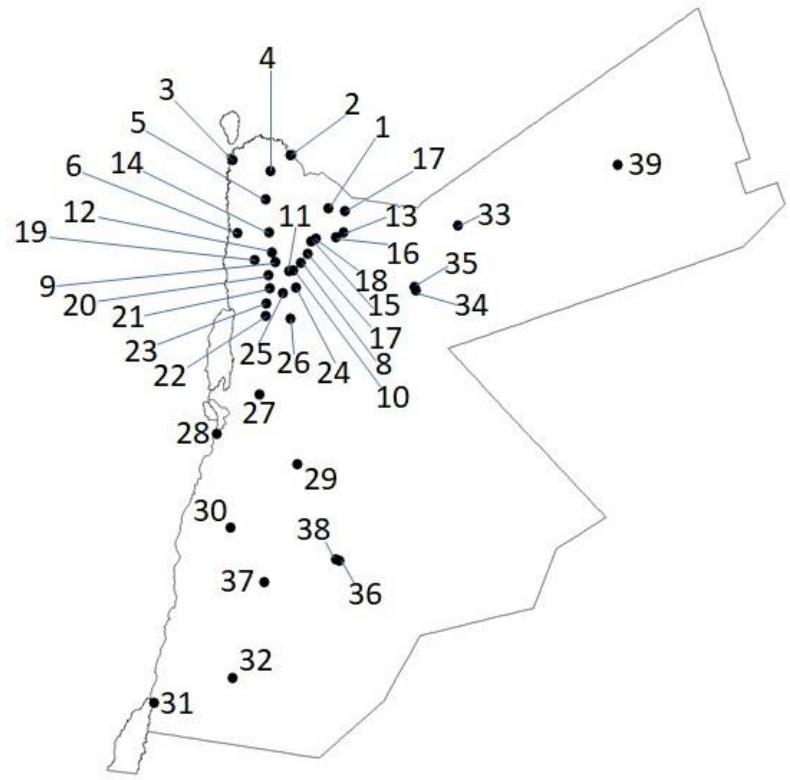
1.2. Study area

This study includes 39 rainfall stations across Jordan, of which the IDs, names and locations are shown in Figure 1. As a summary of their descriptive information, the years of record of the stations range from 22 to 78 years, except one station that has only 9 years of record. In more detail, 15% of the stations have 78 years of record, 33% more than 70, 40% more than 60, and 60% 50 years or more.

All rainfall precipitation records are from October to May, except one record in June for Ras Muneif evaporation station. The number of rainy days in each month. These records are all in October and May, which are the beginning and the end of the rainy season, respectively.

1.3. The significance of the paper

Research in Jordan lacks focus on rainfall stations. One previous research paper was found to study 13 meteorological stations in Jordan [15]. Another paper studied 6 stations [16]. Other research papers were found to study only 3 stations [17, 18]. This paper studies 39 stations across Jordan.



#	ID	Name	#	ID	Name
1	AD0019	Mafraq Airport	21	AN0003	Na'ur
2	AD0021	Turra	22	CC0001	Madaba
3	AD0032	Baqura	23	CC0004	Mushaqqar
4	AE0002	Irbid	24	CD0001	Sahab
5	AH0003	Ras Muneif	25	CD0002	Yaduda
6	AL0010	Deir Alla	26	CD0005	Jiza
7	AL0015	Zarqa	27	CD0010	Rabba
8	AL0016	Ruseifa	28	CF0006	Ghores-Safi
9	AL0018	Jubeiha	29	CF0007	Hasa
10	AL0019	Amman Airport	30	DA0002	Shaubak
11	AL0020	Ain Ghazal	31	ED0001	Aqaba
12	AL0035	Baq'a	32	ED0012	Ram
13	AL0048	Khaldiya	33	F 0002	H5
14	AL0053	King Talal Dam	34	F 0003	Azraq Police Post
15	AL0054	Hashimiya	35	F 0009	Azraq Evap. Station
16	AL0055	Wadi Dhuleil			
17	AL0059	Um El-Jumal	36	G 0002	Jafr Police Post
18	AL0066	Khirebit Es Samra	37	G 0003	Ma'an
19	AM0001	Salt	38	G 0008	Jafr Evap. Station
20	AN0002	Wadi Es-Sir	39	H 0001	H4

Figure 1: The 39 rainfall stations included in this study across Jordan.

Table 1: Number of rainy days in each month for each station

Name	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Name	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May
Mafraq Airport	92	181	312	403	368	238	107	44	Na'ur	86	253	477	537	539	424	176	40
Turra	68	162	294	370	331	260	125	25	Madaba	77	264	455	544	511	405	141	39
Baqura	95	249	381	447	386	320	117	28	Mushaqqar	36	94	185	248	249	148	35	11
Irbid	149	318	513	609	569	483	224	69	Sahab	41	151	252	346	303	198	81	11
Ras Muneif	139	268	433	492	435	393	180	46	Yaduda	6	56	116	125	120	103	36	11
Deir Alla	122	279	458	559	488	411	160	53	Jiza	45	147	275	367	310	227	71	12
Zarqa	68	191	312	414	359	279	95	42	Rabba	58	209	380	473	442	335	121	18
Ruseifa	55	166	303	385	340	264	85	28	Ghores-Safi	25	51	96	145	131	89	38	5
Jubeiha	107	326	536	665	602	508	202	61	Hasa	22	61	68	104	81	82	24	6
Amman Airport	137	354	600	734	712	564	255	99	Shaubak	67	150	278	382	305	251	112	26
Ain Ghazal	4	18	51	57	65	46	22	10	Aqaba	27	47	105	119	87	77	45	9
Baq'a	86	196	350	413	413	316	119	39	Ram	8	14	21	43	24	23	13	2
Khaldiya	29	77	112	173	154	98	33	18	H5	64	162	251	304	310	227	111	49
King Talal Dam	46	146	242	300	310	218	68	21	Azraq Police Post	15	28	42	48	25	40	12	5
Hashimiya	30	94	141	174	182	115	45	15	Azraq Evap. Station	33	79	131	199	153	112	49	20
Wadi Dhuleil	52	134	236	312	301	208	71	23	Jafr Police Post	13	33	37	34	29	30	24	5
Um El-Jumal	49	160	250	320	288	210	85	26	Ma'an	52	86	138	218	171	152	57	26
Khirebit Es Samra	35	87	161	188	190	95	29	8	Jafr Evap. Station	21	19	28	35	32	27	18	4
Salt	105	302	526	618	610	503	196	64	H4	104	154	253	292	272	236	157	81
Wadi Es-Sir	100	277	520	587	545	431	190	41									

2. Methodology

2.1. Goodness-of-fit

Chi-square test is a well-known statistical goodness-of-fit test for testing whether a dataset follows a certain probability distribution. In this study, chi-square test was applied for Mafraq Airport station dataset for both occurrence and amounts model. That is, it was used for testing whether the interarrival times followed the exponential distribution and whether non-zero rainfall precipitation depth data followed the gamma distribution. It was shown that both sets of data followed their corresponding probability distributions except for a simple modification that was needed for the interarrival times' test. This modification was made to exclude zero interarrival times from the model and represent them by a new parameter called zero ratio. This parameter is explained in section 3.3.1.

2.2. Model calibration

2.2.1. Occurrence model

Occurrence model was constructed based on the concept of interarrival time. An interarrival time is the period between two successive rainfall events. In other words, it is the number of dry days between every two successive wet days. According to previous studies, interarrival times follow a statistical probability distribution called exponential distribution. The exponential distribution has one parameter called lambda. Lambda is calculated using the following equations:

$$\lambda = \frac{1}{mean} \tag{6}$$

$$\lambda = \frac{1}{standard\ deviation} \tag{7}$$

Therefore, lambda in this study was calculated using the following equation:

$$\lambda = \frac{1}{2}\left(\frac{1}{mean} + \frac{1}{standard\ deviation}\right) \tag{8}$$

In this study, however, this parameter was found not to be enough for modeling interarrival times. A huge inclination was detected in the zero values of interarrival times. Therefore, a new parameter was created to overcome this error. Lambda was hence used for modeling only interarrival times greater than zero. This means that interarrival times with a minimum of one day were found to follow the exponential distribution. Zero interarrival times were represented by taking a normalized ratio of their number to the total number of interarrival times, which was called zero ratio. Both of those parameters (i.e., lambda and zero ratio) were determined for the real rainfall datasets and the randomly generated synthetic datasets. Percent errors were then calculated for the two parameters. This was done by programming using Python. The code is shown in Appendix A.

2.2.2. Amounts model

After knowing whether a day is dry or wet using the occurrence model, the amounts of rainfall precipitation (i.e., rainfall precipitation depths) are estimated using the amounts model. Thus, the amounts model is only concerned with non-zero rainfall precipitation days (i.e., wet days). Previous studies show that it is acceptable to assume that rainfall precipitation depths for wet days follow the gamma distribution. Gamma distribution has two parameters: alpha and beta, which can be calculated as follows.

$$\alpha = \left(\frac{mean}{standard\ deviation}\right)^2 \tag{9}$$

$$\beta = \frac{(standard\ deviation)^2}{mean} \tag{10}$$

These two parameters were calculated for both the original data and the randomly generated synthetic data, and percent errors were then calculated. This was done by programming using Python. The code is shown in Appendix A.

2.3. Model validation

2.3.1. Occurrence model

2.3.1.1. Graphical method

For visual comparison, a graphical method was used for the validation of the occurrence model for Mafraq Airport station. This was done by imposing the histogram of the interarrival times of the synthetic data on the histogram of those of the observed data. The histogram bars of the observed data are in red, while those of the synthetic data are in blue; thus, the areas where the two histograms overlap are in violet.

2.3.1.2. Percent errors

After calculating the exponential distribution parameters (i.e., lambda and zero ratio) of the interarrival times of the observed data, these parameters were calculated also for the synthetic data. However, the synthetic data parameters are not shown in the results. Instead, the percent errors of the synthetic data parameters based on the observed data parameters are calculated and shown in the results table.

2.3.2. Amounts model

2.3.2.1. Graphical method

For visual comparison, a graphical method was used for the validation of the amounts model for Mafraq Airport station. This was done by imposing the histogram of the non-zero daily rainfall amounts of the synthetic data on the histogram of those of the observed data. The histogram bars of the observed data are in red, while those of the synthetic data are in blue; thus, the areas where the two histograms overlap are in violet.

2.3.2.2. Percent errors

After calculating the gamma distribution parameters (i.e., alpha and beta) of the non-zero daily rainfall amounts of the observed data, these parameters were calculated also for the synthetic data. However, the synthetic data parameters are not shown in the results. Instead, the percent errors of the synthetic data parameters based on the observed data parameters are calculated and shown in the results table.

3. Results and Discussion

3.1. Goodness-of-fit

Goodness-of-fit was tested using the well-known chi-square test for Mafraq Airport station. Interarrival times were successfully fitted to the exponential distribution, and non-zero daily rainfall amounts to the gamma distribution. Results of chi-square tests for interarrival times and non-zero depths for Mafraq Airport station – October are shown in Tables 2–3.

Table 2: Chi-square test for the occurrence model of Mafraq Airport station – October (fitting interarrival times to one-parameter exponential distribution) (1/mean = 0.046062407; 1/standard deviation = 0.043082865; λ = 0.044572636; X^2 = 12.99295; χ^2 = 23.68; $X^2 \leq \chi^2 \Rightarrow H_0$ accepted; Zero frequency = 29; Zero ratio = 0.318681319)

k	f	p	F	$(f-F)^2/F$
2	8	0.085287	5.287814	1.391114
4	6	0.078013	4.83683	0.279721
6	2	0.07136	4.42431	1.328406
8	5	0.065274	4.046972	0.22443
10	5	0.059707	3.701817	0.455257
12	2	0.054615	3.386099	0.567399
14	3	0.049957	3.097308	0.003057
16	6	0.045696	2.833147	3.539866
18	2	0.041799	2.591515	0.135014
20	3	0.038234	2.370492	0.167172
22	1	0.034973	2.168319	0.629506
24	2	0.03199	1.983389	0.000139
26	3	0.029262	1.814231	0.775011
28	0	0.026766	1.6595	1.6595
30	0	0.024483	1.517966	1.517966
More	14	0.262585	16.28029	0.319388
Sum	62	1	62	12.99295

Table 3: Chi-square test for the amounts model of Mafraq Airport station – October (fitting non-zero data points to two-parameter gamma distribution) (mean = 3.031521739; standard deviation = 3.458352503; α = 0.768392041; β = 3.945279982; X^2 = 3.811961161; χ^2 = 15.51; $X^2 \leq \chi^2 \Rightarrow H_0$ accepted)

k	f	p	F	$(f-F)^2/F$
0.5	22	0.209742	19.29624	0.378847
1	9	0.129102	11.87742	0.697084
1.5	10	0.100616	9.256651	0.059694
2	8	0.081893	7.534124	0.028808
2.5	7	0.068028	6.258608	0.087825
3	7	0.057192	5.26169	0.574288
3.5	4	0.048464	4.458652	0.047181
4	2	0.041298	3.799422	0.852214
4.5	2	0.03534	3.251255	0.481549
5	4	0.030339	2.791233	0.523467
More	17	0.197986	18.2147	0.081006
Sum	92	1	92	3.811961

3.2. Model calibration

3.2.1. Occurrence model

Values of lambda and zero ratio for the observed daily rainfall data are shown in Table 4.

3.2.2. Amounts model

Values of alpha and beta for the observed daily rainfall data are shown in Table 4.

3.3. Model validation

3.3.1. Occurrence model

3.3.1.1. Graphical method

A graphical representation for original and synthetic data of Mafraq Airport station is provided in the histograms shown in Figure 2. Visual comparison shows a good similarity between observed and synthetic daily rainfall data except for May due to the lack of the observed records.

3.3.1.2. Percent errors

Percent errors for the exponential distribution parameters (i.e., lambda and zero ratio) were calculated for the interarrival times of the synthetic daily rainfall data. All the percent errors for lambda and zero ratio are less than 10%, which is considered acceptable.

3.3.2. Amounts model

3.3.2.1. Graphical method

A graphical illustration for both non-zero observed data and non-zero synthetic data for Mafraq Airport station is shown as histograms in Figure 3. Visual comparison shows a good similarity between observed and synthetic daily rainfall data.

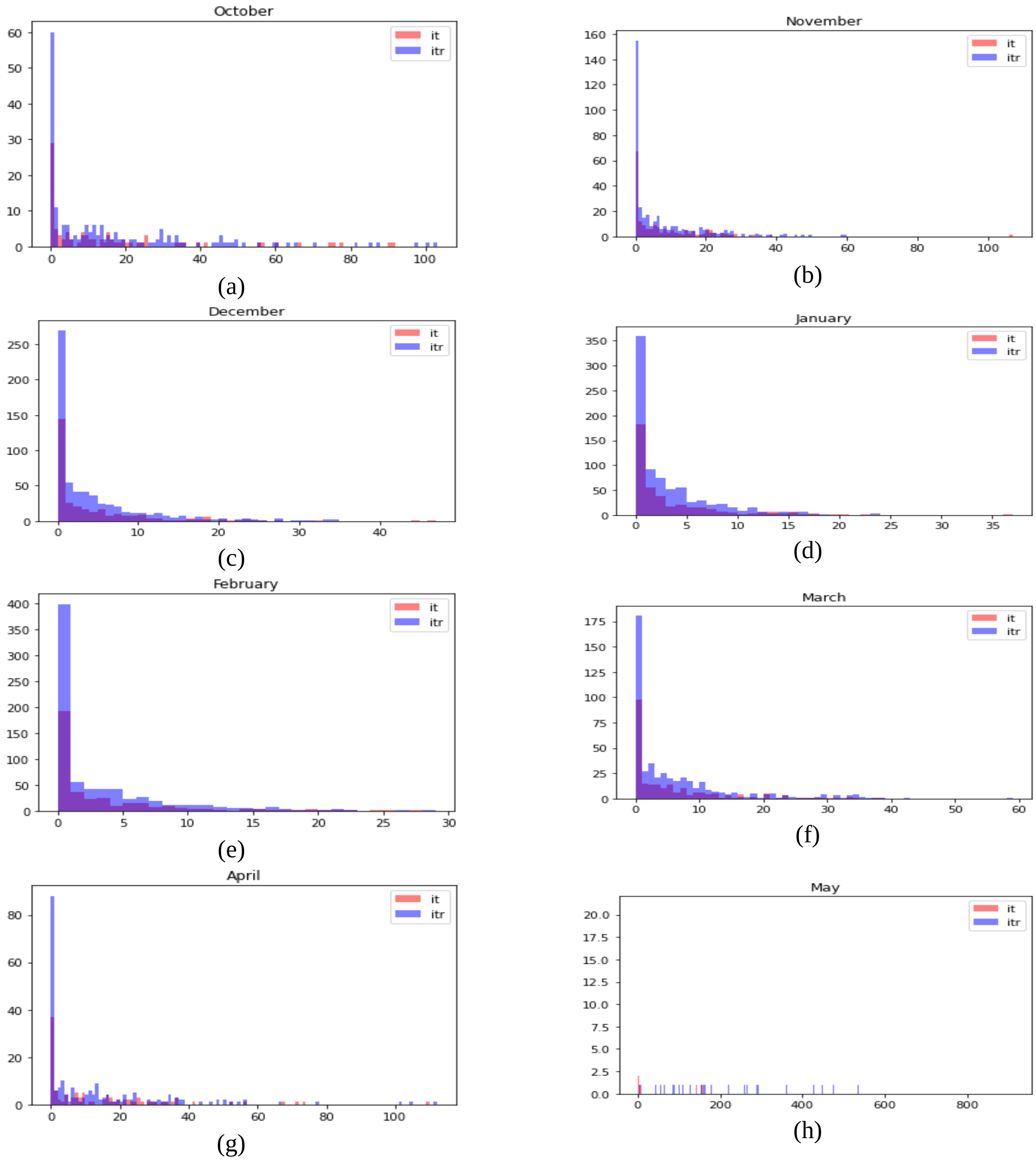


Figure 2: Histogram illustration of observed and synthetic interarrival times for Mafraq Airport station, in (a) October (b) November (c) December (d) January (e) February (f) March (g) April (h) May *(it: observed interarrival times; itr: synthetic (random) interarrival times)*

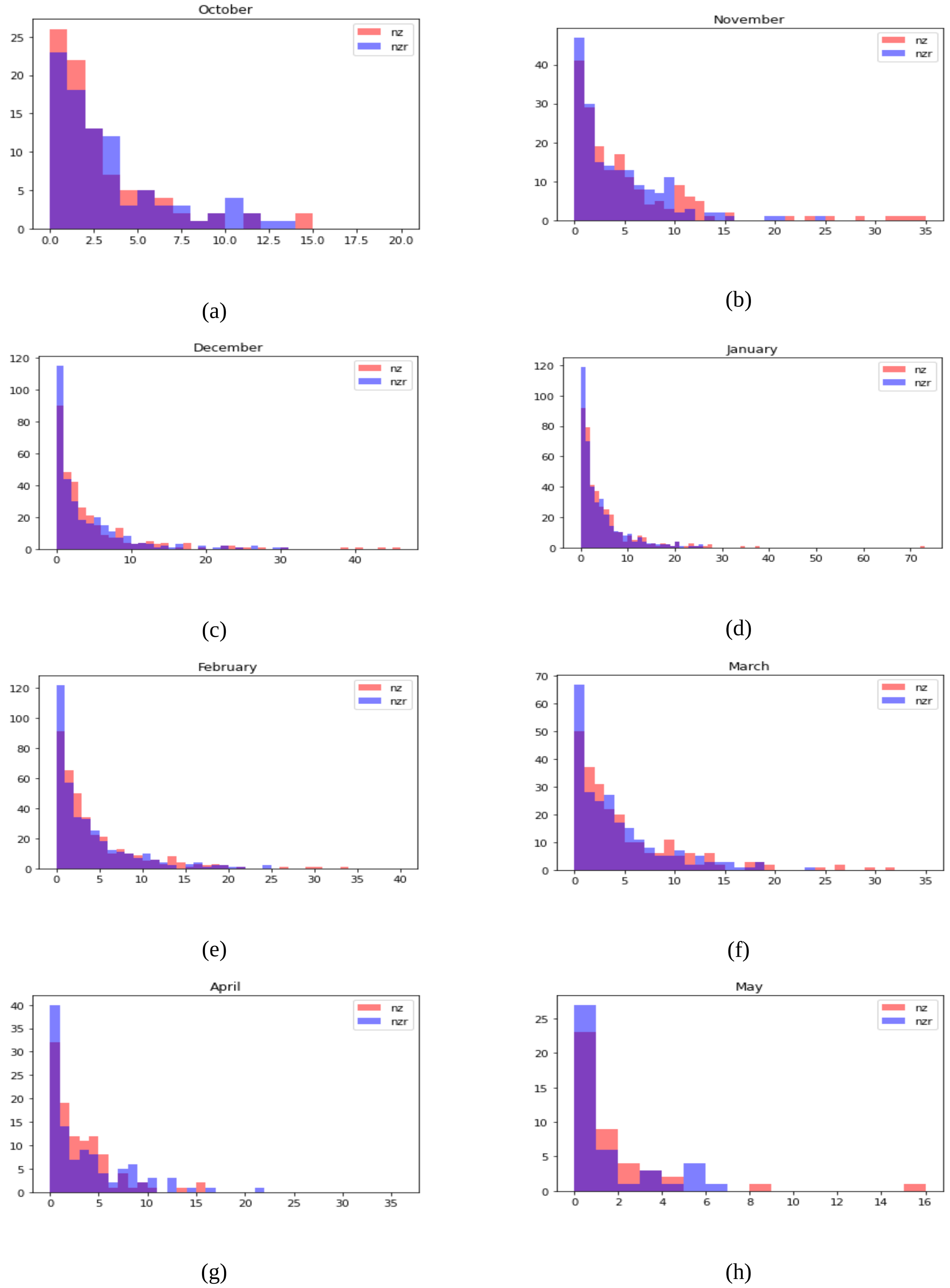


Figure 3 : Histogram illustration of observed and synthetic non-zero rainfall depths for Mafraq Airport station (a) October (b) November (c) December (d) January (e) February (f) March (g) April (h) May *(nz: observed non-zero rainfall depths; nzr: synthetic (random) non-zero rainfall depths)*

3.3.2.2. Percent errors

Percent errors for the exponential distribution parameters (i.e., lambda and zero ratio) were calculated for the interarrival times of the synthetic daily rainfall data. All the percent errors for lambda and zero ratio are less than 10%, which is considered acceptable.

Table 4: Distributional parameters for the observed daily rainfall data

Station Name	Parameter	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May
Mafraq Airport	λ	0.044	0.083	0.135	0.193	0.176	0.112	0.05	0.022
	z	0.319	0.372	0.463	0.452	0.522	0.412	0.346	0.222
	α	0.76	0.648	0.465	0.526	0.647	0.73	0.689	0.339
	β	0.251	0.127	0.102	0.11	0.15	0.146	0.216	0.168
Turra	λ	0.034	0.076	0.121	0.137	0.14	0.093	0.051	0.015
	z	0.25	0.364	0.429	0.496	0.489	0.485	0.435	0.077
	α	0.763	0.922	0.839	0.766	0.724	0.766	0.794	1.109
	β	0.155	0.119	0.099	0.094	0.09	0.09	0.128	0.366
Baqura	λ	0.048	0.11	0.146	0.19	0.153	0.115	0.052	0.015
	z	0.379	0.496	0.551	0.55	0.59	0.536	0.41	0.276
	α	0.288	0.655	0.636	0.519	0.615	0.778	0.727	0.5
	β	0.052	0.08	0.068	0.057	0.075	0.11	0.134	0.075
Irbid	λ	0.057	0.113	0.185	0.213	0.198	0.149	0.083	0.027
	z	0.399	0.487	0.536	0.583	0.599	0.561	0.469	0.186
	α	0.615	0.647	0.555	0.62	0.57	0.719	0.54	0.388
	β	0.134	0.076	0.056	0.064	0.056	0.075	0.089	0.093
Ras Muneif	λ	0.064	0.113	0.17	0.196	0.188	0.157	0.086	0.023
	z	0.406	0.506	0.545	0.596	0.594	0.547	0.428	0.277
	α	0.53	0.611	0.647	0.628	0.608	0.611	0.506	0.49
	β	0.098	0.068	0.06	0.057	0.053	0.057	0.071	0.079
Deir Alla	λ	0.045	0.102	0.155	0.214	0.178	0.14	0.061	0.022
	z	0.352	0.453	0.513	0.556	0.559	0.513	0.415	0.222
	α	0.606	0.471	0.565	0.677	0.631	0.631	0.536	0.557
	β	0.163	0.064	0.079	0.096	0.096	0.102	0.094	0.167
Zarqa	λ	0.02	0.048	0.074	0.098	0.1	0.066	0.032	0.012
	z	0.194	0.356	0.372	0.408	0.376	0.387	0.211	0.186
	α	0.777	0.718	0.484	0.668	0.649	0.812	0.465	0.607
	β	0.21	0.133	0.086	0.126	0.117	0.165	0.131	0.166
Ruseifa	λ	0.017	0.048	0.079	0.096	0.098	0.066	0.025	0.009
	z	0.241	0.307	0.376	0.423	0.374	0.402	0.318	0.241
	α	0.92	0.689	0.637	0.719	0.744	0.733	1.053	0.669

	β	0.188	0.119	0.099	0.12	0.117	0.113	0.241	0.202
Jubeiha	λ	0.031	0.088	0.143	0.184	0.173	0.133	0.061	0.016
	z	0.271	0.474	0.498	0.541	0.543	0.505	0.406	0.323
	α	0.402	0.512	0.598	0.648	0.578	0.736	0.503	0.729
	β	0.06	0.051	0.047	0.049	0.043	0.06	0.067	0.115
Amman Airport	λ	0.041	0.102	0.148	0.193	0.192	0.143	0.072	0.028
	z	0.309	0.462	0.541	0.578	0.601	0.554	0.412	0.25
	α	0.366	0.404	0.432	0.515	0.484	0.509	0.378	0.378
	β	0.111	0.071	0.069	0.078	0.075	0.084	0.107	0.139
Ain Ghazal	λ	0.024	0.037	0.12	0.158	0.165	0.113	0.068	0.044
	z	0	0.529	0.529	0.526	0.594	0.522	0.364	0.091
	α	1.01	1.481	0.462	0.457	0.542	0.583	0.834	1.349
	β	0.43	0.34	0.053	0.052	0.066	0.114	0.353	0.349
Baq'a	λ	0.037	0.088	0.142	0.169	0.167	0.115	0.053	0.019
	z	0.302	0.446	0.501	0.535	0.576	0.532	0.395	0.3
	α	0.866	0.446	0.49	0.624	0.597	0.735	0.684	0.885
	β	0.182	0.053	0.054	0.064	0.066	0.073	0.149	0.246
Khaldiya	λ	0.029	0.06	0.097	0.142	0.142	0.088	0.026	0.017
	z	0.172	0.338	0.357	0.349	0.383	0.278	0.281	0.167
	α	1.107	0.763	0.667	0.514	0.639	0.786	0.624	0.382
	β	0.38	0.166	0.11	0.09	0.124	0.165	0.206	0.092
King Talal Dam	λ	0.026	0.066	0.099	0.103	0.12	0.077	0.039	0.01
	z	0.283	0.473	0.492	0.545	0.558	0.518	0.324	0.318
	α	0.563	0.733	0.713	1.018	0.792	0.924	1.308	2.519
	β	0.107	0.098	0.085	0.112	0.1	0.106	0.311	0.514
Hashimiya	λ	0.029	0.074	0.123	0.129	0.147	0.089	0.042	0.016
	z	0.2	0.404	0.39	0.41	0.467	0.377	0.182	0.125
	α	1.305	0.813	0.712	0.694	0.832	0.72	0.635	1.231
	β	0.481	0.19	0.162	0.119	0.173	0.184	0.186	0.424
Wadi Dhuleil	λ	0.024	0.052	0.077	0.104	0.095	0.07	0.027	0.01
	z	0.288	0.388	0.403	0.449	0.483	0.413	0.338	0.25
	α	0.81	0.72	0.583	0.723	0.527	0.608	0.924	0.907
	β	0.22	0.153	0.128	0.154	0.125	0.142	0.329	0.548
Um El-Jumal	λ	0.03	0.077	0.125	0.154	0.141	0.098	0.035	0.016
	z	0.167	0.35	0.42	0.478	0.497	0.44	0.369	0.074
	α	0.45	0.799	0.79	0.715	0.83	0.747	0.742	0.603
	β	0.125	0.179	0.205	0.175	0.213	0.185	0.241	0.238
Khirebit Es Samra	λ	0.042	0.074	0.134	0.164	0.174	0.072	0.03	0.009
	z	0.147	0.326	0.4	0.457	0.468	0.389	0.103	0.111
	α	1.191	0.876	0.754	1.005	0.828	0.943	1.549	2.072

	β	0.472	0.181	0.182	0.199	0.172	0.249	0.601	0.786
Salt	λ	0.026	0.094	0.153	0.185	0.186	0.143	0.066	0.019
	z	0.365	0.399	0.475	0.519	0.541	0.488	0.352	0.292
	α	0.55	0.638	0.698	0.838	0.76	0.797	0.669	0.94
	β	0.069	0.044	0.041	0.044	0.047	0.051	0.068	0.124
Wadi Es-Sir	λ	0.032	0.079	0.151	0.182	0.194	0.124	0.062	0.016
	z	0.364	0.469	0.512	0.511	0.494	0.494	0.416	0.143
	α	0.567	0.552	0.582	0.732	0.64	0.776	0.479	0.88
	β	0.086	0.046	0.041	0.045	0.044	0.054	0.052	0.118
Na'ur	λ	0.027	0.078	0.148	0.174	0.178	0.116	0.053	0.015
	z	0.294	0.443	0.459	0.488	0.522	0.495	0.381	0.22
	α	0.712	0.565	0.588	0.626	0.689	0.822	0.513	0.657
	β	0.08	0.046	0.045	0.043	0.053	0.065	0.061	0.138
Madaba	λ	0.022	0.076	0.124	0.164	0.175	0.117	0.044	0.014
	z	0.224	0.413	0.466	0.48	0.454	0.458	0.355	0.175
	α	1.12	0.49	0.559	0.786	0.629	0.794	0.781	0.703
	β	0.228	0.052	0.053	0.075	0.055	0.076	0.094	0.098
Mushaqqar	λ	0.04	0.073	0.132	0.194	0.188	0.093	0.03	0.011
	z	0.25	0.404	0.478	0.532	0.569	0.493	0.206	0.25
	α	1.146	0.654	0.484	0.724	0.725	0.546	0.545	0.946
	β	0.19	0.066	0.045	0.069	0.067	0.055	0.065	0.177
Sahab	λ	0.022	0.056	0.088	0.122	0.119	0.075	0.032	0.005
	z	0.22	0.4	0.382	0.445	0.439	0.374	0.346	0.25
	α	0.792	0.588	0.593	0.735	0.702	0.695	0.547	0.648
	β	0.143	0.083	0.058	0.076	0.07	0.06	0.069	0.065
Yaduda	λ	0.005	0.05	0.108	0.115	0.133	0.093	0.047	0.015
	z	0.333	0.455	0.461	0.48	0.467	0.456	0.417	0.167
	α	5.765	0.472	0.671	0.709	1.103	1.069	0.942	3.185
	β	0.997	0.048	0.055	0.067	0.099	0.084	0.09	0.622
Jiza	λ	0.014	0.042	0.066	0.103	0.084	0.062	0.021	0.005
	z	0.267	0.286	0.425	0.383	0.384	0.339	0.268	0
	α	0.87	0.649	0.776	0.831	0.731	0.932	0.596	0.862
	β	0.254	0.087	0.103	0.115	0.095	0.109	0.083	0.15
Rabba	λ	0.021	0.075	0.118	0.163	0.154	0.109	0.041	0.007
	z	0.345	0.438	0.492	0.514	0.533	0.464	0.397	0.263
	α	0.886	0.491	0.49	0.594	0.544	0.597	0.405	0.449
	β	0.184	0.06	0.05	0.057	0.056	0.058	0.044	0.079
	λ	0.011	0.02	0.031	0.044	0.043	0.027	0.016	0.004
Ghores-Safi	z	0.24	0.255	0.316	0.347	0.346	0.337	0.237	0.167
	α	0.27	0.602	0.539	0.667	0.706	0.518	0.405	0.351

	β	0.057	0.161	0.132	0.21	0.234	0.148	0.118	0.047
	λ	0.014	0.036	0.04	0.057	0.05	0.051	0.014	0.005
Hasa	z	0.429	0.267	0.294	0.365	0.338	0.383	0.208	0.143
	α	0.46	0.657	0.353	0.698	0.533	0.837	0.942	1.098
	β	0.128	0.17	0.109	0.218	0.201	0.299	0.322	0.325
	λ	0.035	0.065	0.113	0.162	0.155	0.098	0.044	0.014
Shaubak	z	0.299	0.36	0.451	0.492	0.416	0.454	0.402	0.185
	α	0.562	0.375	0.458	0.506	0.592	0.444	0.453	0.465
	β	0.16	0.059	0.049	0.054	0.063	0.046	0.06	0.076
	λ	0.008	0.018	0.03	0.038	0.031	0.026	0.018	0.003
Aqaba	z	0.185	0.17	0.346	0.185	0.218	0.171	0.133	0.2
	α	0.452	0.384	0.21	0.286	0.223	0.605	0.292	0.324
	β	0.148	0.11	0.048	0.092	0.056	0.205	0.069	0.067
	λ	0.006	0.009	0.014	0.022	0.017	0.014	0.009	0.002
Ram	z	0.143	0.214	0.19	0.357	0.217	0.304	0.231	0.333
	α	0.673	0.945	1.056	1.085	0.858	0.588	2.394	0.681
	β	0.379	0.204	0.322	0.268	0.207	0.076	0.663	0.194
	λ	0.023	0.047	0.077	0.094	0.117	0.071	0.041	0.016
H5	z	0.19	0.34	0.39	0.395	0.381	0.33	0.27	0.22
	α	0.515	0.392	0.65	0.553	0.459	0.433	0.482	1.281
	β	0.135	0.094	0.181	0.193	0.149	0.149	0.15	0.387
	λ	0.01	0.016	0.022	0.024	0.015	0.017	0.009	0.005
Azraq Police Post	z	0.2	0.222	0.286	0.271	0.24	0.4	0.083	0
	α	0.711	1.018	1.159	0.678	1.759	1.344	0.482	1.756
	β	0.187	0.197	0.208	0.165	0.532	0.365	0.095	0.213
	λ	0.022	0.042	0.066	0.106	0.096	0.061	0.026	0.012
Azraq Evap. Station	z	0.094	0.203	0.359	0.377	0.288	0.268	0.265	0.238
	α	0.227	0.55	0.482	0.618	0.4	0.464	0.502	0.551
	β	0.051	0.159	0.14	0.237	0.156	0.136	0.152	0.151
	λ	0.008	0.015	0.015	0.014	0.023	0.014	0.013	0.003
Jafr Police Post	z	0.167	0.212	0.162	0.088	0.069	0.1	0.174	0
	α	0.633	0.828	0.97	1.169	0.507	1.285	0.647	0.752
	β	0.158	0.122	0.197	0.3	0.088	0.257	0.119	0.198
	λ	0.015	0.027	0.045	0.067	0.065	0.05	0.022	0.008
Ma'an	z	0.235	0.233	0.246	0.321	0.24	0.27	0.175	0.222
	α	0.34	0.447	0.423	0.663	0.663	0.524	1	1.398
	β	0.073	0.115	0.126	0.253	0.191	0.138	0.265	0.393
	λ	0.007	0.007	0.01	0.011	0.013	0.009	0.041	0.002
Jafr Evap. Station	z	0.25	0.211	0.25	0.314	0.188	0.259	0.167	0.2
	α	0.573	0.408	0.929	0.353	0.198	0.559	0.722	2.014

	β	0.131	0.088	0.717	0.077	0.085	0.156	0.382	0.666
H4	λ	0.029	0.057	0.083	0.098	0.11	0.073	0.055	0.021
	z	0.262	0.253	0.281	0.349	0.29	0.322	0.312	0.354
	α	0.418	0.454	0.577	0.628	0.475	0.593	0.416	0.607
	β	0.121	0.109	0.176	0.221	0.144	0.186	0.101	0.208

4. Conclusions and Recommendations

4.1. Conclusions

For all the stations included in this study, interarrival times of daily rainfall data can be represented using the one-parameter exponential distribution, except for zero values that can be represented using a ratio between the number of zeros and the total number of values called zero ratio. For all the stations included in this study, non-zero daily rainfall amounts can be represented using the two-parameter gamma distribution.

Goodness-of-fit for one of the stations was tested using chi-square test. Interarrival times were successfully fitted to the exponential distribution, and non-zero daily rainfall amounts to the gamma distribution.

4.2. Recommendations for Future Studies

- Markov chain is recommended to use in the calibration stage of the occurrence model.
- Nonparametric methods (e.g., kernel and nearest-neighbor estimators) are recommended to use instead of the parametric method of assuming probability distributions beforehand.
- Nonparametric methods for data resampling are recommended to use before studying the data.
- Spatial correlations of daily rainfall data are recommended to consider among the meteorological stations.

5. Ethical Statement

We will conduct ourselves with integrity, fidelity, and honesty. We will openly take responsibility for my actions, and only make agreements, which we intend to keep. We will not intentionally engage in or participate in any form of malicious harm to another person or animal.

6. Conflict of Interests

We declare that we have NO conflict of interests in the subject matter or materials discussed in this paper.

7. Data Availability Statement

The data associated with this paper are available with the authors and can be accessed if needed.

References

[1] A. M. Nyongesa, G. Zeng and V. Ongoma, "Non-homogeneous hidden Markov model for downscaling of short rains occurrence in Kenya," *Theoretical and Applied Climatology,* vol. 139, no. 3, pp. 1333-1347, 2020.

[2] R. Mehrotra and A. Sharma, "A semi-parametric model for stochastic generation of multi-site daily rainfall exhibiting low-frequency variability," *Journal of Hydrology,* vol. 335, no. 1–2, pp. 180-193, 2007.

[3] S. Geng, F. W. P. de Vries and I. Supit, "A simple method for generating daily rainfall data," *Agricultural and Forest Meteorology,* vol. 36, no. 4, pp. 363-376, 1986.

[4] C. Gao, M. J. Booij and Y.-P. Xu, "Development and hydrometeorological evaluation of a new stochastic daily rainfall model: Coupling Markov chain with rainfall event model," *Journal of Hydrology,* vol. 589, 2020.

[5] C. Gao, Y.-P. Xu, Q. Zhu, Z. Bai and L. Liu, "Stochastic generation of daily rainfall events: A single-site rainfall model with Copula-based joint simulation of rainfall characteristics and classification and simulation of rainfall patterns," *Journal of Hydrology,* vol. 564, pp. 41-58, 2018.

[6] DSO, "Stochastic Modeling Methods," Dam Safety Office, Department of the Interior, Bureau of Reclamation, Washington D.C., 2003.

[7] F. H. Chiew, R. Srikanthan, A. J. Frost and E. G. Payne, "Reliability of daily and annual stochastic rainfall data generated from different data lengths and data characteristics," in *In MODSIM05 - International Congress on Modelling and Simulation: Advances and Applications for Management and Decision Making, Proceedings,* Melbourne, Victoria, Australia, 2005.

[8] P. Gogumalla and S. K. Tripathi, "Evaluation of Markov Chain Model for Forecasting Precipitation of Uttarakhand Districts," *HCTL Open International Journal of Technology Innovations and Research (IJTIR),* vol. 6, no. 5, pp. 142-151, 2018.

[9] A. Mitra, "Bayesian Approach to Spatio-Temporally Consistent Simulation of Daily Monsoon Rainfall over India," New York, NY, USA, 2017.

[10] A. Sharma and U. Lall, "A nonparametric approach for daily rainfall simulation," *Mathematics and Computers in Simulation,* vol. 48, no. 4–6, pp. 361-371, 1999.

[11] Y. Liu, W. Zhang, Y. Shao and K. Zhang, "A comparison of four precipitation distribution models used in daily stochastic models," *Advances in Atmospheric Sciences,* vol. 28, no. 4, pp. 809-820, 2011.

[12] J. Piantadosi, J. Boland and P. Howlett, "Generating Synthetic Rainfall on Various Timescales—Daily, Monthly and Yearly," *Environmental Modeling & Assessment,* vol. 14, no. 4, pp. 431-438, 2009.

[13] A. Azizah, R. Welastika, A. Nur Falah, B. N. Ruchjana and A. S. Abdullah, "An Application of Markov Chain for Predicting Rainfall Data at West Java using Data Mining Approach," *Conference Series: Earth and Environmental Science,* vol. 303, pp. 12-26, 2019.

[14] R. Leander and T. A. Buishand, "A daily weather generator based on a two-stage resampling algorithm," *Journal of Hydrology,* vol. 374, no. 3–4, pp. 185-195, 2009.

[15] A. Dahamsheh and H. Aksoy, "Structural Characteristics of Annual Precipitation Data in Jordan,"

Theoretical and Applied Climatology, vol. 88, pp. 201-212, 2007.

[16] H. Aksoy and A. Dahamsheh, "Markov chain-incorporated and synthetic data-supported conditional artificial neural network models for forecasting monthly precipitation in arid regions," *Journal of Hydrology,* vol. 562, pp. 758-779, 2018.

[17] H. Aksoy and A. Dahamsheh, "Artificial neural network models for forecasting monthly precipitation in Jordan," *Stochastic Environmental Research and Risk Assessment,* vol. 23, pp. 917-931, 2009.

[18] A. Dahamsheh and H. Aksoy, "Markov Chain-Incorporated Artificial Neural Network Models for Forecasting Monthly Precipitation in Arid Regions," *Arabian Journal for Science and Engineering,* vol. 39, no. 4, pp. 2513-2524, 2014.