Environment and Natural Resources Research; Vol. 7, No. 4; 2017
ISSN 1927-0488 E-ISSN 1927-0496
Published by Canadian Center of Science and Education

Cascade Rainfall Disaggregation Application in U.S. Central Plains

Jurgen D. Garbrecht[1], Rabi Gyawali[1], Robert W. Malone[2] & John C. Zhang[1]

[1] USDA, ARS, Grazinglands Research Laboratory, El Reno, Oklahoma, USA

[2] USDA, ARS, National Laboratory for Agriculture and the Environment, Agroecosystems Management Research, Ames, Iowa, USA

Correspondence: Jurgen Garbrecht, USDA, ARS, Grazinglands Research Laboratory, 7207 West Cheyenne Street, El Reno, Oklahoma 73036, USA. Tel: 405-262-5291. E-mail: Jurgen.Garbrecht@ARS.USDA.GOV

Received: September 30, 2017 Accepted: October 8, 2017 Online Published: October 27, 2017

doi:10.5539/enrr.v7n4p30 URL: https://doi.org/10.5539/enrr.v7n4p30

Abstract

Long-term observations of daily rainfall are common and routinely available for a variety of hydrologic applications. In contrast, observations of 10 or more years of continuous hourly rainfall are rare. Yet, sub-daily rainfall data are required in rainfall-runoff models. Rainfall disaggregation can generate sub-daily time-series from available long term daily observations. Herein, the performance of Multiplicative Random Cascade (MRC) model at disaggregating daily-to-hourly rainfall was investigated. The MRC model was parameterized and validated with 15 years of continuous observed daily and hourly rainfall data at three weather stations in Oklahoma. Model performance, or degree to which the disaggregated rainfall time series replicated observations, was assessed using 46 variables of hourly rainfall characteristics, such as longest wet spell duration, average number of rainfall hours per year, and largest hourly rainfall. Findings include: a) average-type hourly rainfall characteristics were better replicated than single value characteristics such as longest, maximum, or peak hourly rainfall; b) the large number of sub-trace hourly rainfall values (<0.254 mm h^{-1}) generated by the MRC model were not supported by observations; c) the random component of the MRC model led to a variation under 15% of the average value for most rainfall characteristics with the exceptions of the “longest wet spell duration” and “maximum hourly rainfall”; and d) the MRC model produced fewer persistent rainfall events compared to those in the observed rainfall record. The large number of generated trace rainfall values and difficulties to replicate reliably extreme rainfall characteristics, reduces the number of potential hydrologic applications that could take advantage of the MRC disaggregated hourly rainfall. Nevertheless, in most cases, the disaggregated rainfall generated by the MRC model replicated observed average-type rainfall characteristics well.

Keywords: disaggregation, hourly rainfall, cascade model, downscaling, cascade generator

1. Introduction

Many of today’s hydrologic and agricultural crop models require sub-daily weather data to accurately simulate rainfall-runoff and crop growth processes. Long-duration continuous sub-daily weather data at a location of interest are sparse and typically generated internally by the application model (Bisantino et al., 2015; Li et al., 2015; Moriasi et al., 2012; Malone et al., 2010; Semmens et al., 2008; Jones et al., 2003; Bingner at al., 2001; Laflen et al., 1991 and 1987). The need for long-duration hourly rainfall data by agricultural and hydrologic models is expected to increase as hillside, field, and hydrologic processes are investigated at an increasingly finer temporal resolution.

The United States National Weather Service (US NWS) archives daily rainfall observations from thousands of weather stations across the country, many dating back 50 or more years (National Climatic Data Center [NCDC], 2016; Fiebrich & Crawford, 2009; Bonner, 1998). The likelihood of a weather station with a long record of daily rainfall being located within a reasonable proximity of the site of investigation is high. However, long-term continuous *hourly* rainfall observations at or near the site of interest are rare. Cascade disaggregation models represent a technology that may be useful to generate the desired hourly rainfall from available observed daily rainfall records.

Water resources applications of cascade rainfall disaggregation have been published in recent years. Olsson (1988) modeled the temporal structure of rainfall events with a cascade model and found volume and duration of rainfall events, as well as length of rain-free period to be well reproduced. Güntner et al. (2001) applied a cascade model to six weather stations in Brazil and the UK and compared model performance between the two contrasting climates. They found that the cascade disaggregation model replicated a number of rainfall characteristics with high accuracy. Mueller and Haberlandt (2016) successfully used a cascade model to generate sub-hourly rainfall time series to help with the design of urban sewer and drainage systems. Lisniak et al. (2013) incorporated climate circulation patterns in the probabilities of the cascade generator and disaggregated observed current and projected future daily rainfall. A micro-canonical MRC structure for temporal disaggregation of daily to hourly rainfall was explored by Piu et al. (2009); Carsteanu and Foufoula-Georgiou (1996) and Olson et al. (1998). The term “micro-canonical” implies that the cascade model conserves rainfall amount exactly at each cascade level (Lisniak et al., 2013). Thus, all rainfall entering a node is split and the split rainfall transferred in its entirety to the next cascade level where the process of splitting and transferring repeats itself. Rainfall is not gained nor lost at a node as a result of the rainfall splitting process. The sum of the final disaggregated hourly rainfall is exactly the same as the daily rainfall prior to the disaggregation process. It follows that the initial daily rainfall can be reconstructed by aggregation of the disaggregated hourly rainfall (Mueller and Haberlandt, 2016). The above referenced studies suggest that rainfall disaggregation may potentially be a useful tool for hydrological applications. However, the replication of characteristics of observed hourly rainfall by disaggregated daily rainfall remains a challenge and needs further research.

Here the performance of a cascade rainfall disaggregation model was evaluated for weather conditions prevailing in the US Southern Great Plains. The Multiplicative Random Cascade (MRC) rainfall disaggregation model described in Lisniak et al. (2013) was used to generate hourly rainfall from available daily rainfall observations. Daily and hourly rainfall records at three weather stations in Oklahoma were considered for parameterization and validation of the cascade generation parameters. The main objective was to evaluate the performance of the cascade model in terms of generating disaggregated hourly rainfall that compared to a set of observed hourly rainfall characteristics, including average and extreme hourly rainfall values, and clusters and wet spells. Effects of the random component of the disaggregation process, sensitivity of hourly rainfall characteristics to cascade parameters uncertainties, and practical implications of rainfall measurement errors are addressed. Three weather stations with different annual rainfall characteristics were evaluated to determine consistency in the findings of the performance assessment. This study will help clarify the capabilities and limitations of cascade-type disaggregation methodologies to generate hourly rainfall derived from daily rainfall. Spatial rainfall patterns were outside the scope of this study.

2. Materials and Methods

2.1 Rainfall Data

The three weather stations used in this rainfall disaggregation study were chosen based on the high quality and length (>10 years) of available overlapping daily and hourly rainfall records, on the broad range of annual rainfall (720 to 1170 mm/yr), and on a large number of hourly rainfall characteristics (Table 1). Three weather stations that belong to the Oklahoma Mesonet, a network of climate stations operated and maintained by the Oklahoma Climatological Survey were selected (OCS, 2015; McPherson et al., 2008; Illston et al., 2008; Crawford et al., 2007; Brock et al., 1995). Besides the operation and maintenance of Mesonet data, OCS conducts quality control, data base management, and data dissemination (Shafer et al., 2000).

Table 1. Weather station identification, coordinates in decimal degrees (negative values indicates west), elevation, range of rainfall data, and mean annual rainfall.

Station	Station ID	Agency	Coordinates		Elev. AMSL m	Data range	Annual rainfall mm yr^{-1}
			Lat Deg.	Long Deg.			
Idabel, OK	IDAB	OK Mesonet†	33.83	-94.88	110	1994-2015	1170
El Reno, OK	ELRE	OK Mesonet†	35.55	-98.04	420	1994-2015	840
Cheyenne, OK	CHEY	OK Mesonet†	35.55	-99.73	694	1994-2015	720

† rainfall data available at: https://www.mesonet.org/index.php

The rain gages at the weather stations are a standard tipping bucket with rainfall recording accuracy of 0.254 mm (one tip). With regard to observed daily rainfall, the wet or dry day threshold was subjectively selected as twice the accuracy of the rain gage (0.508 mm). Days with observed rainfall over 2 tips (> 0.508 mm) were considered rainy days, and days with observed rainfall of 2 tips or less (≤ 0.508 mm) were considered dry days. With regards to observed hourly rainfall, the lower threshold value was 1 tip (0.254 mm). Observed hourly rainfall smaller than 1 tip “do not exist” because there is not enough rainfall to produce a tip of the bucket to be recorded as rainfall. Both, daily and hourly rainfall records were assumed to reflect steady state conditions.

2.2 Cascade Disaggregation Model

The Multiplicative Random Cascade (MRC) model described in Lisniak et al. (2013) was selected based on subjective considerations of model simplicity, ease of application, model parameterization based on historical data, and conservations of rainfall volume during all phases of the disaggregation. The cascade branching configuration of Lisniak’s disaggregation model was suitable for disaggregation of daily-to-hourly rainfall (Figure 1). Cascade disaggregation parameters included time-scale, rainfall volume, and position properties (Lisniak et al., 2013; Olsson, 1998). A summary of the branching configuration, cascade generator, and parameter dependencies of the MRC model is given in this section. For additional details on the MRC model, the reader is referred to Mueller and Haberlandt, (2016), Lisniak et al. (2013), Güntner et al. (2001), and Olsson et al. (1998).

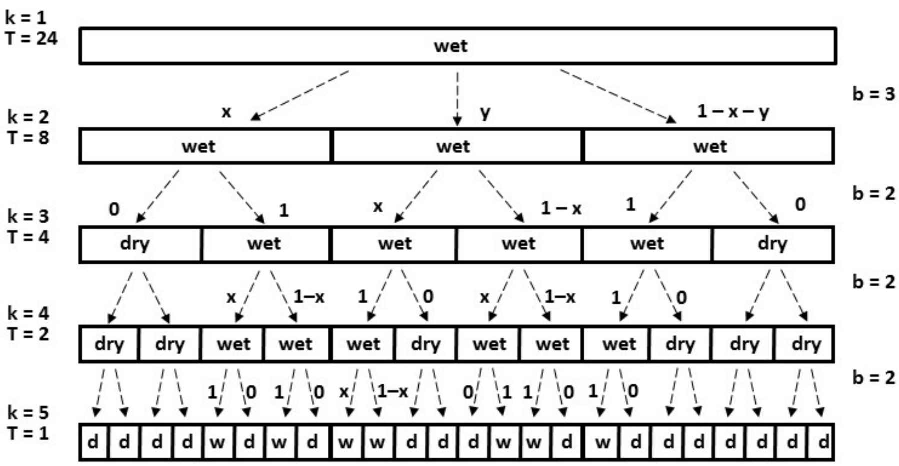

Figure 1. Multiplicative Random Cascade disaggregation model. "k" is cascade level, "T" is time-scale, and "b" is branching number “. Rectangular boxes at each cascade level are referred to as “cells”

Figure adapted from Mueller and Haberlandt (2016).

Cascade disaggregation of daily-to-hourly rainfall involves equally subdividing the daily time increment into successively smaller time increments of 8, 4, 2, and 1 hour durations, called cascade levels. At each cascade level, the rainfall is redistributed into the smaller time increments of the following cascade level according to rules defined by a cascade generator (Güntner et al., 2001).

The branching structure of the MRC model is shown in Figure 1. Temporal disaggregation from daily-to-hourly starts at cascade level, k=1, with time interval, T=24 hours. The first cascade level k=1, with branching b=3 transitions to cascade level k=2 that has three 8-hour time intervals. Subsequent cascade levels, all with branching b=2, transition one cascade level at a time until the twenty-four, 1 hour intervals are reached (Figure 1). During this cascading process branching occurs at wet cells where the rainfall is split according to the generated weights. Generation of weights W for branching wet cells is a random process, where weights W assume a value in the closed interval (0, 1). For branching b=2, the cascade generator assumes one of three states, each state having a given probability of realization:

$$W_1, W_2 = \begin{cases} 1 \; with \; probability \; P(1) \\ 0 \; with \; probability \; P(0) \\ x \; with \; probability \; P(x) \end{cases} \tag{1}$$

where x is a random number between (0,1) from a known probability distribution.

When $W_1 = 1$, the full rainfall amount is assigned to the first branch while the second receives nothing (and vice-versa for $W_1 = 0$). When $W_1 = x$, the rainfall amount is split between both branches. The fraction of the rainfall amount in the first branch is x and the fraction in the second branch is $1 - x$.

For branching b=3, the cascade generator assume one of seven states, with each state having a given probability of realization:

$$W_1, W_2, W_3 = \begin{cases} 1,0,0 & with\ probability\ P(1,0,0) \\ 0,1,0 & with\ probability\ P(0,1,0) \\ 0,0,1 & with\ probability\ P(0,0,1) \\ x, 1-x, 0 & with\ probability\ P(x, 1-x, 0) \\ x, 0, 1-x & with\ probability\ P(x, 0, 1-x) \\ 0, x, 1-x & with\ probability\ P(0, x, 1-x) \\ x, y, 1-x-y & with\ probability\ P(x, y, 1-x-y) \end{cases} \quad (2)$$

where y is a random number between (0,1) from a probability distribution of y conditioned on x. The probability distribution of y is obtained by binning the known histograms of x into 7 equidistant intervals and estimating the empirical probability distribution of y for each bin, given that the corresponding x is inside that interval. The probabilities for the states of the cascade generator and the distributions of x and y are referred to as the parameters of the cascade model.

Grouping of model parameters according to known and/or assumed dependencies generally improves the performance of a cascade model (Olsson, 1998). Dependencies used by Lisniak et al. (2013) are the time resolution, rainfall amount, and position properties. The time resolution classes of the cascade levels are 8, 4, 2 and 1 hours. The two rainfall amount classes considered are “below average” and “above average” rainfall. The four position classes are “isolated” (preceding and succeeding time intervals are dry), “starting” (time interval preceded by dry and succeeded by wet conditions), “ending” (time interval preceded by wet and succeeded by dry conditions), and “contained” (preceding and succeeding time interval are wet). Together they produce 32 time-rainfall-position combinations. Each time-rainfall-position combination is associated with a number of cascade states given by Equations 1 and 2 leading to a total of 128 time-rainfall-position-state combinations.

2.3 Cascade Model Parametrization and Validation

In the context of this study, the term "parameterization" refers to the assignment of probability values to the parameters of the cascade generator. The term "validation" refers to assuring that the daily rainfall disaggregation produces hourly rainfall characteristics that accurately reflect the corresponding observed hourly rainfall characteristics. The simplest form of the MRC model assumes scale-invariance over the five cascade levels and uniform distribution of the $W_{x/x}$. All probability values are assigned during parameterization based on observed hourly rainfall data. Calibration is not required.

Parameterization first requires that the available daily and hourly rainfall time series be divided into a parameterization and validation time period (Table 2). Güntner et al. (2001) designated 2/3 of the years of daily rainfall data for parameterization and the remaining 1/3 years for validation. It is desirable that the rainfall characteristics of the validation and calibration time period be similar. Similarity improves the accuracy, consistency, and reproducibility of the generated hourly rainfall characteristics. The cascade generator parameters are derived directly from the observed hourly rainfall by inverting the cascade structure and aggregating the observed hourly rainfall in reverse order through all cascade levels to result in the original observed daily rainfall (Carsteanu and Foufoula-Georgiou, 1996).

In the validation phase, the probability values of the cascade generator that were determined in the parameterization phase are used to disaggregate the daily rainfall of the validation time period. The characteristics of the resulting disaggregated hourly rainfall are compared to the corresponding characteristics of the observed hourly rainfall and the difference is a measure of performance of the rainfall disaggregation (see next section).

Table 2. Parameterization and validation time periods, number of years of data, mean annual rainfall, and maximum hourly rainfall for each time period

Station	Param. period	# yrs	Annual rain mm yr^{-1}	Max hourly rain mm d^{-1}	Valid period	# yrs	Annual rain mm yr^{-1}	Max hourly rain mm d^{-1}
Idabel	1994-2007	14	1110	51.8	2008-2015	8	1260	59.7
El Reno	1994-2007	14	840	86.1	2008-2015	8	830	59.4
Cheyenne	1994-2007	14	775	83.6	2008-2015	8	630	46.2

2.4 Cascade Model Performance

The performance of the MRC model refers to the degree to which disaggregated rainfall replicates characteristics of the observed rainfall. Model performance involves comparing a variety of observed and disaggregated rainfall variables and characteristics, including rainfall distributions, rainfall clustering, and replication of diurnal rainfall patterns. The measure of performance is the degree of replicability of an observed hourly rainfall characteristic by the disaggregated hourly rainfall, or Relative Error, RE. It is noted that the implicit assumption when using RE for performance measure is that the observed rainfall is always the standard against which the disaggregated rainfall is measured. Thus, the observed data is always the correct value and the error is assumed to be entirely the result of the disaggregation process.

2.5 Rainfall Characteristics

Two sets of rainfall characteristics are used to assess the performance the MRC model. The first set consists of 26 rainfall variables that represent "average-value" type rainfall properties, such as average annual number of hourly rainfall occurrences, average amount of rainfall during wet spells, or, average number of isolated positions (Table 3). Statistics for observed and disaggregated rainfall include mean, standard deviation, sample size, and standard error for all rainfall characteristics.

The performance of the cascade model is assessed by the Relative Error (RE) between observed and disaggregated rainfall. Three ranges of RE are considered. RE range of 0 to 20% defines an "adequate or good" disaggregation, RE range of 20 to 40% defines a "potentially problematic or poor" disaggregation, and RE range of 40% and above defines a "failed" disaggregation. The boundaries of the ranges are chosen to roughly fit about one-third of the RE values in each range. The performance qualifier of each rainfall characteristic is plotted for each station and each rainfall characteristic.

Table 3. Hourly rainfall characteristics used to estimate the replication of observed hourly rainfall (Rainfall Characteristic Data Set 1). POR stands for Period of Record (POR), "Avg" for average, "HR" for hourly rainfall, "#" is number of, "mm" is millimeters, "hr" is hour, "yr" is year, and "%" is percent.

ID	Rainfall characteristic	Units
A1	Avg of (# of HR occurrences per year)	# yr^{-1}
A2	Avg of HR amounts over the POR	mm hr^{-1}
A3	Avg of (HR amounts per year)	mm hr^{-1} yr^{-1}
A4	Avg of (# of high HR occurrences per year)	# yr^{-1}
A5	Avg of the sum of high HR amount per year	mm yr^{-1}
A6	Avg of high HR amount over the POR	mm hr^{-1}
A7	Avg of high HR amount per year	mm hr^{-1} yr^{-1}
A8	Avg of # of low HR occurrences per year	# yr^{-1}
A9	Avg of the sum of low HR amounts per year	mm yr^{-1}
A10A0	Avg of low HR amounts over POR	mm hr^{-1}
A11A1	Avg of low HR amounts per year	mm hr^{-1} yr^{-1}
A12	Avg % of day under rainy conditions over the POR	%
A13	Avg % of day under rainy conditions per year	% yr^{-1}
A14	Avg wet spell amount per year	mm yr^{-1}
A15	Avg of (# clusters of 3 hours and longer per year)	# yr^{-1}
A16	Avg of rainfall amount of clusters of 3 hours and longer over POR	mm
A17	Avg of rainfall duration of clusters of 3 hours and longer over POR	#
A18	Avg of # of isolated HR events per year	# yr^{-1}
A19	Avg of # of contained HR per year	# yr^{-1}
A20	Avg of # of start/end HR per year	# yr^{-1}
A21	Avg of # of isolated high HR per year	# yr^{-1}
A22	Avg of # of contained high HR per year	# yr^{-1}
A23	Avg of # of start/end high HR per year	# yr^{-1}
A24	Avg of # of isolated low HR per year	# yr^{-1}
A25	Avg of # of contained low HR per year	# yr^{-1}
A26	Avg of # of start/end low HR per year	# yr^{-1}

The second set of rainfall data consists of 20 rainfall characteristics that represent "maximum-value" type variables, such as maximum hourly rainfall or number of contained rainfall positions (Table 4). These variables have only one value. The categories and corresponding ranges of RE for this second set of rainfall data are the same as for the first set of rainfall data.

Table 4. Hourly rainfall characteristics used to estimate replication of observed hourly rainfall (Rainfall Characteristic Data Set 2). All rainfall characteristics are determined over the period of record. "#" stands for "number of", "mm" is millimeters, "hr" is hour, and "%" is percent.

ID	Rainfall characteristic	Units
B1	Number of non-zero hourly rainfall occurrences	#
B2	Sum of hourly rainfall amounts	mm
B3	Max hourly rainfall amount	mm hr^{-1}
B4	Number of non-zero low hourly rainfall occurrences	#
B5	Sum of low hourly rainfall amounts	mm
B6	Number of non-zero high hourly rainfall occurrences	#
B7	Sum of high hourly rainfall amounts	mm
B8	Max % of day under rainy conditions	%
B9	Longest wet spell duration (hr)	hr
B10	Largest wet spell amount (mm)	mm
B11	Number of clusters of hourly rainfall lasting 3 hours and longer	#
B12	Number of isolated hourly rainfall occurrences	#
B13	Number of contained R occurrences	#
B14	Number of start/end hourly rainfall occurrences	#
B15	Number of isolated high hourly rainfall occurrences	#
B16	Number of contained high hourly rainfall occurrence	#
B17	Number of start/end high hourly rainfall occurrences	#
B18	Number of isolated low hourly rainfall occurrences	#
B19	Number of contained low hourly rainfall occurrences	#
B20	Number of start/end low hourly rainfall occurrences	#

Rainfall characteristics in Table 3 and 4 are evaluated for the three weather stations. Rainfall characteristics that are similar or related to one another are grouped for each weather station. This is expected to facilitate visual identification of common relationships and disaggregated rainfall characteristics that are better replicated than others. Rainfall characteristics A4 through A7 and B6 and B7 are related to high-flow conditions, and low-flow conditions are covered by A8 through A11, and B4 and B5. Rainfall characteristics A14 through A17 and B9 through B11 address wet spell and cluster properties. All position class rainfall characteristics are A18 through A26 and B12 through B20.

2.6 Hourly Rainfall Distributions

Similarities and dissimilarities in the distributions of observed and disaggregated hourly rainfall over the period of record at the three weather stations are visualized by a probability-of-exceedance (PoE) plot of hourly rainfall. The Wilcoxon-Mann-Whitney test is used to determine whether the distributions of the disaggregated hourly rainfall is significantly different from the observed rainfall distribution.

2.7 Clusters

The capacity of the MRC model to accurately replicate prolonged, uninterrupted hourly rainfall is established by counting the number of clusters of given duration and comparing the frequency of cluster occurrence in the disaggregated hourly rainfall record to the frequency of clusters in the observed hourly rainfall record.

2.8 Diurnal Rainfall Patterns

Diurnal rainfall patterns occur when certain hours of the day systematically receive more or less rainfall than average. Replication of an observed diurnal rainfall pattern by disaggregated rainfall is assessed in two steps. First, the presence of a diurnal rainfall pattern in the observed hourly rainfall is established. This involves filtering the high hour-to-hour variability of the observed rainfall by a 5-point moving average that will help visualize and confirm the presence of potential diurnal patterns. A diurnal pattern is said to be present when the amplitude of the smoothed observed hourly rainfall is larger than the average width of the 90% confidence interval of the observed hourly rainfall. Second, for established diurnal rainfall patterns, the coefficient of variation is used to measure the degree of replicability of observed diurnal rainfall patterns by disaggregated rainfall.

2.9 Random Numbers

Outcomes of the rainfall disaggregation are typically determined by random numbers drawn from a standard uniform distribution. Different sets of random numbers lead to different values of disaggregated hourly rainfall. Estimating the effect of random numbers on the value of rainfall characteristic is necessary for interpreting the variability of rainfall characteristics with those originating from random numbers. The effect of random numbers on disaggregated hourly rainfall characteristics (Table 5) is illustrated for daily rainfall at the El Reno weather station for the 8-year study period (2008-2015). A total of 20 repeat daily rainfall disaggregations are performed each with the same 2008-2015 daily rainfall at the El Reno weather station, with the only difference that each of the disaggregations is driven by a different set of random numbers. Fifteen disaggregated hourly rainfall characteristics are retained from each of the 20 disaggregations and used to calculate the average, standard deviation, coefficient of variation, and relative error of each hourly rainfall.

Table 5. Disaggregated rainfall characteristics used to test sensitivity to different random number sets. "Avg" stands for average, "POR" for period-of-record, "mm" for millimeter, "hr" for hour, "yr" for year, and "#" for number

ID	Rainfall characteristic	Reference to tables 3 and 4	Units
C1	Avg hourly rainfall over the POR	A2	mm hr^{-1}
C2	Avg annual maximum hourly rainfall		mm hr^{-1}
C3	Maximum hourly rainfall	B3	mm hr^{-1}
C4	Avg annual # of high rainfall occurrences	A4	# yr^{-1}
C5	Maximum annual # high rainfall events		# yr^{-1}
C6	Avg annual # of rainfall occurrences	A1	# yr^{-1}
C7	Maximum annual # of rainfall events		# yr^{-1}
C8	Avg # rainfall events below 0.1 mm		# yr^{-1}
C9	Avg % of day under rainy conditions over POR	A12	%
C10	Longest wet spell duration	B9	hr
C11	Avg wet spell duration of clusters of 3 hrs and longer	A17	hr
C12	Avg wet spell rainfall amount of clusters 3hr and >	A16	mm
C13	Avg # isolated rainfall events per year	A18	# yr^{-1}
C14	Avg # contained rainfall events per year	A19	# yr^{-1}
C15	Avg # start/end rain events per year	A20	# yr^{-1}

3. Results

3.1 Rainfall Characteristics

The Relative Error (RE) between observed and disaggregated rainfall is shown in Figure 2 for the average and standard deviation of 26 rainfall characteristics for each of the weather stations. The identification number in the first column in Figure 2 is the key to the rainfall characteristics listed in Table 3. Gray cells in Figure 2 have an RE under 20% (good/adequate disaggregation). Cells with the black and white hash marks have an RE between 21% and 40% (problematic disaggregation). Black cells have an RE greater than 40% (failed disaggregation).

Figure 2 shows a clear difference in the pattern between the RE of the average (left side and for all stations) and the RE of the standard deviation (right side and for all stations). For the RE of the average, there are 19 out of 26 rainfall characteristics (73%) that show good disaggregation results, 4 out of 26 rainfall characteristics (14%) showed a problematic disaggregation at 1 or more stations, and 3 out of 26 rainfall characteristics (11%) failed the disaggregation at all three stations. Rainfall characteristics that performed least were the average number per year of isolated, contained, and start/end positions under high hourly rainfall events (A21 and A23 in Table 3).

For RE of the standard deviation, there was no clear pattern in RE (Figure 2 right side). 42% of 78 cases showed a good disaggregation, 27% a problematic disaggregation, and 31% a failed disaggregation. Weather station Cheyenne exhibited a slightly better disaggregation than at the other two stations with as many as 15 good disaggregation cases and only 4 failed disaggregation cases. Two rainfall characteristics that failed at all three stations are average of low hourly rainfall amounts per year and average number of isolated hourly high rainfall per year (A11 and A21 in Table 3).

With regard to rainfall characteristics of Table 4, 12 out of 20 rainfall characteristics showed a successful disaggregation, and 2 out of 20 rainfall characteristics (number of isolated and start/end occurrences under high hourly rainfall conditions, B15 and B17) failed at all three stations (Figure 3). Rainfall characteristics that displayed failure at 2 or 3 stations included the maximum hourly rainfall and the longest wet spell duration (B3 and B9 in Table 4).

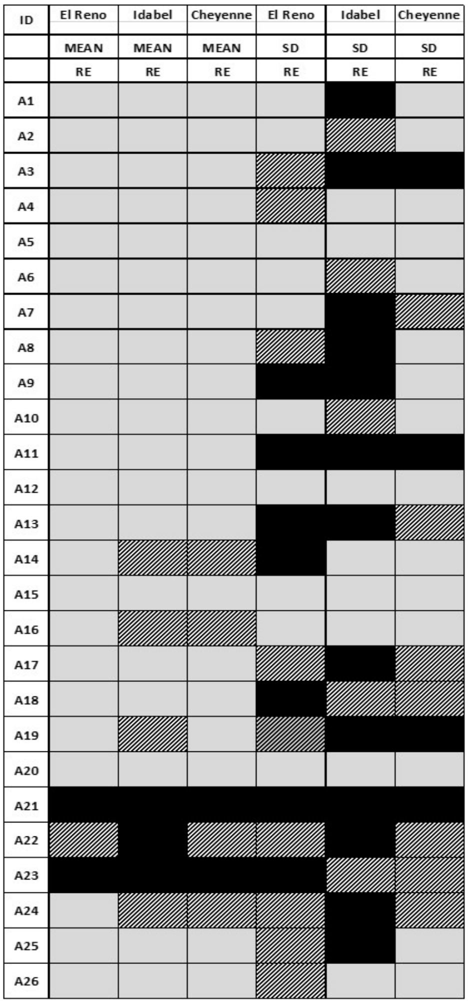

Figure 2. Relative error (RE) of mean and standard deviation between observed and disaggregated hourly precipitation. Categories are based on rainfall characteristics given in Table 3

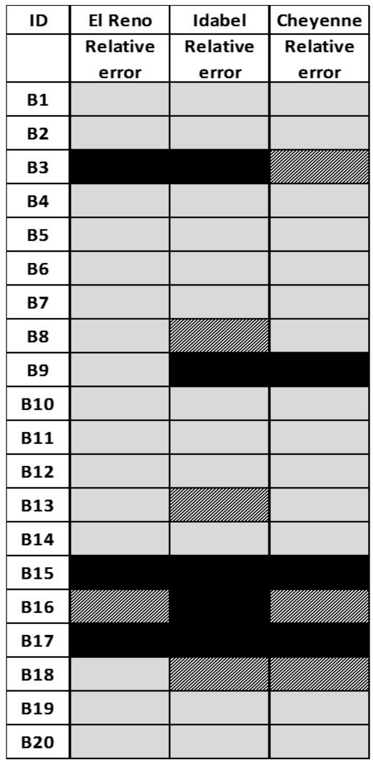

Figure 3. Relative error (RE) between observed and disaggregated hourly rainfall characteristics. Categories are based on values of rainfall characteristics given in Table 4

3.2 Hourly Rainfall Distributions

PoE distributions of observed and disaggregated hourly rainfall are shown in Figure 4. The discrete vertical bars represent the range of observed hourly rainfall. About 30% of observed hourly rainfall is at the lowest observable threshold of 0.254 mm (1 tip). About 45% of the observed hourly rainfall have a rainfall amount of 0.508 mm or less, and 55% have a rainfall amount of 0.762 mm or less. The smooth lines in Figure 4 represent the disaggregated hourly rainfall. About 25% of the hourly rainfall values are lower than 0.254 mm, 35% are lower than 0.508 mm, and 45% are lower than 0.762 mm.

The PoE distributions curves for both observed and disaggregated hourly rainfall are very similar in shape. Only the disaggregated rainfall at Idabel display higher rainfall values in the 20 to 80 percentile range than the other stations. The results of the Wilcoxon-Mann-Whitney test suggest that two samples, observed and disaggregated hourly rainfall, had statistically different hourly rainfall distributions at an alpha level of 0.1 (α=0.1).

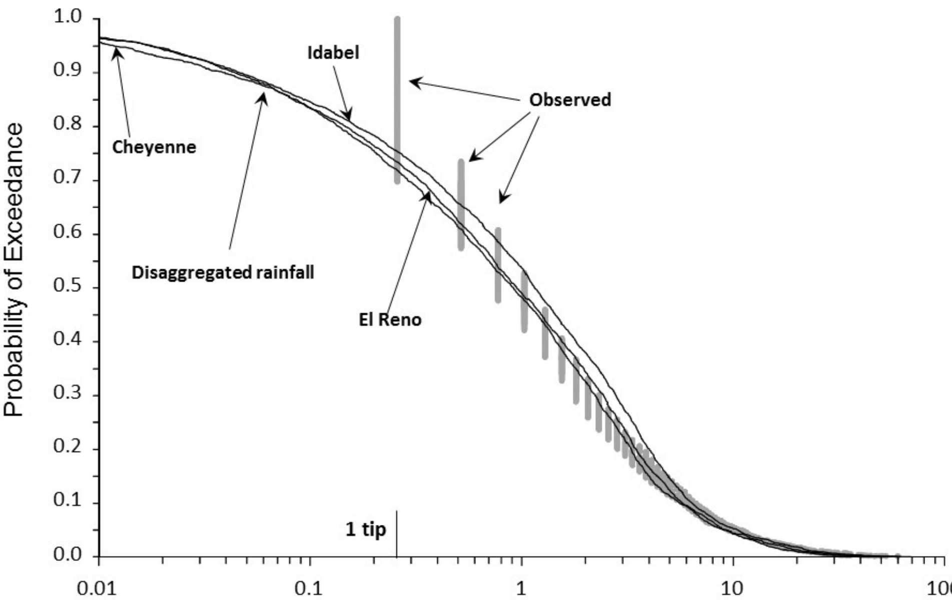

Figure 4. Probability-of-Exceedance curves of observed and disaggregated hourly rainfall for the validation time period. Vertical bars represent observed hourly rainfall. The lowest observable rainfall value is 0.254 mm (1 tip)

3.3 Clusters

Clusters of hourly rainfall are an important rainfall characteristic related to reproducing storm duration. The capacity of the MRC model to accurately replicate prolonged, uninterrupted hourly rainfall is shown in Figure 5.

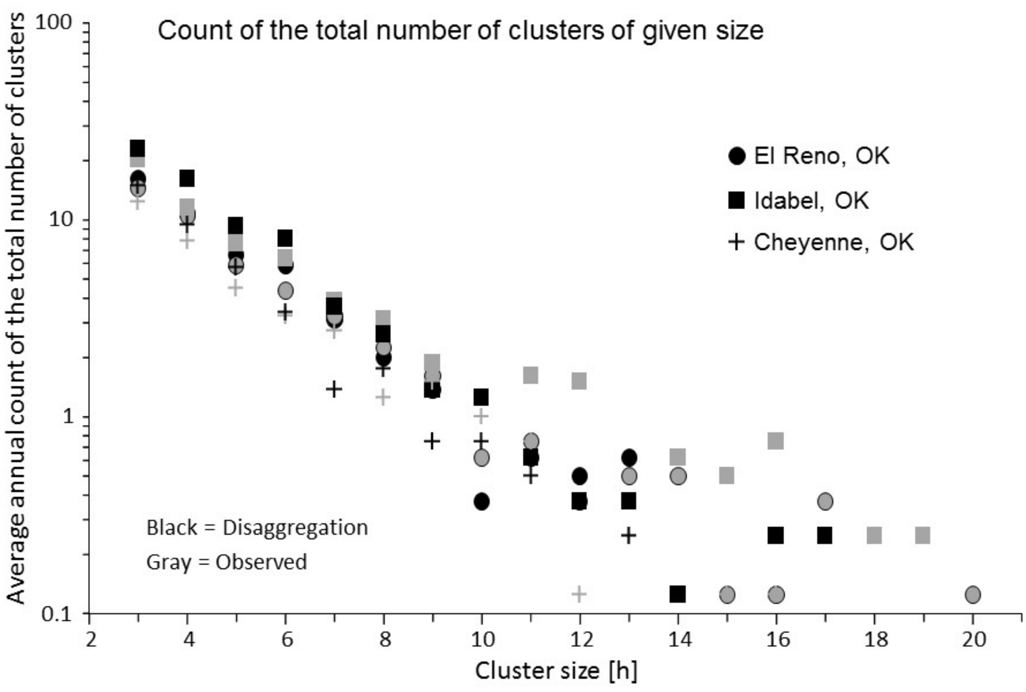

Figure 5. Disaggregated (DIS) and corresponding observed (OBS) average annual number of rainfall clusters (3 or more clusters) of various durations for the validation time period

Increasing cluster duration is accompanied by an exponential decrease in number of clusters of given duration. For weather stations considered in this study, the exponential relationship is stable up to a cluster duration of about 13 hours. Both disaggregated and observed cluster follow the same exponential trend. For cluster duration of 3 to 7 hours, the cluster count of the disaggregated rainfall is in most cases slightly higher than that for the observed rainfall (approximately 10%). For cluster duration above 7 hours, the cluster count of observed rainfall is usually higher than that for disaggregated rainfall. There are very few clusters of disaggregated rainfall of duration of 15 and higher. The close correspondence in number and duration of clusters under 15 hour duration (99% of cases) affirms that the disaggregation process produces clusters of disaggregated rainfall that resembles that of the observed hourly rainfall at the weathers stations. However, few large clusters above 15 hours duration are reproduced by the disaggregation.

3.4 Diurnal Rainfall Patterns

Characteristics of the diurnal rainfall patterns are given in Table 6. Rainfall at Idabel do not display a diurnal rainfall pattern (amplitude of diurnal rainfall pattern smaller than 90% confidence interval). At the other two weather stations a diurnal pattern was present (amplitude of diurnal rainfall pattern larger than confidence interval). The diurnal rainfall pattern at El Reno and Cheyenne were adequately replicated with a coefficient of determination of 0.84 and 0.88, respectively (Figure 6a and 6b).

Table 6. Amplitude of smoothed diurnal rainfall pattern, average width of 90% confidence interval (CI), ratio of amplitude of smoothed diurnal rainfall over average width of 90% CI, and coefficient of correlation of observed versus smoothed disaggregated rainfall

Station Name	Amplitude of smoothed diurnal rainfall [mm]	Avg. width of CI of smoothed rainfall [mm]	Ratio of amplitude over an average width of CI [-]	Coeff. of correlation of observed versus disaggreg. rainfall [-]
Idabel	14.5	26.9	0.54	-0.31
El Reno	22.0	19.9	1.10	0.84
Cheyenne	18.0	18.1	0.99	0.88

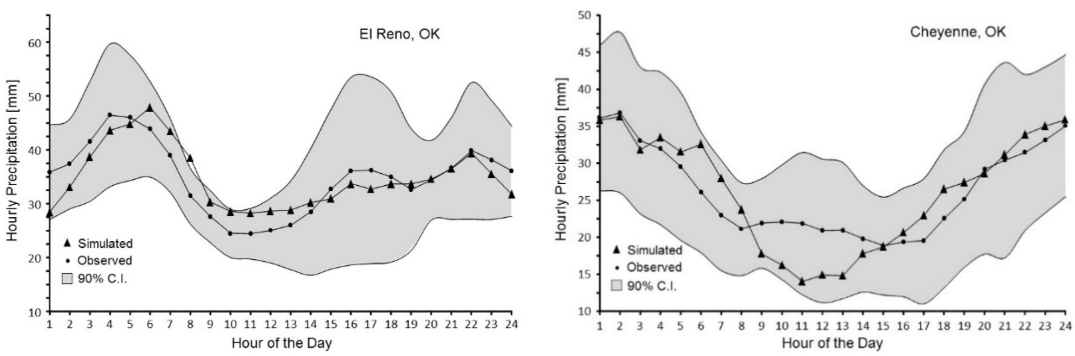

Figure 6. Observed and disaggregated diurnal rainfall pattern at (a) El Reno, Oklahoma, and (b) Cheyenne, Oklahoma. The shaded region represents the 90% confident interval of the observed diurnal rainfall

3.5 Random Number Sensitivity

Twenty repeat rainfall disaggregations were performed with the 2008-2015 El Reno daily precipitation (Table 7). Each repeat disaggregation used different random numbers, leading to twenty different values for each hourly rainfall characteristic given in Table 5. Sensitivity was measured relatively as the rainfall characteristic with the highest standard error, coefficient of variation, root mean squared error, and standard deviation. Rainfall characteristics with lowest standard error, coefficient of variation, root mean squared error, and standard deviation were considered to be least sensitive to random numbers. The two hourly rainfall characteristics out of 15 that were the most sensitive to different random numbers were the maximum hourly rainfall over the POR and the longest wet spell duration over the POR (bold print). Rainfall characteristics that were found to be moderately sensitive to random numbers included the maximum number of high rainfall events over the POR, mean annual number of rainfall events under 0.254 mm, and maximum annual number of rainfall events over the POR. Likewise, the five rainfall characteristics out of 15 that were the least sensitive to random numbers were mean hourly rainfall, mean fraction of the day under rainfall, mean wet spell duration, mean number of hours of rain, and mean number of starting and ending rainfall events (italic print).

Table 7. Average hourly rainfall characteristics produced by 20 repeat rainfall disaggregations using different random numbers. "Mean" in the first column is a mean over the 8 years, whereas "maximum" and "longest" signifies the largest value over the 8 years. Values in bold represent five greatest coefficient of variation (CV), standard error, and RMSE. Values in italic represent the smallest CV, standard error, and RMSE

ID	Rainfall characteristic		Units	Avg	StDev	COV	Std. Error	RMSE
C1	Avg hourly rain	A2	mm hr^{-1}	2.5	*0.06*	*0.023*	*0.01*	*0.0*
C2	Avg annual maximum hourly rain		mm hr^{-1}	43.4	4.11	**0.095**	0.92	3.88
C3	**Maximum hourly rain**	B3	mm hr^{-1}	77.8	**17.10**	**0.220**	**3.82**	**16.60**
C4	Avg annual # of high rain events	A4	# yr^{-1}	94.5	3.70	0.039	0.83	3.13
C5	Maximum annual # high rain events		# yr^{-1}	141.2	**7.35**	0.052	**1.64**	**7.10**
C6	Avg annual # of rain events	A1	# yr^{-1}	331.1	**7.72**	*0.023*	**1.73**	**6.37**
C7	Maximum annual # of rain events		# yr^{-1}	452.9	**24.52**	**0.054**	**5.48**	**23.55**
C8	Avg # rain events below 0.1 mm		# yr^{-1}	51.3	3.71	**0.072**	0.83	3.47
C9	Avg wet portion of rainy day	A12	%	20.7	0.48	*0.023*	*0.11*	*0.34*
C10	**Longest wet spell duration**	B9	hr	19.7	**5.19**	**0.264**	**1.16**	**4.98**
C11	Avg wet spell duration	A17	hr	4.8	*0.10*	*0.021*	*0.02*	*0.10*
C12	Avg wet spell rainfall amount	A16	mm	12.2	*0.41*	0.034	*0.09*	0.39
C13	Avg # isolated rain events	A18	# yr^{-1}	42.6	2.07	0.049	0.46	2.01
C14	Avg # contained rain events	A19	# yr^{-1}	133.7	5.17	0.039	1.16	4.48
C15	Avg # start/end rain events	A20	# yr^{-1}	77.5	2.36	0.030	0.53	2.11

4. Discussion

4.1 Rainfall Characteristics

A closer look at the RE patterns in Figures 2 and 3 revealed that disaggregation A21, A22, A23, B15, B16, and B17 resulted in hourly rainfall that mostly failed to replicate observed rainfall characteristics for both the RE of the mean and standard deviation at the three stations. The six rainfall characteristics relate to the count of position classes (isolated, contained, starting and ending) under high rainfall.

Rainfall characteristics A24, A25, A26, B18, B19 and B20 are also related to the count of position classes with the difference that only low rainfall events are considered. In this case, the disaggregation was successful and produced hourly rainfall that mostly replicated the observed rainfall characteristics. This suggest that properties of "high rainfall" may be the underlying cause that led to the poor disaggregation performance. High rainfall events are much fewer in number (~30% of total), of lower frequency, and cover a broader range of values (2.4 to 56 mm/hr) than low rainfall events (0 to 2.4 mm/hr) making it more difficult to accurately capture high rainfall characteristics in the parameterization of the cascade model.

4.2 Sub-Trace Rainfall Values

The disaggregation process often generates hourly rainfall amounts that are smaller than the rain gauge measuring accuracy (Muller and Haberlandt, 2015; Molnar and Burlando, 2005). The implication is displayed in the PoE plot of the disaggregated hourly rainfall (Figure 4). The smallest disaggregated hourly rainfall is 1/100 of a millimeter, an inconsequential and physically meaningless rainfall amount. Four options are available to resolve or alleviate the effects of a large number of sub-trace rainfall amounts.

Option 1: accept the sub-trace rainfall amounts as calculated.

Option 2: reset sub-trace rainfall amounts to zero (“rain-free” conditions).

Option 3: reset all sub-trace rainfall values to 1 tip (minimum rainfall).

Option 4: prevent the cascade model from splitting rainfall into sub-trace values.

Each of these four options have implications for rainfall disaggregation. Option 1 preserves the mathematically induced sub-trace rainfall amounts. A practitioner would be inclined to set these physically meaningless small rainfall values to zero (Option 2). However, rounding trace values to zero can lead to relevant differences in statistics when calculated with and without trace values. Also, large number of sub-trace rainfall occurrences promotes creation of continuous chains of hourly rainfall, thereby directly affecting the number and duration of rainfall clusters. Resetting the sub-trace rainfall events to zero leads to a rapid brake-up of large clusters. In Option 3, large rainfall clusters and position properties are generally preserved, and basic statistics that depend on number of rainfall events remain unchanged. Similar to Option 2, Option 3 would create noticeable differences in statistics related to rainfall volume and average hourly rainfall. Option 4 would require overriding the probabilities of the cascade generator. The impact that a modified disaggregation model would have on the disaggregated hourly rainfall is unclear as it depends on the type and extent of the modification.

4.3 Uncertainty of Recorded Rainfall

Only rainfall that tip the rain gage bucket are recorded as rain. Any tips due to under- or over-catch, measurement errors, timing of tips, or instrument malfunctions may affect the quality of recorded rainfall amount and subsequent temporal disaggregation. Snow, sleet, and icing represents such a case. The three weather stations considered in this investigation are not heated rain gages. Hence, snowfall, sleet and effects of icing were recorded as rainfall at some later time when the snow/ice melts. This may lead to inconsistencies between weather variables when attempting to synchronize measured rainfall (i.e. snow/ice melt) with other weather related variables or processes. Idabel, El Reno, and Cheyenne stations showed no appreciable snowfall. When snowfall is a relevant fraction of annual precipitation, then rainfall disaggregation should only be performed during the warmer seasons of the year.

4.4 Data Requirements for MRC Parameterization

The length of the observed rainfall record for MRC parameterization is climate dependent. In semi-arid climates rainy days are fewer than in humid climates, and a longer record of observations is needed to accumulate sufficient number of rainy days for MRC parameterization. The number of time-rainfall-position-state combinations of the MRC model must also be considered. For example, if there are 128 different time-rainfall-position-state combinations, as in this study, and if it is assumed that it takes an average of 15 hits per combination to produce reliable probability of occurrence values, then it would take a minimum of 1920 hits to calculate the probability of occurrence of all combinations. Note that the "15 hits" is a subjective value that likely errors on the low side. Increasing the number of hits would require long hourly rainfall data set that are increasingly difficult to come by. For central Oklahoma climate conditions with an average of 5 hits per rainy day (5 hours of actual rainfall per rainy day) and with 92 rainy days per year, a minimum record length of about 5 years of observed hourly rainfall would be required for parameterization of the MRC model. Thus, for a given climate and a given MRC structure (probability combinations), the average number of rainy days per year and the average number of hours of actual rainfall on a rainy day are factors that determine the required number of years of observed hourly rainfall for MRC model parameterization. Another 2 or 3 years of hourly rainfall is needed for model validation.

5. Summary and Conclusions

The performance of a Multiplicative Random Cascade (MRC) model for daily rainfall disaggregation was investigated. The MRC model was parameterized and validated with observed daily and hourly rainfall at three weather stations in Oklahoma. The relative error between disaggregated and observed hourly rainfall was the measure of model performance. Model performance was evaluated for forty six rainfall characteristics. The following are findings of this study on the performance of the MRC disaggregation model under climate characteristics of the Southern Plains of the United States.

- "Average-value" type rainfall characteristics, such as average hourly rainfall or average number of rainfall occurrences, generally produce disaggregated hourly rainfall that replicated more accurately the observed rainfall characteristics compared to "max/min-value" type rainfall characteristics, such as longest wet spell duration or highest hourly rainfall.
- Rainfall characteristics involving rainfall position, such as number of isolated or contained occurrences of rainfall, were poorly replicated by disaggregated hourly rainfall for all stations.
- A diurnal pattern in observed hourly rainfall was replicated at two stations by disaggregated hourly rainfall. However, subjective definition on what constitutes a diurnal pattern was necessary, and the high hour-to-hour rainfall variability required smoothing to reveal the diurnal patterns.

- The cascade models often generates hourly rainfall values below a trace amount that are not present in the observed hourly rainfall record. Yet, the sum of all trace rainfall amounts can exceed 20 to 30% of total rainfall.
- Variability of hourly rainfall characteristics due to the random component of the MRC model was within the range of 0 to 20% (qualified as "adequate or good"). Two exceptions were "maximum hourly rainfall amount" and "longest wet spell duration" which had a relative error of 31% and 44%, respectively.
- For cold seasons when snowfall does not melt on the same day as it fell, the property of water preservation of daily rainfall is violated. Thus, for regions that have a relevant portion of annual precipitation as snow, it is recommended that rainfall disaggregation be restricted to the warm seasons of the year. Snow fall is not an issue when the rain gage is heated.

The main conclusion drawn from this study was that the MRC disaggregation model represents a potential technology to generate hourly rainfall from daily rainfall data. However, extreme rainfall characteristics that are generally of primary interest to the hydrologist, such as peak hourly rainfall or longest wet spell, are also the characteristics most difficult to replicate accurately by disaggregated hourly rainfall.

Acknowledgments

The authors thank Dr. G. Brown at the Biosystems and Agricultural Engineering Department of the Oklahoma State University, Stillwater, Oklahoma, for his insightful and constructive comments, remarks, and suggestions that improved this manuscript. Thanks are also extended to P. Busteed for his assistance in the programing and application of the MRC model, including data processing and drafting of the figures. The authors would like to also thank the editors and anonymous reviewers for their helpful comments.

References

Bingner, R. L., & Theurer, F. D. (2001, March). AnnAGNPS: estimating sediment yield by particle size for sheet and rill erosion. In *Proceedings of the Seventh Interagency Sedimentation Conference, Reno, NV* (Vol. 1, pp. 1-7).

Bisantino, T., Bingner, R. L., Chouaib, W., Gentile, F., & Liuzzi, G. T. (2015). Estimation of runoff, peak discharge and sediment load at the event scale in a medium-size Mediterranean watershed using the AnnAGNPS model. *Land Degradation & Development, 26*, 340-355.

Bonner, W. D. (1998). *Future of the national weather service cooperative observer network*. National Research Council, National Academy Press, Washington, D.C.

Brock, F. V., Crawford, K. C., Elliott, R. L., Cuperus, G. W., Stadler, S. J., Johnson, H. L., & Eilts, M. D. (1995). The Oklahoma Mesonet: a technical overview. *Journal of Atmospheric and Oceanic Technology*, *12*(1), 5-19.

Carsteanu, A., & Foufoula-Georgiou, E. (1996). Assessing dependence amoung weights in a multiplicative cascade model of temporal rainfall. *Journal of Geophysical Research, 101*(D21), 26363-26370.

Fiebrich, C. A., & Crawford, K. C. (2009). Automation: A step toward improving the quality of daily temperature data produced by climate observing networks. *Journal of Atmospheric and Oceanic Technology, 26*, 1246-1260.

Güntner, A.,Olsson, J., Calver, A., & Gannon, B. (2001). Cascade-based Disaggregation of continuous rainfall time series: the influence of climate. *Hydrology and Earth System Sciences Discussions, European Geosciences Union, 5*(2), 145-164. Retrieved from https://hal.archives-ouvertes.fr/hal-00304590

Illston, B. G., Basara, J. B., Fiebrich, C. A., Crawford, K. C., Hunt, E., Fisher, D. K., ... & Humes, K. (2008). Mesoscale monitoring of soil moisture across a statewide network. *Journal of Atmospheric and Oceanic Technology*, *25*(2), 167-182.

Illston, B. G., Basara, J. B., Fischer, D. K., Elliott, R. L., Fiebrich, C. A., Crawford, K. C., Humes, K., & Hunt, E., (2008). Mesoscale Monitoring of Soil Moisture Across a Statewide Network. *J. of Atmos. and Oceanic Tech, 25*, 167-182.

Jones, J. W., Hoogenboom, G., Porter, C. H., Boote, K. J., Batchelor,W. D., Hunt, L. A., … Ritchie, J. T. (2003). DSSAT cropping system model. *European Journal of Agronomy, 18*, 235-265.

Laflen, J. M., Lane, L. J., & Foster, G. R. (1991). The water erosion prediction project - a new generation of erosion prediction technology. *Journal of Soil and Water Conservation, 46*(1), 34-38.

Laflen, J. M., Thomas, A., & Welch R. (1987). *Cropland experiments for the WEPP project*. ASAE paper No. 87-2544, American Society of Agricultural Engineers, St. Joseph, MI.

Li, Z., Luo, C., Xi, Q., Li, H., Pan, J., Zhou, Q., & Xiong, Z. (2015). Assessment of the AnnAGNPS model in simulating runoff and nutrients in a typical small watershed in the Taihu Lake basin, China. *Catena*, *133*, 349-361.

Lisniak, D., Franke, J., & Bernhofer, C. (2013). Circulation Pattern Based Parameterization of a Multiplicative Random Cascade for Disaggregation of Observed and Projected Daily Rainfall Time Series. *Hydrol. Earth Syst. Sci., 17*, 2487-2500.

Malone, R. W., Jaynes, D. B., Ma, L., Nolan, B.T., Meek, D. W., & Karlen, D. L. (2010). Soil-Test N Recommendations Augmented with PEST-Optimized RZWQM Simulations. *Journal of Environmental Quality, 39*(5), 1711-1723.

McPherson, R. A., Fiebrich, C. A., Crawford, K. C., Kilby, J. R., Grimsley, D. L., Martinez, J. E., ... & Melvin, A. D. (2007). Statewide monitoring of the mesoscale environment: A technical update on the Oklahoma Mesonet. *Journal of Atmospheric and Oceanic Technology*, *24*(3), 301-321.

Molnar, P., & Burlando, P. (2005). Preservation of Rainfall Properties in Stochastic Disaggregation by a Simple cascade Model. *Atmospheric Research, 77*, 137-151.

Moriasi, D. N., Wilson, B. N., Douglas-Mankin, K. R., Arnold, J. G., & Gowda, P. H. (2012). Hydrologic and Water Quality models: Use, Calibration, and Validation. *Transactions of the ASABE, 55*(4), 1241-1247

Mueller, H., & Haberlandt, U. (2016). Temporal rainfall disaggregation using a multiplicative cascade model for spatial application in urban hydrology. *Journal of Hydrology*. http://dx.doi.org/10.1016/j.jhydrol.2016.01.031

Müller, H., & Haberlandt, U. (2015). Temporal rainfall disaggregation with a cascade model: from single-station disaggregation to spatial rainfall. *Journal of Hydrologic Engineering*, *20*(11), 04015026.

National Climatic Data Center. (2016). *Climate Data Online (CDO)*. Retrieved from https://www.ncdc.noaa.gov/cdo-web/.

Oklahoma Climatological Survey. (2015). *Oklahoma Climatological Survey*. Retrieved from http://climate.ok.gov/index.php/site/page/about.

Olsson, J. (1998). Evaluation of a Scaling Cascade Model for Temporal Rainfall Disaggregation. *Hydrology & Earth System Sciences, 2*(1), 19-30. http://dx.doi.org/10.5194/hess-2-19-1998

Pui, A., Sharma, A., Mehrotra, R., Sivakumar, B., & Jeremiah, E. (2012). A comparison of alternatives for daily to sub-daily rainfall disaggregation. *Journal of hydrology*, *470*, 138-157.

Semmens, R. G., Goodrich, D. C., Unkrich, C. L., Smith, R. E., Woolhiser, D. A., & Miller, S. N.. (2008). KINEROS2 and the AGWA Modeling Framework. In H. Wheater, S. Sorooshian, & K.D. Sharma (Eds.), *Chapter 5: In Hydrological Modeling in Arid and Semi-Arid Areas* (pp. 49-69). Cambridge University Press, London.

Shafer, M. A., Fiebrich, C. A., Arndt, D. S., Fredrickson, S. E., & Hughes, T. W., (2000). Quality assurance procedures in the Oklahoma Mesonet. *J. Atmos. Oceanic Tech., 17*, 474-494.