

# Weather Generator Effectiveness in Capturing Climate Extremes

Margaret W. Gitau<sup>1</sup>  · Sushant Mehan<sup>1</sup> · Tian Guo<sup>2</sup>

Received: 20 December 2017 / Accepted: 7 March 2018 / Published online: 17 March 2018  
© Springer International Publishing AG, part of Springer Nature 2018

**Abstract** Weather generators are increasingly used in environmental, water resources, and agricultural applications. Given their potential, it is important that weather generators be evaluated, particularly with respect to their ability to capture extreme events. This study was aimed at evaluating weather generator representation of climate extremes with a focus on LARS-WG applied to three stations in the Western Lake Erie Basin, U.S. Generally, LARS-WG captured the number of days with precipitation greater than 50.8 mm (2 in. and 101.6 mm (4 in), 7-day wet sequences, and wet and dry sequences relatively well. The distribution of 1-day maximum precipitation was also generally captured well based on Q-Q plots, although large deviations were seen at the upper tail at one of the stations. The generator greatly underestimated the number of days per year with maximum temperatures greater than 32.2 °C (90 °F) and overestimated the number of days with temperatures less than 0 °C (32 °F). It also underestimated spring and summer values of one-day maximum temperatures across all stations. Fall and winter values were, however, captured fairly well as were seasonal values of one-day minimum temperatures. Overall, the generator performed relatively well in representing extremes within the basin.

**Keywords** Weather generators · Extreme variables · Realizations · Modelling · Water resources · Environmental

## 1 Introduction

Weather generators are increasingly used in environmental, water resources, and agricultural applications. This is especially so where observed data are not consistent or available

---

✉ Margaret W. Gitau  
mgitau@purdue.edu

<sup>1</sup> Department of Agricultural and Biological Engineering, Purdue University, 225 S. University St., West Lafayette, IN 47907, USA

<sup>2</sup> National Center for Water Quality Research, Heidelberg University, Tiffin, OH 44883, USA

in sufficient quantities for these applications (Yu 2003; Fodor et al. 2013; Lemann et al. 2016). In climate related studies, the generators are used to obtain localized predictions based on observed data (Al-Mukhtar et al. 2014) and/or downscale from coarse resolution predictions (Semenov and Barrow 1997; Fatichi et al. 2011; Farzanmanesh et al. 2012). Given this potential, there is the need to evaluate weather generators to determine the suitability of generated data for the intended application. This is particularly so with respect to their ability to capture extreme events given the projected changes in their magnitude and severity and the challenges they present with respect to impact (Klein Tank et al. 2009). A lot of work has been conducted to determine weather generator effectiveness in capturing the primary statistical characteristics of observed data—such as the daily mean, standard deviation, skewness, kurtosis, and percentiles (e.g., Zhang and Garbrecht 2003; Min et al. 2011; Chen and Brissette 2014), and other essential characteristics, as described in Gitau et al. (2017). For example, Mehan et al. (2017) evaluated three commonly used weather generators with respect to reproducing the distribution, occurrence of wet and dry spells, number of snow days, growing season temperatures, and growing degree days in relation to those from observed data.

A relatively less amount of work has been done on generator effectiveness in capturing the extremes with primary works including Kyselý and Dubrovský (2005), Semenov (2008), Hundscha et al. (2009), and Acharya et al. (2017). Extreme metrics assessed in these works include: mean annual maxima of daily precipitation and maximum temperatures; mean annual minima of minimum temperatures; distributions of annual 1-day maxima; lengths of heat and cold waves; 10- and 20-year return periods determined based on Generalized Extreme Value distributions; wet/dry probabilities; and, lengths of wet and dry periods and their distributions. Annual means of maximum precipitation were generally captured well (Semenov 2008) as were annual maxima, multi-day extremes, and heat/cold characteristics for temperatures (Kyselý and Dubrovský 2005). The weather generators evaluated were, however, less effective in capturing 1-day extremes for temperature (Kyselý and Dubrovský 2005; Semenov 2008), while a mixed distribution was better at reproducing 1-day precipitation extremes than the gamma distribution (Hundscha et al. 2009).

Weather generators produce stochastic outputs. Thus, it is unlikely that any two simulation runs (realizations) will produce the same output. This is desirable as it provides a range of possible outputs, making it possible to better capture the range and variability of climate. It also provides opportunities for probabilistic assessments (Kalnay et al. 2006; Wiegel 2012). Several studies have employed the use of multiple realizations, including: 10 (Elliot and Arnold 2001); 30 (Caron et al. 2008); 100 (Hundscha et al. 2009; Min et al. 2011); 250 (Gitau et al. 2012); and, 1000 (Brisson et al. 2015), although the use of one realization is more common (e.g., Zhang and Garbrecht 2003; Farzanmanesh et al. 2012; Fan et al. 2013; Al-Mukhtar et al. 2014). Using too many realizations would be time consuming and computationally expensive, especially where generated data are used in subsequent applications.

In this paper, we present an evaluation of weather generator representation of climate extremes based on work done with the Long Ashton Research Station Weather Generator (LARS-WG; Racsco et al. 1991) on three stations in the Western Lake Erie Basin, U.S., and considering 50 realizations. Since daily weather data are needed to run hydrologic, water quality, and crop-growth models, this study was focused on the suitability of generated values on a daily time scale. We also discuss some important considerations with respect to the use of multiple realizations and with regard to translating the work to data-sparse and non-humid regions.

## 2 Methods

### 2.1 Study Site and Base Data Description

This study was focused on Fort Wayne, Norwalk, and Adrian stations in the Western Lake Erie Basin, part of the U.S. Great Lakes region. These stations, which were adopted from Mehan et al. (2017), were selected based on an analysis of eight stations within the basin, and so as to provide coverage across the basin. Long-term and consistent datasets were available for all three stations making them suitable for use with the analysis. Based on historical data ranging from as far back as July 1887 to November 2017 (MRCC 2017), 1-day maximum precipitation can range from 125.2 mm (4.9 in.) to 229.1 mm (9.0 in.), these having been recorded at Fort Wayne and Norwalk, respectively. One-day maximum temperatures can range from 20.0 °C (68.0 °F) in January to 42.2 °C (108.0 °F) in July while one-day minima can range from -31.1 °C (24.0 °F) in January to 6.7 °C (44.0 °F) in July based on July 1887 to June 2017 records for Fort Wayne and Adrian (MRCC 2017). Historical data (1887-current) on extreme temperatures were not available at Norwalk. For this study, observed data for the period 1966–2015 were used with 100%, 98.8%, and 97.6% availability at Fort Wayne, Norwalk, and Adrian, respectively.

### 2.2 Data Development

The data used for this study were adopted from Mehan et al. (2017). These data included fifty years of observed data for the period 1966–2015 and synthetic data comprising 50 realizations, each 50 years in length, generated from LARS-WG. The synthetic data were developed using different random number seeds to generate each realization, such that each set of 50-year outputs was expected to be different from the others making it possible to represent a range of variability in the generated values (Mehan et al. 2017). An analysis of LARS-WG effectiveness in reproducing essential characteristics of observed data at the three stations was conducted by Mehan et al. (2017). Selected results are provided in Table 1. Based on the analysis, LARS-WG captured the essential characteristics relatively well and was found especially suitable for hydrologic and water resources applications (Mehan et al. 2017).

### 2.3 Analysis of Extremes

A number of extreme metrics were evaluated in this study considering seasonal, annual, and multiday contexts, and also sequences. Seasonal metrics included: the 1-day maximum precipitation; 1-day maximum temperatures; and, 1-day minimum temperatures each of which was considered on a seasonal basis. The 1-day maximum precipitation provides important information for flood risk assessment (Hundecha et al. 2009). Extreme high or extreme low temperatures can lead to crop damage (Semenov 2008) if these occur during the growing season. For example, corn which is an economically important crop in the study area, is severely impacted when temperatures rise above 35 °C (95 °F) or drop below 10 °C (50 °F). For this study, the seasons were defined as: winter (Jan, Feb, Mar); spring (Apr, May, Jun); summer (Jul, Aug, Sep); and, Fall (Oct, Nov, Dec), consistent with definitions in other studies conducted within the basin. Annual metrics included the number of days in a year in which: precipitation exceeded 50.8 mm (2 in.); precipitation exceeded 101.6 mm (4 in.); maximum temperatures exceeded 32.2 °C (90 °F); and, minimum temperatures were below 0 °C (32 °F).

**Table 1** Summary of LARS-WG performance in representing essential characteristics of observed data for Adrian, Norwalk, and Fort Wayne stations. Adapted from Mehan et al. (2017)

|  | Adrian   |               | Norwalk  |               | Fort Wayne |               |
|--|----------|---------------|----------|---------------|------------|---------------|
|  | Observed | LARS-WG**     | Observed | LARS-WG**     | Observed   | LARS-WG**     |
| <b>Precipitation, mm</b>                   |          |               |          |               |            |               |
| Mean                                       | 2.5      | 2.4–2.7       | 2.6      | 2.5–2.8       | 2.6        | 2.5–2.7       |
| Standard deviation                         | 6.6      | 6.5–7.2       | 7.1      | 6.6–8.2       | 6.9        | 6.7–7.4       |
| No rainfall days (%)                       | 66.8     | 66.2–67.7     | 63.6     | 63.4–65.2     | 63.4       | 63.4–65.5     |
| Q99.5 <sup>‡</sup>                         | 41.1     | 40.5–45.2     | 40.1     | 39.3–44.5     | 42.2       | 41.08–46.8    |
| Q100 (Max)                                 | 120.0    | 93.0–120.8    | 229.1    | 124.0–228.4   | 111.8      | 86.9–111.5    |
| > Q99.5<br>(Days/50 years)*                | 92       | 105           | 92       | 103           | 91         | 104           |
| <b>Maximum temperature<sup>†</sup>, °C</b> |          |               |          |               |            |               |
| Mean                                       | 15.0     | 14.9–15.1     | 15.0     | 15.0–15.1     | 15.5       | 15.4–15.6     |
| Standard deviation                         | 11.6     | 11.1–11.3     | 11.5     | 11.0–11.2     | 11.8       | 11.2–11.4     |
| Q0 (Min)                                   | (20.0)   | (19.6)–(15.5) | (22.2)   | (21.7)–(16.7) | (23.9)     | (23.1)–(17.1) |
| Q99.5                                      | 34.4     | 33.8–34.4     | 34.4     | 33.6–34.1     | 35.0       | 33.6–34.3     |
| Q100 (Max)                                 | 40       | 36.9–37.6     | 39.4     | 36.7–37.5     | 41.1       | 37.4–38.8     |
| > Q99.5<br>(Days/50 years)*                | 137      | 74            | 128      | 61            | 64         | 42            |
| <b>Minimum temperature<sup>†</sup>, °C</b> |          |               |          |               |            |               |
| Mean                                       | 3.3      | 3.1–3.3       | 4.5      | 4.4–4.6       | 4.8        | 4.8–5.0       |
| Standard deviation                         | 10.0     | 9.5–9.7       | 10.1     | 9.7–9.8       | 10.3       | 9.8–10.0      |
| Q0   | (30.0)   | (29.9)–(26.9) | (29.4)   | (29.4)–(27.0) | (30.0)     | (30.1)–(27.7) |
| Q99.5                                      | 21.7     | 20.8–21.3     | 22.8     | 20.0–20.3     | 22.8       | 21.9–22.4     |
| Q100                                       | 24.4     | 23.8–24.2     | 26.1     | 24.4–25.3     | 25.6       | 24.5–25.6     |
| > Q99.5<br>(Days/50 years)*                | 73       | 53            | 64       | 70            | 3          | 53            |
| <b>Other</b>                               |          |               |          |               |            |               |
| Snow Days (#/Year)                         | 30       | 31            | 32       | 30            | 33         | 33            |
| Period of optimal growth<br>(Days/Year)    | 52       | 67            | 55       | 67            | 63         | 71            |

\*\* Range of values obtained from 50 different realizations <sup>‡</sup>Q: Quantile; <sup>†</sup> Unless otherwise specified with and asterisk, values in parentheses are negative values

The metrics related to precipitation indicate the extent to which heavy precipitation is experienced in an area. The amounts represent precipitation that is much greater than normal (Karl et al. 1996). Days with minimum temperatures below 0 °C indicate frost days, while 32.2 °C is associated with heat stress and heat injury (NIOSH 2016; Blevins 2011; Epstein and Moran 2006; Griffin undated). Continuous periods in which extreme events occur are of interest due to the sustained impacts on agricultural, environmental, and water resource systems. Continuous metrics considered in this study include: number of periods with 7-day continuous precipitation; number of dry sequences; and, number of wet sequences in the 50-year period considered. The latter two were derived from Mehan et al. (2017) as these had been computed using the same dataset used in this study. Visual comparisons were also employed to allow a more detailed examination of the data and associated distributions. This was done using Q-Q plots and plots of distribution of distributions (Mehan et al. 2017) with a focus on skewness. Q-Q plots are typically used to determine whether the data comes from a particular distribution (Dalgaard 2008; Hensel and Hirsch 2002). For this study, Q-Q plots were developed considering annual 1-day maximum precipitation and temperatures and annual 1-day minimum temperatures from the observed versus simulated datasets. This was done to assess whether the data came from the

same distribution with the expectation that quantiles coming from the same distribution would form a straight line. Thus, the Q-Q plots allowed an evaluation of where and to what extent there were deviations in the data consistent with Hundedcha et al. (2009).

### 3 Results and Discussion

#### 3.1 Performance on Extreme Metrics

Based on the analysis (Table 2), LARS-WG generally underestimated 1-day maximum precipitation across all stations based on the range and median values obtained. However, other extreme values associated with precipitations were captured relatively well, although there was a tendency towards overestimation for the number of days with greater than 50.8 mm (2 in.) precipitation. In the series of 50 years considered for this study, there were 29, 34, and 36 periods with continuous 7-day precipitation recorded from observed data for Adrian; Norwalk, and Fort Wayne, respectively. Values obtained from the generator were slightly lower based on median values (24, 28, 32, respectively). From the observed data, the highest value for maximum one-day precipitation was observed during summer except at Fort Wayne, where the maximum one-day precipitation was recorded in spring during the study period (1966–2015). This differed from the historical data (MRCC 2017) based on which all the highest values occurred in the summer. The generator was, however, able to capture the prevailing pattern in the data. The 1-day maximum precipitation values for the study period were also similar to historical values, except at Fort Wayne where the value was substantially lower (111.8 mm compared with 125.7 mm for the historical). Precipitation events larger than 101.6 mm (4 in.) were rare and these were generally well captured by LARS-WG. The generator captured the occurrence of wet and dry spells relatively well, as detailed in Mehan et al. (2017).

One-day maximum temperatures in the fall and winter were captured relatively well. However, the generator had a tendency to underestimate these values in the spring and summer. The number of days with maximum temperatures greater than 32.2 °C (90 °F) obtained based on simulated data were much lower than (about half of) those obtained from observed data, suggesting a tendency towards simulating fewer extremely hot days. The highest value of 1-day maximum temperature was observed in the spring at all stations during the study period (1966–2015). The generator tended to simulate about the same in both spring and summer. This was consistent with historical values of 1-day maximum temperatures where maximum values were recorded in June and July for Fort Wayne and July for Adrian. Seasonal 1-day minimum temperatures were simulated relatively well at all stations although actual winter values tended to coincide with the absolute minima of the simulated data, hence a tendency to simulate winter temperatures that were warmer than observed. However, the number of days with temperatures less than 0 °C (32 °F) were overestimated, hence overall more frost days being simulated by the generator.

#### 3.2 Visual Comparisons

For annual 1-day maximum precipitation, quantiles from the simulated data matched those from the observed data relatively well at Fort Wayne. Some deviations were observed at the lower end and towards the upper end at Adrian, although the quantiles at the upper end were relatively well matched. For the most part, the plot deviated from the ideal at Norwalk with

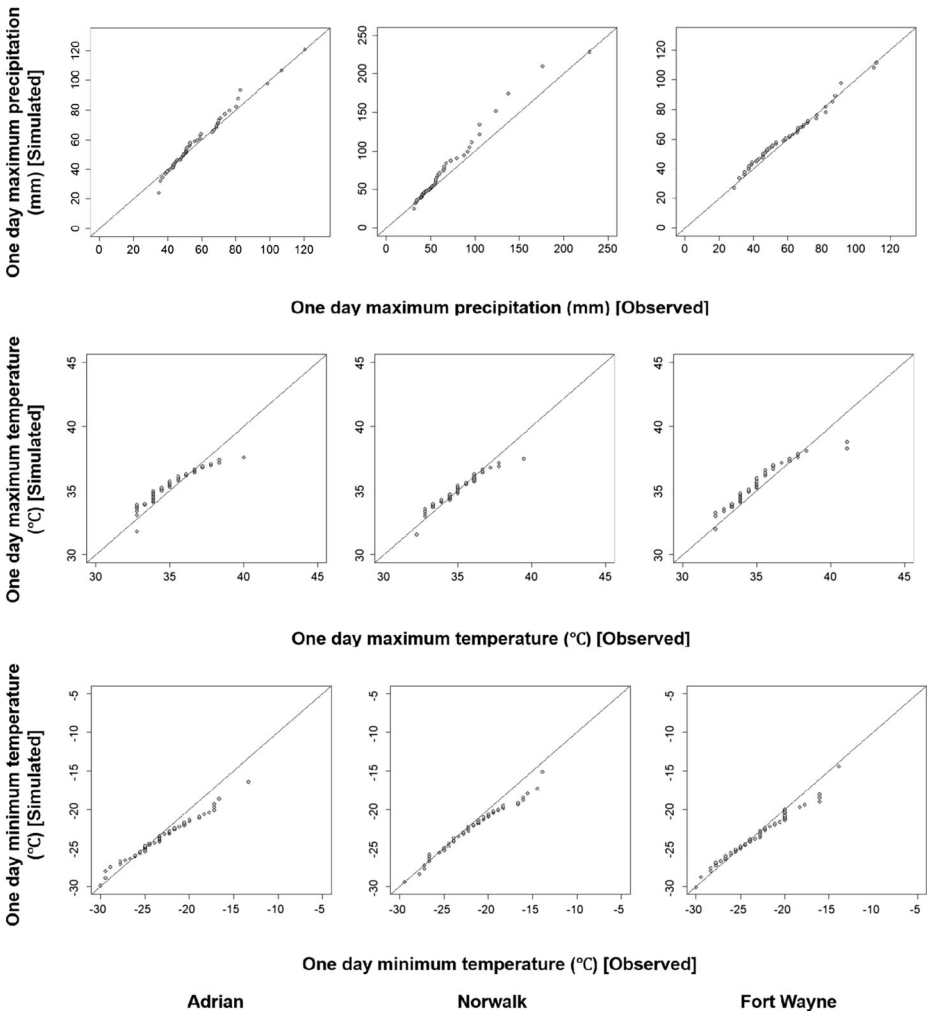
**Table 2** Evaluation of LARS-WG effectiveness in simulating different extreme variables associated with precipitation and temperature in comparison to values obtained from observed data

|  | Adrian   |                       | Norwalk  |                       | Fort Wayne |                       |
|--|----------|-----------------------|----------|-----------------------|------------|-----------------------|
|  | Observed | LARS-WG <sup>†</sup>  | Observed | LARS-WG <sup>†</sup>  | Observed   | LARS-WG <sup>†</sup>  |
|  |          |                       |          |                       |            |                       |
| <b>1-day maximum precipitation (mm)</b>      |          |                       |          |                       |            |                       |
| Winter                                       | 59.7     | (48.4–59.4) 57.2      | 55.9     | (45.1–55.5) 52.5      | 77.0       | (44.7–77.0) 67.9      |
| Spring                                       | 98.6     | (78.2–97.9) 97.9      | 137.7    | (74.1–178.4) 134.3    | 111.8      | (83.0–111.5) 110.8    |
| Summer                                       | 120.4    | (84.0–120.8) 106.7    | 229.1    | (110.7–228.4) 220.5   | 91.2       | (72.1–90.3) 86.5      |
| Fall   | 69.6     | (52.6–69.3) 66.5      | 79.8     | (62.7–78.7) 71.7      | 68.3       | (57.7–75.2) 67.8      |
| Precipitation >50.8 mm (Days/50 years)       | 34       | (34–60) 47            | 40       | (35–59) 48            | 44         | (37–76) 54            |
| Precipitation >101.6 mm (Days/50 years)      | 2        | (0–4) 1               | 6        | (2–15) 9              | 2          | (0–6) 1               |
| 7-day continuous precipitation (Count)       | 29       | (17–38) 24            | 34       | (15–39) 28            | 36         | (19–43) 32            |
| Dry sequence count for 50 years <sup>‡</sup> | 39       | (31–58) 45            | 14       | (11–28) 19            | 18         | (11–39) 23            |
| Wet sequence count for 50 years <sup>‡</sup> | 191      | (148–216) 177         | 230      | (203–271) 232         | 213        | (198–251) 221         |
| <b>1-day max temperature (°C)</b>            |          |                       |          |                       |            |                       |
| Winter                                       | 28.3     | (26.8–29.5) 28.5      | 28.9     | (26.6–29.7) 28.3      | 30.6       | (25.7–30.4) 29.3      |
| Spring                                       | 40.0     | (36.4–37.4) 37.2      | 39.4     | (34.9–36.8) 36.4      | 41.1       | (35.6–38.8) 38.2      |
| Summer                                       | 38.3     | (36.9–37.6) 37.5      | 37.8     | (36.7–37.5) 37.4      | 39.4       | (37.4–38.5) 38.2      |
| Fall   | 32.8     | (29.3–34.2) 32.8      | 32.2     | (29.9–33.9) 32.5      | 31.7       | (30.3–33.1) 31.8      |
| <b>1-day min temperature (°C)</b>            |          |                       |          |                       |            |                       |
| Winter                                       | -30.0    | (-29.9 – -26.9) -29.4 | -29.4    | (-29.4 – -26.2) -28.8 | -30.0      | (-30.1 – -27.3) -29.2 |
| Spring                                       | -15.0    | (-16.1 – -11.1) -14.3 | -13.9    | (-14.6 – -8.4) -12.6  | -13.9      | (-14.7 – -8.3) -11.7  |
| Summer                                       | -2.8     | (-4.9 – -1.5) -3.5    | -1.1     | (-3.3–0.5) -1.6       | -1.7       | (-2.7–2.0) -1.2       |
| Fall   | -25.6    | (-27.6 – -23.9) -26.6 | -27.2    | (-28.0 – -22.6) -26.4 | -27.8      | (-27.9 – -24.4) -27.0 |
| Max temperature > 32.2C (Days/year)          | 12       | (6–8) 7               | 13       | (6–8) 7               | 15         | (6–8) 7               |
| Min temperature < 0 C (Days/year)            | 136      | (142–146) 144         | 123      | (128–131) 130         | 119        | (122–126) 124         |

<sup>†</sup> Values in parentheses are the range of values for the parameter computed from the different realizations. Values outside the parenthesis are the respective median values

<sup>‡</sup> Adapted from Mehan et al. (2017)

larger deviations seen on the upper tail. This was consistent with findings in Mehan et al. (2017) from which the generator tended to overestimate the 99.5 percentile precipitation at this station and the number of days with greater than the 99.5 percentile precipitation, although this latter value was overestimated at all stations (Table 1). The distributions of annual 1-day maximum temperatures were not captured well at any of the stations (Fig. 1). In general, there was a tendency to overestimate values in the lower quantiles and underestimate those in the upper quantiles. Annual 1-day minimum temperatures were generally well captured at the lower end although with a tendency toward underestimating the lower colder temperatures and overestimation at the higher quantiles. Overall, simulated 1-day minimum temperatures were warmer at the lower extremes and cooler at the higher ends relative to observed data. This

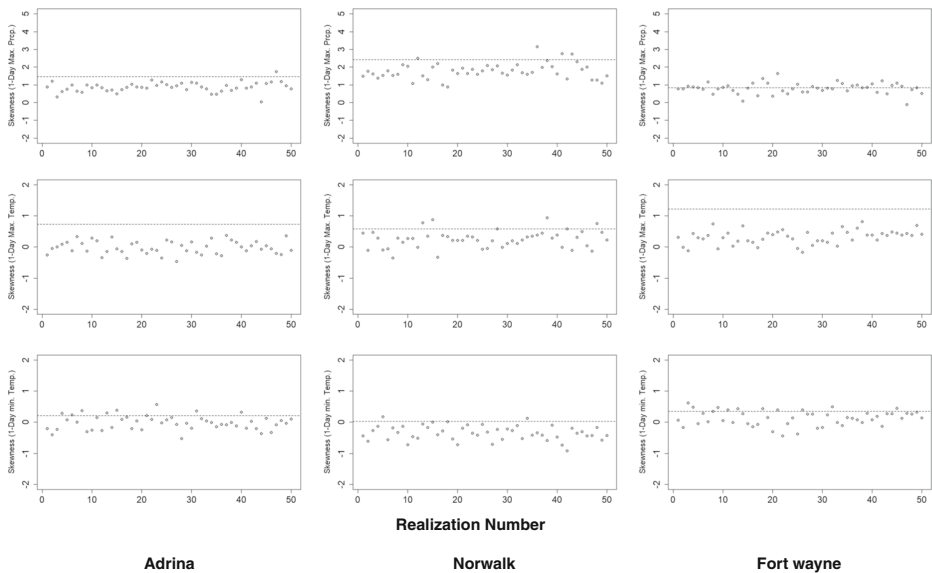


**Fig. 1** Q-Q plots for annual 1-day maximum precipitation, annual 1-day maximum temperatures, and annual 1-day minimum temperatures based on values simulated by LARS-WG for 50 different realization each 50 years in length (maximum value in any one year was used regardless of the realization in which it occurred) for the weather stations at Adrian, MI; Norwalk, OH; and Fort Wayne, IN

pattern was consistent across the stations based on the Q-Q plots (Fig. 1). When considered across realizations, the skewness of annual 1-day maximum precipitation was better captured at Fort Wayne but tended to be underestimated at both Norwalk and Adrian (Fig. 2). The skewness for both annual 1-day maximum temperatures and annual 1-day minimum temperatures was generally underestimated by LARS-WG although simulations were better for minimum temperatures for which values were only slightly underestimated (Fig. 2). The implications of underestimated skewness are more evident with precipitation and translate to a tendency to overestimate mean and median values. This effect is less evident with annual 1-day maximum and minimum temperatures due to the nature of their distributions.

### 3.3 Suitability of Number of Realizations

Fifty realizations were used in this study as this was the number of realizations used by Mehan et al. (2017) from which the data were taken. Based on a comparison of 1, 10, 25, 50, and 100 realizations, Guo et al. (2017) found no added advantage in going beyond 25 realizations given that generator effectiveness in capturing the essential characteristics of observed data did not improve appreciably when more than 25 realizations were used. Guo et al. (2017) also found that 10 realizations were not sufficient to capture essential characteristics of observed data. That said, increasing the number of realizations was not helpful for a particular characteristic if the generator was unable to capture the characteristic. An assessment of data for the Fort Wayne and Norwalk stations using 25 realizations taken at random (Table 3) showed that: annual 1-day maximum precipitation was estimated generally well except at the absolute maximum where it was underestimated at both stations consistent with results from 50 realizations. The 75th percentile value was, however, overestimated at Norwalk; annual 1-



**Fig. 2** Variability of skewness for annual 1-day maximum precipitation, annual 1-day maximum temperatures, and annual 1-day minimum temperatures based on values simulated from the weather generator LARS-WG for 50 different realizations each 50 years in length for the weather stations at Adrian, MI; Norwalk, OH; and Fort Wayne, IN compared to the values calculated from observed data (dashed lines)



**Table 3** Distribution of simulated data for annual 1-day maximum precipitation, annual 1-day maximum temperature, and annual 1-day minimum temperatures as obtained from 25 realizations in comparison to values from observed data at Fort Wayne and Norwalk stations

|                      | 1-day maximum precipitation |                      | 1-day maximum temperature |                      | 1-day minimum temperature |                        |
|----------------------|-----------------------------|----------------------|---------------------------|----------------------|---------------------------|------------------------|
|                      | Observed                    | LARS-WG <sup>†</sup> | Observed                  | LARS-WG <sup>†</sup> | Observed                  | LARS-WG <sup>†</sup>   |
| <b>Fort Wayne</b>    |                             |                      |                           |                      |                           |                        |
| Minimum              | 28.7                        | (27.0–38.5) 32.1     | 32.2                      | (31.8–33.5) 32.8     | -30                       | (-30.0 – -27.6) -29.0  |
| Lower quartile       | 43                          | (43.6–53.2) 47.9     | 33.9                      | (33.9–34.9) 34.4     | -26.1                     | (-26.6 – -24.8) -25.8  |
| Median               | 53.1                        | (52.2–65.4) 57.6     | 34.7                      | (34.6–35.8) 35.2     | -23.1                     | (-24.6 – -21.8) -23.8  |
| Upper quartile       | 69.2                        | (61.5–77.3) 68.9     | 35.7                      | (35.6–37.5) 36.4     | -20                       | (-22.4 – -19.9) -21.5  |
| Maximum              | 111.8                       | (80.4–111.0) 108.4   | 41.1                      | (37.7–38.7) 38.2     | -13.9                     | (-18.8 – -14.9) -17.4  |
| Inter-quartile range | 26.2                        | (14.6–29.4) 20.4     | 1.8                       | (1.5–2.8) 2.1        | 6.1                       | (3.2–5.8) 4.5          |
| <b>Norwalk</b>       |                             |                      |                           |                      |                           |                        |
| Minimum              | 31                          | (25.8–34.7) 29.6     | 32.2                      | (31.8–33.6) 32.8     | -29.4                     | (-29.4 – -27.8) -28.6  |
| Lower quartile       | 42.5                        | (41.3–49.4) 45.4     | 33.9                      | (33.8–34.5) 34.2     | -24.4                     | (-25.5 – -23.4) -24.5  |
| Median               | 54.1                        | (50.6–72.3) 55.4     | 35                        | (34.6–35.6) 34.8     | -21.4                     | (-23.0 – -21.2) -22.0  |
| Upper quartile       | 68.6                        | (68.9–99.4) 85.4     | 36.1                      | (35.3–36.4) 35.7     | -19.3                     | (-21.05 – -19.5) -20.1 |
| Maximum              | 229.1                       | (154.6–228.4) 223.4  | 39.4                      | (36.6–37.5) 37.2     | -13.9                     | (-18.3 – -15.1) -17.0  |
| Inter-quartile range | 26.2                        | (25.9–56.0) 42.0     | 2.2                       | (0.9–2.3) 1.6        | 5.1                       | (3.3–5.6) 4.2          |

<sup>†</sup> Values in parentheses are the range of values for the parameter computed from different realizations. Values outside the parenthesis are the respective median values

day maximum temperatures were also well captured except for the maximum value which was underestimated; and, annual 1-day minimum temperatures were underestimated (warmer) at the lower end and overestimated (colder) at the higher end, also consistent with results from 50 realizations. The interquartile range for annual 1-day minimum temperatures based on simulated values was smaller than that from observed data at the Fort Wayne station. The number of realizations has implications on computational needs especially if the generated data are used in subsequent applications—such as in water resources modeling—thus the need to determine a suitable number of realizations. Oftentimes, a long time series is generated so as to reduce sampling error associated with the stochastic nature of the weather generator (e.g., Wilks 2002; Zhang and Garbrecht 2003; Chen et al. 2012; Eames et al. 2012; Chen and Brissette 2014), sometimes dividing these into sets of equal lengths (pseudo-realizations), e.g., Kou et al. (2007). In the latter case, the entire dataset is essentially a reflection of the single realization and errors may simply be propagated across the dataset as discussed in Gitau et al. (2017). In general, generating a time series that is substantially longer than the base data could result in biases as the sampling of the distribution would be insufficient (Mithen and Black 2011).

### 3.4 General Discussion

In this study, an evaluation of weather generator representation of climate extremes was conducted based on simulated values from LARS-WG and observed data from three stations in the Western Lake Erie Basin, and considering 50 realizations. Generally, LARS-WG captured the number of days with precipitation greater than 50.8 mm (2 in.) and 101.6 mm (4 in.), 7-day wet sequences, and wet and dry sequences relatively well. The distribution of 1-day maximum precipitation was generally captured well with the exception of the upper tail at the Norwalk station. One-day maximum temperatures were also captured relatively well in the fall and winter, although the generator tended to underestimate spring and summer values across all stations. The generator greatly underestimated the number of days per year with maximum temperatures greater than 32.2 °C (90 °F) and overestimated the number of days with temperatures less than 0 °C (32 °F).

When evaluating generator effectiveness, the choice of characteristics to evaluate largely depends on the intended use for the synthetic output. However, 1-day extreme variables (maximum precipitation, maximum temperatures, minimum temperatures), tendency to exceed selected extreme thresholds (number of days in which the threshold is exceeded), occurrence and severity of multi-day events (e.g., 7-day precipitation, wet sequences, dry sequences), and tendency to exceed certain percentile thresholds (e.g., 90<sup>th</sup> percentile precipitation) are commonly evaluated. These can be categorized as absolute, threshold, duration, and percentile based metrics, respectively based on Sillmann et al. (2013). The 90<sup>th</sup> and 10<sup>th</sup> percentiles, for example, have been used to represent extreme hot days and extreme cold days (Vincent et al. 2005), while the 95<sup>th</sup> percentile has been used to represent extreme precipitation (Sillmann and Roeckner 2008). Metrics based on percentiles reflect anomalies relative to the local climate and are, therefore, location-specific (Klein Tank and Können 2003). They, however, have an advantage in that they sample the same part of a distribution regardless of the location, making them especially suitable for comparative analysis across locations (Klein Tank et al. 2009; Zhang et al. 2011). More universal measures include: a 50.8 mm (2 in.) threshold for precipitation; a 30 °C (86 °F) threshold for maximum temperatures, used in relation to heat waves; and, a 32.2 °C (90 °F) threshold for maximum temperatures, which is commonly associated with heat stress as previously discussed. Semenov (2008) defined a heat wave as

two or more days with temperatures greater than 30 °C. A threshold of 0 °C (32 °F) is used for assessment of frost.

Generally, a number of realizations are needed in order to capture the range of variability in climate. The use of multiple realizations is, however, only useful with respect to characteristics that the generator is able to simulate relative to observed data (Guo et al. 2017; Gitau et al. 2017). For example, if a generator has the tendency to underestimate a particular extreme variable, then increasing the number of realizations will result in multiple datasets in which the variable is underestimated. The use of multiple weather realizations in subsequent applications, such as water resources modelling studies, provides opportunities to capture a range of possible outcomes particularly as related to future climate. Such evaluations could, however, become computationally expensive and produce large datasets beyond what can be handled effectively using a standard workstation. An alternative might be to select a best set from the realizations based, for example, on verification methods for paired data such as those described in Jolliffe and Stephenson (2012). However, the time series produced by weather generators and the base input data are not necessarily paired, although this might be implied if the length of the generated data is equal to that of the base data. However, if the length of the generated data (Y) is shorter than that of the base input data (X), generated data could represent any of the many different potential combinations of time series each of length Y, as detailed in Gitau et al. (2017). This might present a challenge where data are generated for use with modelling applications, and especially when model outputs are used in parameter optimization. A challenge may also present where data are extrapolated beyond the length of the base input data as is sometimes done where the interest is in evaluating impacts of factors other than climate or in assessing future scenarios. An important aspect relates to how generator effectiveness would translate with respect to representing future climate. Some key questions to be addressed include how the statistics might change as climate change impacts are realized and considerations that would be needed to allow the generator to be used in a meaningful way to quantify the impacts. Such analysis is, however, beyond the scope of this work.

The stations considered in this study had very rich datasets with long records dating as far back as 1887 and consistent data availability through the study period. This is not often the case. Furthermore, the study site is located in a humid region where rainfall occurs throughout the year. It is not clear how the generator would perform with respect to both essential characteristics and extreme variables in areas with much smaller and less consistent observed datasets and/or in non-humid regions. Possibly the biggest challenge with respect to data-sparse regions is in the ability to derive a suitable set of input parameters. A 20-year period is recommended for LARS-WG (Semenov and Barrow 1997). However, such a record might be difficult to find in some areas and the analysis of extremes might be confounded by the limited knowledge of such events afforded by an inconsistent dataset. In arid and semi-arid regions, precipitation events occur infrequently but are often very severe. Extreme temperatures might also be much higher than what is implied by global thresholds. The occurrence of such events might, thus, be characterised differently than they would in humid regions. Thus, more localized metrics such as percentiles might be more applicable.

## 4 Conclusions

Given the increasing use of weather generators in environmental, water resources, and agricultural applications, and their potential for use in climate impact and climate change

studies, it is important that the generators be evaluated for their suitability for the area in which they would be applied. This is particularly so with respect to their ability to capture extreme events. In this study, LARS-WG was evaluated using data from three stations in the Western Lake Erie Basin. Generally, the generator performed well particularly with respect to capturing wet and dry sequences, winter and fall 1-day maximum temperatures and 1-day minimum temperatures in general. The 50 realizations used in this study were sufficient for the analysis although similar results might have been obtained with 25 realizations based on recent literature. Further work is needed to determine the impact of number of realizations on generator effectiveness in simulating extremes. Further work is also needed to evaluate the generator for use in data-sparse and/or non-humid regions.

**Acknowledgments** A previous shorter version of the paper has been presented in the 10<sup>th</sup> World Congress of the EWRA “Panta Rhei” Athens, Greece, 5-9 July, 2017. This work was made possible in part by funding by USDA National Institute of Food and Agriculture (Project No. IND010639R).

## References

- Acharya N, Frei A, Chen J, DeCristofaro L, Owens EM (2017) Evaluating stochastic precipitation generators for climate change impact studies of New York City’s primary water supply. *J Hydrometeorol* 18(3):879–896. <https://doi.org/10.1175/JHM-D-16-0169.1>
- Al-Mukhtar M, Dunger V, Merkel B (2014) Evaluation of the climate generator model CLIGEN for rainfall data simulation in Bautzen catchment area, Germany. *Hydrol Res* 45(4–5):615–630
- Blevins DW (2011) Water-quality requirements, tolerances, and preferences of pallid sturgeon (*Scaphirhynchus albus*) in the lower Missouri River: U.S. Geological Survey Scientific Investigations Report. Reston, Virginia, pp 2011–5186
- Brisson E, Demuzere M, Willems P, van Lipzig NPM (2015) Assessment of natural climate variability using a weather generator. *Clim Dyn* 44(1–2):495–508
- Caron A, Leconte R, Brissette F (2008) An improved stochastic weather generator for hydrological impact studies. *Can Water Resour J* 33:233–256
- Chen J, Brissette FP (2014) Comparison of five stochastic weather generators in simulating daily precipitation and temperature for the loess plateau of China. *Int J Climatol* 34(10):3089–3105
- Chen J, Brissette FP, Leconte R (2012) WeaGETS– a Matlab-based daily scale weather generator for generating precipitation and temperature. *Procedia Environ Sci* 13:2222–2235
- Dalgaard P (2008) *Introductory statistics with R*. Springer Science+Business Media, LLC, New York
- Eames M, Kershaw T, Coley D (2012) A comparison of future weather created from morphed observed weather and created by a weather generator. *Build Environ* 56:252–264
- Elliot W, Arnold C (2001) Validation of the weather generator CLIGEN with precipitation data from Uganda. *T ASAE* 44:53–58
- Epstein Y, Moran DS (2006) Thermal comfort and the heat stress indices. *Ind Health* 44(3):388–398
- Fan JC, Yang CH, Liu CH, Huang HY (2013) Assessment and validation of CLIGEN-simulated rainfall data for northern Taiwan. *Paddy Water Environ* 11(1–4):161–173
- Farzanmanesh R, Abdullah AM, Shakiba AR, Amanollahi J (2012) Impact assessment of climate change in Iran using LARS-WG model. *Pertanika J Sci Technol* 20(2):299–311
- Fatichi S, Ivanov VY, Caporali E (2011) Simulation of future climate scenarios with a weather generator. *Adv Water Resour* 34(4):448–467
- Fodor N, Dobi I, Mika J, Szeidl L (2013) Applications of the MVWG multivariable stochastic weather generator. *Sci World J* 2013:1–6. <https://doi.org/10.1155/2013/571367>
- Gitau MW, Chiang LC, Sayeed M, Chaubey I (2012) Watershed modeling using large-scale distributed computing in condoor and the soil and water assessment tool model. *Simulation* 88(3):365–380
- Gitau MW, Mehan S, Guo T (2017) Weather generator utilization in climate impact studies: implications for water resources modeling. *European Water* 59:69–75
- Griffin D (undated) Feedlot heat stress checklist. University of Nebraska, Great Plains Veterinary Educational Center. Available at: <http://gpvec.unl.edu/heatdrought/flheatsr.pdf>. Last accessed November 2017

- Guo T, Mehan S, Gitau M, Wang Q, Kuczek T, Flanagan D (2017) Impact of number of realizations on the suitability of simulated weather data for hydrologic and environmental applications. *Stoch Env Res Risk A*. <https://doi.org/10.1007/s00477-017-1498-5>
- Hensel DR, Hirsch RM (2002) Chapter A3: statistical methods in water resources. In: *Techniques of water-resources investigations of the United States Geological Survey, Book 4 Hydrologic Analysis and Interpretation*. U.S. Geological Survey. Available at: <http://water.usgs.gov/pubs/twri/twri4a3/>. Last accessed November, 2017
- Hundecha Y, Pahlow M, Schumann A (2009) Modeling of daily precipitation at multiple locations using a mixture of distributions to characterize the extremes. *Water Resour Res* 45(12). <https://doi.org/10.1029/2008WR007453>
- Jolliffe IT, Stephenson DB (eds) (2012) *Forecast verification: a practitioner's guide in atmospheric science*, 2<sup>nd</sup> edn. Wiley, UK
- Kalnay E, Hunt B, Ott E, Szunyogh I (2006) Ensemble forecasting and data assimilation: two problems with the same solution? In: Palmer T, Hagedorn R (eds) *Predictability of weather and climate*. Cambridge University Press, Cambridge, p 2006
- Karl TR, Knight RW, Easterling DR, Quayle RG (1996) Indices of climate change for the United States. *B Am Meteorol Soc* 77(2):279–292
- Klein Tank AMG, Können GP (2003) Trends in indices of daily temperature and precipitation extremes in Europe, 1946–99. *J Clim* 16(22):3665–3680
- Klein Tank AMG, Zwiers FW, Zhang X (2009): Guidelines on analysis of extremes in a changing climate in support of informed decisions for adaptation. WCDMP-72, WMO/TD-1500, Available at [www.wmo.int/datastat/documents/WCDMP\\_72\\_TD\\_1500\\_en\\_1\\_1.pdf](http://www.wmo.int/datastat/documents/WCDMP_72_TD_1500_en_1_1.pdf). Last accessed February, 2018
- Kou X, Ge J, Wang Y, Zhang C (2007) Validation of the weather generator CLIGEN with daily precipitation data from the loess plateau, China. *J Hydrol* 347(3):347–357
- Kysely J, Dubrovský M (2005) Simulation of extreme temperature events by a stochastic weather generator: effects of interdiurnal and interannual variability reproduction. *Int J Climatol* 25(2):251–269
- Lemann T, Zeleke G, Amsler C, Giovanoli L, Suter H, Roth V (2016) Modelling the effect of soil and water conservation on discharge and sediment yield in the upper Blue Nile basin, Ethiopia. *Appl Geogr* 73:89–101
- Mehan S, Guo T, Gitau M, Flanagan D (2017) Comparative study of different stochastic weather generators for long-term climate data simulation. *Climate* 5. <https://doi.org/10.3390/cli5020026>
- Midwestern Regional Climate Center, MRCC (2017). *Midwest Climate: Climate Summaries*. Available at: [http://mrcc.isws.illinois.edu/mw\\_climate/climateSummaries/climSumm.jsp](http://mrcc.isws.illinois.edu/mw_climate/climateSummaries/climSumm.jsp). 2014–2017. Last accessed November, 2017
- Min YM, Kryjov VN, An KH, Hameed SN, Sohn SJ, Lee WJ, Oh JH (2011) Evaluation of the weather generator CLIGEN with daily precipitation characteristics in Korea. *Asia-Pacific J Atmos Sci* 47(3):255–263
- Mithen S, Black E (2011) *Water, life and civilisation: climate, environment and Society in the Jordan Valley, international hydrology series*. Cambridge University Press, Cambridge
- NIOSH (2016) NIOSH criteria for a recommended standard: occupational exposure to heat and hot environments. By Jacklitsch B, Williams WJ, Musolin K, Coca A, Kim J-H, Turner N. U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health, DHHS (NIOSH) Publication 2016–106, Cincinnati, Ohio
- Racsko P, Szeidl L, Semenov M (1991) A serial approach to local stochastic weather models. *Ecol Model* 57(1–2):27–41
- Semenov MA (2008) Simulation of extreme weather events by a stochastic weather generator. *Clim Res* 35(3):203–212
- Semenov MA, Barrow EM (1997) Use of a stochastic weather generator in the development of climate change scenarios. *Clim Chang* 35(4):397–414
- Sillmann J, Roeckner E (2008) Indices for extreme events in projections of anthropogenic climate change. *Clim Chang* 86(1–2):83–104
- Sillmann J, Kharin VV, Zhang X, Zwiers FW, Bronaugh D (2013) Climate extremes indices in the CMIP5 multi-model ensemble: part 1. Model evaluation in the present climate. *J Geophys Res-Atmos* 118:1716–1733. <https://doi.org/10.1002/jgrd.50203>
- Vincent LA, Peterson TC, Barros VR, Marino MB, Rusticucci M, Carrasco G et al (2005) Observed trends in indices of daily temperature extremes in South America 1960–2000. *J Clim* 18(23):5011–5023
- Wiegel AP (2012) Chapter 8: ensemble forecasts. In: Jolliffe IT, Stephenson DB (eds) *2012 forecast verification: a practitioner's guide in atmospheric science*, 2<sup>nd</sup> edn. Wiley, UK
- Wilks D (2002) Realizations of daily weather in forecast seasonal climate. *J Hydrometeorol* 3:195–207
- Yu B (2003) An assessment of uncalibrated CLIGEN in Australia. *Agric For Meteorol* 119(3):131–148
- Zhang X, Garbrecht JD (2003) Evaluation of CLIGEN precipitation parameters and their implication on WEPP runoff and erosion prediction. *T ASAE* 46(2):311–320
- Zhang X, Alexander L, Hegerl GC, Jones P, Klein Tank A, Peterson TC, Trewin B, Zwiers FW (2011) Indices for monitoring changes in extremes based on daily temperature and precipitation data. *Wiley Interdiscip Rev Clim Chang* 2:851–870. <https://doi.org/10.1002/wcc.147>