

## Research Article

# Modeling Rainfall Variability over Urban Areas: A Case Study for Kuwait

**Jaber Almedeij**

*Civil Engineering Department, Kuwait University, P.O. Box 5969, Safat 13060, Kuwait*

Correspondence should be addressed to Jaber Almedeij, j.almedeij@ku.edu.kw

Received 29 December 2011; Accepted 22 January 2012

Academic Editors: J. Sloan and J. Sun

Copyright © 2012 Jaber Almedeij. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This study examines the spatial and temporal variability of monthly total rainfall data obtained from weather stations located in the urban areas of Kuwait. The rainfall data are analyzed by considering statistics on a seasonal basis and by means of periodogram technique to reveal the periods responsible for the variable pattern. The results demonstrate similarity implying that a point estimate of rainfall data can be considered spatially representative over the urban areas of Kuwait. A sinusoidal model triggering the influence of the detected periods is developed accordingly for the time duration from January 1965 to December 2009. The model is capable of describing the rainfall data with some discrepancies between the actual and calculated values resulting from hidden periods that have not been taken into account. This finding suggests that the ability to construct a more reliable model would require a wider range of historical data to detect the other periods affecting the rainfall pattern.

## 1. Introduction

The development of a model describing the rainfall pattern over a catchment has been a prime focus of hydrological research for many decades. Among the variety of applications, rainfall models have been used to develop storm hyetographs so that to simulate streamflow records [1], assess soil-moisture for crop water requirements [2], estimate water balance for reservoirs and groundwater aquifers [3], and analyze flood frequency to avoid the risk of road ponding especially at stormwater drainage inlets located at in-sag points [4, 5], as road ponding is considered a major contributor to traffic accidents and delays as well as nuisance and possible hazard to pedestrians [6].

For Kuwait, the development of a rainfall model is relevant to establish an efficient rainwater harvesting system that has a potential to provide an additional source of water for agriculture and other applications [7], as the country facing a water shortage problem with only one conventional resource of fresh groundwater located in the north with limited quantities. Kuwait, which is about 18,000 km<sup>2</sup>, is a desert country characterized by long, hot, and dry summer, and short winter. Temperature during summer (winter) reaches an average of 44°C (15°C) during the day time, with

the lowest average falling to 23°C (3°C) during the night. The total amount of rainfall through the year varies between 30 and 250 mm, most of which falls between November and April due to western pressure depressions.

Owing to the high soil's infiltration properties of Kuwait, the collection of surface water runoff through a rainwater harvesting system would be much more efficient in the urban areas than in the desert [8]. The urban catchments of Kuwait, with a total area of about 600 km<sup>2</sup>, have a storm sewer system of separate type where the collected water is currently discharged at the outfalls into the sea. This existing drainage system could be used directly for harvesting, although the collected rainfall would need to be treated more thoroughly before it is used. Studying the economic feasibility of integrating a complete harvesting system for the urban catchments may be accomplished by the estimation of water input using a rainfall model.

A comprehensive analysis of rainfall variability of Kuwait was presented by Marcella and Eltahir [9] at both seasonal and interannual time scales. They concluded that a rainfall model can be developed for the country if the periodic pattern at both seasonal and interannual time scales has been understood. Despite the local weather stations available,

the data obtained are not of sufficient quantity, both spatially and temporally, to carry out a comprehensive analysis of the rainfall pattern. The only historical data consistent is that produced from a weather station located in Kuwait International Airport. Marcella and Eltahir also added rainfall datasets from two other sources, from the Climate Research Unit and the Global Precipitation Climatology Project. Owing to the different techniques adopted to produce the datasets, discrepancies among the rainfall measurements have been noticed. However, the records of Kuwait International Airport consistently fall in between the two other datasets and may accordingly be considered here in an objective manner to provide a quantitative analysis.

The aim here is to model the rainfall pattern over the urban areas of Kuwait. Data available from local weather stations will be employed to examine the spatial variability of rainfall within the urban areas. Then, the temporal rainfall variability will be studied using a more comprehensive data obtained from Kuwait International Airport. A possible model based on considering the spatial and temporal variability of rainfall will then be developed and discussed.

## 2. Rainfall Spatial Variability

The spatial variability of rainfall can be examined over the urban catchments of Kuwait by employing monthly total rainfall data collected from the following weather stations: Jahra, Shwaikh, Salmiyah, Omairia, Kuwait International Airport, and Ahmadi (Figure 1). These stations are located in the urban areas within latitudes 29°20' N and 29°03' N, and longitudes 47°37' E and 48°10' E. As it was mentioned earlier, except that of Kuwait International Airport, the data collected from the weather stations are not of substantial continuity coverage. The only consistent data available for these stations are that within time duration from January 1994 to December 2005.

The average values of monthly total rainfall data can be measured by considering statistics on a seasonal basis using the expression

$$\bar{P} = \frac{1}{N} \sum_{i=1}^N P_{i,\tau}, \quad \tau = 1, \dots, 12, \quad (1)$$

where  $P$  is monthly total precipitation, defined as the total for month  $\tau$  in the given year  $i$ ;  $\bar{P}$  is seasonal mean precipitation; and  $N$  is total number of years. The results are presented graphically in Figure 2. As can be seen, the data possess nearly equivalent seasonal means with small differences of  $\pm 3$  mm. This finding is in agreement with the conclusion drawn by others [9] who found that the average values of rainfall data collected from these stations are not that different since they are sufficiently close in distance, while spatial distribution is evident within a larger scale.

The pattern of the monthly total rainfall can be compared for the different weather stations by plotting the periodogram, which is a Fourier transform of the autocovariance function representing an unsmoothed spectral plot for examining the cyclic structure in the frequency domain [10]. This technique is used to reduce the effect of

the measurement noise and thus detect which frequencies within the range of time are most responsible for the data pattern. Typically, a large peak value shown in a periodogram corresponds to a period that is strongly represented in the time series. For example, a typical periodogram for monthly averaged temperature data can show a period of 12 months implying that 6 months of the year possess considerably lower temperatures than the other 6 months.

Figure 3 provides the periodograms of the rainfall data obtained from the weather stations. As can be seen, not only the seasonal means of the rainfall data are similar, but also the patterns. This is evident in the periodograms for the time duration between 6 and 24 months. This range of time has been considered here for the comparison since the minimum period of 6 months represents the distinct four seasons of the year, that is, spring, summer, autumn, and winter, by which smaller periods become difficult to justify using monthly data; while periods larger than about 24 months can be resolved more conveniently if a longer time series became available. Here, three periods are detected of 6, 12, and 18 months. The 12-month period is more pronounced than the others, with the power value ( $y$  axis) being about 10 times larger. Apparently, the 6- and 12-month periods are related to the seasonal variation of rainfall during the year typically observed in climatological data. The third period of 18 months indicates a variation in rainfall pattern of a longer time. The reason for the existence of this period in the data is unclear and may need further investigation to establish any relation.

## 3. Rainfall Temporal Variability

Owing to the similar seasonal mean and pattern, rainfall data collected for a point estimate can be considered spatially representative over the urban areas of Kuwait. The monthly total data of Kuwait International Airport can thus be considered representative and employed to examine the temporal variability of rainfall in a larger time scale. A full range of data is available with time duration from January 1965 to December 2009. It should be mentioned that some measurements have not been taken in this data from August 1990 to June 1991, corresponding to the Iraqi invasion of Kuwait. To maintain data continuity in terms of time, this lack of information has been handled here by averaging the data by considering the seasonal mean resulting from adding the value of the same month but for the year before and after and then dividing by two.

The monthly total data of Kuwait International Airport for the time from January 1965 to December 2009 are shown in Figure 4(a). Here, a distinct rainfall pattern can be noticed dividing the data into three time intervals of 180 months. The rainfall incidents for the first and third intervals are larger than that for the second. This is evident, given that the corresponding total rainfall amounts through the intervals are 1870, 1420, and 2160 mm, and the seasonal means estimated from (1) are 124.47, 94.93, and 144.25 mm. This pattern suggests the presence of a large period of 360 months that may not appear in the periodogram as its size is

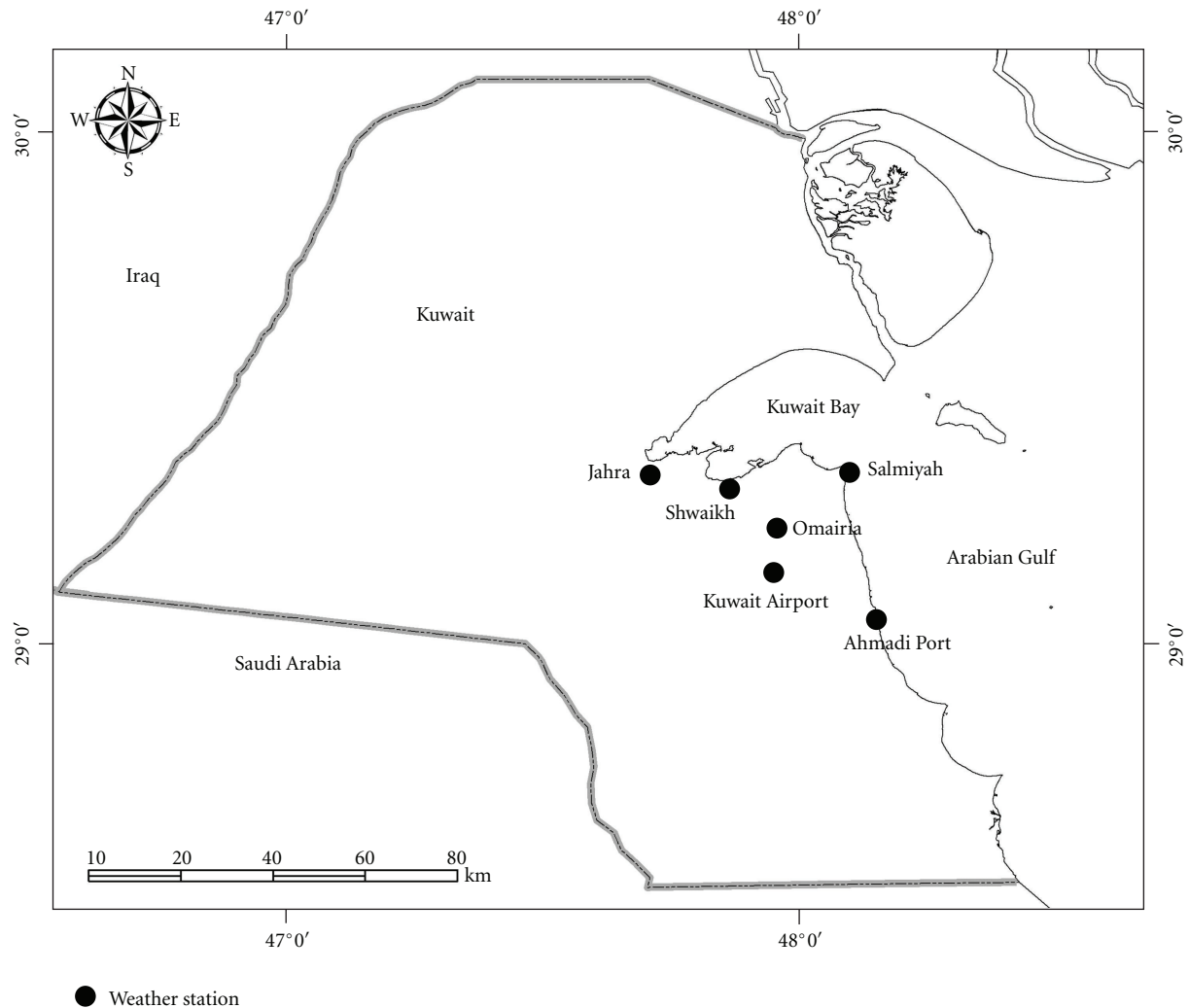


FIGURE 1: Weather stations in urban areas of Kuwait.

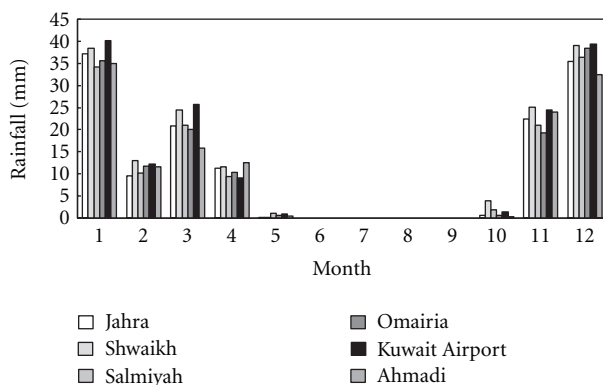


FIGURE 2: Seasonal mean of monthly total rainfall data calculated using (1) for the time duration from January 1994 to December 2005.

so large relative to the size of the data by which there would be no sufficient oscillating movement repetitive over the time interval.

Figure 5 presents the periodogram of the data. If one neglects the peaks shown in the figure for the time less than 6 months, then the following periods dominate: 6, 12, 18, 20, 26, 30, and 42 months. Compared to the periodogram in Figure 3 for the same weather station, more periods appeared as the time scale of the data became longer. The periods of 6, 12, and 18 are similar to that found previously. The period of 20 months can be considered very close to that of the 18 months. The larger periods of 26 and 30 corresponding to patterns of nearly 2 and 2.5 years can either be related to solar activities [11] or to the well-known tropical Quasi-Biennial Oscillation (QBO), which is believed to be functioning as a conduit to transfer solar effects to lower altitudes. The QBO is a dominant natural oscillation in the equatorial lower stratosphere described as the phenomenon of reversal of wind directions; that is, for about one year the prevailing wind direction is easterly, while during the following year it is westerly [12, 13]. The period of 42 months could be either due to solar activities or harmonic correlation with the period of 20 months (i.e.,  $20 \times 2 = 40$ ). As it was expected, the period of 360 months cannot be identified

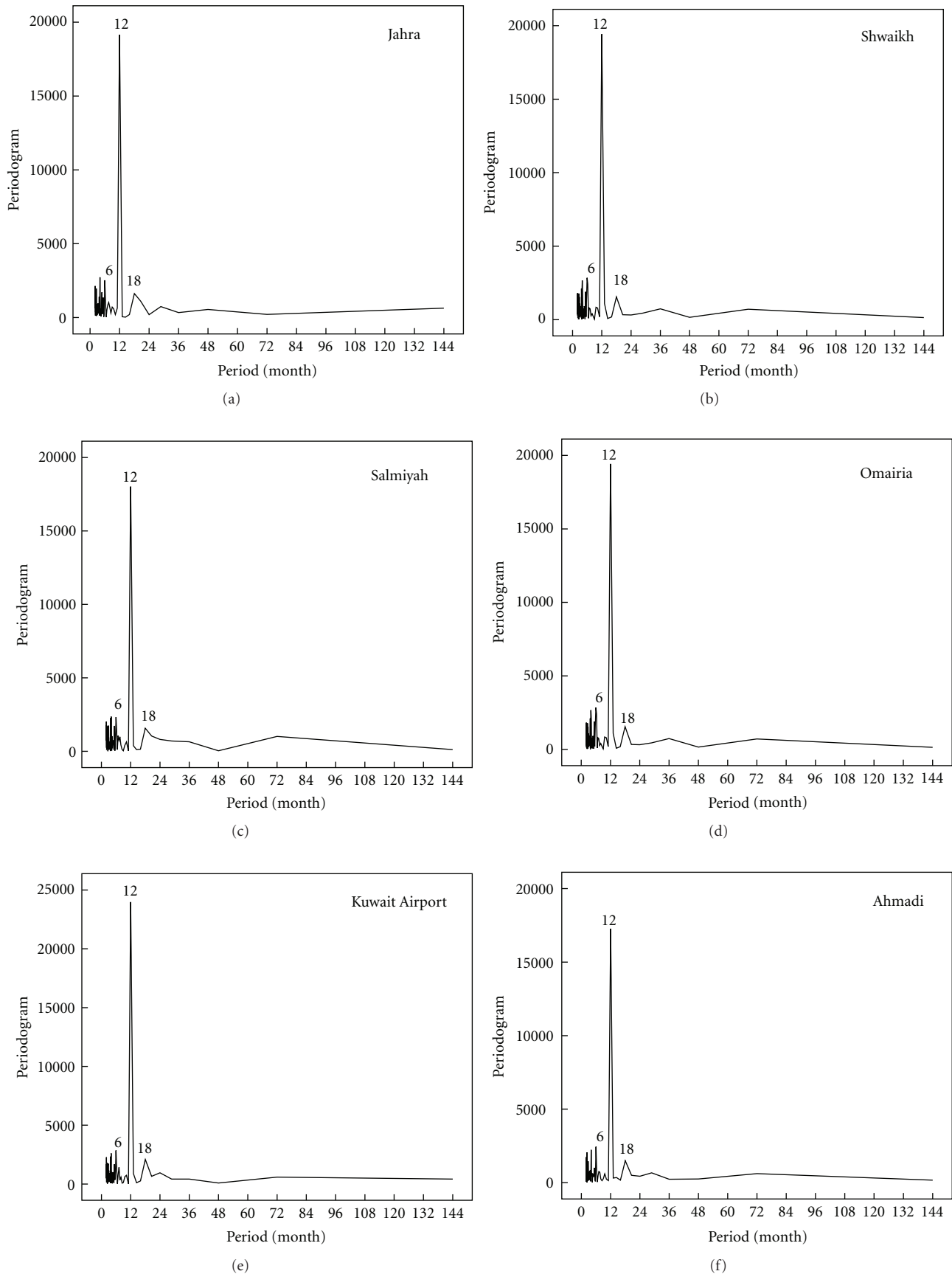


FIGURE 3: Periodograms for the monthly total rainfall data obtained from the weather stations in urban areas for the time duration from January 1994 to December 2005.

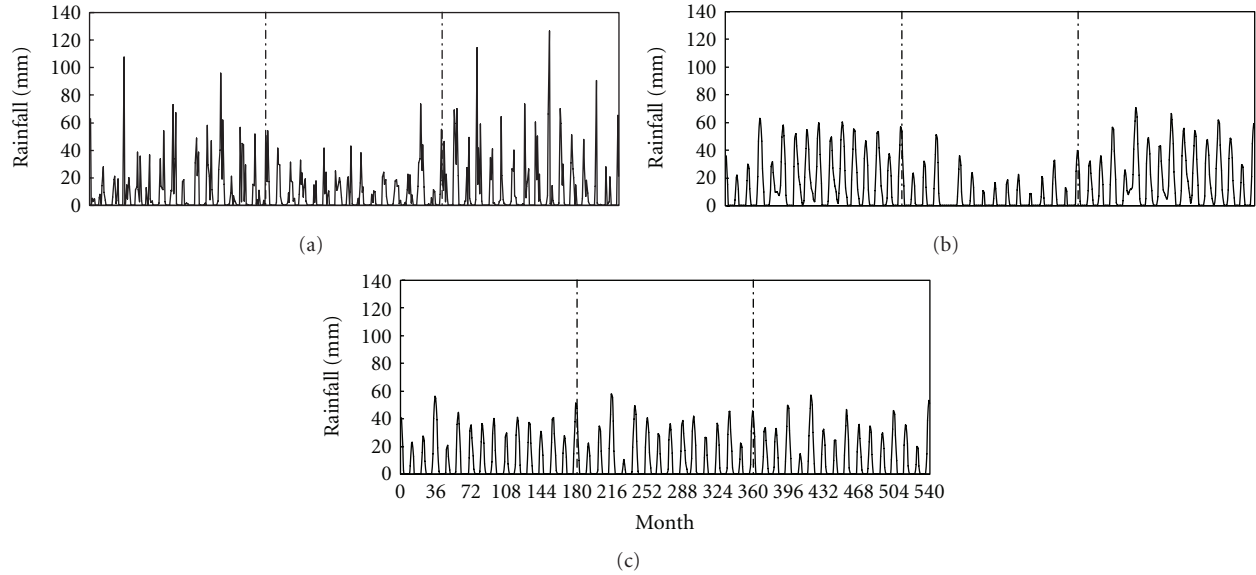


FIGURE 4: Monthly total rainfall for the weather station of Kuwait Airport for the time duration from January 1965 (corresponding to month number 0) to December 2009 (month number 540): (a) actual dataset; (b) calculated records based on periodic model; (c) calculated records based on periodic model with excluding period 360 months.

in the periodogram since its power is not high enough. One may thus conclude that other periods could also exist, although they might not appear explicitly due to insufficient data.

#### 4. Model Development

The data of Kuwait International Airport can be used to model the rainfall variability over the urban areas. Generally, a time series data can be represented by a decomposition model of additive type composed of deterministic and stochastic components. The deterministic component is described by trend and periodic parts, and it can be formulated in a manner that allows exact predictions. The trend describes a long movement of the variable lasting over the entire time of observation, while the periodic part describes oscillating movement repetitive over a specific time interval. The stochastic component of a time series data can never be estimated exactly as it is considered to be made of random effects. Although a stochastic component is taken to be sufficiently stationary in simple time series models, in most commonly considered situations it exhibits complicated statistical correlations.

While visual inspection of the monthly total data obtained from Kuwait International Airport suggests the presence of no trend but only periodic deterministic component, statistical inferences may be used to provide verification. This analysis can be accomplished using the annual total rainfall data. The claim to be tested is that a linear relation exists between the annual rainfall and time at a level of significance  $\alpha = 0.05$ . As can be seen in Figure 6, the  $P$  value for the slope of the fitted linear relation is given as 0.266. This means that

there is a 0.266 probability of obtaining a slope estimate as extreme as or more extreme than the one obtained if the null hypothesis of no linear relation was true. As the  $P$  value is more than the level of significance, the null hypothesis of no linear relation is accepted.

The periodic pattern of monthly rainfall data can now be estimated from the detected periods in the previous section. In general, a time-based data containing a periodic sinusoidal component with a known wavelength can be modeled using Fourier series, which can be expressed for multiperiods as

$$s(t) = \sum_{n=1}^{\infty} \sum_{i=1}^k R_{n,i} \cos(2n\pi f_i t + \theta_{n,i}), \quad (2)$$

where

$$\begin{aligned} R_{n,i} &= \sqrt{a_{n,i}^2 + b_{n,i}^2}, \\ \theta_{n,i} &= \tan^{-1} \left( -\frac{b_{n,i}}{a_{n,i}} \right), \\ a_{n,i} &= 2f_i \int_{L_i}^{L_i+1/f_i} f(x) \cos(2n\pi f_i t) dx, \\ b_{n,i} &= 2f_i \int_{L_i}^{L_i+1/f_i} f(x) \sin(2n\pi f_i t) dx, \end{aligned} \quad (3)$$

where  $s$  is periodic sinusoidal component of rainfall;  $R$  is amplitude of variation;  $f$  is frequency, equal to the inverse of period;  $\theta$  is phase angle; and  $k$  is total number of periodicities. The term  $(2n\pi f t + \theta)$  is measured in radians. As determined from the data,  $k$  value is equal to eight, and the values of  $f$  may be set by the periodic nature of rainfall data, that is,  $f_1 = 1/6$ ,  $f_2 = 1/12, \dots$  cycles per month. The phase angle,  $\theta$ , is necessary to adjust the model so

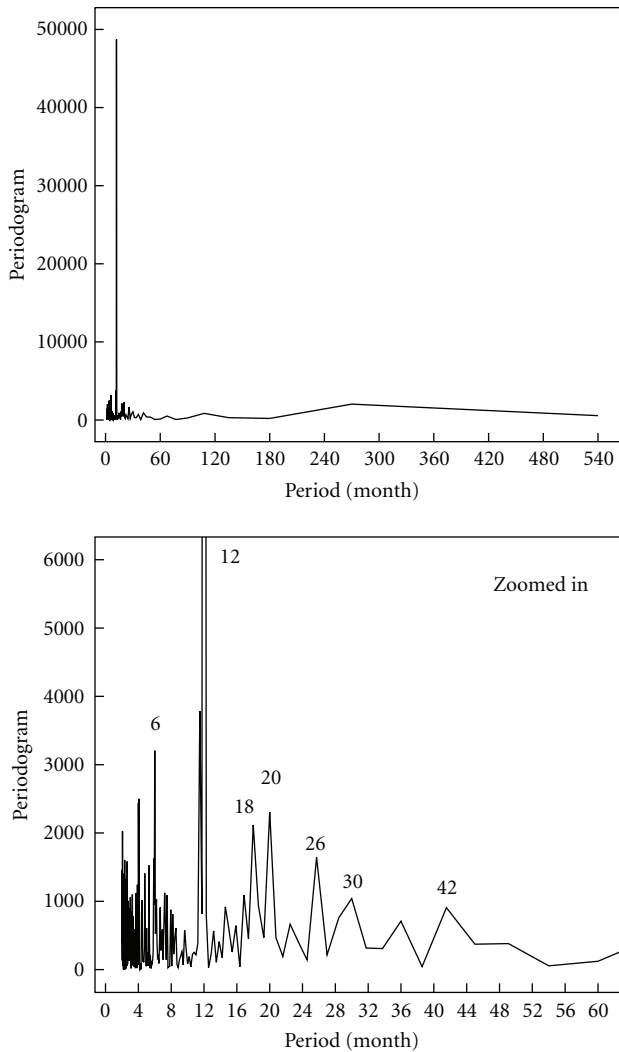


FIGURE 5: Periodogram for the monthly total rainfall data obtained from the weather station of Kuwait International Airport for the time duration from January 1965 to December 2009.

that the cosine function crosses the mean, which is equal to zero for the data, at the appropriate time  $t$ . The difficulty is to determine analytically the function  $f(x)$  in Fourier coefficients. However, because of the existing randomness, a more simple procedure may be followed by determining  $\theta_{n,i}$  and  $R_{n,i}$  by means of numerical optimization. Since the above equation will not be solved analytically, it is more convenient to reduce the number of fitting coefficients. This is achieved by testing a number of  $s(t)$  models each obtained by assuming a different value of  $n$ . Based on Fourier procedure, the larger  $n$  value considered, the higher model accuracy obtained. In this case, however, a higher accuracy would result with a more complicated model form because of the many periods detected in the data. Accordingly, it can be considered for simplicity  $n = 1$ . One difficulty remains is that a cosine function crosses a mean equal to zero will produce positive and negative values. In reality, there should be only positive values, while the negative ones should correspond to

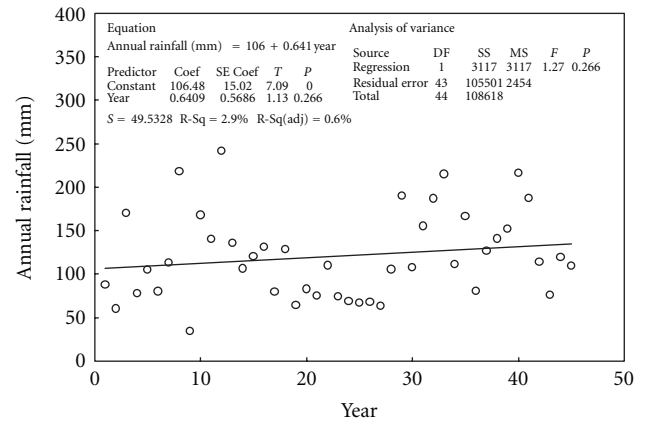


FIGURE 6: Annual total rainfall versus time for the data obtained from the weather station of Kuwait International Airport for the time duration from 1965 (corresponding to year number 1) to 2009 (year number 45). The solid line represents a trend fitted for the data using linear regression.

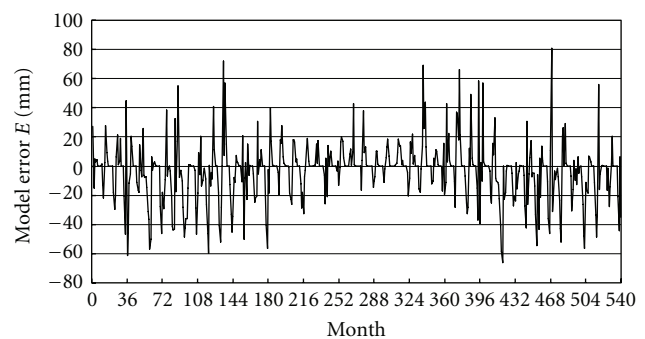


FIGURE 7: Model error  $E$  calculated by subtracting the monthly total rainfall data from the calculated values. The data represent that obtained from the weather station of Kuwait Airport for the time duration from January 1965 (corresponding to month number 0) to December 2009 (month number 540).

zero rainfall. To produce zero rainfall, the binary model is used

$$F(t) = \begin{cases} s(t) & \text{if } s(t) \geq 0, \\ 0 & \text{if } s(t) < 0. \end{cases} \quad (4)$$

Following the above procedure, a model of  $s(t)$  is obtained with coefficients shown in Table 1. The overall rainfall model  $F(t)$  is presented in Figure 4(b).

Diagnostic checking concerns the verification for the adequacy of the fitted model. The accuracy of the model can be evaluated by calculating the residuals  $E$  by subtracting the rainfall data from the calculated values. As seen in Figure 7, the model fits the data with some variation representing the remaining stochastic component of time series. The model



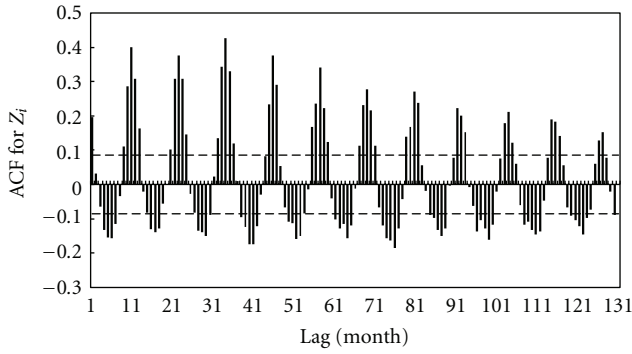


FIGURE 8: Correlogram up to lag 135 for series  $Z_i$ . The dashed lines represent upper and lower bounds with significance level  $\alpha = 0.05$ .

accuracy can be evaluated quantitatively by estimating the mean absolute error

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |E_i|. \quad (5)$$

It is worth mentioning that this statistical criterion for evaluating model accuracy is more convenient for rainfall data with zero values, while other well-known criteria such as the Nash-Sutcliffe equation and the Mean Absolute Percentage Error are inappropriate, as the actual value in their expressions is found in the denominator by which the equations become undefined for zero rainfalls. It is found that the overall model error estimated by (5) is about  $\text{MAE} = 12.6$  mm.

The remaining stochastic component can be tested for any possible persistence and/or random structure. To do this, the residuals are first standardized as follows:

$$Z_i = \frac{E_i - \bar{E}}{\sigma}, \quad (6)$$

where  $\bar{E}$  is mean residual of series  $E_i$ , which is for this data equal to  $\bar{E} = -5.36$  mm; and  $\sigma$  is standard deviation of residual of series  $E_i$ , equal to  $\sigma = 21.3$  mm. The series  $Z_i$  then has zero mean and unit standard deviation. The correlogram  $r$ , also known as the autocorrelation function ACF, can then be obtained for the series  $Z_i$  at lag  $L$  from

$$r_L = \frac{(1/(n-L)) \sum_{i=1}^{n-L} (Z_i - \bar{Z})(Z_{i+L} - \bar{Z})}{(1/n) \sum_{i=1}^n (Z_i - \bar{Z})^2}, \quad (7)$$

where  $\bar{Z}$  is mean of the sample  $n$  values of series  $Z_i$ . The lag  $L$  is usually taken for values up to  $n/4$ . The resulting values of  $r_L$  will range between  $-1$  and  $+1$ . Figure 8 shows the estimated correlogram. Here, two lines of upper and lower confidence bands were drawn with significance level  $\alpha = 0.05$  to examine whether the data are random

$$\text{C.B.} = \pm \frac{z_{1-\alpha/2}}{\sqrt{n}}, \quad (8)$$

where  $z$  is quantile function of standard normal distribution. If the correlogram is higher (lower) than this upper (lower)

TABLE 1: Estimated coefficients for rainfall model  $s(t)$  with  $n = 1$ .

$i$	$R_{n,i}$ (mm)	$f_i$ (1/month)	$\theta_{n,i}$
1	8	1/6	1
2	30	1/12	0.1
3	7	1/18	-0.2
4	10	1/20	-5
5	4	1/26	-0.9
6	6	1/30	-5.8
7	-6	1/42	3.3
8	-20	1/360	1.3

band, the null hypothesis that there is no autocorrelation at and beyond a given lag is rejected at the given significance level. As can be seen, the correlogram exhibits an alternating sequence of positive and negative spikes. Such a pattern is the correlogram signature of a repetitive oscillating movement, suggesting that the remaining stochastic component in this data is due to the presence of other periods that have not been taken into account in the model. Nevertheless, owing to the expected complicated behavior of rainfall and to the insufficient data for this station, the remaining stochastic component may be considered nondeterministic.

The role of hidden periods on data modeling can be highlighted by testing the influence of period of 360 months. As it was mentioned earlier, the 360-month period is not present in the periodogram of monthly total rainfall for the station of Kuwait International Airport during the time from January 1965 to December 2009, although it was considered in the process to estimate the model parameters. If this period had been excluded from the model, then the rainfall pattern would have changed as shown in Figure 4(c). Obviously, the model in this case becomes insufficient to capture the large scale of the data pattern. The overall model error here becomes  $\text{MAE} = 13.2$  mm, which is about 4.7% larger than before. Accordingly, it can be assumed that the spikes shown in Figure 4(a) such as those for months 35, 135, 339, 376, 396, 470, and 518 that have not been simulated sufficiently by the model may result from hidden periods that had not been taken into account. The implication is that the difficulty in fitting rainfall models by this criterion would require a wide range of historical data sufficient to show all possible periods responsible for the variable pattern.

## 5. Conclusions

This study examined the spatial and temporal variability of monthly total rainfall data obtained from weather stations located in the urban areas of Kuwait. It has been shown that a point estimate of rainfall can be considered spatially representative over the urban areas as the weather stations are sufficiently close in distance. This finding has been proved useful to model rainfall variability over the urban areas using a single dataset obtained from Kuwait International Airport for the time from January 1965 to December 2009. A sinusoidal model has been proposed by considering a number of periods of 6, 12, 18, 20, 26, 30, 42, and 360

months. Except for the last one, all the periods have been identified in the rainfall data of Kuwait International Airport using the periodogram technique. The absence of the last period, which is due to insufficient length of time of data, reveals the role of hidden periods on model accuracy. This is confirmed by plotting the correlogram of the standardized residuals for the remaining stochastic component of the model, which exhibits the signature of a repetitive oscillating movement due to the presence of other periods that have not been taken into account. A wider range of rainfall data would thus be necessary to develop a more accurate periodic model.

## Acknowledgment

The author is grateful to the Directorate of Civil Aviation of Kuwait for providing the rainfall data used in this study.

## References

- [1] S. S. Kashid, S. Ghosh, and R. Maity, "Streamflow prediction using multi-site rainfall obtained from hydroclimatic teleconnection," *Journal of Hydrology*, vol. 395, no. 1-2, pp. 23–38, 2010.
- [2] F. Melone, C. Corradini, R. Morbidelli, C. Saltalippi, and A. Flammini, "Comparison of theoretical and experimental soil moisture profiles under complex rainfall patterns," *Journal of Hydrologic Engineering*, vol. 13, no. 12, pp. 1170–1176, 2008.
- [3] D. Jukic and V. Denić-Jukić, "Groundwater balance estimation in karst by using a conceptual rainfall-runoff model," *Journal of Hydrology*, vol. 373, no. 3-4, pp. 302–315, 2009.
- [4] G. M. Foody, E. M. Ghoneim, and N. W. Arnell, "Predicting locations sensitive to flash flooding in an arid environment," *Journal of Hydrology*, vol. 292, no. 1–4, pp. 48–58, 2004.
- [5] J. Almedeij, A. Alsulaili, and J. Alhomoud, "Assessment of grate sag inlets in a residential area based on return period and clogging factor," *Journal of Environmental Management*, vol. 79, no. 1, pp. 38–42, 2006.
- [6] G. K. Young and S. E. Walker, "Risk-cost design of pavement drainage systems," *Journal of Water Resources Planning and Management*, vol. 116, no. 2, pp. 205–219, 1990.
- [7] N. A. Zaghloul and B. L. Al-Mutairi, "Water harvesting of urban Runoff in Kuwait," *Scientia Iranica A*, vol. 17, no. 3, pp. 236–243, 2010.
- [8] K. A. Murtaugh, *Analysis of sustainable water supply options for Kuwait*, M.S. thesis, Massachusetts Institute of Technology, Cambridge, Mass, USA, 2006.
- [9] M. P. Marcella and E. A. B. Eltahir, "The hydroclimatology of Kuwait: explaining the variability of rainfall at seasonal and interannual time scales," *Journal of Hydrometeorology*, vol. 9, no. 5, pp. 1095–1105, 2008.
- [10] P. J. Brockwell and R. A. Davis, *Introduction to Time Series and Forecasting*, Springer, New York, NY, USA, 2002.
- [11] G. A. Bazilevskaya, M. B. Krainev, V. S. Makhmutov, E. O. Flückiger, A. I. Sladkova, and M. Storini, "Structure of the maximum phase of solar cycles 21 and 22," *Solar Physics*, vol. 197, no. 1, pp. 157–174, 2000.
- [12] J. K. Angell and J. Korshover, "Quasi-biennial variations in temperature, total ozone, and tropopause height," *Journal of the Atmospheric Sciences*, vol. 21, no. 5, pp. 479–492, 1964.
- [13] R. J. Reed, W. J. Campbell, L. A. Rasmussen, and R. G. Rogers, "Evidence of a downward propagating annual wind reversal in the equatorial stratosphere," *Journal of Geophysical Research*, vol. 66, no. 3, pp. 813–818, 1961.



