An Investigation of Warm-Season Spatial Rainfall Variability in Oklahoma City: Possible Linkages to Urbanization and Prevailing Wind

LAUREN M. HAND* AND J. MARSHALL SHEPHERD

Department of Geography, The University of Georgia, Athens, Georgia

(Manuscript received 8 May 2008, in final form 24 July 2008)

ABSTRACT

This study used 9 yr (1998–2006) of warm-season (June–September) mean daily cumulative rainfall data from both the Tropical Rainfall Measuring Mission (TRMM) Multisatellite Precipitation Analysis and rain gauge stations to examine spatial variability in warm-season rainfall events around Oklahoma City (OKC). It was hypothesized that with warm-season rainfall variability, under weakly forced conditions, a rainfall anomaly would be present in climatological downwind areas of OKC. Results from both satellite and gauge-based analyses revealed that the north-northeastern (NNE) regions of the metropolitan OKC area were statistically wetter than other regions. Climatological sounding and reanalysis data revealed that, on average, the NNE area of OKC was the climatologically downwind region, confirming that precipitation modification by the urban environment may be more dominant than agricultural/topographic influences on weakly forced days. The study also established that satellite precipitation estimates capture spatial rainfall variability as well as traditional ground-based resources do. TRMM products slightly underestimate the precipitation recorded by gauges, but the correlation improves dramatically when the analysis is restricted to mean daily rainfall estimates from OKC urban grid cells containing multiple gauge stations ($R^2 = 0.878$). It was also quantitatively confirmed, using a relatively new concentration factor analysis, that prevailing wind–rainfall yields were consistent with the overall framework of an urban rainfall effect. Overall, the study establishes a prototype method for utilizing satellite-based rainfall estimates to examine rainfall modification by urbanization on global scales and in parts of the world that are not well instrumented with rain gauge or radar networks.

1. Introduction

In 2008, more than one-half of the world’s population lived in urban areas, and this fraction could balloon to 81% by 2030 (UNFPA 2007). “Urban footprints” spread well beyond the immediate vicinity of cities, affecting local- to global-scale atmospheric composition, surface energetics, water and carbon cycle processes, and ecosystem development. The urban heat island (UHI) is well studied (Oke 1982), and the Intergovernmental Panel on Climate Change report (Trenberth et al. 2007) highlighted the growing body of research linking urban-related processes and regional precipitation. Increases in precipitation due to urban effects are typically observed at distances between 30 and 75 km from the city center (Landsberg 1970; Sanderson and Gorski 1978; Bornstein and Lin 2000; Shepherd et al. 2002; Shepherd 2006; Mote et al. 2007). An idealized diagram based on Shepherd et al. (2002) showing the region of highest rainfall increases due to urban effects is shown in Fig. 1. Areas 25–75 km downwind of the city center and within a 125° sector will typically experience the greatest increase in rainfall due to the urban region.

Despite recent results (Shepherd 2005; Mote et al. 2007; Stallins and Rose 2008; Rose et al. 2008) providing evidence that convection and precipitation can be enhanced or initiated by urban regions, the debate over the role of urban environments on precipitation is ongoing. Kaufmann et al. (2007) suggested that urbanization in the Pearl River Delta of China has reduced local precipitation because of changes in surface hydrology. Guo et al. (2006) found decreased cumulative rainfall around Beijing, China. Smaller cloud droplet size distributions and suppressed rainfall have been shown to occur because of increased aerosol concentrations from anthropogenic sources over and downwind of urban areas (Rosenfeld 1999, 2000; Borys et al. 2003; Givati

Corresponding author address: Marshall Shepherd, Dept. of Geography, The University of Georgia, GG Building, Rm. 106, Athens, GA 30602.

E-mail: marshgeo@uga.edu

DOI: 10.1175/2008JAMC2036.1

© 2009 American Meteorological Society
and Rosenfeld 2004). Recent studies (van dan Heever and Cotton 2007; Rosenfeld et al. 2007; Jin et al. 2005) are beginning to shed light on the possible role of giant cloud condensation nuclei (CCN) (enhancement) and smaller CCN (suppression).

Motivation

Lowry (1998) discussed several potential problems with methods and inferences used in many historical studies of urban-induced precipitation. He noted that the urban effects being sought usually cannot be distinguished from other influences, such as local topography and temporal changes in the relative frequencies of different synoptic weather types. Such uncertainties highlight the need for more observational and modeling research in this area (Dabberdt et al. 2000).

This research takes a unique approach to examine urbanization and rainfall modification. The study uses 9 yr (1998–2006) of warm-season (June–September) mean daily rainfall accumulation obtained from both the Tropical Rainfall Measuring Mission (TRMM) (Kummerow et al. 2000; Huffman et al. 2007) Multi-satellite Precipitation Analysis (TMPA) and ground-based gauge stations to examine spatial variability in warm-season rainfall events around Oklahoma City (OKC), Oklahoma, as a function of prevailing wind and urban land cover. To be more specific, the objectives are 1) to spatially quantify the variability of warm-season rainfall in the Oklahoma City area under weak large-scale conditions, 2) to assess the viability of using satellite approaches relative to ground-based data to observe urban rainfall anomalies, and 3) to investigate Oklahoma City’s rainfall variability as a function of prevailing wind regimes. An overarching goal of this study is to establish a prototype method in a well-instrumented location for using satellite-based rainfall in rapidly urbanizing areas of the globe not well instrumented with rain gauge or radar networks.

It is hypothesized that the climatologically downwind region is the preferred location for rainfall events—in particular, heavy and extremely heavy cumulative rainfall events, on weakly forced days. Based on this hypothesis, the location of cumulative and extreme rainfall anomalies is dependent on the prevailing wind regime relative to urban land cover. A new concentration factor analysis illustrates a relationship between the location of cumulative rainfall and wind direction.

2. Background

a. Urban effects on precipitation in the Great Plains region

The southern Great Plains have played host to a relatively small number of studies investigating the enhancement of precipitation by urban effects. A recent study by Niyogi et al. (2006) simulated a mesoscale event in OKC (Fig. 2). Their land surface model, which included an urban canopy model, showed that OKC concentrated the precipitation downwind of the urban region. In a similar way, increased rainfall amounts of 25 mm were resolved 55 km downwind of OKC using a high-resolution mesoscale model to simulate convection observed during the International H2O Project (Simpson 2006). Studies have also examined the space–time structure of extreme storm rainfall in the southern plains through empirical analyses of radar and rain gauge data (Smith et al. 1994). Their findings illustrate that the climatological rain gauge network poorly represents isolated heavy rainstorms.

b. Characteristics of Oklahoma City

A potential shortcoming of any study that attempts to link rainfall modification with urban areas is the difficulty of separating topographic and other effects (e.g., sea-breeze circulations) from urban effects (Shepherd et al. 2002). OKC was selected for the study because of its well-defined central city, moderate size, relatively flat terrain (roughly 391 m), well-characterized climatological behavior, lack of topographic or water barriers, and wealth of supporting meteorological instrumentation (Fig. 3). Surface observations from the Oklahoma Mesonet (http://www.mesonet.org/sites/) (Shafer et al.
2000) and the Oklahoma Climatological Survey (http://climate.ocs.ou.edu) provide valuable datasets with a broad spatial coverage.

OKC is located in central Oklahoma and has an estimated population of 528,042 (U.S. Census Bureau 2005). Roughly 1.3 million residents live within the greater metropolitan limits of OKC. The city has a total land area of 1608.8 km² (621.2 mi²), making it the third largest city in America based strictly on geographic area. The U.S. Census Bureau (2005) data on urbanized

![High-resolution simulations using the Coupled Ocean–Atmosphere Mesoscale Prediction System to investigate the impact of urban and land vegetation processes on the prediction of the mesoscale convective systems observed on 30 Jul 2003 in the vicinity of OKC (from Niyogi et al. 2006).](a) 1015 UTC
(b) 1100 UTC
(c) 1130 UTC
(d) 1400 UTC

**FIG. 2.** High-resolution simulations using the Coupled Ocean–Atmosphere Mesoscale Prediction System to investigate the impact of urban and land vegetation processes on the prediction of the mesoscale convective systems observed on 30 Jul 2003 in the vicinity of OKC (from Niyogi et al. 2006).
areas ranks OKC 12th among urbanized areas with greatest sprawl (1970–90).

Mean annual rainfall over the region is characterized by a pronounced east–west gradient, with annual accumulations of less than 400 mm in western Oklahoma and greater than 1200 mm in eastern Oklahoma (Smith et al. 1994). A portion of warm-season rainfall variability may be attributed to the agricultural lands and irrigation practices in OKC’s surrounding areas (Niyogi et al. 2006). More observational and modeling work is needed to distinguish urban influences from agricultural influences.

The seasonal distribution of extreme rainstorms in the southern plains has shown pronounced peaks in late spring and early autumn. However, high levels of moisture and convective instability are most common during the warm season (Bradley and Smith 1994). Thunderstorms are common across Oklahoma during the summer months and occur primarily as a result of highly localized events. A northward-retreating jet stream leads to broad areas of high pressure with weak synoptic forcing. The steering-level flow (i.e., 700 hPa; Shepherd et al. 2002) is typically southwesterly (Fig. 4).

c. Field studies in Oklahoma City

Joint Urban 2003 focused on the understanding of atmospheric processes within the urban environment to advance knowledge about movement of contaminants in and around cities and into and within building interiors (Allwine 2004). The resulting data have been used to improve, refine, and verify computer models that simulate the atmospheric transport of contaminants in urban areas. Joint Urban 2003 specifically selected intensive operation periods on rain-free days (J. Basara 2007, personal communication); however, there is clearly a need to consider rain days in future experiment designs.

3. Research design, data, and method

Nine years of mean daily rainfall estimates from both ground- and satellite-based datasets were employed.
The study focused on only the warm season months (June–September) of 1998–2006. The 9-yr period reflects the availability of TRMM data at the time of study. The overwhelming consensus from the Metropolitan Meteorological Experiment (METROMEX; Braham et al. 1981) and other efforts is that urban effects are most pronounced during warm-season months (Huff and Changnon 1972; Changnon et al. 1991; Jauregui and Romales 1996), when the UHI-induced mesoscale circulation is more dominant and can significantly alter boundary layer processes.

Mean daily rainfall estimates from gauge and satellite estimates were analyzed to quantify the spatial variability of precipitation around OKC under weak large-scale conditions. To demonstrate opportunities to observe multiple urban rainfall anomalies over an extensive area using satellite data, a statistical analysis was performed comparing the area-averaged TRMM-based and point-source rain gauge estimates. Further, a relatively new statistical method, the concentration factor (CF), was utilized to investigate the linkage between OKC’s spatial rainfall variability and prevailing wind regime. These methods will be described in greater detail later. Results from these analyses were used to characterize the possible role of urban land cover on the precipitation variability.

a. Study area

For this study, the OKC region is defined by a 75 km by 75 km square grid (5625 km²). The study area was divided into nine equally averaged 25-km grid squares corresponding to the 0.25° (~25 km) resolution of the TRMM product. The center point for the grid is located at 35.482°N latitude and 97.535°W longitude, the heart of OKC’s central business district (CBD). The central grid cell represents the urban core (CBD), and the remaining cells denote the eight directional components (NW, N, NE, E, SE, S, SW, and W) surrounding the urban environment. The area was delineated using the U.S. Geological Survey National Map viewer (http://nationalmap.gov), and based upon the 2001 National Land Cover Database (Fig. 5). Areas of red and light gray are representative of light to heavy urban land cover in OKC. A distinction is made between the primary land uses around OKC. Land to the west (east) is primarily agricultural (forested).

b. Data

1) AIRMASS-TYPE DATA

The spatial synoptic classification (SSC) method (http://sheridan.geog.kent.edu/ssc.html) was employed to determine the type of synoptic environment during the warm season (June–September) of 1998–2006 in OKC. SSC classifies seven different weather types: dry polar (DP), dry moderate (DM), dry tropical (DT), moist polar (MP), moist moderate (MM), moist tropical (MT), and moist tropical plus (MT⁺). Transitional (TR) is applied to days during a period of transition from one air mass to another.

Only days reflecting an MT or MT⁺ (hereinafter MT) air mass were used in this study. Atmospheric instability and convective activity are most common under these conditions. The warm and humid MT air masses represent an atmosphere that is “synoptically benign,” limiting precipitation mechanisms to convective development (no other forms of synoptic forcing). Of the total of 1098 days in the study period, 476 days (43%) were classified as having the MT air mass. The decision to stratify the data by SSC classification is consistent with one of Lowry’s key recommendations for urban
studies: stratification and disaggregation of data to avoid merging effects between dissimilar synoptic systems.

2) SATELLITE DATA

Satellite-based daily rainfall amounts from the TMPA were acquired from the Goddard Distributed Active Archive Center. TMPA is a 3-hourly, 0.25° (~25 km) product based on the TMPA described in Huffman et al. (2007). This study used the merged 3B42 daily-accumulation version of the TMPA, composed of available microwave (e.g., TRMM Microwave Imager, Special Sensor Microwave Imager, Advanced Microwave Scanning Radiometer, and Advanced Microwave Sounding Unit) and calibrated infrared (IR) estimates. Note that the merged products like TMPA are so relatively new that there is a lack of intercomparison/validation studies in the literature. This makes our study a valuable contribution in itself. Ebert et al. (2007), in a 1-yr study, showed that satellite products like TMPA perform better in summer months. However, that study compared a real-time version of the TMPA product that we applied. Herein, we describe a research version of the TMPA that contains a gauge adjustment. The resolution of the gauge network is coarse [e.g., the 2.5° Global Precipitation Climatology Center monthly gauge dataset described by Rudolf et al. (1994)], and so it is encouraging that, even with this coarse adjustment, our results using 0.25° data products are still very accurate.

3) RAIN GAUGE DATA

Ground-based daily precipitation data were extracted for 11 stations from the TD3200 Cooperative Summary of the Day database of the National Climatic Data Center (NCDC) (Table 1). Selection of stations was based on spatial coverage and quality control criteria. The stations chosen for the analysis were located within 37.5 km of OKC in lieu of the 25-km-resolution grid of the merged TRMM 3B42 product and the 12.5 km extending from the central point of the urban cell. An effort was made to acquire the greatest possible number of stations across the study area, having at least one station in each grid cell.
The selected stations did not have more than 15% of days missing over the warm season during the 1998–2006 study period. Three stations from the Oklahoma Mesonet were utilized to supplement gaps in spatial coverage when NCDC cooperative observing (coop) stations failed to meet the selection criteria. The locations of all 14 stations are shown in Fig. 6.

4) RADIOSONDE DATA

Sounding data for the Norman, Oklahoma, (OUN) sounding site were obtained from the radiosonde data archive (raob) produced jointly by NCDC and the Forecast Systems Laboratory (http://raob.fsl.noaa.gov). The raob database merges NCDC TD6301 observations with observations from the Integrated Global Radiosonde Archive. The 1200 UTC (0700 central daylight time) sounding was used in this study to determine daily prevailing wind direction at 700 hPa. An acceptable threshold of 10% missing data over the study period was established. As noted earlier, the predominant 700-hPa wind direction during the warm-season months is from the southwest. The frequency of occurrence for each wind direction is shown in Table 2.

c. Research method

1) SPATIAL DISTRIBUTION OF RAINFALL IN OKLAHOMA CITY

It is hypothesized that cumulative rainfall—in particular, heavy-rainfall days and extremely heavy rainfall days—will not be uniformly distributed around OKC on weakly forced days but rather will exhibit a preferred region of development in the climatological urban downwind region as determined by analysis of the Norman sounding. To quantify daily rainfall accumulation, region-specific cumulative rainfall categories were established. Cumulative daily precipitation data from the Norman and OKC Will Rogers coop stations were filtered conditionally to include only those days with measurable rainfall. These two stations were selected because they had long records and a fixed location over the study period. Using the conditional daily precipitation amounts, descriptive statistics and frequency distribution curves were produced to examine the natural and quartile breaks in the rainfall data.

Threshold values (Table 3) were designated in four accumulation categories. Precipitation amounts within the first and second quartiles (below the median) were classified as low-accumulation days. Moderate accumulation was assigned to values falling between the second and third quartiles (50th–75th percentile). Values above the upper quartile but below the 90th percentile were classified as heavy accumulation and those above the 90th percentile were designated as extreme.

Daily precipitation estimates from the TRMM satellite and gauge locations were acquired and filtered by warm-season and spatial synoptic classifications. In grid cells having multiple gauges, the average of both locations was used to represent the mean daily rainfall for that cell. Frequency distributions of percent occurrence of daily rainfall totals and other variables were generated for each accumulation class. The urban rainfall ratio URR, as described in Shepherd et al. (2002), was also computed for each grid cell:

$$URR = \frac{R_i}{R_{BG}},$$

where $R_i$ is the mean daily accumulated rainfall at a grid box over the 9-yr study period and $R_{BG}$ is the average
2) CORRELATION ANALYSIS

It is hypothesized that the TRMM-based estimates will be relatively well correlated with point-source rain gauge estimates. Point-source gauge methods may not be appropriate for analysis of spatial distributions of rainfall, which, unlike temperature or moisture fields, tends to be highly variant, especially during warm-season months. The point-source/area-average problem is well known in precipitation remote sensing communities and is an ongoing area of research.

Comparisons between satellite-estimated precipitation and rain gauge data have typically used linear regression analysis as the main analytic tool for assessment (Barrett et al. 1994). Two correlation analyses were conducted for this study to establish the relative accuracy of a point-source (gauge) versus areally averaged (satellite) rainfall estimate. A scatterplot of mean daily rainfall estimates from the nine TRMM 25-km grid cells versus the corresponding estimates from the gauge stations was prepared to determine how well (or poorly) the rainfall distribution from the TRMM data correlates with the results of the gauge data. Because the satellite estimate of precipitation is area averaged, a second plot was created using only grid cells that contain multiple gauges. With this restriction, the correlation was limited to five grid blocks: the N, NE, S, SW, and NW cells.

Various statistical parameters can be used to measure the strength of the statistical relationship between the estimated (satellite) and the reference (gauge) values. The statistical criteria used in this study are 1) the bias or absolute error, which measure the difference

<table>
<thead>
<tr>
<th>Direction</th>
<th>Days</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>42</td>
<td>9.31</td>
</tr>
<tr>
<td>NE</td>
<td>30</td>
<td>6.65</td>
</tr>
<tr>
<td>E</td>
<td>31</td>
<td>6.87</td>
</tr>
<tr>
<td>SE</td>
<td>30</td>
<td>6.65</td>
</tr>
<tr>
<td>S</td>
<td>87</td>
<td>19.29</td>
</tr>
<tr>
<td>SW</td>
<td>111</td>
<td>24.61</td>
</tr>
<tr>
<td>W</td>
<td>58</td>
<td>12.86</td>
</tr>
<tr>
<td>NW</td>
<td>62</td>
<td>13.75</td>
</tr>
</tbody>
</table>

TABLE 2. The frequency of occurrence of the eight wind direction classes at 700-hPa level for warm-season, airmass-filtered days (n = 476).
between the mean satellite estimate and the mean gauge estimate, 2) the linear correlation coefficient $R$, which measures the strength and direction of the linear relationship between satellite and gauge estimates, 3) the root-mean-square error (RMSE, which is an absolute measure of the distance between estimate and reference, and 4) the coefficient of determination $R^2$, which is a statistical measure of how well the regression line approximates the reference (gauge) estimates. The range of $R^2$ values between 0.33 and 0.65 was selected as an acceptable (moderate) degree of correlation based upon the interpretation presented by McGrew and Monroe (2000).

3) RELATING AIRFLOW DIRECTION AND DAILY RAINFALL

Given that wind direction is an important factor for determining the geographical distribution of precipitation relative to an urban area, the final segment explores the relationship between prevailing wind flow and rainfall around OKC. It is hypothesized that there will be a relationship between the location of maximum cumulative rainfall and wind direction. Further, we hypothesize that the location of maximum cumulative rainfall will be coincident with the climatological mean downwind region. A statistical method called the concentration factor was used to examine the linkage between OKC’s spatial rainfall variability and prevailing wind regime. The CF provides a simple and efficient method to quantify wind regime–rainfall relationships for multiple locations in a single parameter rather than using complicated statistical functions or multiple analyses. This relatively new method was used to examine, given weak large-scale forcing, what prevailing wind flow regime would lead to the most active rainfall production in different regions around the city. Because of the topographical attributes of OKC, our assumption is that no preferred region for enhanced rainfall exists around OKC under specific prevailing wind flows. If one does exist, it is due to urban effects that have been previously documented in the literature.

Daily radiosonde soundings from the Norman location were used to determine mean prevailing flow at 700 hPa (steering level). Because rainfall record length for the gauge stations varied, the CF analyses were restricted to daily precipitation estimates from the merged 3B42 TRMM-based product.

The concept of the CF is to relate spatial variations in rainfall amount and intensity to frequency of occurrence of a parameter. Phillips and McGregor (2001) sought relationships between airflow and rainfall for multiple sites in the British Isles (their Fig. 3.4). The CF is presented in Eq. (2):

$$CF = 100 \left( \frac{R_f - N_f}{NR_d - N} \right),$$

where $R_f$ is the total amount (mm) of precipitation falling on all of the days in the sample with a given flow direction, $R_d$ is the long-term mean daily rainfall (1998–2006), $N_f$ is the number of days with a given flow direction, and $N$ is the total number of days. A few bad or missing sounding points at 700 hPa reduced the total number of days to $N = 451$. If the CF is positive (negative), then the rainfall percentage is greater (lower) than the percentage frequency of occurrence of that particular flow group, hence the term concentration factor. The CF analysis was applied to each sector to identify what wind flow at 700 hPa is most common for heavy-extreme-rainfall days.

4. Results

a. Rainfall variability and correlation analysis

Warm-season, airmass-filtered mean daily rainfall totals from both satellite and gauge estimates were computed for each grid block to investigate the spatial variation in daily rainfall totals around OKC. In Fig. 7, mean daily rainfall totals at each gauge location are represented as proportional circles. The size of a circle reflects the actual rainfall value, the radius being proportional to the magnitude of mean daily rainfall. Satellite and area-averaged gauge estimates of mean daily rainfall for the period 1998–2006 are shown in Fig. 8.

Both rainfall methods indicate that the N–NE regions of the metropolitan OKC have higher daily rainfall totals relative to other areas. The larger daily accumulations in the N–NE tier are consistent with hypothesized downwind urban rainfall effect because the warm season is dominated by southwesterly flow. Gauge estimates generally recorded slightly larger mean daily rainfall values over the study period than did satellite estimates. Statistical significance of these differences is tested later.

Daily rainfall accumulation from both satellite and gauge estimates was averaged over the study period and ranked across the nine grid cells (Table 4). From the table it is evident that the largest difference between the

| Table 3. Region-specific cumulative rainfall categories for OKC. |
|----------------------|--------|-------|
| Rainfall classification | mm    | in.   |
| Low                  | 0–5   | 0.0–0.2 |
| Moderate             | 6–13  | 0.2–0.5 |
| Heavy                | 14–26 | 0.5–1.0 |
| Extreme              | >26   | >1.0  |
two estimates occurs in the SE cell. The SE region represents the region of lowest mean daily rainfall under satellite estimates but the third-largest under gauge estimates. The most likely explanation for this discrepancy is that the one gauge at Stella experienced a large rainfall event that was smoothed out by the TMPA area average. A review of descriptive statistics reveals that the maximum daily total at Stella was 61.72 mm whereas the maximum daily total in the TMPA SE cell was only 47.70 mm. Stella also exhibits a larger standard deviation than the SE cell. In addition, when we averaged only the days for which TMPA and Stella had observations, the mean values were essentially the same. It was important for us to show the liability of missing gauge data, and therefore we chose the approach of averaging over the full range of Sheridan classified days. The two gauge stations with the highest missing data percentages (Arcadia and Norman) are located in grid cells (NE and S, respectively) with more than one gauge station.

The ranking of the grid cells by mean daily rainfall is most consistent for regions with multiple gauge locations, which provide a better representation of area-averaged precipitation for comparison with TRMM estimates. With the exception of the E region, the largest differences in rainfall estimates occur in grid cells with a single gauge location: the SE, W, and urban blocks. In addition, larger mean daily rainfall totals are recorded by gauge estimates in single gauge cells.

Another way to portray the deviation from a background value is the URR presented in Shepherd et al. (2002). The URR [Eq. (1)] calculates the ratio of mean accumulated rainfall at a particular grid box to the mean over the entire extent. Figure 9 displays the results as a percent difference from the area average. Mean daily accumulation in the N and E tiers exceeds that of the mean of the area average under both estimates. Conversely, S and SW sectors have mean daily accumulations below the area average. The poor spatial coverage of gauges in the W and urban sectors (single-gauge cells) is the most likely cause of the differences in rainfall estimates observed there.

1) CORRELATION ANALYSIS

Mean daily rainfall estimates from the TRMM 0.25° (~25 km) grid cells were compared with the corresponding estimates from the gauge stations to determine how well (or poorly) the rainfall distribution
from the TRMM data correlates with the gauge data. Figure 10 is the scatterplot of satellite versus gauge precipitation estimates for all nine grid cells in the study area. Because the satellite estimate of precipitation is area averaged, a second plot was created using only grid cells that contain multiple gauges. Under these criteria, the correlation was limited to five grid blocks: the N, NE, S, SW, and NW cells (Fig. 10). The linear regression
line and its corresponding equation and $R^2$ value are shown.

Over the study area, the TRMM product slightly underestimates the precipitation recorded by gauges, but the mean bias ($-0.18 \pm 0.24$) is well within the range (from $-1.0$ to $1.0$ mm day$^{-1}$) noted by Tian et al. (2007). The regression lines for both precipitation estimates are located below a 1:1 line; however, it is clear that better agreement is achieved for the lower end of the cumulative totals (i.e., $<2.00$ mm day$^{-1}$). Restricting the correlation of satellite and gauge estimates to grid cells with multiple gauge locations (five-cell correlation) reduces the underestimation problem associated with the TRMM product and decreases the scatter about the regression line with increasing rainfall estimates.

The coefficient of determination $R^2$ for the nine-cell regression is 0.175. The correlation improves dramatically when the SE cell is not included in the regression analysis. An apparent outlier, the SE cell experienced the largest difference in mean daily rainfall between satellite and gauge estimates, resulting from a likely extreme event directly over the one gauge as noted earlier. With this data point removed, the $R^2$ value improves to 0.668 ($R = 0.817$). Further improvement occurs when the analysis is restricted to mean daily rainfall estimates from grid cells containing multiple gauge stations. The $R^2$ value for the five-cell linear regression is 0.878, an increase from the previous assessment. Both absolute error (bias) and RMSE are lower for the spatially averaged, five-cell correlation. The relatively high correlation and moderate error statistics of the five-cell analysis establish the accuracy of the TRMM area-averaged (satellite) rainfall estimates versus point-source (gauge) estimates and the viability of using satellite approaches globally. Further, the results are consistent with the limited validation studies in the literature. Tian et al. (2007) evaluated the TMPA over a 3-yr period using National Oceanic and Atmospheric Administration stage-IV data (primarily radar derived) and the National Centers for Environmental Prediction (NCEP) Climate Prediction Center near-real-time daily precipitation analysis and also found correlations ranging from 0.60 to 0.75 during the summer months at scales of less than 5 days. At seasonal time scales (e.g., summer), TRMM-based estimates showed very good correlation (0.7–0.9) and low bias (from $-1$ to $1$ mm day$^{-1}$) in the Great Plains region.

In summer, the errors in rain occurrence in the satellite estimates do not necessarily correspond to errors in rain amount, suggesting that errors in rainfall intensity play a larger role. Another factor that applies to all satellite techniques is that over land the algorithms can only detect rainfall in convection with cold temperatures and ice near the cloud top. Thus, satellite products may not detect shallow “warm” rain events that gauges would discern, leading to an underestimation of precipitation accumulation.

### TABLE 4. Daily rainfall accumulation (mm) from both satellite and gauge estimates averaged over the study period and ranked across the nine grid cells.

<table>
<thead>
<tr>
<th>Grid cell</th>
<th>TRMM Rank</th>
<th>Gauge Rank</th>
<th>No. of gauges</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urban</td>
<td>1.92</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>N</td>
<td>2.01</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>NE</td>
<td>2.11</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>E</td>
<td>1.89</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>SE</td>
<td>1.59</td>
<td>9</td>
<td>3</td>
</tr>
<tr>
<td>S</td>
<td>1.59</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>SW</td>
<td>1.70</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>W</td>
<td>1.95</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>NW</td>
<td>1.85</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>Area avg</td>
<td>1.84</td>
<td>2.02</td>
<td></td>
</tr>
</tbody>
</table>

### 2) FREQUENCY ANALYSIS

The frequency distribution of percent occurrence of daily rainfall totals in the ranges 0–5 (low), 6–13 (moderate), 14–26 (heavy), and >26 (extreme) mm was produced. The results from both satellite and gauge estimates indicate that the majority of rain (approximately 90%) that falls over OKC during the warm season is in the low category in all sectors. Moderate-, heavy-, and extreme-rainfall days occur less than 10% of the time, respectively, and can occur in any region. The mean percent of occurrence for heavy and extreme events was 4.97% (5.20%) for satellite (rain gauge) estimates in all sectors. Ebert et al. (2007) noted that heavy rainfall was more difficult than light rain for satellite products to diagnose. However, the satellite product is likely more reliable for capturing spatial variability of area-averaged rainfall. Nevertheless, both methods properly capture the approximate 5% occurrence of heavy events well.
We calculated the \( z \) scores of percent occurrence (heavy events) for each sector and plotted them in Fig. 11. The \( z \) score indicates by how many standard deviations an observation is above or below the mean. Examining the satellite and gauge data, it is apparent that the only sectors that exceed the mean by more than 1 standard deviation are N, NE, and E. This result indicates that a greater percentage of heavy rainfall days are evident in the NE, N, and E sectors, consistent with their location downwind of the urban area, under a mean southwesterly flow regime, as conceptualized in Fig. 1.

In a similar way, the percentage of total warm-season rainfall over the study period in each of the four rainfall classes (not shown) revealed, as expected, that the majority of the total rainfall in OKC over the study period is received from heavy- and extreme-cumulative-rainfall days. The greatest percentage of the total rainfall over the study period comes from extreme-rainfall days under gauge estimates. Satellite estimates show the largest percentages of total rainfall distributed more or less equally between the heavy and extreme daily-rainfall classes.

3) Statistical Analysis

First, an unpaired \( t \) test at the 95% confidence interval was conducted to verify whether differences in mean daily rainfall (gauge measured) between a particular location and the regional average are significant. The results indicate that the larger mean daily rainfall that the N cell receives is statistically significant (\( p \) value = 0.09).

In a similar way, an unpaired \( t \) test at the 95% confidence interval was conducted for daily rainfall estimates retrieved by satellite. The results indicate that the larger mean daily rainfall that the NE cell receives is statistically significant (\( p \) value = 0.081). The lower mean daily rainfall receipt in the S and SE cells is also statistically significant (\( p \) values = 0.09). The differences
in mean daily rainfall of the remaining grid cells showed no statistical significance. Livezey and Chen (1983) have discussed the issues related to testing significance of multiple site data. We caution the readers that our significance values could be slightly inflated based on assumptions that we made concerning the multiple-site problem (i.e., expected means equal for each sector, and nonadjustment of the anomaly box under consideration).

b. Concentration factor analysis

1) SPATIAL VARIATIONS IN WIND DIRECTION RELATIONSHIPS AT 700 hPA

For the final component of the study, the 700-hPa level was selected to explore the relationship between prevailing wind flow and rainfall. The CF analysis (Fig. 12) displays the spatial variations in rainfall yields under each wind direction class. It is an efficient method for investigating airstream–rainfall relationships for multiple locations. Descriptive statistics of CF for all sample days with that particular wind direction over the nine-box region are shown in Table 5. If the CF of a grid cell is positive (negative), then its rainfall percentage at that location is greater (lower) than the percentage frequency of occurrence of a particular flow regime.

(i) The northerly group (N)

All nine locations under this airflow group recorded a negative concentration factor, which suggests that this is generally a dry flow regime for Oklahoma City. The SE and S cells were the two wettest locations, having the smallest negative concentration factors. It is important to note that downwind is a relative term based on 700-hPa wind direction. Therefore, the downwind location of these cells relative to the urban area is consistent with observations from previous studies (e.g., Rose et al. 2008), indicating the signature of the “urban effect” in rainfall patterns over and downwind of metropolitan areas. The N cell (upwind for this airflow group) was the driest location under this airflow group as conceptualized in Fig. 1.

(ii) The northeasterly group (NE)

All CFs were negative, the most negative being −5.37 in the N block, which is also a dry flow regime for Oklahoma City. The wettest locations under the north-easterly group were in the S and SW cells. Again, these locations would be the relative downwind areas of the urban environment under this flow regime.

(iii) The easterly group (E)

The E regimes, represented by large negative CFs, produced the least amount of rainfall at all locations (with the exception of the NW cell) relative to the flow’s frequency of occurrence (mean CF = −6.55). All stations recorded negative concentration factors (Fig. 12), ranging from −6.76 to −6.08. The NW block was the wettest location under this class (i.e., the CF was more than 2 standard deviations above the mean for this airflow group).

(iv) The southeasterly group (SE)

The SE regimes have a range of negative concentration factors similar to the E group. The airflow’s mean concentration factor (−6.36) is the second lowest of the eight wind directions. A limited range of CF values for SE and E flows illustrates the lack of spatial variability in the rainfall receipt under these flows. Having the smallest standard deviation of all flow directions (0.15), cumulative rainfall under SE flows differs only modestly between locations.

(v) The southerly group (S)

The S airstreams produced more rainfall relative to the flow’s frequency of occurrence than N, NE, E, and SE airstreams, despite a mean CF of −1.96 (dry). The positive concentration factors of the S, SW, and W cells are, however, unexpected, in light of their upwind location. The N cell (CF = −7.98), contrary to its downwind location, was actually the driest location, with the lowest negative CF recorded under any of the flow groups. The S flows produce marked spatial variations in rainfall receipt across OKC. With CF values ranging from moderately positive (CF = 2.16) to a low negative, this wind direction class has a substantially higher

![FIG. 11. The z scores for percent occurrence of rainfall on heavy and extreme days for satellite (gray bars) and gauge (black bars).](image-url)
standard deviation (Table 5) than do flows with N and E components.

(vi) The southwesterly group (SW)

The SW airstreams generate more rainfall at all locations relative to the flow's frequency of occurrence, with a mean concentration factor of 11.71. Concentration factors range from 8.82 to 16.35, making the SW group one of only two (W is the second) airflow groups for which the minimum CF is a positive value. The wettest location under this flow group is the NE cell, having a concentration factor that is 2 standard deviations above the mean. The higher rainfall accumulation of the NE cell under the SW group is consistent with its
geographical location downwind of OKC, but it is interesting that a second area of large rainfall totals appears in the upwind SW cell.

(vii) The westerly group (W)

The W flow conditions are the most likely to produce the highest daily rainfall totals; however, they are also the most varied (i.e., large standard deviation). All locations recorded positive concentration factors, ranging from 6.48 to 20.74. The high concentration factor (CF > 20) for the N cell is the largest recorded CF across all airflow groups. The urban and downwind E cells were also locations of high rainfall accumulation. The SW and S cells had the lowest concentration factors, consistent with their upwind locations.

(viii) The northwesterly group (NW)

With the exception of the W and NW cells, all CFs were negative, the most negative being −4.82 in the S cell. Although still a dry airflow group (mean CF = −1.77), NW airstreams are wetter relative to N and E flows. Higher rainfall amounts (positive CFs) were generated in the NW and W cells than expected. The driest locations were in the S and SW cells.

2) STATISTICAL ANALYSIS

These results reveal that marked spatial variations in rainfall yields are evident across the eight airflow types. A chi-square test of independence (Tidwell 1976) was conducted to assess whether these differences are real. This test was selected because chi-square is an effective method to determine whether a truly significant difference exists between a set of observed frequencies and the corresponding expected frequencies. Two null hypotheses were proposed: there is no association between wind direction and rainfall totals at a given location ($H_0 \ 1$) and there is no association between location and rainfall totals under a given wind direction class ($H_0 \ 2$).

Observed and expected rainfall totals (mm) over the 451-day sample were used. The expected values were computed by assuming that each wind direction class would generate a rainfall total that is absolutely proportional to its frequency of occurrence. CFs cannot be used here because each null hypothesis is testing the degree of departure from an expected CF of 0, resulting in the chi-square statistic being undefined. Because each calculated value of chi-square greatly exceeds its critical value, the two null hypotheses can be rejected: the differences noted in rainfall totals between wind direction classes at 700 hPa and the nine grid cells are real and are not simply a product of chance (Tables 6, 7).

3) SUMMARY

Study results revealed statistically significant spatial variations in rainfall yields across the eight airflow types. From the CF analysis at the 700-hPa level, the highest amounts of rainfall were generally observed from SW and W flows. SW and W were the most frequent wind flow regimes, representing 25% and 20% of days, respectively. These flow regimes were also associated with higher spatial variability (large standard deviation) in cumulative rainfall. In contrast, E and SE flows were the driest and least frequent, each flow regime representing less than 7% of days in the study period. The N and NE cells of the study area received the most precipitation under W and SW flows, and the lowest daily accumulation from other flow regimes. The rainfall pattern associated with the SW wind regime, the most frequent direction, generated more rainfall at all locations relative to the flow’s frequency of occurrence. With the most frequent prevailing flow (SW, W), storms may be translated to the NE, enhanced over OKC, or initiated over OKC before traversing to the NE sector of the study area where the major rainfall anomaly was identified earlier by statistical analysis of the rainfall data. Huff and Changnon (1972) found that heavier rainstorms over Illinois tend to move from the southwest. Alternatively, the S and SW cells of OKC had the lowest mean daily rainfall values. These sectors are located in a region that is seldom downwind of the urban

<table>
<thead>
<tr>
<th>Flow regime</th>
<th>Mean</th>
<th>Std dev</th>
<th>Min value</th>
<th>Max value</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northernly</td>
<td>−2.86</td>
<td>1.27</td>
<td>−4.23</td>
<td>−1.00</td>
<td>3.23</td>
</tr>
<tr>
<td>Northeasterly</td>
<td>−4.35</td>
<td>0.81</td>
<td>−5.37</td>
<td>−2.96</td>
<td>2.41</td>
</tr>
<tr>
<td>Easterly</td>
<td>−6.55</td>
<td>0.20</td>
<td>−6.76</td>
<td>−6.08</td>
<td>0.68</td>
</tr>
<tr>
<td>Southeasterly</td>
<td>−6.36</td>
<td>0.15</td>
<td>−6.60</td>
<td>−6.18</td>
<td>0.42</td>
</tr>
<tr>
<td>Southerly</td>
<td>−1.96</td>
<td>3.62</td>
<td>−7.98</td>
<td>2.16</td>
<td>10.14</td>
</tr>
<tr>
<td>Southwesterly</td>
<td>11.71</td>
<td>2.40</td>
<td>8.82</td>
<td>16.35</td>
<td>7.53</td>
</tr>
<tr>
<td>Easterly</td>
<td>12.14</td>
<td>4.53</td>
<td>6.48</td>
<td>20.74</td>
<td>14.25</td>
</tr>
<tr>
<td>Northwesterly</td>
<td>−1.77</td>
<td>2.88</td>
<td>−4.82</td>
<td>3.70</td>
<td>8.51</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Quadrant</th>
<th>Chi-square statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urban</td>
<td>389.99</td>
</tr>
<tr>
<td>N</td>
<td>551.79</td>
</tr>
<tr>
<td>NE</td>
<td>428.03</td>
</tr>
<tr>
<td>E</td>
<td>300.35</td>
</tr>
<tr>
<td>SE</td>
<td>257.43</td>
</tr>
<tr>
<td>S</td>
<td>212.21</td>
</tr>
<tr>
<td>SW</td>
<td>223.89</td>
</tr>
<tr>
<td>W</td>
<td>241.14</td>
</tr>
<tr>
<td>NW</td>
<td>240.15</td>
</tr>
</tbody>
</table>
area and are, therefore, influenced little, if any, by the urban environment. Modeling work is still needed to determine which urban factors (e.g., destabilization, roughness, aerosols) are most important for altering rainfall.

Concentration factors were summed across the nine grid cells and then ranked (Table 8) to reveal the airflow groups that yield similar CFs when spatially averaged across the OKC study area. The W and SW airstreams are the most likely to produce the highest daily precipitation (mm day$^{-1}$). The W and SW groups had the highest spatially averaged CF, but westerlies are more likely to produce higher precipitation yields than are southerlies. The S group was also drier than the NW group, despite the higher absolute moisture-holding capacity of S airstreams.

It was interesting to find that S and SW flows were often associated with an “upwind” anomaly. This is consistent with Rose et al. (2008) and D. Niyogi (2008, personal communication) who have both noted occasional upwind anomalies. Our CF analysis is consistent with expectations for an urban rainfall effect, but it is important to recall that urban-initiated or urban-enhanced rainfall will not always be exclusively a function of urban land cover on weakly forced days. Factors such as atmospheric moisture, stability, and aerosols likely play some role.

5. Conclusions

The study confirmed, using satellite and rain gauge estimates, that the N–NE area of the metropolitan OKC area was, climatologically, the wettest region on weakly forced days. Given that the western side of OKC is slightly more elevated, one might expect higher rainfall amounts west of the city. However, the results show the climatologically downwind N and NE sectors to be the anomalous regions, suggesting that precipitation modification by the urban environment is more significant than an orographic effect.

Mean daily rainfall is most consistent between satellite and gauge estimates for regions with multiple gauge locations, which provide a better representation of area-averaged precipitation for comparison with TRMM estimates. The t tests of the gauge estimates showed statistically higher ($p$ value $= 0.09$) mean daily rainfall in the N cell relative to the area average. Similar results from satellite estimates identified the NE cell as the anomalous region.

The TRMM product slightly underestimates the precipitation recorded by gauges but is well within previously published error bars. The correlation improves dramatically when the analysis is restricted to mean daily rainfall estimates from grid cells that contain multiple gauge stations (from 0.175 for the nine-cell regression to 0.878 for the five-cell regression). Satellite precipitation estimates show promising spatial correlation with the rain gauge networks. However, the results suggest that caution should be applied when using the estimates for analyses of extreme precipitation events (i.e., light rain or extremely heavy rain).

The CF analysis not only revealed clear climatological associations between wind direction and rainfall yield, but also suggested that downwind anomalies were consistent with previous hypotheses related to the urban rainfall effect. In general, the results provide support for enhancement of rainfall climatologically downwind from the urban region. Climatologically, the 700-hPa SW flows yield the greatest rainfall amounts. The W and SW 700-hPa flows produced high mean daily rainfall totals in the N and NE cells, where the urban anomaly apparently maximizes and is climatologically downwind of the metropolitan area. We also found evidence of relative downwind anomalies under less common prevailing flow regimes (e.g., northerly flow–south anomaly) consistent with Rose et al. (2008).

Future directions

The correlation analysis emphasized the limitations of a sparse rain gauge network for urban studies. An interpolation scheme such as kriging could provide a

---

**Table 7.** The 700-hPa chi-square statistics for $H_0$. The critical value of the test statistic at the 0.05 level (degrees of freedom $= 8$) is 26.12 (Ball and Buckwell 1986, p. 251).

<table>
<thead>
<tr>
<th>Wind direction class</th>
<th>Chi-square statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northerly</td>
<td>82.45</td>
</tr>
<tr>
<td>Northeasterly</td>
<td>228.91</td>
</tr>
<tr>
<td>Easterly</td>
<td>481.64</td>
</tr>
<tr>
<td>Southeasterly</td>
<td>468.23</td>
</tr>
<tr>
<td>Southerly</td>
<td>65.77</td>
</tr>
<tr>
<td>Southwesterly</td>
<td>448.80</td>
</tr>
<tr>
<td>Westerly</td>
<td>1012.85</td>
</tr>
<tr>
<td>Northwesterly</td>
<td>56.34</td>
</tr>
</tbody>
</table>

**Table 8.** Concentration factors summed and ranked across the nine grid cells for each 700-hPa wind direction class.

<table>
<thead>
<tr>
<th>Wind direction class</th>
<th>Summed CF</th>
<th>CF rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northerly</td>
<td>−25.76</td>
<td>5</td>
</tr>
<tr>
<td>Northeasterly</td>
<td>−39.14</td>
<td>6</td>
</tr>
<tr>
<td>Easterly</td>
<td>−58.99</td>
<td>8</td>
</tr>
<tr>
<td>Southeasterly</td>
<td>−57.22</td>
<td>7</td>
</tr>
<tr>
<td>Southerly</td>
<td>−17.61</td>
<td>4</td>
</tr>
<tr>
<td>Southwesterly</td>
<td>105.42</td>
<td>2</td>
</tr>
<tr>
<td>Westerly</td>
<td>109.27</td>
<td>1</td>
</tr>
<tr>
<td>Northwesterly</td>
<td>−15.97</td>
<td>3</td>
</tr>
</tbody>
</table>
better set of areal rainfall values from point values for future studies, but such approaches have well-known error as well. Implementing ground-based and TRMM precipitation radar observations could provide additional information to improve the validation of TMPA daily rainfall accumulations.

Given that OKC is a rapidly urbanizing area, the inner nine cells may not properly capture the downwind regime. Further study using a denser gauge network and a wider area would verify (or refute) the presence of any anomalous regions at a larger scale. The OKC Micronet will be completed in 2008 and will consist of 30–40 stations deployed across OKC (with a higher density near the CBD) that will measure weather variables, including precipitation. From three to five Oklahoma Mesonet stations (Brock et al. 1995) will also be strategically located within the metropolitan area.

It would also be useful to conduct a longer-term analysis that captures the preurban period of OKC. Here, the motivation would be to determine whether the “downwind” anomaly still existed during periods of less urbanized land cover. In the past, data availability has made this type of study challenging, and so atmosphere-land models are often applied.

Acknowledgments. The authors acknowledge Dr. Thomas L. Mote, Dr. Andrew Grundstein, the American Meteorological Society Industry Fellowship, the NASA PMM Program, and the Oklahoma Climate Survey, as well as the anonymous reviewers.

REFERENCES


