

Some variations of the rainfall in Mexico City from 1954 to 1988 and their statistical significance

JOSÉ LUIS BRAVO, ENRIQUE AZPRA, VÍCTOR ZARRALUQUI and CARLOS GAY
Centro de Ciencias de la Atmósfera, Universidad Nacional Autónoma de México, Circuito de la Investigación Científica s/n, Ciudad Universitaria, 04510 México, D.F.
Corresponding autor: J. L. Bravo; e-mail: jlbravo@atmosfera.unam.mx

Received October 7, 2013; accepted August 18, 2014

RESUMEN

Se estudian algunas características de la precipitación, tales como el principio y final de la estación lluviosa, así como la distribución espacial y temporal de la lluvia, usando datos de precipitación de 23 estaciones localizadas en el Distrito Federal para el periodo 1954-1988. Se definen cuatro variables relacionadas con su comportamiento: el día del primer y último evento de 10 mm de precipitación acumulados y el porcentaje de precipitación acumulada hasta los días julianos 150 y 275. Con estas variables se hace un análisis de varianza y de componentes principales, el cual permite establecer que la precipitación en el Distrito Federal (que incluye parte de la Ciudad de México y áreas circundantes) responde uniformemente en el tiempo a la presencia de fenómenos meteorológicos de mesoescala. Los resultados indican que al principio y final del periodo de estudio la precipitación fue menor que en los años intermedios; adicionalmente, la temporada de lluvias se establece de oeste a este y se retira de este a oeste, causando una temporada más larga en el suroeste que en el noreste. La distribución espacial muestra la mayor precipitación en la parte suroeste debido al efecto de montaña, y la menor en la parte noreste.

ABSTRACT

Some characteristics of the precipitation such as the beginning and end of the rainy season, as well as their spatial and temporal distributions are studied. The study included data from 23 climatological stations located in the Federal District for the period 1954-1988. Four variables related to the beginning and end of the rainy season were defined, namely the day of the first and last event with 10 mm of accumulated precipitation and the percentage of accumulated precipitation until Julian days 150 and 275. An analysis of variance and principal components was made with these variables, establishing that rainfall in the Federal District (which includes part of Mexico City and its surrounding areas) responds homogeneously through time to the presence of mesoscale meteorological phenomena. The results indicate that precipitation was lower at the beginning and end of the study period, and higher in the middle; additionally, the rainy season starts from west to east and ends from east to west, causing a longer season in the southwest than in the northeast. The spatial distribution shows the greatest rainfall in the southwest due to the effect of the nearby mountains, and the lowest in the northeast.

Keywords: Rainy season, beginning, end, spatial distribution.

1. Introduction

Precipitation is a key link in the global water cycle and a proxy for climate change; therefore, proper assessment of the impact of urban environment on precipitation (land use, aerosols and thermal properties) will be increasingly important in the ongoing climate

concerns such as diagnostics and prediction, weather forecasting, fresh water resource management and urban planning design, to name a few. These facts are particularly critical if current projections for global urban growth are accurate as mentioned by Shepherd (2005). The study is intended to improve the diagnosis

and prediction of weather in the Federal District. The city is located in the southwestern part of the Mexico Basin, with the exception of the northeast, which is surrounded by mountains reaching altitudes of up to 4000 masl (Fig. 1). The city extends in an approximate radius of 30 km from its downtown area (19.4 °N, 99.1 °W).

Mexico City is the largest urban center in the country and one of the largest in the world. Since 1950 it has undergone an enormous growth in extension, industrial establishments and therefore population size (Table I). Such expansion determines the land use change experienced by the Mexico City Metropolitan Area. Those changes cause the phenomenon known as “heat island” and associated effects (Jáuregui, 1997, 2000, 2004). According to Huff and Changnon (1973), the heat island effect has increased convective precipitation. In large urban areas, as in the Federal District, changes in the albedo of the surface may result in a warmer surface, thus a more unstable atmosphere and adequate conditions for intense storms to occur. More particles in the atmosphere acting as condensation nuclei may also result in more intense convective activity.

In recent years an interest in the behavior of rainfall in urban population centers has arisen, and the capital

Table I. Population in the Federal District and its growth rate

Census	Population	Rate
1980	8 831 079	2.4%
1990	8 235 744	−0.7%
2000	8 605 239	0.4%
2005	8 720 916	0.2%
2010	8 851 080	0.3%

Source: INEGI, 2010, 2013.

city of Mexico does not escape such interest. The purpose of this work is to estimate the variability of precipitation according to statistical analysis as well as to discuss dynamical aspects related to this variability.

In the Federal District the largest amount of rainfall occurs during spring and summer, while during the period from November to April the arrival of cold air fronts coming from northern latitudes results in scattered rainfall events, as mentioned by Jáuregui (2000), Pérez (2004), and Jáuregui and Heres (2008). The rainy season starts with light and isolated events, caused by the invasion of moisture from the Pacific and Atlantic oceans, and the arrival of cold fronts during the last part of the winter. Precipitation

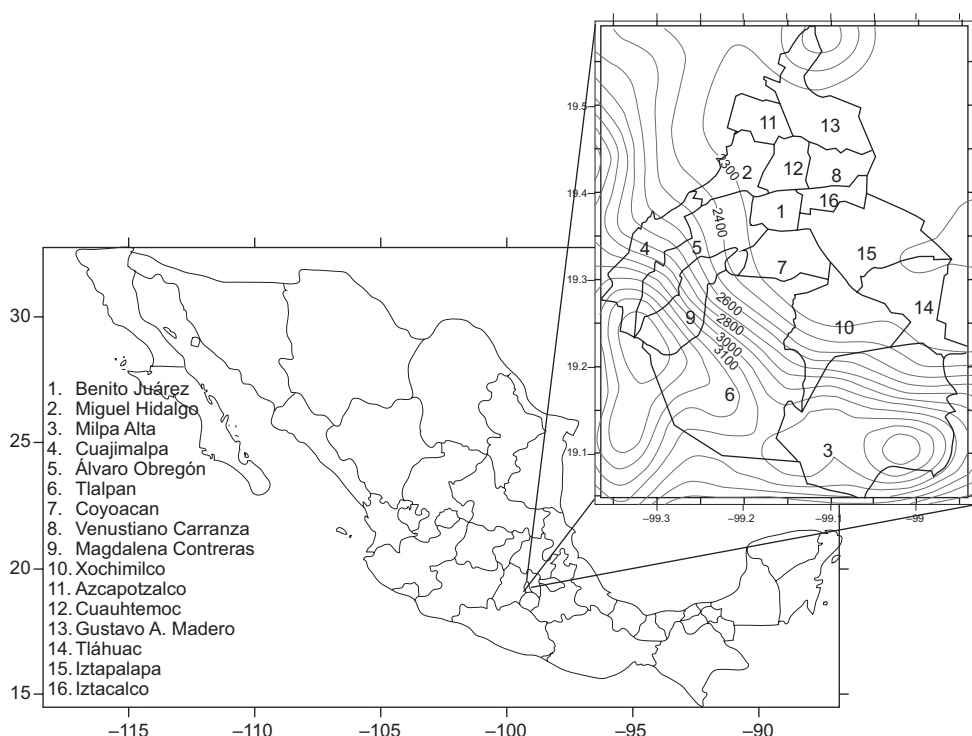


Fig. 1. Orography of the Mexico City basin and Mexico City boroughs (altitude in meters).

gradually becomes more frequent and intense until the trade winds reach the southern half of the country in the summer. These are due to the two basic patterns on the synoptic scale: dry, westerly flow from November to April with anti-cyclonic conditions, and moist easterly flows due to the trade winds for the other half of the year (de Foy *et al.*, 2005). Tropical cyclones from both oceans may produce some rain events from May 15 to November 30 (Azpra *et al.*, 2001).

2. Materials and methods

We used data from climatological stations included in the climate database CLICOM (CLImate COMputing, <http://clicom-mex.cicese.mx>) created in 1985 and managed by the Servicio Meteorológico Nacional (Mexican National Weather Service). This digital platform contains daily precipitation data, among other parameters and meteorological variables. The database contains reports from 64 stations collected between 1922 and 2010 for the Federal District, although not all stations have simultaneous observations during the entire period. Figure 2 shows the number of available stations between 1922 and 2010.

To work with a robust data set and given that loss of data for a continuous period may skew annual averages, we selected stations that had at least 355 observations per year. Most stations show information from 1954 to 1988, therefore we considered only the 23 stations with at least 16 years of observations. Those before 1954 do not present regular observations, and after 1988 several were not working. Values of annual averages of daily precipitation are

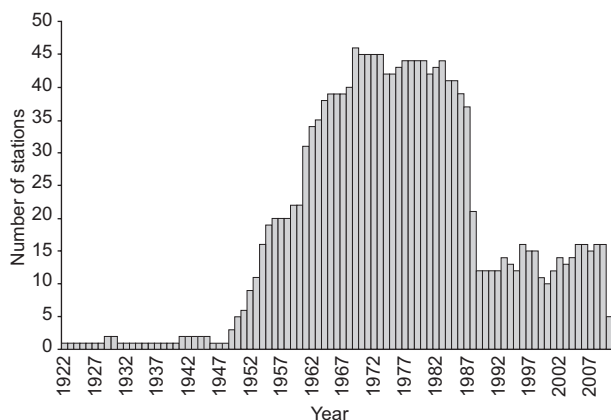


Fig. 2. Number of available stations between 1922 and 2010.

placed in a matrix of 23 columns (stations) and 35 rows (years); the density of observations is 85%. The 15% missing data are predominantly random and were substituted with the corresponding column mean, a procedure that minimizes bias in analyses without loss of information, just as a wise deletion procedure would do (Pigott, 2001). Additionally, it was required that no missing days of information were present within the first or last three months, to avoid losing the first or last event of rain.

The principal components analysis method (Morrison, 1990; Jackson, 1991) was used in order to orderly extract the uncorrelated pieces of information of a multivariate data set. We also used the Kruskal-Wallis nonparametric method (Conover, 1980) to reject the hypotheses that precipitation variations, in time and space, are due to chance only.

2.1 Principal components analysis for annual precipitation

Principal components analysis was performed with data from the aforementioned matrix. As already stated, to avoid reducing the amount of stations due to missing data, or minimizing the introduction of inexistent dependences, 15% of the missing data was completed with the average of the corresponding column (station). We calculated the eigenvalues of the covariance matrix. The results of the analysis show that the first principal component is the one that explains most of the variance (40.9%). The remaining components individually only explain a small proportion of the variance matrix, as seen in Figure 3.

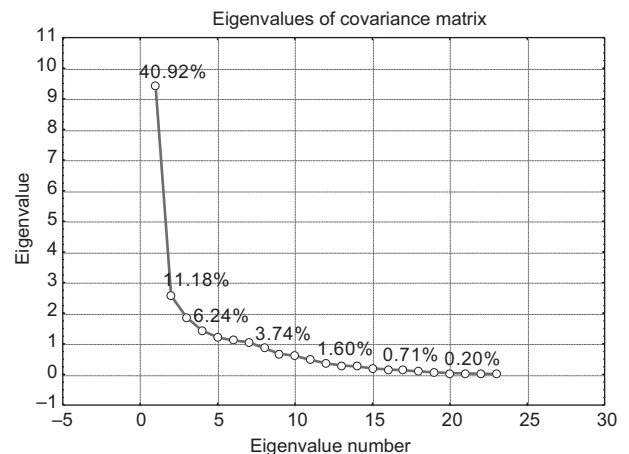


Fig. 3. Proportion of variance explained by each principal component.

The first principal component explains a substantial proportion of variance relative to the other components; this implies that there is only one principal independent piece of information in the behavior of annual precipitation in the study area. The yearly mean for stations (means for rows of the data matrix) is highly correlated with the first principal component ($r = 0.949$); indeed, the mean is identified with the first principal component, because the r value is close to 1, the difference is due to randomness. The use of the mean instead of the first component has the advantage of being a linear combination with easy physical interpretation.

2.2 Interannual variation of total precipitation

As mentioned before, means are representative values to calculate the trend of precipitation for the entire region. The principal components analysis supports this approximation. The result is shown in Figure 4.

After grouping precipitation by years and stations, a Kruskal-Wallis test (K-W) was made in order to test if variability within years is significant with respect to variability among years. The test resulted highly significant ($p < 0.001$) indicating a high probability that the variation of rainfall through time is not only random. Precipitation differences from year to year should be related to natural phenomena not to randomness alone. The precipitation showed a relative calculated maximum in 1974 equivalent to an average of 2.3 mm/day (840 mm per year) versus 1.9 mm/day

(694 mm annually) and 2.15 mm/day (785 mm per year) at the beginning and end of the sampling period respectively. The observed maximum and minimum values were 2.79 mm/day in 1976 (1019 mm/year) and 1.60 mm/day (585 mm/year) in 1957.

2.3 Spatial variation of total precipitation

Figure 5 shows the spatial distribution of the annual amount of precipitation in the Federal District. It was calculated averaging the values of total precipitation for the period 1954-1988 for each station. Ordinary Kriging interpolation was used for the graphical representation.

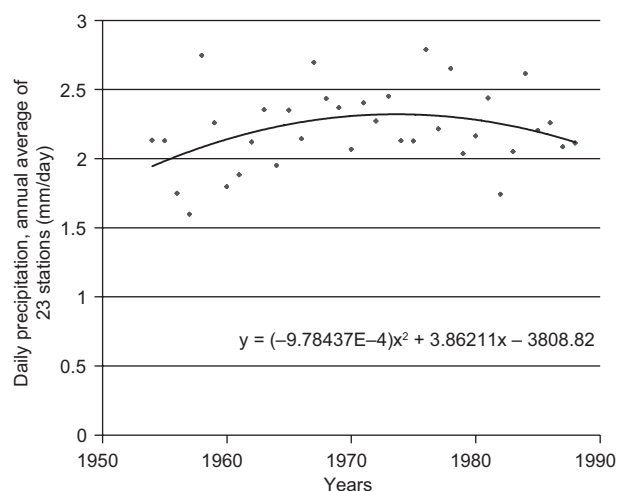


Fig. 4. Trend of daily precipitation, annual averages for 23 stations from 1954 to 1988.

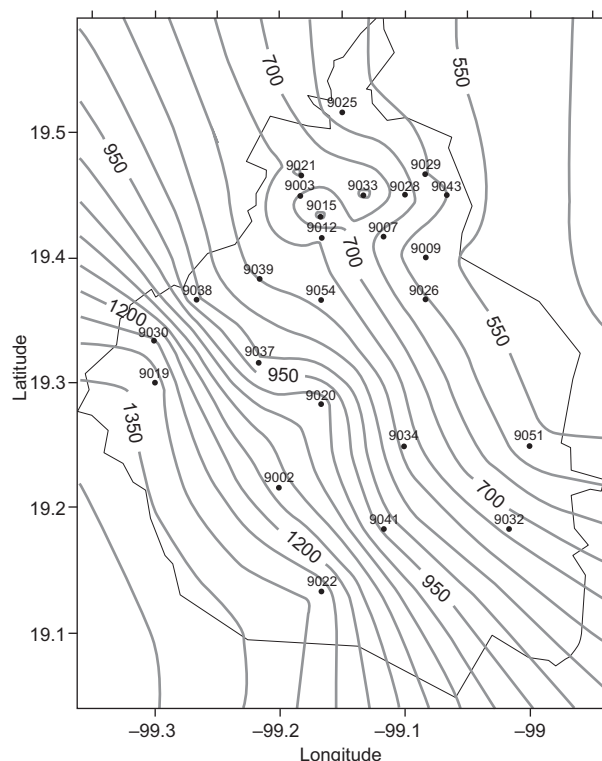


Fig. 5. Spatial distribution of the annual precipitation amount (mm) for the period 1954-1988. The 23 stations selected are displayed.

As in the case of variation with time, the K-W tests were carried out classifying the annual average of daily precipitation by stations. This test, which compares the variability within stations with the variability among stations, was also very significant ($p < 0.001$), indicating that precipitation varies significantly among stations in space and that variations are not only random.

Low precipitation is observed in the northeastern area, increasing rapidly southwestward, where it presents the highest values due to the Ajusco-Chichinauhtzin mountain range located in the southwestern Mexico City basin, which is characterized by high altitudes. This is in agreement with the results of Magaña *et al.* (2003) and Estrada *et al.* (2009). The ultimate explanation involves the prevailing winds, which blow from northeast to southwest (Jáuregui, 1988; Raga and LeMoine, 1996; Sedema, 2001), raising at the windward side of the mountains and cooling adiabatically, with the contained humidity precipitating mostly at the piedmont, as explained by Bluestein (1992). This behavior is known as orographic rain effect (Mosiño and García, 1974).

2.4 Estimation of the beginning and end of the rainy season

The concept of rainy season has many definitions, which refer to its amount (Alfaro, 1999), the period when the rain exceeds evapotranspiration (Yao, 1981), and some other characteristics. In this work, two approaches are used:

1. The Julian day on which 10 mm have accumulated in a single day or in consecutive days (Yaakov, 2004). Similarly, the last Julian day for the same amount of rain. The former is related to the presence of precipitation at the beginning of the rainy season caused by moist air contribution to the Mexico basin and the invasion of the last cold fronts of the winter season. The latter is related to the precipitation at the end of the rainy season in autumn and the return of cold fronts at the winter season, when there is little contribution of moisture from the oceans. These variables will be denoted as 10i and 10f, respectively. There was no restriction to the number of days for accumulation; however, the accumulation occurred at most in three days.

2. The second variable is obtained by standardizing the graphs of cumulative rainfall so that at the end of the year the value of total precipitation must equal 1.00. Using these graphs the accumulated proportional precipitation until Julian day 150 is calculated and called “variable 150”. The variable related with the end of the rainy season is analogously calculated as the accumulated proportional precipitation until Julian day 275 and is called variable 275. This procedure was repeated for each year and each climatic station. These days were chosen somewhat arbitrarily

because, after Julian day 150, the increase of accumulated precipitation is almost linear and stops at day 275, approximately; they correspond to May 30 (May 29) and October 2 (October 1), respectively, for normal years (leap years).

Figure 6 shows an example of a total standardized precipitation graph for the Cincel station (9007) in the Cuauhtémoc borough, between 1954 and 1988. Note the large variability, for example, on Julian day 150; precipitation in the station varies from 6 to 30%. For Julian day 275, precipitation goes from 78 to 99%. Such percentages indicate whether the “rainy season” is ahead or is delayed. Variability is a well-known characteristic of precipitation that hampers the determination of statistical significance for the observed behavior.

We considered that the rainy season started in advance when the accumulated precipitation for day 150 was above 0.20. Similarly, when the rainfall cumulative percentage for day 275 was close to 1.00, we established that the end of the rainy season occurred earlier.

The distribution of the 10i and 10f events in Figure 7 clearly indicates the behavior of precipitation in Mexico City: the beginning of the “rainy season” is denoted by the accumulation of points around day 110, whereas the end is marked by increasing dispersion around day 290. The “dry season” is manifested by scattered events before and after these dates; their existence is not continuous. This indicates that the variables 150 and 10i are suitable parameters for estimating the onset of the rainy season. The 10i events previous to Julian day 110 indicate that the amount of accumulated precipitation

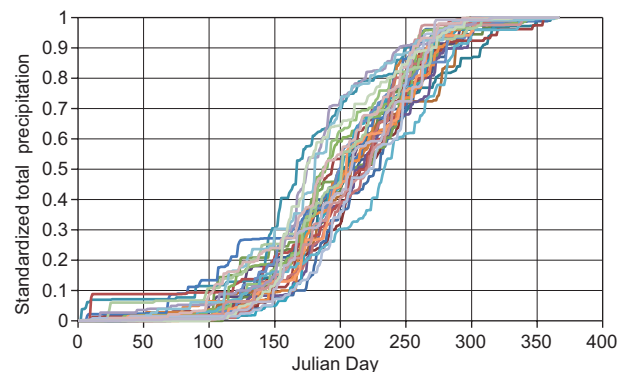


Fig. 6. Standardized accumulated total precipitation for the period 1954–1988 at the Cincel station.

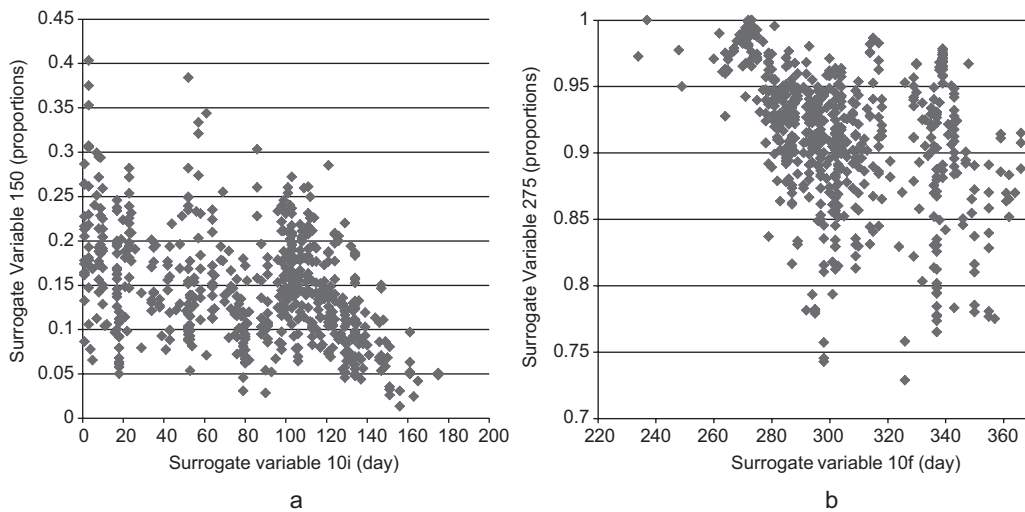


Fig. 7. (a) Variable 150 vs. 10i. (b) Variable 275 vs. 10f

has a greater variability (0.07 to 0.40); on the other hand, if the first 10i event occurs around Julian day 110 or later, the amount of accumulated rain up to Julian day 150 has been low (< 0.25).

2.5 Principal components analysis for surrogate variables

As in the case of total precipitation, having once obtained the corresponding surrogate variables, we used principal components analysis to obtain eigenvalues for the variance-covariance matrix. We found that, in the four surrogate variables, only one principal component was sufficient to explain most of the variance; the other principal components account only for a small proportion of the variance-covariance matrix. This implies that the points are distributed in an elongated (first principal component) hyper-ellipse with its principal axis in the first eigenvector direction, which suggests that the surrogate variables 10i, 150, 10f and 275 are closely correlated along time among stations; for example, if during a certain year day 10i occurs early in a climatological station, it also occurs early in other stations within the area; and if precipitation is low until day 150, it will be low for other stations. To confirm this hypothesis, the correlation matrix for the four variables was calculated. Each variable has 253 correlations for the 23 stations. The results showed 221, 244, 137 and 236 positive correlations with significance $p < 0.05$. As in the annual precipitation, a linear combination of its values is enough to explain most of the variance. As an example of the abovementioned, the proportion

of variance explained by each principal component for accumulated precipitation until Julian day 275 is shown in Figure 8.

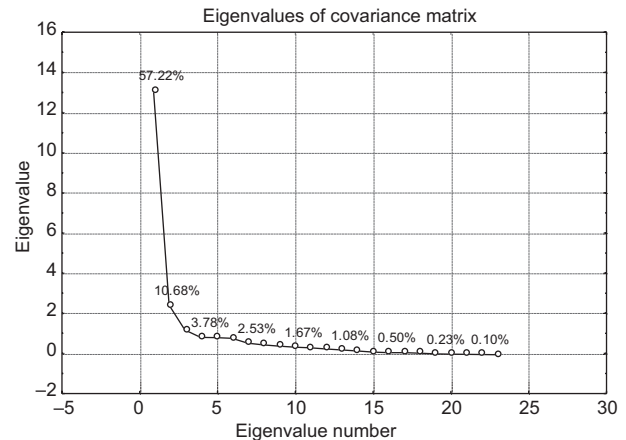


Fig. 8. Proportion of variance explained per principal component for variable 275.

The surrogate variables are strongly correlated among stations in the sample because the beginning and end of the rainy season is related to mesoscale phenomena affecting all stations within the study area. The variations for each particular station are due to the features of each station determining its own microclimate.

2.6 Temporal dependence of the surrogate variables

To justify that differences among years are significant, and therefore attempt to fit a linear trend, we

Table II. Results of the K-W tests for temporal dependence.

Temporal dependence					
Beginning	10i	K-W is significant. Temporal tendency is not significant	End	10f	K-W is significant (trend to finish earlier the rainy season) ($b = -0.507$, $p \leq 0.064$) ⁺
	150	K-W is significant. Temporal tendency is not significant		275	K-W is significant (variable 275 increases with time) ($b = 0.00160$, $p \leq 0.0274$) [*]

⁺5% significance; ^{*}10% significance.

first accomplished a non-parametric ANOVA (K-W) test. The results are shown in Table II and the significances in Appendix 1.

Based on the results of Table II, the trend of the surrogated variables averaged for stations in the Federal District during the 35 years period (1954-1988) is shown in Figures 9a, b.

Figure 9a has a negative trend for the variable 10f from 1954 to 1988, pointing to an earlier end of the season towards the end of the period. Figure 9b shows a positive trend for variable 275, which means that a greater amount of precipitation occurred before day 275. The behavior of these two variables is consistent and indicates that “the end of the rainy season” takes place earlier every year producing a shorter rainy season. For variables 10i and 150, trends are not significant. Also, variable 10i shows the greater variance among the surrogate variables, producing uncertainty in the start of the rainy season.

2.7 Spatial variation of the rainy season

Non-parametric ANOVA (K-W) tests were conducted to assess whether the differences between stations are significant when compared to the variability within stations. These tests were only significant for the variables 150 and 10f. The results are shown in Table III and Appendix 1.

Spatial distributions of variables 10i and 10f averaged for each station during 1954-1988 are displayed in Figure 10. The left panel (10i) shows that precipitation starts earlier in the southwestern part of the Federal District and later in the northeastern, the Julian day values range from 60 (in the southwest) to 95 (in the northeast). For the variable 10i, K-W test is not significant.

In Figure 10, the right panel shows that in the same zone where the rainy season starts earlier (southwest) it ends later, producing a longer length

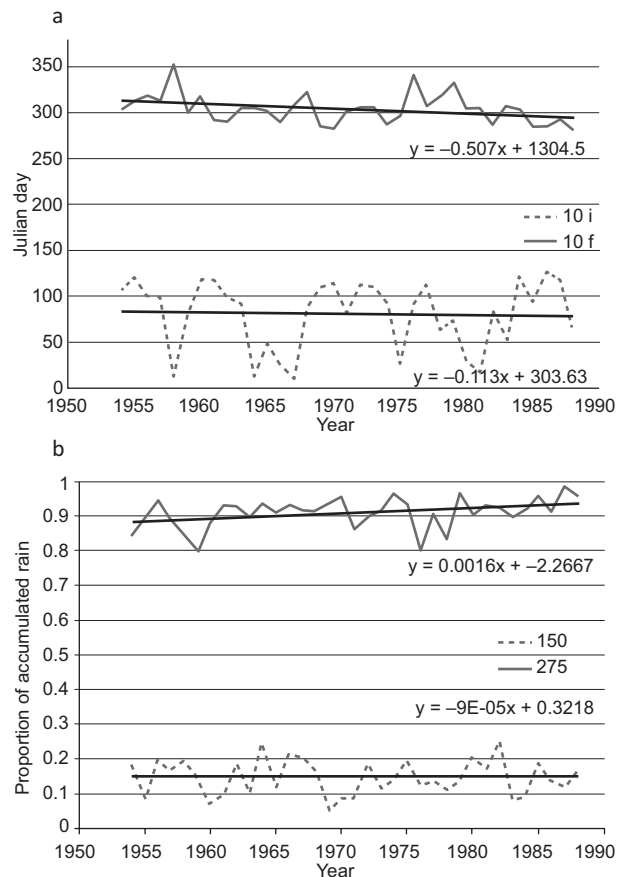


Fig. 9. Average variables between 1954 and 1988. (a) Julian day for variables 10i and 10f. (b) Proportion of accumulated precipitation for variables 150 and 275.

of the season. On the other hand, in the eastern part of the basin the season ends earlier; Julian days range between 290 and 320, and the K-W test for this variable is significant. This implies that the difference in days between the start and end of the season in the northeastern part of the Federal District is minor and therefore the extension of the season is minor too.

Cuajimalpa boroughs, and the minimum precipitation of 550 mm/year was recorded in the northeast area at the Gustavo A. Madero, Venustiano Carranza, Iztacalco, Tláhuac and Iztapalapa boroughs. The spatial variation of precipitation closely resembles the orography: low (high) altitudes correspond to low (high) precipitation. The orographic effect is enhanced by the dominant winds, which blow from northeast to southwest.

The distribution of points in Figure 7a, b indicates that rainfall events before day 110 and after day 290 are intermittent; this behavior is caused by invasions of cold air that eventually lead to precipitation. This period corresponds to the so-called dry season.

The rainy season is set with the arrival of trade winds to the southern part of the country. This meteorological system causes continuous flow of moisture, which enters the study area giving rise to the establishment of steady rains as seen in Figure 7a, b between days 110 and 290. This effect is reinforced by the moisture brought by tropical cyclones.

The variability of extended rains is larger at the beginning than at the end of the year, therefore the final period of the rainy season is sharper and clearer. The variables 10f and 275 indicate that the end of the season in the studied interval is shortened by approximately 0.5 days per year.

The beginning of the rainy season is set from west to east and the end from east to west, influenced mainly by the mountain effect. This micro-climate effects would be due to the particular relief of the basin as suggested by Estrada *et al.* (2009).

Acknowledgements

The authors would like to thank M. in G. Elda Luyando for her useful help in the preparation of the references, to Mr. Alfonso Salas and M. in G. María de Lourdes Calderón for the elaboration of some maps, and to M. Sc. Francisco J. Villicaña for the fruitful discussions we held with him. Special thanks to two anonymous reviewers for their insightful suggestions.

References

- Alfaro E. J. and L. Cid, 1999. Análisis de las anomalías en el inicio y el término de la estación lluviosa en Centroamérica y su relación con los océanos Pacífico y Atlántico tropical. *Top. Meteor. Oceanog.* **6**, 1-13.
- Azpra R. E., A. G. Carrasco, O. Delgado and C. Villicaña, 2001. *Los ciclones tropicales de México*. Instituto de Geografía, UNAM/Plaza y Valdés, Mexico, 120 pp. (Temas selectos de geografía de México I [M. E. Hernández Cerda, coord.]).
- Bluestein H. B., 1992. *Synoptic-dynamic meteorology in mid-latitudes*. Vol. I. *Principles of kinematics and dynamics*. Oxford University Press, 431 pp.
- Conover W. J., 1980. *Practical nonparametric statistics*. John Wiley & Sons, New York, 592 pp.
- Estrada F., A. Martínez-Arroyo, A. Fernández-Eguarte, E. Luyando and C. Gay, 2009. Defining climate zones in Mexico City using multivariate analysis. *Atmósfera* **22**, 175-193.
- De Foy B., E. Caetano, V. Magaña, A. Zitácuaro, B. Cárdenas, A. Retama, R. Ramos, L. T. Molina and M. J. Molina, 2005. Mexico City basin wind circulation during the MCMA-2003 field campaign. *Atmos. Chem. Phys.* **5**, 2267-228.
- Huff F. A. and S. A. Changnon Jr., 1973. Precipitation modification by major urban areas. *Bull. Amer. Met. Soc.* **54**, 1220-1231.
- INEGI, 2010. Demografía del Distrito Federal. Instituto Nacional de Estadística Geografía e Informática, Mexico. Available at: http://cuentame.inegi.org.mx/monografias/informacion/df/poblacion/dinamica.aspx?tema=me_
- INEGI, 2013. Tasa de crecimiento media anual de la población por entidad federativa, 1990 a 2010. Instituto Nacional de Estadística Geografía e Informática, Mexico. Available at: <http://www3.inegi.org.mx/sistemas/sisept/Default.aspx?t=m-demo148&s=est&c=29>.
- Sedema, 2001. Informe Climatológico Ambiental. Cuenca del Valle de México. Secretaría del Medio Ambiente, Gobierno del Distrito Federal, Mexico. Available at: http://www.sma.df.gob.mx/sma/download/archivos/informe_climatologico_ambiental_cuenca_valle_mexico.pdf.
- Jackson J. E., 1991. *A user's guide to principal components*. John Wiley and Sons, New Jersey, 569 pp. (Wiley Series in Probability and Mathematical Statistics).
- Jáuregui E., 1988. Local wind and air pollution interaction in the Mexico basin. *Atmósfera* **1**, 131-140.
- Jáuregui E., 1997. Heat island development in Mexico City. *Atmos. Environ.* **31**, 3821-3831, doi:10.1016/S1352-2310(97)00136-2.
- Jáuregui E., 2000. *El clima de la ciudad de México*. 1ª ed. Plaza y Valdés, Mexico, 129 pp.

- Jáuregui E., 2004. Impact of land-use changes on the climate of the Mexico City Region. *Investigaciones Geográficas* **55**, 46-60.
- Jáuregui E. and M. A. Heres, 2008. El clima/bioclima de un parque periurbano de la Ciudad de México. *Investigaciones Geográficas* **67**, 101-112.
- Magaña V., J. Pérez and M. Méndez, 2003. Diagnosis and prognosis of extreme precipitation events in the Mexico City Basin. *Geofis. Int.* **41**, 247-259.
- Morrison D. F., 1990. *Multivariate statistical methods*. McGraw-Hill, New York.
- Mosiño A. P. and E. García, 1974. Climates of North America. In: *World survey of climatology*, vol. 2 (R. Bryson and F. Kenneth Hare. eds.). Elsevier, pp. 345-404.
- Pérez J. B., 2004. Pronóstico numérico del tiempo para el Valle de Mexico. M.Sc. thesis. Centro de Ciencias de la Atmósfera, UNAM, Mexico.
- Pigott T. D., 2001. A review of methods for missing data. *Educ. Res. Eval.* **7**, 353-383. Available at: <http://www.stat.uchicago.edu/~eichler/stat24600/Admin/Missing-DataReview.pdf>.
- Raga G. B and L. LeMoine, 1996. On the nature of air pollution dynamics in Mexico City. Part I: Non-linear analysis. *Atmos. Environ.* **30**, 3987-3993.
- Shepherd J. M., 2005. A review of current investigations of urban-induced rainfall and recommendations for the future. *Earth Interact.* **9**, 1-27.
- Yaakov A., H. Kutiel and H. Lavee, 2004. Analysis of beginning, end, and length of the rainy season along a Mediterranean-arid climate transect for geomorphic purposes. *J. Arid Environ.* **59**, 189-204.
- Yao A., 1981. Agricultural climatology. In: *World survey of climatology*, vol. 3 (E. Landsberg, Ed.). Elsevier Scientific Publishing Company, pp. 189-298.

Appendix 1. Significances for the Kruskal-Wallis test and regressions.

	K-W grouped by years	K-W grouped by stations	Regression grouped by stations
10i	$p \leq 0.001$	$p \leq 0.34$	$p \leq 0.88$
10f	$p \leq 0.001$	$p \leq 0.04$	$p \leq 0.001$
150	$p \leq 0.001$	$p \leq 0.002$	$p \leq 0.73$
275	$p \leq 0.001$	$p \leq 0.87$	$p \leq 0.001$