

Department of Geography, Mainz University, Germany

An Analysis of Regional and Intra-annual Precipitation Variability over Iran using Multivariate Statistical Methods

M. Domroes, M. Kaviani, and D. Schaefer

With 5 Figures

Received November 24, 1997 Revised July 17, 1998

Summary

The temporal and spatial precipitation regime of Iran was analysed using multivariate analyses of monthly mean precipitation records for 71 stations. A Principal Component Analysis was applied to the correlation matrix in order to describe the intra-annual variations of precipitation. The Principal Component scores were mapped to visualize the spatial structure of the three derived precipitation regimes. By applying an agglomerative clustering (WARD) of the three Principal Component scores, five homogeneous spatial clusters, representing five precipitation regions, were developed. The intra-annual types of precipitation distribution, shown by the five clusters, are described and discussed.

1. Introduction

Because of its low annual precipitation, Iran is commonly regarded as a country of dry conditions. This is a climatic simplification, however, precipitation varies over space and time and this highlights the existence of large diversities of precipitation over Iran (Fig. 1).

Besides spatial differences, intra-annual variations of precipitation also occur (Kaviani, 1988). The complex structure of the precipitation distribution over Iran derives from the vast area of the country (about $1,648,000 \text{ km}^2$), its wide latitudinal extent $(25^{\circ}$ to 40° N) and its pronounced relief. The two highest mountain systems, Zagros and the Northern Highlands,

which rise up to 4,557 m and 5,670 m above sea level, respectively, strikingly affect the temporal and spatial patterns of precipitation (Ehlers, 1980). Zagros, in the western part of Iran, is the most elevated mountain range, extending from North-West to South-East, while the Northern Highlands (Talish and Alburz) are West-to-East-stretching along the southern Caspian Sea. Both mountain systems are located across the paths of the prevailing moisture-loaden Westerlies and Northerlies. In addition, different atmospheric circulation systems are linked.

In this paper, multivariate statistical analyses of precipitation records available for Iran are used to study regional and intra-annual precipitation variability: (1) to identify temporal and spatial types of precipitation, and (2) to define homogeneous spatial groups or precipitation regions, respectively.

2. Data

Monthly precipitation records for a 31-year observation period (1957–1987) are compiled for 71 stations from the Iranian Meteorological Organisation. The stations under study are distributed all over Iran, although their spatial density is low and uneven over some parts of the country, see Fig. 2, Table 1.

Fig. 1. Mean annual totals of precipitation (mm) over Iran, 1957±1987

Fig. 2. Location of the 71 Iranian meteorological stations under study

Table 1. Stations under Study with their Latitude, Longitude (in Decimal Degrees) and Elevation

| | No. Station | Lat., \mathcal{N} | Long., $^{\circ}E$ | Elev., m |
|----|---------------|---------------------|--------------------|----------|
| 1 | Ghaturchay | 45.25 | 38.85 | 950 |
| 2 | Moshiran | 47.52 | 38.70 | 653 |
| 3 | Khoy | 44.97 | 38.55 | 1,107 |
| 4 | Astara | 48.87 | 38.43 | -20 |
| 5 | Ardebil | 48.28 | 38.25 | 1,350 |
| 6 | Tabriz | 46.28 | 38.13 | 1,364 |
| 7 | Orumie | 45.08 | 37.53 | 1,313 |
| 8 | Osnoor | 47.90 | 37.50 | θ |
| 9 | Baranduzch | 45.22 | 37.48 | 1,300 |
| 10 | Anzali | 49.47 | 37.47 | -26 |
| 11 | Rasht | 49.60 | 37.25 | -7 |
| 12 | Gonbadekav | 55.17 | 37.25 | 150 |
| 13 | Lahijan | 50.00 | 37.18 | -2 |
| 14 | Miandoab | 46.10 | 36.97 | 1,314 |
| 15 | Ramsar | 50.67 | 36.90 | -20 |

3. Methods

To study the regional and intra-annual precipitation variability over Iran, two multivariate statistical analyses are used which are usually applied together (Leber et al., 1995; McGregor, 1993; Mills, 1995).

3.1 Principal Component Analysis

Principal Component Analysis (PCA) is a useful and widely applied tool in climatology: A large data matrix can be reduced to some important factors (Principal Components, PCs) which often can be physically interpreted. PCA can be used in climate change studies in order to detect recent changes while performing time series analyses, and also for ordinary data reduction of a large data matrix (Schaefer, 1996). After having decided which variables are to be analysed and the standardisation of the values, the correlation matrix is built. PCA is often performed with a correlation matrix. Therefore, the Kaiser-Meyer-Olkin measure of sampling adequacy (MSA) can be computed for a decision whether the correlation matrix is suitable for PCA. The MSA values in the anti-image correlation matrix can be interpreted as follows and therefore also the suitability of the variables for PCA (Backhaus et al., 1994):

 $MSA \geq 0.9$: "marvelous", MSA>0.8: "meritorious", $MSA \geq 0.7$: "middling", $MSA \geq 0.6$: "mediocre", MSA>0.5: "miserable" MSA<0.5: "unacceptable".

The number of the extracted PCs is determined by applying the Kaiser criterion. Accordingly all PCs with eigenvalues >1 are extracted. Additionally, a scree-plot can be drawn for a visual check of the number of PCs to be extracted. To improve the interpretation of the unrotated PCA results the PCs are rotated using the orthogonal VARIMAX rotation (Bortz, 1993). This rotation method is common where a Cluster Analysis is to follow PCA because the PCs are uncorrelated and the assumptions of the Cluster Analysis are satisfied (McGregor, 1993). The standardized PC scores are calculated using the regression method.

3.2 Cluster Analysis

For the spatial classification of precipitation over Iran an agglomerative hierarchical Cluster Analysis (CA) was applied, using the Ward algorithm. As the variables must not be correlated with each other, PCA is usually prefixed (Backhaus et al., 1994). In a first stage, the squared Euclidian distance is calculated, based on the extracted orthogonal PC scores from the PCA. This is a similarity function which is often combined with the Ward clustering method (McGregor, 1993).

In a second stage the clusters are formed. There are different but not fully satisfactory methods for determining the number of clusters; the decision is up to the individual. However, the definition of the number of clusters or regions can be supported by analysing the Dendrogram or the Agglomeration schedule. Additionally, a graphical test, the Elbow criterion can be applied: The coefficients from the Agglomeration schedule and the number of clusters are drawn in a line chart and can be interpreted visually (Backhaus et al., 1994). In practice, it is useful to compare the results of different solutions of the cluster membership and to find out the most appropriate solution for the respective question formulation.

The homogeneity of the derived clusters can be studied with analysis of variance and the computed $ETA²$ values. $ETA²$ is the ratio of the variance between groups to the total variance and describes the explained variance (Backhaus, 1994). The $ETA²$ values can be interpreted as the values of the explained variance (r^2) derived from the analysis of correlation: If the computed $ETA²$ values for the variables are high, the corresponding clusters can be considered homogeneous.

4. Results and Discussion

4.1 Principal Components and the Spatial Distribution of the Scores

In order to analyse the orthogonal structure of monthly mean precipitation a PCA is performed with the data matrix $(71[*]12)$: 71 cases which correspond to the number of stations, and 12 variables, corresponding to the 12 months of the

Table 2. Eigenvalues, Explained Variance and Cumulative Variance Corresponding to the Complete Solution of 12 PCs Derived from the Correlation Matrix of Monthly Precipitation Data of 71 Iranian Stations

| PC | Eigenvalue | $%$ of | % of Cumulative | | | |
|----------------|------------|----------|-----------------|--|--|--|
| | | Variance | Variance | | | |
| 1 | 9.01261 | 75.1 | 75.1 | | | |
| 2 | 1.41066 | 11.8 | 86.9 | | | |
| 3 | 1.08053 | 9.0 | 95.9 | | | |
| $\overline{4}$ | 0.15879 | 1.3 | 97.2 | | | |
| 5 | 0.14418 | 1.2 | 98.4 | | | |
| 6 | 0.06400 | 0.5 | 98.9 | | | |
| 7 | 0.04679 | 0.4 | 99.3 | | | |
| 8 | 0.03394 | 0.3 | 99.6 | | | |
| 9 | 0.02856 | 0.2 | 99.8 | | | |
| 10 | 0.01090 | 0.1 | 99.9 | | | |
| 11 | 0.00597 | 0.0 | 100.0 | | | |
| 12 | 0.00308 | 0.0 | 100.0 | | | |

year. After standardisation and calculation of the correlation matrix the Kaiser-Meyer-Olkin measure of sampling adequacy (MSA) is computed showing MSA-values > 0.8 except for April and May (0.71 and 0.79, respectively). The correlation matrix can, however, be applied for a PCA. The initial statistics are then computed (see Table 2).

According to the Kaiser criterion and the Scree plot (not shown) three PCs are extracted which explain 75.1% of the total variance for the first, 11.8% for the second and 9.0% for the third PC, respectively. Thus, the first three PCs account for 95.9% of the total variance. The communalities of all the variables in the analysis are > 0.9 .

The PC loadings must be analysed in order to interpret the PCA results. High loadings (>0.7) indicate good correlations between the variables and the PCs. After the VARIMAX rotation the first three PCs show high loadings (>0.7) for June-November, December-March and April-May, respectively (Table 3). For March, November and December, relatively high PC loadings can be observed for several of the PCs, showing a transitional character for these months which occur between the seasons. Thus the PC loadings of the monthly mean precipitation totals reproduce three main types of intra-annual variation of precipitation over Iran:

 First Principal Component (PC1): Summer precipitation type (June-November)

Table 3. VARIMAX Rotated PC Loadings Derived from the Correlation Matrix of Monthly Precipitation Data of 71 Iranian Stations

| | PC ₁ | PC ₂ | PC ₃ |
|-----------|-----------------|-----------------|-----------------|
| January | .30666 | .92161 | .12359 |
| February | .40520 | .82507 | .31957 |
| March | .37493 | .70368 | .55164 |
| April | .03107 | .35722 | .91176 |
| May | .43889 | .05557 | .87727 |
| June | .90100 | .15369 | .34380 |
| July | .89262 | .35583 | .18998 |
| August | .89102 | .41276 | .13201 |
| September | .88455 | .39457 | .13893 |
| October | .87100 | .42933 | .18852 |
| November | .74889 | .57402 | .24309 |
| December | .54330 | .79344 | .16712 |

- Second Principal Component (PC2): Winter precipitation type (December-March)
- Third Principal Component (PC3): Spring precipitation type (April-May)

The computed PC scores are mapped to identify the spatial distribution of the PCs. Kriging is used to create a regularly spaced grid $(0.5^*0.5^{\circ})$ based on the irregularly spaced stations in the study area (Burrough and McDonnell, 1998). Thereafter, isolines are drawn to demonstrate the spatial structure of the computed PCs. Due to standardisation, the mean of the PC scores is 0, and the standard deviation 1.

PC1 which explains 44.9% of the total variance (after the VARIMAX rotation) indicates a summer peak of precipitation and loads heavily from June to November. Over large parts of Iran the PC1-score values are slightly negative (about -1), while, on the other hand, positive PC1-scores occur over smaller parts (Fig. 3a). High positive PC1-scores, expressing large totals of precipitation in summer, are clearly established over northern parts of Iran. The highest (3.81) PC1-score value observed at Anzali describes the largest total of summer precipitation showing a standard deviation about four times higher than the mean summer precipitation over Iran. The southwest corner of the Caspian Sea experiences the highest values due to the occurrence of (northeastern) local sea breezes that are moisture-loaden from the Caspian Sea. The PC1-scores have negative values (about -1)

Fig. 3a, b, c. Spatial distribution of VARIMAX rotated PCscores (PC1: 3a, PC2: 3b, PC3: 3c)

over most of the western and southern parts of Iran indicating low summer precipitation totals. The lowest total is observed at Emamgheis (-1.54) .

PC2 which accounts for 31.5% of the total variance (after the VARIMAX rotation), loads heavily from December to March. Two regions in Iran experience high winter precipitation: Zagros Highlands and the southwest corner of the Caspian Sea (Fig. 3b). During this season two main reasons for precipitation are worth mentioning: (1) Depressions, with their origin over the Mediterranean Basin leading to precipitation, especially at Zagros. (2) The high-pressure centre over Asia that leads to easterlies and northeasterlies. Consequently the southeastern part of Caspian Sea receives its highest precipitation in winter. The highest PC2-score is experienced at Anzali (3.01) indicating the highest winter precipitation while the lowest is at Khoy (-1.89) . Apart from the above mentioned two areas with high positive values, the PC2-scores are negative (about -1) in all other parts of Iran, indicating low winter totals of precipitation.

PC3 which explains 19.5% of the total variance (after the VARIMAX rotation), loads heavily from April to May. High PC3 scores are observed over the northern and northwestern parts of Iran (Fig. 3c). The cyclonic activity that occurs during winter is reduced in spring. Furthermore, the strong High over Asia in winter tends to weaken. Consequently, precipitation along the Caspian Sea coast decreases (compared to winter). On the other hand, temperatures rise rapidly in spring, and convectional rainfall may frequently occur because of intense heating of the surface. The highest PC3-score (2.42) , reflecting the highest spring precipitation, occurs at Saghez, and the lowest (-1.64) at Busher. All other regions experience low spring precipitation. The difference between the highest and lowest PC3-scores is small indicating a comparably low spatial variability for spring precipitation.

4.2 Classification of Homogeneous Precipitation Regions

On the basis of a hierarchical clustering of the three PC scores from all 71 stations in the study, different solutions of the cluster membership were analysed. The coefficient in the Agglomeration Schedule increases strongly from a four to a five cluster solution which indicates

Fig. 4. The five Iranian precipitation regions deduced after the WARD clustering method applied to the VARIMAX rotated PC-scores

remarkable differences between the clusters. As a result, and by interpretation of the Dendrogram and the Elbow criterion (not shown), five clusters or homogeneous precipitation regions are identi fied, respectively (Fig. 4). This solution seems to be the most appropriate result. Descriptive statistics (mean, standard deviation) are calculated for all clusters in order to describe the intraannual variation of precipitation (Table 4, Fig. 5). In addition the $ETA²$ values derived from the analysis of variance are given in Table 4. The $ETA²$ values for the variables vary between 0.69 and 0.92. Therefore the explained variances vary between 69 and 92% underlining the homogeneity of the separate regions (Table 4). The mean

values for All-Iran were also computed to allow a comparison with the cluster means.

Cluster 1, comprising 13 stations, is located in the northwestern part of Iran and the eastern section on the shore of the southern Caspian Sea. This area experiences a mean annual precipitation total of 379 mm with a relatively high standard deviation of 163 mm. Under the strong influence of the mid-latitude Westerlies, a great number of depressions lead to high winter precipitation, though its maximum (about 50 mm) occurs in spring (April and May). While cyclonic activity declines during spring, depressions migrating to the Northeast are blocked by high pressure centres over the Caspian Sea resulting in high precipitation totals in the area of Cluster 1. In addition, a great atmospheric instability derives from the rise of temperature over the whole of Iran producing heavy convectional rainfall which is sometimes even associated with hailstorms.

Cluster 2, comprising seven stations, is developed along the Caspian Sea shores of Iran. This cluster experiences the highest mean annual precipitation of 1,331 mm. Precipitation totals are comparably high showing a relatively homogeneous distribution throughout the year. The lowest totals occur in June and July (with 49 and 32 mm, respectively), and the highest in September, October and November, with 165, 250 and 183 mm, respectively. As the main causes of heavy precipitation in the Caspian Sea area in September, October and November both the formation and establishment of the Asian High

Table 4. Mean Precipitation Values, Standard Deviation (Std dev) and $ETA²$ for the five Clusters (of Fig. 4) and the All-Iran Mean; Values given in mm

| | | Jan | Feb | Mar | Apr | May | Jun | Jul | Aug | Sep | Oct | Nov | Dec | Year |
|------------------|---------|------|------|------|------|------|----------|-----------------------------|----------|----------|----------------|------------|------|-------|
| Cluster 1 | Mean | 35 | 33 | 48 | 52 | 47 | 20 | 7 | 9 | 15 | 38 | 40 | 33 | 379 |
| | Std dev | 18 | 18 | 19 | 10 | 9 | 11 | 8 | 14 | 20 | 26 | 18 | 21 | 163 |
| Cluster 2 | Mean | 118 | 106 | 101 | 55 | 50 | 49 | 32 | 78 | 165 | 250 | 183 | 143 | 1,331 |
| | Std dev | 33 | 22 | 19 | 19 | 14 | 14 | 7 | 18 | 53 | 61 | 56 | 42 | 285 |
| Cluster 3 | Mean | 64 | 56 | 63 | 60 | 31 | 2 | | | | 18 | 39 | 53 | 388 |
| | Std dev | 33 | 22 | 17 | 16 | 13 | 3 | | | | 8 | 16 | 18 | 106 |
| Cluster 4 | Mean | 23 | 19 | 22 | 21 | 10 | | Ω | θ | | 5 | 10 | 16 | 130 |
| | Std dev | 8 | 7 | 9 | 9 | 6 | | θ | | | $\overline{4}$ | 6 | 8 | 51 |
| Cluster 5 | Mean | 67 | 41 | 38 | 21 | 6 | Ω | | | Ω | 5 | 23 | 46 | 249 |
| | Std dev | 21 | 12 | 16 | 9 | 3 | | $\mathcal{D}_{\mathcal{L}}$ | | Ω | 3 | 12 | 17 | 72 |
| ETA ² | | 0.69 | 0.71 | 0.71 | 0.69 | 0.79 | 0.85 | 0.85 | 0.89 | 0.89 | 0.92 | 0.85 | 0.77 | 0.88 |
| All-Iran | Mean | 53 | 43 | 48 | 39 | 25 | 10 | 5 | 10 | 19 | 38 | 42 | 46 | 379 |
| | Std dev | 34 | 29 | 28 | 21 | 19 | 16 | 10 | 24 | 52 | 75 | 53 | 40 | 357 |

Fig. 5. Annual variation of precipitation for the five clusters (of Fig. 4) and the All-Iran mean

and a thermal depression over the Caspian Sea must be considered. The pressure gradients lead to Easterlies and Northeasterlies that become moisture-loaden when crossing the Caspian Sea. The peak of precipitation is, therefore, observed in the southwestern part of the coastal stretches of the Caspian Sea. Another reason for the precipitation maximum from September to November is the northern track of very strong cyclonic activity, which is far greater than the activity over the southern track, also needs to be considered.

Cluster 3, comprising 17 stations, is established over the western part of Zagros and the north-eastern part of Iran. The mean annual precipitation total accounts for 388 mm which is comparable to Cluster 1. However, this cluster shows fairly homogeneous precipitation, expressed by lower standard deviations and a different intra-annual variation of precipitation. This cluster is characterized by most precipitation occurring in winter and originating from Mediterranean depressions that reach Iran far from the west. Further to the East, cyclonic activity and precipitation totals decrease.

Cluster 4, being the biggest and comprised of 20 stations, covers the whole central Desert Basin of Iran, including the southern part of Albors and an area of Zagros. Mean annual precipitation is very low (130 mm), due to the well-developed Lee-effect in the east of Zagros, causing decreasing precipitation further eastwards. Like the other areas, this cluster receives precipitation mostly in winter and spring, and hardly any precipitation can be observed from June to September.

Cluster 5, comprising 14 stations, is located in the southern and southwestern part of Iran, near the Arabian Gulf. This region is strongly affected by the southern track of Mediterranean cyclones in winter, leading to frequent precipitation and reasonably high mean annual totals of 249 mm which are distinctly higher than those in Cluster 4. Against the peak of precipitation in winter, there is hardly any precipitation experienced from June to September.

5. Conclusion

In the first part of the paper, temporal and spatial precipitation regimes over Iran were analysed by applying PCA to the monthly mean precipitation records from 71 stations over a 31 observation period (1957–1987). Three PCs were extracted, explaining 95.6% of the total variance. The VARIMAX rotated PCs show for Iran three types of intra-annual variation of precipitation with different main precipitation seasons. As a result, most of Iran experiences dry conditions during summer (PC1) when only the Caspian Sea area records relatively high precipitation. Mostly orographic precipitation can be observed due to local sea breezes that carry moisture from the Caspian Sea. In summer, the southwest corner is the wettest area of the Caspian Sea.

In general, winter represents the main rainy season over Iran (PC2). Most parts receive more than half of the annual precipitation totals during winter. The highest totals can be observed over the Zagros Highlands and in the southwest corner of the Caspian Sea. Winter precipitation originates from Mediterranean depressions.

During spring (PC3), most parts of Iran experience dry conditions. Only the northern parts record high precipitation totals, with the highest values in the northern and record northwestern parts which are mostly due to convective rainstorm activity.

In the second part of the paper a CA was applied to the PC-scores: Five homogeneous clusters were derived, establishing regions with well pronounced intra-annual precipitation variability. On the one hand, these clusters are well separated from each other according to their intra-annual precipitation variations, yet on the other hand, clear spatial connections can be observed. The stations of Cluster 1 are located in the northwestern part of Iran and in the southeastern part of the coast of the Caspian Sea. The annual precipitation total is comparatively high (379 mm) and its intra-annual variations relatively even. The southwestern part of the Caspian Sea (Cluster 2) experiences highest precipitation totals in Iran throughout the year, resulting in the mean annual total of 1,331 mm. Cluster 3 records only slightly higher totals than Cluster 1, but the intra-annual variation varies. While the summer is very dry, winter is the main rainy season. Depressions originating over the Mediterranean basin lead to substantial precipitation totals, indicating a dominantly cyclonic origin for precipitation. Cluster 4 occupies the central Desert Basin of Iran, recording a mean annual total of precipitation of only 130 mm and reflecting the driest cluster of Iran. This region is located on the leeward sides of the two main mountain systems. Precipitation mostly occurs during winter. The same pattern occurs for the intra-annual variation for Cluster 5, which occupies the southern and southwestern parts of Iran. Precipitation during winter contributes most of the annual total (249 mm).

Acknowledgements

This paper results from a joint research project of all three authors which was carried out at the Department of Geography, Mainz University, Germany, during a Research Fellowship under the German Academic Exchange Service (DAAD) granted to Prof. Dr. Kaviani at Mainz University in July and August 1997.

References

- Backhaus, K., et al., 1994: Multivariate Analysemethoden. Berlin: Springer, 594 pp.
- Beyraghdar, G., 1994: The climate of Iran and shortage of water resources. Journal of Iran I.M.O., Scientific and Technical Journal, 22, 38-46 (in Farsi).
- Bortz, J., 1993: Statistik für Sozialwissenschaftler. Berlin: Springer, 753 pp.
- Burrough, P. A., McDonnell, R. A., 1998: Principles of Geographical Information System. Oxford: Oxford University Press, 333 pp.
- Domroes, M., Peng, G., 1988: The Climate of China. Berlin: Springer, 361 pp.
- Domroes, M., Ranatunge, E., 1992: The orthogonal structure of Monsoon rainfall variation over Sri Lanka. Theor. Appl. Climatol., 46, 109-114.
- Ehlers, E., 1980: Iran. Grundzüge einer geographischen Landeskunde. Wissenschaftliche Länderkunden, Bd. 18. Darmstadt: Wissenschaftliche Buchgesellschaft, 596 pp.
- Fisher, W. B., (ed.), 1968: The Cambridge History of Iran. Vol. 1. The Land of Iran. Cambridge: Cambridge University Press, 784 pp.
- Kaviani, M., 1988: Statistical analysis on precipitation regime in Iran. Roshd, Journal of Geography, $13, 4-12$ (in Farsi).
- Kaviani, M., 1995: Climatic research of precipitation regime in Province Isfahan. Isfahan University Research Bulletin, 6, 55-64 (in Farsi).
- Leber, D. et al., 1995: Climate classification of the Tibet autonomous region using multivariate statistical methods. Geo Journal, 37.4, 451-473.
- McGregor, G. R., 1993: A multivariate approach to the evaluation of climatic regions and climatic resources of China. Geoforum, 24, 357-380.
- Mills, G. F., 1995: Principal component analysis of precipitation and rainfall regionalisation in Spain. Theor. Appl. Climatol., 50, 169-183.
- Schaefer, D., 1996: Uni- und multivariate statistische Untersuchungen zu rezenten Klimaänderungen in Sri Lanka. Ph.D. thesis, Department of Geography, Mainz University, Germany, 226 pp.

Authors' address: M. Domroes, M. Kaviani and D. Schaefer, Department of Geography, Mainz University, D-55099 Mainz, Germany.