Summer monsoon rainfall patterns over South Korea and associated circulation features

B.-J. Kim, R. H. Kripalani, J.-H. Oh, and S.-E. Moon

With 6 Figures

Received October 30, 2000
Revised November 12, 2001

Summary
Seasonal summer monsoon (June through August) rainfall patterns over South Korea are classified by an objective method using data for a 40-year period (1961–2000). The rainfall patterns are represented by the percentage departures from the normal rainfall of 12 stations spread uniformly over South Korea. The statistical technique employed is the k-means (KM) clustering method. The Euclidean distance has been used as a measure of similarity between the patterns.

Four dominant types are obtained by this method. Inter-correlations among the types suggest that the dominant patterns are distinct. The summer monsoon rainfall shows an increasing trend. Investigation of the physical processes associated with these patterns using NCEP/NCAR Reanalysis data clearly reveals contrasting circulation features associated with the dominant types during the summer monsoon period. In particular, contrasting circulation features are related to the position, shape and strength of the North Pacific Subtropical High.

1. Introduction
The most significant weather phenomenon during the summer season (June through August: JJA) over the Korean peninsula is the quasi-stationary front extending from south China to southern Japan. This front is called 'Changma' in Korea, 'Mei-yu' in China and 'Baiu' in Japan. This front is located along the northern and northwest periphery of the subtropical anticyclone over the North Pacific. The North Pacific Subtropical High greatly influences the climate over East Asia. The low-level jet at the northwestern edge of this anticyclone transports a large amount of water vapor into East Asia (Tao and Chen, 1987; Kurihara, 1989; Ding, 1994). At the southern edge of this High is the western Pacific warm pool. The influence of the anomalous heating over the warm pool on the atmospheric circulation and precipitation over East Asia is also well documented (e.g. Huang and Sun, 1992).

The joint spatio-temporal behaviour of the rainfall fields can be characterized by pattern recognition methods. The recurring rainfall patterns can be identified by these methods. These dominant pattern types can represent both the climate and the climate variability. The commonly used classification methods for the meteorological fields are:

(i) map-to-map (MM) correlation method (Lund, 1963)
(ii) k-means (KM) clustering method (Kruizinga, 1979)

Though empirical orthogonal functions (EOFs) (Kutzbach, 1967) are not designed specifically to find clusters, some studies have used EOFs for cluster analysis (e.g. Kulkarni et al., 1992). Some work on the delineation of the precipitation
patterns over South Korea using principal component analysis has been done previously (e.g., Seo and Joung, 1982; Ho and Kang, 1988). Classification of the rainfall patterns over South Korea using the above three procedures yielded similar dominant patterns. However the results of the KM-method appear reliable especially for the synoptic climatology/composite techniques.

In view of the above, there are two problems of interest here. What are the spatial patterns of rainfall over South Korea? Are there specific large-scale circulation features, which are associated with these patterns? Hence in the present study the principal features of the spatial distribution of the seasonal (JJA) percentage departures from normal rainfall patterns as realized during the 40-year period (1961–2000) are examined by applying a statistical technique and investigating the lower-tropospheric circulation features associated with these patterns using NCEP/NCAR Reanalysis data.

2. Data

(i) Summer monsoon rainfall over South Korea accounts for 50–60% of the annual rainfall, hence season (JJA) rainfall for 12 stations uniformly spread over South Korea for the 40-year period (1961–2000) has been used. Figure 1 shows the position of the Korean peninsula in the Asian domain and the location of the 12 stations. Figure 2 depicts the normal rainfall pattern and the pattern of variability (standard deviation). These patterns are similar to the patterns shown in earlier studies (e.g. Ho and Kang, 1988). In this study the spatial pattern for each year is represented by the percentage departure from the normal rainfall of each station.

(ii) A time series of Korean monsoon rainfall (KMR) for the period 1961–2000 has been prepared by simple arithmetic average of the 12 stations. The most dominant EOF shows loadings of the same sign over the entire country (not shown), justifying the average of the 12 stations as representative of KMR.

(iii) Gridded 2.5° by 2.5° latitude/longitude monthly geopotential heights for the lower tropospheric 850 hPa level have been extracted from the NCEP/NCAR (National Centers for Environmental Prediction/ National Center for Atmospheric Research) Reanalyses data set (Kalnay et al., 1996) for the period 1961–1998.

3. The KM-method

In this method a class is represented by the mean of all patterns belonging to that class.

(i) Establish the number \( k \) of classes. Choose, at random, \( k \) patterns from the complete data set and assign one to each class. At this stage the mean pattern of every class is equal to its respective chosen pattern.

(ii) Assign each pattern of the complete data set to the class whose mean is nearest according to the distance measure \( \delta_{(i,j)} \) given by

\[
\delta_{(i,j)} = \sum_{i,j=1}^{m} (X_i - X_j)^2
\]

This is the usual Euclidean distance between two \( m \)-dimensional vectors.

(iii) Compute new mean patterns for each class.

(iv) Repeat (ii) and note whether there are any patterns that change class. If so, then repeat (iii) and (iv). The above-described procedure converges quickly.

Instead of the Euclidean distance, the Mahalanobis distance \( \Delta^2 \) can also be used as a measure of similarity between patterns. Mahalanobis (1930) defined the squared distance between two multivariate normal populations as

\[
\Delta^2 = (\mu^{(1)} - \mu^{(2)})^T \Sigma^{-1} (\mu^{(1)} - \mu^{(2)})
\]

where \( \mu^{(1)} \) and \( \mu^{(2)} \) are the mean vectors and \( \Sigma \) is the variance-covariance matrix.

Using the univariate analogy for \( \Delta^2 \), the distance square between two univariate populations can be defined as (Anderson, 1984)

\[
\delta_{ij} = (X_i - X_j)^2 / \sigma, \quad i, j = 1, 2, 3 \ldots k; \quad i \neq j
\]

where \( X_k \) is the mean of the \( k \)th pattern

\[
\sigma = \text{standard deviation of the difference} (X_i - X_j)
\]

Thus

\[
X_k = \frac{1}{m} \sum_{j=1}^{m} X_{kj} \quad k = 1, 2, \ldots n
\]

and

\[
\sigma = \sqrt{\text{var}(X_i) + \text{var}(X_j) - 2 \text{cov}(X_i, X_j)}
\]