

# Spatial and temporal rainfall patterns in Central Dry Zone, Myanmar - A hydrological cross-scale analysis

Mya Thandar Toe<sup>1,\*</sup>, Mamoru Kanzaki<sup>1</sup>, Tsung-Hsun Lien<sup>2</sup>, and Ke-Sheng Cheng<sup>2,3</sup>

<sup>1</sup>Graduate School of Agriculture, Kyoto University, Kyoto, Japan

<sup>2</sup>Department of Bioenvironmental Systems Engineering, National Taiwan University, Taipei City, Taiwan

<sup>3</sup>Master Program in Statistics, National Taiwan University, Taipei City, Taiwan

## Article history:

Received 15 December 2015

Revised 7 February 2016

Accepted 15 February 2016

## Keywords:

Climatological monsoon break, Orographic effect, Semivariogram, K-means cluster analysis, Principal component analysis

## Citation:

Toe, M. T., M. Kanzaki, T.-H. Lien, and K.-S. Cheng, 2017: Spatial and temporal rainfall patterns in Central Dry Zone, Myanmar - A hydrological cross-scale analysis. *Terr. Atmos. Ocean. Sci.*, 28, 425-436, doi: 10.3319/TAO.2016.02.15.01(Hy)

## ABSTRACT

Myanmar (Burma) is traditionally an agriculture-based country. However, irrigation is not available in most of its agricultural lands. This study focuses on the Central Dry Zone (CDZ), which is the driest part of Myanmar. Exploratory data analysis, semivariogram analysis and modeling, K-means cluster analysis and principal component analysis were conducted in this study to investigate the general CDZ climatology. The spatial and temporal rainfall variation patterns of different scales, including daily, event-scale and monthly rainfall are studied. A climatological monsoon break divides the wet season into two peaks. The monsoon break is the result of different climate dynamics – May-to-June period monsoon southwesterly and August-to-October period tropical cyclone vorticity. Rainfall stations in different clusters identified by the K-means cluster analysis reflect the orographic effect and different climate dynamics, which influence the spatial and temporal rainfall variation patterns in the CDZ. Principal component analysis results for average monthly rainfall reveals that the first principal component mainly explains the spatial variabilities in average monthly rainfall in the CDZ. The second principal component explains the seasonal (temporal) variation in average monthly rainfall. It was found that during the wet season, spatial rainfall variations in the CDZ are more significant than the seasonal (temporal) rainfall variation. Understanding the spatial and temporal variability in CDZ rainfall can provide valuable information on potential water availability in both the time and spatial domains, which will then enable making sound cropland planning and forest management decisions.

## 1. INTRODUCTION

Myanmar (Burma), located in continental Southeast Asia, is traditionally an agricultural country. The agro-sector remains the major contributor to Myanmar's economy. In general, agricultural practices across the country depend heavily on the amount and timing of the summer monsoon precipitation (Roy and Roy 2011). Only 17% of the total agricultural land in Myanmar is currently under irrigation (Central Statistical Organization 2008). Compared to its neighboring countries, there are fewer publications on the climatology of Myanmar. Roy and Kaur (2000) studied the monthly precipitation of Myanmar, particularly the monsoon months (June to September), and delineated several

distinctive homogenous precipitation zones across the country based on the annual precipitation in monsoon months. Of those zones, the Central Dry Zone (CDZ) is known to receive the lowest rainfall in the country, with a mean annual rainfall of less than 680 mm. The CDZ region-wide average annual rainfall accounts for 3.2% of the country-wide average annual rainfall, has no more than 50 rainy days annually and has a precipitation (P) to potential evapotranspiration (PET) ratio from 0.5 - 0.65 (Cho 1999). As a result, the CDZ is classified as an arid zone (Roy and Kaur 2000), but with a semi-arid to dry-subhumid climate as defined by the rainfall index (Gratzfeld 2003) and aridity index (P/PET) by UNEP (1992).

The CDZ is home to 34% of the total population of Myanmar (The Union of Myanmar 2005). It contains vast

\* Corresponding author  
E-mail: myathandartoe@gmail.com

areas of agricultural land due to its relatively flat elevation. The agricultural lands are cultivated mainly with drought-tolerant crops. However, the crop yields are usually constrained by the limited and irregular nature of the CDZ rainfall (Poe 2011). Although the Ayeyarwady (Irrawaddy) River, the longest river in Myanmar, flows through the CDZ, irrigation from the river has not been extended to most of the agricultural lands in the region. As a result, spatial and temporal variations in CDZ rainfall are crucial to the livelihoods of CDZ inhabitants. In recent decades increasing expansion of agricultural settlements has occurred in the CDZ. While recognizing the need for agricultural expansion, it is also vitally important to implement sound cropland planning and forest management to avoid the potential adverse effects of deforestation and crop-season water shortages. Understanding the spatial and temporal variability in CDZ rainfall can provide valuable information on potential water availability in both the time and spatial domains, which will then enable making sound cropland planning and forest management decisions.

Climate zone delineation is often conducted by considering the mean annual precipitation, mean annual temperature and highest or lowest monthly precipitation and temperature (Kottek et al. 2006; Peel et al. 2007). However, for cropland planning and forest management purposes, the amounts and spatial and temporal variation characteristics of seasonal and monthly rainfalls are of major concern. In addition, while agricultural irrigation is not available in most of the CDZ areas, the development of potential irrigation projects requires hydrological analysis involving daily to event-scale rainfalls. Although the CDZ as a whole is considered a homogenous precipitation zone based on the amounts of annual wet season precipitation (Roy and Kaur 2000), it has been realized that the spatial and tem-

poral rainfall patterns of different scales (including daily, events, monthly and seasonal rainfall) need to be considered in making decisions regarding choosing suitable crops and managing forest and crop areas within the CDZ. Therefore, this study aims to investigate the spatial and temporal rainfall variation patterns at different scales in the CDZ of Myanmar. Findings from this study will provide useful information and data for making sound cropland planning and forest management decisions.

## 2. STUDY AREA AND RAINFALL DATA

The CDZ is located in the central part of Myanmar and encompasses an approximate area of 54390 km<sup>2</sup>. The area is shielded by high mountains along its west and east boundaries, with the Irrawaddy River flowing from north to south (see Fig. 1). Although there are a total of 54 rainfall stations in the CDZ, most rainfall stations have very short recording lengths for daily rainfall. Hourly rainfall measurements are not available at most rainfall stations. Thus, fifteen years (2000 - 2014) of daily rainfall datasets available at rainfall stations listed in Table 1 were used in this study. Each station represents an administrative unit referred to locally as a township. The datasets collected were supported by the Dry Zone Greening Department, Myanmar, which has a mandate to green the entire CDZ. The data were originally collected mainly from the township meteorological stations operated by the Department of Meteorology and Hydrology as well as from agro-meteorological stations operated by the Department of Agriculture and Irrigation in townships without meteorological stations (e.g., Thayet Township). The elevations (above the mean sea level) of the rainfall stations vary from -10 m at Station 50 to 283 m at Station 30. Generally

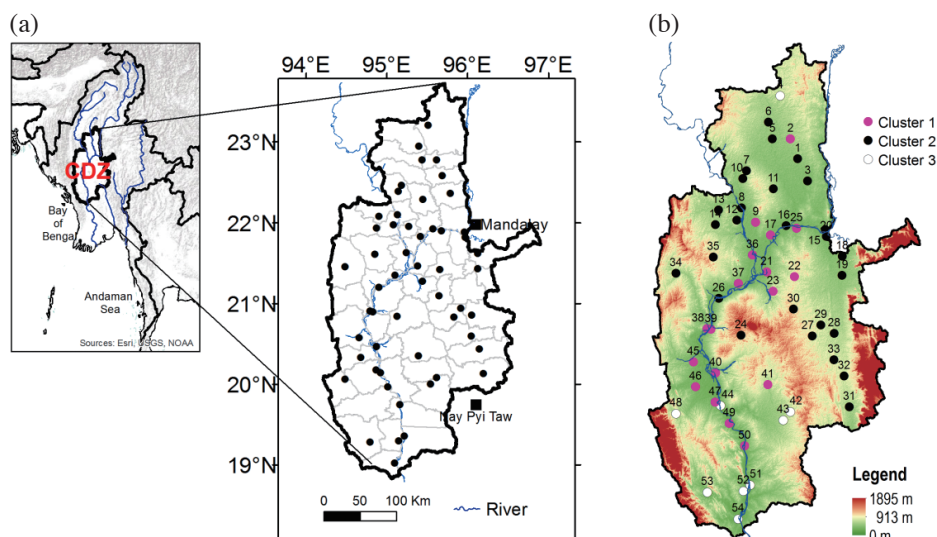


Fig. 1. (a) Location map of the Central Dry Zone (CDZ). (b) Topographic map of the CDZ. Dots of different colors represent rainfall stations in the CDZ. Rainfall stations in different clusters are based on the K-means cluster analysis results described in the section “Delineating unique climate zones in the CDZ”. (Color online only)

Table 1. Rainfall stations in the CDZ.

Name	ID	Latitude	Longitude	Elev (m-msl)
Shwebo	1	19.36	95.22	109
Khin-U	2	22.29	95.45	113
Wetlet	3	22.39	95.15	81
Kanbalu	4	20.90	94.83	177
Ye-U	5	21.96	95.27	89
Teze	6	22.47	95.18	117
Debayin	7	19.02	95.10	119
Monywa	8	23.21	95.51	44
Chaung-U	9	22.78	95.62	84
Butalin	10	20.84	95.13	118
Ayadaw	11	21.62	96.13	152
Salingyi	12	20.14	94.92	93
Yinmabin	13	21.10	95.65	119
Pale	14	20.83	95.83	135
Sagaing	15	20.18	94.87	66
Myinmu	16	19.29	94.79	89
Myaung	17	19.97	95.01	64
Kyaukse	18	22.10	95.13	82
Myittha	19	21.61	94.85	87
Tada-U	20	21.83	95.42	78
Myingyan	21	21.47	95.38	53
Natogyi	22	21.93	95.58	228
Taungtha	23	21.43	96.13	125
Kyaukpadaung	24	20.08	95.62	632
Nganzun	25	20.35	95.39	143
Nyaung-U	26	21.42	95.66	68
Meiktila	27	21.90	95.68	242
Thazi	28	20.06	94.48	176
Wundwin	29	21.20	94.91	204
Mahlaing	30	21.35	95.10	283
Tatkon	31	21.94	94.87	159
Yamethin	32	21.45	94.48	207
Pyawbwe	33	20.33	94.68	196
Pauk	34	20.60	96.05	156
Myaing	35	21.89	95.96	264
Yesagyo	36	20.58	94.66	65
Pakokku	37	21.98	95.09	67
Seikphyu	38	20.91	94.79	58
Chauk	39	22.59	95.69	67
Yenangyaung	40	19.75	95.16	88
Natmauk	41	21.82	95.97	162
Myothit	42	20.13	96.20	207
Taungdwingyi	43	20.00	95.55	140
Magway	44	21.28	95.45	21
Salin	45	22.95	95.40	74
Pwinbyu	46	19.30	95.15	50
Minbu	47	20.86	96.05	46
Ngaphe	48	22.37	95.79	58
Minhla	49	20.94	95.92	59
Sinbaungwe	50	20.44	96.15	-10
Aunglan	51	20.47	94.88	36
Thayat	52	21.63	95.24	74
Mindon	53	22.78	95.44	69
Kamma	54	22.08	94.90	28

speaking, the CDZ has a very distinctive dry pattern (from November to April of the next year) and wet (May to October) seasons, with the wet season accounting for approximately 97% of the annual total rainfall. In particular, from December to March there is very little rain at all stations. Thus, in this study we focus on the spatial and temporal variation in rainfalls in the wet season only. Daily rainfall at all stations is further partitioned into individual storm events using a minimum inter-event-time of one-day period and event-total rainfall at individual stations was collected. The mean and standard deviation of the event duration are 2.17 and 1.83 days, respectively. The mean and standard deviation of event-total rainfall are 27.70 and 34.59 mm, respectively.

The mean and standard deviations of 15 years of daily rainfalls at individual stations were calculated. Station Khin-U was chosen as the reference station since the mean daily rainfall for the May-July period (the MJJ period) and the August-October period (the ASO period) at Station Khin-U were closest to the mean daily rainfalls for all stations in the study area. The following criteria were adopted to extract daily rainfalls at all stations for subsequent daily rainfall semivariogram analysis:

- (1) Identifying days in which the daily rainfall at Station Khin-U falls within the  $(0.25 \pm 0.075)$ ,  $(0.5 \pm 0.075)$ , and  $(0.75 \pm 0.075)$  range of cumulative probabilities. Daily rainfall within these ranges represents low, median and high rainfall intensities, respectively.
- (2) Selecting days for which more than half of the rainfall stations recorded larger-than-zero rainfall on the same day.
- (3) Collecting daily rainfall occurred in the above selected days for all rainfall stations.

The empirical cumulative distribution function of daily rainfall for the MJJ period and the ASO period at Station Khin-U are shown in Fig. 2. The daily low, median and high intensity rainfall corresponds to (3.50 - 5.08 mm), (7.11 - 11.42 mm), and (14.48 - 24.50 mm) ranges for the MJJ period and (3.42 - 6.60 mm), (9.07 - 13.02 mm), and (17.77 - 32.27 mm) ranges for the ASO period, respectively.

### 3. METHODOLOGY

An exploratory data analysis was conducted on rainfall observations at individual rainfall stations to provide a general description of the CDZ climatology. In order to investigate the spatial and temporal rainfall patterns at various scales, three major techniques were employed in this study. Semivariogram analysis and modeling was conducted to investigate the spatial variation structures across various rainfall scales, including daily, event-scale and monthly rainfalls. The K-means cluster analysis was used to group individual rainfall stations into a suitable number of groups, using site-specific average monthly rainfalls in the wet season (May to

October). Principal component analysis of site-specific average monthly rainfalls was then conducted to summarize the spatial and temporal variations in average monthly rainfalls in the CDZ and to characterize the major attributes of individual components.

#### 4. CLIMATOLOGY OF THE CENTRAL DRY ZONE

The seasonal variation in average monthly rainfalls in the CDZ is shown in Fig. 3. The CDZ has a very distinctive pattern of dry and wet seasons, with nearly all annual rainfall amounts occurring in the wet season. In particular, approximately 60% of the annual rainfall occurs in the August to October period. In the wet season there seems to be a decreasing monthly rainfall trend from May to July and an increasing monthly rainfall trend from July to September. The monthly rainfall in July is significantly lower than that of other months in the wet season. This suggests that a shift in rain patterns may occur in July and rainfall in the MJJ period and the ASO period may have different characteristics.

A similar monthly rainfall pattern has also been observed in Thailand (Takahashi and Yasunari 2006, 2008). Takahashi and Yasunari (2006) coined the term “*climatological monsoon break*” (CMB) to describe a transition period that changes rainfall systems from the monsoon southwesterlies to tropical depressions over Indochina. The summer monsoon over the Indochina Peninsula occurs from May to October. Takahashi and Yasunari (2006) found that the CMB divides the monsoon season over Thailand into two peaks and surrounding large-scale atmospheric circulation features before and after the CMB also differ. During the first rainy peak (in late May) the rainfalls are mainly induced by the monsoon southwesterlies, whereas the second rainy peak rainfalls (in mid-September) are induced by tropical cyclonic vorticity. Since Myanmar is a neighboring country of Thailand and the CMB patterns observed in the CDZ in Myanmar and in Thailand are nearly identical, it is believed that MJJ period and ASO period rainfalls in the CDZ are caused by different climate dynamics, i.e., the monsoon southwesterlies and tropical cyclones, respectively. In the

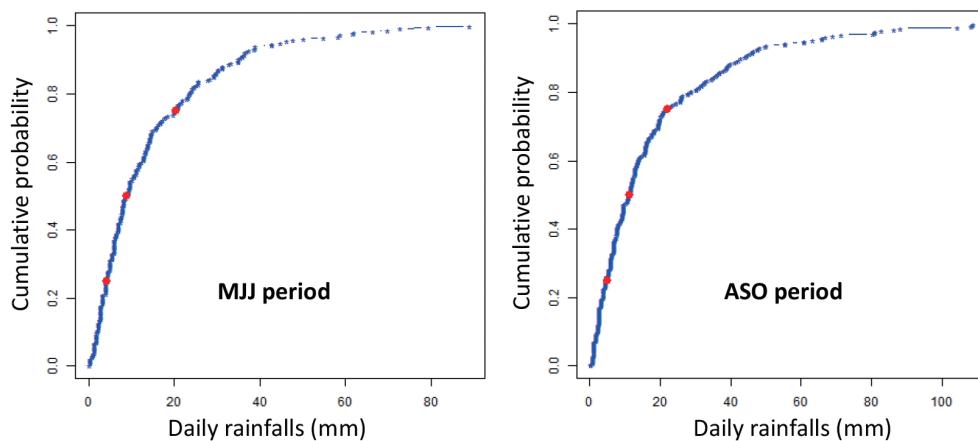


Fig. 2. Empirical cumulative distribution functions of daily rainfall for the May-July (MJJ) period and the August-October (ASO) period at Station Khin-U based on 15 years (2000 - 2014) of daily rainfall. Red dots correspond to daily rainfall in the 0.25, 0.5, and 0.75 cumulative probabilities. (Color online only)

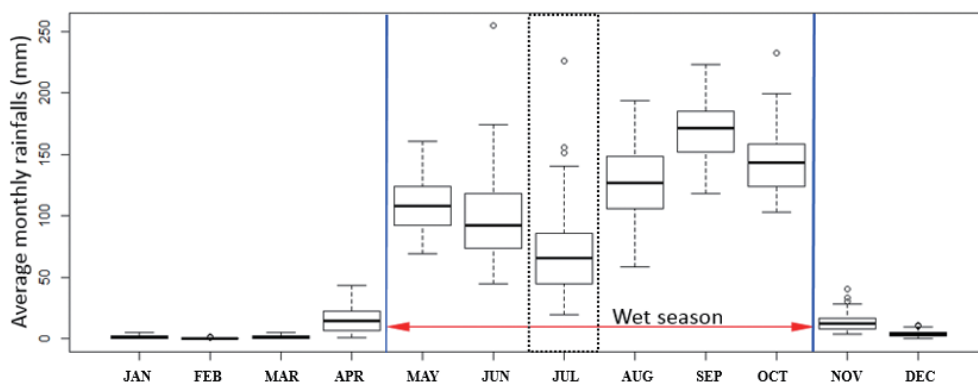


Fig. 3. Seasonal variation in average monthly rainfall in the Central Dry Zone of Myanmar. Boxplots are based on average monthly rainfall for 54 rainfall stations. The outliers (marked by open circles) are determined using 1.5 multiples of the inter-quartile range. (Color online only)

following section, we further investigate the effects of different climate dynamics on the spatial variation characteristics of rainfalls in the MJJ and ASO periods.

**5. SPATIAL VARIABILITIES OF RAINFALLS OF DIFFERENT SCALES**

Eighty-six days of daily rainfall were collected based on the criteria described in the study area and rainfall data section. Among these data for 29, 26, and 31 days of low, median and high rainfall intensities were collected, respectively. The spatial variability in rainfalls of different scales, including daily, event-scale and monthly rainfalls were analyzed through semivariogram modeling.

The semivariogram is widely used in many hydrological, environmental and meteorological studies to characterize the spatial variability in physical quantities under investigation (Journel and Huijbregts 1978; Hughes and Lettenmaier 1981; Bastin et al. 1984; Lebel et al. 1987; Cheng et al. 2003, 2008). Let  $\{Z(x_i), i = 1, 2, \dots, n\}$  represent the values of a variable of interest (rainfall in our study) at different locations  $x_i, i = 1, 2, \dots, n$ . Spatial distribution or variation in the variables is considered a random field and the semivariogram of a stationary random field is defined as

$$\gamma(|x_i - x_j|) = \frac{1}{2} Var[Z(x_i) - Z(x_j)] = \frac{1}{2} E[Z(x_i) - Z(x_j)]^2 \quad (1)$$

where  $\gamma(h)$  is the semivariogram and  $E(Z)$  and  $Var(Z)$  respectively represent the expected value and variance in the random variable  $Z$ . The semivariogram of a stationary random field has three important parameters: the *influence range*, *sill*, and the *nugget*. The characteristics of these semivariogram parameters can be summarized as follows (Cheng et al. 2008):

The *influence range* is the minimum distance  $h = |x_i - x_j|$  beyond which the two random variables  $Z(x_i)$  and  $Z(x_j)$  become independent. Small values in the influence range imply high degree of spatial variability for the random field under investigation.

As the distance between  $x_i$  and  $x_j$  increases, the semivariogram will reach an asymptotic *sill*, which is numerically the same as the variance in the random variable  $Z(x)$ .

Theoretically, the semivariogram should pass through the origin, however, due to measurement errors or variations in  $Z(x)$  at a very small spatial scale, an abrupt jump in the semivariogram near the origin may be observed. This is called the *nugget effect* and the departure of  $\gamma(h)$  from the origin is the *nugget* value.

Storm events of various seasons and intensities may have different spatial variation structures; therefore, it may not always be realistic to adopt a unique semivariogram for all storm events, irrespective of the rainfall intensities and

seasonal and meteorological conditions (Cheng et al. 2003). Generally speaking, daily rainfalls of higher intensities tend to have higher degrees of spatial variabilities. Therefore, daily or event rainfalls of higher intensities correspond to higher sill values and a set of semivariograms with different sill values will result. Bastin et al. (1984) suggested a *climatological semivariogram* model of the following form:

$$\gamma(m, h) = \alpha(m) \cdot \gamma^*(h) \quad (2)$$

where  $m$  is an index for individual storm events,  $h$  is the Euclidian distance and  $\alpha(m)$  is the scaling factor. With this structure, all the time (or event) nonstationarity is concentrated on the scaling factor  $\alpha(m)$ , and the scaled component  $\gamma^*(h)$  is time (or event) invariant and is called the *dimensionless semivariogram*. The scaling factor in Eq. (2) is equivalent to the sill or the variance in the rainfall field and the dimensionless semivariogram characterizes the influence range and small-scale variabilities (Cheng et al. 2008). In this study, an exponential semivariogram model of the following form is adopted to fit the experimental dimensionless semivariogram for daily and event rainfalls:

$$\gamma(h) = C_0 + \omega \left[ 1 - \exp\left(-\frac{h}{a}\right) \right] \quad (3)$$

In the above equation,  $C_0$  is the nugget value and  $(C_0 + \omega)$  represents the sill which should be close to 1.0, and the influence range is approximately  $3a$ .

**5.1 Semivariogram of Daily Rainfalls**

Dimensionless experimental semivariograms of daily rainfalls of the MJJ and ASO periods were fitted to the following exponential models (see Fig. 4):

$$\gamma(h) = 0.54 + 0.47 \left[ 1 - \exp\left(-\frac{h}{32}\right) \right] \quad (\text{the MJJ period}) \quad (4)$$

$$\gamma(h) = 0.54 + 0.50 \left[ 1 - \exp\left(-\frac{h}{42}\right) \right] \quad (\text{the ASO period}) \quad (5)$$

where distance  $h$  is in km units. Both semivariograms have a nugget value of 0.54 which indicates that localized spatial variations (approximately within 10 km range) in daily rainfalls cannot be monitored using the current network of rainfall stations in the study area. Both models have sill values  $(C_0 + \omega)$  very close to 1.0, as can be anticipated for dimensionless semivariograms. In addition, the influence range of daily rainfalls of the MJJ period (96 km) is significantly smaller than the influence range of daily rainfalls of the ASO period (126 km). Such results are consistent with

the findings of Takahashi and Yasunari (2006) – daily rainfalls in the ASO period are influenced by large scale tropical cyclones. In contrast, during the MJJ period, there are almost no tropical cyclones in the region and daily rainfalls are mostly affected by monsoon southwesterlies from the Indochina Peninsula.

## 5.2 Semivariogram of Event-Total Rainfalls

Dimensionless experimental semivariograms of event-total rainfalls of the MJJ and ASO periods were fitted to the following exponential models (see Fig. 5):

$$\gamma(h) = 0.45 + 0.49 \left[ 1 - \exp\left(-\frac{h}{35}\right) \right] \quad (\text{the MJJ period}) \quad (6)$$

$$\gamma(h) = 0.62 + 0.39 \left[ 1 - \exp\left(-\frac{h}{48}\right) \right] \quad (\text{the ASO period}) \quad (7)$$

Characteristics of the dimensionless semivariograms for the daily and event-total rainfalls are similar, except that event-total rainfalls are associated with larger influence ranges (105 km for the MJJ period and 144 km for the ASO period) due to their longer integration time (i.e., durations of storm events which are almost all less than 10 days).

## 5.3 Semivariogram of Monthly Rainfalls

In order to investigate the spatial variabilities in monthly rainfalls, the monthly rainfalls at individual rainfall stations were calculated for each year. Since rainfalls for the MJJ period and the ASO period were induced by different climate dynamics, semivariogram modeling of monthly rainfalls were conducted by considering site-specific monthly rainfalls for the MJJ and the ASO periods, respectively. Unlike semivariograms for rainfalls of daily and event scales, semivariograms for monthly rainfalls of the MJJ and the ASO periods were fitted to the following linear models (see Fig. 6):

$$\gamma(h) = 0.6529 + 0.0021h \quad (\text{the MJJ period}) \quad (8)$$

$$\gamma(h) = 0.4259 + 0.0035h \quad (\text{the ASO period}) \quad (9)$$

As the distance  $h$  increases, the fitted linear semivariograms also increase and do not reach a sill. The power and linear semivariogram models indicate the existence of a long-range spatial correlation or the random field under investigation has not been observed at a sufficiently large enough distance for the semivariogram to reach its sill. A semivariogram may also increase without reaching a sill if a spatial variation trend is present. Monthly rainfalls are the integration of daily rainfalls over a one-month period and thus have significantly higher spatial correlation and larger influence ranges than daily or event-scale rainfalls. Monthly rainfalls in the CDZ also exhibit a clear spatial variation trend, as demonstrated in Fig. 7. Thus, the larger influence range and presence of a trend in spatial variation for the monthly-scale rainfalls jointly contribute to the linear semivariogram modeling of monthly rainfalls in the CDZ.

## 6. DELINEATING UNIQUE CLIMATE ZONES IN THE CDZ

Even though the CDZ is broadly classified as having an arid or semi-arid climate, cropland planning and forest management in the CDZ requires more detailed climate zone delineation by considering the amounts and spatial and temporal variation characteristics of seasonal and monthly rainfalls. Therefore, K-means cluster analysis was conducted using the average monthly rainfalls for individual months in the wet season as classification features. A six-value vector of the average monthly rainfalls (from May to October) at a particular station was considered as a sample for K-means cluster analysis, with a total of 54 samples. The classification features were normalized with respect to the means and standard deviations of individual features (i.e., average monthly rainfalls of individual months) before implementing the K-means cluster analysis in order to eliminate the

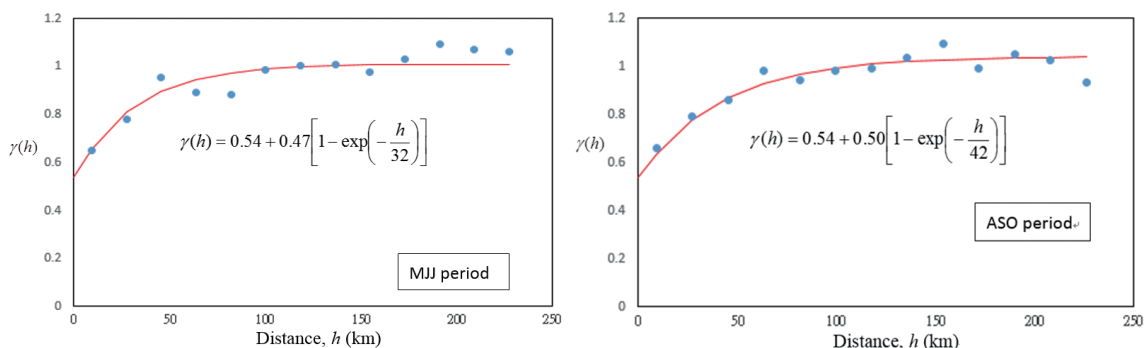


Fig. 4. Dimensionless semivariograms for daily rainfalls in the May-July (MJJ) period and August-October (ASO) period. (Color online only)

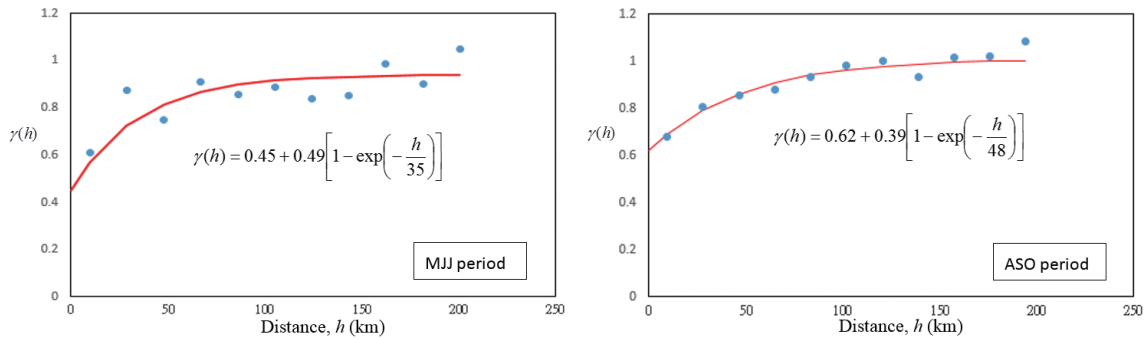


Fig. 5. Dimensionless semivariograms for event-total rainfalls for the May-July (MJJ) period and August-October (ASO) period. (Color online only)

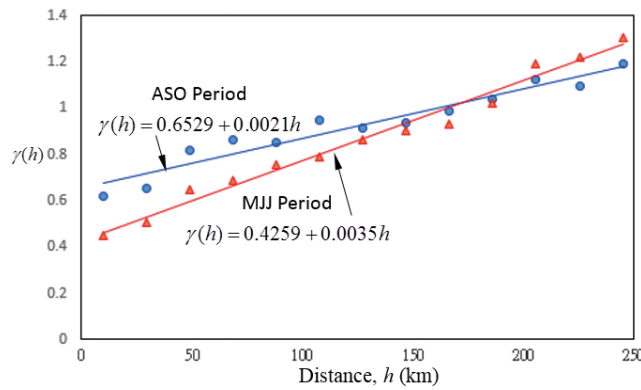


Fig. 6. Dimensionless semivariograms for monthly rainfalls for the May-July (MJJ) period and August-October (ASO) period. (Color online only)

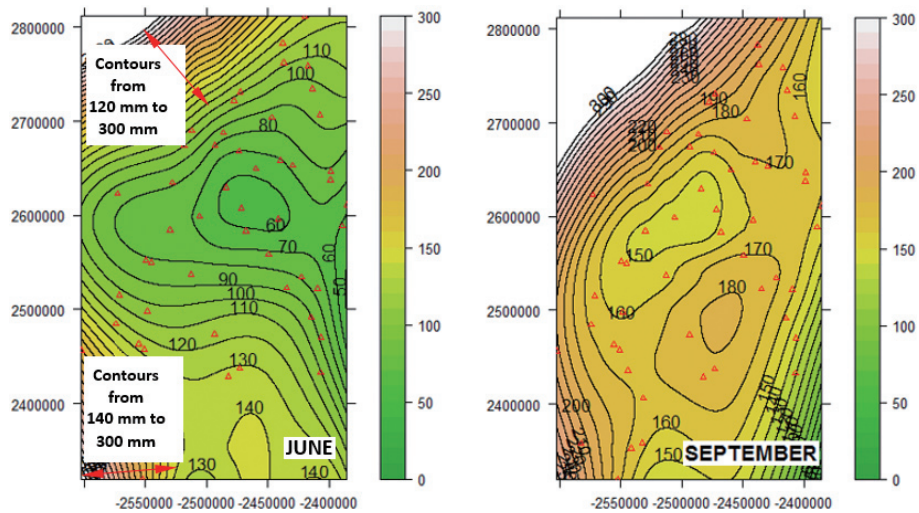


Fig. 7. Contours of average monthly rainfalls (in mm) for June and September in the CDZ. A spatial variation trend can be observed in either plot. (Color online only)

effect of magnitude differences in classification features.

K-means cluster analysis implementation requires that the number of clusters to extract be specified in advance. Hartigan (1975) suggested the following Hartigan indicator for choosing the optimal number of clusters:

$$H = \left( \frac{\text{WithinSS}(k)}{\text{WithinSS}(k+1)} - 1 \right) \cdot (n - k - 1) \quad (10)$$

In the above equation,  $\text{WithinSS}(k)$  and  $\text{WithinSS}(k+1)$  are the within-cluster sum of squares for the  $k$  and  $k+1$  cluster cases, respectively.  $n$  is the number of samples (data points) for cluster analysis. If  $H$  is greater than 10, it is then justifiable to add the extra group. Figure 8 illustrates changes in the Hartigan indicator when the number of clusters increases. Using the average monthly rainfall for May to October as the classification features, the Hartigan indicator suggests 3 clusters for the K-means cluster analysis. The rainfall stations assigned to individual clusters by the K-means cluster analysis are shown in different colors in Fig. 1. Among the 54 rainfall stations, 18, 27, and 9 stations were grouped into Clusters 1, 2, and 3, respectively. Rainfall stations in Cluster 1 are mostly located along the Irrawaddy River, while stations in Clusters 2 and 3 are roughly scattered in the central to north and southern CDZ regions, respectively. There is also a ridge in the lower central CDZ, which stretches in the southeast toward the North West direction and roughly separates the stations in Clusters 2 and 3. Histograms of the average monthly rainfalls for June, July, and September (months of peak average monthly rainfalls and the monsoon break occurrences) of individual clusters are shown in Fig. 9. An overall comparison of the average monthly rainfalls for different clusters reveals that the rainfall stations in Cluster 1 generally receive the least amount of rainfall among the three clusters, possibly due to the orographic effect. Rainfall stations in Cluster 3 are mostly located at the south end of the CDZ and are more influenced by monsoon southwesterlies and the tropical cyclonic vorticity from the South China Sea, resulting in having higher average monthly rainfall than the stations in other clusters. Figure 10 shows the spatial and temporal variation in average monthly rainfalls. In addition to the higher average monthly rainfall in Cluster 3 and lowest rainfall in Cluster 1, differences in the spatial variation patterns of the MJJ and ASO periods can also be observed. Compared to the ASO period, the MJJ period has higher spatial variabilities in average monthly rainfall. This is also consistent with the semivariogram modeling results, which yields smaller and larger influence ranges for the MJJ and ASO period semivariograms, respectively.

## 7. PRINCIPAL COMPONENT ANALYSIS OF AVERAGE MONTHLY RAINFALLS

Principal component analysis (PCA) has been widely

used in the fields of hydrology, meteorology, and climatology to simplify large data sets for better understanding of climatological precipitation patterns (Ehrendorfer 1987; Baeriswyl and Rebetez 1997; Roy and Kaur 2000; Singh 2006). In this study, PCA was also conducted on the average monthly observed rainfall at different rainfall stations.

Each original variable (average monthly rainfall for a particular month in the wet season) was normalized with respect to its mean and standard deviation and then all normalized variables were used for the PCA. The first three principal components account for 88% of the total variation. The percentages of the total variation in average monthly rainfall explained by the first, the second and the third principal components are 54, 22, and 12%, respectively. Principal component loads (PC loads) represent correlation coefficients between the principal component scores (PC scores) and the normalized original variables. PC loads measure the importance of each variable in accounting for the variability in the principal components.

PC loads of the first three principal components are shown in Fig. 11. For the first principal component (PC1), PC loads for the normalized average monthly rainfalls in May to October are nearly the same (particularly, for June to September). This implies that PC1 mainly explains the spatial variabilities in average monthly CDZ rainfalls. This interpretation can also be illustrated by comparing frequency histograms of PC scores for different clusters (see Fig. 12). Histograms of PC1 scores for rainfall stations in different clusters are significantly different. Particularly, PC1 scores for Clusters 1 and 3 are well separated whereas PC1 scores for Cluster 2 cannot be completely separated from Clusters 1 and 3. Thus, cluster-specific histograms of PC2 scores show that Clusters 1 and 3 have nearly the same range (since variations between the two clusters have been completely accounted for by PC1 and PC2 scores for Cluster 2 show a certain degree of separation from the PC2 scores for Clusters 1 and 3. Finally, with most variations among the three clusters accounted for by the first two principal components, PC3 scores for the three clusters fall mostly in the same range of (-2, 2).

PC loads for PC2 show negative correlation with the average monthly rainfalls for June and July and positive correlation with the average monthly rainfalls for September and October. The temporal variation pattern in PC2 loads also resembles the temporal variation pattern of the average monthly rainfalls. These indicate that the second principal component mainly explains the seasonal variation in average monthly rainfalls, although it also accounts for the spatial variations among Cluster 2 and the other two clusters, which have not been accounted for by the first principal component.

The fact that PC1, which mainly explains the spatial variabilities in average monthly rainfall, accounts for significantly higher percentages of the total variation than PC2 (54 to 22%), which mainly explains the seasonal variation in average monthly rainfall, indicating that during the





Fig. 8. Hartigan indicator versus number of clusters in the K-means cluster analysis using the average monthly rainfall (May to October) as classification features. (Color online only)

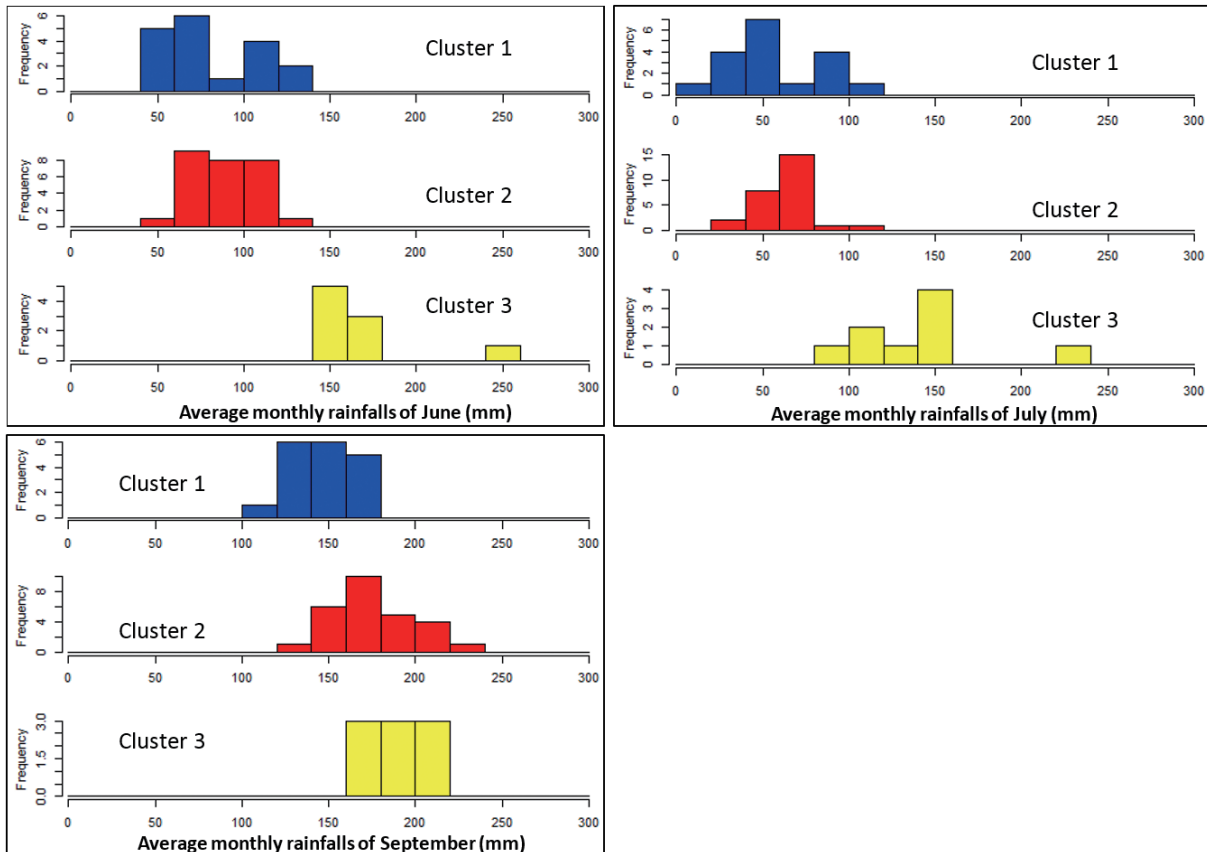


Fig. 9. Histograms of average monthly rainfalls for June, July, and September for rainfall stations in different clusters. (Color online only)

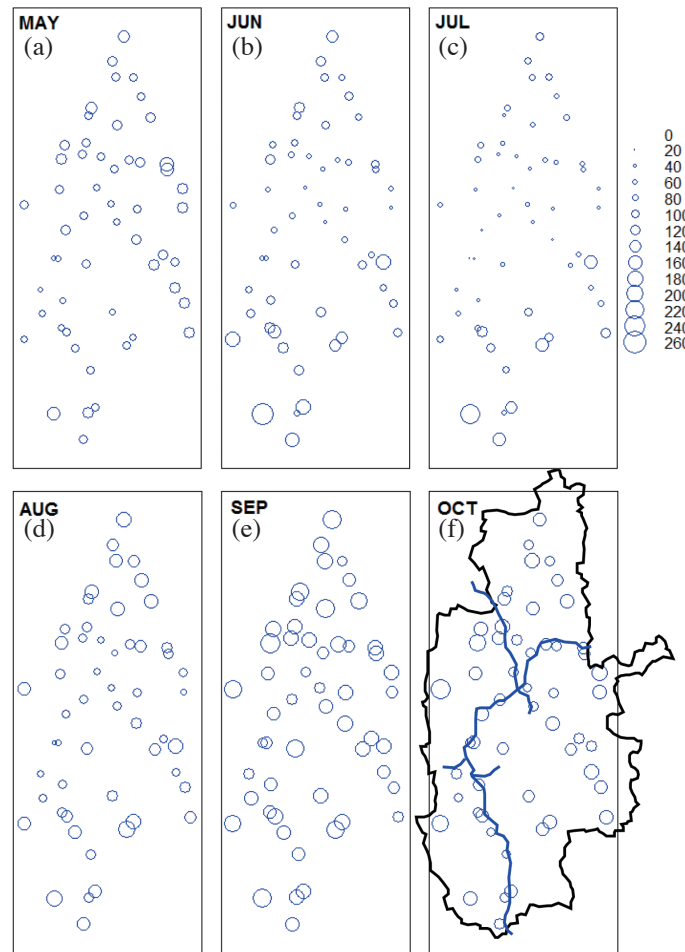


Fig. 10. Spatial and temporal variations in average monthly rainfall (a) - (f). CDZ boundary and the Irrawaddy River are shown in Panel (f). (Color online only)

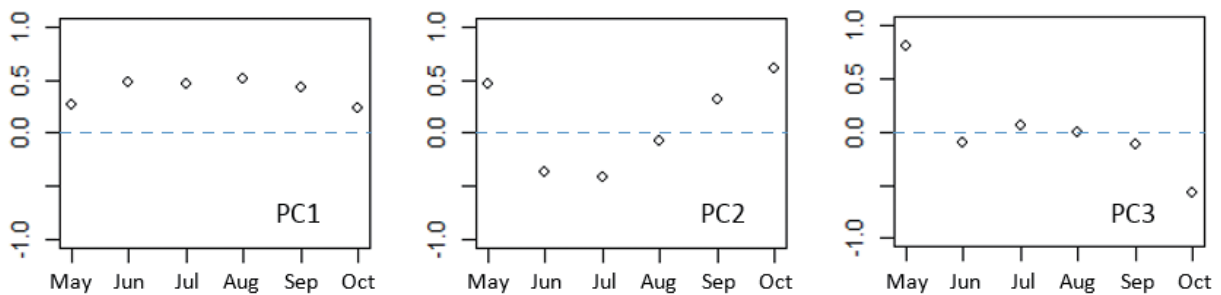


Fig. 11. Principal component loads for the normalized average monthly rainfall for the first three principal components (PC1, PC2, and PC3). (Color online only)

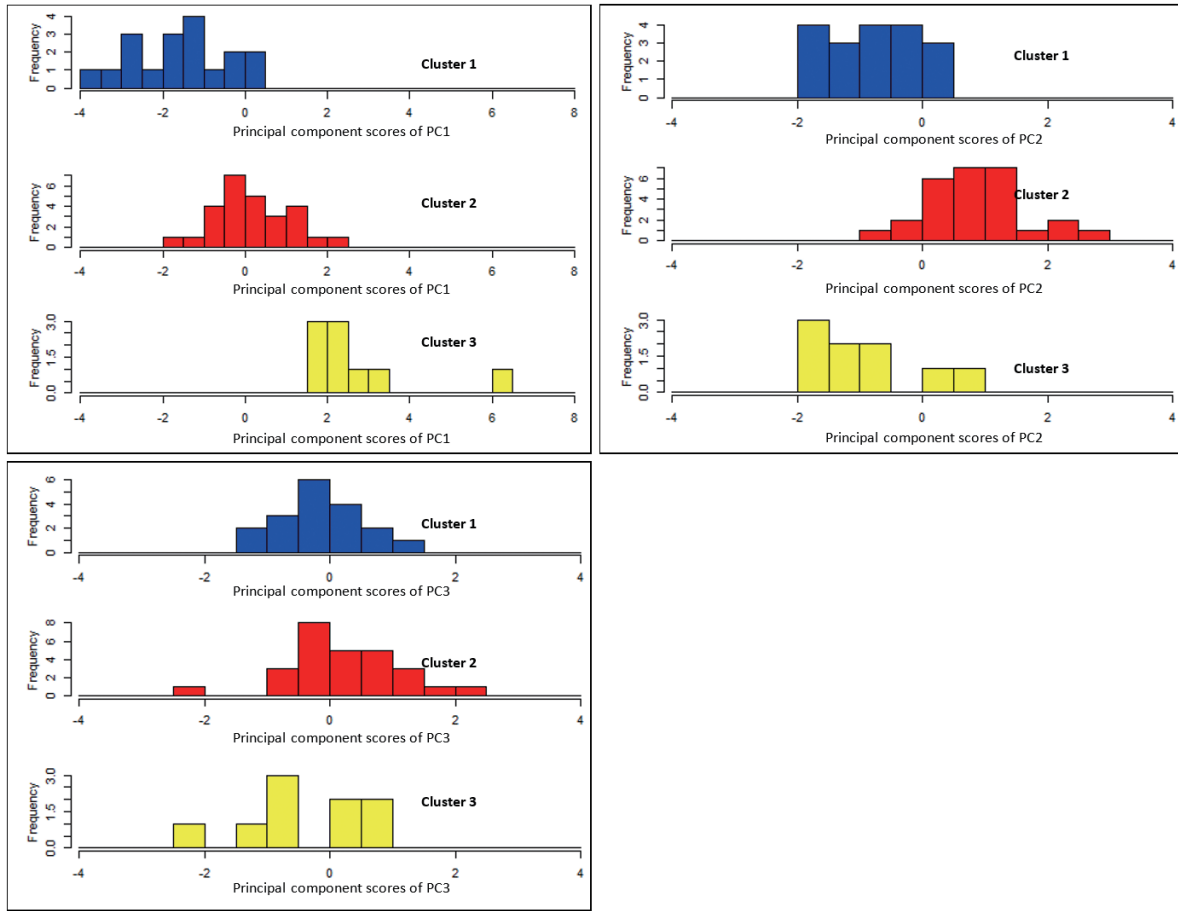


Fig. 12. Frequency histograms of PC scores for rainfall stations in different clusters. (Color online only)

wet season spatial rainfall variations in the CDZ are more significant than the seasonal (temporal) rainfall variations.

PC loads for PC3 have near-zero values for June, July, August and September and yet significantly positive (0.80) and negative (-0.57) loads for May and October, respectively. Since the onset and the ending of the wet season may not always fall exactly at the beginning of May and the end of October, the relatively minor effect of changing the onset and ending time of the wet season on average monthly rainfalls for May and October may be accounted for by PC3.

### 8. SUMMARY AND CONCLUSIONS

This study analyzed rainfall data on different scales, including daily, event-scale, and monthly rainfalls in the CDZ of Myanmar to characterize their spatial and temporal variation patterns. Exploratory data analysis, semivariogram analysis and modeling, K-means cluster analysis and the principal component analysis were conducted to investigate the spatial and temporal variation patterns in the CDZ rainfalls. A few concluding remarks are given as follows:

(1) The CDZ has a very distinctive pattern of dry and wet seasons, with nearly all annual rainfall occurring in the

wet season. A climatological monsoon break occurs at the end of July and divides the wet season into two peaks. The monsoon break is the result of different climate dynamics - monsoon southwesterly of the MJJ period and tropical cyclonic vorticity of the ASO period.

- (2) Rainfalls in the MJJ period, which are mostly affected by monsoon southwesterlies, have smaller influence ranges and higher degrees of spatial variability compared to rainfalls in the ASO period, which are influenced by large scale tropical cyclones. Monthly rainfalls have significantly higher spatial correlations and larger influence ranges than daily or event-scale rainfalls. Notably, the larger influence range and presence of a trend in spatial variation of the monthly-scale rainfalls jointly contribute to the linear semivariogram modeling of monthly rainfalls in the CDZ.
- (3) Rainfall stations in the three clusters identified by the K-means cluster analysis reflect the orographic effect and different climate dynamics that influence the spatial and temporal variation patterns in rainfall. The Cluster 1 rainfall stations generally receive the least amount of rainfall among the three clusters due to the orographic effect. Cluster 3 rainfall stations have higher average

monthly rainfall than stations in other clusters and are more influenced by the monsoon southwesterlies and the tropical cyclonic vorticity from the South China Sea.

- (4) The first three principal components account for 88% of the total variation. The first principal component mainly explains the spatial variabilities of average monthly rainfalls in the CDZ. The spatial rainfall variations are more significant than the seasonal (temporal) rainfall variations during the wet season.
- (5) The second principal component mainly explains the seasonal variation in the average monthly rainfall, although it also accounts for the spatial variations among Cluster 2 and the other two clusters, which have not been accounted for by the first principal component.

**Acknowledgements** This work was supported by JSPS KAKENHI Grant Number 23248055. The authors are grateful to Professor Eiji Nawata for sharing the fund. We also thank the Central Dry Zone Greening Department, Myanmar for supporting the rainfall data. The editor and anonymous reviewers are appreciated for their helpful comments and suggestions.

## REFERENCES

- Baeriswyl, P. A. and M. Rebetez, 1997: Regionalization of precipitation in Switzerland by means of principal component analysis. *Theor. Appl. Climatol.*, **58**, 31-41, doi: 10.1007/BF00867430. [[Link](#)]
- Bastin, G., B. Lorent, C. Duqué, and M. Gevers, 1984: Optimal estimation of the average areal rainfall and optimal selection of rain gauge locations. *Water Resour. Res.*, **20**, 463-470, doi: 10.1029/WR020i004p00463. [[Link](#)]
- Central Statistical Organization, 2008: Statistical Yearbook, The Government of the Republic of the Union of Myanmar, Ministry of National Planning and Economic Development.
- Cheng, K. S., C. Wei, Y. B. Cheng, and H. C. Yeh, 2003: Effect of spatial variation characteristics on contouring of design storm depth. *Hydrol. Process.*, **17**, 1755-1769, doi: 10.1002/hyp.1209. [[Link](#)]
- Cheng, K. S., Y. C. Lin, and J. J. Liou, 2008: Rain-gauge network evaluation and augmentation using geostatistics. *Hydrol. Process.*, **22**, 2554-2564, doi: 10.1002/hyp.6851. [[Link](#)]
- Cho, H. A., 1999: Climate Trend of Central Myanmar, Workshop on Greening of Central Dry Zone, Forest Research Institute, Yezin, Myanmar, September 20-21, 1999.
- Ehrendorfer, M., 1987: A regionalization of Austria's precipitation climate using principal component analysis. *Int. J. Climatol.*, **7**, 71-89, doi: 10.1002/joc.3370070107. [[Link](#)]
- Gratzfeld, J., 2003: Extractive Industries in Arid and Semi-Arid Zones: Environmental Planning and Management, IUCN, Gland, Switzerland and Cambridge, U.K., 122 pp.
- Hartigan, J. A., 1975: Printer graphics for clustering. *J. Stat. Comput. Simul.*, **4**, 187-213, doi: 10.1080/00949657508810123. [[Link](#)]
- Hughes, J. P. and D. P. Lettenmaier, 1981: Data requirements for kriging: Estimation and network design. *Water Resour. Res.*, **17**, 1641-1650, doi: 10.1029/WR017i006p01641. [[Link](#)]
- Journel, A. G. and C. J. Huijbregts, 1978: Mining Geostatistics, Academic Press, London and New York, 600 pp.
- Kottek, M., G. Jürgen, B. Christoph, R. Bruno, and R. Franz, 2006: World map of the Köppen-Geiger climate classification updated. *Meteorol. Z.*, **15**, 259-263, doi: 10.1127/0941-2948/2006/0130. [[Link](#)]
- Lebel, T., G. Bastin, C. Obled, and J. D. Creutin, 1987: On the accuracy of areal rainfall estimation: A case study. *Water Resour. Res.*, **23**, 2123-2134, doi: 10.1029/WR023i011p02123. [[Link](#)]
- Peel, M. C., B. L. Finlayson, and T. A. McMahon, 2007: Updated world map of the Köppen-Geiger climate classification. *Hydrol. Earth Syst. Sci.*, **11**, 1633-1644, doi: 10.5194/hess-11-1633-2007. [[Link](#)]
- Poe, C. A., 2011: Food Security Assessment in the Dry Zone Myanmar, WFP, Food Security Analysis Services, 29 pp.
- Roy, N. S. and S. Kaur, 2000: Climatology of monsoon rains of Myanmar (Burma). *Int. J. Climatol.*, **20**, 913-928, doi: 10.1002/1097-0088(20000630)20:8<913::AID-JOC485>3.0.CO;2-U. [[Link](#)]
- Roy, S. S. and N. S. Roy, 2011: Influence of Pacific decadal oscillation and El Niño southern oscillation on the summer monsoon precipitation in Myanmar. *Int. J. Climatol.*, **31**, 14-21, doi: 10.1002/joc.2065. [[Link](#)]
- Singh, C. V., 2006: Pattern characteristics of Indian monsoon rainfall using principal component analysis (PCA). *Atmos. Res.*, **79**, 317-326, doi: 10.1016/j.atmosres.2005.05.006. [[Link](#)]
- Takahashi, H. G. and T. Yasunari, 2006: A climatological monsoon break in rainfall over Indochina-A singularity in the seasonal march of the Asian summer monsoon. *J. Climate*, **19**, 1545-1556, doi: 10.1175/JCLI3724.1. [[Link](#)]
- Takahashi, H. G. and T. Yasunari, 2008: Decreasing trend in rainfall over Indochina during the late summer monsoon: Impact of tropical cyclones. *J. Meteorol. Soc. Jpn.*, **86**, 429-438, doi: 10.2151/jmsj.86.429. [[Link](#)]
- The Union of Myanmar, 2005: National Action Programme of Myanmar to Combat Desertification in the Context of United Nations Convention to Combat Desertification (UNCCD), Ministry of Forestry, Myanmar.
- UNEP, 1992: World Atlas of Desertification, London; Baltimore: Edward Arnold.