Nonlinear, Intraseasonal Phases of the East Asian Summer Monsoon: Extraction and Analysis Using Self-Organizing Maps

JUNG-EUN CHU
Division of Earth Environmental System, Pusan National University, Busan, Korea

SAJI N. HAMEED
Center for Advanced Information Science and Technology, University of Aizu, Aizu-Wakamatsu City, Japan

KYUNG-JA HA
Division of Earth Environmental System, Pusan National University, Busan, Korea

(Manuscript received 8 September 2011, in final form 3 April 2012)

ABSTRACT

The hypothesis that regional characteristics of the East Asian summer monsoon (EASM) result from the presence of nonlinear coupled features that modulate the seasonal circulation and rainfall at the intraseasonal time scale is advanced in this study. To examine this hypothesis, the authors undertake the analysis of daily EASM variability using a nonlinear multivariate data classifying algorithm known as self-organizing mapping (SOM).

On the basis of various SOM node analyses, four major intraseasonal phases of the EASM are identified. The first node describes a circulation state corresponding to weak tropical and subtropical pressure systems, strong upper-level jets, weakened monsoonal winds, and cyclonic upper-level vorticity. This mode, related to large rainfall anomalies in southeast China and southern Japan, is identified as the mei-yu–baiu phase. The second node represents a distinct circulation state corresponding to a strengthened subtropical high, monsoonal winds, and anticyclonic upper-level vorticity in southeast Korea, which is identified as the changma phase. The third node is related to copious rain over Korea following changma, which we name the post-changma phase. The fourth node is situated diagonally opposite the changma mode. Because Korea experiences a dry spell associated with this SOM node, it is referred to as the dry-spell phase.

The authors also demonstrate that a strong modulation of the changma and dry-spell phases on interannual time scales occurs during El Niño and La Niña years. Results imply that the key to predictability of the EASM on interannual time scales may lie with analysis and exploitation of its nonlinear characteristics.

1. Introduction

The East Asian summer monsoon (EASM) is well known for its intraseasonal and interannual variability. The rainy season known as mei-yu in China, baiu in Japan, and changma in Korea manifests regional differences that are prominent on the intraseasonal time scale. The primary monsoon rainband associated with mei-yu and baiu develops in mid June in the lower regions of the Yangtze River valley and in southern Japan, respectively (Kang et al. 1999; Wang et al. 2007). Conversely, the Korean rainy season involves two separate rainy periods: a main surge in mid June, known as the changma period, and a secondary surge in mid August to early September, known as the post-changma period (Kim et al. 2010; Wang et al. 2007). The modulation of regional rainy seasons of the EASM including mei-yu, baiu, and changma, is regarded as meridional shifts of the heavy rain belt, which may be a local manifestation of a northward-propagating climatological intraseasonal oscillation (CISO) (Wang and Xu 1997). Although the CISO displays regular phase changes with respect to the calendar year, the northward propagation of the summer monsoon intraseasonal
oscillation (ISO) are not always phase locked (Wang and Rui 1990), so the structure and propagation of the summer monsoon ISO tend to be more irregular (Yoo et al. 2010). Understanding the regional nature of the EASM, particularly the underlying nonlinear characteristics, may provide important insight into the difficulties in the seasonal-to-interannual predictability of EASM precipitation (Kang et al. 2002; Kim et al. 2008; Lee et al. 2011).

Most previous studies have examined intraseasonal variations embedded in the EASM by using statistical methods such as covariance analysis (Kang et al. 1999; Lau and Chan 1986) and multichannel singular spectrum analysis (Krishnamurthy and Shukla 2007). However, these linear analyses are limited in their ability to describe the full set of characteristics of monsoon subseasonal variation (Yoo et al. 2010). Thus, nonlinear analysis has recently gained popularity. Yoo et al. examined the spatial patterns of discrete rainfall states of the Asian summer monsoon intraseasonal phases derived from a hidden Markov model (HMM). However, a lack of explicit constraints to control the time scale of intraseasonal phases causes the HMM to simulate more high-frequency variability than is actually observed (Jones 2009; Yoo et al. 2010). Chattopadhyay et al. (2008) adopted a self-organizing map (SOM) to objectively identify and explore nonlinear characteristics of the Indian monsoon ISO. The aforementioned studies have focused on the near-tropical ISO. A similar examination of the ISO within the more subtropical belt of the EASM domain was attempted by Chu and Ha (2011) but they only focused on a methodological approach for monsoon intraseasonal phases. However, a dynamical analysis and interpretation of intraseasonal phases was not undertaken.

In this paper, we adopt the SOM methodology—a type of unsupervised, artificial neural network—to objectively identify nonlinear phases embedded within the EASM circulation. In addition, we explore associated circulation states, teleconnections, and their manifestations on Korean rainfall. The main advantage of the SOM approach is that it provides a concise description of the main patterns of variability while accommodating nonlinearity in the data. Further, it provides a powerful visualization of the underlying data structure and relations among the main modes of variability. A SOM analysis, together with a set of generalized patterns produced from the input data, describes the multidimensional distribution function of the data (Hewitson and Crane 2002; Kohonen 1990). The input data is a series of circulation state vectors comprising six important daily indices representing subtropical high pressure regions, lower- and upper-level wind vectors, and vertical and horizontal wind shear. Our SOM analysis yields a nonlinear classification of the continuum of atmospheric state vectors while preserving both the underlying probability distribution function of the data and the topological relationship between the states.

Scale interactions between interannual and intraseasonal modes of variability are aspects of EASM variability that carry important implications for predictability. In their study of the relationship between the El Niño–Southern Oscillation (ENSO) and ISO, Tam and Lau (2005) examined the propagation and growth/decay characteristics of ISO in various phases of ENSO on the basis of a lag correlation technique. Yun et al. (2010) discovered a significant lag correlation between interannual variability of northward propagating ISO, which has a quasi-biennial time scale through preceding and concurrent summers, and ENSO events. Teng and Wang (2003) demonstrated that the ENSO affects the northwestward-propagating ISO mode in the western North Pacific by changing the mean circulation through the vertical wind shear mechanism. Moreover, Yoo et al. (2010) showed the interannual modulation of the ISO associated with ENSO by employing a nonhomogeneous HMM. Hence, an additional aspect of this study examines the relationship between the nonlinear circulation states identified in our study and interannual climate modes, particularly the ENSO.

The rest of the paper is organized in the following manner. In section 2 we describe the data and briefly explain the main ideas behind SOM methodology. In section 3 we identify the intraseasonal phases and present underlying circulation patterns derived from SOM. Section 4 examines the relationship between intraseasonal phases and ENSO, and section 5 contains a summary and conclusions.

2. Data and methods

a. Observational datasets

The data used for the analysis of summer monsoon rainfall were collected from 1997 to 2008 by the daily Global Precipitation Climatology Project (GPCP). Because the daily GPCP began in October 1996, no earlier data were available. Data used for the analysis of large-scale circulation characteristics, including the geopotential height and horizontal components of winds at selected pressure levels, were collected from 1979 to 2008 by the National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalysis (Kalnay et al. 1996). In addition, daily precipitation data collected at 79 synoptic stations located throughout the Korean Peninsula were used to examine regional rainfall amount.
TABLE 1. Description of the six East Asian summer monsoon circulation indices: $U$ is zonal and $V$ is meridional wind and $Z$ is geopotential height.

<table>
<thead>
<tr>
<th>Index</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>CI500H</td>
<td>$Z_{850} {25^\circ-35^\circN, 135^\circ-152.5^\circE}$</td>
</tr>
<tr>
<td>CI850U</td>
<td>$U_{850} {32.5^\circ-37.5^\circN, 127.5^\circ-147.5^\circE}$</td>
</tr>
<tr>
<td>CI850V</td>
<td>$V_{850} {32.5^\circ-37.5^\circN, 127.5^\circ-147.5^\circE}$</td>
</tr>
<tr>
<td>RM2</td>
<td>$U_{200} {40^\circ-50^\circN, 110^\circ-150^\circE} - U_{200} {25^\circ-35^\circN, 110^\circ-150^\circE}$</td>
</tr>
<tr>
<td>SI</td>
<td>$U_{850} {5^\circ-15^\circN, 90^\circ-130^\circE} - U_{850} {5^\circ-15^\circN, 90^\circ-130^\circE}$</td>
</tr>
<tr>
<td>WNPMI</td>
<td>$U_{850} {20^\circ-30^\circN, 110^\circ-140^\circE}$ - $U_{850} {20^\circ-30^\circN, 110^\circ-140^\circE}$</td>
</tr>
</tbody>
</table>

b. East Asian summer monsoon indices

The EASM was sometimes restricted to the subtropical monsoon that prevails over eastern China north of 20°N, Korea, Japan, and adjacent marginal seas (Chen et al. 2000; Mao and Wu 2006). However, the EASM domain in this paper covers a large area from tropical regions including part of the western North Pacific to extratropical regions so as to include the western North Pacific subtropical high (WNPSH). The WNPSH is an important component in a coupled circulation–convection system of the EASM with its role of moisture transport and linkage between the ENSO and EASM.

The evolution of regional summer monsoon rainfall is accompanied and characterized by large-scale changes in the WNPSH, low-level winds, and the associated moisture transport from the Indo-Pacific warm pool, upper-level Asian jet, vertical shear, and planetary-scale teleconnection patterns of circulation, among other features (Ha et al. 2005; Ha and Lee 2007). Because of the similarity between the seasonal mean and dominant ISO mode (Ha et al. 2005), large-scale circulation indices can be used for the indices of the dominant intraseasonal mode. On the basis of dynamical consistency and regional relevance of precipitation intraseasonal phases, we chose six daily monsoon indices derived from circulation fields, excluding moisture fields (Table 1). The daily standardized anomalies of these six indices were constructed by subtracting the climatological daily mean: index characteristics are described below.

(i) The CI500H index represents a stationary anticyclone over the western North Pacific (Ha et al. 2005).

(ii) The CI850U and CI850V indices represent low-level jets from the southwest flow over the Asian monsoon associated with a strong changma. Advection of moist and warm air by low-level winds is essential for generating convective instability and sustaining convective activity (Ninomiya 1980).

(iii) The RM2 index, proposed by Lau et al. (2000), represents upper-level vorticity, a prominent feature of which is the northward advance of the WNPSH, which causes the axis of the climatological subtropical jet to migrate northward by about 10°–15° latitude. Lau et al. argued that this process represents a remarkable response of the subtropical upper-level flow to tropical heating in the Southeast Asian region.

(iv) The SI index represents the vertical shear index, as described in Wang et al. (1998). The vertical shear of zonal wind represents the zonal thermal winds between 850 and 200 hPa that result from north–south and land–sea thermal contrasts.

(v) The western North Pacific monsoon index (WNPMI) represents the difference of 850-hPa westerlies between a southern region (5°–15°N, 110°–130°E) and a northern region (20°–30°N, 110°–140°E). This latitudinal differential westerly index reflects not only the strength of the tropical westerlies but also the intensity of the low-level vorticity associated with the Rossby wave response to the Philippine Sea convective heat source (Wang et al. 2001).

c. Self-organizing map

1) THE SOM ALGORITHM

The self-organizing mapping (SOM) consists of two layers: input and competitive. The input layer is fully connected to the competitive layer of map nodes (Fig. 1). When an input is presented, the output nodes compete to represent the pattern. The SOM’s two processes are that of training and mapping. The key to the ability of the SOM to extract patterns is the way that it learns; this is embodied in its training algorithm. Artificial neural networks learn by an iterative process, whereby input data are presented successively to the networks. The initial step in this iterative procedure is to randomly distribute nodes in the data space. Thus, in our two-dimensional example, nodes were initially a random cloud of points in two dimensions. After successive presentation of the input data to the network, the nodes approach the positions that best represent the input data. The number of nodes (output patterns) is defined by the user and is dependent on the level of detail desired in the analysis. In our study, we applied $3 \times 3$ nodes physically considering active, break, and normal basic states and their underlying substates of ensemble mean, above state, and below state (Chattopadhya et al. 2008). The reason why we used a $3 \times 3$ SOM map is that the $3 \times 3$ map most
effectively distinguishes the four intraseasonal phases. To ensure the robustness of the SOM analysis, we examined the sensitivity of SOM sizes (figure not shown). Since SOM classification includes both major and subsequent modes, a small number of nodes may superimpose the transitional properties of intraseasonal phases on the major modes. To include both major and subsequent modes together, a larger number of nodes than that of major nodes is needed. If the SOM is larger than the $3 \times 3$ map, then the major modes are not clearly distinguished from adjacent modes (figure not shown).

Mathematically, the principal goal choosing the number of clustering is to maximize similarity within clusters and minimize similarity between clusters. Therefore, based on consideration of mathematical optimization and the physical requirement of identifying distinct patterns, a configuration of $3 \times 3$ states is chosen.

The general SOM training algorithm is outlined below; mathematical details can be found elsewhere (Kohonen 1997). All input vectors are fully connected to nodes in the competitive layer, and the nodes are uniquely defined by a reference vector consisting of weighting coefficients. Adjustment of the reference vectors to the input vector, an essential part of SOM, is achieved through a user-defined iterative cycle. This adaptation minimizes the Euclidean distance (EUD) between the reference vector for any $j$th node $W_j$ and the input data vector $X$; that is,

$$EUD = |X - W| = \sqrt{\sum_{i=1}^{n} (X_i - W_i)^2}.$$

The first input sample is then compared with each node in the competitive layer. The node with the least Euclidean distance between itself and the input vector is known as the winning node. During the iterative process, the winner node updates the reference vector and its associated weights together with those of neighbor nodes within the neighborhood radius. Since each node has to be adjusted relative to its neighbor, inclusion of the neighborhood makes the SOM classification nonlinear. The training cycle may be continued $n$ times, and updating equations are described as

$$W_j(n+1) = \begin{cases} W_j(n) + c(n)[X(n) - W_j(n)], & j \in R_j(n) \\ W_j(n), & \text{otherwise}. \end{cases}$$

In this reference vector for the $j$th node for the $n$th training cycle, $X(n)$ is the input vector; $R_j(n)$ is the predefined neighborhood around node $j$; and $c(n)$ is the neighborhood kernel, which defines the neighborhood. The neighborhood kernel may be a monotonic decreasing function of $n$ [$0 < c(n) < 1$], known as a bubble, or it may be of Gaussian type as

$$\alpha(n) \exp \left[ -\frac{||r_j - r_i||^2}{2\sigma^2(n)} \right],$$

where $\alpha(n)$ and $\sigma(n)$ are constants monotonically decreasing with $n$. Here $\alpha(n)$ is the learning rate, which determines the velocity of the learning process, while $\sigma(n)$ is the amplitude, which determines the width of the
neighborhood kernel. We used Gaussian type as the neighborhood kernel. In addition, \( r_i \) and \( r_j \) are the coordinates of the nodes \( j \) and \( i \), respectively, in which the neighborhood kernel is defined. Throughout training, the learning rate and size of the update neighborhood—the update radius—decrease, leading to progressively refined initial generalized patterns. Finally, the SOM consists of a number of patterns characteristic of the data, with similar patterns nearby and dissimilar patterns farther apart. After the training process, the final map, or reference vector, is completed. The mapping process distributes each input vector to a corresponding reference vector based on its similarity, such as the least Euclidean distance. In this way, the nodes in a self-organizing map compete to most effectively represent the particular input sample.

2) IMPLEMENTATION OF SOM

An input vector contains six components of circulation indices for a particular day (Fig. 1). Similarly, the corresponding reference vector has six weighting coefficients. Once we obtained classifications using the SOM algorithm, the dates from June to August (JJA) over 30 years from 1979 to 2008 were collected at each node. Thus, the number of input samples was 2760, representing 30 \( \text{yr} \times 92 \) days JJA, which were finally mapped onto a two-dimensional \((3 \times 3)\) lattice. Each node contains a reference vector consisting of six indices and clustered dates. The composite of classified dates provided a spatial structure of each phase; that is, if the summer monsoon ISO is a convectively coupled oscillation, each pattern should be strongly related to a particular phase of the precipitation oscillation. In addition to spatial pattern, the clustered dates add temporal information.

### Table 2. Mean days per event, represented by bold type. Percent frequency of days and the probability of no transition at each node are in parentheses and braces, respectively.

<table>
<thead>
<tr>
<th>((1, 1))</th>
<th>((1, 2))</th>
<th>((1, 3))</th>
</tr>
</thead>
<tbody>
<tr>
<td>6 (18.4%) {78.9}</td>
<td>3 (7.1%) {52.0}</td>
<td>4 (12.6%) {69.2}</td>
</tr>
<tr>
<td>3 (9.2%) {44.9}</td>
<td>(2, 1) {2, 2} {2, 3}</td>
<td>(2, 3)</td>
</tr>
<tr>
<td>(3, 1) {3, 2} {3, 3}</td>
<td>3 (7.8%) {35.5}</td>
<td></td>
</tr>
<tr>
<td>4 (14.7%) {69.2}</td>
<td>3 (8.4%) {51.1}</td>
<td>5 (18.1%) {75.4}</td>
</tr>
</tbody>
</table>

3. Results

a. Identification of intraseasonal phases derived from SOM

A composite of classified dates was performed in section 3a to detect the geographical rainfall structure of each node. To include information of basic statistics of intraseasonal phases, we showed the mean days per event, percent frequency of days, and probability of no transition at each node (Table 2). Here, the number of events was determined by counting the total number of times the data records were mapped consecutively to a particular node with no break. Mean days per event were defined by averaging the number of consecutively mapped days per event. Frequency of days was defined as the number of days clustered in a particular node divided by the total number of days used in the classification (30 \( \text{yr} \times 92 \) days yr\(^{-1}\)). The probability of no transition, also expressed in percentage, is the probability that, when an input vector corresponding to a particular day is mapped to a node, the next day will be mapped again to the same node. Thus, for the \((1, 1)\) node, \(78.9\%\) of the cases are successively projected onto the node. Similarly, it is of the same order for the node \((3, 3)\) and is lowest for \((2, 2)\). This implies that, when a day is attached to a \((1, 1)\) or \((3, 3)\) node, the next day shows the highest probability of clustering at the same node; the chance is lowest for the node \((2, 2)\). Further, it can be seen that mean days per event is highest for the \((1, 1)\) node with 6 days and \((3, 3)\) node with 5 days and the corresponding percentages of frequencies of days clustered at these nodes are also higher. Comparing to results based on the Indian summer monsoon suggested by Chattopadhyay et al. (2008), there are four nodes that are sustainable and have greater portion, while two major nodes are found for the Indian monsoon ISO.

The statistics described above can be used to examine the typical time scales of variability for any chosen node. It is well known that the intraseasonal oscillation of the EASM has a broadband spectrum ranging from 20 to 60 days. It has been proposed that the northward propagating oscillation exhibits dominant periodicities in the 30–60-day (Tsou et al. 2005; Wang and Xu 1997; Yun et al. 2009) or 20–50-day (Mao et al. 2010) time scales over the North Pacific and East Asian region. Another periodicity of 10–20-day oscillation controls the behavior of the South China Sea (SCS) summer monsoon and Yangtze rainfall for most years (Chen and Chen 1995; Mao and Chan 2005). This broadband nature of the frequency spectrum may be due to nonlinear interaction between the dominant periodicities and higher and lower periodicities. In comparison with ISO periodicities mostly derived by filtered outgoing longwave radiation (OLR) or precipitation anomalies, we present the periodicity of an ISO event using a persistence of each intraseasonal phase from discretized dates obtained through SOM analysis. The persistence is represented as mean days per event in Table 2. Assuming that one full cycle of the intraseasonal phase is an episode, the total number of days per episode is 33, which corresponds to the most
dominant periodicity of an ISO over the East Asian region. Thus, the aforementioned results demonstrate the quantitative estimate of the ISO within a season available in various sources and allows for further application of the SOM to study the monsoon ISO.

On the basis of intraseasonal phases derived from the SOM algorithm, the spatial patterns of four nodes including (1, 1), (3, 1), (3, 3), and (1, 3) were considered as major modes; their underlying dynamical fields are suggested in section 3c.

b. Classification of precipitation states

The composite precipitation corresponding to the clustered dates of four major nodes are shown in Fig. 2 [full figures of nine nodes can be found in Chu and Ha (2011)]. In (1, 1) the zonally elongated rainfall in southeast China and southern Japan is similar to an onset structure called mei-yu and baiu. Following a counterclockwise direction, the (2, 1) node shows a northward-shifted center of rainfall over southern Korea and southeastern Japan (figure not shown). The (3, 1) and (3, 3) nodes represent a changma-like pattern with copious rainfalls over Korea. The observed data from Korean synoptic stations also demonstrate that the rainfall averages in the nodes are almost three times higher than those in opposite nodes (not shown). While the regions over 25°N show similar patterns, distinct differences among changma-proper nodes can be found over the subtropical western Pacific regions. Dry conditions appear along the western North Pacific high in the (3, 1) node, and wet conditions dominate over the subtropical western North Pacific in the (3, 3) node. Temporal analysis shows that...
the (3, 3) phase occurs following the changma season rains. It is interesting to note that SOM distinguishes the secondary peak of Korean rain, which has been recently regarded as the postchangma season. In the (1, 3) node, continental regions experience a dry spell associated with the condition, while oceanic areas have scattered rainfall distribution. The northward propagation of the convective center can be seen by following the panels counterclockwise starting from (1, 1) in Fig. 2.

Until now, we have found that SOM effectively captured the regional characteristics of various phases in intraseasonal monsoon precipitation. However, this method does not clearly explain the temporal evolution between different intraseasonal phases. Figure 3 shows the number of clustered days in JJA for 30-yr periods in four major nodes so that the seasonal variation can be seen [full figures of nine nodes can be found in Chu and Ha (2011)]. For example, if each case for 1 June is clustered in the (1, 1) node from the entire 30-yr periods, it will be shown as 30 for 1 June. It is evident that the early stage of summer tends to be clustered at the (1, 1) node, and the maximum days for each cluster shows the movement along the counterclockwise direction starting from the (1, 1) node. According to Fig. 3b, many portions of days for the (3, 1) node are concentrated in mid-June to late July; this period is equivalent to the changma season. The maximum number of clustered days in the (1, 1) node and (3, 1) node are 1 June and 4 July, respectively.

In the (3, 3) node, days appeared in late June and gradually increased in early August. August is regarded as the prime season for tropical cyclones such as typhoons. Many days for August are divided into (3, 3) and (1, 3) nodes. Although the number of days of the (2, 1) and (2, 3) nodes are evenly distributed throughout the summer, not much variance is evident (figure not shown). It can be found that there is specific preference for each node during the summer season and oscillating features of the nodes. This result also indicates that each mode can be viewed as one component of the monsoon ISO that is phase locked to the seasonal cycle to a certain degree.

On the basis of the precipitation patterns and evolutionary history of nine nodes, four major nodes—(1, 1), (3, 1), (3, 3), and (1, 3)—were named mei-yu–baiu, changma mode, postchanga, and dry-spell modes, respectively. In addition, we performed empirical orthogonal function (EOF) analysis (Fig. 4). It was found that EOF1 resembles the mei-yu–baiu mode, while EOF2 is similar to the changma mode on the basis of pattern correlation coefficients of 0.65 and 0.30. However, the postchanga and dry-spell modes are not shown in EOF analysis, which indicates that the SOM can capture the distinguished patterns between changma and postchanga modes and the terminated monsoon precipitation structure in the dry-spell mode. It is clear that the SOM technique, through the use of many large-scale circulation parameters, is able to capture the low-frequency subseasonal variability
of rainfall over East Asia. In this study, we use six largescale monsoon indices—including CI500H, CI850U, CI850V, RM2, SI, and WNPMI—selected as predictors. To ensure the robustness of the SOM analysis, we examined the sensitivity of predictors to SOM classification by removing each index from the others. The results exhibit that almost identical patterns of the four major modes can be captured even though one predictor is removed (figure not shown). Most of the pattern correlation coefficients (PCCs) between four major modes from original experiment and those from sensitivity experiments show higher than 0.9 values. However, the dry-spell mode is rather sensitive to CI500H with its relatively lower PCC (=0.69). Considering that PCC = 0.69 is still a substantial value, the set of six large-scale indices can be reasonable predictors. The patterns of dynamical field and interpretation associated with each mode are subsequently discussed in detail.

c. Large-scale circulation related to intraseasonal phases

The large-scale patterns of several other dynamical variables associated with the four major modes identified by the SOM techniques are noted in Fig. 5. The 200-hPa zonal wind and the 850-hPa wind anomalies are presented to observe extension of the upper-level jet stream and low-level moisture transport, respectively. Rossby wave propagation is described by the 500-hPa geopotential height. Based on an analysis of the various SOM nodes, we identified four major intraseasonal
phases of the EASM located at the far corners of the SOM. These four nodes correspond to two major circulation patterns with opposite phases.

In the mei-yu–baiu mode, a zonally elongated jet stream is conspicuous, which represents a circulation state corresponding to weak tropical and subtropical pressure systems over the western Pacific, weakened monsoonal winds, and cyclonic upper-level vorticity over the Asian jet exit region. However, the vertical wind shear is large with stronger westerly winds in the upper troposphere (Fig. 5). This effect is also linked to relatively warmer conditions over the Indian Ocean produced by a heat-induced high and cooler condition over the Asian continents (figure not shown). This meridional temperature gradient can reinforce the jet stream through a thermal wind relationship.

The changma mode occurs with a distinct circulation state corresponding to a strengthened subtropical high, monsoonal winds, and anticyclonic upper-level vorticity over southeast Korea. However, the vertical shear is weak with a weaker upper-level westerly associated with a weaker and northward-shifted subtropical jet stream. Advection of moist, warm air by low-level winds is essential for generating convective instability and sustaining convective activity (Ninomiya 1980; Ha et al. 2005). The cold, dry inflow from the north and warm, moist air produced by the WNPSH demonstrate the convective instability that provides reasonably intense precipitation over the Korean Peninsula. The upper- and lower-level circulation features of the changma mode correspond to the strong changma patterns discussed by Ha et al. (2005). Another interesting feature is the presence of a weakened tropical high pressure system extending from the South China Sea to the Philippines.

A mirror image of the node representing the mei-yu–baiu phase can be observed in the circulation vector at its diagonally opposite corner [refer to Fig. 3 in Chu and Ha (2011)]. Temporal analysis shows that this phase occurs after the changma season rains and the midsummer dry period. Copious rains occur over Korea during this period, known as the postchangma phase. The prominent circulating feature of the mode is the northeastward advance of the WPH and convective activity over the subtropical western Pacific. The main effect of the northward advance of the western North Pacific subtropical high causes the axis of the climatological subtropical jet to migrate northward by about 10°–15° latitude (Lau et al. 2000), which represents a remarkable response of the subtropical upper-level flow to tropical heating over the western Pacific. A wave train pattern can be found from the Philippine Sea to the west coast of North America, which is considered as a convectively coupled Rossby wave-like system triggered by anomalous convective activity over the tropical western North Pacific (Hsu and Weng 2001; Mao et al. 2010).

The dry-spell mode is also diagonally opposite the changma mode and features a mirror image of the circulation vector. A low pressure anomaly develops over the subtropical western Pacific, while a high pressure anomaly intensifies northeast of the low pressure anomaly. The southwestward transport of moisture from the Pacific Ocean increases precipitation near the South China Sea. This process also terminates moisture transport from the equatorial Pacific into East Asia, which in turn creates dry conditions in Korea.

4. Relationship between ENSO and intraseasonal phases

a. Lead–lag correlations between ENSO events and intraseasonal phases

On the interannual time scale, the intraseasonal phases can be affected by slowly varying “external” components such as ENSO. The interannual relationship between ENSO and intraseasonal phases will help to overcome uncertainty in the prediction of interannual variability (IAV). Various studies have been performed on the lead–lag relationship between the tropical Pacific SST and the East Asian monsoon system (Chang et al. 2000; Lau and Weng 2001; Wu et al. 2003; Lau and Wang 2006; Lee et al. 2005; Wang et al. 2000). Typically, it has been considered that rainfall of the EASM tends to be enhanced following the preceding El Niño system, which has a mature phase during the boreal winter, December–February (DJF). El Niño persistently influences circulation and rainfall anomalies in East Asia through the following summer, JJA. However, the relationship between equatorial eastern Pacific SST anomalies and rainfall in East Asia remains a controversial issue. Chen et al. (1992) argued that significant correlations could not be detected between eastern Pacific SST anomalies and the EASM. These diverse results imply that interannual variation of the EASM is probably influenced by complex air–sea–land and tropical–extratropical interactions (Wang et al. 2000).

The ENSO teleconnection is broadly characterized by anomalous Philippine Sea anticyclonic results from a Rossby wave response to suppressed convective heating (Wang et al. 2000; Wu et al. 2003). To support our hypothesis that the intraseasonal phases of the EASM are related to ENSO, we constructed a composite difference diagram of simultaneous summertime (JJA) 850-hPa geopotential height anomalies and the preceding wintertime SST (DJF) for years with high and
low occurrences of the four modes (Fig. 6). To obtain the high and low occurrence years, we normalized the annual number of clustered days. If the normalized annual number of days exceeds one standard deviation, the year is regarded as a high occurrence year and, if it is below minus one standard deviation, the year is regarded as a low occurrence year.

As shown in Fig. 6, the high occurrence years for the mei-yu–baiu mode and changma mode are significantly related to the anomalous anticyclone over the western North Pacific (Figs. 6a and 6b), while those for the post-changma mode and dry-spell mode are associated with cyclonic circulation (Figs. 6c and 6d). Another interesting feature is the positive (negative) values over the central equatorial Pacific for the mei-yu–baiu mode and changma mode (dry-spell mode). It implies that the intraseasonal phases are somewhat connected to the tropics. The evidence of a relationship between tropical SST and extratropical intraseasonal phases can also be found in Figs. 6e–h. Although the mei-yu–baiu mode and post-changma mode are not significantly correlated to the thermal condition over the equatorial eastern Pacific during the preceding winter, there is a distinct difference between the four major modes. The high occurrence years for the mei-yu–baiu mode and changma mode tend to be related with the El Niño–like pattern over the equatorial eastern Pacific during the preceding winter, while those for the postchangma mode and dry-spell mode exhibit a La Niña–like pattern. The thermal condition over the Kuroshio Extension region is rather significant for the mei-yu–baiu mode. Similar patterns with opposite features are identified in the changma mode and dry-spell mode. Significant positive (negative) SST can be found over the equatorial central Indian Ocean and eastern Pacific for the changma mode (dry-spell mode).
The seasonal evolution of the relationship between the equatorial eastern Pacific SST and each intraseasonal phase of the EASM is examined by lead–lag correlations of the four major modes. The interannual variability of these modes—mei-yu–baiu, changma, post-changma, and dry spell—is depicted by using the annual number of clustered days. A time series of the seasonal mean Niño-3 index from a lead time of DJF to a lag time of July–September (JAS) was prepared to calculate the correlation coefficient (CC). Figure 7 shows the lead–lag CC between the Niño-3 index and the four major modes. The mei-yu–baiu and changma modes, which occur in early summer, positively correlate with the eastern Pacific SST during the preceding winter. This relationship is maintained until the following spring, but it is not significant after March–May (MAM). For all lead–lag periods from DJF to JAS, the CC between the Niño-3 index and the changma mode is twice that for the mei-yu–baiu. On the contrary, the postchangma and dry-spell modes, which occur in later summer, show negative correlation with the preceding winter-to-spring Niño-3 index. CCs for these modes start with a similar value as that in DJF and the same as that in February–April (FMA). Although an abrupt decline is shown in the dry-spell mode, the CC for the postchangma mode is more persistent. The opposite lead–lag CCs among the four major modes indicate that particular monsoon phases are favorable to ENSO.

b. ENSO impacts on intraseasonal phases

In the previous sections, we demonstrated that particular monsoon phases are favored by preceding ENSO events. The reason why the response of the four intraseasonal modes to ENSO is not linear is fundamentally due to the nonlinear atmospheric response to warm and cold ENSOs. The composite circulation fields or the four major modes for the years with the preceding wintertime El Niño and La Niña show an asymmetric structure and demonstrate the nonlinear relationship between intraseasonal phases and ENSO.


The mean number of days per event clustered during ENSO years for the four major modes is shown in Fig. 8. Of the 92 clustered days in one year (in JJA) 16.9, 13.6, 16.7, and 11.6 were for the mei-yu–baiu, changma, post-changma, and dry-spell modes, respectively. It is interesting to note that the mean number of days for the El Niño events increased in the mei-yu–baiu and changma modes by 38% and 45%, respectively. On the contrary, these results decreased in the postchangma and dry-spell modes by 32% and 58%, respectively, which indicates that the mei-yu–baiu and changma modes are favored by the preceding winter El Niño. However, the La Niña association appears to be different. No specific preference for the mei-yu–baiu or postchangma mode is indicated by a preceding La Niña event, although the changma mode (dry-spell mode) tends to occur less (more) frequently through winter equatorial eastern SST cooling. Thus, indications on the modulation of variation by external components such as ENSO could aid prediction of the nonlinear monsoon precipitation intraseasonal oscillation (ISO) over East Asia.

5. Summary and conclusions

Nonlinear variability of monsoon rainfall creates difficulties in predicting the intraseasonal precipitation of the EASM. We hypothesized that the summer monsoon intraseasonal phases are a convectively coupled oscillation, and hence, it should be possible to identify the phases of rainfall oscillation by using large-scale circulation parameters. However, the relationship between rainfall and circulation is nonlinear; therefore, an effective method for isolating the commonality among parameters is necessary such that various phases of nonlinear convectively coupled intraseasonal oscillation are detected. For this reason, we adopted a nonlinear pattern recognition algorithm known as self-organizing mapping (SOM) in this study. Unlike linear techniques, SOM is capable of identifying various intraseasonal
phases of the EASM, including their evolutionary histories. This advantage of the SOM will provide extended-range prediction of intraseasonal monsoon precipitation.

We used six large-scale circulation indices to describe the intraseasonal phases of the EASM (Table 1). The daily large-scale dynamical indices used as the SOM algorithm input parameters demonstrate that it captures the temporal evolution and spatial patterns associated with different intraseasonal phases of the monsoon rainfall (Fig. 2). This result proves the strength of the SOM technique in isolating the nonlinear coupled states and establishes that the monsoon intraseasonal phases are a nonlinear coupled oscillation. On the basis of an analysis of the various SOM nodes, we identified four major intraseasonal phases of the EASM, which were positioned at the far corners of the SOM. The first node described a circulation state corresponding to weak tropical and subtropical pressure systems, weakened monsoonal winds, and cyclonic upper-level vorticity. However, the vertical wind shear was large with stronger westerly winds in the upper troposphere. This mode, which is related to large rainfall anomalies in southeast China and southern Japan, occurred several weeks prior to the onset of changma rains in Korea. Based on its various characteristics, we identified this mode as the mei-yu–baiu phase. The second node selected for this analysis represented the changma over Korea and occurred with a distinct circulation state corresponding to a strengthened subtropical high, monsoonal winds, and anticyclonic upper-level vorticity in southeast Korea. However, the vertical shear was weak with a weaker upper-level westerly associated with a weaker and northward-shifted subtropical jet stream. Another interesting feature is the presence of weakened tropical high pressure systems extending from the South China Sea into the Philippines. The third node is related to copious rains over Korea, which we termed the postchangma phase. Temporal analysis showed that this phase occurred after the changma season rains and the midsummer dry period. The fourth node was diagonally opposite of the changma mode and featured a mirror image of the circulation vector. Because Korea experienced a dry spell associated with this SOM node, we named it the dry-spell phase.

In addition, we considered the modulation of monsoon intraseasonal characteristics by external components such as ENSO to provide assistance in the prediction of intraseasonal monsoon precipitation.

**FIG. 8.** Mean annual number of days clustered at the four major nodes: number of days clustered at each node for El Niño years (La Niña years) is described with dark (light) gray bars. The left (right) number in the upper-left corner is an increased or decreased percentage of mean number of days for El Niño (La Niña) years compared to the total mean.
monsoon precipitation over East Asia. We discovered that the mei-yu–baiu mode and changma mode are favored by the preceding winter El Niño. However, a different La Niña association was apparent. No specific preference for the mei-yu–baiu mode or postchangma mode was detected by the preceding La Niña event, although the changma mode (dry-spell mode) tended to be less (more) frequent through winter equatorial eastern SST cooling. The results have great implications in improving the predictability of interannual variability, which is controlled by the nonlinear and chaotic monsoon intraseasonal oscillation.

Acknowledgments. This work was supported by the National Research Foundation of Korea (NRF) (Grant 2011-0021927, GRL), funded by the Korea government (MEST). Saji Hameed acknowledges support from APEC Climate Center through a collaborative research grant.

REFERENCES