An Objective Identification Technique for Regional Extreme Events

FUMIN REN,* DONGLIN CUI,*,+ ZHIQIANG GONG,* YANJIAO WANG,* XUKAI ZOU,*
YIPING LI,# SHIGONG WANG,+ AND XIAOLING WANG*

* *Lab for Climate Studies, China Meteorological Administration, Beijing, China*
\+ *College of Atmospheric Sciences, Lanzhou University, Lanzhou, China*
\# *Lanzhou Institute of Arid Meteorology, China Meteorological Administration, Lanzhou, China*

(Manuscript received 30 August 2011, in final form 5 April 2012)

ABSTRACT

An extreme weather and climate event does not only mean that an extreme occurs at an individual point (station), but more generally it has a certain impacted area and duration, which means that it is a regional extreme event (REE). How to identify a REE is the basis for studies in this area. An objective identification technique for REE (OITREE), which is based on the model of ''the string of candied fruits,'' is proposed in this study. This technique consists of five steps: to select a daily index for individual points (stations), to partition natural daily abnormality belts, to distinguish the event's temporal continuity, to establish an index system for regional events, and to judge extremity for regional events. In the index system developed specially for regional events, there are five single indices, namely extreme intensity, accumulated intensity, accumulated area, maximum impacted area and duration, as well as an integrated index and the spatial location. In this study, the proposed method was first applied to examine four types of REEs in China: heavy precipitation, drought, high temperature, and low temperature. Results show that the technique is skillful in identifying REEs, demonstrating the usefulness of the proposed method in detecting and studying of REEs and operational application.

1. Introduction

There has been notable progress in extreme events research in the recent three decades. In the early 1980s, Karl et al. (1984) first showed their interests in studies of extreme temperatures and diurnal temperature range. Subsequently, a number of studies, such as Karl et al. (1986, 1991) and Horton (1995), focused on temperature range trends, asymmetric diurnal temperature change, and changes in maximum and minimum temperatures. The research community's attention to precipitation extremes came later than that to temperature extremes, with one early example being a study of extreme precipitation in Japan by Iwashima and Yamamoto (1993). There has also been an increase in the number of observational studies on drought published in the last decade or so. However, drought index studies were commenced much earlier, with Penman (1948) developing an aridity index and Palmer (1965) proposing the Palmer drought severity index (PDSI).

Past achievements in extreme indices might be summarized in two aspects: 1) the 27 indices for temperature and precipitation extremes, which were recommended by the Commission for Climatology (CCl)/Climate Variability and Predictability (CLIVAR)/the World Meteorological Organization–Intergovernmental Oceanographic Commission (WMO–IOC) Joint Technical Commission for Oceanography and Marine Meteorology (JCOMM) Expert Team on Climate Change Detection and Indices (ETCCDI) (http://cccma.seos.uvic.ca/etccdi/index.shtml); and 2) several commonly used drought indices including PDSI.

A further analysis of the above existing indices and indicators suggests that almost all of them only refer to extremes of individual points (stations or sometimes grid points). It is well known that extreme weather and climate events, such as the 2003 boreal summer heat wave in Europe and the 2009/10 drought in Southwest China, are generally regional phenomena, which involve certain impacted areas and durations. Then, the question before us is how to identify and characterize regional extreme events?

In the past 10 years, attention to regional extreme events increased gradually. Shiau and Shen (2001) noted that a useful index for a drought event is a joint indicator

---

Corresponding author address: Dr. Fumin Ren, Lab for Climate Studies/CMA, Beijing 100081, China.
E-mail: fmren@163.com

DOI: 10.1175/JCLI-D-11-00489.1

of its duration and severity. Biondi et al. (2005, 2008) proposed a new stochastic model for quantifying climate episodes in long time series with three random variables: duration, magnitude, and peak value. Zhang and Zhou (2006) and Wang et al. (2007) proposed concepts of (natural) drought process for individual points. Min and Qian (2008) and Huang and Qian (2009) analyzed regional characteristics and persistence features of temperature and precipitation extremes. Andreadis et al. (2005), Sheffield and Wood (2007), and Sheffield et al. (2009) analyzed severe regional drought events using monthly soil moisture data from a simulation of the terrestrial water cycle with the variable infiltration capacity (VIC) land surface model. Tang et al. (2006), Qian et al. (2011) and Ding and Qian (2011) tried to identify different regional extreme events using different definitions which were mainly subjective. At a workshop organized by the World Climate Research Programme (WCRP) (Zolina et al. 2011), some advanced metrics for regional extreme weather and climate extremes were discussed.

Based on the above review, it is very clear that the key issue for studying regional extreme events is to develop methods/techniques to identify the regional extreme events. However, with a large variety of regional extreme events, e.g., drought, heavy precipitation, and high and low temperatures, how can we find out something common in the processes and structures of these extreme events to come up with an objective approach to identify them?

To answer this question, we focused on methodologies for objectively identifying regional extreme events. After several years of explorative study (Ren et al. 2010), an objective approach for identification of regional extreme events has been developed, which is initially applied to studies of regional extreme events in China. This paper attempts to document the method and sum up the findings.

2. Research ideas

Our preliminary analysis suggests that different kinds of regional extreme events have the following common structural characteristics: they are regional events which affect a certain domain and lasting for a specific time period. Therefore, identification of extreme events first calls for identification of a regional event with a certain extent of impact and a specific duration, that is, a common regional event.

An evolving regional event has two structural characteristics: 1) a regional event always has a given duration (days); and 2) on a daily anomaly chart of climate elements (e.g., temperature anomaly) during the duration, every day the regional event may cover a specific area, namely a certain impacted area. When the entire evolution is seen as a whole, it is found that the evolutionary process is strikingly similar to Tanghulu-chuan, which is a popular traditional winter snack in northern China and means a string of candied fruits, or a string of dried figs in some parts of the west! As shown in Fig. 1, the left panel is the model of a candied fruit string, and the right panel shows the evolution process of a regional event which represents the 2009/10 Southwest China drought—a string of daily impacted areas, where each ‘‘candied fruit’’ is equivalent to a daily impacted area, and when all daily impacted areas were strung together, a complete regional event is identified.

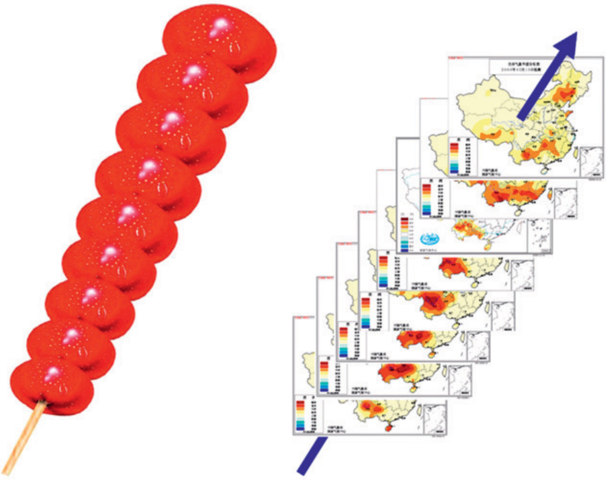

FIG. 1. A sketch of the variations of a regional event: the model of (left) a string of candied fruits and (right) a string of daily impacted areas that represents the 2009/10 Southwest China drought.

The above analysis also shows that to identify a regional event, it is necessary to address two key techniques. The first one is how to separate natural daily abnormality belts. With this technique, the separation of daily impacted areas can be achieved. The second is how to identify the temporal continuity of an event. With the latter, the daily natural abnormality belts can be reasonably put together into a ‘‘string.’’

After the identification of regional events, an index system is also required to be set up to evaluate each event and then reasonably identify the extreme events, which are then defined as regional extreme events.

Now it is important to draw a common definition of a regional event in this study. A regional event is a weather and climate event that has a certain impacted area and a specific duration, derived from a string of daily impacted areas. It is easy to understand that regional events can vary greatly in intensity, impacted area, or duration. For a large area such as China, from the point of the whole region, only regional events that have large values of intensity, impacted area, or duration can be called a China-wide (or China's) regional event, while events which have small values of intensity, impacted area or duration can be

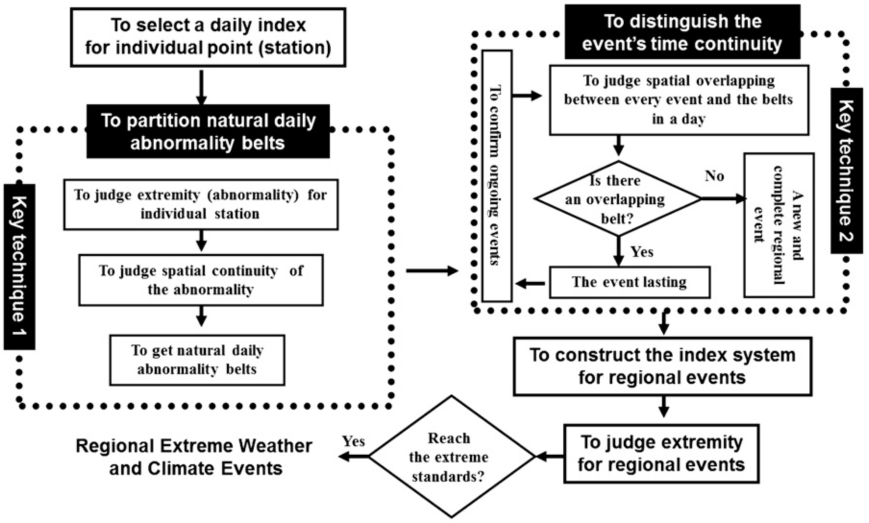

FIG. 2. The general flowchart for the objective identification technique of regional extreme weather and climate events (OITREE).

treated as weak events for China. However, some events which are weak in a national context may be significant over smaller areas; for example, in a subregion such as Guizhou province, a weak event for China may still be a significant regional event for Guizhou.

3. The objective identification technique

Based on the above idea, together with the natural rainbelt separating technique in partitioning tropical cyclone rainfall (Ren et al. 2001, 2007), a general flowchart for the objective identification technique for regional extreme events (OITREE) is developed (Fig. 2). The technique includes five steps: 1) to select a daily index for each individual point (station), 2) to partition natural daily abnormality belts, 3) to distinguish the event's temporal continuity, 4) to construct the index system for regional events, and 5) to judge extremity for regional events.[1] For each regional extreme event considered, step 1 is to select a suitable daily index that can be an element or a special index for an individual point (station). In step 2, based on the definition of abnormality occurring at each station and spatial structure analysis of the abnormality, natural daily abnormality belts can be partitioned every day during the time period concerned. In step 3, by comparing the spatial distributions of natural daily abnormality belts between adjacent days, the event's temporal continuity can be determined, and then a regional event can be constructed. Based on the above steps, regional events can be identified for the time period concerned. Then, step 4 is to construct the index system for regional events, which includes five single indices and an integrated index. Finally, according to the distribution of the integrated index, regional extreme events can be defined as those with a large value of the integrated index beyond a certain threshold.

Among these steps, the steps to partition natural daily abnormality belts and to distinguish each event's temporal continuity are the two key techniques. Partitioning natural daily abnormality belts (hereinafter referred to as "abnormality belts") starts from a structural analysis of the abnormality distribution. Then, the abnormality distribution can be separated into different natural daily abnormality belts. This step includes three substeps–to judge abnormality for individual stations, to judge spatial continuity of the abnormality and to get natural daily abnormality belts. Meanwhile, distinguishing the event's temporal continuity starts from the analysis of the spatial distributions of different natural daily abnormality belts between adjacent dates by analyzing their spatial overlap to evaluate the temporal continuity of an event, which is further divided into three substeps—to confirm ongoing events, to judge spatial overlap between every ongoing event and the belts in each day, and to identify the ongoing event's temporal continuity.

The index system for regional events, which consists of five single indices and an integrated index, is required to be specifically established. The five technical steps are described in more detail as follows:

a. To select a daily index for individual points (stations)

In case of a particular regional event considered, first it is necessary to select an appropriate index for

---

[1] In this paper abnormality is defined as the quality or state of being abnormal, generally surpassing a certain threshold based on statistics.

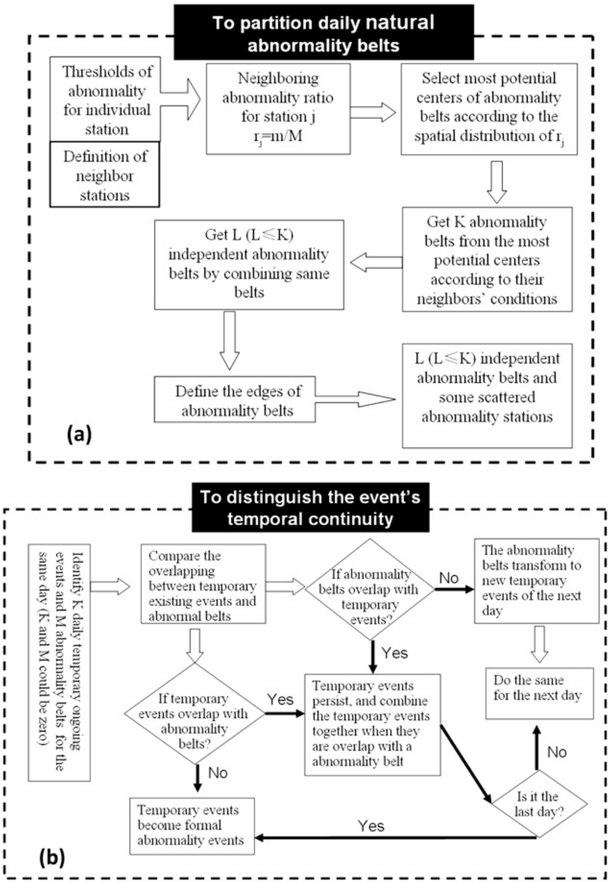

FIG. 3. The flowcharts for the two key techniques of OITREE: (a) the first one—to partition natural daily abnormality belts, and (b) the second one—to distinguish the event's temporal continuity.

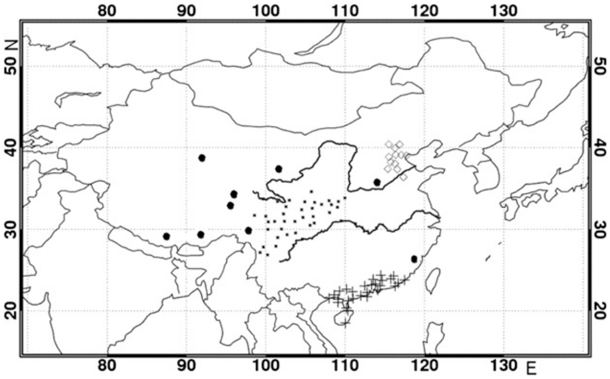

FIG. 4. An example of partitioning natural daily abnormality belts. Three independent abnormality belts were represented by different symbols of "◇", "■," and "+," while the symbol of "●" represents scattered abnormality station.

individual stations. Usually it may be selected from a commonly used set of indices. For a regional drought event, a daily composite meteorological drought index (e.g., CI or PDSI) can be chosen. For a regional high-temperature event, daily maximum temperature may be selected. In addition, for some special regional events such as ice storms, dust storms, or heavy snow, other indices may be used, or new indices for individual stations can be derived when necessary.

b. To partition natural daily abnormality belts

Figure 3a shows the flowchart for partitioning natural daily abnormality belts. Based on selection of a daily index (T) for individual points (stations), thresholds (T_t) of abnormality for individual stations are defined, for example, 90% (or 10%) percentile or other values. Generally, to ensure an adequate continuity in identifying temporal continuity for a regional event, the thresholds need not be very strict.

For station j, if T_j at the station is higher (or lower) than the threshold T_{tj} in a given day, it means abnormality occurs at the station on the day. Suppose distribution of abnormality for the day is shown by all the stations in Fig. 4.

The detailed analysis process to partition natural daily abnormality belts is presented as follows:

For each station j ($j = 1, 2, \ldots, N$), the first step is to define its neighbor stations, which refer to those within the fixed distance (e.g., 200 km) to the candidate station. The fixed distance needs to be defined carefully to make sure every station has at least one neighbor station. Based on this definition, the following analysis can be done.

1) TO CALCULATE NEIGHBORING ABNORMALITY RATIO

For station j with abnormality, its neighboring abnormality ratio is as follows:

$$r(j) = m/M, \quad j = 1, \ldots, n, \tag{1}$$

where M and m represent the numbers of its neighbor stations and those with abnormality, respectively. Obviously $r(j)$ varies between 0.0 and 1.0. For a station j without abnormality, $r(j)$ is defined as zero.

2) TO SELECT POTENTIAL CENTERS OF ABNORMALITY BELTS

For station j with abnormality, when $r(j)$ is greater or over a certain threshold value R_0, it would belong to a certain abnormality belt.

Therefore, this process includes three steps: 1) $r(j)$ is ranked in descending order from the largest to the smallest value; 2) the station with the largest $r(j)$ is selected as the first potential center of abnormality belt (hereinafter refer to as an abnormality belt center); and 3) the remaining ($N-1$) stations are scrutinized, and any station will be defined as an abnormality belt center, only when it satisfies the following two conditions:

$$r(j) > R_0 \tag{2}$$

and

$$d > d_c, \tag{3}$$

where d is the minimal distance between station j and any one of the previously selected abnormality belt centers, and d_c is a constant value (e.g., 300 km). To help identify all the independent abnormality belts, R_0 should not be too large and usually stays between 0.3 and 0.5. Suppose K abnormality belt centers can be identified with the above method.

3) TO DEFINE MAJOR FEATURES OF ABNORMALITY BELTS

The following steps are performed for each station of the K most potential centers of abnormality belts.

Step 1: The station may be regarded as a new abnormality belt l, only if it does not fall into any of already identified abnormality belts. Otherwise, the same process applies to the next abnormality belt center.

Step 2: If the station j belongs to the abnormality belt l, then any one (j_0) of its neighbor stations that is not affiliated to any of the already identified abnormality belts belongs to the abnormality belt l if $r(j_0) \geq R_0$.

Step 3: For those stations that have just been identified as part of the abnormality belt l, step 2 is repeated until no new neighbor station which meets the conditions can be identified; then return to step 1.

It is assumed that L ($L \leq K$) independent abnormality belts can be separated by the above method.

4) TO DEFINE THE EDGES OF ABNORMALITY BELTS PRELIMINARILY

This process only applies to those abnormality stations that do not fall into any of the L abnormality belts. For each such station, the following 2 steps are taken.

Step 1: Calculate the numbers of its neighbor stations that belong to each of the L abnormality belts: $N_c(1), N_c(2), \ldots, N_c(L)$.

Step 2: Define $I_{\max}$ as the value of I which gives the maximum value of $N_c(I)$. The station is then considered to belong to the abnormality belt $l_{\max}$. If $N_c(I)$ is zero for all values of I, the station is a scattered abnormality station that does not belong to any abnormality belt.

5) TO FINALLY DEFINE THE EDGES OF ABNORMALITY BELTS

Based on preliminarily defining the edges of the L abnormality belts, to get more reasonable edges, the process described in paragraph "4)" can be repeated once or several times.

At this point, the L independent abnormality belts (e.g., as shown in Fig. 4, three abnormality belts with different symbols) and some scattered abnormality stations (as shown in Fig. 4 by the symbol ●) have been separated successfully.

c. *To distinguish the event's temporal continuity*

From the above analysis, a regional event can be treated as "a string of daily abnormality belts", including both beginning and ending dates as well as information about daily abnormality belts during the event. Figure 3b shows the flowchart for the second key technique of OITREE—to distinguish the event's temporal continuity. The key technique includes four steps.

1) TO IDENTIFY DAILY ONGOING EVENTS AND ABNORMALITY BELTS

From the first day of the analysis period, this analysis needs to go on every day. When the first day which has abnormality belt(s) appears, the abnormality belt(s) will be treated as ongoing event(s) for the next day(s), and the analysis goes on.

For any given day, the ongoing events and the abnormality belts need to be defined. The ongoing events are considered as the events which were taking place on the previous day. The reason for calling them "ongoing events" is that the previous day is temporarily taken as an ending date for those events. On any given day, if the number of ongoing events K carrying over from the previous day is zero for that day, then all the abnormality belts will become new ongoing events for the next day; if the number of abnormality belts M is zero, then all the existing ongoing events will be finalized as formal events; if both K and M are zero on a day, then the identification process will continue to the next day.

2) TO IDENTIFY THE OVERLAP BETWEEN DAILY ONGOING EVENTS AND ABNORMALITY BELTS

Whether there exists an overlap between a daily ongoing event and an abnormality belt can be identified by comparing the spatial distributions of the abnormality belt and the impacted area of the ongoing event in the previous day. If there exists a common area between a daily ongoing event and an abnormality belt, which means overlap exists.

3) TO JUDGE THE ONGOING EVENT'S TEMPORAL CONTINUITY

Based on identification of the overlap between daily ongoing events and abnormality belts, it can be determined that an ongoing event continues when it overlaps with a (or

TABLE 1. Index system for regional weather and climate events. Note that these indices are classified into three grades: the grade I indices are for the whole event process, grade II indices provide daily extreme values, while grade III indices present station extreme values. Here, K is the number of days for duration, and J is the number of impacted stations for the whole event process, while J_k is the number of impacted stations for day k and S_k is the distribution of the J_k impacted stations and Area(S_k) is the area of S_k. Here, T_{ki} is the value of the daily index for station i on day k, and $T_{ki}|_c$ is its threshold. Meanwhile, T_{kj} and $T_{kj}|_c$ are similar to T_{ki} and $T_{ki}|_c$, respectively, when from the point of the number of impacted stations for the whole event process.

Index's name		Grade I (for event)	Grade II (for day k)	Grade III (for station j)
Single indices	Extreme intensity (I_1)	$I_1 = \underset{k=1,K}{\mathrm{Max}}(I_{1k})$	$I_{1k} = \underset{i=1,J_k}{\mathrm{Max}}(T_{ki})$	$I_1\|_j = \underset{k=1,K}{\mathrm{Max}}(T_{kj})$
	Accumulated intensity (I_2)	$I_2 = \sum_{k=1}^{K} I_{2k}$	$I_{2k} = \sum_{i=1}^{J_k}(T_{ki} - T_{ki}\|_c)$	$I_2\|_j = \sum_{k=1}^{K}(T_{kj} - T_{kj}\|_c)$
	Accumulated impacted area (A_s)	$A_s = \sum_{k=1}^{K} A_k$	$A_k = Area(S_k)$	
	Maximum impacted area (A_m)	$A_m = \mathrm{Area}\left(\bigcup_{k=1}^{k} S_k\right)$		
	Lasting period (duration; D)	$D = K$		
Intergrated index (integrated intensity; (Z)		$Z = F(I_1, I_2, A_s, A_m, D)$ Scheme: to standardize I_1, I_2, A_s, A_m, and D first, and then weighted sum	$Z_k = f(I_{1k}, I_{2k}, A_k)$ Scheme: using the parameters in the formulation Z to standardize I_{1k}, I_{2k}, and A_k first, and then weighted sum	
Spatial location		1) distributions of station extreme intensity ($I_1\|_j$) and station accumulated intensity ($I_2\|_j$) 2) maximum impacted area and its geometric center, the centers of gravity with weights of $I_1\|_j$ and $I_2\|_j$	Distributions of daily impacted area and its geometric center	

several) given abnormality belt(s), and its impacted area for the current day is the area (or the sum of the areas) of the given abnormality belt(s); when it does not overlap with any abnormality belt, it means that the ongoing event has ended, and it is finalized as a formal event.

When multiple ongoing events overlap with an abnormality belt on any given day, these events should be combined together. When an abnormality belt does not overlap with any ongoing event, it becomes a new ongoing event for the next day.

4) Special treatment for the last day

When above analysis proceeds to the final day under study, all ongoing events are finalized as formal events.

It should be noted that, for some kinds of regional events (e.g., regional heavy precipitation), short (e.g., 1 ~ 2 days) interruptions are allowed to occur during an event process. In this case, for an ongoing event in the analysis, its ending date is temporarily taken as the last date on which the ongoing event has an impacted area.

d. To establish an index system for regional events

An index system for regional weather and climate events as shown in Table 1 is proposed according to an analysis of the features of regional events. The index system is as follows.

1) Single indices

To characterize regional events in different aspects, first five single indices that capture different characteristics of a regional event were proposed. The five single indices are extreme intensity (I_1), accumulated intensity (I_2), accumulated impacted area (A_s), maximum impacted area (A_m), and duration (D). The five formulae are as follows:

$$D = K \tag{4}$$

where K is the duration of a regional event in days.

$$A_s = \sum_{k=1}^{K} A_k = \sum_{k=1}^{K} \mathrm{Area}(S_k), \quad \text{and} \tag{5}$$

$$A_m = \mathrm{Area}\left(\bigcup_{k=1}^{K} S_k\right), \tag{6}$$

where S_k is the distribution of the J_k impacted stations for day k and Area(S_k) means the area of S_k.

$$I_1 = \underset{k=1,K}{\mathrm{Max}}\,[\underset{i=1,J_k}{\mathrm{Max}}\,(T_{ki})], \quad \text{and} \tag{7}$$

$$I_2 = \sum_{k=1}^{K}\sum_{i=1}^{J_k}(T_{ki} - T_{ki}|_c), \tag{8}$$

where J_k is the number of impacted stations for day k and T_{ki} is the value of the daily index for station i on day k, while $T_{ki}|_c$ is the threshold of T_{ki} for defining abnormality.

At the same time, these indices are classified into three grades: the first-grade indices are for the whole event process, including process extreme value, accumulated process intensity, accumulated process impacted area, maximum process impacted area, and duration. The second-grade indices provide a daily extreme value (I_{1k}), accumulated daily intensity (I_{2k}) and daily impacted area (A_k):

$$I_{1k} = \underset{i=1,J_k}{\mathrm{Max}}\,(T_{ki}), \quad \text{and} \tag{9}$$

$$I_{2k} = \sum_{i=1}^{J_k}(T_{ki} - T_{ki}|_c). \tag{10}$$

The third-grade indices present extreme values for a station during the duration of the event, including station extreme intensity ($I_1|_j$) and station accumulated intensity ($I_2|_j$):

$$I_1|_j = \underset{k=1,K}{\mathrm{Max}}\,(T_{kj}), \quad \text{and} \tag{11}$$

$$I_2|_j = \sum_{k=1}^{K}(T_{kj} - T_{kj}|_c), \tag{12}$$

where T_{kj} and $T_{kj}|_c$ are defined analogously to T_{ki} and $T_{ki}|_c$, respectively, but considered from the perspective of the number of impacted stations for the whole event process.

2) The integrated index

The integrated index, that is, the integrated intensity, is defined on the basis of the single indices. The integrated index can be classified into two grades.

The first-grade integrated index is for the whole event process and it was developed as a function of the five single indices. After a number of tests and comparisons, a calculation named as "standardization weighted summation" was chosen as the best scheme, namely, weighted I_1, I_2, A_s, A_m, and D is summed after they are standardized respectively. The formula is as follows:

$$Z = F(I_1, I_2, A_s, A_m, D) = e_1 I_1^{\%} + e_2 I_2^{\%} + e_3 A_s^{\%} + e_4 A_m^{\%} + e_5 D^{\%}, \tag{13}$$

where $I_1^{\%}$, $I_2^{\%}$, $A_s^{\%}$, $A_m^{\%}$ and $D^{\%}$ are standardized I_1, I_2, A_s, A_m and D, respectively, and e_1, e_2, e_3, e_4 and e_5 are their coefficients. Undoubtedly, the integrated index is constructed based upon research interest and purpose, so

TABLE 2. Values of the parameters of the OITREE method for regional heavy precipitation events in China.

Parameter's name	Code	Meaning	Value
The daily index for individual points (stations)	T	It depends on what kind of regional event concerned and can be an element or a special index.	Daily precipitation
Threshold for the daily index	T_t	Abnormality occurs at individual points (stations) only when T is beyond T_t.	The 95% percentile of T
Threshold for the neighbor station distance	d_0	For a point (station) all the stations within d_0 are defined as its neighbor stations.	250 km
Threshold for the neighboring abnormality ratio	R_0	An abnormality station can be defined as an abnormality belt center if its ratio bigger than R_0.	0.3
Threshold for number of abnormality stations in a belt	M_0	If number of abnormality stations in a belt is not smaller than M_0, the belt is an abnormality belt.	5
Maximum length of time for an interruption in an event	M_{gap}	Interruptions which are not longer than M_{gap} are allowed to occur during an event process.	1 day
Direction indicator	Idirec	If T being bigger than T_t means abnormality, Idirec equals "1," otherwise Idirec equals "−1."	1
The five weight coefficients for grade I integrated index	e_1, e_2, e_3, e_4, and e_5	The five weight coefficients in the formula $Z = F(I_1, I_2, A_s, A_m, D) = e_1 I_1^{\%} + e_2 I_2^{\%} + e_3 A_s^{\%} + e_4 A_m^{\%} + e_5 D$.	0.04, 0.27, 0.26, 0.223 and 0.20
The index and its threshold for defining a regional event for a whole specific region	Integrated intensity, Z_0	An event can be defined as a regional event for the whole specific region if Z is bigger than Z_0, otherwise the event is only a weak event for the whole specific region.	Z, −0.3
Thresholds for classifying the regional events for the whole specific region	Z_1, Z_2, Z_3	The three thresholds can be applied in defining the four intensity categories by proportion 10% (extreme, $Z \geq Z_1$), 20% (severe, $Z_1 > Z \geq Z_2$), 40% (moderate $Z_2 > Z \geq Z_3$), and 30% (slight, $Z_3 > Z$).	1.1, 0.18, −0.48

how to gain the five weighting coefficients is a very important step. It is suggested that an objective method, such as considering the correlation between the specific research problem and each single index, be applied in this step. To gain such an integrated index Z such that the greater the intensity of the event the greater the Z value, the five coefficients, e_1, e_2, e_3, e_4 and e_5 need to be carefully determined, to make sure the five products—$e_1I_1^{\%}$, $e_2I_2^{\%}$, $e_3A_s^{\%}$, $e_4A_m^{\%}$ and $e_5D^{\%}$—have the same nature: the larger the product value, the greater the Z value.

The second-grade category index refers to daily integrated intensity:

$$Z_k = f(I_{1k}, I_{2k}, A_k) = e_1I_{1k}^{\%} + e_2I_{2k}^{\%} + e_3A_k^{\%}. \tag{14}$$

Its scheme follows the standardization weighted average summation, and it is suggested that its coefficients and parameters for standardization be the same of these for the first-grade category.

3) SPATIAL LOCATION

The spatial location can tell where a regional event occurs and is also an indispensable important parameter for the event. The spatial location is generally divided into two kinds—spatial distribution and center point location.

The spatial distribution includes spatial distribution of the event's maximum impacted area, spatial distribution of station extreme intensity ($I_1|_j$), and spatial distribution of station accumulated intensity ($I_2|_j$) within the maximum impacted area.

In addition, it is sometimes important to use a point location parameter, namely the center point location, to represent a regional event. Potential definitions of the center point location include the geometric center of the event's maximum impacted area, and the centers of gravity with weights of station extreme intensity and station accumulated intensity.

e. To judge extremity for regional events

Upon completion of the above steps described in 3a–3d, extreme events can be identified from all detected weather and climate events for a specific region. This step includes two substeps.

Step one is to define regional events for a certain region. As the OITREE method can identify all the events occurring in the region including these that may last only one day or cover a much smaller area when compared to the entire region interested, generally it is important and necessary to define only those stronger events as regional events for the whole specific region. The way to do this is suggested as following: according to the specific region interested and the distribution of an index from the five single indices and the first-grade integrated index Z, a threshold of the index can be determined empirically and only an event with a large value that exceeds that threshold can be defined as a regional event for the whole specific region, to exclude those which are weaker and can be called "weak events" for the whole specific region but may be a regional event for a smaller region.

Step two is to identify regional extreme events. According to the distribution of the integrated index, three thresholds can be specified according to the four intensity categories by proportion 10% (extreme), 20% (severe), 40% (moderate), and 30% (minor). Then the top events (e.g., top 10%) of the selected regional events are defined as regional extreme weather and climate events.

4. Application for regional weather and climate events in China

The OITREE technique was applied in a study of four types of regional extreme events—regional heavy precipitation events, meteorological drought events, high temperature events, and low temperature events—in China during 1961–2010.

a. Regional heavy precipitation events

To help understand the application of the OITREE method, Table 2 provides the values of the parameters of the method for identifying regional heavy precipitation events in China. At step 1, daily precipitation was selected as the daily index for individual points (stations). There are four parameters in step 2: the threshold for the daily index is set to the 95th percentile of the data during 1981–2010, while threshold for the neighbor station distance, threshold for the neighboring abnormality ratio and threshold for number of abnormality stations in a belt are 250 km, 0.3 and 5 respectively. In step 3, maximum length of time for an interruption in an event is set to 1 day, whereas normally its value would be generally set to zero, which means no interruption allowed. There are several parameters in step 4: direction indicator is "1" because of the nature of the daily index, while the five weight coefficients for the grade I integrated index are 0.04, 0.27, 0.26, 0.23, and 0.20 respectively. In addition, the three weight coefficients for the grade II integrated index are set to 0.04, 0.27, and 0.26, which are the same as the first three of the five weight coefficients for grade I integrated index. In step 5, the index and its threshold for defining a regional event for the whole of China are integrated intensity and −0.3 respectively, while the thresholds for classifying the regional events for the whole of China are 1.1, 0.18, and −0.48 respectively.

All the above parameters can be set only after substantial additional analysis such as experiential values, tests, and special calculations and may be varied depending on the particular application of interest. For example, the five weight coefficients are defined according to the ratio of correlation coefficients between the area of farmland affected by floods (independent data to this study) and the five single indices. For the index and its threshold for defining a regional event for the whole China, we chose the integrated intensity (Z) and the value -0.3 after some comparison tests with single indices and analyzing the distribution of the integrated intensity (Z).

According to the results, the top three extreme regional precipitation events in China during the past 50 years were the following. 1) Number 1 event lasted 15 days from 13–27 June 1998 occurring in the middle and lower reaches of the Yangtze River and the area south to it, with the integrated index being 8.19. 2) Number 2 event occurred in South China and southern part of central China during 12–20 June 1994, with the integrated index being 4.62. 3) Number 3 event occurred along the middle and lower reaches of the Yangtze River in southern central China during 24 June–2 July 1999, with the integrated index being 4.52. Ding (2008) indicated that extreme floods occurred with serious economic losses and a large number of deaths during each of the above three time periods and regions.

In addition, Fig. 5 shows results for the number 1 regional extreme heavy precipitation event in detail. Figure 5a is the distribution of the maximum impacted area of the event during 13–27 June 1998, while Figs. 5b–p display the distributions of the daily impacted areas during the period. It is easy to see that during the period the daily impacted area varies day by day and the whole event influenced a large area in the middle and lower reaches of the Yangtze River and the area south to it. The results look quite reasonable.

Figure 6a presents the variations of annual frequency of China's regional heavy precipitation events during 1961–2010. China experienced 363 regional heavy precipitation events in this period, which could be classified into 4 intensity categories by proportion 10%, 20%, 40%, and 30%, corresponding to the following: 36 extreme heavy precipitation events, 73 severe heavy precipitation events, 145 moderate heavy precipitation events, and 109 minor heavy precipitation events respectively. Figure 6a shows variations in annual frequency of China's regional heavy precipitation events. The variations suggest a weak upward trend, with obvious interdecadal variations including high frequencies in the late 1980s and 1990s.

b. Other regional events

154 regional drought events were identified in China during 1961–2010. The annual frequency of regional droughts has an obvious but insignificant upward trend. There are also obvious interdecadal variations with below-normal frequencies in the 1990s and above-normal frequencies in the 2000s (Fig. 6b). According to an analysis of the integrated index, the widespread drought event in Northern China in 1998/99 ranked top on record, lasting for about 230 days (from September 1998 to May 1999).

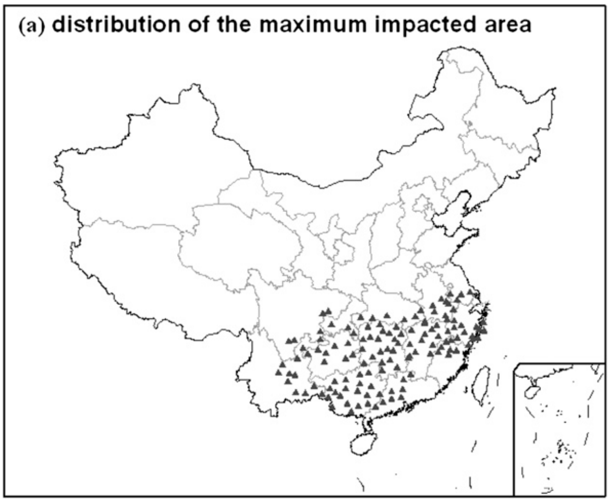

FIG. 5. Results for the number 1 regional extreme heavy precipitation event in China: (a) distribution of the maximum impacted area of the regional heavy precipitation event during 13–27 Jun 1998; and (b)–(p) distributions of the daily impacted areas from 13–27 Jun 1998, respectively.

During 1961–2010, China experienced 182 regional high temperature events, and the frequency of regional high temperature events displayed an obvious but insignificant upward trend, with obvious interdecadal variations of high frequencies in early 1960s, later 1990s, and 2000s (Fig. 6c).

There were 183 regional low temperature events in China during 1961–2010. The annual frequency has a significant decreasing trend. There are also obvious interdecadal variations with high frequencies in the period before 1987 and low frequencies in the period after 1987 (Fig. 6d).

5. Summary and discussions

Through the above analyses and discussions, the conclusions are given below.

1) An Objective Identification Technique for Regional Extreme Events (OITREE) has been developed, which consists of five components: (i) selecting a daily index for individual point (station); (ii) partitioning natural daily abnormality belts; (iii) distinguishing the event's temporal continuity; (iv) establishing an index system for regional events; and (v) judging

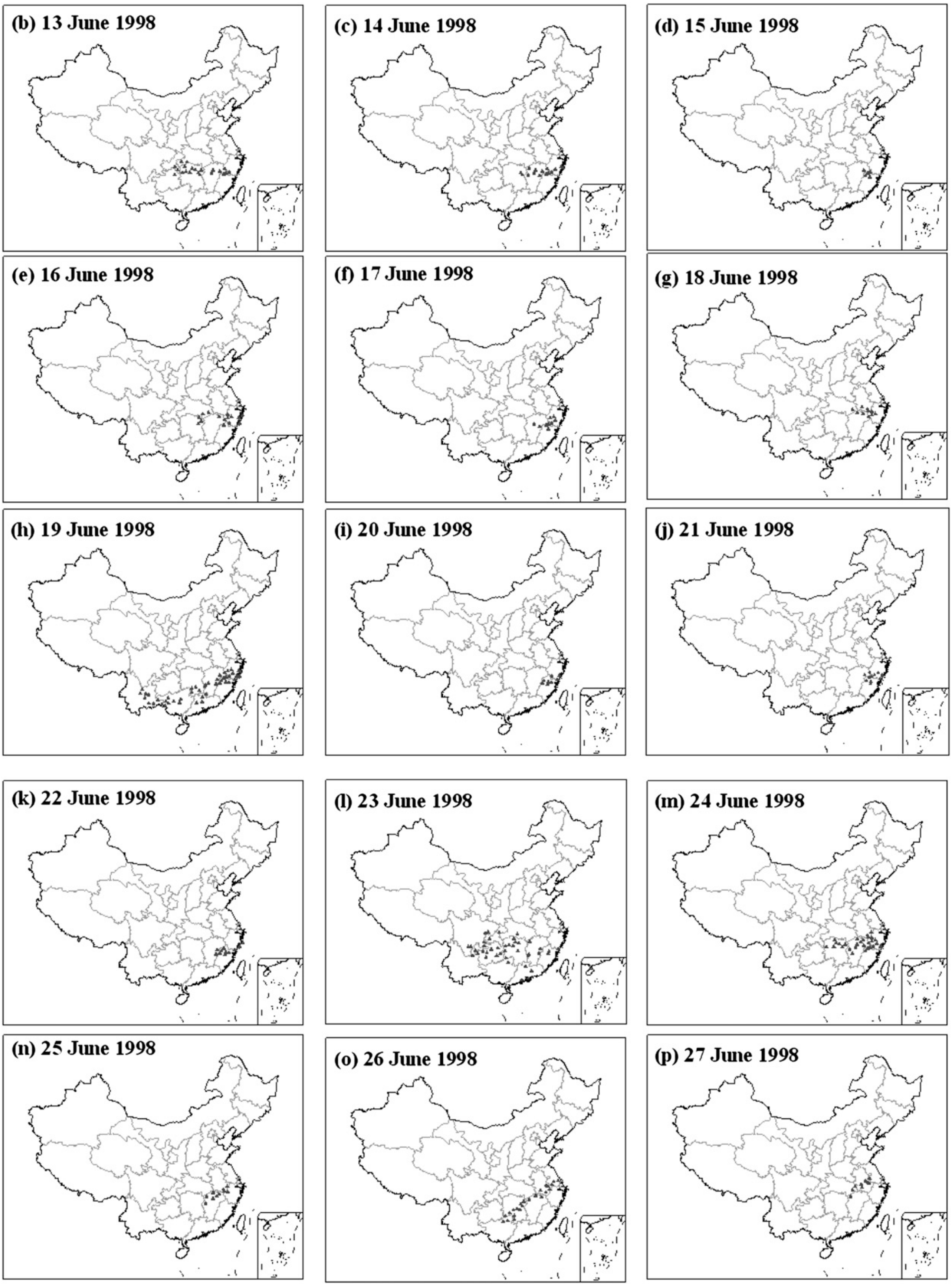

FIG. 5. (*Continued*)

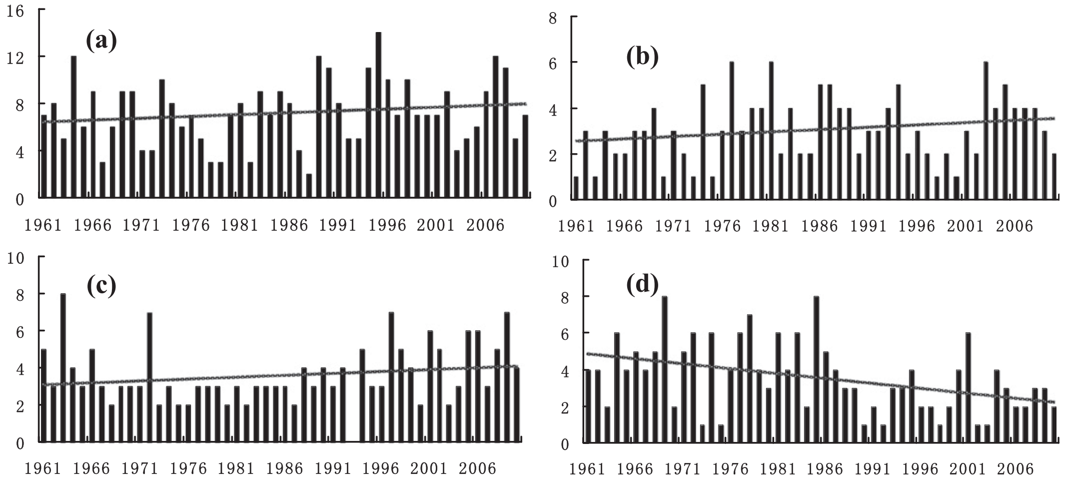

FIG. 6. Variations of annual frequencies of four types of China's regional extreme events during 1961–2010, with the lines representing the trends: (a) regional heavy precipitation events, (b) regional drought events, (c) regional high temperature events, and (d) regional low temperature events.

extremity for regional events. The core techniques are those described in (ii) and (iii); the component mentioned in (iv) was specially developed based on the features of regional events, which includes five single indices: extreme intensity, accumulated intensity, accumulated area, maximum impacted area and duration, as well as an integrated index.

2) OITREE shows skillful capability in identifying regional extreme events. It can objectively and automatically capture daily impacted areas of a regional event for its duration, and reasonably put them in a ''string'' to shape an entire regional event.
3) OITREE has preliminarily been applied to study four types of regional extreme events: drought, heavy precipitation, and high- and low-temperature events in China. The main conclusions are as follows: during 1961–2010, China experienced 154 regional drought events, 363 regional heavy precipitation events, 182 high temperature events, and 183 low temperature events. Increasing trends that are statistically insignificant exist in the annual frequencies of regional drought events, regional heavy precipitation events, and high temperature events, while a significant decreasing trend exists in the annual frequency of low temperature events.

However, some issues about the OITREE method need to be acknowledged or discussed. First, although station data are used in this study, the method can be applied to other forms of data. Gridded data, if available and of high quality, are potentially more suitable to be applied in the method. No matter what kind of data are used, it is important for the network density of the data to be approximately constant during the analysis time period to make results homogeneous, especially when station data are used. Second, even though the OITREE method in the form presented in this paper is generated from an analysis of daily data, the method could also be applied to data of other time scales such as weekly, monthly, seasonal, etc. Third, the OITREE method is broadly objective, but some additional analysis is required to determine suitable values for some parameters such as the distance for defining neighbor stations, the weighting coefficients of the single indices in the formula of the integrated index, the event thresholds, etc. Fourth, although the five single indices can characterize different aspects of regional events, some of them, especially the accumulated intensity (I_2) and the accumulated area (A_s), are not independent of the other indices. Fifth, the underlying index, namely the daily index for individual points (stations), is very important for the analysis. To gain an ideal result, the underlying index needs to be chosen carefully. Lastly, compared with some existing identification methods of regional extreme events, for example, Andreadis et al. (2005), Tang et al. (2006), Qian et al. (2011), and Ding and Qian (2011), the OITREE method has four distinctive features: (i) theoretical basing on the model of a string of candied fruits, (ii) being a common method suitable to identify almost all the kinds of regional extreme events, (iii) being more objective with more detailed introduction, and (iv) more easily being applied in climate operations with a special website of monitoring regional extreme events in China (http://cmdp.ncc.cma.gov.cn/extreme/dust.php?product=dust_moni&monicat=regional) in Chinese being near completion.

Study of regional extreme weather and climate events is becoming a new frontier of extreme weather and climate events study. The Commission for Climatology

(CCl) of the World Meteorological Organization (WMO) also gave more attention to research in this field, and has set up a task team on definitions of extreme climate events in 2010 (WMO 2010).[2]

This study suggests that an initial success has been achieved in developing an objective technique for identification of regional extreme events. Although there is still room for improvements in this technique, we believe that these findings could shine light on the research of regional extreme weather and climate events as well as their operational applications.

Acknowledgments. The authors would like to express their sincere thanks to Dr. Blair Trewin and another reviewer for the very helpful suggestions and comments to improve the manuscript greatly. Special thanks go to Prof. C. P. Chang, Prof. Liguang Wu, Prof. Deliang Chen, Prof. N.-C. Lau, Prof. Francis W. Zwiers, Prof. Manola Brunet, Dr. Tsz-cheung Lee, Prof. Jinhai He, Dr. Dushmanta Pattanaik, and Dr. Changzheng Liu for their great support and encouragements to this study. This work was supported by the National Natural Science Foundation of China (Grant 41175075) and the Chinese Ministry of Science and Technology Project (Grants N2010CB950501 and 2007BAC29B04).

REFERENCES

Andreadis, K. M., E. A. Clark, A. W. Wood, A. F. Hamlet, and D. P. Lettenmaier, 2005: Twentieth-century drought in the conterminous United States. *J. Hydrometeor.,* **6,** 985–1001.

Biondi, F., T. J. Kozubowski, and A. K. Panorska, 2005: A new model for quantifying climate episodes. *Int. J. Climatol.,* **25,** 1253–1264.

——, ——, ——, and L. Saito, 2008: A new stochastic model of episode peak and duration for eco-hydro-climatic applications. *Ecol. Model.,* **211,** 383–395.

Ding, T., and W. Qian, 2011: Geographical patterns and temporal variations of regional dry and wet heatwave events in China during 1960-2008. *Adv. Atmos. Sci.,* **28,** 322–337.

Ding, Y., 2008: *Encyclopedia of Meteorological Disasters in China–Zonghe* (in Chinese). Meteorological Press, 222 pp.

Horton, B., 1995: Geographical distribution of changes in maximum and minimum temperatures. *Atmos. Res.,* **37,** 101–117.

Iwashima, T., and R. Yamamoto, 1993: A statistical analysis of the extreme events: Long-term trend of heavy daily precipitation. *J. Meteor. Soc. Japan,* **71,** 637–640.

Karl, T. R., G. Kukla, and J. Gavin, 1984: Decreasing diurnal temperature range in the United States and Canada from 1941–80. *J. Climate Appl. Meteor.,* **23,** 1489–1504.

——, ——, and ——, 1986: Relationship between temperature range and precipitation trends in the United States and Canada from 1941–80. *J. Climate Appl. Meteor.,* **25,** 1878–1886.

——, and Coauthors, 1991: Global warming: Evidence for asymmetric diurnal temperature change. *Geophys. Res. Lett.,* **18,** 2253–2256.

Min, S., and Y. Qian, 2008: Regionality and persistence of extreme precipitation events in China (in Chinese). *Adv. Water Sci.,* **19,** 763–771.

Palmer, W. C., 1965: Meteorological drought. U.S. Weather Bureau Research Paper 45[R], 58 pp.

Penman, H. L., 1948: Natural evaporation from open water, bare soil, and grass. *Proc. Roy. Soc. London,* **A193,** 120–146.

Qian, W., X. Shan, and Y. Zhu, 2011: Ranking regional drought events in china for 1960–2009. *Adv. Atmos. Sci.,* **28,** 310–321.

Ren, F., B. Gleason, and D. Easterling, 2001: A numerical technique for partitioning cyclone tropical precipitation. *J. Tropical Meteor.,* **17,** 308–313.

——, Y. Wang, X. Wang, and W. Li, 2007: Estimating tropical cyclone precipitation from station observations. *Adv. Atmos. Sci.,* **24,** 700–711.

——, C. Donglin, G. Zhiqiang, W. Yanjiao, and Z. Xukai, 2010: Explorative study on objective identification technique for regional persistent extreme events. *Chinese Meteorological Administration Laboratory of Climate Study's 2009 Annual Conference Proceedings,* Chinese Meteorological Administration, 52–54.

Sheffield, J., and E. F. Wood, 2007: Characteristics of global and regional drought, 1950–2000: Analysis of soil moisture data from off-line simulation of the terrestrial hydrologic cycle. *J. Geophys. Res.,* **112,** D17115, doi:10.1029/2006JD008288.

——, K. M. Andreadis, E. F. Wood, and D. P. Lettenmaier, 2009: Global and continental drought in the second half of the twentieth century: Severity–area–duration analysis and temporal variability of large-scale events. *J. Climate,* **22,** 1962–1981.

Shiau, J. T., and H. W. Shen, 2001: Recurrence analysis of hydrologic droughts of differing severity. *J. Water Resour. Plan. Manage.,* **127,** 30–40.

Tang, Y., J. Gan, L. Zhao, and K. Gao, 2006: On the climatology of persistent heavy rainfall events in China. *Adv. Atmos. Sci.,* **23,** 678–692.

Wang, Z., X. Zhu, D. Liu, C. Yang, and Y. Gu, 2007: An aridity index based on the natural process of aridity and its application. *J. Nanjing Instit. Meteor.,* **30,** 134–139.

WMO, 2010: Report of the Meeting of the Management Group of the Commission for Climatology. WMO Rep. CCl/MG/2010, 61 pp. [http://www.wmo.int/pages/prog/wcp/ccl/mg/documents/MeetingReportFinal.pdf.]

Zhang, Q., and X. Zou, 2006: Classification of meteorological drought category. *GB/National Standards of China,* Standards Press of China, 1–17.

Zolina and Coauthors, 2011: WCRP (GEWEX/CLIVAR) and UNESCO (IHP) Metrics and Methodologies of Estimation of Extreme Climate Events Workshop. WCRP Informal/Series Rep. 9/2011, 39 pp. [Available online at http://www.wcrp-climate.org/documents/Report_9_2011_%20Paris.pdf.]

---

[2] Partly because of this study, Dr. Fumin Ren was appointed as the lead of the Task Team.