Asia-Pac. J. Atmos. Sci., 53(1), 51-61, 2017
DOI:10.1007/s13143-017-0004-9
pISSN 1976-7633 / eISSN 1976-7951

Multiple Aspects of Northern Hemispheric Wintertime Cold Extremes as Revealed by Markov Chain Analysis

Hye-Sil Kim[1], Yong-Sang Choi[1,2], Joo-Hong Kim[3], and WonMoo Kim[1,4]

[1]*Ewha Womans University, Seoul, Korea*
[2]*Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA*
[3]*Korea Polar Research Institute, Incheon, Korea*
[4]*APEC Climate Center, Busan, Korea*

(Manuscript received 17 March 2016; accepted 25 August 2016)

Abstract: High-impact cold extremes have continued to bring devastating socioeconomic losses in recent years. In order to explain the exposure to cold extremes more comprehensively, this study investigates multiple aspects of boreal winter cold extremes, i.e., frequency, persistence, and entropy (Markovian descriptors). Cold extremes are defined by the bottom 10th percentile of daily minimum temperatures during 1950-2014 over the northern hemisphere. The spatial and temporal distributions of Markovian descriptors during 65 years are examined. Climatological mean fields show the spatial coincidence of higher frequency, shorter persistence, and higher entropy of cold extremes, and vice versa. In regard to the temporal variations over six representative regions of North America, Europe, and Asia, all regions share a decreasing tendency of frequency with the increases in regional winter mean temperature. By contrast, persistence and entropy show their intrinsic decadal variability depending on regions irrespective of the regional temperature variability, which give different information from frequency. Therefore, the exposure to cold extremes would not simply decrease with regional warming. Rather these results indicate that the descriptors with multiple aspects of the extremes would be needed to embrace the topical features as well as the holistic nature of cold extremes.

Key words: Cold extremes, Markov chain analysis, multiple descriptors for extremes, decadal variations of cold extremes

1. Introduction

Extreme weather events such as torrential rainfall, deadly storms, and heat/cold waves have received continual attention. This is because these events are not only responsible for extreme socioeconomic losses, but also life-threatening (Easterling et al., 2000; Meehl et al., 2000; Kunkel et al., 2004; Zhang et al., 2005). Among them, the hot/cold extreme events and their changes under global warming have been one of the major issues of climate science because temperature itself is the most basic indicator of climate change. The latest reports by the Intergovernmental Panel on Climate Change (IPCC) indicated that the frequency of hot extremes is very likely to increase, and that of cold extremes (CEs) would decrease with global warming (IPCC, 2013). Although the reduction of CEs over the globe is probable, many people have still been killed by or suffered from extremely cold weather in recent years under a warmer global climate. For example, Eurasia experienced a severe, long-lasting cold spell in the winter of 2011/12, which was the worst at least within the last 26 years in Central and Eastern Europe, causing the deaths of more than 650 people in Eastern Europe (Blunden et al., 2013). Similarly, in North America, record-breaking frigid temperatures lasted for several days during the 2013/14 winter (Herrling et al., 2015). These unforeseen extreme events signify some important factors are missing in the current diagnosis of CEs.

Recent studies thereby attempted to investigate various characteristics of CE other than the frequency since the frequency of CEs solely is insufficient to understand human exposure and vulnerability to CEs (Aguilar et al., 2005; Meehl et al., 2000; Lorenz, 2010). These studies, however, used another single descriptor of CEs defined for their purpose inquiring consistency with what is revealed by different descriptors in other studies. For example, the decrease in the occurrence number of CEs and the increase in longer lasting CEs should be explained in a consistent manner. Thus it gives the necessity of the present study systematically defining metrics for the multiple aspects of CEs.

We propose three multiple climate descriptors to characterize CEs: frequency, persistence, and entropy. The first aspect describes the *frequency* of the CE. It is well known that the frequency of CEs can reflect regional climatic variability (Parry and Carter, 1985). Another aspect is the *persistence* of the CE. In fact, the persistence of the CE could be used as a factor to determine the degree of risk (Parry et al., 2001), given the same intensity of CE. The persistence also includes valuable information on the timescale of the governing dynamical atmospheric processes of CE. The last aspect is the *entropy* of the CE. This aspect is needed for diagnosing the degree of an order of sequential CEs. High irregularity in order of occurrences brings much more damage due to difficulties in prediction, as it results in CEs that appear to occur abruptly with no time for any countermeasures. Thus, quantifying

Corresponding Author: Yong-Sang Choi, Engineering B353, Department of Atmospheric Science and Engineering, Ewha Womans University, 52, Ewhayeodae-gil, Seodaemun-gu, Seoul 03760, Korea.
E-mail: ysc@ewha.ac.kr

irregularity can be a new criterion for evaluating the degree of risk and potential damages from CEs.

To understand these multiple aspects of the nonlinear system, we apply a novel statistical method called Markov chain analysis (Hill et al., 2004) to characterize CEs. Hill et al. (2004) first applied Markov chain analysis to ecological communities. The first application in the meteorological field was conducted by Mieruch et al. (2010), in which regions were categorized according to the three Markovian climate descriptors using temperature and water vapor. They demonstrated that Markov chain analysis has a substantial advantage when simultaneously analyzing multiple aspects of chaos time series from nonlinear dynamics.

This study presents a 64-year climatology for the above three descriptors to examine the multiple aspects of the spatial characteristics of CEs. To this, we show that Markov chain analysis is applicable. The time variations of these three descriptors are then analyzed under the regional climate constraints. We also investigate the pattern of changes in the CE descriptors along with the regional winter mean temperature changes. All analyses in this study were conducted over the area in the northern hemisphere during the extended wintertime (November to March).

2. Data and Definitions of Wintertime Cold Extreme Events (CEs)

We used HadGHCND, the gridded station data at 3.75° × 2.5° resolutions, produced by the UK Met Office Hadley Centre for 1950-2014. This data set is composed of 23,000 ground-based station observations, without any satellite or aircraft measurements. It has been created using an angular distance weighting approach to interpolate the station data onto the required grid (Caesar et al., 2006). The daily minimum surface temperature was chosen as the variable for analyzing CEs during the 1950/51-2013/14 winters. Grid boxes that had more than 90 missing days are excluded (corresponding to about 1% of the total number of days).

The definition of CE is of vital importance for the extreme weather analysis. Previous studies defined extreme weather in various ways such as the most extreme value of a period, a constant difference from the seasonal normal, a change point at the tail of the empirical probability distribution of temperatures, or a particular temperature (Domonkos and Piotrowicz, 1998). This diversity reflects the choice of criteria depends on the content of one's inquiry. For example, human health is affected by the persistence of moderate cold extremes (Rocklöv et al., 2011) and a plant's phenology is determined by the critical temperature during the winter season (Gutschick and BassiriRad, 2003). For the purpose of diagnosis of the temperature variation, percentile thresholds (Jones et al., 1999; Horton et al., 2001) are widely used. The top or bottom 10th percentiles of temperatures are acknowledged to indicate unusually hot or cold events, respectively (IPCC, 2013). Accordingly, this study adopts the bottom 10th percentile of daily minimum temperatures as the threshold for CEs.

Figure 1 shows the bottom 10th percentile of daily minimum temperatures. The thresholds overshoot 10°C in low latitudes, and the CE in that region is certainly not a "disastrous" event. Thus, only the grid boxes northward of 25°N are included in the analyses. As shown in Fig. 1, inland regions have lower values than the coastal areas of the same latitude, as the small heat capacity of the drier far-inland regions reduces the daily minimum temperatures.

To more explicitly examine the regional characteristics of the decadal variability of CEs, we selected two sub-continental regions per continent in the northern hemisphere that have distinct climatological characteristics (refer to section 4): i.e., Western North America (WNA; 140°W-170°W, 55°N-75°N) and Eastern North America (ENA; 50°W-85°W, 25°N-50°N) for the continent of North America, Northern Europe (NEU; 10°W-40°E, 48°N-75°N) and Southern Europe and Mediterranean (SEM; 10°W-40°E, 35°N-48°N) for the continent of Europe, and Central Asia (CAS; 55°E-100°E, 30°N-55°N) and East Asia (EAS; 100°E-135°E, 30°N-55°N) for the continent of Asia.

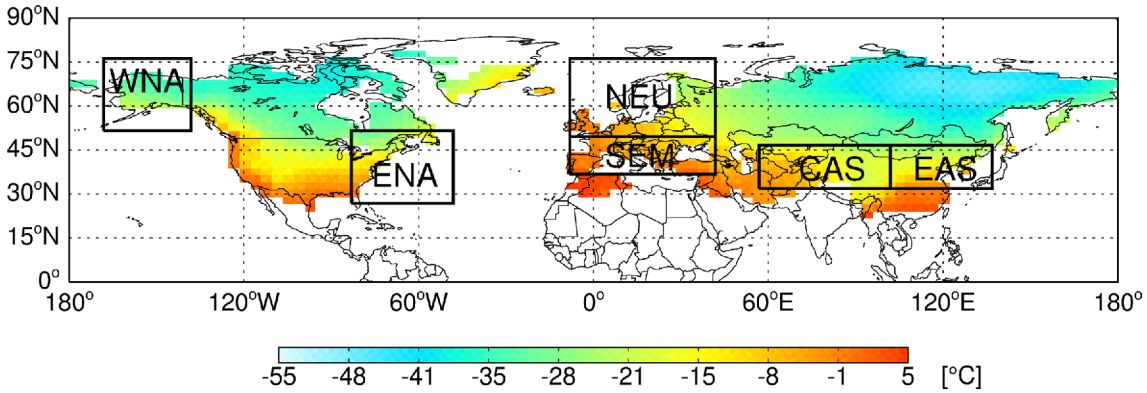

Fig. 1. The temperature corresponding to the bottom 10th percentile of the daily surface temperatures during the 1950-2014 wintertime at the 3.75° × 2.5°-grid. The six sub-regions (having distinct regional characteristics in average values of frequency, persistence, and entropy as shown in Fig. 3) partitioned in this study for quantitative analysis are superimposed as boxes.

3. Markovian CE descriptors

We applied a Markov chain analysis, which is a time- and state-discrete stochastic process (Norris, 1998). Specifically, the dichotomous (CE or non-CE in this study) events are described by a two-state, first-order Markov chain (Wilks, 1995). An advantage of this method is to represent the nonlinear system simply by a transition matrix **P** (demonstrated below) that represents the transition probability from one state to another (Mieruch et al., 2010). This Markovian transition matrix can give multiple physical information such as frequency, persistence, and entropy. Stationarity of data is a necessary assumption for applying the Markov chain analysis. To strictly keep the stationarity, the data need to be deseasonalized and detrended beforehand. However, it has been shown that such a strict prerequisite of stationarity is not required for statistical analysis. Nevertheless, our dichotomous data well fulfill weak stationarity within each 10-year running window in that the time series show (a) a constant mean and standard deviation and (b) the invariance of autocorrelation (Wilks, 1995).

Let us consider, for example, the following sequence of 12 days, consisting of only two types of events, i.e., the CE (hereafter denoted as C) and the non-CE (hereafter denoted as N):

$$\text{CNNNCCCNCNNN} \tag{1}$$

The stationary distribution, $\boldsymbol{\pi}$, for C and N is defined by the ratio of the occurrence number of each state (n_j, where j = C or N) to that of all possible states ($\Sigma_j n_j = n_N + n_C \equiv n_t$, where n_t is the total data number), as follows:

$$\boldsymbol{\pi} = \begin{bmatrix} \pi_C \\ \pi_N \end{bmatrix} = \begin{bmatrix} \frac{n_C}{n_t} \\ \frac{n_N}{n_t} \end{bmatrix} = \begin{bmatrix} \frac{5}{12} \\ \frac{7}{12} \end{bmatrix}. \tag{2}$$

The transition matrix **P** is constructed for the transition probabilities from the state of the present day to the state of the next day. Here, given the two states, C and N, four transitions are possible; a normal day following a normal day (P_{NN}), a CE day following a normal day (P_{CN}), a normal day following a CE day (P_{NC}), and a CE day following a CE day (P_{CC}). In other words, all components of P's are conditional probabilities of

$$P_{ij} = \Pr\{X_{t+1} = i | X_t = j\}, \text{ where } i,j = \{\mathrm{C},\mathrm{N}\}. \tag{3a}$$

It should be noted that the sample space for X_{t+1} consists of only two - C and N, given that X_t is either C or N. This means that $P_{NN} + P_{CN}$ and $P_{NC} + P_{CC}$ should be unity (Wilks, 1995). They are estimated as

$$P_{ij} = \frac{n_{ij}}{\Sigma_i n_{ij}} \cong \frac{n_{ij}}{n_j}, \text{ where } i,j = \{\mathrm{C},\mathrm{N}\} \tag{3b}$$

and n_{ij} denotes the number of transitions from j for the present day to i for the next day. The denominator term in Eq. (3b), $\Sigma_i n_{ij}$, differs depending on the first and last states of finite data. For the calculation efficiency, we accept the approximation made in (3b) that holds for a large sample size and do not count the last state of the given sample data in the approximated denominator term, n_j. This approximation is useful because there is no available data following the last state to be incorporated into the counts in numerators (Wilks, 1995). The four transition probabilities constitute the transition matrix:

$$\mathbf{P} = \begin{bmatrix} P_{CC} & P_{CN} \\ P_{NC} & P_{NN} \end{bmatrix} \cong \begin{bmatrix} \frac{n_{CC}}{n_C} & \frac{n_{CN}}{n_N} \\ \frac{n_{NC}}{n_C} & \frac{n_{NN}}{n_N} \end{bmatrix} = \begin{bmatrix} \frac{2}{5} & \frac{2}{6} \\ \frac{3}{5} & \frac{4}{6} \end{bmatrix}. \tag{3c}$$

From the two factors, **P** and $\boldsymbol{\pi}$, three descriptors, frequency (Fr), persistence (Pe), and entropy (En) can be derived. Note that n_N is six not seven in our approximated formula, as the last state is not considered.

The three descriptors provide successional properties of extreme weather such as the event number, the persistence period, and the entropy of CEs. Thus, these descriptors retain their independent meanings as detailed below.

a. Frequency

Returning to example (1), the frequency of C or N among the 12 total days is the simplest factor that represents the system's statistical property. In this study, the frequency is defined as the number of same consecutive states (i.e., CCC) in a given period, as they are normally generated by the same physical process and can be regarded as one event (IPCC, 2013). Consequently, the Fr value of C in example (1) is 3 per 12 days by counting the number of C groups (*not* by counting the number of C events). Fr values can then be expressed in terms of P_{CC} and π_C:

$$\mathrm{Fr} = n_{NC} = \frac{n_{NC}}{n_C} \cdot \frac{n_C}{n_t} \cdot n_t = P_{NC} \cdot \frac{n_C}{n_t} \cdot n_t = (1 - P_{CC}) \cdot \pi_C \cdot n_t,$$

when the last state is N; (4a)

$$\mathrm{Fr} = n_{NC} + 1 = (1 - P_{CC}) \cdot \pi_C \cdot n_t + 1, \text{ when the last state is C;} \tag{4b}$$

where the equations adopted $\pi_C = n_C/n_t$ in Eq. (2), $P_{CC} = n_{CC}/n_C$ in Eq. (3b), and $P_{NC} + P_{CC} = 1$.

These formulae indicate that the Fr values of C are determined from the occurrence probability of the consecutive C state (P_{CC}), the stationary distribution of C (π_C), and the total data number (n_t). In Eqs. (4a) and (4b), the increase in π_C and n_t, and the decrease in P_{CC} lead to an increase in the Fr value. This is illustrated in Fig. 2a, which shows Fr values (contour lines) as a function of P_{CC} (x-axis) and π_C (y-axis). A maximum Fr value appears for $\pi_C = 1$ and $P_{CC} = 0$, while a minimum Fr value appears for $\pi_C = 0$ and $P_{CC} = 1$.

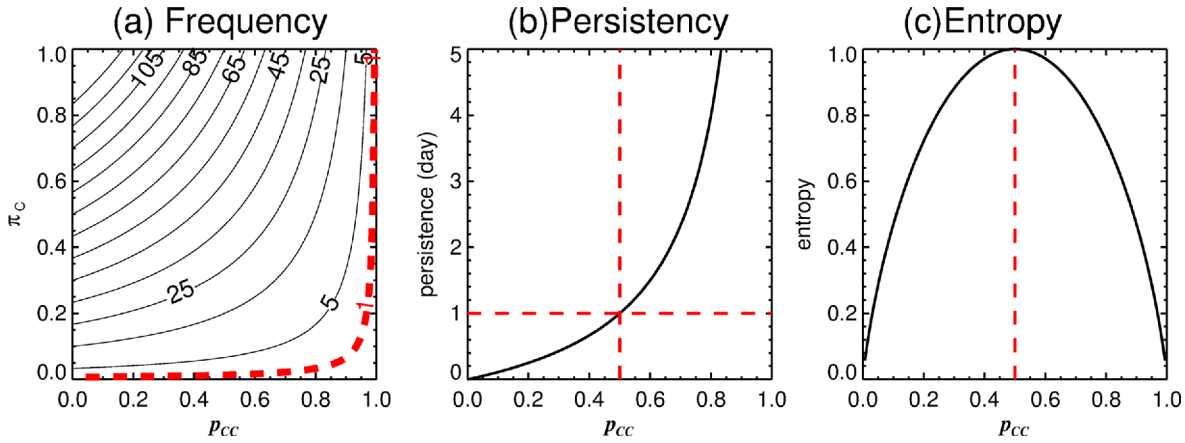

Fig. 2. Diagram of (a) Frequency, (b) Persistence, and (c) Entropy values representing the component of the transition matrix (P_{CC}) and stationary distribution (π_C), based on Eqs. (4a), (5a), and (6), respectively. The red dashed line indicates Markovian descriptors with unity.

b. Persistence

The length of each C group (i.e., persistence) is also considered as a critical factor that characterizes regional CE because it can partly represent the degree of risk caused by extreme events. In example (1), the persistence of the three C groups differs. The first and third C groups have a C state only for one day, whereas the second C group has a C state for two more days after the day of occurrence. Persistence is defined as the average number of consecutive C events over all C groups. That is, the Pe value of C in example (1) is 2/3, by dividing the total number of consecutive C events (i.e., 0 for C, and 2 for CCC) by that of the C groups (3). Pe can also be expressed in terms of the components of **P** and $\boldsymbol{\pi}$, P_{CC} and π_C:

$$\mathrm{Pe} = \frac{n_{CC}}{Fr} = \frac{n_{CC}}{n_C}\cdot\frac{n_C}{n_t}\cdot\frac{n_t}{Fr} = \frac{P_{CC}\cdot\pi_C}{(1-P_{CC})\cdot\pi_C} = \frac{P_{CC}}{1-P_{CC}}, \tag{5a}$$

when the last state is N;

$$\mathrm{Pe} = \frac{n_{CC}}{Fr} = \frac{P_{CC}\cdot\pi_C}{((1-P_{CC})\cdot\pi_C + 1/n_t)}, \tag{5b}$$

when the last state is C

Here $\pi_C = n_C/n_t$ in Eq. (2), $P_{CC} = n_{CC}/n_C$ in Eq. (3b), and $P_{NC} + P_{CC} = 1$ is utilized. Equation (5b) can be approximated to Eq. (5a) when n_t is large enough. As n_t in this study is around 1,500, it is sufficient to consider Pe as a function of P_{CC} only. In this case, the Pe value increases hyperbolically with P_{CC} (see Fig. 2b).

c. Entropy

The last Markovian descriptor is entropy. It is traditionally a measure of the disorder of the closed system (Styer, 2000) in physics, but it has been widely used for representing the degree of complexity (Pincus, 1991), chaos or irregularity in a system (Inouye et al., 1991) in various fields. To the climate data, Mieruch et al. (2010) applied entropy as follows.

$$En_{system} = \Sigma_j \pi_j \times En_j, \text{ and } En_j = \Sigma_i - P_{ij} log_n P_{ij} \tag{6}$$

where n = the number of states and En_j is the entropy of j state.

In this study, we adopt Eq. (6) to diagnose the irregularity of sequential CEs from the past 64-winter data. Thus we only refer En_C (i.e., entropy of cold extreme events) instead of the total En_{system}; hereafter, En simply indicates En_C. Taking the relation '$P_{NC} + P_{CC} = 1$' from Eq. (3b), En (i.e., En_C) eventually becomes a function of P_{CC}. Additionally, we divided $\log\Sigma i$ (= 2) for normalizing the entropy (Mieruch et al., 2010):

$$\begin{aligned}\mathrm{En} &= (1/\log_2) \times (-P_{NC}\log_2 P_{NC} - P_{CC}\log_2 P_{CC}) = \\ &(-1/\log_2) \times ((1-P_{CC})\log_2(1-P_{CC}) + P_{CC}\log_2 P_{CC})\end{aligned} \tag{7}$$

Mathematically En varies between 0 and 1. This value is the objective index to represent the degree of irregularity of a given system. For example, a system with En = 0 (En = 1), a specific event occurs perfectly regularly (irregularly) as the En-P_{CC} relation indicates in Fig. 1c. When CE never persists ($P_{CC} = 0$) or CE eternally lasts ($P_{CC} = 1$), we can say that CE system is perfectly regular. On the other hand, in the case of unexpected persistence of CEs ($P_{CC} = 0.5$ only in the two-state example), the CE system can be regarded as perfectly irregular. Comparing Figs. 2b and 2c, it can be inferred that the Pe-En relation is positive when $P_{CC} < 0.5$, while negative when $P_{CC} > 0.5$. It should be noted that this relation is derived from our specific choice of the C-state entropy and two-state of Markov chain. Otherwise, such a simple relation between Pe and En cannot be simply derived.

The three descriptors (Fr, Pe, and En) for the CEs were calculated for each grid in the northern hemisphere during the winters of 1950/51-2013/14. When calculating the components of **P** and $\boldsymbol{\pi}$, we excluded the day on which a missing value

appeared. The attribution of a single extreme event to anthropogenic climate change is difficult to prove (IPCC, 2013). However, the statistics of CEs over at least 10 years can be reasonably associated with regional winter mean temperature change that may or may not be a consequence of anthropogenic influences. We, therefore, applied the Markov chain analysis moving 10-year period, such as 1950/51-1959/60, 1951/52-1960/61, … 2003/04-2012/13, and 2004/05-2013/14. In particular, the Fr values calculated for each 10-year period were later divided by 10 to indicate the averaged Fr value per winter, which we omit 'per winter' in the unit of Fr for brevity.

4. Climatology and Time Variations of the CE Descriptors

Figure 3 shows the spatial patterns of the averaged three descriptors (Fr, Pe, and En) of CEs. The averaged Fr is about five events in the northern hemisphere, but large regional differences exist. Relatively large Fr values are widely distributed over southern Canada and the United States (6-8 events), and sporadically over the northern parts of Europe and the eastern part of Asia (i.e., downstream of the Tibetan Plateau). Relatively small Fr values, fewer than two events, are found over Central Asia including upstream of the Tibetan Plateau. Corresponding to the temperature variability, inland regions have more frequent cold extremes than coastal regions. It is noted that Fr values indicate the number of the consecutive CE events (not the number of CE days) during a winter. For instance, Fr is 1 if CE days persist for a winter. In

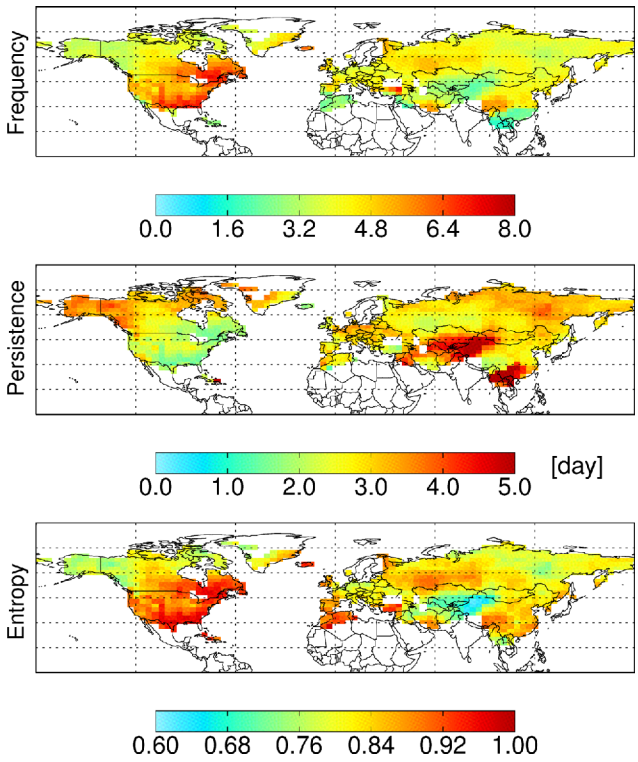

Fig. 3. Averaged values in 54 data sets of frequency, persistence, and entropy of CEs.

general, the Pe values show almost the opposite spatial pattern to the Fr values (middle panel). Thus, the regions (Alaska, the southern parts of Europe, and Central Asia) where Fr values are relatively small show large Pe values in this two-state application. Quantitatively, CEs in these regions last longer than three days, while CEs in southern Canada and the United States last shorter than two days. Finally, En values have the opposite spatial pattern to that of Pe values (the bottom panel). That is, the averaged P_{CC} values larger than 0.5 (the averaged persistence day > one day, refer to section 3 for details) play a decisive role in reducing the entropy of CEs over the most regions. Simply put, climatological means show that higher frequency, shorter persistence, and higher entropy of CEs coincide spatially.

To examine the regional characteristics of CEs' time variation, we took an area average of each descriptor (Fr, Pe, and En) over the six regions presented in Fig. 1. Since the trend assessment of CEs differs based on the analyzed period (Walsh et al., 2001), we focus on examining the decadal variations (solid lines in Fig. 4) of three descriptors rather than quantifying the regression slope. The Fr values (Fig. 4a) suggest that CEs over all regions have occurred relatively less frequently in recent decades. While the higher Fr values were observed around the 1960s-70s, the values became lower in the 2000s for all the six regions. The average of the Fr changes among the six regions is −0.042 Fr per moving ten winters; SEM and EAS have the smallest (−0.023 Fr per moving ten winters) and the largest (−0.067 Fr per moving ten winters) changes, respectively. This globally decreasing tendency of the frequency was the overall conclusion of previous studies (IPCC, 2013).

In addition to the trend, the regional characteristics of Fr reveal decadal variations, with the modality of their temporal variation differing among the regions. The winter climate in WNA is known to be modulated by the large-scale teleconnection associated with various North Pacific sea surface temperature (SST)-related climate modes, such as the Pacific Decadal Oscillation (PDO) and the Northeast Pacific mode (Zhang et al., 1997; Hartmann, 2015). In WNA, the higher Fr appears during the 1970s and 1990s, and the lower Fr occurs in the 1980s and 2000s. This large decadal variability in WNA reflects the combined role of the North Pacific decadal climate modes. Over the ENA, the Fr values of CEs increased during the 1950s, remained high until the 1970s, and gradually have decreased since the late 1970s. The low values in the 1950s and 1990s are consistent with the observed variability over ENA in the previous studies (Rogers and Rohli, 1991; Walsh et al., 2001; Portis et al., 2006), whereas the decreasing trend since the 1960s is inconsistent with the previous studies concluding no evidence of a decreasing trend in the United States (Walsh et al., 2001; Portis et al., 2006; Westby et al., 2013). Based on the previous studies, the decadal variation of CEs over ENA can be affected by the Pacific-North American pattern (PNA)-related teleconnection associated with the North Pacific climate modes (Rogers and Rohli, 1991; Hartmann,

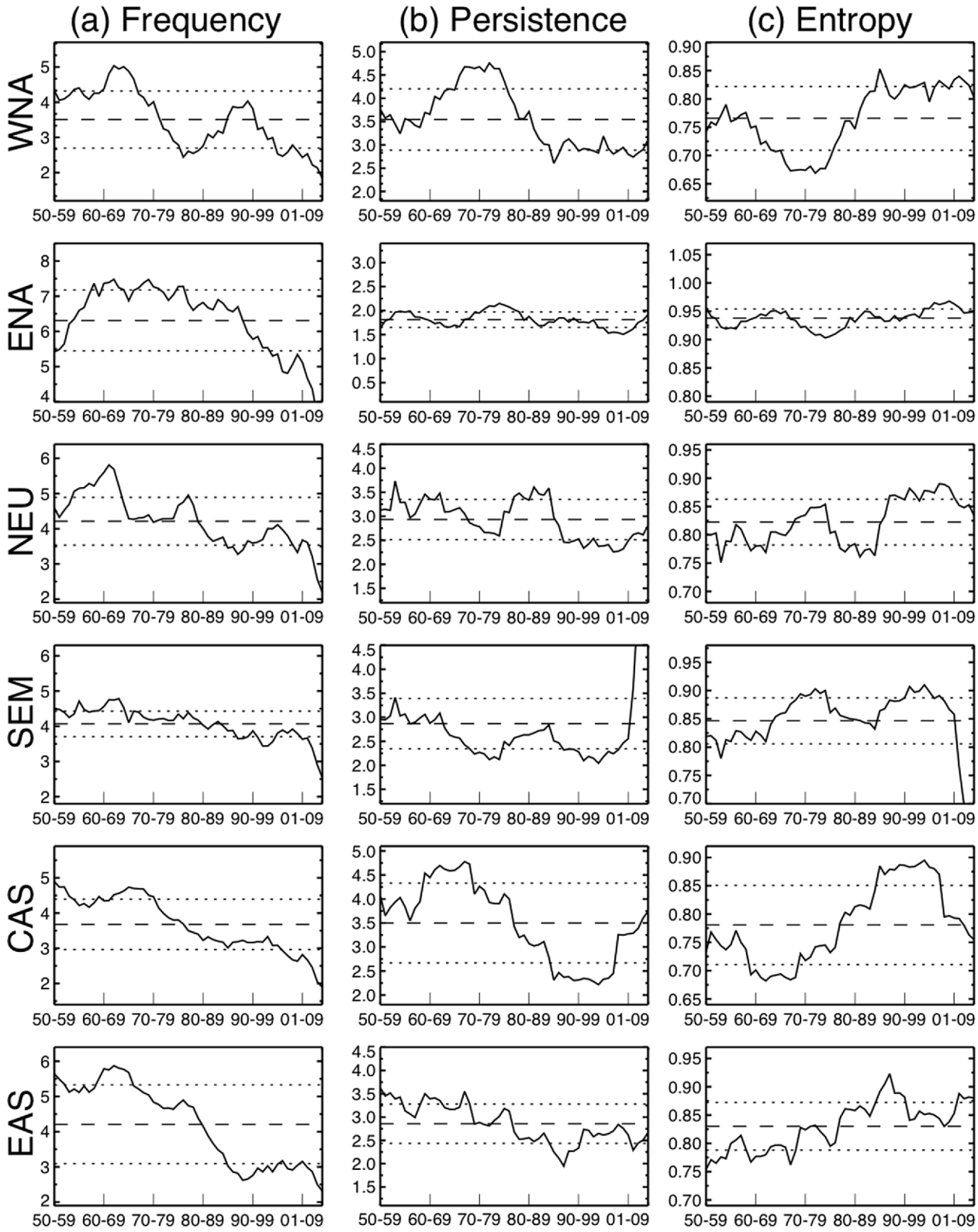

Fig. 4. Time series of CE descriptors for each of the six regions, which are identical to the regions in Fig. 1; Fr (a, the first column), Pe (b, the second column), En (c, the third column). The *x*-axis represents the analyzed period for moving 10-year periods; e.g. 50-59 indicates the period from 1950 to 1959 used to calculate the descriptor. Note that to clearly show the decadal variability (solid lines) of each descriptor, the y-axis of CEs for one region is centered at the mean (dashed lines) and ranges with the same standard deviation (± 1 standard deviations appear as dotted lines) as that for the other regions.

2015), and also by the Arctic Oscillation (AO)/North Atlantic Oscillation (NAO) (Thompson and Wallace, 1998; Walsh et al., 2001). In Europe, on the other hand, the Fr values of CEs over NEU and SEM have different temporal changes. CEs over NEU exhibit substantial decadal variability, whose frequency is lower and the amplitude is larger than those over SEM, while CEs over SEM display little decadal variation, but experience an abrupt and massive reduction during the recent period. In Asia, the Fr values over CAS and EAS also have similar decadal variability; they decreased up to the 1990s and started to decline again up to the recent period. This variability is consistent with the decadal variability over the northeast Asia shown in Woo et al. (2012). They demonstrated that the frequency of CEs (high Fr values) over Asian regions is governed by the AO phase (negative).

Figure 4b shows that the long-term variations of the Pe values are distinctive from those of the Fr values (correlation coefficient, r = –0.29 between Fr and Pe over all regions). In North America, the large Pe values over WNA are shown for the period of 1970s-80s, with a stagnant period after the 1990s of Pe ≈ three days. On the other hand, the Pe values over ENA not only have the smallest values but also remain almost invariant, independent of the Fr values. In Europe, both NEU and SEM show similar variability; they decrease up to the 1980s; a higher Pe appears in the 1990s and starts to decline again. As stated in the introduction, Europe underwent a record-breaking severe winter during the 2012/13 period, and we can confirm that these events are well captured in the Pe over SEM. In Asia, Pe over CAS shows similar variability to the Fr values up to the 2000s, but it began to increase in a recent decade, similar to SEM. Though weaker than CAS, Pe over EAS also stopped decreasing in the 1990s. Note that the recent increases in the Pe values over SEM, CAS, and EAS in Fig. 4b seem to be related to the hot research topic about the role of amplified Arctic warming in CEs in mid-latitudes in recent years (e.g., Francis and Vavrus, 2012; Kim et al., 2014). Although it is still debate whether the amplified Arctic warming could cause more persistent CEs in mid-latitudes (Barnes, 2013; Screen and Simmonds, 2013; Overland et al., 2015), the data clearly show the recent upswing of Pe in some regions as opposed to the decrease in Fr.

The time variations of En (Fig. 4c) are opposite to those of Pe (r = –0.93 between Pe and En over all regions). The small (large) entropy of CEs is associated with long (short) persistence. For example, En over ENA exhibited the lowest values during the period of 1970s-1980s, but maintained higher values of around 0.83 for 1990s-2010s. In other words, En values also have different behaviors from Fr depending on regions, which is similar to the Pe values.

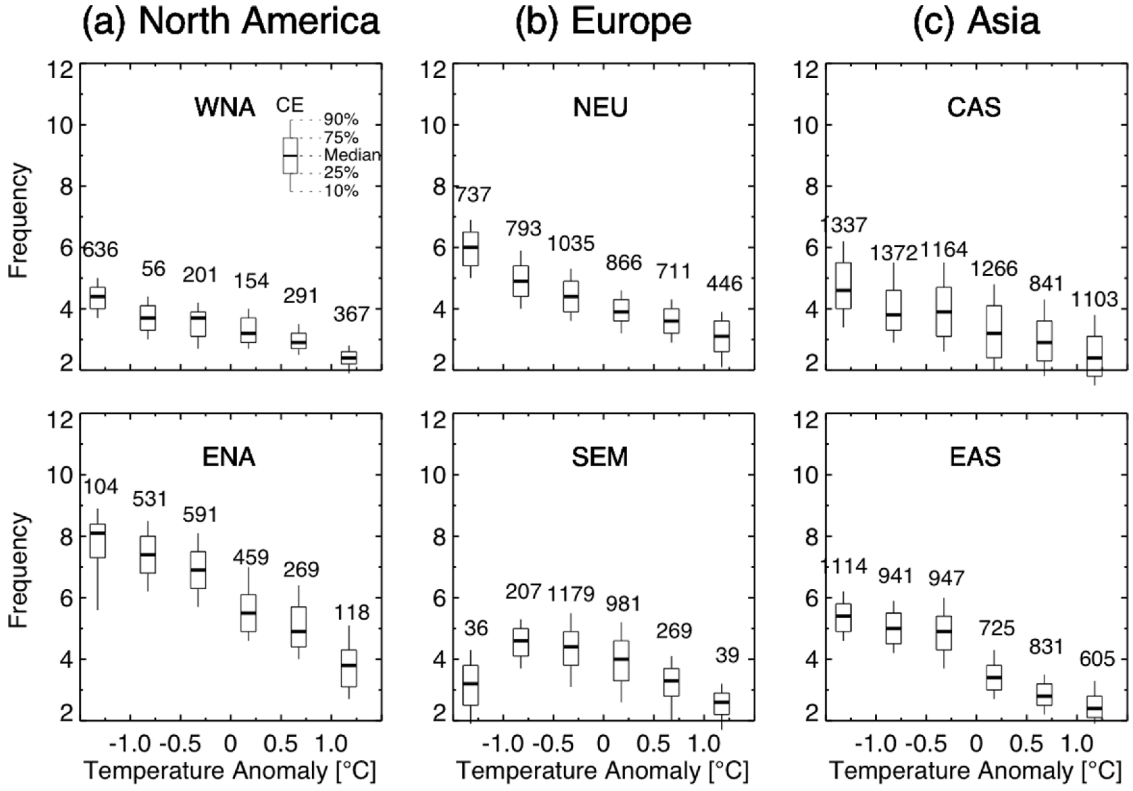

Fig. 5. Box plot showing the distributions of Fr values depending on regional winter temperature anomalies (it is calculated by subtracting the mean of the 54 values of winter mean temperature from each winter mean temperature) in six selected domains which defined in Fig. 1. Each box includes information about the 90th, 75th, 25th, and 10th percentile, and the median value event numbers of CEs. Sample numbers are written over each box.

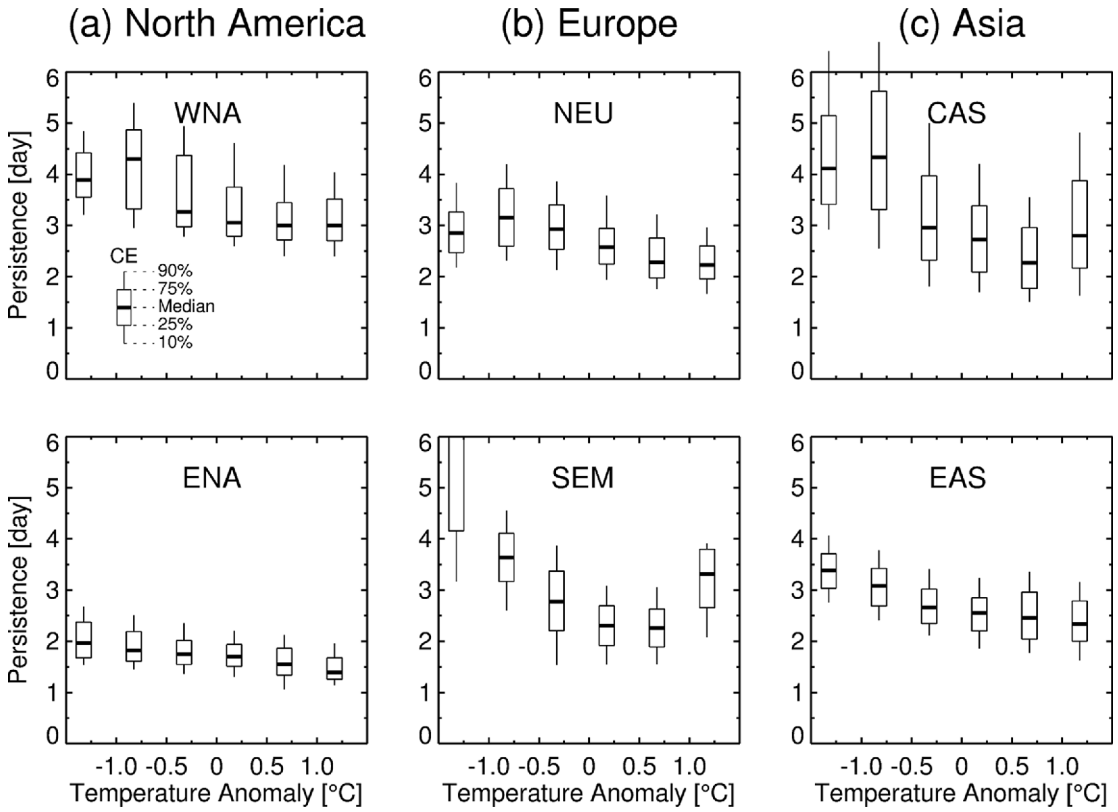

Fig. 6. As for Fig. 5, but with Pe values.

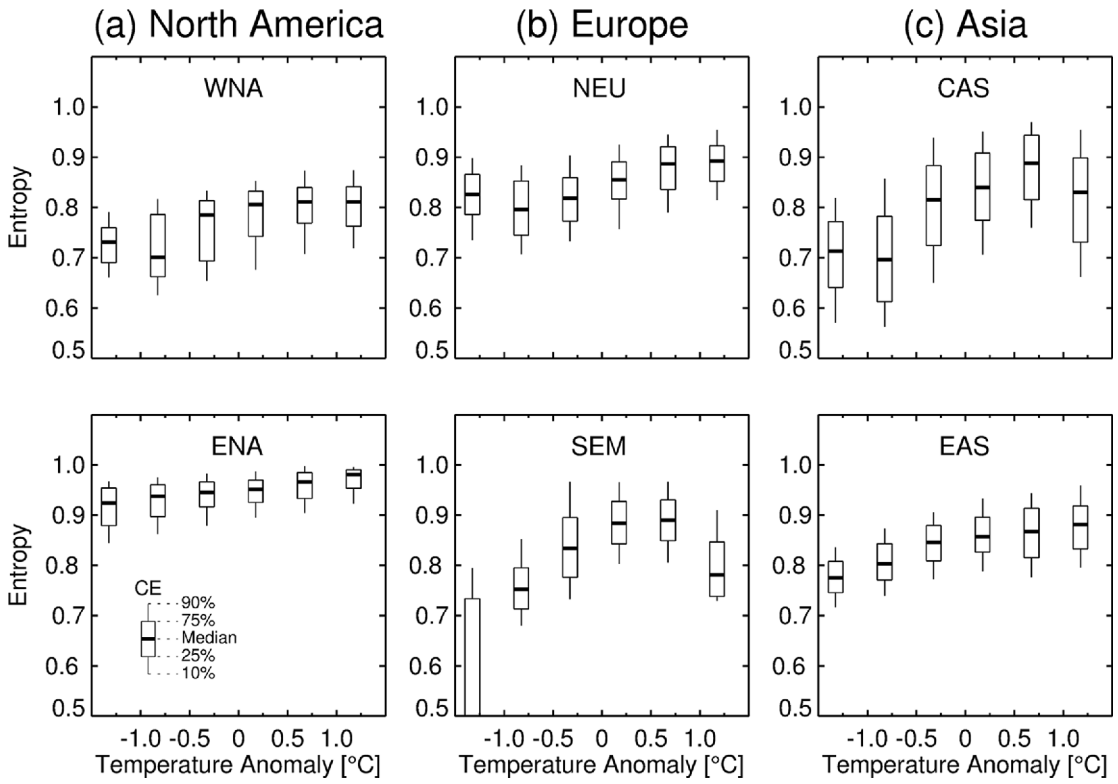

Fig. 7. As for Fig. 5, but with En values.

5. Relationship with regional wintertime temperature

Are these climate descriptors closely associated with regional climate change? In this section, we investigate the relationship between CE descriptors and winter mean temperature anomalies (T′) for each region. To maintain the coherency with CE descriptors, we used 54 sets of moving 10-year-average daily minimum temperature anomalies departed from the mean value of the entire period. The 54 CE descriptors at each grid in the six domains (Fig. 1) were binned into six ranges of T′ values: $< -1.0^{o}$C, -1.0 to -0.5^{o}C, -0.5 to 0^{o}C, 0 to 0.5^{o}C, 0.5 to 1^{o}C, and $\geq 1^{o}$C. We then used the box plots to present the statistics of the CE descriptor samples for each T′, showing the 90th, 75th, 50th (median), 25th, and 10th percentile values.

Figure 5 shows the box plots of the Fr values for the CEs against T′ over the six regions. Their median values (the black thick solid lines), as well as 10th to 90th percentile ranges commonly decrease with increasing T′ over all six regions. It is expected by simply shifting of the T′ distribution to the warmer side. In North America (Fig. 5a), the Fr values over WNA are uncorrelated with the regional temperature at $T' < 0$, but they start to decrease monotonically (5 to 1) with increasing T′ at $T' > 0$. On the other hand, the Fr values over ENA monotonically decrease (8 to 4) with increasing T′ at all T′ ranges. In fact, this dependence on T′ over ENA is the largest among the six regions. NEU in Europe (Fig. 5b) shows that Fr also decreases with T′, whereas SEM has a particular pattern whereby relatively small Fr groups are formed in the lowest T′ category. In Asia, the Fr value decreases with T′ with a similar decreasing rate in both the CAS and EAS of Asia (Fig. 5c).

Next, the Pe values of CEs are investigated in relation to T′ (Fig. 6). A change in Pe with an increase in T′ shows apparent regional dependency. In North America (Fig. 6a), Pe values over WNA are maintained over three at $T' > 0^{o}$C, whereas those over ENA decrease with increasing T′ from 2 to 1.5 (days). Over NEU in Europe (Fig. 6b), Pe slightly decreases from about three to two (days) with increasing T′. In SEM, the maximum Pe appears for the lowest T′, decreases until $T' < 0.5^{o}$C, but increases again for $T' > 0.5^{o}$C. The CAS of Asia (Fig. 6c) shows a similar Pe-T′ relation with WNA and NEU; however, when $T' > 0.5^{o}$C, Pe increases again similar to SEM, while that over EAS monotonically decreases with T′. In summary, the Pe-T′ relation is regionally much more diverse than the Fr-T′ relation.

Finally, the En-T′ relation is shown in Fig. 7. Similar to Pe, the En values behave differently depending on the regions. In spite of the regional dependency, the highest median values of En appear in the condition of the highest T′, which implies that it can become harder to forecast successional CEs when regional warming occurs (WNA, ENA, NEU, and EAS). However, for regions such as SEM and CAS, the highest median values are observed when $0 < T' < 1.0^{o}$C. When $T' > 1.0^{o}$C, the highest regional anomaly temperature categories, the En values tend to be lower. Thus, we cannot be confident that successional CEs, in general, would have higher entropy under the regional warming.

6. Conclusions and Discussion

We have shown that the persistence or entropy can behave differently from the frequency of events in different time scales and different regions. One of the primary factors to evaluate vulnerability to climate change is known as the 'exposure' (Secretariat, 2005). To accurately assess the level of exposure, various characteristics of extreme events as well as their changes with regional warming need to be understood. In that sense, examining the frequency of extreme events, the primary focus in previous studies, may not be sufficient to diagnose realistic regional exposure. Viewed in this light, this study offers substantive information to improve the measurement of the level of exposure. We propose multiple climate descriptors to characterize CEs, such as frequency, persistence, and entropy over land areas in the northern hemisphere.

As a straightforward method for analyzing these descriptors, this study carefully used Markov chain analysis for the period 1950-2014. It is the first trial to apply Markov chain to the extreme climate research. Thereby we deliberately dealt with only two states, for which the conventional approaches (individually calculating each descriptor) independent of Markov chain analysis (i.e., the transition matrix) can obtain the same results. In fact, obtaining multiple descriptors from the conventional approaches is harder for more than two states that are different climate variables or diverse dynamical phases in a variable. As Mieruch et al. (2010) noted, this Markov chain analysis will be powerful for incorporating characteristics of more than two states. Our study, therefore, is suggestive of a practical methodology for the future study and provides an applicability to the expand extreme states (i.e., incorporation of extremely hot and cold days, heavy precipitation and strong winds, or many thresholds in extreme temperatures and so on). The insight to the climate system with many states by Markov chain cannot be replaced with the results from traditional methods. Thus, our application of the Markov chain analysis lays the foundations for future researches on nonlinear extreme events in the future.

Climatologically, higher frequency, shorter persistence, and higher entropy coincide in the same region (Fig. 3). The relation among the climate descriptors is relatively simple, as our current application is based on the two stated definition with the bottom 10th percentile threshold. For a more complex request for multiple levels of extreme events, such a simple relation may not be achieved. However, the simplicity of our methodology allows that our study can be easily extended to cover various definitions of CEs.

We categorized six regions with climatologically distinct characteristics and examined the temporal changes of the descriptors (Fig. 4). The noticeable feature is that time variations of frequency are uncorrelated with those of persistence. In 1970s-80s, the CEs show more occurrences and longer

persistence. In the 2000s, on the other hand, the persistence of CEs is increased or stagnated depending on the regions although the frequency reduces. This strongly implies that the single descriptor cannot capture the holistic nature of CEs. It is expansive to study dynamical mechanisms or causes for the observed decadal variability of frequency and persistence.

The variations of entropy that represent irregularity of sequential CEs were found to be strongly linked with those of persistence of CEs (Figs. 4b vs. 4c). This linkage is not only because we adopted the first Markov chain with only two states, but also because of the simple definition of entropy. Despite the high dependence of entropy upon persistence, the entropy has its meaning that can diagnose the irregularity of extreme events. As an example of a more advanced application of entropy, we can apply the multiple extreme indicators (e.g., intense heat waves, heavy rainfall or severe droughts) together. In this advanced example, the entropy values synthesize the degree of vulnerability from irregular weather events and then can be practically adopted for planning on severe weather phenomina.

It is worth discussing previous studies showing the large-scale climate variability in comparison with the variability of Markovian physical descriptors. For example, the decadal variability of frequency in eastern North America (Fig. 4b) had a similarity with the variability of cold air outbreaks shown in previous studies (Rogers and Rohli, 1991; Walsh et al., 2001; Portis et al., 2006). The studies could involve the decadal variability of the PNA teleconnection associated with the North Pacific climate modes (e.g., PDO, Northeast Pacific mode, etc.) and the AO/NAO. The recent increases of persistence over the Southern parts of Europe and central/eastern Asia (Fig. 4b) can be understood in the context of the Arctic and mid-latitude teleconnection (Francis and Vavrus, 2012; Kim et al., 2014), though the dynamical mechanism is still controversial (Barnes, 2013; Screen and Simmonds, 2013; Overland, 2015). By concisely summarizing the multiple aspects of extreme events of the nonlinear weather system, the Markovian descriptors may provide a valuable starting point for further studies combining statistics and dynamics. Although this study did not explore the underlying dynamical behavior related with the Markovian descriptors of CEs, the future study will extend to more deepened dynamical interpretation in relation to blocking wave patterns and large-scale climate modes.

Finally, we investigated how these changes in CEs descriptors were associated with the regional winter mean temperature anomaly (T′). The sensitivity of the frequency of CEs to T′ depends on the region, but the frequency decreased with regional warming. This is because the higher T′ results in shifting of the temperature distribution to the warmer side, as reported in previous studies (Meehl et al., 2000). Persistence, however, is uncorrelated to T′ in general. The entropy of CEs also does not have a clear linear relation with T′ over all regions. Therefore, with respect to the persistence and entropy of CEs, it is hard to determine whether socioeconomic losses or the level of exposure would diminish or enhance in the future.

Markov chain analysis is a useful method to understand the characteristics of a complex weather system including extreme events. There is, of course, a limitation in examining weather features using our simple definition of CEs. Weather is a phenomenon of a continuous system, so several features cannot be captured well when events are categorized to two simple discrete states. The appropriate definition to represent "extreme" can differ depending on the purpose and the region. Similarly, some N events that break the cluster of C can be an issue and can influence the estimate of the cluster characteristics (Ferro and Segers, 2002). We found, however, that the relative spatial distribution or the long-term variation of CEs was not sensitive to the definition of CEs.

Despite these limitations, our study suggests that CEs should essentially be characterized by multiple descriptors to improve our understanding of the impacts and causes of CEs, ensuring the utility of the Markov chain analysis to the extreme events. We also found more informative results when CEs are defined in a different way with multi-state. For more profound understanding and better prediction of extremes, we suggest examining the characteristics of extremes in more specific regions (smaller than those analyzed here), to account for the regional variation.

Acknowledgements. This study was supported by the Korea Meteorological Administration Research and Development Program (KMIPA2015-6110), RP-Grand 2015 of Ewha Womans University, Korea, and the research project (PE16100) of the Korea Polar Research Institute. Y.-S. Choi acknowledges the support by the Jet Propulsion Laboratory, California Institute of Technology, sponsored by the *National Aeronautics and Space Administration* (NASA). W. Kim is currently supported by APEC Climate Center. We appreciate UK Met Office for providing the daily minimum surface temperatures used in this paper at their data server (http://www.metoffice.gov.uk/hadobs/hadghcnd/download.html). We also thank Ms. Soo Jin Shin for her initial work.

Edited by: Masahiro Watanabe

References

Aguilar, E., and Coauthors, 2005: Changes in precipitation and temperature extremes in Central America and Northern South America, 1961-2003. *J. Geophys. Res.,* **110**, D23107, doi:10.1029/2005JD006119.

Barnes, E. A., 2013: Revisiting the evidence linking Arctic amplification to extreme weather in midlatitudes. *Geophys. Res. Lett*., **40**, 4728-4733, doi:10.1002/grl.50880.

Blunden, J., and Coauthors, 2013: State of the climate in 2012. *Bull. Amer. Meteor. Soc.*, **94**, 1-238.

Caesar, J., L. Alexander, and R. Vose, 2006: Large-scale changes in observed daily maximum and minimum temperatures: Creation and analysis of a new gridded data set. *J. Geophys. Res.*, **111**, D05101, doi:10.1029/2005JD006280.

Domonkos, P., and K. Piotrowicz, 1998: Winter temperature characteristics

in Central Europe. *Int. J. Climatol.*, **18**, 1405-1417, doi:10.1002/(SICI)1097-0088(19981115)18:13<1405::AID-JOC323>3.0.CO;2-D.

Easterling, D. R., G. A. Meehl, C. Parmesan, S. A. Changnon, T. R. Karl, and L. O. Mearns, 2000: Climate extremes: Observations, modeling, and impacts. *Science*, **289**, 2068-2074, doi:10.1126/science.289.5487.2068.

Ferro, T. A., and J. Segers, 2003: Influence for clusters of extreme values. *J. Roy. Stat. Soc.*, **65**, 546-556, doi:10.1111/1467-9868.00401.

Francis, J. A., and S. J. Vavrus, 2012: Evidence linking Arctic amplification to extreme weather in mid-latitudes. *Geophys. Res. Lett.*, **39**, L06801, doi:10.1029/2012GL051000.

Gutschick, V. P., and H. BassiriRad, 2003: Extreme events as shaping physiology, ecology, and evolution of plants: toward a unified definition and evaluation of their consequences. *New. Phytol*., **160**, 21-42, doi:10.1046/j.1469-8137.2003.00866.x.

Hartmann, D. L., 2015: Pacific sea surface temperature and the winter of 2014. *Geophys. Res. Lett.*, **42**, 1894-1902, doi:10.1002/2015GL063083.

Herring, S. C., M. P. Hoerling, J. P. Kossin, T. C. Peterson, and P. A. Stott, 2015: Explaining extreme events of 2014 from a climate perspective. *Bull. Amer. Meteor. Soc.*, **96**, 1-172, doi: 10.1175/BAMS-ExplainingExtremeEvents2014.1.

Hill, M. F., J. D. Witman, and H. Caswell, 2004: Markov chain analysis of succession in a rocky subtidal community. *American Nat*., **164**, 46-61, doi:10.1086/422340.

Horton, E. B., C. K. Folland, and D. E. Parker, 2001: The changing incidence of extremes in worldwide and Central England temperatures to the end of the twentieth century. *Climatic Change*, **50**, 267-295, doi:10.1023/A:1010603629772.

Inouye, T., K. Shinosaki, H. Sakamoto, S. Toi, S. Ukai, A. lyama, Y. Katsuda, and M. Hirano, 1991: Quantification of EEG irregularity by use of the entropy of the power spectrum. *Electroencephalogr. Clin. Neurophysiol.* **79**, 204-210.

IPCC, 2013: *Climate Change 2013: The Physical Science Basis: Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, 1535 pp.

Jones, P. D., E. B. Horton, C. K. Folland, M. Hulme, D. E. Parker, and T. A. Basnett, 1999: The use of indices to identify changes in climatic extremes. *Climatic Change*, **42**, 131-149, doi:10.1023/A:1005468316392.

Kim, B.-M., S.-W. Son, S.-K. Min, J.-H. Jeong, S.-J. Kim, X. Zhang, T. Shim, and J.-H. Yoon, 2014: Weakening of the stratospheric polar vortex by Arctic sea-ice loss. *Nat. Commun.*, **5**, 4646, doi:10.1038/ncomms5646.

Kunkel, K. E., D. R. Easterling, K. Hubbard, and K. Redmond, 2004: Temporal variations in frost-free season in the United States: 1895-2000. *Geophys. Res. Lett.*, **31**, L03201, doi:10.1029/2003GL018624.

Lorenz, R., E. B. Jaeger, and S. I. Seneviratne, 2010: Persistence of heat waves and its link to soil moisture memory. *Geophys. Res. Lett.,* **37**, L09703, doi:10.1029/2010GL042764.

Meehl, G. A., and Coauthors, 2000: An introduction to trends in extreme weather and climate events: Observations, socioeconomic impacts, terrestrial ecological impacts, and model projections. *Bull. Amer. Meteor. Soc.*, **81**, 413-416, doi:10.1175/1520-0477(2000)081<0413:AITTIE>2.3.CO;2.

Mieruch, S., S. Noël, H. Bovensmann, J. P. Burrows, and J. A. Freund, 2010: Markov chain analysis of regional climates. *Nonlin. Proc. Geophys.*, **17**, 651-661, doi:10.5194/npg-17-651-2010.

Norris, J. R., 1998: *Markov Chains.* 1st pbk. ed. Cambridge University Press, 237 pp.

Overland, J., J. A. Francis, R. Hall, E. Hanna, S.-J. Kim, and T. Vihma, 2015: The melting Arctic and midlatitude weather patterns: are they connected? *J. Climate*, **28**, 7917-7932, doi:10.1175/JCLI-D-14-00822.1.

Parry, M., and Coauthors, 2001: Millions at Risk: Defining critical climate change threats and targets. *Global environ. Change*, **11**, 181-183, doi:10.1016/S0959-3780(01)00011-5.

Parry, M. L., and T. R. Cater, 1985: The effect of climatic variations on agricultural risk. *Climatic Change*, **7**, 95-110, doi: 10.1007/BF00139443.

Pincus, S. M., 1991: Approximate entropy as a measure of system complexity. *Proc. Natl. Acad. Sci*., **88**, 2297-2301.

Portis, D. H., M. P. Cellitti, W. L. Chapman, and J. E. Walsh, 2006: Low-frequency variability and evolution of North American cold air outbreaks. *Mon. Wea. Rev.*, **134**, 579-597, doi:10.1175/MWR3083.1.

Rocklöv, J., K. Ebi, and B. Forsberg, 2011: Mortality related to temperature and persistent extreme temperatures: a study of cause-specific and age-stratified mortality. *Occup. Environ. Med.*, **68**, 531-536, doi:10.1136/oem.2010.058818.

Rogers, J. C., and R. V. Rohli, 1991: Florida citrus freezes and polar anticyclones in the Great Plains. *J. Climate*, **4**, 1103-1113, doi:10.1175/1520-0442(1991)004<1103:FCFAPA>2.0.CO;2.

Screen, J. A., and I. Simmonds, 2013: Exploring links between Arctic amplification and mid-latitude weather. *Geophys. Res. Lett.*, **40**, 959-964, doi:10.1002/grl.50174.

Styer, D. F., 2000: Insight into entropy. *American J. Phys*., **68**, 1090-1096.

Secretariat, U. N. F. C. C. C., 2005: Compendium on methods and tools to evaluate impacts of and vulnerability and adaption to climate change. Final draft report, 143-155.

Thompson, D. W. J., and J. M. Wallace, 1998: The Arctic Oscillation signature in the wintertime geopotential height and temperature fields. *Geophys. Res. Lett.*, **25**, 1297-1300, doi:10.1029/98GL00950.

Walsh, J. E., A. S. Phillips, D. H. Portis, and W. L. Chapman, 2001: Extreme cold outbreaks in the United States and Europe, 1948-99. *J. Climate*, **14**, 2642-2658, doi:10.1175/1520-0442(2001)014<2642:ECOITU> 2.0.CO;2.

Westby, R. M., Y.-Y. Lee, and R. X. Black, 2013: Anomalous temperature regimes during the cool season: long-term trends, low-frequency mode modulation, and representation in CMIP5 simulations. *J. Climate*, **26**, 9061-9076, doi:10.1175/JCLI-D-13-00003.1.

Wilks, D. S., 1995: *Statistical Methods in the Atmospheric Science*. Academic Press, 284-300.

Woo, S. H., B. M. Kim, J. H. Jeong, S. J. Kim, and G. H. Lim, 2012: Decadal changes in surface air temperature variability and cold surge characteristics over northeast Asia and their relation with the Arctic Oscillation for the past three decades (1979-2011). *J. Geophys. Res.*, **117**, D18117, doi:10.1029/2011JD016929.

Zhang, Y., J. M. Wallace, and D. S. Battisti, 1997: ENSO-like interdecadal variability: 1900-93. *J. Climate*, **10**, 1004-1020, doi:10.1175/1520-0442(1997)010<1004:ELIV>2.0.CO;2.

Zhang, X., and Coauthors, 2005: Trends in Middle East climate extreme indices from 1950 to 2003. *J. Geophys. Res.*, **110**, D22104, doi:10.1029/2005JD006181.