



Intensity-duration-frequency relationship of WBGT extremes using regional frequency analysis in South Korea[☆]

Ju-Young Shin, Kyu Rang Kim^{*}, Jong-Chul Ha

Applied Meteorology Research Division, National Institute of Meteorological Sciences, South Korea



ARTICLE INFO

Keywords:

Extreme high temperature
Heat-related stress
Heatwave
Wet-bulb globe temperature
Return level

ABSTRACT

The risk levels of heat-related extreme events need to be estimated for prediction and real-time monitoring to mitigate their impacts on air quality, public health, the ecosystem, and critical infrastructure. Many countries have adopted meteorological variable base thresholds for assessing the risk level of heat-related extreme events. These thresholds provide an approximate risk level for a specific event but do not consider its intensity and duration in the risk assessment. The current study provides a statistical tool to assess the risk of heat-related extreme events while concurrently considering their intensities and durations based on the wet-bulb globe temperature (WBGT). To this end, the intensity-duration-frequency (IDF) relationship of the extreme WBGT in South Korea was derived. Regional frequency analysis was employed to understand the IDF relationship. Return levels of heat-related extreme events in South Korea were calculated and their characteristics were investigated based on the annual maximum WBGT observations. The results showed that the IDF relationship could provide the risks of heat-related extreme events while concurrently considering their intensities and durations. The extreme WBGT in South Korea was used to categorize two regions such as coastal and inland based on their statistical characteristics. The return levels of the annual maximum WBGT events were found to vary largely by location. The return levels corresponding to 32 °C with 3-h duration for stations in the coastal and inland regions ranged from 1- to 100-years and 3- to 1000-years, respectively. Mean values of return levels for heatwave events in Seoul, Incheon, Daejon, Gwangju, Daegu, and Busan were 2.8-, 8.4-, 15.3-, 2.8-, 1.6-, and 2.2-years, respectively. The return levels of heatwaves for the warmer cities are smaller than those for cooler cities. The return levels of the heatwave events in South Korea showed a significant increasing trend in several cities, supporting the notion that the impact of heatwave events on South Korea might become more severe in the future.

1. Introduction

Global warming trends indicate that many regions will undergo significant warming that could have negative impacts on air quality, public health, the ecosystem, and critical infrastructure (Trenberth and Josey, 2007). Many studies have reported that extreme weather events, such as heat waves, will continue to become increasingly frequent as well as more severe because of global warming (Meehl and Tebaldi, 2004; Perkins et al., 2012; Willett and Sherwood, 2012; Sherwood and Fu, 2014). A large number of heatwaves have already occurred in different regions around the world causing considerable damage (Zhang et al., 2013; Chien et al., 2016; Green et al., 2016; Lee et al., 2018; Lim

et al., 2019). Projections under climate change scenarios show the possibility of more intense and frequent extreme heat events occurring in the future (Kim et al., 2018; Lee and Min, 2018; Newth and Gunasekera, 2018). Thus, monitoring and predicting the risk of the heat-related extreme events is critical to mitigate their impacts (Ebi and Schmier, 2005; Toloo et al., 2013; Zhang et al., 2014; Jänicke et al., 2020).

For an ongoing or expected heatwave event, it is crucial to estimate its risk or risk level that is magnitude of risk categorized by a metric table for prediction and real-time monitoring to mitigate its impact (Bobb et al., 2011; Lowe et al., 2011). In addition, the risk or risk level of future heat-related stress events for outdoor activities should be

[☆] This work was funded by the Korea Meteorological Administration Research and Development Program “Advanced Research on Biometeorology” under Grant (KMA2018-00620). This work did not employ any information about human subjects or animal experiment.

* Corresponding author. Applied Meteorology Research Division, National Institute of Meteorological Sciences, 33, Seohobuk-ro, Seogwipo-si, Jeju-do, 63568, South Korea.

E-mail address: krk9@kma.go.kr (K.R. Kim).

estimated to plan appropriate work times (Grundstein et al., 2015). It is, however, difficult to construct a heat-related risk assessment model to assess the impact on mortality because of the complexities associated with the relevant factors, such as meteorological, demographical, behavioral, and regional factors, as well as data availability (Medina-Ramón and Schwartz, 2007; Gascon et al., 2016; Son et al., 2016; Kim and Kim, 2017). There are many definitions of the heatwave, and the heat-related risk assessment models provide different results depending on the use of heatwave definition (Robinson, 2001; Smith et al., 2013; Scalley et al., 2015; Lee et al., 2016). Such a model also needs to be annually updated for calibration because of the large variabilities of the various factors (Guo et al., 2012). Certain early heat warning systems use weather prediction and heat-related risk assessment models (Ebi et al., 2004; Åström et al., 2015; Matzarakis et al., 2020). Because of the aforementioned limitations of early warning systems, meteorological variable base thresholds were widely employed for assessing the risk level of heat-related events. Since the risk level categories of heat-related extreme events in the threshold model are predefined, it only provides an approximate risk level for the specific event. For example, all events for which the values of the temperature indicator exceed a threshold have the same risk level. Therefore, practical assessments of the risk level of heat-related extreme events based on meteorological variables are essential, particularly for developing countries.

Most of threshold approach that provides risk levels based on a metric table using the predefined thresholds cannot concurrently consider the intensity and duration of a heat-related event, but the mortality-based risk assessment model can consider both aspects (Lee et al., 2016; Xu et al., 2018). This is important because the risks of heat-related events are strongly associated with their respective intensities and durations (Nairn and Fawcett, 2015; Scalley et al., 2015; Lee et al., 2016). For example, heat wave definitions used in previous studies differ in three respects: temperature indicator, intensity, and threshold (Robinson, 2001; Smith et al., 2013; Kent et al., 2014; Chen et al., 2015; Xu and Tong, 2017; Xu et al., 2018). Thus, heatwave events are now represented in terms of their respective intensities and durations (Raei et al., 2018). Hence, the consideration of these two factors in the risk assessment would lead to reliable predictions of heat-related extreme events.

Frequency analysis has been widely used for the risk assessment of extreme events, such as floods, extreme precipitation, and extreme temperatures (Katz et al., 2002; Sveinsson et al., 2002; Shin et al., 2015). The intensity-duration-frequency (IDF) relationship can be derived from the frequency analysis of the variable of interest (Koutsoyiannis et al., 1998; Javelle et al., 2002; Ouarda et al., 2019). In the IDF relationship, the return level (also termed as the return period or the average recurrent interval) of the event is considered as the magnitude of risk. This approach can assess the risk of a specific event as well as consider its intensity and duration. The IDF relationship has been employed in several previous studies on heatwave events (Khalil et al., 2005; Nogaj et al., 2006; Zahid et al., 2017; Raggad, 2018). These studies used the annual maximum air temperature as a temperature indicator as well as the multiple daily durations of the events. The thermal environment consists of various meteorological variables, such as air temperature, solar radiation, humidity, and wind speed, which can affect human thermoregulation (Barnett et al., 2010; Blazejczyk et al., 2012; Kamp-

The WBGT is not only a standard indicator for heat-related stress within sub-daily durations, but is also used as the temperature indicator in the definition of a heatwave (Parsons, 1999; Heo et al., 2019). Thus, using the WBGT as the temperature indicator is logical because of its association with heat-related events of short and long durations. However, since additional meteorological variables needs in calculation of WBGT unlike air temperature, the WBGT observations may be available in limited areas.

Given this background, the current study presents a new statistical tool to assess the risk of heat-related extreme events while concurrently considering their intensities and durations based on the observed WBGT. The IDF relationship of the annual maximum WBGT (AMWG) is used as the statistical tool in this study. Regional frequency analysis (RFA) is employed to obtain the IDF relationship. The return levels of AMWG events in South Korea are calculated, and their characteristics are investigated. The approach proposed in this study can provide an alternative way to assess the risk of the heat-related extreme events based on observations of meteorological variables. Thus, it enhances our capacity to mitigate the impacts of heat-related extreme events via monitoring and prediction of their risk. To the best of our knowledge, statistical characteristics of WBGT extremes have not been previously explored via frequency analysis. In addition, the RFA has not been used in the frequency analysis of extreme temperatures. The results of this study will improve our understanding of the use of WBGT extremes in the risk assessment of heat-related extreme events.

2. Methods

2.1. Data

WBGT is an index that is calculated as the weighted average of air temperature, natural wet bulb temperature, and black globe temperature (Yaglou and Minard, 1957) as follows:

$$\text{WBGT} = 0.7T_w + 0.2T_g + 0.1T_a \quad (1)$$

where T_w , T_g , and T_a are the natural wet bulb temperature, black globe temperature, and air temperature, respectively. As the black globe temperature is not regarded as a standard element, it is often not measured at weather stations. In addition, as the wet bulb temperature is measured by a specific hygrometer type at weather stations, its data availability is often inconsistent for different recording periods and stations. Thus, empirical and physical estimation models using other meteorological variables have been applied to obtain WBGT estimates (Epstein and Moran, 2006; Liljegren et al., 2008; Lemke and Kjellstrom, 2012). In this study, the empirical estimation model suggested by the Korea Meteorological Administration (KMA) is used to obtain the WBGT for South Korea. This model requires T_a and T_w , and has been known to perform well in the estimation of the WBGT in South Korea (Lee et al., 2019). The WBGT model proposed by the KMA is as follows:

$$\text{WBGT} = -0.2442 + 0.55399T_w + 0.45535T_a - 0.0022T_w^2 + 0.00278T_wT_a \quad (2)$$

T_w in Equation (2) is calculated using the equation suggested by Stull (2011), as seen below.

$$T_w = T_a \tan^{-1} [0.151977(RH + 8.313659)^{1/2}] + \tan^{-1}(T_a + RH) - \tan^{-1}(RH - 1.67633) + 0.00391838RH^2 \tan^{-1}(0.023101RH) - 4.686035 \quad (3)$$

mann et al., 2012). The wet-bulb globe temperature (WBGT) is a temperature indicator that combines the effects of air temperature, humidity, solar radiation, and wind speed (Yaglou and Minard, 1957).

where RH is the relative humidity. Hourly mean T_a and RH observations made at 91 stations in South Korea were used in this study (data.kma.go.kr). All these stations are maintained and operated by the KMA. Hence,

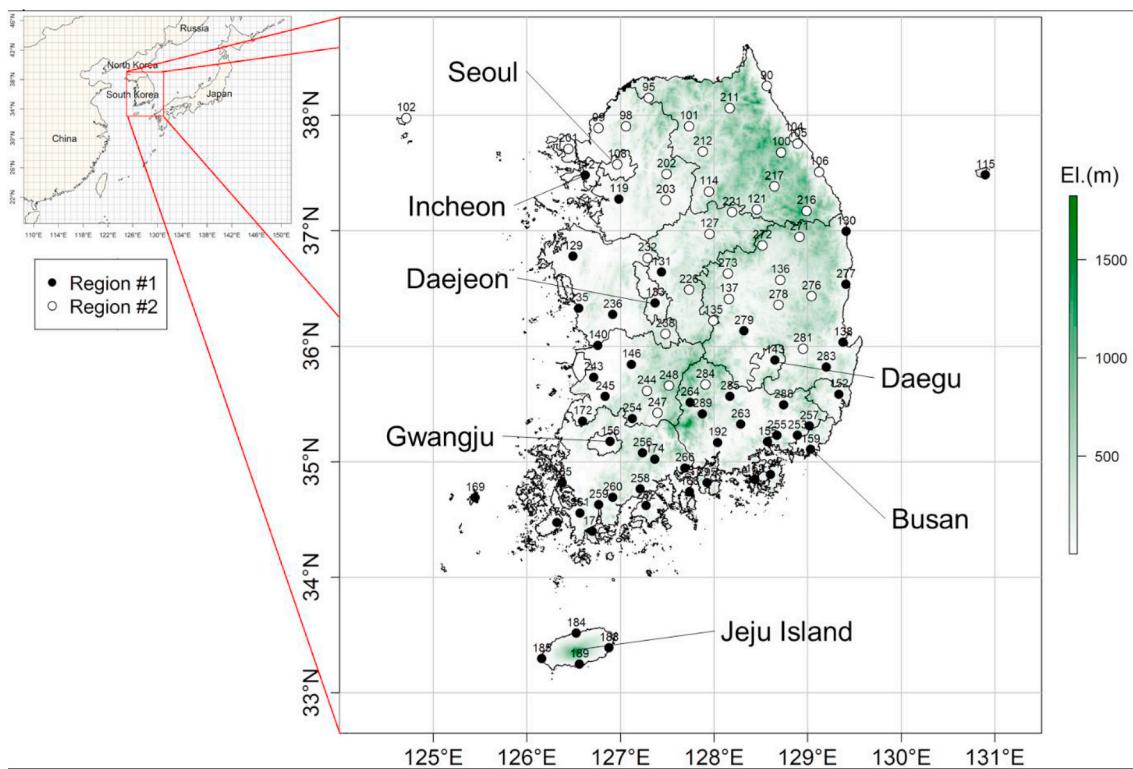


Fig. 1. Locations of the studied weather stations with elevation. Note that closed and open circles indicate the stations belonging to region #1 and #2 which stations in each region are statistically homogenous based on statistical analysis, respectively. The detailed information about the weather stations is presented in Table 1.

the quality of the data would be credible. The locations of the used stations are presented in Fig. 1, and their information (name, code, record length, latitude, longitude, and elevation) appears in Table 1. The observed data of the stations ranged from 8 to 19 years.

The hourly WBGT estimates were obtained using Equation (2) with observations of T_a and RH, following which the AMWG was calculated for 50 durations. The employed durations are briefly described in Table 2. For instance, the AMWG for the given duration is the maximum value of the averaged WBGT for a moving window size for that duration within a certain year.

For the sub-daily temporal scale, 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10 h are used, and 1–40 d are employed for the daily temporal scale. The AMWGs for the sub-daily durations represent a heat-related hazard for outdoor activity and labor, while the AMWGs for the daily durations represent the risk magnitude of a heatwave (Parsons, 2006; Grundstein et al., 2014; Heo et al., 2019).

A number of definitions are used for the intensity (Raei et al., 2018; Royé et al., 2020). In the current study, the intensity of a heat-related extreme event is the mean of the hourly WBGT for the given duration, because this value is strongly associated with heat-related risk, has high accessibility, and is applicable for most heat-related events (Barnett et al., 2010; Vaneckova et al., 2011). To conduct the RFA, the hourly mean air temperature, dew point temperature, relative humidity, vapor pressure, precipitation, and solar radiation observations were also collected from the 91 weather stations.

2.2. Regional frequency analysis

The number of AMWG data of 8–19 years is insufficient to obtain a reliable estimation using at-site frequency analysis. Therefore, in the current study, RFA was employed as the methodology of the frequency analysis. RFA has been widely used in the frequency analysis of extreme events because it can provide a quantile estimation for ungauged sites, and the results of the frequency analysis are robust under the assumption

that the distribution models of the transformed sample data are similar at different sites in a statistically homogeneous region (Javelle et al., 2002; Drissia et al., 2019). In addition, this method provides more accurate estimation than at-site frequency analysis when the number of samples is insufficient (Hosking and Wallis, 1997). When the amount of data available is limited, the parameter estimates for the frequency distribution model are often inaccurate and have major uncertainties (Kyselý et al., 2011; Liang et al., 2017). The inaccurate parameter estimation will lead to an inaccurate quantile estimation, particularly for the high return level. To overcome the shortage of data, the RFA adopts the concept of a regional growth curve with a site-specific index or parameter. If the amount of data available is low, the RFA returns a more reliable and accurate quantile estimation than the conventional frequency analysis (at-site frequency analysis). When the amount of available data is sufficient to obtain reliable results, conventional frequency analysis would be more appropriate than RFA. Brief descriptions of, and information for, the RFA are presented in the following subsections.

2.2.1. Index flood method

Several RFA methods have been employed in frequency analysis of extreme events, e.g. region of interest, quantile regression and index flood (IF) methods (Kyselý et al., 2011; Chebana et al., 2014). The purpose of many of the RFA methods, such as region of interest and quantile regression, is to obtain the quantile at ungauged stations, as such, these methods may not reduce the uncertainty in the frequency analysis. The IF method can be employed to achieve the two goals of the RFA: 1) obtaining quantiles at ungauged stations and 2) overcoming the shortage of available data (Hosking and Wallis, 1997). In this study, the IF method proposed by Hosking and Wallis (1997) was used as the RFA methodology. As mentioned previously, the RFA has often been selected as an alternative frequency analysis method for extreme events when the number of samples is small. Application of the IF method can lead to a more robust analysis of the WBGT in South Korea as it can reduce the

Table 1

Information for the used stations.

Code	Name	Record length (year)	Lat. (°)	Lon. (°)	El. (m)	Code	Name	Record length (year)	Lat. (°)	Lon. (°)	El. (m)
90	Sokcho	19	38.25	128.57	18.1	202	Yangpyeong	19	37.49	127.49	47.3
95	Cheorwon	19	38.14	127.30	155.5	203	Icheon	19	37.26	127.48	80.1
98	Dongducheon	19	37.90	127.06	115.6	211	Inje	19	38.06	128.17	200.2
99	Paju	17	37.88	126.77	30.6	212	Hongcheon	19	37.68	127.88	140.0
100	Daegwallyeong	19	37.67	128.72	772.6	216	Taebaek	19	37.17	128.99	712.8
101	Chuncheon	19	37.90	127.74	76.5	217	Jeongseon	9	37.38	128.65	307.4
102	Baengnyeongdo	18	37.97	124.71	36.0	221	Jecheon	19	37.16	128.19	259.8
104	Bukgangneung	11	37.80	128.86	78.9	226	Boeun	19	36.49	127.73	175.0
105	Gangneung	19	37.75	128.89	26.0	232	Cheonan	19	36.76	127.29	81.5
106	Donghae	19	37.50	129.12	39.9	235	Boryeong	19	36.33	126.56	15.5
108	Seoul	19	37.57	126.97	85.7	236	Buyeo	19	36.27	126.92	11.3
112	Incheon	19	37.47	126.63	69.0	238	Geumsan	19	36.11	127.48	170.4
114	Wonju	19	37.33	127.95	148.6	243	Buan	19	35.73	126.72	12.0
115	Ulleungdo	19	37.48	130.90	222.4	244	Imsil	19	35.61	127.29	247.0
119	Suwon	19	37.27	126.99	34.8	245	Jeongeup	19	35.56	126.84	69.8
121	Yeongwol	19	37.18	128.46	240.6	247	Namwon	19	35.42	127.40	132.5
127	Chungju	19	36.97	127.95	116.3	248	Jangsu	19	35.66	127.52	406.5
129	Seosan	19	36.77	126.49	28.9	251	Gochanggun	11	35.43	126.70	54.0
130	Uljin	19	36.99	129.41	50.0	252	Yeonggwang	11	35.28	126.48	37.2
131	Cheongju	19	36.63	127.44	58.7	253	Gimhae	10	35.23	128.89	53.3
133	Daejeon	19	36.37	127.37	68.9	254	Sunchang	13	35.37	127.13	127.0
135	Chupungnyeong	19	36.22	129.00	243.7	255	Bukchangwon	10	35.23	128.67	48.8
136	Andong	19	36.57	128.71	140.1	256	Juam	9	35.08	127.24	74.6
137	Sangju	17	36.41	128.16	96.2	257	Yangsan	9	35.31	129.02	14.9
138	Pohang	19	36.03	129.38	3.9	258	Bosung	19	34.76	127.21	2.8
140	Gunsan	19	36.01	126.76	23.2	259	Gangjin Gun	19	34.63	126.77	12.5
143	Daegu	19	35.88	128.65	53.5	260	Jangheung	19	34.69	126.92	45.0
146	Jeonju	19	35.84	127.12	61.4	261	Haenam	9	34.55	126.57	16.4
152	Ulsan	19	35.58	129.34	82.0	262	Goheung	9	34.62	127.28	51.9
155	Changwon	19	35.17	128.57	37.6	263	Uiryeong	8	35.32	128.29	14.2
156	Gwangju	19	35.17	126.89	72.4	264	Hamyang	19	35.51	127.75	151.2
159	Busan	19	35.11	129.03	69.6	266	Gwangyang	19	34.94	127.69	86.7
162	Tongyeong	19	34.85	128.44	32.3	268	Jindogun	19	34.47	126.26	5.4
165	Mokpo	19	34.82	126.38	38.0	271	Bongwhoa	9	36.94	128.91	324.3
168	Yeosu	19	34.74	127.74	64.6	272	Yeongju	19	36.87	128.52	210.8
169	Heuksando	19	34.69	125.45	76.5	273	Mungyeong	19	36.63	128.15	170.6
170	Wando	19	34.40	126.70	35.2	276	Cheongsong	19	36.43	129.04	206.2
172	Gochang	8	35.35	126.60	52.0	277	Yeongdeok	19	36.53	129.41	40.6
174	Suncheon	8	35.02	127.37	165.0	278	Uiseong	9	36.36	128.69	81.8
175	Jindo	17	34.47	126.32	478.5	279	Gumi	19	36.13	128.32	48.9
184	Jeju	19	33.51	126.53	20.5	281	Yeongcheon	19	35.98	128.95	93.8
185	Gosan	19	33.29	126.16	71.5	283	GyeongJu	19	35.82	129.20	39.2
188	Seongsan	19	33.39	126.88	20.3	284	Geochang	19	35.67	127.91	226.0
189	Seogwipo	19	33.25	126.57	49.0	285	Hapcheon	19	35.57	128.17	32.0
192	Jinju	19	35.16	128.04	30.2	288	Miryang	19	35.49	128.74	11.2
201	Ganghwa	19	37.71	126.45	47.8						

Table 2

Summary of employed durations.

Category	Durations	Number of durations	Target event
Sub-daily	1–10 h	10	Outdoor activities and labor in daytime
Daily	1–40 d	40	Heatwave events

uncertainty derived from the small number of samples. In the IF method, the sample data are transformed by an index. In the current study, the mean of the extreme WBGT was used as the index. The transformed AMWG (tAMWG) was calculated using Equation (4) (Hosking and Wallis, 1997).

$$t\text{AMWG}_{i,d}^j = \frac{\text{AMWG}_{i,d}^j}{\overline{\text{AMWG}}_d^j} \quad (4)$$

where $t\text{AMWG}_{i,d}^j$, $\text{AMWG}_{i,d}^j$, and $\overline{\text{AMWG}}_d^j$ are the transformed i th AMWG at the j th station and the d th duration, the i th AMWG at the j th station and the d th duration, and the mean of the AMWG at the j th station and the d th duration, respectively. The tAMWG is used as the random

variable in the RFA.

The IF method adopts the regional probability distribution concept, namely that the probability distributions for the variable of interest at different stations in a statistically homogeneous region are identical. The regional probability distribution for a given duration can be expressed by Equation (5) (Hosking and Wallis, 1997).

$$y = F_d(t\text{AMWG}_d | \theta_d^R) \quad (5)$$

where y , $F_d(\cdot)$, and θ_d^R are the non-exceedance probability with respect to a given $t\text{AMWG}_d$, a given cumulative distribution function for duration d , and the regional parameters of the regional probability distribution for duration d , respectively. The quantile function of the regional probability distribution, namely the regional growth curve, is given by Equation (6). The quantile at a specific station can be calculated using Equation (7).

$$\widehat{t\text{AMWG}}_d = F_d^{-1}(y | \theta_d^R) \quad (6)$$

$$\widehat{\text{AMWG}}_d^j = \overline{\text{AMWG}}_d^j \times \widehat{t\text{AMWG}}_d = \overline{\text{AMWG}}_d^j \times F_d^{-1}(y | \theta_d^R) \quad (7)$$

The return level of $\widehat{\text{AMWG}}_d^j$ is defined as $\frac{1}{1-y}$. Detailed information about the IF method and its procedure can be found in Hosking and

Wallis (1997). The IF method is briefly described in the following subsection.

2.2.2. Identification of homogeneous regions

A cluster analysis must be carried out to delineate a homogenous region. The performance of various clustering methods such as K-means, ward, self-organizing map, and the fuzzy C-means method, were evaluated for delineation of statistically homogeneous regions in RFA (Lin and Chen, 2006; Satyanarayana and Srinivas, 2008; Yang et al., 2010; Ngongondo et al., 2011; Rahman et al., 2013b; Basu and Srinivas, 2014; Nam et al., 2015; Kulkarni, 2017). The research community does not consider any of the particular clustering methods for RFA superior (Abdi et al., 2017). All clustering methods have pros and cons for identifying clusters depending on the employed data sets. The K-means clustering method is one of the most popular methods and has been frequently employed in cluster analysis because of its simplicity and ease of interpretation (Nagpal et al., 2013). In addition, because the K-means clustering method has one parameter (k) that represents the number of clusters, it may provide a more consistent performance for other data sets than other methods with a larger number of parameters that can be tuned by the user. Even the parameter k can be optimized by the indices for measuring the suitability of the cluster analysis result. Application of the K-means clustering method leads to more consistent results that can be easily reproduced in other studies. Thus, the K-means clustering method was used as the clustering algorithm in the current study (MacQueen, 1967). Sixty-eight indices that can represent the characteristics of extreme WBGT events were used in the clustering analysis. These indices are presented in Table 3. The silhouette index was employed to find the optimal number of clusters (Rousseeuw, 1987). Because the statistical characteristics of extreme WBGT events and the AMWGs in the same clusters (based on the results of the cluster analysis) can be different, the homogeneity of the AMWGs in the same clusters should be evaluated. Hosking and Wallis (1997) suggested using the H index to measure the heterogeneity of data sets in a cluster based on regional L-moment ratios. The H index is given by the following equation:

$$H = \frac{(V - \mu_V)}{\sigma_V} \quad (8)$$

where V , μ_V , and σ_V are the weighted standard deviation of the at-site sample L-coefficient of variation (L-CV), mean of the simulated V , and standard deviation of the simulated V , respectively. V is calculated using Equation (9).

$$V = \left\{ \frac{\sum_{j=1}^N n^j (t^j - t^R)^2}{\sum_{j=1}^N n^j} \right\}^{\frac{1}{2}} \quad (9)$$

where N , n^j , t^j , and t^R are the number of stations, the number of samples at the j th station, the L-CV at the j th station, and the regional L-CV ($= \frac{\sum_{j=1}^N n^j t^j}{\sum_{j=1}^N n^j}$), respectively. μ_V and σ_V were obtained by simulating a large number of realizations for the region with the same number of stations. In the simulation, the kappa distribution was fitted to the sample sets of

each station, and then the realizations were resampled from the fitted kappa distribution. The simulated values of V were obtained from the simulated realizations for the region. The region was regarded as “acceptably homogeneous” if $H < 1$, “possibly heterogeneous” if $1 \leq H < 2$, and “definitely heterogeneous” if $H \geq 2$.

2.2.3. Parameter estimation method for regional probability distribution

There are two main reasons for adopting the RFA method: 1) To obtain reliable results with a limited amount of data (<30), and 2) to estimate the risk level corresponding to the variable of interest in locations where the variable is temporally or irregularly recorded, or no data is available (Hosking and Wallis, 1997; Javelle et al., 2002; Chebana et al., 2014). In this study, regional frequency analysis is employed due to the former reason. The regional parameter can be estimated by the L-moment (equivalent to probability weighted moment), and maximum likelihood methods (Hosking and Wallis, 1997; Sveinsson et al., 2001). The L-moment method provides superior performance for estimating parameters of the frequency distribution model with less data than the moment and maximum likelihood methods (Hosking et al., 1985; Raynal-Vellaseñor, 2012). The L-moment method can accurately represent the statistical characteristics of a population with a limited amount of data due to its parameter estimation of frequency distribution. The recorded length of AMWG data in the current study ranged from 8 to 19. These numbers are insufficient to estimate the quantile corresponding to a high-risk level. Hence, there is uncertainty surrounding the quantile estimates with respect to the high-risk level using conventional frequency analysis. The application of the L-moment method leads to a more accurate regional parameter estimation due to its characteristics, this will provide a reliable quantile estimation, particularly with respect to high risk levels. In addition, because the index flood method was derived from the L-moment method, the application of the L-moment method can avoid the deviations and implementations of other metrics, such as the H measure and Z statistic, based on other parameter estimation methods. The L-moment method is appropriate for estimating regional parameters of frequency distribution for extreme WBGT in this study. Thus, the regional parameters were estimated based on the L-moment in the current study. The L-moment is the linear combination of moments based on order statistics. The regional L-moment ratios were employed for the regional parameter estimation. The m th-order regional L-moment ratios were calculated using Equation (10) (Hosking and Wallis, 1997).

$$t_m^R = \frac{\sum_{j=1}^N n^j t_m^j}{\sum_{j=1}^N n^j} \quad (10)$$

where t_m^j is the m th order L-moment ratio at the j th station. The 2nd-, 3rd-, and 4th-order L-moment ratios are the L-CV, L-coefficient of skewness, and L-coefficient of kurtosis, respectively. The regional parameters of the regional probability distribution were estimated based on the sample L-moment ratios. In the current study, the generalized logistic (GLO), generalized extreme value (GEV), Pearson-type 3 (PE3), and generalized normal (GNO) distributions were tested as the probability distribution candidates. These probability distribution models have been broadly used for modeling extreme events (Rahman et al.,

Table 3
Indices for representing the characteristics of each station.

Category	Variable	Number of indices
Geographical information	Latitude, Longitude, and Elevation	3
Mean	Temperature, Dew point temperature, Relative humidity, Wind speed, Vapor pressure, Precipitation, Solar radiation, and WBGT	8
Standard deviation	Temperature, Dew point temperature, Relative humidity, Wind speed, Vapor pressure, Precipitation, Solar radiation, and WBGT	8
Daily maximum	Temperature, Dew point temperature, Relative humidity, Wind speed, Vapor pressure, Precipitation, Solar radiation, and WBGT	8
Daily minimum	Temperature, Dew point temperature, Relative humidity, Wind speed, Vapor pressure, Precipitation, Solar radiation, and WBGT	5
Monthly mean	WBGT	12
Monthly maximum	WBGT	12
Monthly minimum	WBGT	12

2013a; Hu et al., 2019; Igor et al., 2019). The probability density function and information about the four distribution models are presented in Appendix A. In addition, detailed information about the parameter estimation method for these distributions using L-moments can be found in Hosking and Wallis (1997).

2.2.4. Selection of the appropriate probability distribution model

When using RFA, the appropriate distribution model must be identified for the region. In RFA, one distribution represents all the stations in that region. Thus, the conventional goodness-of-fit tests and measures are not suitable for the goodness-of-fit of distribution models. Hosking and Wallis (1997) suggested using the Z measure (Z^{DIST}) to evaluate the goodness-of-fit of the fitted distribution, which has scale and shape parameters, to the sample data sets in a region. Z^{DIST} is the bias-corrected and standardized distance between a sample's regional L-kurtosis and the theoretical L-kurtosis of the distribution model corresponding to a given L-skewness. The distribution models can successfully reproduce L-skewness because of the existence of the shape parameter. Hence, L-kurtosis would be a good indicator of the appropriateness of the distribution model to the observed data sets. Z^{DIST} can be calculated using Equation (11).

$$Z^{DIST} = \frac{(t_4^{DIST} - t_4^R + B_4)}{\sigma_4} \quad (11)$$

where t_4^{DIST} , B_4 , and σ_4 are the theoretical L-kurtosis of the distribution model corresponding to a given L-skewness, the bias in t_4^R , and the standard deviation of t_4^R , respectively. B_4 and σ_4 can be obtained from the simulation. When $|Z^{DIST}| \leq 1.64$, the distribution model is an adequate fit for the regional probability distribution. Detailed information about the simulations for B_4 and σ_4 can be found in Hosking and Wallis (1997). The number of realizations in the simulation in the current study was 3000. In addition, the L-moment ratio diagram was used to evaluate the goodness-of-fit of the fitted probability distribution model to the sample data. The L-moment ratio diagram presents the theoretical L-moment ratios of the probability distribution models. Since this parameter allows a comparison of the L-moment ratios of the sample data to the theoretical L-moment ratios of the distribution models, the overall goodness-of-fit of the distribution to the sample data can be assessed.

3. Results

3.1. Homogeneous regions for WBGT

The silhouette index was calculated from 2 to 20 clusters using the K-means clustering method for the employed meteorological indices that represent the characteristics of the WBGT at each station. The results are presented in Fig. S1 (in Supplement). When the number of clusters equaled 2, the value of the silhouette index was the highest. Thus, the number of homogeneous regions should be limited to one or two regions. We tested the appropriateness for one and two regions. Because 12 of 91 stations were discordant as per the discordancy measures (not shown in this study), they were excluded from the RFA if one region was selected. On the other hand, all the stations could be used in the RFA if two regions were selected. In addition, the result of cluster analysis was compared to the Koppen climate classification map in South Korea to verify physiogeographic consistency. Based on the results of clustering analysis for two regions, the one region located in coastal area is classified as warm continental climate/humid continental climate, humid subtropical climate, and warm oceanic climate/humid subtropical climate while another region located in inland area is classified warm continental climate/humid continental climate (Peel et al., 2007; Yun et al., 2012). The employed clustering analysis would provide physiogeographically reasonable results. Hence, two regions were adopted for the RFA in the current study. To carry out the RFA, the sample sets of the stations in the same region should be statistically homogeneous. To assess the heterogeneity of the selected regions, the heterogeneity measures for the AMWG data in the 2 regions for 50 durations were calculated (Fig. 2). The figure shows that the values of the all the H measures for the durations were lower than 1. These results support that the AMWG data in the same region for all the tested durations were not heterogeneous. Hence, the selection of regions via the K-mean method and the silhouette index was reasonable, and the same selection was then applied for the RFA of the AMWG. Regarding the sub-daily calculations, the values of the H measures were found to be similar in each region. The values of H measures for daily durations are differed depending on the duration. A large negative H measure indicates a strong correlation between stations. Since the AMWGs for long durations at different stations were very similar, large negative values were observed for the H measure at the daily temporal scale. The regions the stations belong to are presented in Fig. 1. The first region (region #1) covered the coastal area, while the second region (region #2) included inland and mountainous areas. The two regions presented in Fig. 1 were employed in the RFA of the WBGT in the current study.

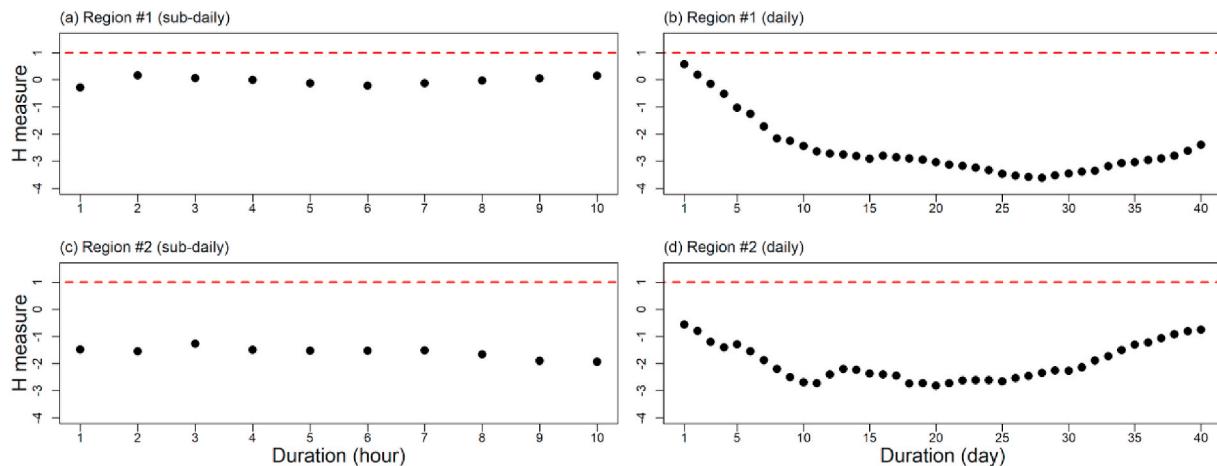


Fig. 2. Heterogeneity measure (H measure) for each duration and region based on the annual maximum WBGT. The black dots and red dash lines indicate the H measure corresponding to the duration and the threshold to determine whether the data sets are statistically heterogeneous or not, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

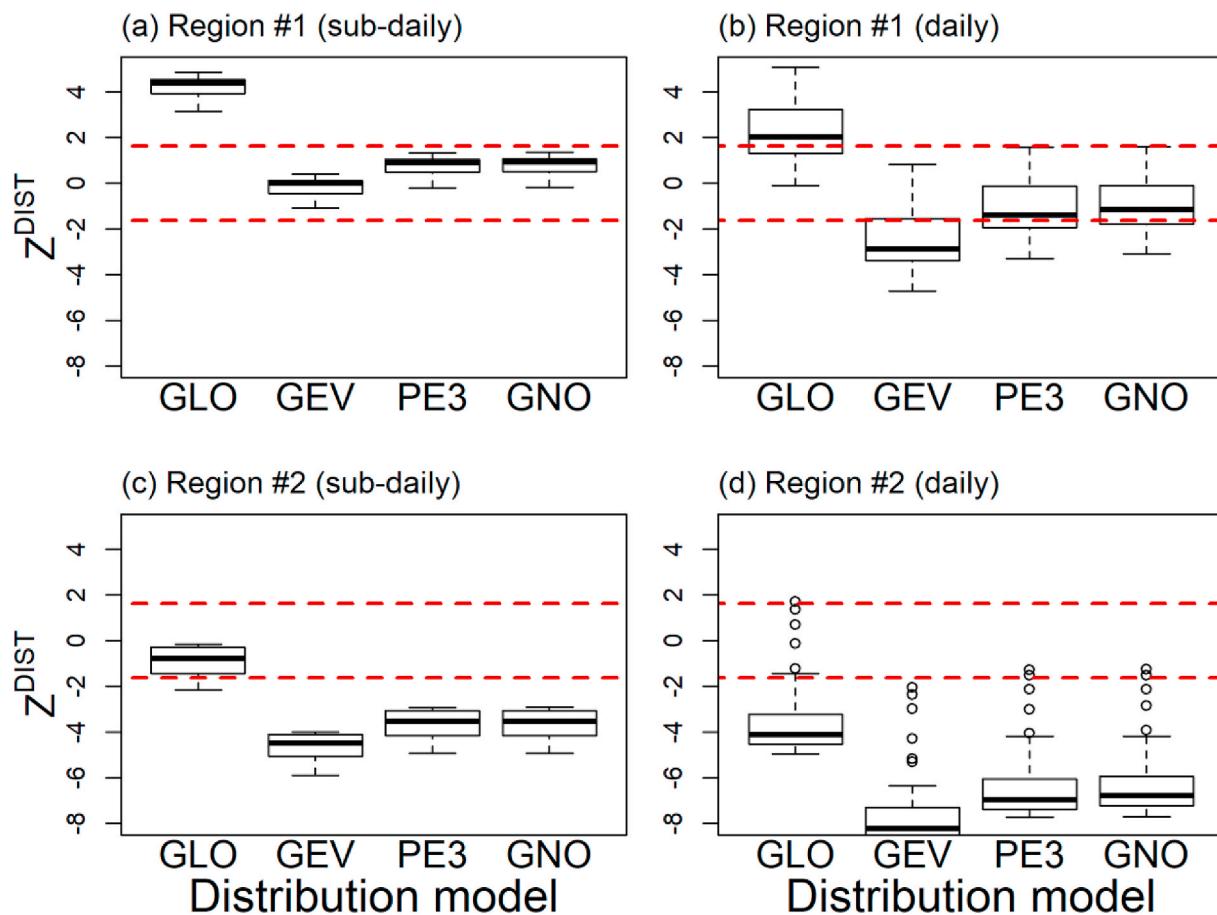


Fig. 3. Boxplots of Z^{DIST} measures for the tested probability distribution and each region. Note that the red lines indicate the acceptance thresholds for the distribution models based on Z^{DIST} . GLO, GEV, PE3, and GNO indicate the generalized logistic, generalized extreme value, Pearson-type 3, and generalized normal distributions, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

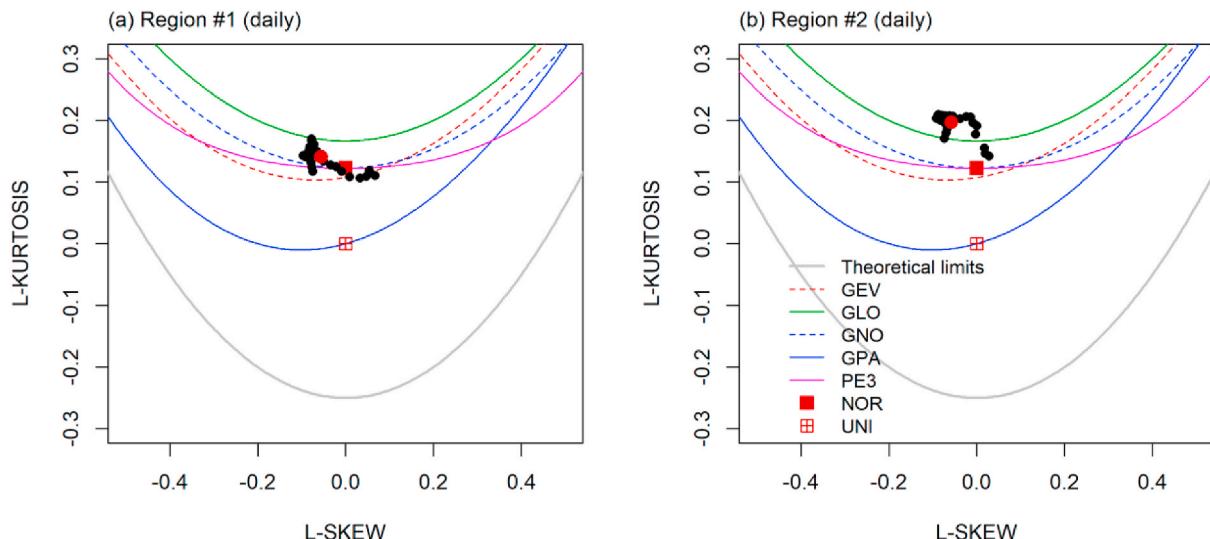


Fig. 4. L-moment ratio diagram for regions #1 (a) and #2 (b) based on the daily temporal scale, respectively. The black and red circles indicate the L-moment ratios for each duration and the mean of the L-moment ratios, respectively. GEV, GLO, GNO, GPA, PE3, NOR, and UNI indicate the generalized extreme value, generalized logistic, generalized normal, generalized Pareto, Pearson-type 3, normal, and uniform distributions, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

3.2. Intensity-duration-frequency curve of extreme WBGT

The values of Z^{DIST} for the candidate distribution models, regions, and durations were computed and are presented via boxplots in Fig. 3. Each box includes Z^{DIST} for all the durations (10 for the sub-daily durations and 40 for the daily durations). When Z^{DIST} of the candidate distribution was between -1.64 and 1.64 , the distribution was considered to have sufficient goodness-of-fit for the sample data in a given region. For region #1, the GEV, PE3, and GNO distributions provided sufficient goodness-of-fit for the sub-daily durations. For the daily durations, the GNO distribution seemed to be the most appropriate among the candidate distributions. The PE3 distribution was also considered as an applicable distribution model for region #1 where is the coastal area. Unlike the sub-daily durations, the GEV distribution provided insufficient goodness-of-fit for the daily durations. For region #2, the GLO distribution was the most appropriate distribution model for the sub-daily durations. All boxes for the candidate distribution models in Fig. 3 (d) are located outside the red dashed lines (acceptance thresholds). This result indicates that the four distribution models may not successfully reproduce the statistical characteristics of the AMWG data in region #2 for the daily durations. The GLO distribution seems to be the most appropriate model among the four candidate distribution models.

While the distribution models for the sub-daily durations in the two regions were easily selected based on Z^{DIST} , the selection of the distribution model for the daily durations using Z^{DIST} was ambiguous. To identify the most appropriate distribution for the daily durations, the L-moment ratio diagram was employed (Fig. 4). The red circles are the means of the L-moment ratios for all daily durations. As shown in Fig. 4 (a), the red circle is located near the blue dashed and purple solid lines representing the statistical characteristics of the GNO and PE3 distributions, respectively. Because the distance between the red circle and the blue dashed line is the shortest, the GNO distribution is the most appropriate model for the daily durations in region #1. For region #2, the means of the L-moment ratios are located above the green solid line representing the GLO distribution. Since they are the closest to the green solid line, the GLO distribution would be the most appropriate distribution for the daily duration in region #2. Therefore, in the current study, the GNO and GLO distributions were selected for the RFA of regions #1 and #2, respectively.

The IDF relationship is often described by IDF curves. The IDF curve at a specific station can be obtained from the RFA. Because one growth curve (an inverse function of the regional probability distribution) applies for one region in the RFA, each region has a distribution model, and the shapes of the distribution models of individual stations in the same region are very similar. Investigating the difference between the statistical characteristics of the AMWG in different regions is critical to understand the IDF relationship of the AMWG in South Korea.

For comparison between regions #1 and #2, the regional means of the IDF curves in the two regions for the sub-daily durations are presented in Fig. 5. When the duration or intensity increases, the return levels increase. The return levels of the same WBGT values for region #1 were smaller than those for region #2. While the differences between the return levels of regions #1 and #2 for the not high WBGTs (e.g., 30 and 31 °C) were small, the differences for the high WBGTs (e.g., 32 and 33 °C) were very large. For example, the return level of 33 °C for region #2 was approximately thrice as large than that for region #1. The high return level for region #2 was due to the lower value of the WBGT in this region compared to that in region #1. The regional means of the IDF curves in the two regions for the daily durations are presented in Fig. 6. Similar to the results for the sub-daily durations, the return levels increased when the duration or intensity increased. The slopes of the return levels over the daily durations were steeper than those of the sub-daily durations. For instance, the return level of 28 °C was approximately 10 yr at 1 d and approximately 3310 yr at 40 d for region #2. Thus, the difference between the return levels at 1 and 40 d was 3300 yr. For region #1, the differences were smaller than those for region #2, but they were still larger than that for the sub-daily durations. The return levels of the same WBGT value for region #1 were much smaller than those for region #2.

3.3. Return levels of heat stress and heatwave in South Korea

To investigate the differences in the characteristics between the two regions, the regional means of the IDF curves of the two regions were compared. The risk levels of each station were also different, and they can be estimated by the approach proposed in the current study. For a comparison of the return levels between stations, the return levels of 32 °C for the sub-daily durations (1–3 h) at all the studied stations in South Korea are presented in Fig. 7. Because the current study focuses on the

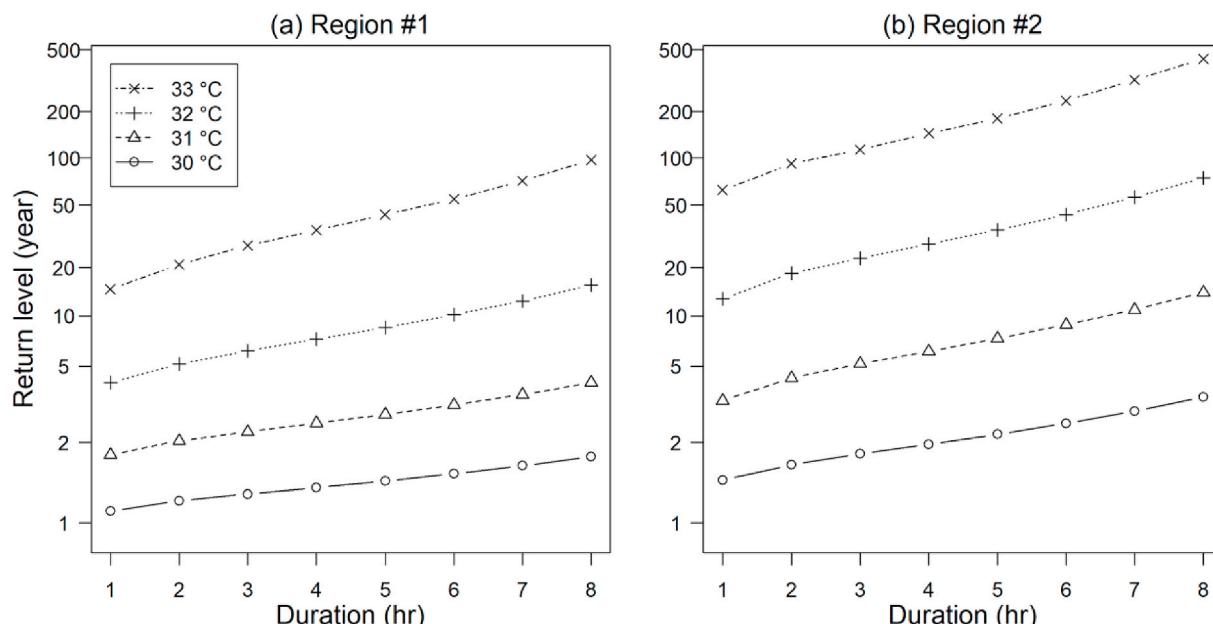


Fig. 5. Regional means of the IDF curve of the annual maximum WBGT for sub-daily durations in each region in South Korea.

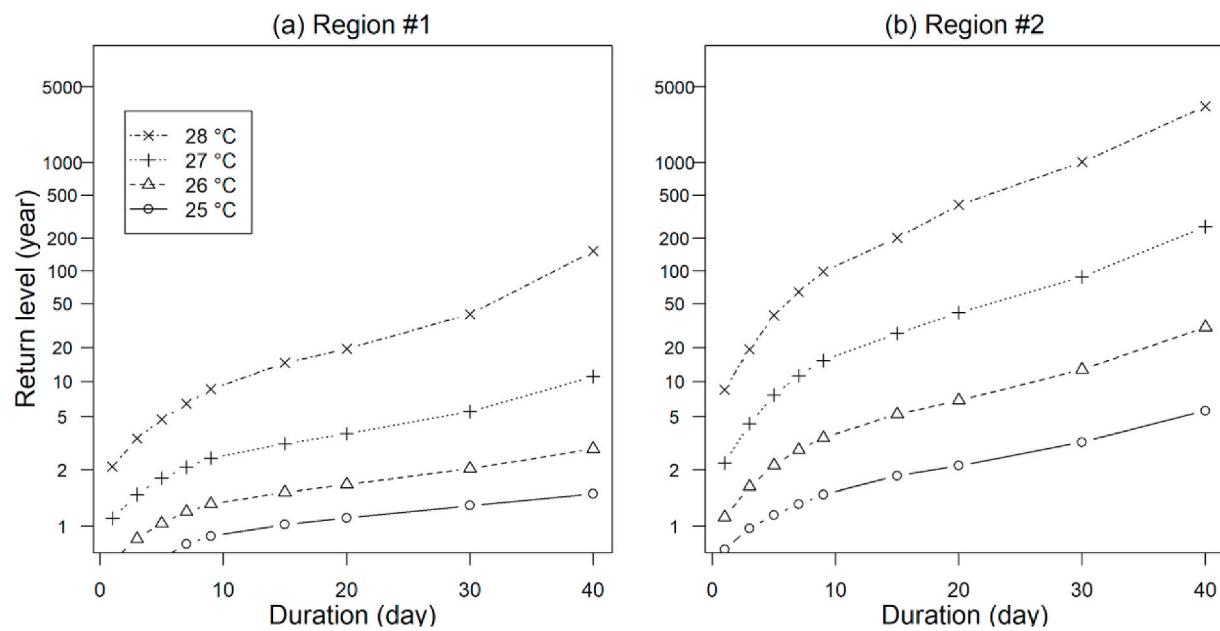


Fig. 6. Regional means of the IDF curve of the annual maximum WBGT for daily duration in each region in South Korea.

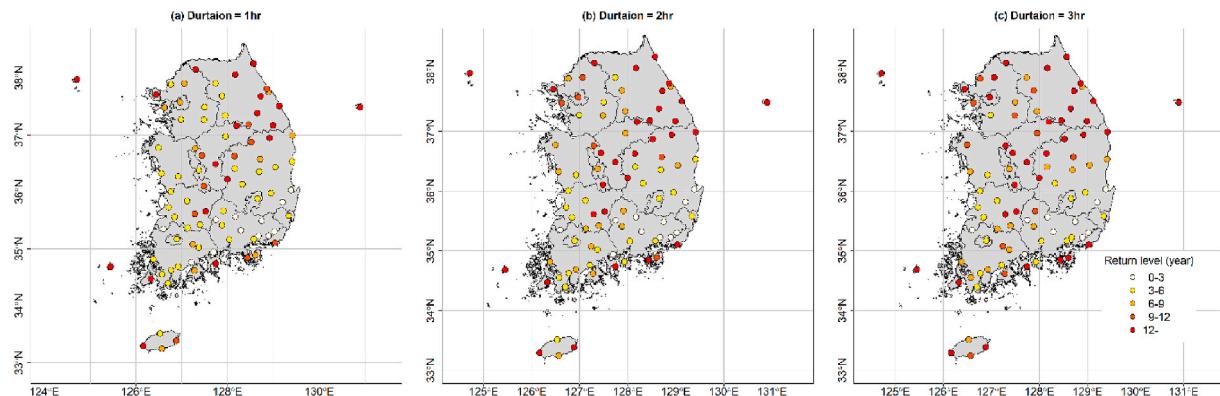


Fig. 7. Return levels of 32 °C for 1- to 3-h durations at the studied stations in South Korea.

analysis of extreme WBGT events, 32 °C, which is the maximum threshold of heat stress for outdoor work as per the ISO guideline (Parsons, 1999), was selected as the representative intensity for extreme WBGT events in Fig. 7. As shown in Fig. 7, the return levels differ depending on the locations of the stations. The stations in the southern region showed low return levels, while those in the mid-inland and northeastern regions had high return levels. Because the humidity in the mid-inland region was low and the air temperatures in the northeastern region were low due to high elevation, the stations in these regions showed low WBGT values. Because WBGT events higher than 32 °C are rare, the return level of 32 °C for WBGT was high. Thus, it is noteworthy that the statistical characteristics of extreme WBGT events can be heterogeneous and very different within a small area. For example, Jeju island has a volcanic terrain and is located in the southern region of South Korea. Four weather stations exist on this island. The return levels of these four stations differed within the small area. The changes over durations also differed depending on the locations of the stations. When the duration increased, the changes in the return levels in the southern region were small, while those in the other regions were relatively larger.

The definition of heatwave based on the WBGT was adopted by the suggestion of Heo et al. (2019) that was based on the relationship between the WBGT and mortality in South Korea. The heatwaves are

defined as occurring over two or more consecutive days with a daily maximum WBGT at or above 28 °C. The return levels of the heatwave events that occurred in six metropolitan cities: Seoul (#108), Incheon (#112), Daejeon (#133), Gwangju (#156), Daegu (#143), and Busan (#159) in South Korea from 2000 to 2018 are presented in Fig. 8. The event numbers on the x-axis are ordered based on the time sequence. For example, the largest event number indicates the latest event. Detailed information about the heatwave events (such as time, duration, mean, and return level) is provided in Tables S1–S6 (in Supplement). The temporal variations of the return levels for the heatwave events in South Korea were found to be very large. For example, the return levels for the heatwave events at Deajeon's stations ranged from less than 2 yr to higher than 1000 yr. Most of the return levels for the heatwave events in South Korea were lower than 2 yr. The latter result is expected because several heatwave events occur in the country within a year.

Based on Fig. 8, the return levels of the heatwave events at six cities in South Korea seem to have increased. To examine the trend in the risk levels of the heatwave events, the Mann–Kendall (MK) test and t-test for linear trends at the 5% significance level were employed in the current study (Shin et al., 2019). The return levels of certain heatwave events (e.g., 100, 200, and 1000 yr) in some cities were excessively large as compared to the other return levels. Because including a few large values in the trend detection test can spoil the results of the test, the

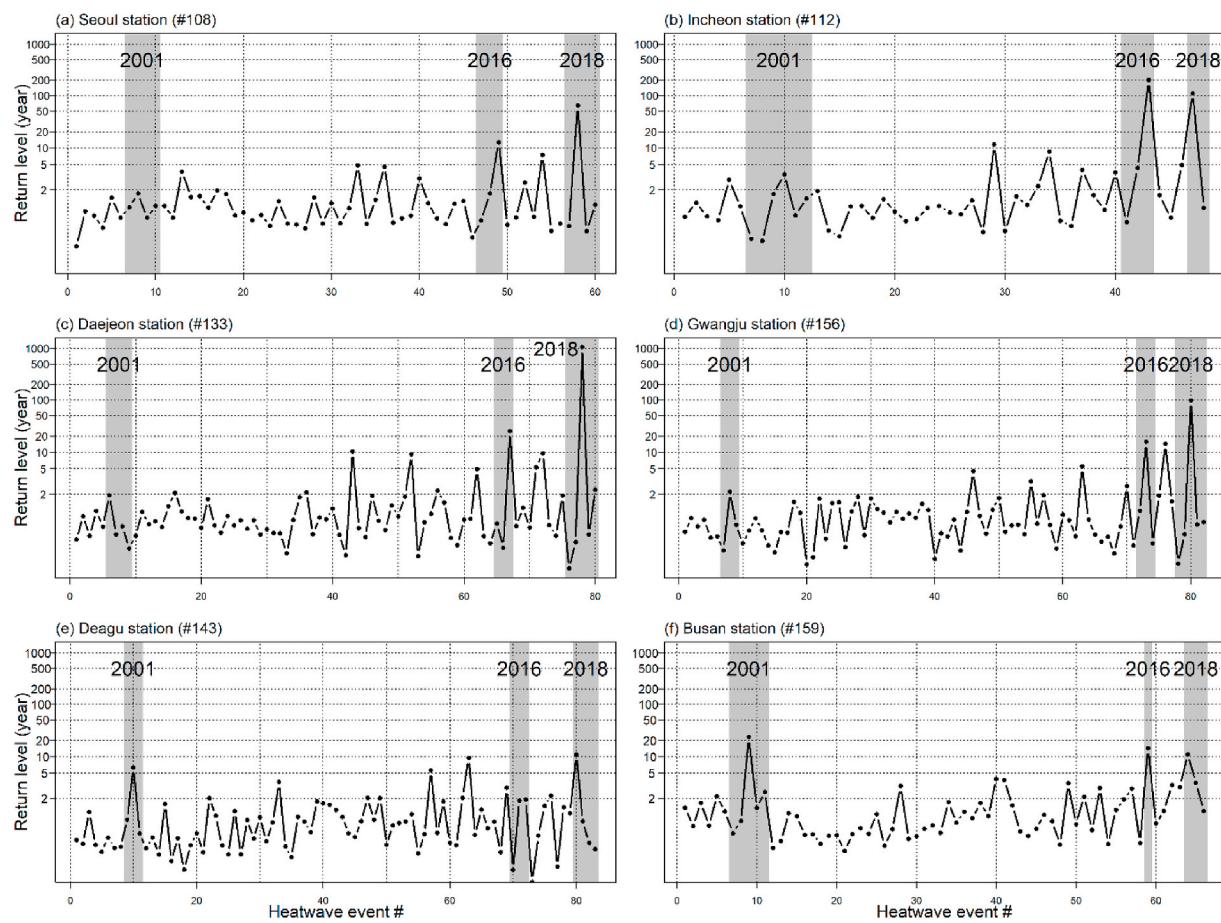


Fig. 8. Return level of heatwave events based on the daily maximum WBGT for six cities in South Korea from 2000 to 2018. Note that the event numbers on the x-axis are ordered as per the time sequence, and the grey shaded areas indicate major heat event years (2001, 2016, and 2018) in South Korea.

non-exceedance probability (y) was employed instead of the return levels to avoid excluding these events from the analysis. Fig. 9 presents the non-exceedance probabilities of the heatwave events based on the daily maximum WBGT and the results of the trend detection tests for six cities in South Korea from 2000 to 2018. Statistically significant trends were detected in the non-exceedance probability series at five stations (Incheon, Daejeon, Gwangju, Daegu, and Busan). Significant trends were detected by both tests for Incheon, Daegu, and Busan. This result indicates that the non-exceedance probabilities and return levels of the heatwave events in South Korea increased over time. These results also imply that heatwave events may become more severe in South Korea in the future.

4. Discussion

The proposed approach in this study could provide the magnitude of risk (return level) of heatwave events or heat-related stress while concurrently accounting for intensity and duration unlike the threshold approach. In addition, this approach can be easily implemented in any region for which the meteorological data are available since the proposed approach uses meteorological variables only. The proposed approach also provides quantification of the risk level for heat-related extreme events. For example, the return level of heatwave event #78 at Daejeon in 2018 is 1000 yr (see Fig. 8). Based on the threshold approach, this event appears similar to the other heatwave events. The proposed approach, however, can identify this heatwave event as an extraordinary one that requires emergency attention. Hence, the proposed approach may help decision makers to better understand and monitor heat-related extreme events. Because the proposed approach

quantifies the risk level based on statistics of the meteorological characteristics for heat-related extreme events, it should not be used to directly estimate the impact of heat-related risk on mortality and health issues. Thus, this approach may be beneficial as an add-on to warning systems using numerical weather prediction or real-time monitoring, and is an appropriate choice when mortality data are unavailable.

Previous studies used the annual maximum air temperature as the temperature indicator for heat-related extreme events (Khaliq et al., 2005; Ouarda and Charron, 2018; Mazdiyasni et al., 2019) due to the accessibility of data and a sufficient recording period for the assessment. Although many studies reported that air temperature is a good temperature indicator for assessing the heat-related risk (Medina-Ramón and Schwartz, 2007; Vaneckova et al., 2011; Allen and Sheridan, 2018), its ability is limited due to complex mechanisms in human thermal comfort (Epstein and Moran, 2006; Malchaire, 2006). Thermal comfort indices are a better alternative to pragmatically assess the heat-related risk based on the heat-related stresses (Collier et al., 2017). For example, Kang et al. (2020) reported that perceived temperature (PT) would be the most appropriate temperature indicator out of air temperature, WBGT, and PT, to assess heat-related risk based on mortality rate in South Korea. In addition, this study showed that the thermal comfort indices consistently represented the heat-related risk for the descriptive and predictive analyses. In this study, the WBGT was employed as a temperature indicator for heat-related extreme events. Since WBGT can account for various factors relating to thermal comfort, the frequency analysis of WBGT extremes leads to a more accurate risk level for heat-related risks on people. Hence, the derived IDF relationship of the current study may be a more appropriate tool to assess and represent the risk levels of the heat-related extreme events than air

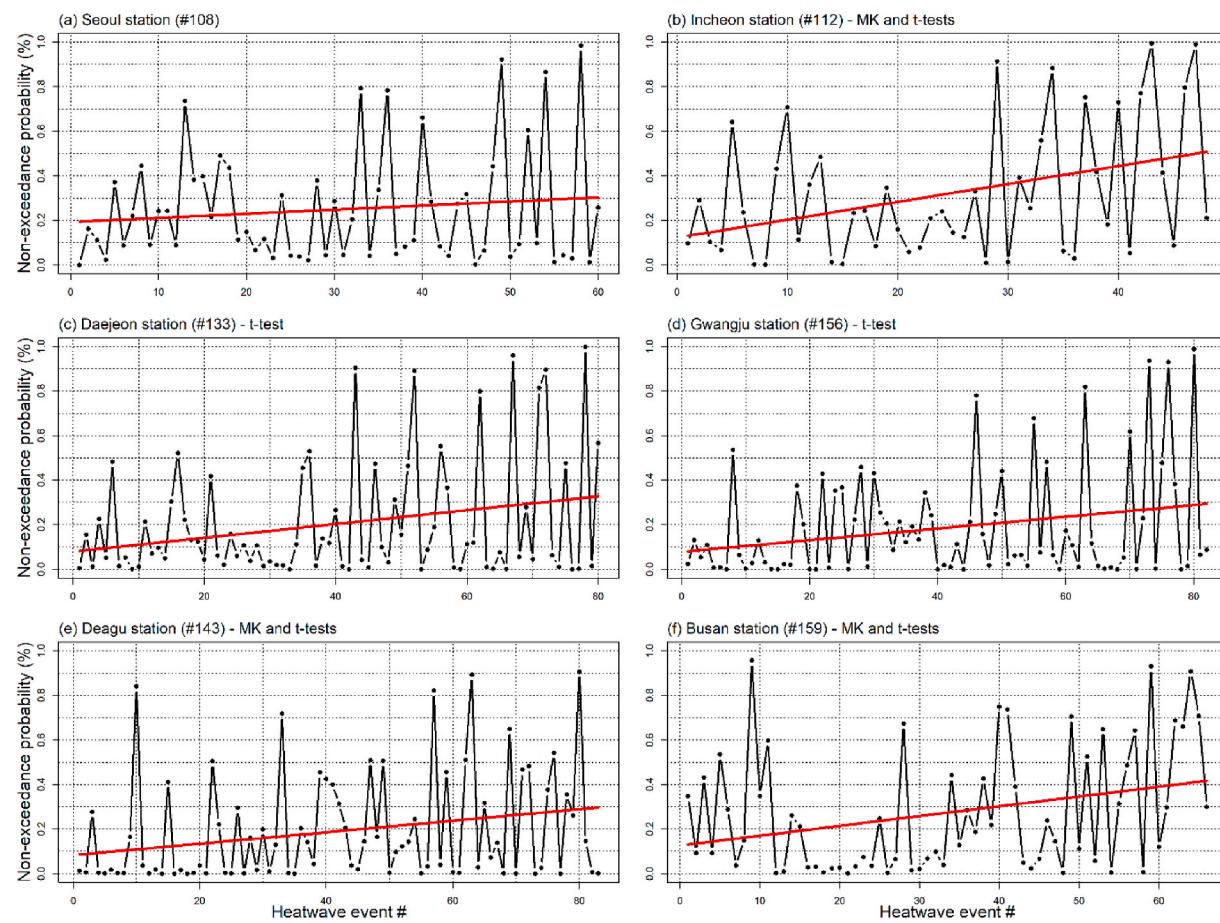


Fig. 9. Non-exceedance probability of heatwave events based on the daily maximum WBGT for six cities in South Korea from 2000 to 2018. Note that the event numbers on the x-axis are ordered as per the time sequence, and the solid red lines indicate a linear trend in the time series of the non-exceedance probability. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

temperature. The WBGT represents the limited capacity of thermal comfort in some conditions, for instance, if the evaporation of sweat is restricted (Budd, 2008). The thermal comfort indices such as PT, physiological equivalent temperature (PET), and universal thermal climate index (UTCI), based on the energy balance between the human body, would solve these limitations (Höppe, 1999; Jendritzky et al., 2012; Staiger et al., 2012). As mentioned above, the use of PT may be a more appropriate choice for assessing heat-related health risk than air temperature and WBGT in South Korea based on Kang et al. (2020). The IDF relationship of the extreme PT would be a useful tool for assessing heat-related health risk in South Korea. In the calculation of the PT, a larger number of meteorological variables than WBGT are required. Thus, it is difficult to obtain the PT data over a long period and also across the nation because only a small number of stations record the meteorological variables used in the calculation of the PT. In the stations where the PT can be calculated, the applicability of the relationship of IDF for the PT extremes should be investigated and its characteristics should be compared to those for the WBGT extremes. However, the RFA is an alternative method to obtain the IDF relationship of the PT extreme as well as the extremes of other physiology-based thermal indices. This can be used when the number of stations is sufficient for the RFA and the PT data within these stations are statistically homogeneous.

While the number of heatwave events was sufficient to perform the trend detection analysis in the current study, the number of recording years was insufficient to investigate the change in the risks of the heatwave events. Hence, the results of this study only provided a snapshot for the given period. To investigate changes in the return levels or the risk of heatwave events over time, the frequency and trend

detection analyses should be carried out with data sets observed over several decades. Note that although the risk level of a heatwave increases statistically as per the results, this does not necessarily equate to an increased impact of the heatwaves on public health, because various factors, such as income, age, region, and acclimation, are also associated with the overall impact (Wu et al., 2014; Armstrong et al., 2017; Chen et al., 2017).

While the return level of a heat-related extreme event undoubtedly depends on its intensity and duration, many other factors are also associated with the risk. The approach in the current study suffers from a limitation with regard to modeling the dependence structure between the intensity and duration of the heat-related extreme events, because this work involves a univariate analysis. To assess the risk of heat-related extreme events while concurrently considering the intensity and duration, multivariate frequency analysis has been employed (Mazdiyasi et al., 2017, 2019; Ouarda and Charron, 2018). However, multivariate analysis requires a larger number of samples than univariate analysis to obtain reliable results. In the current study, RFA was employed due to the shortage of available AMWG data. Thus, the applicability of the multivariate analysis for WBGT data in South Korea should be tested in the future when a sufficient number of sample data is available for the analysis.

When the amount of data is sufficient to reliably carry out classical frequency analysis that uses the data at the target station, the RFA loses its advantage of providing an accurate quantile estimation with a limited amount of data. The RFA calculates the frequency distribution for the data of interest at a target station by pooling information from nearby stations with similar physiographical and meteorological

characteristics. As such, the frequency distribution obtained by the RFA does not represent the true frequency distribution at the target station, although it is very similar. Classical frequency analysis with a sufficient amount of data may provide a more accurate frequency distribution since the pooled data sets in the RFA lead to loss of target station information. Classical frequency analysis outperforms the RFA when the amount of available data at an individual station is greater than 50, based on the simulation study (Hosking and Wallis, 1997), and in this instance, classical frequency analysis is also carried out and compared to the RFA.

RFA is preferable to compensate for the uncertainty of a small sample size. RFA uses a more complicated procedure and a larger number of assumptions than classical frequency analysis. As such, RFA also has uncertainties, particularly in terms of the RFA procedure, such as variable selection for clustering analysis, the clustering algorithm, parameter estimation method, and selection of the best fit distribution. The selection of variables that represent the physiographical and meteorological characteristics of stations in clustering analysis depends on the experiences and knowledge of the users. Thus, the results of clustering analysis will be impacted by the selection of variables. In addition, there is no standard clustering algorithm for identifying statistically homogeneous regions. The results of the clustering analysis will also be altered depending on which clustering algorithm is used. Selections of variables and clustering algorithms will largely impact the stations near the borderline between clusters (or regions in the RFA) because when regions are altered it will change the distribution model. Hence, characteristics of the AMWG at the stations near the borderline will be carefully analyzed and investigated. For the stations located centrally, the selections may lead to small impacts on the quantile estimation.

Hosking and Wallis (1997) suggested the use of the L-moment method (known as probability weighted moment) for the regional parameter estimation in the IF method. Hosking et al. (1985) showed that the L-moment method leads to a superior performance over the moment and maximum likelihood methods in classical frequency analysis when the amount of data is limited. If RFA is used when the sample size is small, the L-moment method would be a rational choice to estimate the regional parameters as it also provides an accurate performance in the classical frequency analysis with the small sample sizes. Alternative parameter estimation methods can be used to estimate regional parameters with the alternative RFA methodology. For example, the population IF method used the maximum likelihood method to estimate regional parameters (Sveinsson et al., 2001). These methods were preferred when a complicated formulation, e.g. a large number of parameters and application of site specific locations and scale parameters, is adopted in the RFA. Since the purpose of applying RFA is to overcome limited AMWG data, the L-moment method may provide the most robust regional parameter estimation.

The most appropriate distributions for the AMWG events of the sub-daily and daily durations may differ for South Korea. As shown in Fig. 3, the distributions of the Z^{DIST} values are different for the sub-daily and daily durations in the same region. For instance, the AMWG events for the sub-daily and daily durations in region #1 may follow the GEV and GNO distributions, respectively. To ensure a consistent analysis, the GNO distribution was employed to analyze the AMWG events for all the durations in the current study. It is possible to use different distribution models for the sub-daily and daily durations. The occurring mechanisms of the extreme events may differ for the AMWG events of the sub-daily and daily durations. The AMWG events for the sub-daily durations may be associated with the diurnal cycles and local extremes, while those for the daily durations may correspond to long-term oscillations or variations. Hence, the distribution models may differ due to different characteristics depending on the durations. Also note that because the distribution model is supposed to be a population distribution in the multivariate analysis, the multivariate frequency analysis may improperly reproduce these different statistical characteristics of AMWG events for short and long durations. Thus, multivariate frequency analysis that

can simultaneously consider these characteristics should be used for risk assessments of extreme WBGT events in South Korea.

5. Conclusions

The IDF relationship of the extreme WBGT in South Korea was derived in the current study. This IDF relationship can provide the risks of heat-related extreme events while concurrently considering their intensities and durations based on the return levels of the AMWG. The IDF relationship can quantify the heat-related risk from extreme events, unlike the conventional threshold approach. In addition, the methodology employed in this study can be a universal tool to assess risk of heat-related extreme events over different durations, ranging from sub-daily to multiple days. Combining the IDF relationship with numerical weather prediction in heat-related extreme events can help the decision maker prepare for upcoming events and mitigate their impacts on public health, agriculture, the ecosystem, and the environment of South Korea. RFA can be applicable in assessing heat-related risk based on frequency analysis of WBGT, particularly when the amount of data is limited. RFA can increase our capacity to assess the risk of heat-related extreme events through various thermal comfort indices such as heat index, PT, PET, and UTCI. Thermal comfort indices such as PT, PET, and UTCI are based on the energy balance between the human body and need a number of variables in their calculations. Therefore, obtaining observations of these indices over a long period, e.g. 30 years, is difficult. RFA is a better alternative to risk assess heat-related extreme events based on physiology-based thermal comfort indices when the number of stations are sufficient.

The dependence structure between duration and intensity was not considered in the frequency analysis. Thus, duration return levels, that are not considered in AMWG calculation, cannot be estimated from the developed IDF relationship. Since the WBGT is a meteorology-based thermal comfort index, the relationship between human health and WBGT needs to be identified by the heat-related risk assessment. Since the relationship was not identified in the current study, the heat-related health risk cannot be assessed by the developed IDF relationship. The physiology-based thermal comfort indices can directly represent the risk of the heat-related events as it considers the energy balance between atmospheric conditions and the human body. Thus, application of the physiology-based thermal comfort in the heat-related risk assessment based on return level is important to mitigate the adverse impacts of the heat-related extreme events on human health. In addition, the risks of heat-related events for specific activities could not be quantified based on estimation of the return level of the event. Thus, the relationship between the return levels and the impacts of heat-related extreme events needs to be investigated to enhance our understanding of the risks from the heat-related impacts on various activities in the future.

Credit author statement

Ju-Young Shin, Conceptualization, Methodology, Software, Data curation, Visualization, Validation, Writing - original draft and Writing-Reviewing and Editing, Investigation. Kyu Rang Kim. Conceptualization, Validation, Writing - original draft, Writing- Reviewing and Editing, Investigation, Supervision. Jong-Chul Ha, Project administration and Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was funded by the Korea Meteorological Administration

Research and Development Program “Advanced Research on Biometeorology” under Grant (KMA2018-00620).

Appendix A. Probability distribution models

The generalized extreme-value (GEV) distribution is composed of three parameters: location (ξ), scale (α), and shape (k). Its probability density function (PDF) is given by following formula:

$$f_{\text{GEV}}(x) = \alpha^{-1} e^{-(1-k)t-e^{-t}}, t = \begin{cases} -\kappa^{-1} \log \left\{ 1 - \frac{k(x-\xi)}{\alpha} \right\}, & k \neq 0 \\ \frac{(x-\xi)}{\alpha}, & k = 0 \end{cases} \quad (\text{A.1})$$

where x is a random variable. The domains of the GEV distribution are $-\infty < x \leq \xi + \alpha/k$ if $k > 0$, $-\infty < x < \infty$ if $k = 0$, and $\xi + \alpha/k \leq x < \infty$ if $k < 0$.

The generalized logistic (GLO) distribution also contains three parameters: location (ξ), scale (α), shape (k). Its PDF form is presented in Equation (A.2).

$$f_{\text{GLO}}(x) = \frac{\alpha^{-1} e^{-(1-k)t}}{(1+e^{-y})^2}, t = \begin{cases} -\kappa^{-1} \log \left\{ 1 - \frac{k(x-\xi)}{\alpha} \right\}, & k \neq 0 \\ \frac{(x-\xi)}{\alpha}, & k = 0 \end{cases} \quad (\text{A.2})$$

where x is a random variable. The domains of the GEV distribution are $-\infty < x \leq \xi + \alpha/k$ if $k > 0$, $-\infty < x < \infty$ if $k = 0$, and $\xi + \alpha/k \leq x < \infty$ if $k < 0$.

The Pearson-type III (PE3) distribution (also known as the three-parameter gamma distribution) contains three parameters: location (μ), scale (σ), and shape (γ). If $\gamma \neq 0$, let $\sigma = 4/\gamma^2$, $\beta = 0.5\sigma|\gamma|$, and $\xi = \mu - 2\sigma/\gamma$. The PDF form of the PE3 distribution is given by the following formula:

$$f_{\text{PE3}}(x) = \frac{(x-\xi)^{\alpha-1} e^{-(x-\xi)/\beta}}{\beta^\alpha \Gamma(\alpha)} \quad (\text{A.3})$$

where x is a random variable. The domains of the PE3 distribution are $\xi \leq x < \infty$ if $\gamma > 0$ and $-\infty \leq x < \infty$ if $\gamma = 0$.

The generalized normal (GNO) distribution (also known as the three-parameter log-normal distribution) contains three parameters: location (ξ), scale (α), and shape (k). Its PDF is given by the following formula:

$$f_{\text{GNO}}(x) = \frac{e^{kt-t^2/2}}{\alpha\sqrt{2\pi}}, t = \begin{cases} -\kappa^{-1} \log \left\{ 1 - \frac{k(x-\xi)}{\alpha} \right\}, & k \neq 0 \\ \frac{(x-\xi)}{\alpha}, & k = 0 \end{cases} \quad (\text{A.4})$$

where x is a random variable. The domains of the GNO distribution are $-\infty < x \leq \xi + \alpha/k$ if $k > 0$, $-\infty < x < \infty$ if $k = 0$, and $\xi + \alpha/k \leq x < \infty$ if $k < 0$.

Appendix B. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.envres.2020.109964>.

References

- Abdi, A., et al., 2017. Regional frequency analysis using Growing Neural Gas network. *J. Hydrol.* 550, 92–102. <https://doi.org/10.1016/j.jhydrol.2017.04.047>.
- Allen, M.J., Sheridan, S.C., 2018. Mortality risks during extreme temperature events (ETEs) using a distributed lag non-linear model. *Int. J. Biometeorol.* 62, 57–67. <https://doi.org/10.1007/s00484-015-1117-4>.
- Armstrong, B., et al., 2017. Longer-term impact of high and low temperature on mortality: an international study to clarify length of mortality displacement. *Environ. Health Perspect.* 125, 107009. <https://doi.org/10.1289/EHP1756>.
- Åström, C., et al., 2015. Developing a heatwave early warning system for Sweden: evaluating sensitivity of different epidemiological modelling approaches to forecast temperatures. *Int. J. Environ. Res. Publ. Health* 12, 254–267. <https://doi.org/10.3390/ijerph120100254>.
- Barnett, A.G., et al., 2010. What measure of temperature is the best predictor of mortality? *Environ. Res.* 110, 604–611. <https://doi.org/10.1016/j.envres.2010.05.006>.
- Basu, B., Srinivas, V.V., 2014. Regional flood frequency analysis using kernel-based fuzzy clustering approach. *Water Resour. Res.* 50, 3295–3316. <https://doi.org/10.1002/2012wr012828>.
- Blazejczyk, K., et al., 2012. Comparison of UTCI to selected thermal indices. *Int. J. Biometeorol.* 56, 515–535. <https://doi.org/10.1007/s00484-011-0453-2>.
- Bobb, J.F., et al., 2011. A bayesian model averaging approach for estimating the relative risk of mortality associated with heat waves in 105 U.S. Cities. *Biometrics* 67, 1605–1616. <https://doi.org/10.1111/j.1541-0420.2011.01583.x>.
- Budd, G.M., 2008. Wet-bulb globe temperature (WBGT)—its history and its limitations. *J. Sci. Med. Sport* 11, 20–32. <https://doi.org/10.1016/j.jssams.2007.07.003>.
- Chebana, F., et al., 2014. Regional frequency analysis at ungauged sites with the generalized additive model. *J. Hydrometeorol.* 15, 2418–2428. <https://doi.org/10.1175/jhm-d-14-0060.1>.
- Chen, K., et al., 2015. Influence of heat wave definitions to the added effect of heat waves on daily mortality in Nanjing, China. *Sci. Total Environ.* 506–507, 18–25. <https://doi.org/10.1016/j.scitotenv.2014.10.092>.
- Chen, T., et al., 2017. Time-series analysis of heat waves and emergency department visits in atlanta, 1993 to 2012. *Environ. Health Perspect.* 125, 057009 <https://doi.org/10.1289/EHP44>.
- Chien, L.-C., et al., 2016. Spatiotemporal analysis of heat and heat wave effects on elderly mortality in Texas, 2006–2011. *Sci. Total Environ.* 562, 845–851. <https://doi.org/10.1016/j.scitotenv.2016.04.042>.
- Collier, R.J., et al., 2017. A 100-Year Review: stress physiology including heat stress. *J. Dairy Sci.* 100, 10367–10380. <https://doi.org/10.3168/jds.2017-13676>.

- Drissia, T.K., et al., 2019. Flood frequency analysis using L moments: a comparison between at-site and regional approach. *Water Resour. Manag.* 33, 1013–1037. <https://doi.org/10.1007/s11269-018-2162-7>.
- Ebi, K.L., Schmierer, J.K., 2005. A stitch in time: improving public health early warning systems for extreme weather events. *Epidemiol. Rev.* 27, 115–121. <https://doi.org/10.1093/epirev/mxi006>.
- Ebi, K.L., et al., 2004. Heat watch/warning systems save lives: estimated costs and benefits for philadelphia 1995–98. *Bull. Am. Meteorol. Soc.* 85, 1067–1074. <https://doi.org/10.1175/bams-85-8-1067>.
- Epstein, Y., Moran, D.S., 2006. Thermal comfort and the heat stress indices. *Ind. Health* 44, 388–398. <https://doi.org/10.2486/indhealth.44.388>.
- Gascon, M., et al., 2016. Residential green spaces and mortality: a systematic review. *Environ. Int.* 86, 60–67. <https://doi.org/10.1016/j.envint.2015.10.013>.
- Green, H.K., et al., 2016. Mortality during the 2013 heatwave in England – how did it compare to previous heatwaves? A retrospective observational study. *Environ. Res.* 147, 343–349. <https://doi.org/10.1016/j.envres.2016.02.028>.
- Grundstein, A., et al., 2014. The geography of extreme heat hazards for American football players. *Appl. Geor.* 46, 53–60. <https://doi.org/10.1016/j.apgeog.2013.10.007>.
- Grundstein, A., et al., 2015. Regional heat safety thresholds for athletics in the contiguous United States. *Appl. Geor.* 56, 55–60. <https://doi.org/10.1016/j.apgeog.2014.10.014>.
- Guo, Y., et al., 2012. High temperatures-related elderly mortality varied greatly from year to year: important information for heat-warning systems. *Sci. Rep.* 2, 830. <https://doi.org/10.1038/srep00830>.
- Höppe, P., 1999. The physiological equivalent temperature – a universal index for the biometeorological assessment of the thermal environment. *Int. J. Biometeorol.* 43, 71–75. <https://doi.org/10.1007/s004840050118>.
- Heo, S., et al., 2019. Comparison of health risks by heat wave definition: applicability of wet-bulb globe temperature for heat wave criteria. *Environ. Res.* 168, 158–170. <https://doi.org/10.1016/j.envres.2018.09.032>.
- Hosking, J.R.M., Wallis, J.R., 1997. *Regional Frequency Analysis: an Approach Based on L-Moments*. Cambridge University Press, Cambridge.
- Hosking, J.R.M., et al., 1985. Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. *Technometrics* 27, 251–261. <https://doi.org/10.1080/00401706.1985.10488049>.
- Hu, C., et al., 2019. A modified regional L-moment method for regional extreme precipitation frequency analysis in the Songliao River Basin of China. *Atmos. Res.* 230, 104629. <https://doi.org/10.1016/j.atmosres.2019.104629>.
- Igor, L., et al., 2019. Regional flood frequency analysis based on L-moment approach (case study tisza river basin). *Water Resour.* 46, 853–860. <https://doi.org/10.1134/S009780781906006X>.
- Jänicke, B., et al., 2020. A simple high-resolution heat-stress forecast for Seoul, Korea: coupling climate information with an operational numerical weather prediction model. *Int. J. Biometeorol.* <https://doi.org/10.1007/s00484-020-01893-1>.
- Javelle, P., et al., 2002. Development of regional flood-duration-frequency curves based on the index-flood method. *J. Hydrol.* 258, 249–259. [https://doi.org/10.1016/S0022-1694\(01\)00577-7](https://doi.org/10.1016/S0022-1694(01)00577-7).
- Jendritzky, G., et al., 2012. UTCl—why another thermal index? *Int. J. Biometeorol.* 56, 421–428. <https://doi.org/10.1007/s00484-011-0513-7>.
- Kamppmann, B., et al., 2012. Physiological responses to temperature and humidity compared to the assessment by UTCl, WGBT and PHS. *Int. J. Biometeorol.* 56, 505–513. <https://doi.org/10.1007/s00484-011-0410-0>.
- Kang, M., et al., 2020. Event-based heat-related risk assessment model for South Korea using maximum perceived temperature, wet-bulb globe temperature, and air temperature data. *Int. J. Environ. Res. Publ. Health* 17, 2631.
- Katz, R.W., et al., 2002. Statistics of extremes in hydrology. *Adv. Water Resour.* 25, 1287–1304. [https://doi.org/10.1016/S0309-1708\(02\)00056-8](https://doi.org/10.1016/S0309-1708(02)00056-8).
- Kent, S.T., et al., 2014. Heat waves and health outcomes in Alabama (USA): the importance of heat wave definition. *Environ. Health Perspect.* 122, 151–158. <https://doi.org/10.1289/ehp.1307262>.
- Khalil, M.N., et al., 2005. Frequency analysis and temporal pattern of occurrences of southern Quebec heatwaves. *Int. J. Climatol.* 25, 485–504. <https://doi.org/10.1002/joc.1141>.
- Kim, E.J., Kim, H., 2017. Effect modification of individual- and regional-scale characteristics on heat wave-related mortality rates between 2009 and 2012 in Seoul, South Korea. *Sci. Total Environ.* 595, 141–148. <https://doi.org/10.1016/j.scitotenv.2017.03.248>.
- Kim, Y.-H., et al., 2018. Multi-model event attribution of the summer 2013 heat wave in Korea. *Weather Clim. Extremes.* 20, 33–44. <https://doi.org/10.1016/j.wace.2018.03.004>.
- Koutsoyiannis, D., et al., 1998. A mathematical framework for studying rainfall intensity-duration-frequency relationships. *J. Hydrol.* 206, 118–135. [https://doi.org/10.1016/S0022-1694\(98\)00097-3](https://doi.org/10.1016/S0022-1694(98)00097-3).
- Kulkarni, A., 2017. Homogeneous clusters over India using probability density function of daily rainfall. *Theor. Appl. Climatol.* 129, 633–643. <https://doi.org/10.1007/s00704-016-1808-8>.
- Kysely, J., et al., 2011. Comparison of regional and at-site approaches to modelling probabilities of heavy precipitation. *Int. J. Climatol.* 31, 1457–1472. <https://doi.org/10.1002/joc.2182>.
- Lee, D.-G., et al., 2018. Effects of heat waves on daily excess mortality in 14 Korean cities during the past 20 years (1991–2010): an application of the spatial synoptic classification approach. *Int. J. Biometeorol.* 62, 575–583. <https://doi.org/10.1007/s00484-017-1466-2>.
- Lee, J.-S., et al., 2019. Evaluating the accuracies of the WBGT estimation models and their onsite applicability in Korea. *J. Korea Soc. Hazard Mitig.* 19, 53–63. <https://doi.org/10.9798/KOSHAM.2019.19.4.53>.
- Lee, S.-M., Min, S.-K., 2018. Heat stress changes over east asia under 1.5° and 2.0°C global warming targets. *J. Clim.* 31, 2819–2831. <https://doi.org/10.1175/jcli-d-17-0449.1>.
- Lee, W.K., et al., 2016. Added effect of heat wave on mortality in Seoul, Korea. *Int. J. Biometeorol.* 60, 719–726. <https://doi.org/10.1007/s00484-015-1067-x>.
- Lemke, B., Kjellstrom, T., 2012. Calculating workplace WBGT from meteorological data: a tool for climate change assessment. *Ind. Health* 50, 267–278. <https://doi.org/10.2486/indhealth.MS1352>.
- Liang, Y., et al., 2017. L-Moment-Based regional frequency analysis of annual extreme precipitation and its uncertainty analysis. *Water Resour. Manag.* 31, 3899–3919. <https://doi.org/10.1007/s11269-017-1715-5>.
- Liljegeen, J.C., et al., 2008. Modeling the wet bulb globe temperature using standard meteorological measurements. *J. Occup. Environ. Hyg.* 5, 645–655. <https://doi.org/10.1080/15459620802310770>.
- Lim, Y.-H., et al., 2019. Estimation of heat-related deaths during heat wave episodes in South Korea (2006–2017). *Int. J. Biometeorol.* 63, 1621–1629. <https://doi.org/10.1007/s00484-019-01774-2>.
- Lin, G.-F., Chen, L.-H., 2006. Identification of homogeneous regions for regional frequency analysis using the self-organizing map. *J. Hydrol.* 324, 1–9. <https://doi.org/10.1016/j.jhydrol.2005.09.009>.
- Lowe, D., et al., 2011. Heatwave early warning systems and adaptation advice to reduce human health consequences of heatwaves. *Int. J. Environ. Res. Publ. Health* 8, 4623–4648. <https://doi.org/10.3390/ijerph8124623>.
- MacQueen, J., 1967. Some methods for classification and analysis of multivariate observations. In: *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1. Statistics. University of California Press, Berkeley, Calif, pp. 281–297.
- Malchaire, J.B.M., 2006. Occupational heat stress assessment by the predicted heat strain model. *Ind. Health* 44, 380–387. <https://doi.org/10.2486/indhealth.44.380>.
- Matzarakis, A., et al., 2020. The heat health warning system in Germany—application and warnings for 2005 to 2019. *Atmosphere* 11, 170. <https://doi.org/10.3390/atmos11020170>.
- Mazdiyasi, O., et al., 2017. Increasing probability of mortality during Indian heat waves. *Sci. Adv.* 3, e1700066. <https://doi.org/10.1126/sciadv.1700066>.
- Mazdiyasi, O., et al., 2019. Heat wave intensity duration frequency curve: a multivariate approach for hazard and attribution analysis. *Sci. Rep.* 9, 14117. <https://doi.org/10.1038/s41598-019-50643-w>.
- Medina-Ramón, M., Schwartz, J., 2007. Temperature, temperature extremes, and mortality: a study of acclimatization and effect modification in 50 US cities. *Occup. Environ. Med.* 64, 827–833. <https://doi.org/10.1136/oem.2007.033175>.
- Meehl, G.A., Tebaldi, C., 2004. More intense, more frequent, and longer lasting heat waves in the 21st century. *Science* 305, 994–997. <https://doi.org/10.1126/science.1098704>.
- Nagpal, A., et al., 2013. Review based on data clustering algorithms. In: *2013 IEEE Conference on Information & Communication Technologies*, pp. 298–303.
- Nairn, J.R., Fawcett, R.J.B., 2015. The excess heat factor: a metric for heatwave intensity and its use in classifying heatwave severity. *Int. J. Environ. Res. Publ. Health* 12, 227–253.
- Nam, W., et al., 2015. Delineation of the climatic rainfall regions of South Korea based on a multivariate analysis and regional rainfall frequency analyses. *Int. J. Climatol.* 35, 777–793. <https://doi.org/10.1002/joc.4182>.
- Newth, D., Gunasekera, D., 2018. Projected changes in wet-bulb globe temperature under alternative climate scenarios. *Atmosphere* 9, 187. <https://doi.org/10.3390/atmos9050187>.
- Ngongondo, C.S., et al., 2011. Regional frequency analysis of rainfall extremes in Southern Malawi using the index rainfall and L-moments approaches. *Stoch. Environ. Res. Risk Assess.* 25, 939–955. <https://doi.org/10.1007/s00477-011-0480-x>.
- Nogaj, M., et al., 2006. Amplitude and frequency of temperature extremes over the North Atlantic region. *Geophys. Res. Lett.* 33. <https://doi.org/10.1029/2005gl024251>.
- Ouarda, T.B.M.J., Charron, C., 2018. Nonstationary temperature-duration-frequency curves. *Sci. Rep.* 8, 15493. <https://doi.org/10.1038/s41598-018-33974-y>.
- Ouarda, T.B.M.J., et al., 2019. Non-stationary intensity-duration-frequency curves integrating information concerning teleconnections and climate change. *Int. J. Climatol.* 39, 2306–2323. <https://doi.org/10.1002/joc.5953>.
- Parsons, K., 2006. Heat stress standard ISO 7243 and its global application. *Ind. Health* 44, 368–379. <https://doi.org/10.2486/indhealth.44.368>.
- Parsons, K.C., 1999. International standards for the assessment of the risk of thermal strain on clothed workers in hot environments. *Ann. Occup. Hyg.* 43, 297–308. <https://doi.org/10.1093/annhyg/43.5.297>.
- Peel, M.C., et al., 2007. Updated world map of the Köppen-Geiger climate classification. *Hydroclim. Earth Syst. Sci.* 11, 1633–1644. <https://doi.org/10.5194/hess-11-1633-2007>.
- Perkins, S.E., et al., 2012. Increasing frequency, intensity and duration of observed global heatwaves and warm spells. *Geophys. Res. Lett.* 39. <https://doi.org/10.1029/2012gl053361>.
- Raei, E., et al., 2018. GHWR, a multi-method global heatwave and warm-spell record and toolbox. *Sci. Data.* 5, 180206. <https://doi.org/10.1038/sdata.2018.206>.
- Raggad, B., 2018. Stationary and non-stationary extreme value approaches for modelling extreme temperature: the case of riyadh city, Saudi arabia. *Environ. Model. Assess.* 23, 99–116. <https://doi.org/10.1007/s10666-017-9588-9>.

- Rahman, A.S., et al., 2013a. A study on selection of probability distributions for at-site flood frequency analysis in Australia. *Nat. Hazards* 69, 1803–1813. <https://doi.org/10.1007/s11069-013-0775-y>.
- Rahman, M.M., et al., 2013b. Regional extreme rainfall mapping for Bangladesh using L-moment technique. *J. Hydrol. Eng.* 18, 603–615. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0000663](https://doi.org/10.1061/(ASCE)HE.1943-5584.0000663).
- Raynal-Vellaseñor, J.A., 2012. Estimation procedures for the general extreme value distribution for the maxima: alternative PWM method. *J. Hydrol. Eng.* 17, 909–922. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0000525](https://doi.org/10.1061/(ASCE)HE.1943-5584.0000525).
- Robinson, P.J., 2001. On the definition of a heat wave. *J. Appl. Meteorol.* 40, 762–775. [https://doi.org/10.1175/1520-0450\(2001\)040<0762:otdoh>2.0.co;2](https://doi.org/10.1175/1520-0450(2001)040<0762:otdoh>2.0.co;2).
- Rousseeuw, P.J., 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* 20, 53–65. [https://doi.org/10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7).
- Royé, D., et al., 2020. Heat wave intensity and daily mortality in four of the largest cities of Spain. *Environ. Res.* 182, 109027. <https://doi.org/10.1016/j.envres.2019.109027>.
- Satyanarayana, P., Srinivas, V.V., 2008. Regional frequency analysis of precipitation using large-scale atmospheric variables. *J. Geophys. Res. Atmos.* 113 <https://doi.org/10.1029/2008jd010412>.
- Scalley, B.D., et al., 2015. Responding to heatwave intensity: excess Heat Factor is a superior predictor of health service utilisation and a trigger for heatwave plans. *Aust. N. Z. J. Publ. Health* 39, 582–587. <https://doi.org/10.1111/1753-6405.12421>.
- Sherwood, S., Fu, Q., 2014. A drier future? *Science* 343, 737–739. <https://doi.org/10.1126/science.1247620>.
- Shin, J.-Y., et al., 2019. Spatial and temporal variations in rainfall erosivity and erosivity density in South Korea. *Catena* 176, 125–144. <https://doi.org/10.1016/j.catena.2019.01.005>.
- Shin, J.-Y., et al., 2015. Heterogeneous mixture distributions for modeling multisource extreme rainfalls. *J. Hydrometeorol.* 16, 2639–2657. <https://doi.org/10.1175/jhm-d-14-0130.1>.
- Smith, T.T., et al., 2013. Heat waves in the United States: definitions, patterns and trends. *Climatic Change* 118, 811–825. <https://doi.org/10.1007/s10584-012-0659-2>.
- Son, J.-Y., et al., 2016. Urban vegetation and heat-related mortality in Seoul, Korea. *Environ. Res.* 151, 728–733. <https://doi.org/10.1016/j.envres.2016.09.001>.
- Staiger, H., et al., 2012. The perceived temperature – a versatile index for the assessment of the human thermal environment. Part A: scientific basics. *Int. J. Biometeorol.* 56, 165–176. <https://doi.org/10.1007/s00484-011-0409-6>.
- Stull, R., 2011. Wet-bulb temperature from relative humidity and air temperature. *J. Appl. Meteorol. Climatol.* 50, 2267–2269. <https://doi.org/10.1175/jamc-d-11-0143.1>.
- Sveinsson, O.G.B., et al., 2001. Population index flood method for regional frequency analysis. *Water Resour. Res.* 37, 2733–2748. <https://doi.org/10.1029/2001wr000321>.
- Sveinsson, O.G.B., et al., 2002. Regional frequency analysis of extreme precipitation in northeastern Colorado and fort collins flood of 1997. *J. Hydrol. Eng.* 7, 49–63. [https://doi.org/10.1061/\(ASCE\)1084-0699\(2002\)7:1\(49\)](https://doi.org/10.1061/(ASCE)1084-0699(2002)7:1(49)).
- Tolo, G., et al., 2013. Are heat warning systems effective? *Environ. Health* 12, 27. <https://doi.org/10.1186/1476-069X-12-27>.
- Trenberth, K.E., Josey, S.A., 2007. Observations: surface and atmospheric climate change. In: Solomon, S., et al. (Eds.), *Climate Change 2007: The Physical Science Basis: Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, UK, pp. 235–336.
- Vaneckova, P., et al., 2011. Do biometeorological indices improve modeling outcomes of heat-related mortality? *J. Appl. Meteorol. Climatol.* 50, 1165–1176. <https://doi.org/10.1175/2011jamc2632.1>.
- Willett, K.M., Sherwood, S., 2012. Exceedance of heat index thresholds for 15 regions under a warming climate using the wet-bulb globe temperature. *Int. J. Climatol.* 32, 161–177. <https://doi.org/10.1002/joc.2257>.
- Wu, J., et al., 2014. Estimation and uncertainty analysis of impacts of future heat waves on mortality in the eastern United States. *Environ. Health Perspect.* 122, 10–16. <https://doi.org/10.1289/ehp.1306670>.
- Xu, Z., et al., 2018. Heatwave and health events: a systematic evaluation of different temperature indicators, heatwave intensities and durations. *Sci. Total Environ.* 630, 679–689. <https://doi.org/10.1016/j.scitotenv.2018.02.268>.
- Xu, Z., Tong, S., 2017. Decompose the association between heatwave and mortality: which type of heatwave is more detrimental? *Environ. Res.* 156, 770–774. <https://doi.org/10.1016/j.envres.2017.05.005>.
- Yagliou, C.P., Minard, D., 1957. Control of heat casualties at military training centers. *AMA Arch Ind Health* 16, 302–316.
- Yang, T., et al., 2010. Regional frequency analysis and spatio-temporal pattern characterization of rainfall extremes in the Pearl River Basin, China. *J. Hydrol.* 380, 386–405. <https://doi.org/10.1016/j.jhydrol.2009.11.013>.
- Yun, K.-S., et al., 2012. Changes in climate classification and extreme climate indices from a high-resolution future projection in Korea. *Asia-Pacific Journal of Atmospheric Sciences* 48, 213–226. <https://doi.org/10.1007/s13143-012-0022-6>.
- Zahid, M., et al., 2017. Return levels of temperature extremes in southern Pakistan. *Earth Syst. Dynam.* 8, 1263–1278. <https://doi.org/10.5194/esd-8-1263-2017>.
- Zhang, K., et al., 2014. Using forecast and observed weather data to assess performance of forecast products in identifying heat waves and estimating heat wave effects on mortality. *Environ. Health Perspect.* 122, 912–918. <https://doi.org/10.1289/ehp.1306858>.
- Zhang, Y., et al., 2013. Risk factors for direct heat-related hospitalization during the 2009 Adelaide heatwave: a case crossover study. *Sci. Total Environ.* 442, 1–5. <https://doi.org/10.1016/j.scitotenv.2012.10.042>.