

## K-평균 군집분석을 이용한 동아시아 지역 날씨유형 분류

조영준\* · 이현철 · 임병환 · 김승범

기상청 국립기상과학원 관측예보연구과

(접수일: 2019년 7월 2일, 수정일: 2019년 9월 5일, 게재확정일: 2019년 9월 26일)

### Classification of Weather Patterns in the East Asia Region using the K-means Clustering Analysis

Young-Jun Cho\*, Hyeon-Cheol Lee, Byunghwan Lim, and Seung-Bum Kim

*Observation and Forecast Research Division, National Institute of Meteorological Sciences, Seogwipo, Korea*

(Manuscript received 2 July 2019; revised 5 September 2019; accepted 26 September 2019)

**Abstract** Medium-range forecast is highly dependent on ensemble forecast data. However, operational weather forecasters have not enough time to digest all of detailed features revealed in ensemble forecast data. To utilize the ensemble data effectively in medium-range forecasting, representative weather patterns in East Asia in this study are defined. The k-means clustering analysis is applied for the objectivity of weather patterns. Input data used daily Mean Sea Level Pressure (MSLP) anomaly of the ECMWF ReAnalysis-Interim (ERA-Interim) during 1981~2010 (30 years) provided by the European Centre for Medium-Range Weather Forecasts (ECMWF). Using the Explained Variance (EV), the optimal study area is defined by 20~60°N, 100~150°E. The number of clusters defined by Explained Cluster Variance (ECV) is thirty ( $k = 30$ ). 30 representative weather patterns with their frequencies are summarized. Weather pattern #1 occurred all seasons, but it was about 56% in summer (June~September). The relatively rare occurrence of weather pattern (#30) occurred mainly in winter. Additionally, we investigate the relationship between weather patterns and extreme weather events such as heat wave, cold wave, and heavy rainfall as well as snowfall. The weather patterns associated with heavy rainfall exceeding  $110 \text{ mm day}^{-1}$  were #1, #4, and #9 with days (%) of more than 10%. Heavy snowfall events exceeding  $24 \text{ cm day}^{-1}$  mainly occurred in weather pattern #28 (4%) and #29 (6%). High and low temperature events ( $> 34^\circ\text{C}$  and  $< -14^\circ\text{C}$ ) were associated with weather pattern #1~4 (14~18%) and #28~29 (27~29%), respectively. These results suggest that the classification of various weather patterns will be used as a reference for grouping all ensemble forecast data, which will be useful for the scenario-based medium-range ensemble forecast in the future.

**Key words:** Weather pattern, East Asia region, k-means clustering analysis, medium-range forecast, extreme weather

\*Corresponding Author: Young-Jun Cho, Observation and Forecast Research Division, National Institute of Meteorological Sciences, 33, Seohobuk-ro, Seogwipo-si 63568, Korea.  
Phone: +82-64-780-6613, Fax: +82-64-738-6513  
E-mail: genesis19y@korea.kr

## 1. 서 론

기상예보는 예보대상기간에 따라 크게 초단기, 단기, 중기, 그리고 장기예보로 구분된다. 여기서 초단기, 단기, 그리고 중기예보는 각각 6시간, 3일, 그리고 10일 이내의 예보대상기간으로 정의되고, 장기예보는 10일 이상의 기간으로 정의된다. 기상예보는 현재의 상태를 이해할 수 있는 기상관측으로부터 시작된다. 기상관측에서 나타난 특징을 분석하여 현재와 미래의 상태를 진단하고 예상하며, 또한 수치모델 입력자료로 활용되어 미래를 예측할 수 있기 때문이다. 대한민국 기상청은 지상 및 고층기상관측, 기상레이더, 기상위성, 이동형 기상관측차량, 기상항공기, 해양 기상관측선(기상1호), 그리고 수치예보모델 등을 운영하여 예보대상기간, 기상현상, 그리고 현상의 시공간적 규모에 따라 적절하게 활용하고 있다.

건설, 농작, 항공, 항만 물류 등 사회 여러 분야에서 예보 정확도 향상과 예보기간 확장 등 기상정보에

대한 사회적 요구사항이 증가하고 있다. 현재 중기예보 시 양상블(ensemble) 예측자료의 의존도가 높은 실정이다. 기상청은 수치모델의 초기상태를 달리하여 다양한 양상을 자료를 생산하여 예보판단 자료로 활용하고 있다. 그러나 예보관이 중기예보를 생산할 때, 전체 양상을 결과를 상세하게 분석하기에는 시간적으로 한계가 존재한다. 따라서 양상을 예측자료의 활용성을 높이기 위해 간소화된 양상을 예측자료의 제공이 요구된다. 따라서 유사 양상을 자료를 그룹화하여 간소화된 시나리오 기반의 예보판단 자료 산출이 필요하다.

날씨유형 연구는 폭염, 호우, 가뭄, 대기질 등 다양한 기상현상에서 연구가 수행되어 왔다(Park and Lee, 1998; Goodess and Jones, 2002; Esteban et al., 2005; Hoinka et al., 2006; Vicente-Serrano and López-Moreno, 2006; Casola and Wallace, 2007; Kysely, 2007; Huth et al., 2008; Fleig et al., 2010; Mukougawa and Mabuchi, 2012; Hisashi et al., 2013; Hsu and Cheng,

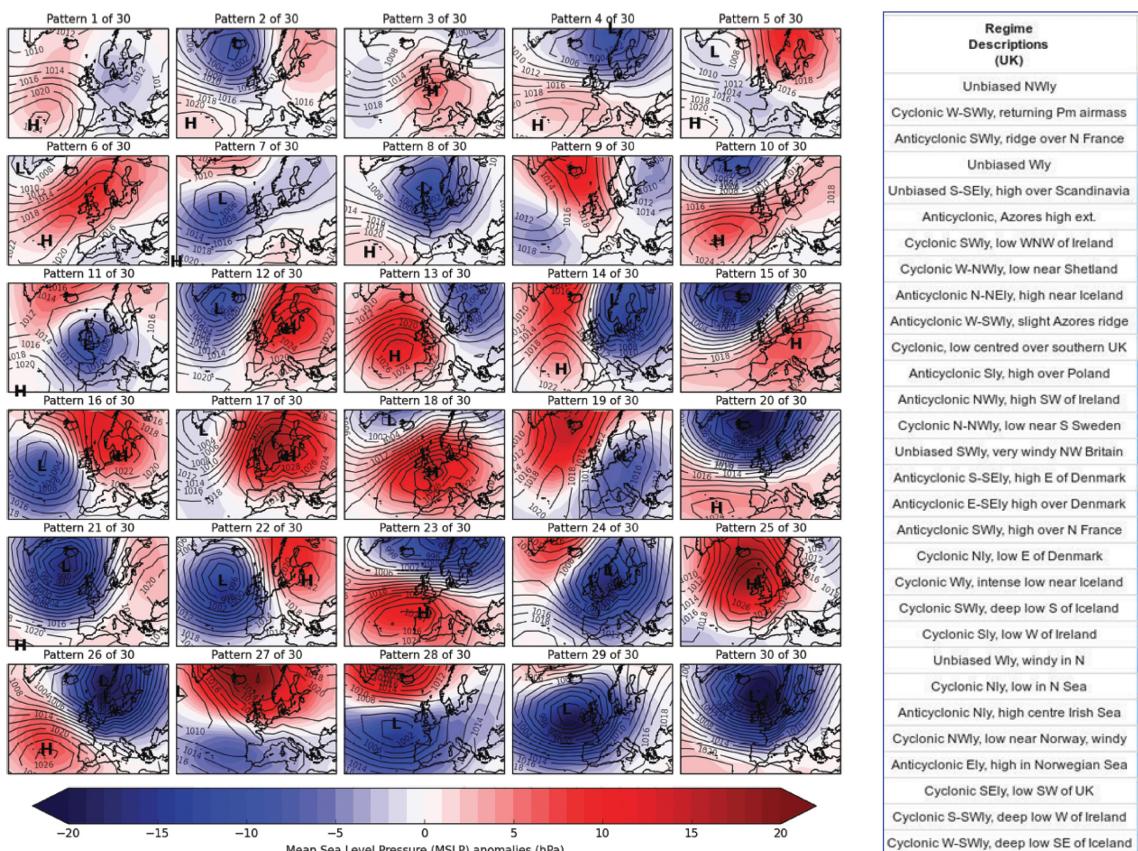


Fig. 1. Definitions of 30 weather patterns. Shaded and contour indicate the Mean Sea Level Pressure (MSLP) anomalies and MSLP mean values (2 hPa intervals), respectively (Neal et al., 2016).

2016; Zhang et al., 2016). Yoon et al. (2018)은 종관 기상관측장비(Automated Synoptic Observing System, ASOS)의 지상기온과 K-평균 군집분석(K-means clustering analysis)을 이용하여 한반도 폭염과 관련된 세가지 날씨유형을 분류하였고, 유형별로 폭염의 공간적 분포를 조사하였다. 이 연구결과는 날씨유형을 통해 폭염 지속시간, 영향 범위 등 예보에 직접적으로 활용될 수 있다. Lim and Seo (2018)는 인공신경망(artificial neural network)기법인 자기조직화지도(Self-Organizing Map, SOM; Kohonen, 1982) 기법을 이용하여 한반도 여름철 극한 온도와 관련된 북서태평양과 북대서양의 해수면온도 아노말리 두 가지 중요 원격상관패턴을 분류하였다.

Ohba et al. (2015)은  $150 \text{ mm day}^{-1}$  이상 강수와 관련된 날씨유형을 분류하였고, 지역별로 상위 50위에 해당되는 집중호우 날씨유형 발생빈도를 조사하였다. Nguyen-Le et al. (2017)은 큐슈 북부지역의 집중호우 예보 정확도 향상을 위해 장마기간 동안  $100 \text{ mm day}^{-1}$  이상의 집중호우 사례에 대해 날씨유형 연구를 수행하였다. 대기질 날씨유형 연구로 Zhang et al. (2016)은 중부 북부지역의 계절별 대기질 변화를 분석하였고, Hsu and Cheng (2016)은 대만 윈린(Yunlin)의 Particle Matter 2.5 (PM 2.5) 농도에 영향을 미치는 기상조건을 조사하기 위해 바람장 자료를 이용해 군집분석을 수행하였다.

영국 기상청(Met Office)에서는 북대서양과 유럽 지역을 포함하는  $70^{\circ}\text{W} \sim 50^{\circ}\text{E}$ ,  $25^{\circ}\text{N} \sim 70^{\circ}\text{N}$  영역(North Atlantic and European, NAE)에 대해 30개 날씨유형을 분류하여 중기예보 도구(medium-range forecasting tool) 및 수치모델 검증자료로 활용하고 있다(Fig. 1; Neal et al., 2016). 좌측 상단이 1번 유형, 우측 하단이 30번 유형을 나타내고, 유형 번호가 증가할수록 겨울철에 주로 발생하였다. 또한 유형에 따라 기압의 배치와 기류의 특징을 요약하였다. 30개 날씨유형 선정 시 객관성을 위해 K-평균 군집분석을 사용하였고, 유형의 일관성을 확보하기 위해 1850~2003년(154년) 기간 동안의 장기간 European and North Atlantic daily to MULti-decadal climATE variability (EMULATE) MSLP (EMSLP) 자료를 활용하여 연구가 수행되었다.

이 연구결과를 기초로 현상에 대한 분석과 예측 정확도 향상을 위해 영국기상청에서는 호우 및 가뭄, 산사태, 영국 해안가의 홍수 예보, 해양기상 등 다양한 예보 활용연구를 시도하고 있다(Neal et al., 2018; Richardson et al., 2018; Robbins et al., 2018; Steele et al., 2018). Neal et al. (2018)은 영국 해안가 홍수와 관련된 특정 날씨유형에 대한 정보를 활용하면 상대적으로 정확한 중기, 장기예보가 가능함을 보였다. Richardson et al. (2018)은 객관적으로 분류된 날씨유

형을 이용하여 유럽 강수량 및 가뭄 경향 분석을 수행하였다. 이 연구결과는 예보 기초 정보로 활용하고 있고, 날씨유형 정보가 대규모 대기순환 분석에 유용함을 확인하였다. Robbins et al. (2018)은 7~30일 사이에 발생할 스코틀랜드 산사태 조기 경보를 위해 특정 날씨유형에서 발생하는 산사태를 분석하였다. 또한 파고, 해류 변화 등 해상 날씨정보를 제공하기 위한 연구가 날씨유형을 기초로 수행되었다(Steele et al., 2018).

한반도에서는 대설, 강풍, 가뭄, 폭염, 그리고 집중호우 등 특정 위험기상이 계절에 따라 뚜렷하게 발생한다. 따라서 장기간의 기상자료를 이용하여 계절별로 한반도 날씨를 결정하는 기압배치 등 동아시아 지역의 날씨유형을 이해하는 것이 중요하다. 그러나 과거 날씨유형 연구는 주로 위험기상과 관련된 사례 중심으로 수행되었고, 객관적 분석기법을 활용한 날씨유형 분류 연구는 아직까지 미흡한 실정이다. 본 연구에서는 장기간 재분석 자료와 객관적 분석기법(K-means clustering)을 활용하여 동아시아 지역의 대표 날씨유형을 분류하는데 목적이 있다. 이러한 결과는 기상청 중기예보 시 양상을 자료의 효율적인 활용을 위해 시나리오 기반의 양상을 예측자료 생산을 위한 자료로 활용될 것이다.

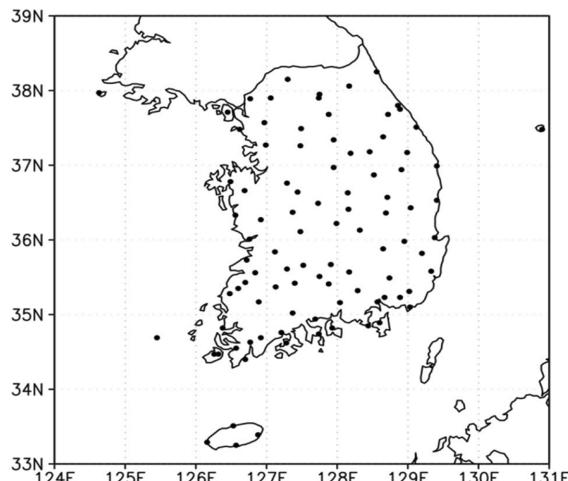
## 2. 자료 및 방법

중기예보 시 양상을 예측결과의 분류기준이 되는 동아시아 지역 대표 날씨유형 분류를 위해 유럽중기 예보센터(European Centre for Medium-Range Weather Forecasts, ECMWF)에서 제공하는 ERA-Interim (ECMWF ReAnalysis-Interim; Dee et al., 2011) 평균해면기압 (Mean Sea Level Pressure, MSLP) 자료를 이용하였다. 기간은 1981~2010년(30년)이고, 시공간적 분해능은 각각 일 간격 그리고  $1.5^{\circ}$  격자해상도이다(Table 1). 군집분석에 사용된 입력자료는 해면기압 값의 계절적 차이를 제거하기 위해 해면기압 아노말리(해당일-해당일 30년 평균)를 사용하였다. 또한 아노말리 계산 시 평년치는 일 변동성을 줄이기 위해 1-2-1 이항필터링(binomial filtering) 기법을 적용하였다(Fereday et al., 2008).

연구방법은 객관적 군집분석 방법 중 K-평균 군집분석(K-means clustering analysis)을 활용하였다. K-평균 군집분석은 비계층적(non-hierarchical) 군집분석 기법 중 하나로 계산 속도가 빠르고 대용량 자료를 활용하여 군집을 분류하는데 효과적인 것으로 알려져 있다. K-평균 군집분석은 k개의 중심 값과 각 개별 데이터간의 거리를 계산하여 중심에서 가장 가까운 k 개의 군집으로 묶는 알고리즘이다(MacQueen, 1967;

**Table 1.** The information of dataset used in this study.

Dataset	Period	Variables	Spatial resolution
ERA-Interim		Daily Mean Sea Level Pressure (MSLP)	1.5 degree
ASOS	1981~2010	Daily precipitation Maximum fresh snow depth Daily maximum and minimum temperature	-

**Fig. 2.** Locations of 95 Automated Synoptic Observing Systems (ASOSs) over South Korea.

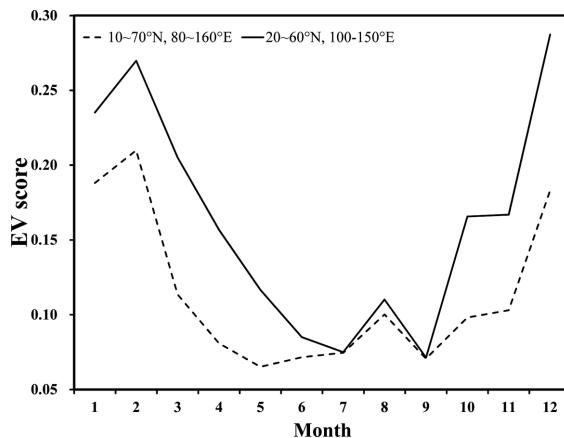
Everitt, 1993; Lee and Park, 1999; Choi et al, 2009).

위험기상과 관련된 날씨유형을 분석하기 위해 1981~2010년(30년) 동안 95개 지점의 ASOS (Automated Synoptic Observing System) 자료를 이용하였다(Fig. 2). 본 연구에서는 극한 강수 그리고 기온과 관련된 위험기상 사례를 선정하기 위해 ASOS의 일 강수량, 최심 신적설, 일 최고기온, 그리고 최저기온 기상변수를 사용하였다(Table 1). 위험기상 선정기준은 극값 1%에 해당하는 사례로 선정되었다. 극값 1%는 일 강수량, 최심 신적설, 일 최고기온, 그리고 최저기온이 각각  $> 110 \text{ mm day}^{-1}$ ,  $> 24 \text{ cm day}^{-1}$ ,  $> 34^\circ\text{C}$ , 그리고  $< -14^\circ\text{C}$ 에 해당하는 수치이다.

### 3. 동아시아지역 대표 날씨유형

#### 3.1 최적 연구영역

동아시아 지역의 대표 날씨유형을 선정하기 위해 크게 두 가지 영역이 제시되었다. 제시된 두 가지 영역은 기상청 분석일기도 영역( $20\sim60^\circ\text{N}$ ,  $100\sim150^\circ\text{E}$ , 좁은 영역)과 약 5일 동안 기압계 변화를 감시할 수 있는  $10\sim70^\circ\text{N}$ ,  $80\sim160^\circ\text{E}$ (넓은 영역)이다. 최적 연구영역을 설정하기 위해 EV (Explained Variation)를 계산

**Fig. 3.** Comparison of EV score between wide area ( $10\sim70^\circ\text{N}$ ,  $80\sim160^\circ\text{E}$ ) and narrow area ( $20\sim60^\circ\text{N}$ ,  $100\sim150^\circ\text{E}$ ).

하였다(Beck et al., 2016). 본 연구에서는 종관 기단의 성질을 잘 반영하는 기온에 대해 EV를 조사하였다. 즉, EV 값을 이용하여 제시된 두 영역 중 남한 기온(일 평균기온)의 변동성을 가장 잘 설명하는 영역을 최적 연구영역으로 선정하였다. 여기서 남한 기온은 30년 동안의 연구기간과 공간 대표성을 고려하여 강릉, 서울, 청주, 대전, 대구, 광주, 부산 7개 지점 일 평균기온의 평균값으로 정의하였다.

EV는 다음 식을 이용하여 계산된다. 여기서 K와 N은 각각 군집번호(1~30번)와 날짜를 의미한다. 본 연구에서는 군집수를 30개( $k = 30$ )로 설정하여 계산하였다.  $N_k$ 는 군집별 빈도수를 나타내고,  $\bar{a}_k$ 는 각 군집에 해당하는 일 평균기온의 평균,  $a_i$ 는 일 평균기온, 그리고  $\bar{a}$ 는 전체 일 평균기온의 평균값을 의미한다. 일 반적으로 EV 값이 클수록 기온과 같은 기상변수의 변동성을 잘 설명함을 의미한다.

$$EV = \frac{\sum_{k=1}^K N_k (\bar{a}_k - \bar{a})^2}{\sum_{i=1}^N (a_i - \bar{a})^2}$$

Figure 3은 제시된 두 영역에 대해 월별 EV 값을 보여준다. EV 값은 여름철을 제외한 계절에서 상대적으로 높은 수치를 보였고, 상대적으로 좁은 영역(분

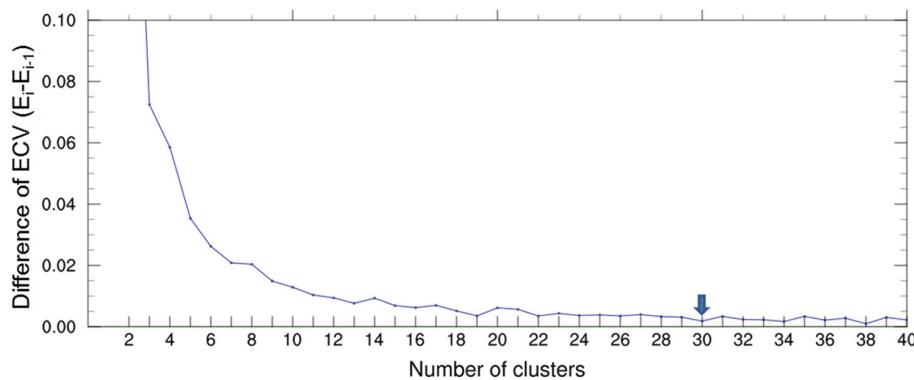


Fig. 4. Difference of ECV ( $E_i - E_{i-1}$ ) on the number of cluster.

석일기도 영역)에서 더 높은 수치를 보였다. 이러한 특징은 겨울철의 경우, 기단의 변동성(강도 및 확장)에 따라 남한 기온이 크게 영향을 받는 특징에 기인하는 것으로 판단된다. 최적 연구영역을 선정하기 위해 본 연구에서는 좁은 영역(분석일기도 영역)과 넓은 영역(5일 예보 고려) 중 EV 값이 상대적으로 큰 분석일기도 영역( $20\sim60^{\circ}\text{N}$ ,  $100\sim150^{\circ}\text{E}$ )을 선정하였다. 이러한 결과는 영국기상청 연구결과(위/경도 범위: 35/50도)와 유사한 범위를 보였다(Neal et al., 2016).

### 3.2 최적 군집 개수 선정

선정된 연구영역(동아시아 지역)에서 나타나는 날씨 유형의 군집수를 결정하기 위해 ECV (Explained Cluster Variance)를 계산하였다. 일반적으로 최적 군집수는 ECV 값이 포화되는 위치로 정의된다(Philipp et al., 2007; Beck and Philipp, 2010; Hoffmann and Schlünzen, 2013). ECV는 다음과 같이 정의하였다.

$$WSS = \sum_{k=1}^{k=n} \sum_{i=1}^{i=Nk} (x_{ki} - c_k)^2$$

$$ECV = 1 - \frac{WSS}{TSS}$$

여기서 WSS (Within cluster Sum of Squares)는 군집수에 따른( $k = 1\sim\infty$ ) 해면기압 차이의 합을 나타내고,  $c_k$ 는 군집별 대표 해면기압,  $x_{ki}$ 는 군집에 해당하는 일별 해면기압을 의미한다. TSS (Total Sum of Squares)는 군집수가 1 ( $k = 1$ )일 때, 해면기압 차이의 합을 나타낸다. ECV는 1에서 WSS와 TSS의 비를 뺀 값으로 WSS의 군집수( $k$ )를 순차적으로 증가시켜 계산하게 된다.

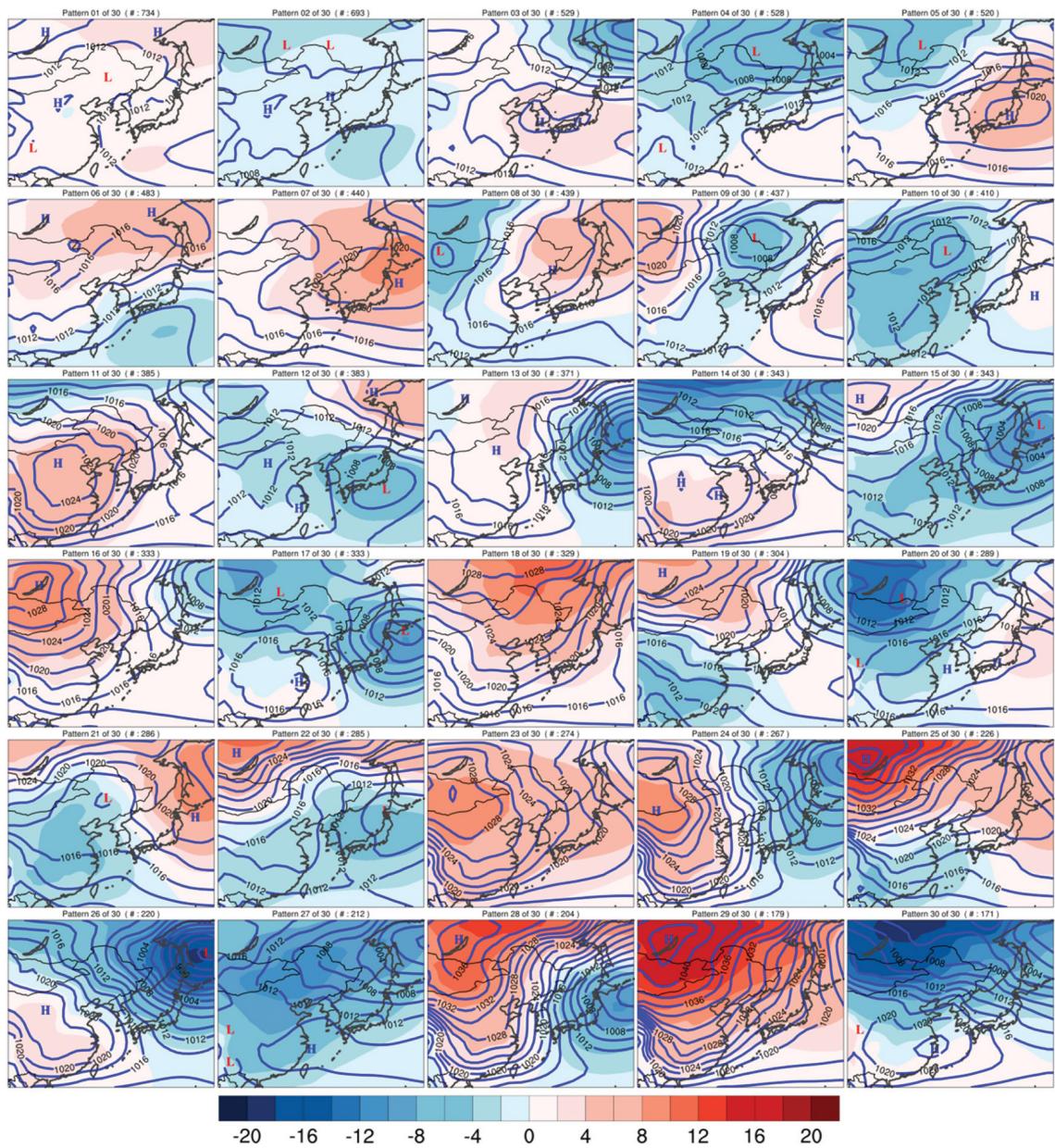
Figure 4는 군집수 증가에 따른 ECV의 차이값( $E_i - E_{i-1}$ )을 보여준다. 군집수가 2 ( $k = 2$ )일 때, ECV는 0.224 수치를 보였고, 군집수가 증가할수록 점차 감소

하여 군집수가 30 ( $k = 30$ )일 때, ECV는 약 0.002로 감소폭이 가장 큰 ECV 값의 1% 이하에 해당하는 수치를 보였다. 이후 가장 큰 ECV 값 1% 이내 수준에서 일정한 수치를 유지하였다. 본 연구에서는 가장 큰 ECV 값의 1% 이하 수치가 처음으로 나타나는 30개 ( $k = 30$ )를 최적 군집수로 정의하였다(Fig. 4 화살표).

### 3.3 동아시아 지역 대표 날씨유형

지난 30년(1981~2010년) 동안의 ERA-I 평균해면기압 자료와 K-평균 군집분석을 활용하여 동아시아 지역을 대표하는 30개 날씨유형을 선정하였다(Fig. 5). 발생빈도가 높은 순서에 따라 유형 1번에서 유형 30번까지 순차적으로 순서를 배치하였다. 음영과 실선은 각각 해면기압 아노말리와 해면기압을 의미한다. 일반적으로 해면기압의 강도는 낮은 번호로 분류된 날씨유형의 경우, 해면기압 아노말리의 강도가 상대적으로 약한 반면 높은 번호로 분류된 날씨유형은 아노말리의 강도가 상대적으로 강하게 나타났다. Table 2는 날씨유형에 따라 월별 발생빈도를 보여준다. 낮은 날씨유형 번호일수록 전 계절에 걸쳐 발생하였지만 여름철의 발생빈도가 상대적으로 높게 나타났다. 반면 날씨유형 번호가 높을수록 겨울철의 발생빈도가 다른 계절에 비해 상대적으로 높았다.

날씨유형별 해면기압의 특징은 날씨유형 1번의 경우, 연구지역 전체에서 양의 해면기압 아노말리가 분포하였고, 한반도는 고기압의 영향권에 있는 것으로 분석되었다. 발생빈도는 상대적으로 여름철(6~9월)에 약 56%로 높게 나타났다. 반면 날씨유형의 번호가 높아질수록 여름철의 발생빈도가 감소하였고, 주로 겨울철에 발생빈도가 최대 4~5%로 여름철에 비해 상대적으로 높게 나타났다. 해면기압 특징도 아노말리의 강도가 다른 유형에 비해 강하고 나타났고, 대부분의 날씨유형에서 겨울철 기압계 특징을 보여주는 서고동 저형 기압 배치를 보였다. 이러한 결과를 통해 계절



**Fig. 5.** Weather patterns over the East Asia region. Contour line and Shaded indicate MSLP and MSLP anomaly, respectively.

별로 주로 발생하는 날씨유형을 확인할 수 있었다. 이러한 특징을 바탕으로 날씨유형에 따라 풍계와 주요 기압 위치 등 주요 특징을 Table 3에 요약하였다.

중위도 지역에 위치한 한반도는 호우, 대설, 폭염, 그리고 한파 등 계절에 따라 뚜렷한 위험기상이 발생 한다. 본 연구에서는 연구기간(30년) 동안 과거 남한 지역에서 관측된 95개 지점 ASOS 자료를 이용하여 날씨유형에 따른 위험기상의 발생일수를 조사하였다.

위험기상은 극한 강수현상인 호우와 대설 그리고 극한 기온현상인 폭염과 한파에 대해 분석하였다. Figure 6은 날씨유형에 따른 총 발생빈도(상단 그림)와 위험 기상의 발생일수(중하단 그림)를 보여준다. 위험기상 사례 선정 시 사용된 기준치는  $> 110 \text{ mm day}^{-1}$ ,  $> 24 \text{ cm day}^{-1}$ ,  $> 34^\circ\text{C}$ , 그리고  $< -14^\circ\text{C}$ 로 극값 1%에 해당하는 수치이다. 95개 ASOS 지점 중 1개 지점이라도 해당 기준에 도달하면 위험기상 사례로 선정하

**Table 2.** Monthly frequency according to weather pattern (Group Number; GN).

G N	Month												Ave Occ. (%)
	1	2	3	4	5	6	7	8	9	10	11	12	
1	2.7	2.7	2.6	2.8	5.1	12.3	15.2	18.7	9.4	4.5	2.8	1.3	6.7
2	1.8	4.5	3.0	3.0	4.6	11.8	17.3	10.8	8.3	5.0	2.4	3.2	6.3
3	2.4	2.9	3.4	3.7	6.6	8.9	8.9	7.0	5.6	3.0	2.9	2.7	4.8
4	3.4	2.9	2.7	5.4	4.8	7.3	7.0	7.7	5.7	3.8	4.0	3.0	4.8
5	4.8	2.9	4.3	3.9	5.8	5.4	4.2	2.7	7.4	7.1	4.3	4.0	4.8
6	2.6	1.8	2.6	3.6	6.7	6.4	7.6	7.1	6.6	3.2	3.2	1.4	4.4
7	2.6	4.3	2.9	4.4	4.6	4.3	4.7	5.2	3.2	4.8	3.3	3.8	4.0
8	4.3	3.9	3.9	4.8	4.6	3.8	3.9	2.8	3.2	5.7	4.2	3.0	4.0
9	2.4	3.0	3.1	3.8	5.7	5.8	4.2	5.3	4.3	4.4	2.9	3.0	4.0
10	4.1	4.3	4.6	4.3	3.9	2.4	2.0	4.0	3.4	3.2	4.7	4.0	3.7
11	4.1	4.5	3.9	4.8	3.4	2.6	1.7	2.8	2.9	3.7	4.0	4.0	3.5
12	1.9	2.6	3.2	3.4	4.8	5.2	3.9	3.7	6.6	2.8	1.6	2.3	3.5
13	4.0	4.1	3.1	3.0	3.1	4.0	3.6	3.1	3.0	3.3	3.3	3.1	3.4
14	4.3	3.9	4.6	3.0	2.9	0.7	0.8	1.2	2.1	3.3	6.1	4.7	3.1
15	3.4	3.6	4.3	2.9	3.9	2.2	3.6	2.9	2.6	2.3	3.8	2.3	3.1
16	2.8	2.6	2.8	4.0	2.5	2.6	2.7	2.6	4.3	3.6	2.2	3.9	3.0
17	3.2	2.9	4.3	3.2	3.6	2.4	1.4	1.7	3.0	3.7	3.7	3.4	3.0
18	4.2	3.7	4.5	3.4	2.4	2.2	1.6	1.1	3.0	3.3	3.2	3.4	3.0
19	2.6	2.7	3.3	4.4	2.9	2.3	0.8	1.5	3.3	2.4	3.7	3.4	2.8
20	4.0	5.0	2.7	3.6	1.8	0.4	0.2	0.5	1.2	2.8	4.4	5.2	2.6
21	4.7	2.5	3.2	3.1	2.2	1.2	1.1	1.2	1.1	3.0	3.4	4.5	2.6
22	2.7	2.5	2.6	3.4	3.0	2.0	1.4	1.9	2.1	4.0	3.0	2.6	2.6
23	3.2	4.4	4.4	2.6	1.8	1.8	1.0	2.0	1.8	2.0	1.9	3.2	2.5
24	3.2	3.8	1.7	2.1	3.0	1.0	0.5	1.1	2.0	3.0	3.0	4.8	2.4
25	3.7	3.0	3.8	2.0	1.6	0.2	0.0	0.0	0.8	3.6	3.4	2.8	2.1
26	3.3	1.6	2.6	2.8	2.0	0.3	0.4	0.7	1.4	3.2	3.3	2.4	2.0
27	2.9	3.8	3.9	2.4	1.1	0.2	0.3	0.9	0.9	1.5	2.9	2.6	1.9
28	3.0	3.2	2.6	2.9	0.8	0.0	0.0	0.0	0.4	2.0	2.3	5.2	1.9
29	3.8	3.7	2.5	1.4	0.3	0.0	0.1	0.0	0.1	0.8	4.3	2.8	1.6
30	3.9	2.9	2.9	1.8	0.5	0.0	0.0	0.1	1.1	1.6	4.1	1.6	

였다. 여기서 위험기상 발생일수(%)는 해당 유형의 총 발생일수와 극값 1%에 해당하는 일수의 비율로 계산되었다.

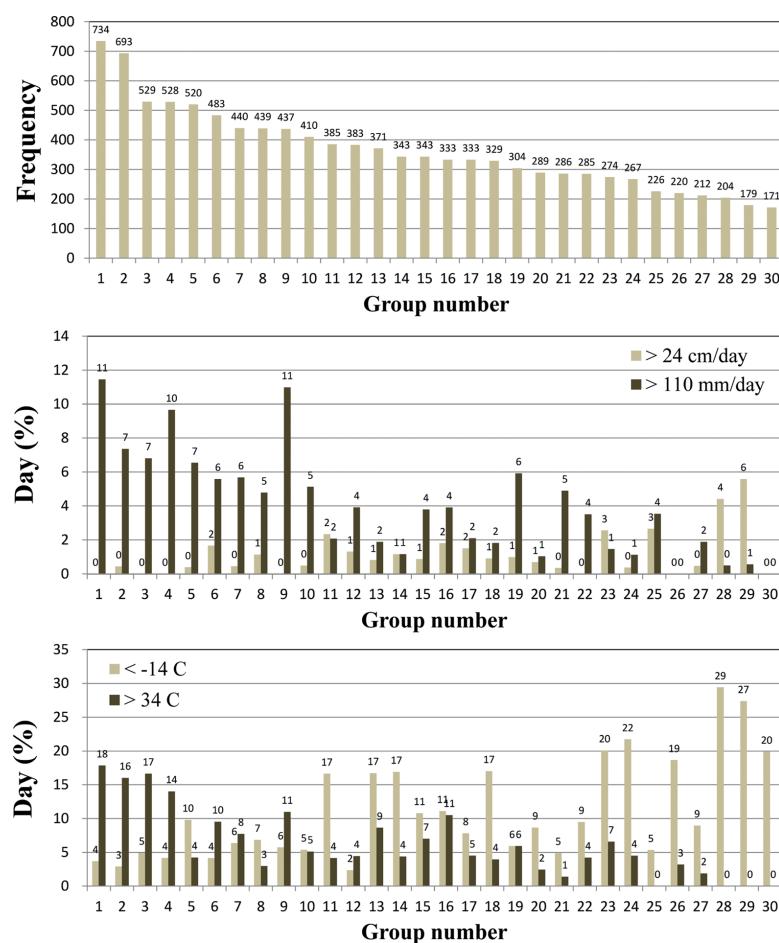
Figure 6의 상단 그림은 날씨유형에 따른 총 발생빈도로써 날씨유형 1번에 해당하는 빈도는 734일이고, 유형 30번은 171일로 분석되었다. 중앙과 하단에 위치한 그림은 각각 날씨유형에 따른 극한 대설과 호우 그리고 극한 기온 발생일수(%)를 보여준다. 일 강수량 110 mm를 초과하는 극한 호우는 유형 1번과 9번에서 약 11% 그리고 유형 4번에서 약 10% 순으로 나타났고, 일수로는 각각 84일, 48일, 그리고 51일 발

생하였다. 기압배치의 특징으로 유형 1번은 양의 해면기압 아노말리가 영역 전체에 분포하였고, 한반도에 고기압이 자리잡은 형태를 보였다. 따라서 이 유형과 관련된 호우는 열적 대류성 강수와 관련된 것으로 추정된다. 유형 9번과 4번은 한반도 북쪽에 저기압이 위치한 형태로 전선 또는 장마전선에 동반된 호우로 판단된다.

극한 대설현상( $> 24 \text{ cm day}^{-1}$ )은 날씨유형 28번과 29번에서 각각 약 4%(9일)와 6%(10일)의 발생일수(%)를 보였다. 기압계의 특징으로는 두 유형 모두 북서쪽에 강한 양의 해면기압 아노말리가 위치하였고, 이

**Table 3.** Characteristics of 30 weather patterns over the East Asia region.

GN	Description	GN	Description
1	Weak South-Westerly	16	Northerly, high of Siberia
2	Weak North-Westerly	17	Cyclonic North-Westerly, low E. of Japan
3	Anticyclonic	18	Anticyclonic Northerly, high N. of Korea
4	Cyclonic South-Westerly	19	Unbiased Easterly
5	Southerly, high of Japan	20	Westerly, low of Mongolia
6	Unbiased North-Easterly	21	Cyclonic Southerly
7	Anticyclonic, high N. of Japan	22	Cyclonic North-Westerly
8	North-Easterly, high N. of Korea	23	Anticyclonic North-Westerly, high of Mongolia
9	Cyclonic Westerly, low N. of Korea	24	Northerly, high and low in West and East
10	Cyclonic South-Westerly, low N. of Korea	25	Easterly, high of Siberia
11	Anticyclonic North-Westerly, high of China	26	Cyclonic North-Westerly, Deep low N.E. of Japan
12	Cyclonic North-Westerly, low of Japan	27	Cyclonic Westerly
13	Unbiased Northerly, low N.E. of Japan	28	Northerly, high and low in Siberia and Japan
14	North-Westerly, high of China	29	Anticyclonic North-Easterly, deep high of Siberia
15	Cyclonic Northerly, low N.E. of Japan	30	Unbiased Westerly

**Fig. 6.** Frequency (day; upper), day (%) of extreme precipitation (middle) and temperature (bottom) events according to weather pattern. Day (%) is calculated by the ratio of extreme days to total day by weather pattern.

위치와 동일하게 강한 고기압이 위치하였다. 이 기압 패턴은 전형적인 겨울철 한반도 대설과 관련된 유형으로써 고기압의 확장 방향에 따라 영동지역 또는 서해안지역에 강설현상이 발생하게 된다. 즉, 고기압이 동해상 또는 서해상으로 확장되는 방향에 따라 강설 현상이 영동 또는 서해안 지역으로 구분되어 발생된다.

극한 고온현상( $> 34^{\circ}\text{C}$ ) 그리고 저온현상( $< -14^{\circ}\text{C}$ )과 관련된 날씨유형을 분석하였다. 분석결과, 고온현상은 날씨유형 1~4번에 약 14~18%(73~131일)로 분포하였고, 저온현상은 28번과 29번 날씨유형에서 29%(60일) 그리고 27%(49일)의 발생일수(%)를 보였다. 극한 기온과 관련된 기압 배치는 주로 고온의 경우, 북태평양 고기압의 확장과 관련 있는 것으로 분석되었고, 저온은 시베리아 고기압의 확장과 관련이 있었다.

#### 4. 요약 및 결론

본 연구에서는 1981~2010년(30년) 동안의 ERA-I MSLP 아노말리 자료와 K-평균 군집분석을 활용하여 동아시아 지역의 대표 날씨유형을 분류하였다. 최적 연구영역과 군집수를 설정하기 위해 EV와 ECV를 계산하였다. 최적 연구영역은 제시된 두 영역 중 EV 수치가 상대적으로 높은 값을 보인  $20\text{--}60^{\circ}\text{N}$ ,  $100\text{--}150^{\circ}\text{E}$ (기상청 분석일기도 영역)을 선정하였다. 최적 군집수는 ECV 값이 포화되는 30개( $k = 30$ )로 정의되었다. 선정된 연구영역에 대해 동아시아 지역을 대표하는 30개 날씨유형을 선정하였다(Fig. 5). 날씨유형 1번에서 유형 30번까지 발생빈도가 높은 순서에 따라 순차적으로 배치하였다. 날씨유형 번호가 증가할수록 해면 기압 아노말리의 강도가 강하게 나타났고, 반면 발생빈도는 상대적으로 낮았다. 낮은 날씨유형 번호는 전계절에 걸쳐 발생하였지만 주로 여름철에 많이 발생하였다. 날씨유형 1번의 경우 주로 여름철(6~9월)에 발생빈도가 약 56%로 높게 발생하였다. 반면 높은 날씨유형 번호(25번 이상)는 겨울철에 발생빈도가 상대적으로 높았다.

날씨유형에 따른 위험기상의 발생일수를 조사하였다. 여기서 위험기상은 연구기간(30년) 동안 관측된 극값 1%에 해당하는 극한 강수현상과 극한 기온현상이다. 호우와 관련된 날씨유형은 1, 4, 9번 유형으로 10%(84, 51, 48일) 이상의 발생일수를 보였고, 주로 열대 대류성 강수현상 그리고 전선에 동반된 강수현상으로 판단된다. 또한 대설과 관련된 날씨유형은 28, 29번 유형으로 4%(9일) 그리고 6%(10일) 발생일수를 보였고, 시베리아 고기압과 관련이 있었다. 고온과 저온현상과 관련된 유형은 각각 날씨유형 1~4번 그리고 28~29번과 관련이 있었고, 발생일수는 각각 약 14~18%(73~131일) 그리고 27~29%(49~60일)로 분석되었

다. 이러한 고온과 저온현상은 주로 여름철 북태평양 고기압의 확장 그리고 겨울철 시베리아 고기압의 확장과 관련이 있었다.

이 연구 결과를 통해 계절별로 주로 발생하는 날씨 유형을 확인할 수 있었고, 위험기상과의 관련성이 높은 날씨유형을 확인할 수 있었다. 이러한 결과는 향후 기상청 중기예보 시 시나리오 기반의 앙상블 예측 자료 생산을 위해 활용될 것으로 기대된다. 그러나 본 연구에서는 일년 동안의 모든 날짜에 대해 군집분석을 수행하였기 때문에 폭염, 한파, 호우, 그리고 대설 등 위험기상 현상에 대해 직접적으로 적용하기에는 한계가 있을 것으로 판단된다. 향후 계절을 구분하거나 또는 위험기상을 현상별로 구분하여 날씨유형을 상세하게 분류하는 연구가 수행되어야 할 것이다.

#### 감사의 글

이 연구는 기상청 국립기상과학원 「위험기상에 대한 분석·예보의 융합기술 고도화」(1365003081)의 지원으로 수행되었습니다.

#### REFERENCES

- Beck, C., and A. Philipp, 2010: Evaluation and comparison of circulation type classifications for the European domain. *Phys. Chem. Earth*, **35**, 374-387, doi:10.1016/j.pce.2010.01.001.
- \_\_\_\_\_, \_\_\_, and F. Streicher, 2016: The effect of domain size on the relationship between circulation type classifications and surface climate. *Int. J. Climatol.*, **36**, 2692-2709, doi:10.1002/joc.3688.
- Casola, J. H., and J. M. Wallace, 2007: Identifying weather regimes in the wintertime 500-hPa geopotential height field for the Pacific-North American sector using a limited-contour clustering technique. *J. Appl. Meteor. Climatol.*, **46**, 1619-1630, doi:10.1175/jam2564.1.
- Choi, Y.-J., Y.-R. Jung, H.-J. Kim, K.-R. Kim, S.-Y. Lee, and H.-R. Lee, 2009: A study on regionalization by K-means method and time series pattern of precipitation areas in South Korea. *J. Korean Data Anal. Soc.*, **11**, 2761-2772 (in Korean with English abstract).
- Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.*, **137**, 553-597, doi:10.1002/qj.828.
- Esteban, P., P. D. Jones, J. Martín-Vide, and M. Mases, 2005: Atmospheric circulation patterns related to heavy snowfall days in Andorra, Pyrenees. *Int. J. Climatol.*, **25**, 319-329, doi:10.1002/joc.1103.

- Everitt, B., 1993: *Cluster analysis. 3rd edit.*, Halsted Press, 170 pp.
- Fereday, D. R., J. R. Knight, A. A. Scaife, C. K. Folland, and A. Philipp, 2008: Cluster analysis of north Atlantic-European circulation types and links with tropical Pacific sea surface temperatures. *J. Climate*, **21**, 3687-3703, doi: 10.1175/2007jcli1875.1.
- Fleig, A. K., L. M. Tallaksen, H. Hisdal, K. Stahl, and D. M. Hannah, 2010: Inter-comparison of weather and circulation type classifications for hydrological drought development. *Phys. Chem. Earth*, **35**, 507-515, doi: 10.1016/j.pce.2009.11.005.
- Goodess, C. M., and P. D. Jones, 2002: Links between circulation and changes in the characteristics of Iberian rainfall. *Int. J. Climatol.*, **22**, 1593-1615, doi:10.1002/joc.810.
- Hisashi, K., N. Yoshinori, and T. Seiji, 2013: Objective classification of the sea level pressure distribution pattern in East Asia: Analysis of the cold half of the year. *Geogr. Rev. Japan Ser. A*, **86**, 95-114, doi:10.4157/grj.86.95 (in Japanese with English abstract).
- Hoffmann, P., and K. H. Schluenzen, 2013: Weather pattern classification to represent the urban heat island in present and future climate. *J. Appl. Meteor. Climatol.*, **52**, 2699-2714, doi: 10.1175/jamc-d-12-065.1.
- Hoinka, K. P., C. Schwierz, and O. Martius, 2006: Synoptic-scale weather patterns during Alpine heavy rain events. *Q. J. R. Meteorol. Soc.*, **132**, 2853-2860, doi: 10.1256/qj.05.239.
- Hsu, C.-H., and F.-Y. Cheng, 2016: Classification of weather patterns to study the influence of meteorological characteristics on PM2.5 concentrations in Yunlin County, Taiwan. *Atmos. Environ.*, **144**, 397-408, doi:10.1016/j.atmosenv.2016.09.001.
- Huth, R., C. Beck, A. Philipp, M. Demuzere, Z. Ustrnul, M. Cahynová, J. Kyselý, and O. E. Tveito, 2008: Classifications of atmospheric circulation patterns: recent advances and applications. *Ann. N. Y. Acad. Sci.*, **1146**, 105-152, doi:10.1196/annals.1446.019.
- Kohonen, T., 1982: Self-organized formation of topologically correct feature maps. *Biol. Cybern.*, **43**, 59-69, doi:10.1007/bf00337288.
- Kyselý, J., 2007: Implications of enhanced persistence of atmospheric circulation for the occurrence and severity of temperature extremes. *Int. J. Climatol.*, **27**, 689-695, doi:10.1002/joc.1478.
- Lee, D.-K., and J.-G. Park, 1999: Regionalization of summer rainfall in South Korea using cluster analysis, *Asia-Pac. J. Atmos. Sci.*, **35**, 511-518 (in Korean with English abstract).
- Lim, W.-I., and K.-H. Seo, 2018: Investigation on characteristics of summertime extreme temperature events occurred in South Korea using self-organizing map. *Atmosphere*, **28**, 305-315, doi:10.14191/Atmos.2018.28.3.305 (in Korean with English abstract).
- MacQueen, J., 1967: Some methods for classification and analysis of multivariate observations. *Proc. fifth Berkeley Symp. on Math. Statist. and Prob.*, **1**, 281-297.
- Mukougawa, H., and M. Mabuchi, 2012: Characteristics of atmospheric circulation related to wintertime temperature variation over the Far-East. *Annals of Disas. Prev. Res. Inst.*, **55**, 247-253 (in Japanese with English abstract).
- Neal, R., D. Fereday, R. Crocker, and R. E. Comer, 2016: A flexible approach to defining weather patterns and their application in weather forecasting over Europe. *Meteor. Appl.*, **23**, 389-400, doi:10.1002/met.1563.
- \_\_\_\_\_, R. Dankers, A. Saulter, A. Lane, J. Millard, G. Robbins, and D. Price, 2018: Use of probabilistic medium- to long-range weather-pattern forecasts for identifying periods with an increased likelihood of coastal flooding around the UK. *Meteorol. Appl.*, **25**, 534-547, doi:10.1002/met.1719.
- Nguyen-Le, D., T. J. Yamada, and D. Tran-Anh, 2017: Classification and forecast of heavy rainfall in northern Kyushu during Baiu season using weather pattern recognition. *Atmos. Sci. Lett.*, **18**, 324-329, doi: 10.1002/asl.759.
- Ohba, M., S. Kadokura, Y. Yoshida, D. Nohara, and Y. Toyoda, 2015: Anomalous weather patterns in relation to heavy precipitation events in Japan during the Baiu season. *J. Hydrometeorol.*, **16**, 688-701, doi:10.1175/jhm-d-14-0124.1.
- Park, J.-G., and D.-K. Lee, 1998: Cluster analysis and development mechanism of explosive cyclones in East Asia. *Asia-Pac. J. Atmos. Sci.*, **34**, 523-547 (in Korean with English abstract).
- Philipp, A., P. M. Della-Marta, J. Jacobbeit, D. R. Fereday, P. D. Jones, A. Moberg, and H. Wanner, 2007: Long-term variability of daily north Atlantic-European pressure patterns since 1850 classified by simulated annealing clustering. *J. Climate*, **20**, 4065-4095, doi: 10.1175/jcli4175.1.
- Richardson, D., H. J. Fowler, C. G. Kilsby, and R. Neal, 2018: A new precipitation and drought climatology based on weather patterns. *Int. J. Climatol.*, **38**, 630-648, doi:10.1002/joc.5199.
- Robbins, J., R. Dankers, C. Dashwood, K. Lee, R. Neal,

- and H. Reeves, 2018: Early warning of landslides in Scotland using probabilistic weather pattern forecasts. *Proc. 20th EGU General Assembly 2018*, Vien, Austria, EGU2018, 7496.
- Steele, E., R. Neal, R. Dankers, N. Fournier, K. Mylne, P. Newell, A. Saulter, A. Skea, and J. Upton, 2018: Using weather pattern typology to identify calm weather windows for local marine operations. *Conf. Offshore Tech.*, OTC-28784-MS, Houston, Texas, USA, doi:10.4043/28784-MS
- Yoon, D., D.-H. Cha, G. Lee, C. Park, M.-I. Lee, and K.-H. Min, 2018: Impacts of synoptic and local factors on heat wave events over southeastern region of Korea in 2015. *J. Geophys. Res.: Atmos.*, **123**, 12081-12096, doi:10.1029/2018JD029247.
- Vicente-Serrano, S. M., and J. I. López-Moreno, 2006: The influence of atmospheric circulation at different spatial scales on winter drought variability through a semi-arid climatic gradient in Northeast Spain. *Int. J. Climatol.*, **26**, 1427-1453, doi:10.1002/joc.1387.
- Zhang, Y., A. Ding, H. Mao, W. Nie, D. Zhou, L. Liu, X. Huang, and C. Fu, 2016: Impact of synoptic weather patterns and inter-decadal climate variability on air quality in the North China Plain during 1980-2013. *Atmos. Environ.*, **124**, 119-128, doi:10.1016/j.atmosenv.2015.05.063.