

# Use of RegCM gridded dataset for thunderstorm favorable conditions analysis over Poland—climatological approach

Ewelina Walawender<sup>1</sup> · Rafał Kielar<sup>1</sup> · Zbigniew Ustrnul<sup>2</sup>

Received: 9 September 2014 / Accepted: 27 August 2015 / Published online: 12 September 2015  
© The Author(s) 2015. This article is published with open access at Springerlink.com

**Abstract** The paper analyzes equivalent data for a low density meteorological station network (spatially discontinuous data) and poor temporal homogeneity of thunderstorm observational data. Due to that, a Regional Climate Model (RegCM) dataset was tested. The Most Unstable Convective Available Potential Energy index value (MUCAPE) above the 200 J kg<sup>-1</sup> threshold was selected as a predictor describing favorable conditions for the occurrence of thunderstorms. The quality of the dataset was examined through a comparison between model results and soundings from several aerological stations in Central Europe. Good, statistically significant (0.05 significance level) results were obtained through correlation analysis; the value of Pearson's correlation coefficient was above 0.8 in every single case. Then, using methods associated with gridded climatology, data series for 44 weather stations were derived and an analysis of correlation between RegCM modeled data and in situ thunderstorm observations was conducted with coefficients in the range of 0.75–0.90. The possibility of employing the dataset in thunderstorm climatology analysis was checked via a few examples by mapping monthly, seasonal, and annual means. Moreover, long-term variability and trend analysis along with modeled MUCAPE data were tested. As a result, the RegCM modeled MUCAPE gridded dataset was proposed as an easily available, suitable, and valuable predictor for thunderstorm

climatology analysis and mapping. Finally, some limitations are discussed and recommendations for further improvements are given.

## 1 Introduction

Thunderstorms are regarded as commonly occurring atmospheric phenomena over the area of Poland and Central Europe in the summer season. Because of their intensity, violent nature, as well as socio-economic consequences, thunderstorms are also regarded as an extreme phenomenon whose spatial and temporal variability pattern is an important element in analysis and risk assessment for a given geographic area.

In order to conduct a reliable climatologic analysis of thunderstorm occurrence probability as well as to investigate long-term variability, it is essential to obtain data which are temporally (at least a 30-year period) and spatially continuous. The homogeneity of the data series is crucial in this process also.

As it is well known, thunderstorms are a phenomenon which is discontinuous both spatially and temporally. In spite of the development of meteorological radars and modern methods of lightning detection as well satellite products, climatological long-term data for Poland are based on visual observations made at weather stations.

The strongly local nature of thunderstorms, their spatial variability, and their subjectivity of visual observations made at weather stations make traditional observational data difficult to interpolate. Moreover, due to the low density of the weather station network, reliable interpolation and map creation are nearly impossible.

This issue was raised by the IPCC et al. (2001). Attention was called to the lack of sufficient information on the course of such small-scale phenomena needed to analyze trends in thunderstorm occurrence with respect to progressing climate change.

✉ Ewelina Walawender  
ewelina.walawender@imgw.pl

<sup>1</sup> Institute of Meteorology and Water Management–National Research Institute, 14 Borowego St., 30-215 Krakow, Poland

<sup>2</sup> Institute of Geography and Spatial Management, Jagiellonian University, 7 Gronostajowa St., 30-387 Krakow, Poland

The report stressed also the growing need for using climate models with proper spatial resolution to assess patterns and analyze spatial patterns relying on homogenous and spatially continuous data (Goody et al. 2002; Trenberth et al. 2002).

Recent years in research on spatial and temporal patterns of thunderstorm occurrence prove that data from observational data reanalyses can be used successfully (Bengtsson et al. 2004, 2007; Dee et al. 2011). Because of their homogenous character, reanalysis serve as a valuable background for the long-term analysis of the variability of atmospheric phenomena (Bengtsson et al. 2004, 2007; Dee et al. 2011). For this reason, reanalysis data were also used in the assessment of the variability of thunderstorm phenomena occurrences in the most recent IPCC report (2013). However, they are not sufficient as far as spatial resolution is concerned, and consequently, numerous attempts in employing downscaling methods and various types of climatic models have been made (e.g., Cavazos and Hewitson 2005; Yoshimura and Kanamitsu 2008; Dettinger 2013).

Data derived from a model does not contain information on a specific thunderstorm occurrence; it may however be used for determining conditions that favor thunderstorm occurrences over a given area, which has been confirmed by numerous works in this field (e.g., Marsh et al. 2007, 2009; Trapp et al. 2007, 2010; Lombardo and Colle 2010; Allen et al. 2011; Gensini et al. 2014a). Of course, one needs to approach results with a certain dose of criticism towards analytical errors resulting from errors in observational data and errors that follow the implemented method of modeling (Thorne and Vose 2010).

Nevertheless, modeled data, following validation, give the possibility to create a temporally and spatially continuous dataset, so that the climatology of conditions that favor thunderstorms can be reconstructed on various scales, including a global scale (Brooks et al. 2003; Riemann-Campe et al. 2011), continental (Romero et al. 2007), or a regional or national scale (Schneider and Dean 2008; Gensini and Ashley 2011; Holley et al. 2014).

The main aim of this paper is an evaluation of the possibility of the use of data obtained from the Regional Climate Model (RegCM) in order to determine conditions that favor thunderstorm occurrences in Poland in the period 1966–2010. Additionally, possible use in climatological analysis on various temporal scales including long-term variability analysis was investigated.

## 2 Data and methodology

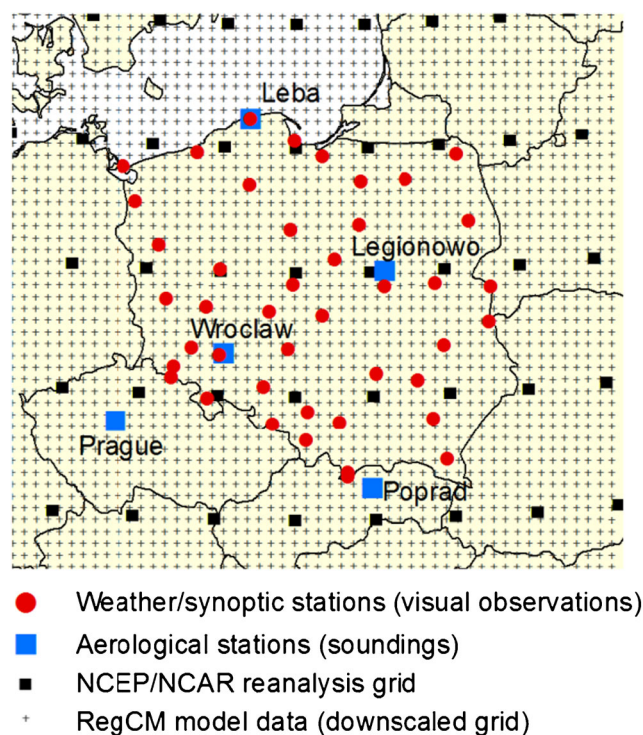
### 2.1 Observations vs. modeled data

In the period 1966–2010, there were 44 weather stations (synoptic type) in Poland that possessed almost continuous records

of thunderstorm observations. On the scale of a relatively large country (ca. 313,000 km<sup>2</sup>), the number of stations is definitely insufficient to create fully reliable climatologic maps with spatially continuous coverage. Additionally, the lack of a regular spread of the stations (especially in the south and east of the country; Fig. 1) results in the need for significant data extrapolation, which causes further errors.

Because of the abovementioned factors as well as the qualitative character of observational data (apart from duration of a phenomenon, it is difficult to determine its intensity through measurement), modeled data were analyzed instead. In addition, it was decided to conduct a climatologic analysis of conditions that favor convective phenomena occurrences, as we cannot derive data on the actual thunderstorm occurrence from a model.

To obtain spatially continuous data, data from the NCEP/NCAR (United States National Centers for Environmental Prediction–National Center for Atmospheric Research) (Kalnay et al. 1996; Kistler et al. 2001) were used. Data with a temporal resolution of 6 h were used for the surface level (2 m a.s.l.) and for 17 isobaric levels in the atmosphere. Because of insufficient spatial resolution (2.5° grid; Fig. 1) on the regional scale, the data were used as inputs in the process of downscaling (Wilby and Wigley 1997) carried out with the use of the RegCM model, version 4.1.1 (Elguindi et al. 2011).



**Fig. 1** Location of weather stations (meteorological and aerological) used in the paper's analysis in comparison to NCEP/NCAR reanalysis and RegCM model gridded data

The model, thanks to improved parameters of atmospheric physics and data on water resources as well as land cover, allows for the modeling of weather conditions for a domain on a regional scale (Giorgi and Anyah 2012). Thus, homogeneous data for the surface layer (2 m) and 23 upper levels (with 50 mb as the uppermost) with improved temporal and spatial resolution (3 h and ca. 20 km, respectively) in a regular grid were obtained.

Validation of the modeled data was completed using aerological soundings derived from five upper air measurement stations: Łeba (Łeba), Wrocław (Wrocław) and Legionowo in Poland, Poprad-Ganowce (Poprad) in Slovakia, and Praha-Libus (Prague) in the Czech Republic (Fig. 1).

## 2.2 MUCAPE as a predictor for thunderstorm occurrence

To determine the conditions that favor thunderstorm occurrences in Poland, it was decided that the CAPE parameter (Convective Available Potential Energy,  $\text{J kg}^{-1}$ ) would be used, considering both forecasters' experience and numerous international studies on its use as a convective phenomenon predictor (e.g., Rasmussen and Blanchard 1998; Brooks et al. 2003; Craven and Brooks 2004; Manzato 2005; DeRubertis 2006; Brooks et al. 2007; Adams and Souza 2009; Kunz et al. 2009; Riemann-Campe et al. 2011) as well as studies on the long-term variability of convective conditions (Ye et al. 1998; Trapp et al. 2007; Brooks 2013).

CAPE provides information on atmospheric instability by determining the amount of potentially available convective energy (positive instability energy value in an analyzed vertical profile). Testing several CAPE index derivatives allowed to select MUCAPE (Most Unstable CAPE,  $\text{J kg}^{-1}$ ) as the final predictor. MUCAPE is the index determining the instability energy computed for an air parcel at the most unstable isobaric level (the level at which the CAPE index has the highest value in the analyzed profile). It is used in thunderstorm detection as an index that is less sensitive to fluctuations in surface moisture and the boundary layer environment than CAPE, which considers buoyancy over the depth of the atmosphere (Lombardo and Colle 2010; Allen et al. 2011).

As a threshold value determining possible thunderstorm occurrence, the value of MUCAPE  $>200 \text{ J kg}^{-1}$  was selected. This choice was influenced by an analysis of thunderstorm occurrences in Poland as well as an analysis of studies that have addressed this issue. While it is true that even CAPE  $>100 \text{ J kg}^{-1}$  is regarded as a threshold designed to ensure that the environment can be characterized as being convective (Brooks et al. 2003; Kunz et al. 2009; Gensini et al. 2014b), in the USA, however, thunderstorms appear when CAPE is mainly about  $500 \text{ J kg}^{-1}$  (Rasmussen and Blanchard 1998). In Europe, the same type of comparison shows lower CAPE values during stormy weather, for example,  $260 \text{ J kg}^{-1}$  in Spain and  $210\text{--}400 \text{ J kg}^{-1}$  in Switzerland (Siedlecki 2009).

In the study conducted for The Netherlands, Haklander and Van Delden (2003) determined a value of MUCAPE  $\geq 168 \text{ J kg}^{-1}$  as an optimal threshold for thunderstorm probability forecasting. In the few studies carried out for thunderstorm conditions in Poland, a value of CAPE  $>300 \text{ J kg}^{-1}$  has been used most often (Malinowska 2011). However, this value eliminates most cases of thunderstorms that occur outside of the main thunderstorm season. Brooks et al. (2003) confirm that lower values of CAPE can be accompanied by severe thunderstorms. The threshold value of  $200 \text{ J kg}^{-1}$  selected for this study appears to be optimal for climatologic analysis using an annual approach, especially because the intensity of the thunderstorm phenomenon is not the main subject of this paper, but rather the temporal and spatial potential of conditions that favor the occurrence of convection.

Data modeled via the RegCM were first used to calculate MUCAPE values and then were validated in relation to the occurrence of favorable (extreme) parameters. All grid points for a chosen domain were analyzed for each of the eight observational terms for all days in the period 1966–2010. To compute the MUCAPE index, 23 sigma areas (areas of model calculations) were taken into account. For every given level, a CAPE index value was calculated using the equation proposed by Moncrieff and Miller (1976):

$$\text{CAPE} = g \int_{z_{LFC}}^{z_{EL}} \left( \frac{T_{vp} - T_{ve}}{T_{ve}} \right) dz$$

where

$g$ —acceleration due to gravity

$z_{LFC}$ —height of the level of free convection

$z_{EL}$ —height of the equilibrium level (neutral buoyancy)

$T_{vp}$ —virtual temperature of the specific parcel

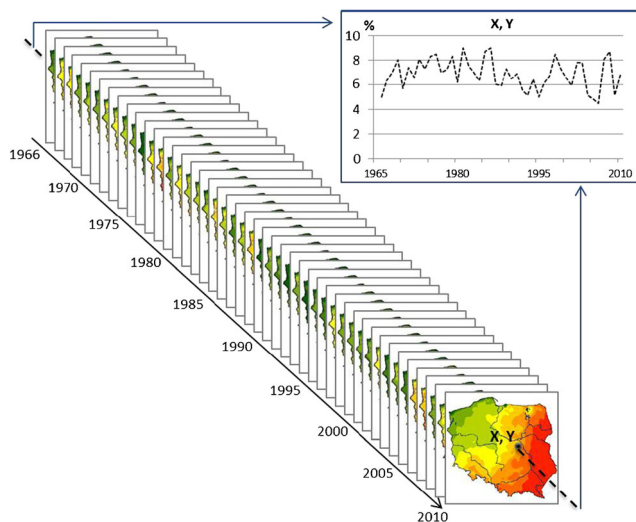
$T_{ve}$ —virtual temperature of the environment

$z$ —height

Next, the maximum for all 23 levels for a given point in a set time period was taken as MUCAPE.

## 2.3 Mapping methods and gridded climatology possibilities

Basing on the set MUCAPE threshold, the number of cases that met the “appropriate” conditions was calculated for every point of a regular grid for each month of a given year in the period 1966–2010. Then, values were computed for a mean monthly, seasonal, and annual frequency. Next, they were interpolated with the use of radial basis functions to create a continuous raster dataset. In this way, more than 550 maps of frequency of MUCAPE  $>200 \text{ J kg}^{-1}$  occurrence over the territory of Poland were generated. By using spatial analysis tools provided by geographic information systems (GIS), it was possible to determine long-term variability of the analyzed index for any selected site (Fig. 2).



**Fig. 2** Gridded climatology approach: variability of annual frequency (%) of MUCAPE  $>200 \text{ J kg}^{-1}$  (1966–2010) at a selected XY point derived through the extraction of data from 45 annual frequency maps

The solutions used in the study follow methodology associated with so-called gridded climatology (Tveit et al. 2000, 2005; Ustrnul 2001; Ustrnul and Czekierda 2006; Hahn and Warren 2007; Perčec Tadić 2010, Cecil et al. 2014) and offer new possibilities in the case of insufficient in situ data. These methods are helpful especially in long-term variability and trend analysis.

### 3 Results

#### 3.1 Modeled and reanalysis data for MUCAPE climatology analysis over Poland

The evaluation of possible use of modeled data was preceded by RegCM data quality validation. The RegCM monthly number of days with MUCAPE  $>200 \text{ J kg}^{-1}$  was calculated and correlated with the same parameter computed from upper air data (aerological data/soundings) derived from five

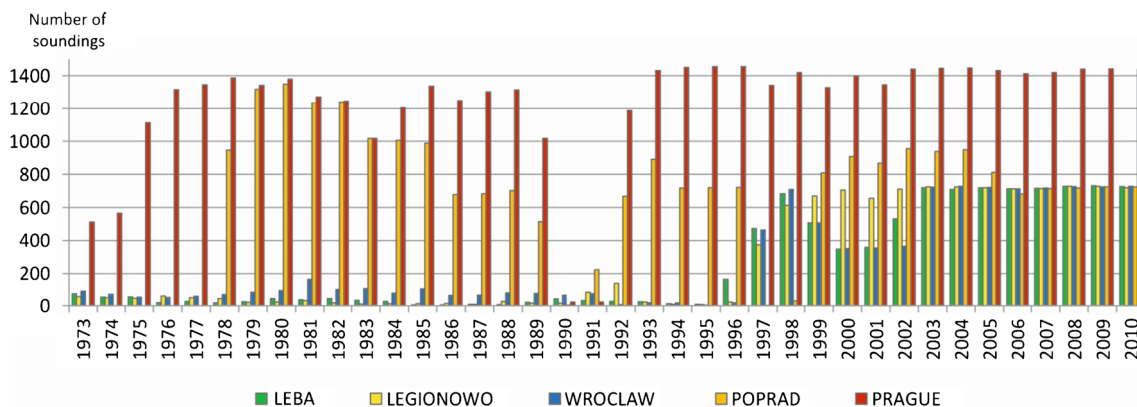
aerological stations: three Polish (Wrocław, Leba, and Legionowo) and two located nearby (Prague and Poprad) (Fig. 1). The Prague data series was particularly valuable because of the number of soundings conducting daily at the station (every 6 h at 00, 6, 12, and 18 UTC in contrast to two soundings obtained at 00 and 12 UTC at other stations), which seemed to be closest to the modeled data temporal resolution. All soundings were obtained from the University of Wyoming sounding archive (University of Wyoming 2014). Figure 3 shows the completeness of the data available at the Wyoming database.

Due to data gaps, three time intervals were selected for correlation analysis. The basic study period for all the stations analyzed was 1997–2010 (Table 1a). Eight years (2003–2010) with the most complete data were considered separately (Table 1b). Because of best data availability, results for Prague (1973–2010) and Poprad (1978–2010) were tested for long time intervals (more than 30 years taking into account all gaps, Table 1c).

The data from the Polish weather stations were correlated additionally with the monthly number of thunderstorms observed at the nearest station.

The relationship between datasets was examined with the use of Pearson's correlation coefficient (PCC).

All of the calculated PCC values indicate a strongly or very strongly positive relationship between MUCAPE calculated from RegCM data and soundings ( $\text{PCC} > 0.8$ , cf. *RegCM* vs. *soundings* in Table 1a, b and c). The lowest values were obtained for the Wrocław station ( $\text{PCC} = 0.812$ ) and for Leba ( $\text{PCC} = 0.836$ ) over the period 1997–2010, which can be explained by a smaller number of soundings in comparison to other stations (Fig. 3). The analysis of the complete data set from 2003 to 2010 indicated a stronger relationship (0.888, Leba and 0.894, Wrocław). In both cases, the strongest relationship between RegCM data and aerological data was observed for Poprad data, even though the number of soundings was, most of the time, smaller than that for Prague.



**Fig. 3** Annual number of aerological soundings available from the University of Wyoming (2014) database for weather stations included in the analysis (1973–2010 period)

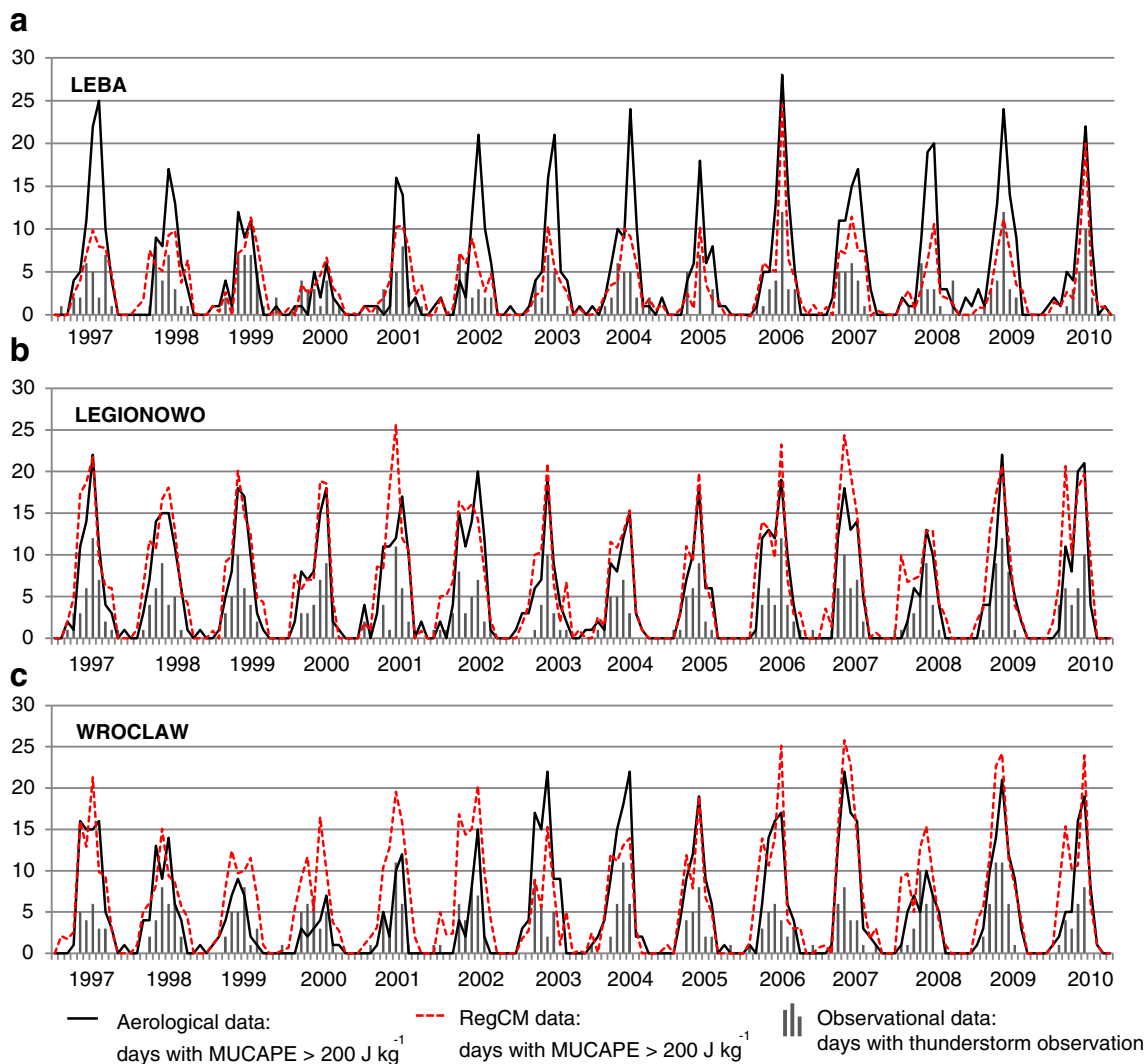


**Table 1** Pearson correlation coefficients between the number of cases derived from modeled RegCM data, soundings (monthly number of days with MUCAPE  $> 200 \text{ J kg}^{-1}$ ), and observations (monthly numbers of days with thunderstorm occurrence): *A*—1997–2010 data series; *B*—2003–2010 data series; *C*—more than 30 years data series

(A) PCC (1997–2010)	LEBA	LEGIONOWO	WROCLAW	PRAGUE	POPRAD
RegCM vs. soundings	0.836	0.918	0.812	0.918	0.949
RegCM vs. observations	0.838	0.890	0.818	–	–
Soundings vs. observations	0.760	0.901	0.773	–	–
(B) PCC (2003–2010)	LEBA	LEGIONOWO	WROCLAW	PRAGUE	POPRAD
RegCM vs. soundings	0.888	0.926	0.894	0.921	0.953
RegCM vs. observations	0.849	0.917	0.818	–	–
Soundings vs. observations	0.848	0.917	0.817	–	–
(C) PCC (>30 years)	PRAGUE (1973–2010)	POPRAD (1978–2010)			
RegCM vs. soundings	0.885	0.900			

A strong correlation ( $\text{PCC} > 0.8$ ) was also found between the results from RegCM and thunderstorm observational data (*RegCM vs. observations* in Table 1a, b). Values of the correlation coefficient do not vary significantly between the 1997–2010 and 2003–2010 time periods, which differ from the

comparison of observational data with soundings (*Soundings vs. observations* in Table 1a, b). For Polish stations, in both cases, the strongest relationship was observed for Legionowo. It should be mentioned that during the period of the homogeneous number of available aerological soundings (Table 1b),



**Fig. 4** Long-term variability (1997–2010) of the monthly number of days with MUCAPE  $> 200 \text{ J kg}^{-1}$  derived from the RegCM model and aerological soundings compared to days with thunderstorm observations at weather stations in Leba (a), Legionowo (b), and Wroclaw (c)

obtained PCC values are close to identical in the case of the investigation of the relationship between observational data both with the RegCM data and soundings. Therefore, it can be argued that the derived RegCM data set is quite reliable and can be implemented in further analyses.

Such a conclusion can be confirmed also by charts of monthly patterns with conditions that favor thunderstorm occurrences versus recorded observations of thunderstorms (Fig. 4). The RegCM data demonstrates temporal variability of convective phenomena quite well, and only in the case of Leba (Fig. 4a), convective phenomena are presented by this data as less frequent when compared to the aerological data. The cause of this difference may lay in the spatial resolution of the RegCM data (ca. 20 km, see section 2.1), which makes it impossible to present local climate characteristics. A larger number of thunderstorms were noted in Leba in comparison with the rest of the Polish coastline. For this reason, the station in Leba, as the only coastal one, was included into a different thunderstorm region by Bielec-Bąkowska (2003 and 2013).

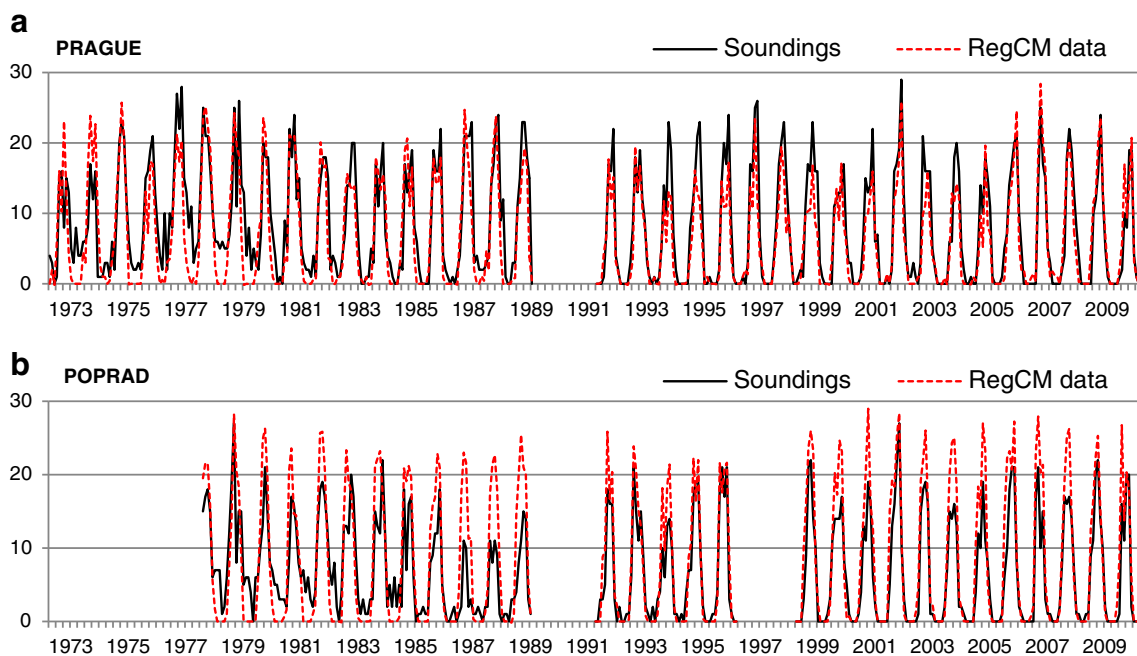
Taking into account the long-term correlation analysis for Prague and Poprad, a strong relationship between data sets and a slightly higher correlation coefficient for Poprad were also observed (Table 1c). However, by analyzing the monthly pattern of the number of days with MUCAPE  $>200 \text{ J kg}^{-1}$  (Fig. 5), it can be observed that aerological data for Prague is quantitatively similar to the modeled data (Fig. 5a), which almost certainly was caused by more sounding data available per day (conducted every 6 h). In Poprad (706 m a.s.l.), located at the foot of the High Tatra Mountains, the environmental

conditions for convection are more favorable. Therefore, a conclusion can be drawn that aerological data are being underestimated (Fig. 5b), whereas results obtained with the RegCM data represent the actual frequency of conditions that favor thunderstorm occurrences (as these could occur during the day in-between the aerological soundings carried out).

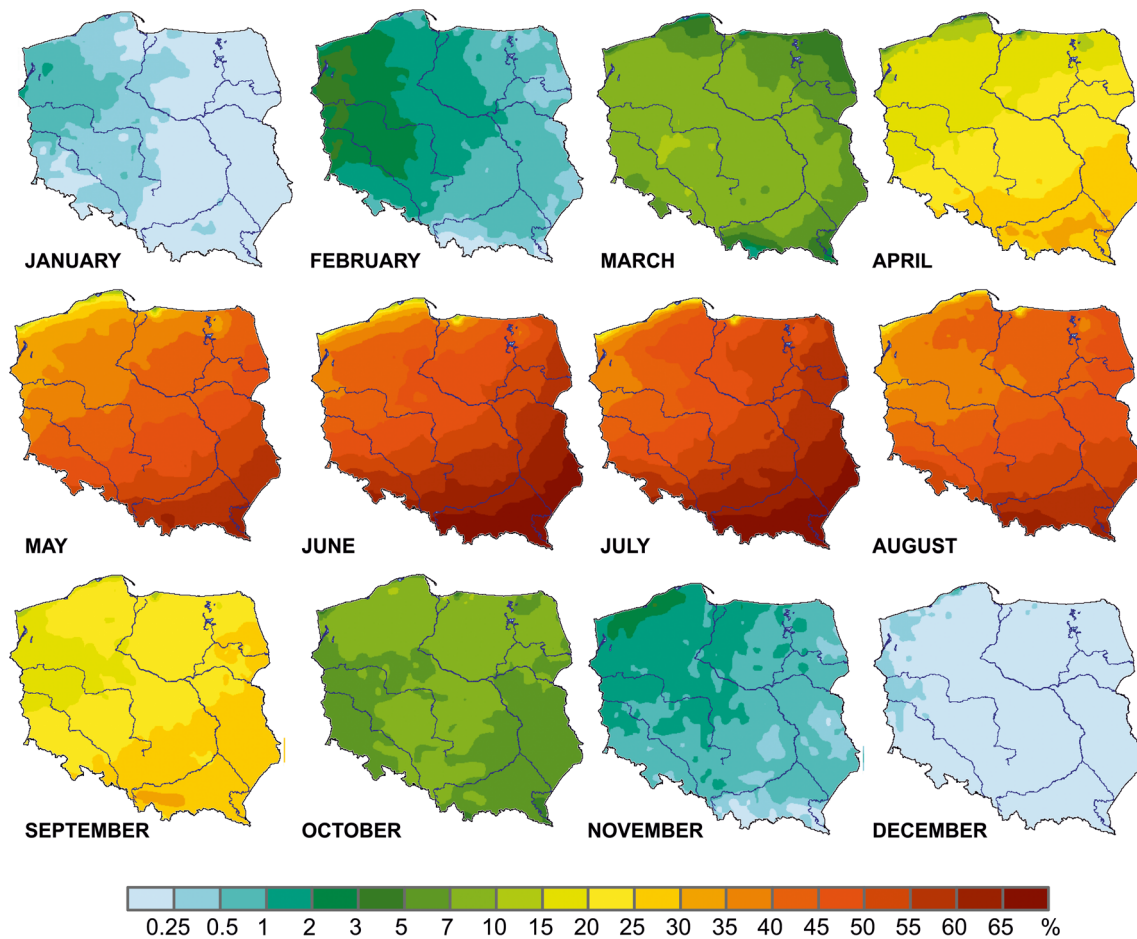
### 3.2 Use of MUCAPE RegCM dataset for spatial analysis and mapping—mean monthly, annual, and seasonal thunderstorm frequency

Using explanatory data analysis methods, RegCM modeled data were tested for its possible use in spatial analyses and the creation of climatologic maps as a predictor of thunderstorm occurrences in Poland. The use of the data as an explanatory variable was investigated for three temporal scales: monthly, seasonal (thunderstorm season from May to August), and annual.

Modeled data for the entire 1966–2010 period were used as an input into GIS zonal statistical methods, which made it possible to generate maps of average monthly probability of days with conditions that favor thunderstorm occurrence in every month of the year (Fig. 6). The output maps can be regarded as climate maps (45-year analysis) presenting temporal and spatial variability of convective conditions in Poland. The maps quite accurately represent actual winter season thunderstorm occurrences in Poland. During the winter season, the spatial pattern of thunderstorm occurrence is reversed—the most frequent observations are in the



**Fig. 5** Long-term variability of the monthly number of days with MUCAPE  $>200 \text{ J kg}^{-1}$  derived from the RegCM data and aerological soundings in Prague (a) and Poprad (b) (period of data accessibility from the University of Wyoming (2014))



**Fig. 6** Monthly frequency (%) of days with MUCAPE  $>200 \text{ J kg}^{-1}$  (RegCM dataset, 1966–2010)

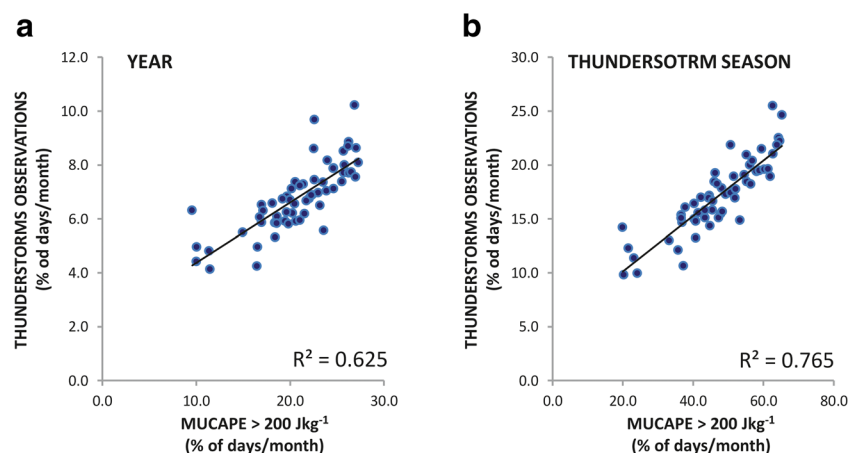
northwestern part of the country, with smaller frequency towards the southeast. The frequency is very small, usually 1–2 % and close to 0 % in December and January. During the warm half of the year, more frequent thunderstorm occurrence in the southeastern part of Poland is typical, with a maximum occurring between May and August (Bielec-Bąkowska 2003 and 2013, Kolendowicz 2004).

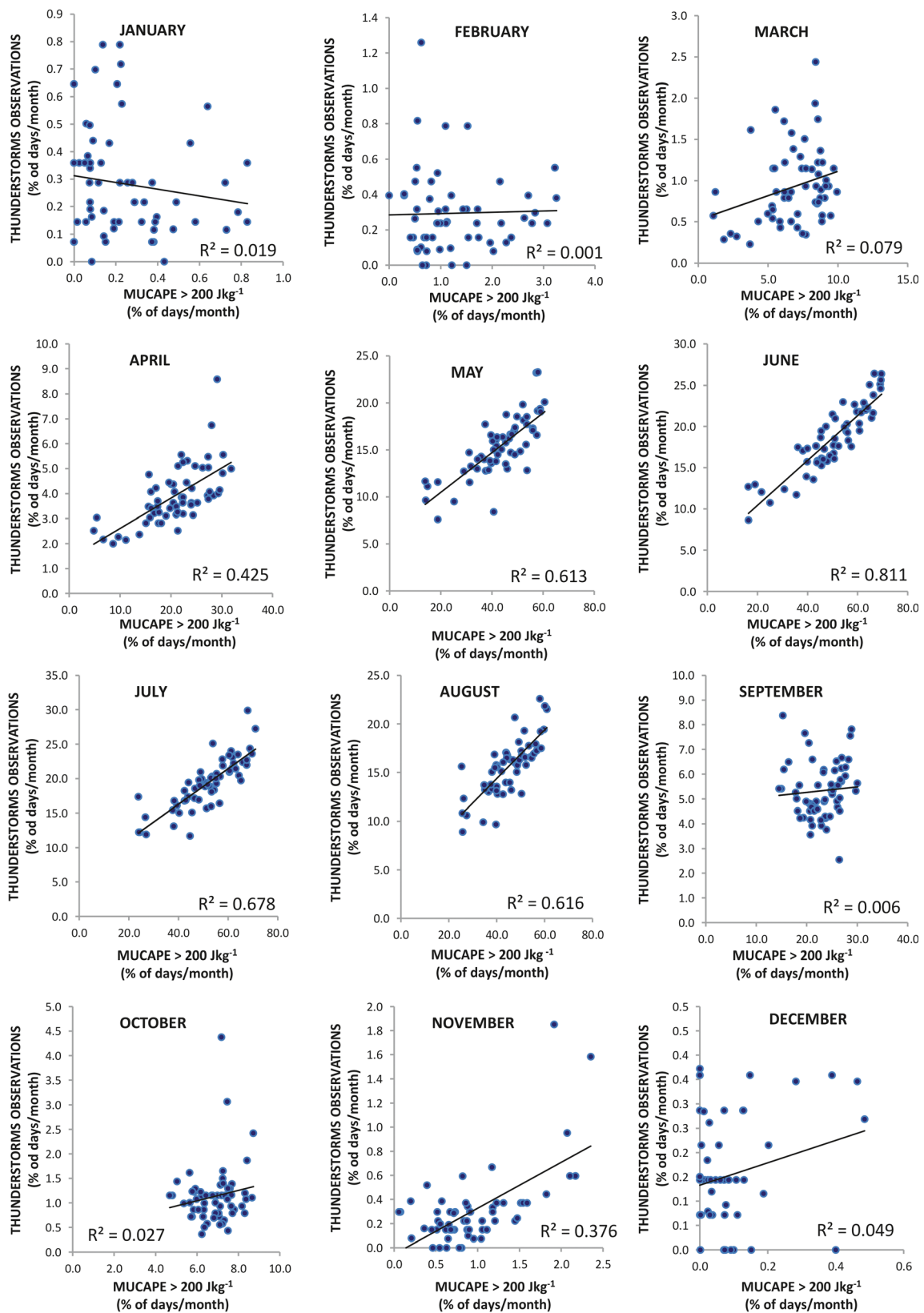
In order to conduct a validation and confirm the possibility of using RegCM data as a variable that explains thunderstorm

occurrences in Poland, the relationship between the modeled data and the mean monthly frequency of thunderstorm observations at these stations was investigated. Values of the monthly mean frequency of occurrence of days with MUCAPE  $>200 \text{ J kg}^{-1}$  were attributed to the sites of the weather stations.

To present the characteristics of the relationships between the data sets, scatterplots were used (Figs. 7 and 8). For data sets that did show a positive linear dependence, the value of the coefficient of determination was additionally computed ( $R^2$ ).

**Fig. 7** Relationship between the average number of days with thunderstorm observations at weather stations and average number of days with favorable conditions (MUCAPE  $>200 \text{ J kg}^{-1}$ ): **a** annual, **b** during the thunderstorm season (May to August);  $R^2$  coefficient of determination





**Fig. 8** Relationship between the average monthly number of days with thunderstorm observations at weather stations and number of days with favorable conditions (MUCAPE > 200 J kg<sup>-1</sup>); R<sup>2</sup> coefficient of determination



RegCM data showed a relatively strong dependence on observational data for the entire year and for the thunderstorm season—viewed both as a whole (Fig. 7) and when divided into separate months from May to August (Fig. 8). On a monthly scale, the best match between modeled data and thunderstorm observational data was found for June ( $R^2 > 0.8$ ) and July ( $R^2$  close to 0.7), which happens to be when thunderstorm activity is at its maximum and is characterized by a quite regular spatial pattern. A completely random distribution and no defined relationship are shown in scatterplots for the non-thunderstorm season, especially for the winter months. Because of the low frequency of thunderstorm observations at that time, the results should not be treated as a proof of the poor value of modeled data, especially due to the fact that thunderstorm occurrence in Poland in this period can be described as a random event (Kolendowicz 2004). Considering all of the results, it may be acknowledged that the data set provided by RegCM can be a valuable explanatory variable in numerous spatial analyses.

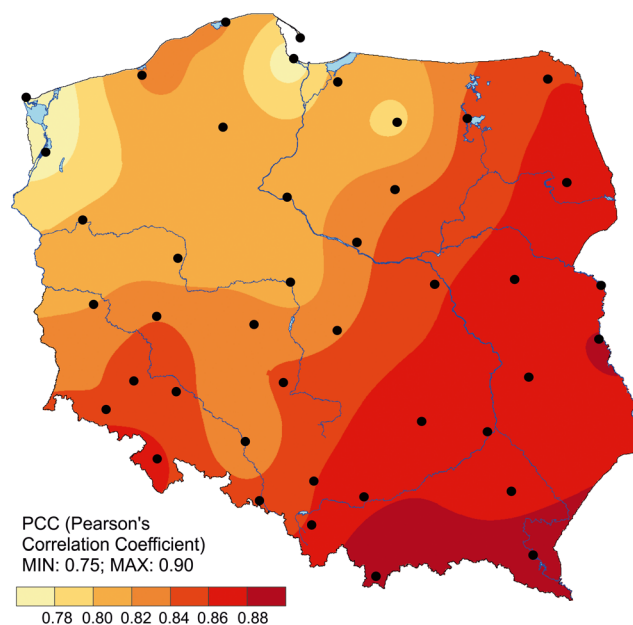
### 3.3 Use of the MUCAPE RegCM dataset for long-term spatial variability

In the last step of the research process, the possibility of using obtained modeled data for long-term variability spatial analysis was investigated.

The results of validation carried out with the use of aerological data (soundings) presented in chapter 2.1 confirmed that the derived gridded data can be successfully used to present MUCAPE index long-term variability. It can be suggested that the data can equally accurately show long-term variability of any other CAPE derived index, which makes it an exceptionally valuable material when the number of upper-air stations is small, as it is in the case of Poland.

To check the spatial relationship between long-term modeled and in situ data, correlation coefficients were calculated for 44 weather stations in Poland. As it was investigated in terms of long-term annual variability, a series consisting of 540 elements, monthly values for 45 years ( $45 \times 12$ ), were compared. PCC values for all 44 points were statistically significant at  $p = 0.001$  level and varied from 0.75 in the Baltic Sea region to 0.90 in the southeastern part of the country. In the second step, PCC values were interpolated to obtain a spatially continuous map (Fig. 9).

An attempt was also made to check how PCC values varied spatially in particular periods of the year. Because of the revealed lack of a linear relationship and infrequent occurrence of thunderstorms in winter seasons (Fig. 8), the analysis was conducted for the period with the most active “thunderstorm months” (May, June, July, and August) and as a whole season (May to August) as well as for the year, taking into account all component data for the 45-year long period (Table 2). Monthly data from June were the most consistent for modeled



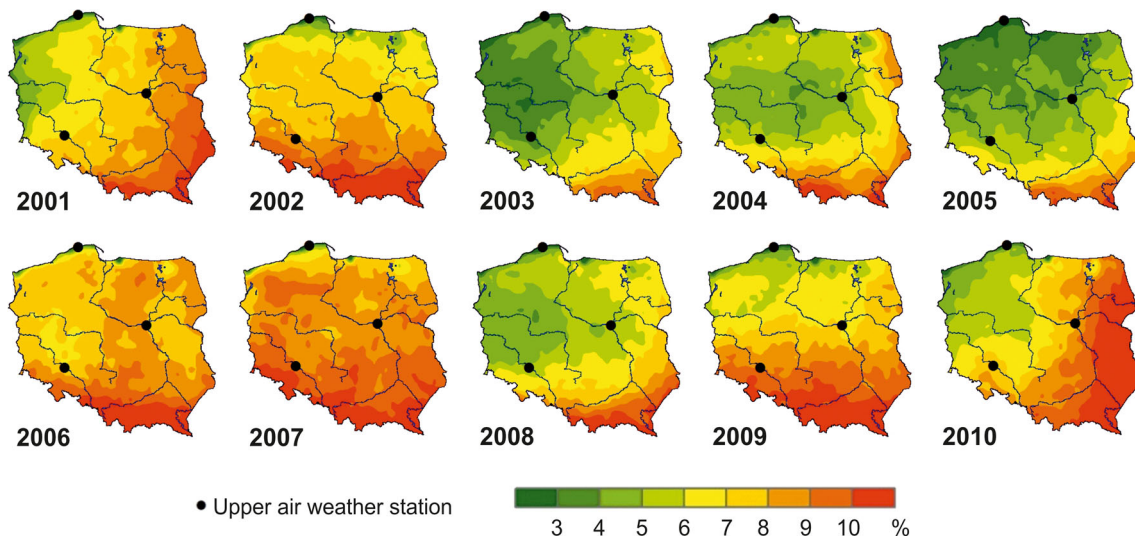
**Fig. 9** Spatially interpolated Pearson's correlation coefficients for the RegCM-derived data series and observational data from 44 weather stations (long-term variability of the monthly number of days with MUCAPE  $>200 \text{ J kg}^{-1}$  or thunderstorm observation, 1966–2010)

and observational data (statistically significant results for all studied sites, PCC median = 0.54). The data for the rest of the studied months as well as for the thunderstorm season were also characterized by a statistically significant relationship for the majority of the weather stations ( $>80\%$ , PCC median about 0.40). Slightly worse (but sufficiently valuable still) results were obtained for the comparison of data series for the whole year, as PCC was statistically significant in 72.1% of cases (PCC median = 0.38).

The research results clearly show that RegCM data on conditions favoring thunderstorms can provide a valuable input into various climatologic analyses. First, as continuous gridded data, they can become the basis for climate maps. Thanks to its geospatial format the dataset can be easily processed for very many iterations. Figure 10 shows the example of climate maps of the annual frequency of days with

**Table 2** Long-term (1966–2010) data series correlation results: A—percentage (%) of weather stations with statistically significant PCC values at  $p = 0.05$  significance level; B—median PCC value; C—maximum PCC value

Time period	A (% of stations) $p = 0.05$ (PCC $> 0.294$ )	B median PCC	C maximum PCC
May	84.1	0.47	0.71
June	100.0	0.54	0.68
July	88.6	0.51	0.70
August	86.4	0.44	0.76
May to August	81.8	0.39	0.67
Year	72.1	0.38	0.72



**Fig. 10** Annual frequency maps of days with  $\text{MUCAPE} > 200 \text{ J kg}^{-1}$ : examples generated for the 2001–2010 period (RegCM dataset, 1966–2010)

$\text{MUCAPE} > 200 \text{ J kg}^{-1}$  for the first decade of the twenty-first century. Undoubtedly, one should be concerned that the obtained gridded data cannot be interpreted as a direct quantitative representation of the occurrence of thunderstorms. Every map shows the annual frequency of days with conditions that favor convection; hence, they most often exceed the number of observed thunderstorm occurrences. However, spatial and temporal patterns are maintained.

#### 4 Discussion

Knowledge on spatial and temporal patterns of extreme phenomena is a key factor in risk analysis on various levels. Thunderstorms are one of the most evident examples of such phenomena. Therefore, the analysis of thunderstorms' climatologic aspects at various spatial and temporal scales and the study of long-term variability are exceptionally important. Yet, multiple difficulties emerge during their analysis, as thunderstorms are an example of phenomena that occur locally (small-scale phenomena). Climate maps based only on the results of visual observation performed at weather stations usually contain errors that result from the interpolation of spatially discontinuous data. Therefore, there is a need for additional variables that explain the distribution of thunderstorm phenomena over time and in spatial terms. Today, the development of modeling techniques and easy access to data from reanalysis allows one to obtain spatially continuous information on indices that help characterize atmospheric instability for a vertical profile; such indices are usually available for data from aerological soundings done only at a few stations in the country. One of the available predictors is MUCAPE (Most Unstable CAPE) instability index. In the study of the 45-year period (1966–2010), the index appears to be a very good predictor for both the spatial and temporal distribution in an annual, seasonal, and, for some months, monthly

occurrences of thunderstorms in Poland. Thanks to numerical modeling and GIS (geographic information systems) techniques acquired data can serve as the basis for complex analyses both on a national and regional scale in accordance with gridded climatology methods.

The purpose of the conducted analysis was to test the relevance of data obtained using the Regional Climate Model in order to determine conditions that favor the occurrence of days with thunderstorms. The possible use of gridded data (derived from the RegCM model) in mean annual, seasonal, and monthly analysis was also studied. Additionally, long-term variability and trend analysis were investigated. Above all, it can be stated that modeled data can be successfully used for expanding the base of information provided by very limited aerological data (only a few stations in Poland) and it can be a valuable source of data for studies on CAPE climatology for the territory of Poland.

Furthermore, most of the cases of the relationship between modeled and observational data that were studied show satisfactory Pearson's correlation coefficient values (statistically significant at least at  $p = 0.05$ ). Only analyses for months other than those of the thunderstorm season did not show the usefulness of data derived from RegCM. This may have been caused by an exceptionally low frequency as well as randomness in thunderstorm occurrences in this time period in Poland, especially when the whole data set is taken into account. It may be argued that modeled data reflect the seasonal diversity of thunderstorm occurrences particularly well—with a minimum in December and maximum in June and July.

It may be concluded that the method used in this study, which consists of the selection of favorable convective ambient conditions for describing thunderstorms' spatial and temporal occurrence patterns, is a good extension from observation data, especially in research studies on a regional scale and spatial analyses based on gridded climatology. It can be particularly helpful for analysis in areas with few observational

points or when such points do not represent the diversity of the whole area. Spatially continuous gridded data allow also to use them as environmental variables in order to explain the spatial distribution of thunderstorms in Poland, and this can serve as a basis for interpolation in regression kriging methods (Hengl 2007), especially on the scale of a year or the months in the thunderstorm season. However, it needs to be remembered that modeled data are averaged and are not always sufficient to fully reflect local diversity of the geographic environment (e.g., most of currently used models still do not possess accurate information on relief and land cover).

The results suggest also the possibility of using modeled data to describe long-term variability and to conduct trend analysis of thunderstorm occurrence in Poland, which is crucial in recent climate change research.

In summary, the selection of the MUCAPE index as a measure of the potential number of days with thunderstorm occurrence turned out to be helpful and satisfactory.

As far as certain limitations of the analysis are concerned, the validation was conducted via the point to point approach. The use of buffer zones around weather stations may probably improve the results, as not all of the thunderstorms in the given area are observed at a weather station.

Nevertheless, the paper serves as a baseline for further research and reveals possible improvements, including additional predictors of convective phenomena and improved climate models.

**Acknowledgments** Presented results were partly (RegCM downscaling) prepared thanks to the ISOK project (IT system for country protection against extreme hazards; POIG.07.01.00-00-025/09) which was financed by the European Fund of Regional Development under the Operational Programme Innovative Economy and the Polish Government.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Adams DK, Souza EP (2009) CAPE and convective events in the southwest during the North American monsoon. *Mon Weather Rev* 137: 83–98
- Allen JT, Karoly DJ, Mills GA (2011) A severe thunderstorm climatology for Australia and associated thunderstorm environments. *Aust Meteorol and Oceanogr J* 61:143–158
- Bengtsson L, Hagemann S, Hodges KI (2004) Can climate trends be calculated from reanalysis data? *J Geophys Res* 109:D11111
- Bengtsson L, Arkin P, Berrisford P, Bougeault P, Folland CK, Gordon C, Haines K, Hodges KI, Jones P, Källberg P, Rayner N, Simmons AJ, Stammer D, Thorne PW, Uppala SM, Vose RS (2007) The need for a dynamical climate reanalysis. *Bull Am Meteorol Soc* 88:495–501
- Bielec-Bąkowska Z (2003) Long-term variability of thunderstorms occurrence in Poland in the 20th century. *Atmos Res* 67-68:35–52
- Bielec-Bąkowska Z (2013) Thunderstorms and hails in Poland. *Prace Geograficzne* 132:99–132 in Polish; summary in English
- Brooks HE, Lee JW, Craven JP (2003) The spatial distribution of severe thunderstorm and tornado environments from global reanalysis data. *Atmos Res* 67-68:73–94
- Brooks HE, Anderson AR, Riemann K, Ebberts I, Flachs H (2007) Climatological aspects of convective parameters from the NCAR/NCEP reanalysis. *Atmos Res* 83:294–305
- Brooks HE (2013) Severe thunderstorms and climate change. *Atmos Res* 123:129–138
- Cavazos T, Hewitson BC (2005) Performance of NCEP-NCAR reanalysis variables in statistical downscaling of daily precipitation. *Clim Res* 28:95–107
- Cecil DJ, Buechler DE, Blakeslee RJ (2014) Gridded lightning climatology from TRMM-LIS and OTD: dataset description. *Atmos Res* 135-136:404–414
- Craven JP, Brooks HE (2004) Baseline climatology of sounding derived parameters associated with deep moist convection. *Nat Weather Dig* 28:13–24
- Dee DP, Källén E, Simmons AJ, Haimberger L (2011) Comments on “Reanalyses suitable for characterizing long-term trends”. *Bull Amer Meteor Soc* 92:65–70
- DeRubertis D (2006) Recent trends in four common stability indices derived from U.S. Radiosonde Observations. *J of Clim* 19:309–323
- Dettinger MD (2013) Projections and downscaling of 21st century temperatures, precipitation, radiative fluxes and winds for the Southwestern US, with focus on Lake Tahoe. *Clim Chang* 116: 17–33
- Elguindi N, Bi X, Giorgi F, Nagarajan B, Pal J, Solmon F, Rauscher S, Zakey A, Giuliani G (2011) Climatic model RegCM user manual, Version 4.1. Trieste, Italy
- Gensini VA, Ashley WS (2011) Climatology of potentially severe convective environments from the North American regional reanalysis. *Electron J of Severe Storms Meteorol* 6:1–40
- Gensini VA, Mote TL, Brooks HE (2014a) Severe-thunderstorm reanalysis environments and collocated radiosonde observations. *J of Appl Met and Clim* 53:742–751
- Gensini VA, Ramseyer C, Mote TL (2014b) Future convective environments using NARCCAP. *Int J Climatol* 34:1699–1705
- Giorgi F, Anyah RO (2012) The road towards RegCM4. *Clim Res* 52:3–6
- Goody R, Anderson J, Karl T, Miller RB, North G, Simpson J, Stephens G, Washington W (2002) Why monitor the climate? *Bull Am Meteorol Soc* 83:873–878
- Hahn CJ, Warren SG (2007) A gridded climatology of clouds over land (1971–96) and ocean (1954–97) from surface observations worldwide. *Numeric Data Product NDP-026E*. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee
- Haklander AJ, Van Delden A (2003) Thunderstorm predictors and their forecast skill for the Netherlands. *Atmos Res* 67-68:273–299
- Hengl T (2007) A practical guide to geostatistical mapping of environmental variables. Office for Official Publications of the European Communities, Luxembourg
- Holley DM, Dorling SR, Steele CJ, Earl N (2014) A climatology of convective available potential energy in Great Britain. *Int J Climatol* 34(14):3811–3824. doi:10.1002/joc.3976
- IPCC (2001) Climate change 2001: the scientific basis. In: Houghton JT, Ding Y, Griggs DJ, Noguer M, van der Linden PJ, Dai X, Maskell K, Johnson CA (eds) Contribution of the working group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA
- IPCC (2013) Climate change 2013: the physical science basis. In: Stocker TF, Qin D, Plattner GK, Tignor M, Allen SK, Boschung J, Nauels A,



- Xia Y, Bex V, Midgley PM (eds) Contribution of working group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA
- Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woollen J, Zhu Y, Chelliah M, Ebisuzaki W, Higgins W, Janowiak J, Mo KC, Ropelewski C, Wang J, Leetmaa A, Reynolds R, Jenne R, Joseph D (1996) The NCEP/NCAR 40-year reanalysis project. *Bull Amer Meteor Soc* 77:437–470
- Kistler R, Kalnay E, Collins W, Saha G, White G, Woollen J, Chelliah M, Ebisuzaki W, Kanamitsu M, Kousky V, van den Dool H, Jenne R, Fiorino M (2001) The NCEP–NCAR 50-year reanalysis: monthly means CD-ROM and documentation. *Bull Amer Meteor Soc* 82: 247–267
- Kolendowicz L (2004) Thunderstorm regions in Poland. *Geogr Polonica* 77:89–95
- Kunz M, Sander J, Kottmeier C (2009) Recent trends of thunderstorm and hailstorm frequency and their relation to atmospheric characteristics in southwest Germany. *Int J Climatol* 29:2283–2297
- Lombardo KA, Colle BA (2010) The spatial and temporal distribution of organized convective structures over the northeast and their ambient conditions. *Mon Weather Rev* 138:4456–4474
- Malinowska M (2011) Variability of chosen instability indices in Poland in XXI century. *Prace i Studia Geograficzne* 47:97–107
- Manzato A (2005) The use of sounding-derived indices for a neural network short-term thunderstorm forecast. *Weather and Forecast* 20:896–917
- Marsh PT, Brooks HE, Karoly DJ (2007) Assessment of the severe weather environment in North America simulated by a global climate model. *Atmos Sci Lett* 8:100–106
- Marsh PT, Brooks HE, Karoly DJ (2009) Preliminary investigation into the severe thunderstorm environment of Europe simulated by the Community Climate System Model 3. *Atmos Res* 93:607–618
- Moncrieff MW, Miller MJ (1976) The dynamics and simulation of tropical cumulonimbus and squall lines. *Q J R Meteorol Soc* 120:373–394
- Perčec Tadić M (2010) Gridded Croatian climatology for 1961–1990. *Theor Appl Climatol* 102:87–103
- Rasmussen EN, Blanchard DO (1998) A baseline climatology of sounding-derived supercell and tornado forecast parameters. *Weather and forecast* 13:1148–1164
- Riemann-Campe K, Fraedrich K, Lunkeit F (2011) Global climatology of convective available potential energy (CAPE) and convective inhibition (CIN) in ERA-40 reanalysis. *Atmos Res* 93:534–545
- Romero R, Gayà M, Doswell CA (2007) European climatology of severe convective storm environmental parameters: a test for significant tornado events. *Atmos Res* 83:389–404
- Schneider RS, Dean AR (2008) A comprehensive 5-year severe storm environment climatology for the continental United States. Preprints, 24th Conf. Severe Local Storms, Savannah GA
- Siedlecki M (2009) Selected instability indices in Europe. *Theor Appl Climatol* 96:85–94
- Thorne PW, Vose RS (2010) Reanalyses suitable for characterizing long-term trends: are they really achievable? *Bull Amer Meteor Soc* 91: 353–361
- Trapp RJ, Diffenbaugh NS, Brooks HE, Baldwin ME, Robinson ED, Pal JS (2007) Changes in severe thunderstorm environment frequency during the 21st century caused by anthropogenically enhanced global radiative forcing. *P Natl Acad Sci USA* 104:19719–19723
- Trapp RJ, Robinson ED, Baldwin ME, Diffenbaugh NS, Schwedler BRJ (2010) Regional climate of hazardous convective weather through high-resolution downscaling. *Clim Dynam* 37:677–688
- Trenberth KE, Karl TR, Spence TW (2002) The need for a systems approach to climate observations. *Bull Am Meteorol Soc* 83:873–878
- Tveito OE, Førland E, Alexandersson H, Drebs A, Jónsson T, Tuomenvirta H, Vaarby-Laursen E (2000) Nordic climate maps. KLIMA Report. 06/01, DNMI, Oslo
- Tveito OE, Bjordal I, Skjelvag AO, Aune B (2005) A GIS-based agro-ecological decision system based on gridded climatology. *Meteorol Appl* 12:57–68
- University of Wyoming (2014): Atmospheric soundings. <http://weather.uwyo.edu/upperair/sounding.html>. Accessed 11 August 2014
- Ustrnul Z (2001) Gridded data and atmospheric circulation. *Rocznik Fizycznogeograficzny* 6:7–12 in Polish, summary in English
- Ustrnul Z, Czekerda D (2006) Spatial-temporal methods in the climatological research (based on the example of Poland). *Roczniki Geomatyki* 4:147–160
- Wilby RL, Wigley TML (1997) Downscaling general circulation model output: a review of methods and limitations. *Progr in Phys Geogr* 21:530–548
- Ye B, Del Genio AD, Lo KKW (1998) CAPE variations in the current climate and in a climate change. *J of Clim* 11:1997–2015
- Yoshimura K, Kanamitsu M (2008) Dynamical global downscaling of global reanalysis. *Mon Weather Rev* 136:2983–2998