

Judit Bartholy\*, Rita Pongrácz, Zsófia Molnár  
Eötvös Loránd University, Budapest, Hungary

## 1. INTRODUCTION

According to the IPCC Third Assessment Report (2001) regional impacts of global climate change tend to intensify extremes and contribute to more frequent occurrence of the extreme climatological events. In order to provide regional climate scenarios for any particular area past climate tendencies and climatological extremes must be analyzed. Since regular observations being reliable cover only a century or two, other historical information from documentary sources must be taken into consideration for the earlier periods.

In historical documents, diaries, annals, farm and estate records, logs, account rolls, letters, newspapers or other written reports, unusual and extreme events are mentioned and possibly recorded more or less objectively. These historical sources can be used to evaluate the occurrence, duration and geographical location of extreme climatic events of the past centuries when no or only scarce instrumental time series are available (Bradley and Jones, 1992). Antal Réthly (1879-1975), the Hungarian meteorologist, professor and director of the former National Meteorological and Earth Magnetism Institute did a pioneer work in documentary research on the meteorological extremes occurred in Central/Eastern Europe. He collected the historical documents (in their original form) related to meteorology into a four-volume-long book (around 2500 pages altogether), titled "Meteorological events and natural disasters in Hungary" (Réthly, 1962; 1970; Réthly and Simon, 1999). In order to facilitate the detailed analysis of these documentary sources a special code system using hierarchical subclasses has been defined and applied to the approximately 15000 collected climatological items. The applied code system distinguishes three main categories of climate information: temperature, precipitation and wind related events, containing about 3800, 10000, and 1300 information items, respectively. Furthermore, the three level subclass system involves 10 second-level classes and 61 third-level classes. Besides the event classification the coded database contains full geographical information about the location of the meteorological events, e.g., the code for the settlement itself, the geographical coordinates, and the identification number of the subregion.

Spatial and temporal distribution of precipitation, temperature, and wind related climate events and extremes has been evaluated using settlement and subregional scales. Also, geographical distribution of

extreme climatological events has been mapped. Annual and seasonal time series have been analyzed for the Carpathian Basin, and compared to other reconstructed temperature and precipitation index time series from other geographical locations in Central/Eastern Europe. Different methodology was applied to the Réthly collection by Racz (1999). He mainly focused on the different time scales (monthly, seasonal, and annual) from a historical point of view, and especially, on the Little Ice Age period.

## 2. CODE SYSTEM AND DIGITALIZATION

Our research goal has included the digitalization of the documents containing meteorological information, as well, as the temporal and spatial evaluation of the events collected by Réthly (Bartholy et al., 2001; Pongracz and Bartholy, 2002). In order to accomplish detailed analysis of the Réthly documentary collection, a special hierarchical code system has been defined and applied. During this digitalization process altogether 14159 climate-related sources have been analyzed. Since one document may contain information on several types of climatological parameter, the total number of climate related items from the Réthly collection and the classified data are not equal.

The applied code system distinguishes three main categories of climate information: temperature, precipitation and wind related events, containing 3820, 10046, 1321 information items, respectively. The second and the third levels of the classification describe further the event type assigning specific code to it as shown in Fig. 1. presenting the frequency distribution of the events associated with the specific codes of climate information. In case of temperature cold conditions dominate, while in case of wind most of the archive records mention the strength. Precipitation information takes 66% of the total collection and the most often reported event is the 'rain' (not the extreme categories, i.e., heavy rain, extended heavy rain, rainstorm), which can be explained by the source types (many estate records and account rolls) and by the agricultural importance of water. Other frequent classes of precipitation include thunderstorm, hail, flood, and drought. The largest portion of the historical database relates to precipitation events, which are classified into 32 groups, e.g., flood, drought, snowstorm, heavy rain, thunderstorm, hail, fog, wet year, etc. Then, temperature related information is classified into 14 groups, e.g., cool summer, severe

\* Corresponding author address: Judit Bartholy, Eötvös Loránd Univ., Dept. of Meteorology, Pf. 32. Budapest, Hungary, H-1518; e-mail: [bari@ludens.elte.hu](mailto:bari@ludens.elte.hu)

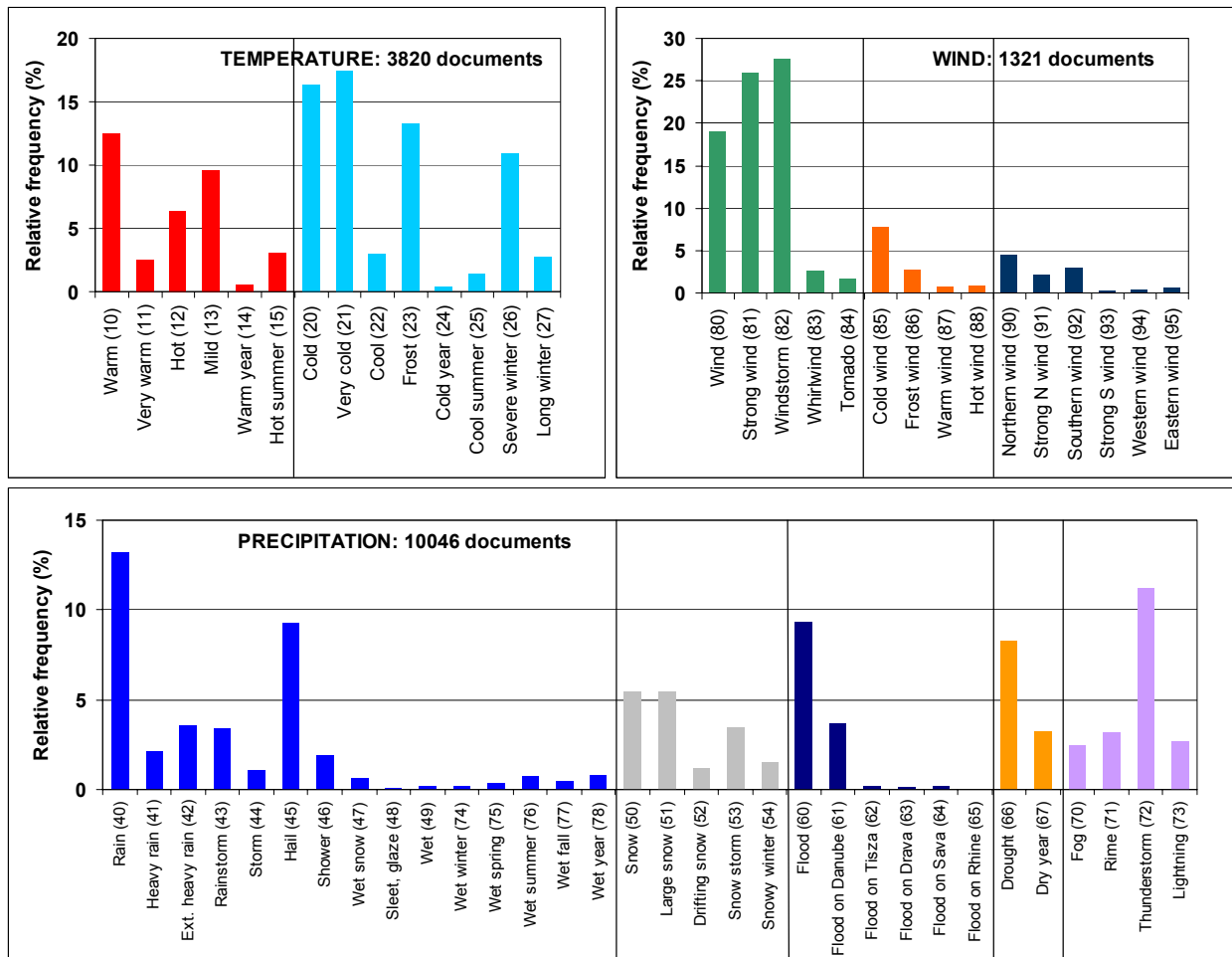


Fig. 1. Distribution of temperature, precipitation, and wind related documentary sources in the Réthly collection (corresponding codes are indicated in parentheses).

frost period, extreme cold winter, long winter period, cold year, warm year. Finally, wind information is classified into 15 groups, e.g., hot wind, cold wind, frost wind, strong wind, windstorm, tornado. Since documentary sources are based on personal impression of historical events and therefore they are subjective. In general, these reports relate to unusual and extreme meteorological phenomena. Some parts of the Réthly-data consist of agricultural yields (mostly cereal and vine) that refer to climate information, e.g., poor and sour grapevine implies cold and wet climate conditions.

Besides the event classification the coded database contains full geographical information about the location of the meteorological events, e.g., the code for the settlement itself, the geographical coordinates, and the identification number of the subregion. The subregions, their area, and the associated codes are presented in Fig. 2 for the five main regions of the Carpathian Basin. Documentary

sources sometimes provide information on very large areas, which have been also coded (some examples are listed in the last group of the figure). During the past centuries dramatic historical changes have occurred in the Carpathian region resulting in wars, human migration waves, border modification, regime changes. Therefore, geographical identification of some settlements caused severe problems since very often the names changed several times and the oldest names are forgotten, also, some small rural settlements depopulated long time ago and disappeared. However, usually the subregion could be determined based on the context of the chronicles, so these historical documentary items have been recorded in the final database, too.

Based on the classification described above the coded datasets have been analyzed separately for the main meteorological parameters (Pongracz and Bartholy, 2002; Bartholy et al., 2003). Fig. 3. presents the spatial distribution of temperature, precipitation and

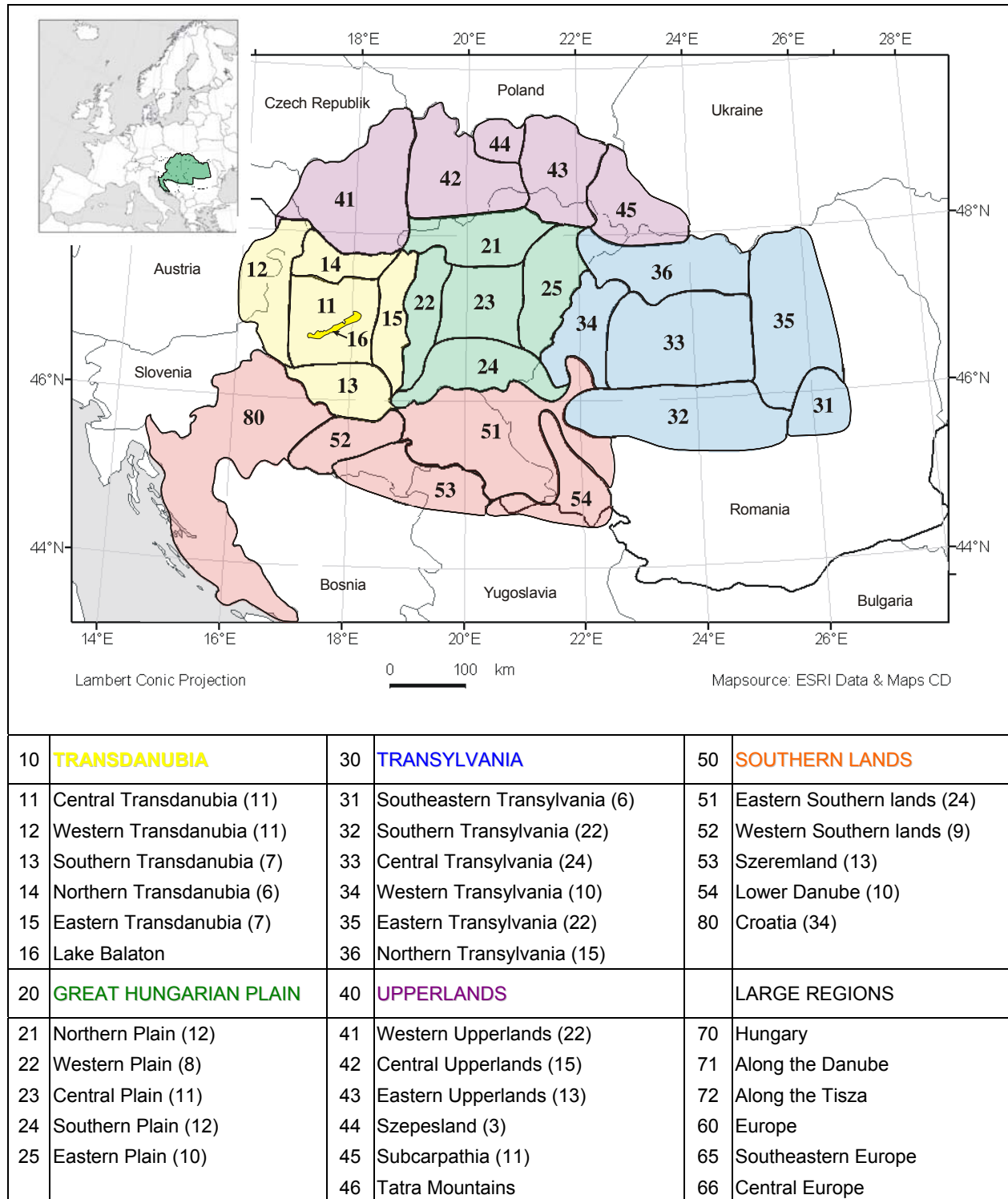


Fig. 2. Subregions of the Carpathian Basin applied to the analysis of the Réthly collection  
(After the name of the subregion the area is indicated in  $10^3 \text{ km}^2$  units.)

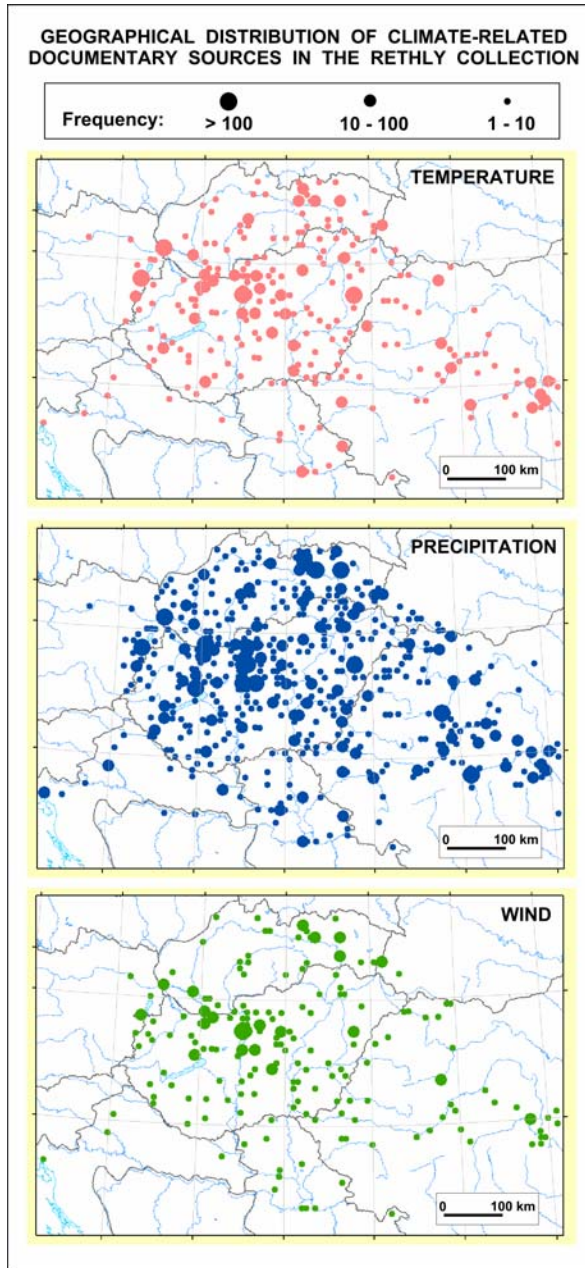


Fig. 3. Geographical frequency distribution of temperature, precipitation, and wind related documentary sources in the Réthly collection (number of documents is indicated by the size of the circles).

wind related data (upper, middle, and lower panel of the figure, respectively). Each geographical location providing documentary information related to the given climate parameters appears as a circle on the maps. The size of the circle represents the total number of related records during the entire period (using three frequency categories: 1-10, 10-100, more than 100 items). These maps illustrate also that the precipitation-related reports dominate the Réthly

collection and its spatial distribution is the most representative for the Carpathian Basin.

In this paper results for temperature and precipitation (since they are the most important climatic factors) are presented in the next two sections.

### 3. TEMPERATURE RELATED SOURCES

First, spatial distribution of temperature related data is analyzed that is followed by a temporal evaluation of the time series on seasonal and annual scales.

Fig. 4. provides details on the spatial distribution of the temperature related historical documents. Similarly to the previous maps, geographical locations recording any temperature related historical information are indicated by circles. Their size corresponds to the total number of temperature records, while their division indicates the ratio of the reported cold and warm conditions (blue and red sector, respectively). In spite of the good spatial coverage of temperature related data, historical settlements with high cultural background dominate. Cold events are documented more frequently, namely, 65% of the entire temperature related database is associated to cold conditions, while 35% of the records indicate warm conditions. The same stands for the five main regions of the Carpathian Basin (ratios of the cold and warm conditions are about 63-69% and 31-37%, respectively).

Temperature related information is separated into four seasonal datasets containing 1797, 1050, 751, and 571 records for winter (DJF), spring (MAM), summer (JJA), and autumn (SON), respectively. In order to perform the temporal analysis on seasonal scale moderate and extreme conditions have been separated for the warm and cold temperature related records (corresponding codes are shown in Fig. 1). Definition of moderate and extreme warm/cold conditions depends on the season since, for instance, the same code in winter and in summer may represent moderate and extreme event, respectively (e.g., code 23 denoting frost indicate normal event in winter, while extremely unusual in summer). If absolute frequency values of moderate and extreme temperature events are displayed the earlier periods with less historical sources would be underrepresented. On the other hand, relative distribution by itself may introduce a large decrease in confidence. Thus, a reliability index has been defined based on the total number of data in a given 50 year period using the mean ( $m$ ) and the standard deviation ( $s$ ) as follows. Intermediate reliability indicates that the number of temperature related historical records is between  $(m-0.3 \cdot s)$  and  $(m+0.3 \cdot s)$ , while low and high reliability implies less than  $(m-0.3 \cdot s)$  and more than  $(m+0.3 \cdot s)$  occurrences during 50 years, respectively.

Fig. 5 presents these reliability classes and the relative distribution of seasonal extreme and moderate

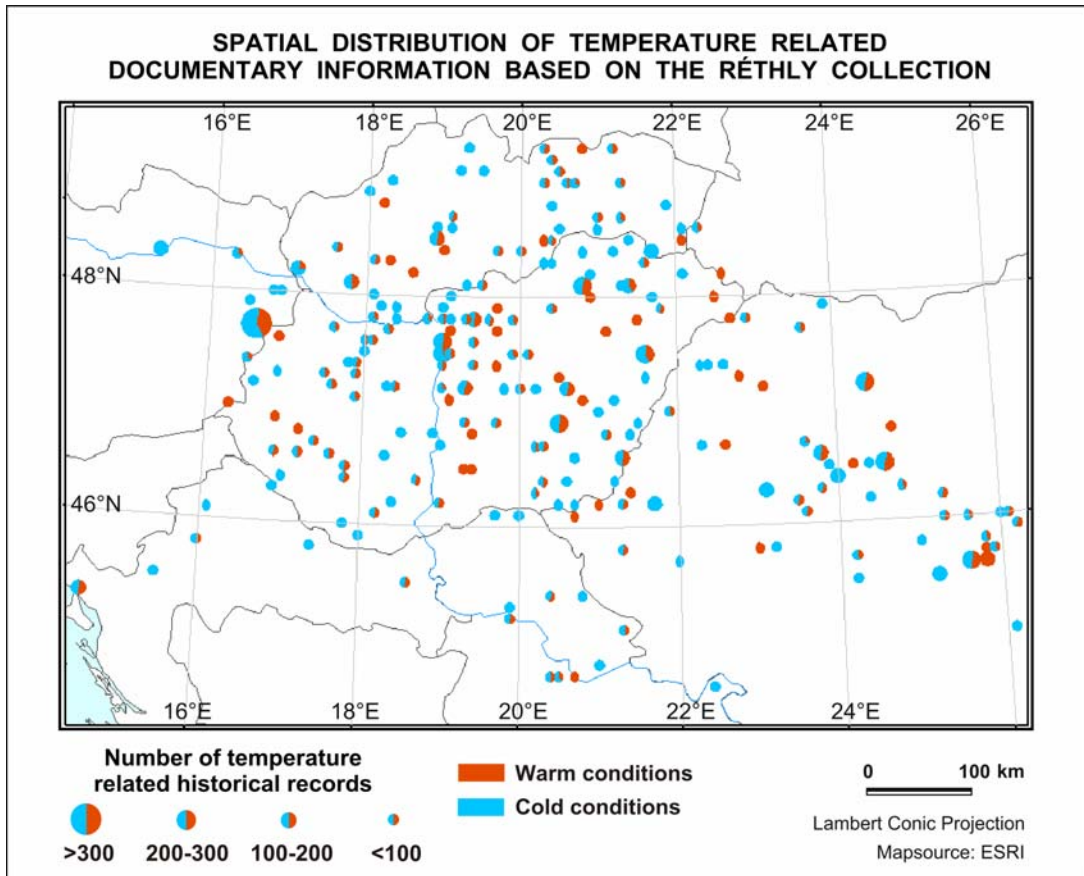


Fig. 4. Spatial distribution of temperature related documentary information. Circles indicate the number of total records and the ratio of the cold and warm conditions

temperature conditions for the 500-1900 period based on the 50 year frequency. Because of the scarce data in the earlier centuries the first two histograms consider all the 50 year periods from 500 to 1000, and from 1000 to 1350. Evidently, reliability increases with more data from the more recent periods, except the last 50 years (1850-1900) when number of written documentary records is decreasing due to the spread of regular meteorological observations instead of historical reports.

Distribution of the recorded temperature conditions highly depends on seasons. Winter is dominated by the extreme cold conditions (usually more than 50% of the total data). Spring distribution performs larger similarities than autumn, but instead of extreme colds, extreme and moderate cold conditions altogether possess the large portion of the data. In the meanwhile, summer distribution differs from the other three seasons since usually more than half of the historical documents contain information about extreme warm conditions during summer. A warmer episode between 1500 and 1600 can be detected in

each season (especially in summer). Furthermore, the recent warming trend after 1700 is the most pronounced in winter.

In order to synthesize the temperature related information from the Réthly collection a historical dataset reconstruction for the Carpathian Basin is accomplished on annual scale for the period 1650-1900 and 1200-1900, respectively. Panels A, B, and C of Fig. 6 present different stages of the reconstructed annual temperature index. The frequency of warm and cold reports are plotted on panel A, where positive values are used for warm conditions and negative values for cold conditions. The sum time series of the signed frequencies is smoothed by the 20 year moving average. Then, these raw data are adjusted according to the assigned intensity values (from -3 to +3) of the cold/warm events. The adjusted frequency data and the 20 year moving average of the summed time series are shown on panel B. Finally, the sum of the adjusted frequency is normalized by the total annual number of temperature related historical data resulting in the temperature index shown on panel C. These



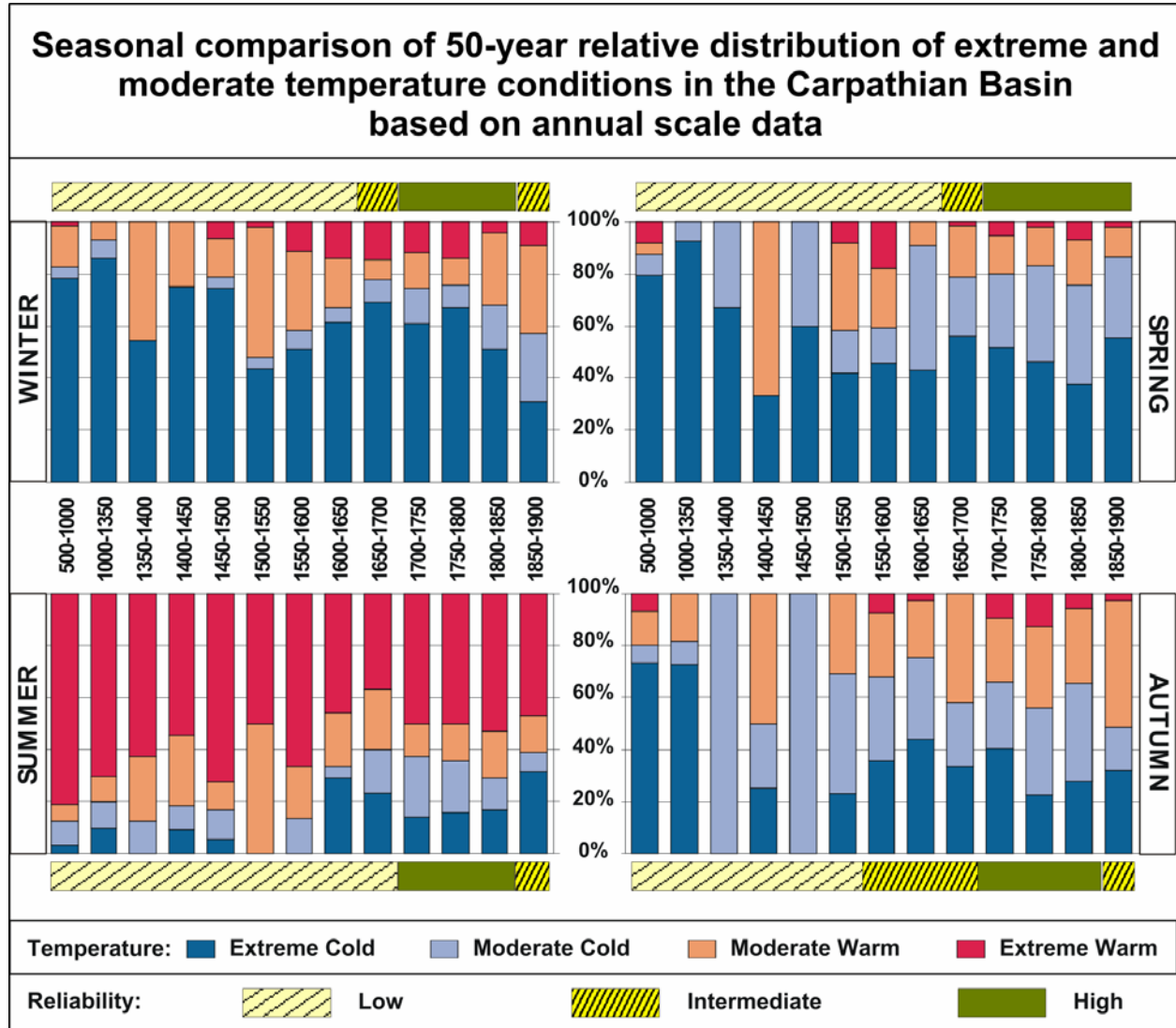


Fig. 5. Seasonal 50-year relative frequency of extreme and moderate temperature conditions in the Carpathian Basin. Reliability represents the number of total data in a given 50 year period.

reconstructed time series are compared to the graduated temperature index (GTI) for Switzerland (Pfister, 1993) shown on panel D. Originally, Pfister et al. (1994) calculated monthly GTI values (ranging between -3 and +3) that are averaged for the year. GTI is calibrated with proxy data and instrumental observations, the reference period is 1901-1960.

Single year anomalies of the above mentioned index time series usually are not coincident in case of the Carpathian Basin and Switzerland because of their considerable distance. However, warm and cold periods (on the 20 year moving average scale) perform

good agreement, especially, cold periods of 1800-10 and 1840-50, and warm periods of 1820-30 and 1865-75. Furthermore, the warming trend of the 1700s may be considered as part of the climate recovery period in the Carpathian Basin at the end of the Little Ice Age (Lamb, 1969; 1995; Luterbacher, 2001), while in Switzerland two warm episodes can be recognized instead

Further comparisons to other reconstructed data from the region are planned. Also, a temporal analysis of time series separated for sensitive subregions of the Réthly collection is needed.

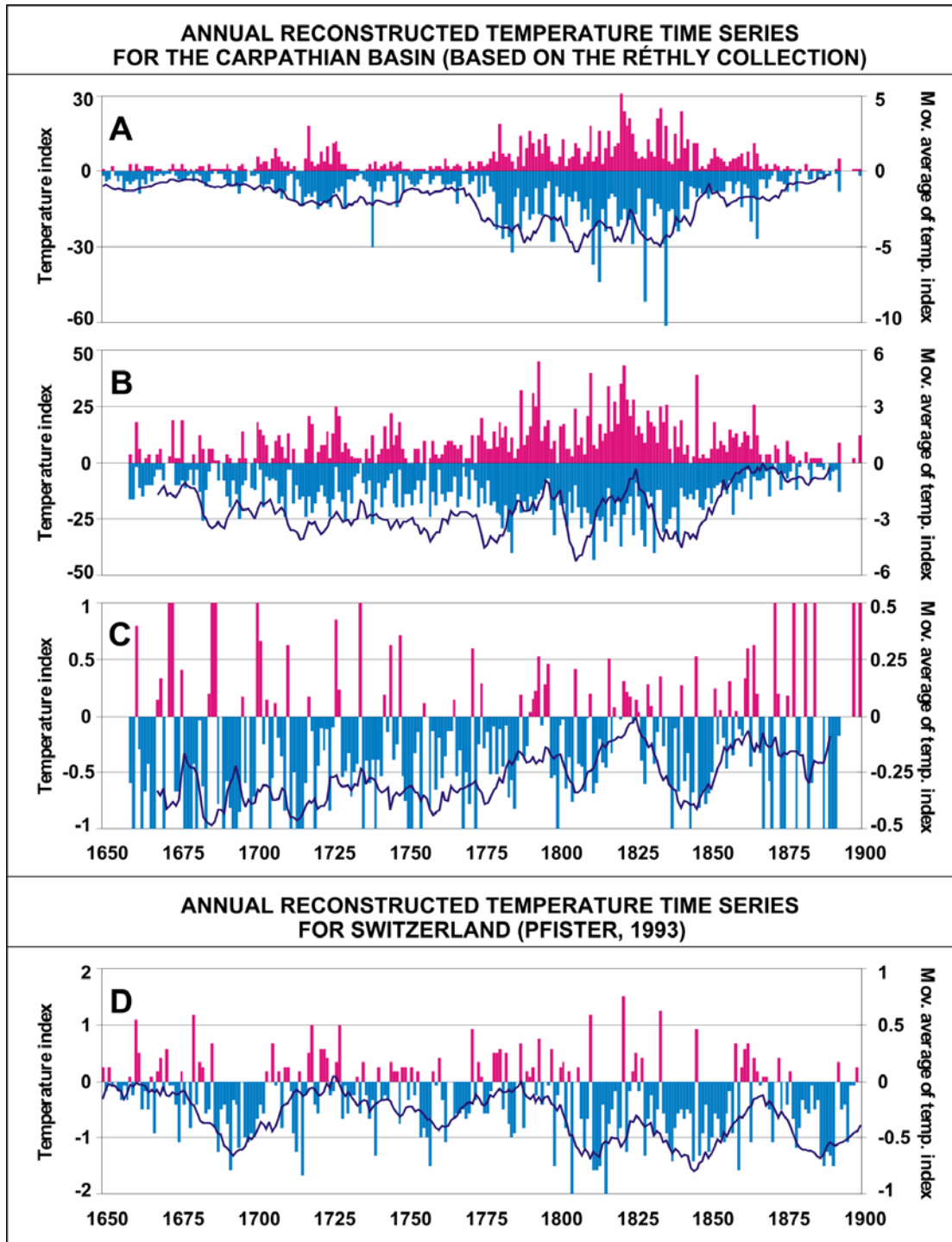


Fig. 6. Historical annual temperature indices and their 20-year moving averages for the 1650-1900 period.  
A: Frequency values of documents from the Réthly collection reporting warm and cold temperature events from the Carpathian Basin. B: Adjusted temperature index using intensity values for the recorded warm/cold events.  
C: Normalised temperature index taking into account the total annual number of temperature related historical data.  
D: Graduated temperature index for Switzerland (Pfister, 1993), see more detailed description in the text.

#### 4. PRECIPITATION RELATED SOURCES

Compared to temperature related data, almost three times more historical records associated with precipitation are available from the Réthly collection. This implies more geographical locations (about twice as many as the number of settlements where temperature related information was recorded); therefore in case of precipitation the spatial analysis presented in this paper is based on the subregion scale instead of the settlements themselves (Bartholy et al., 2003). First, seasonal time series of 50-year relative frequency distribution of the main precipitation related code groups are presented (together with reliability information). Then, some special events and code composites of precipitation documents are

spatially analyzed using subregional frequency factors.

After seasonal separation of the dataset, the following analysis is accomplished for winter, spring, summer, and autumn, as shown in Fig. 7. The graphs summarize all the precipitation related reports (except the group of 'others': fog, rime, thunderstorm, and lightning) on a seasonal scale using the 50 year relative frequency values. The 28 specialized codes form four main groups of precipitation related events, namely, droughts, river floods, snow events, and wet conditions (details of the corresponding codes are listed in Fig. 1). Similarly to the temperature related analysis, the first two histograms consider all the 50 year periods from 500 to 1000, and from 1000 to 1350 because of the small numbers of data

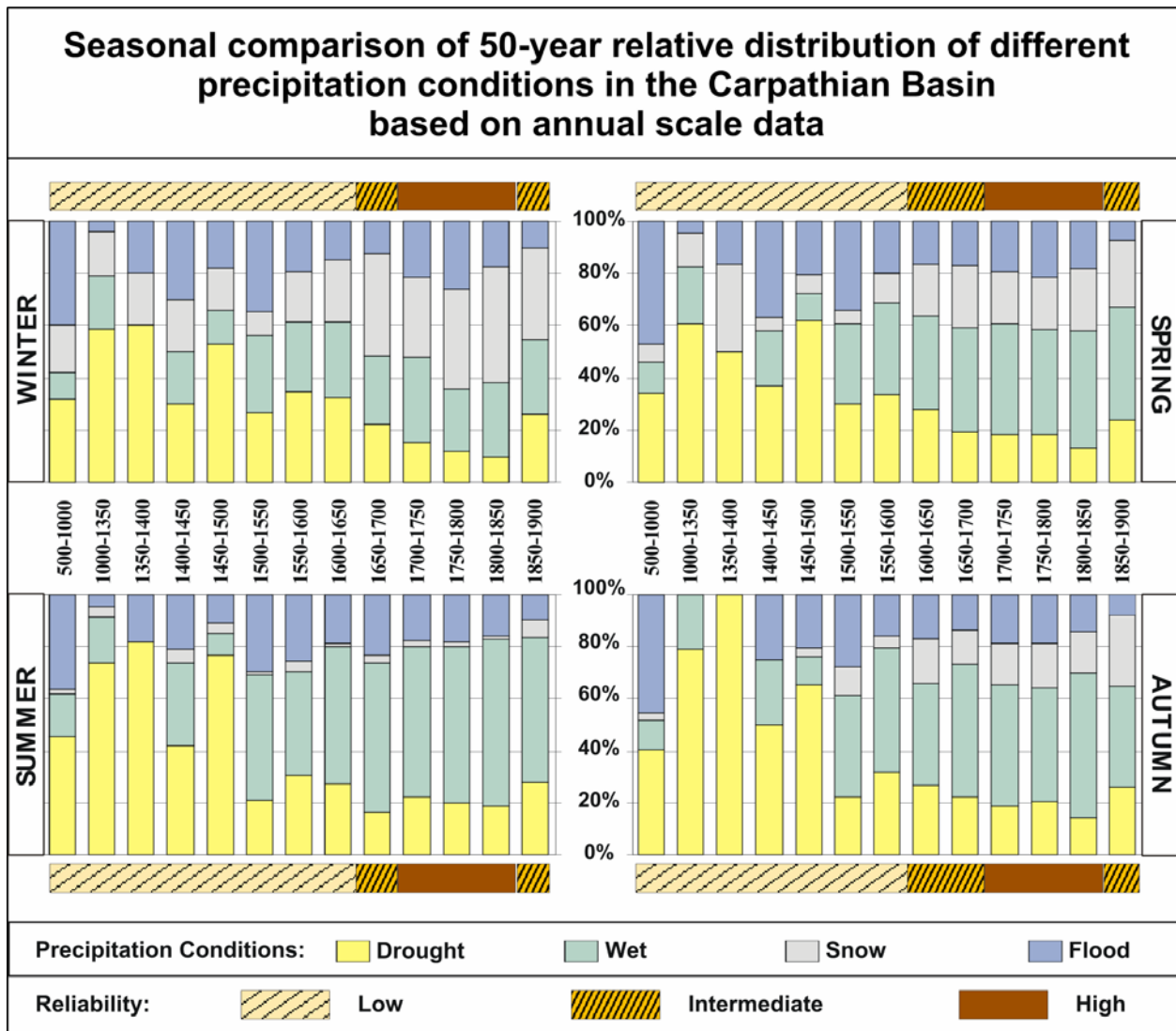


Fig. 7. Seasonal 50-year relative frequency of extreme and moderate precipitation conditions in the Carpathian Basin. Reliability represents the number of total data in a given 50 year period.



Until the 1500s, extreme events (i.e., flood and drought) dominate the precipitation information, while afterwards their contribution decreases from 69% to 33% on average. Evidently, special seasonal structures appear because of the significant seasonality of snow presence. On the one hand, this may be explained by the disastrous floods and droughts, which were often accompanied by famine, food shortage, and epidemics during the early centuries. These consequences affected the past societies so much that they were recorded in several historical sources together with the meteorological events, while more common snow events and wet conditions are less important from a chronicle point of view. On the other hand, the possible reasoning must consider the low reliability index values in the early centuries.

Drought events dominate the historical sources in two 50 year periods, 1350-1400 and 1450-1500, when most of the country wide droughts were recorded throughout the two millennia. Fig. 7 suggests that these extremes occurred mainly in summer and autumn. After 1500, the decreasing tendency in the relative occurrence of droughts is the most pronounced in winter, while it almost disappears in summer.

In order to analyze the spatial distribution of precipitation related event occurrences, a subregional event frequency factor ( $SFF_k$ ) is defined using the ratio of the numbers of documents and the spatial extension of subregions as follows,

$$SFF_k(i) = X_k(i) / t(i),$$

where  $X_k(i)$  is the number of event  $k$  ( $k$  denotes the event code shown in Fig. 1) records from the settlements located in subregion  $i$ , and  $t(i)$  is the area of subregion  $i$  (presented in Fig. 2). The quintiles of the  $SFF_k$  values determine five classes of spatial distribution: "very rarely", "sometimes", "several times", "frequently", "very often". Because of the  $SFF$  definition, subregions with large uninhabited areas may be underrepresented in the analysis, but high elevated spots are rarely found in Eastern Central Europe (only the High Tatra Mountains, and some parts of the Eastern Carpathian Mountains). Based on this process, maps for each precipitation codes, and also, for thematic selection of codes have been generated. In this paper only some of them are presented, namely, maps for hail ( $SFF_{45}$ ) and thunderstorm ( $SFF_{72}$ ) in Fig. 8, and furthermore, maps for drought ( $SFF_{66,67}$ ), rain ( $SFF_{40}$ ), severe snow ( $SFF_{51,52,53}$ ), heavy rain and storm ( $SFF_{41,42,43,44}$ ) in Fig. 9. In these figures the subregions are colored according to the  $SFF_k$  classification, while relative frequency values indicate the occurrences of event  $k$  in the subregions relative to the total number of the event  $k$  records from the Carpathian Basin.

The upper panel of Fig. 8 presents the annual distribution of two special events (hail and thunderstorm), which may cause severe damages on agricultural production. Monthly relative frequencies of hail are provided for the 17<sup>th</sup>, 18<sup>th</sup>, and 19<sup>th</sup> centuries in the three pie-charts of the left side, while the same is shown for thunderstorm on the right side. Before the 1600s, the number of historical sources is too sparse; therefore this temporal analysis is based on the three above mentioned centuries. Both events occur mostly in late spring and in summer when convective processes dominate the weather in this region. Specifically, 83-88% of the hails are reported in May, June, July and August, while 61-67% of the thunderstorms are recorded in these four months. However, considerable shifts can be detected from one century to another. Historical records on hail in June increased from 24% in the 1600s to 32% in the 1800s, while in August the relative frequency decreased from 18% to 10% during the same period. Similarly, in case of thunderstorm, the relative occurrence in June increased from 14% to 19%, while it decreased in July from 23% to 15% between the 17<sup>th</sup> and 19<sup>th</sup> century.

The spatial analysis presented in the lower panel of Fig. 8 is based on the entire database (the first document containing information on thunderstorm is from 173 AD in the Western Upperlands subregion, while on hail it is from July 1423 in the Western Transdanubia, city Sopron). Besides the spatial distribution of the historical data,  $SFF$  classification also provides area representativity of the event occurrences. Hails and thunderstorms occurred "frequently" and "very often" in Transdanubia and the Great Hungarian Plain, which may be explained by the high agricultural sensitivity of these regions (agricultural sources recorded these events often, while other type of historical documents did not pay so much attention to them since their effect is less important on other human activities).

Furthermore, four spatial distribution maps are presented in Fig. 9 providing frequency information on code composites for the entire database. In order to represent different general precipitation related events, extreme and severe climatological codes have been aggregated. Drought conditions (codes 66 and 67) are frequently recorded in the Great Hungarian Plain area and other hilly subregions where agricultural activity may be important. Frequent severe snow conditions are mainly associated with mountains and hilly subregions, e.g., Southeastern Transylvania, Western Upperlands where the Lower Tatra Mountains are located, and the entire Transdanubia. On the right side of Fig. 9 analysis of rains and extreme rain/storm conditions indicates that hilly subregions in Transdanubia and the Upperlands and also parts of the Great Hungarian Plain are dominated by large  $SFF$  values.

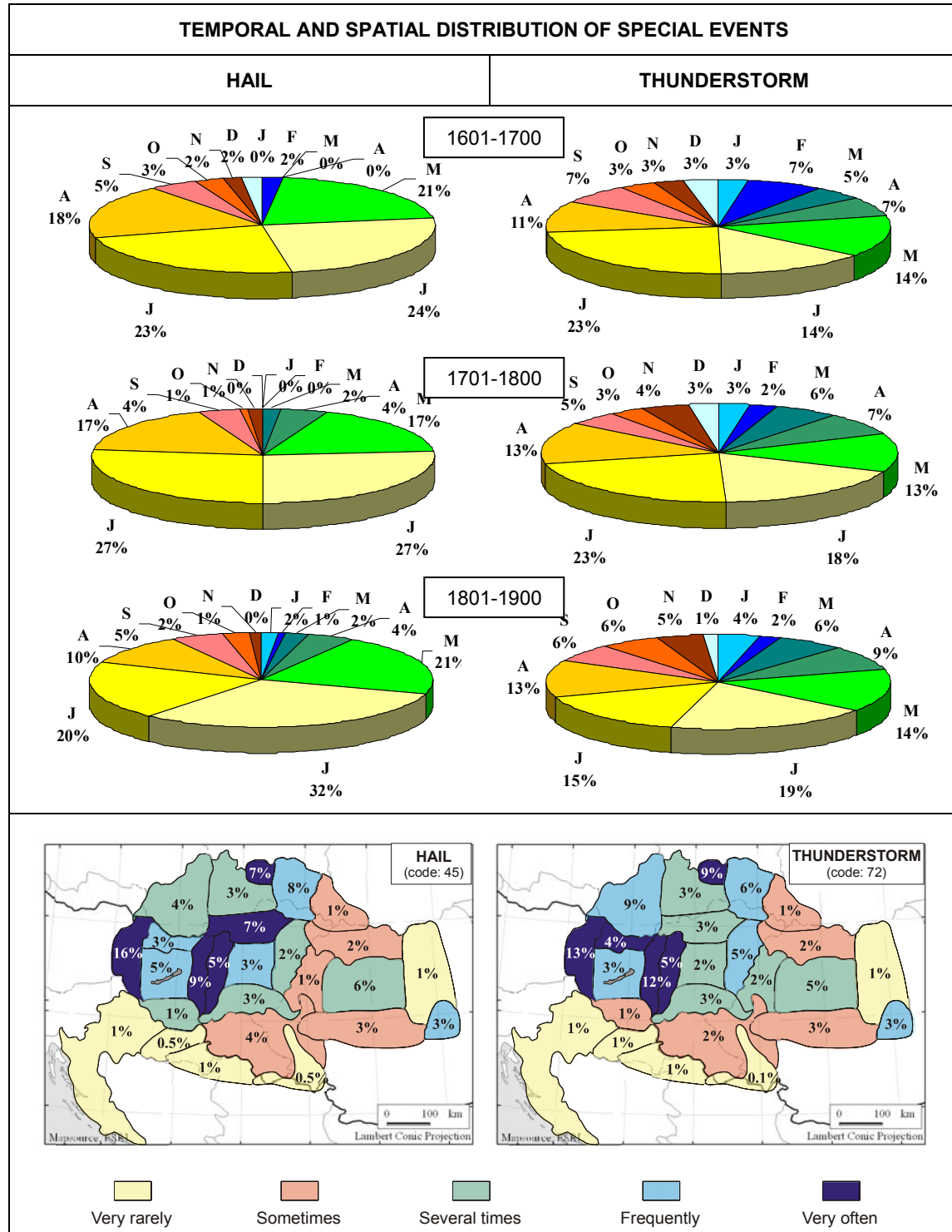


Fig. 8. Annual distribution of hails and thunderstorms in the 17th, 18th, and 19th centuries (upper panel, monthly relative frequencies), and their spatial distribution for the entire period (lower panel, subregions are colored according to the five classes of  $SSF_{45}$  and  $SSF_{72}$ , while frequency numbers of the maps indicate the event occurrences)

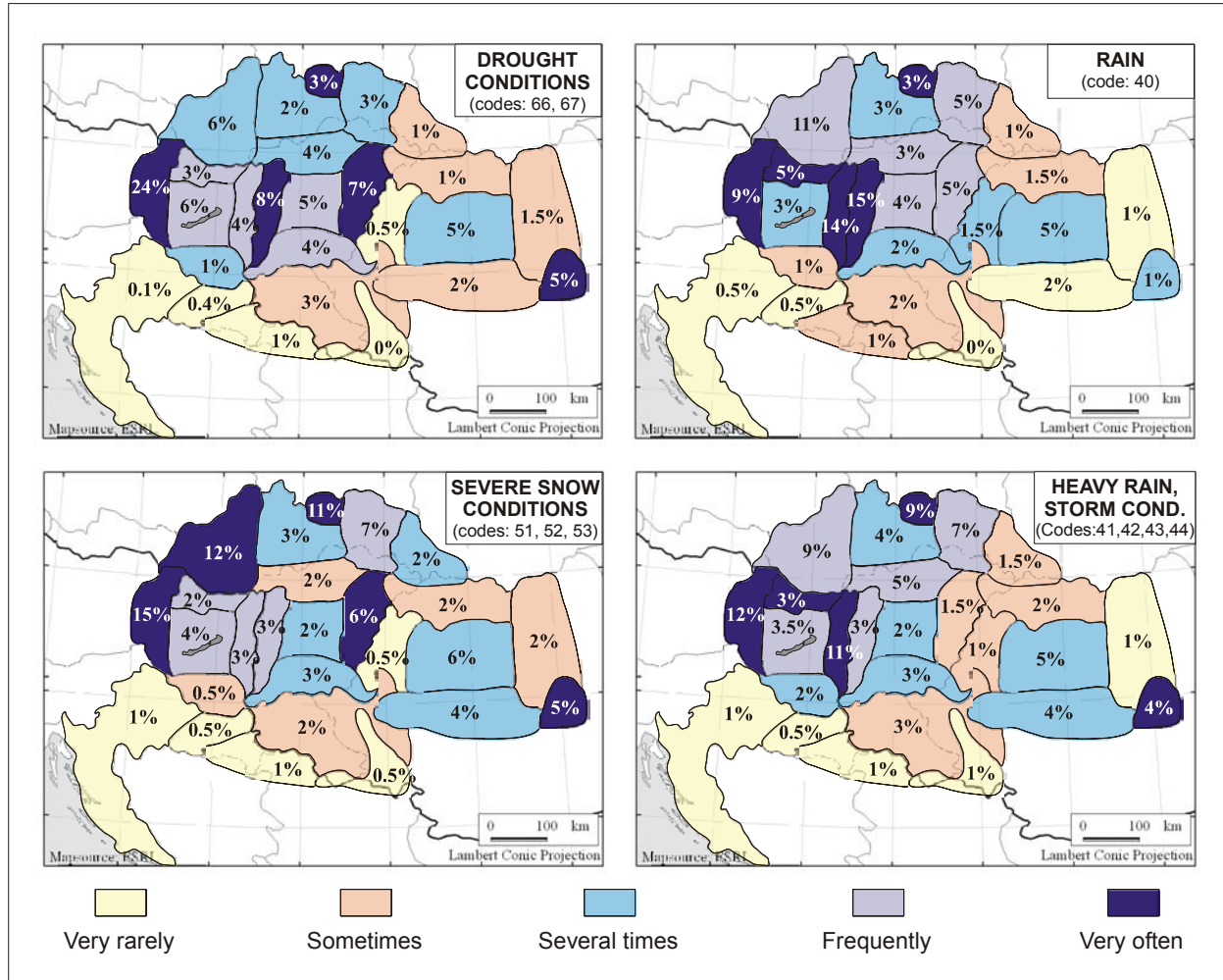


Fig. 9. Spatial distribution of precipitation related information based on thematic selection. Subregions are colored according to the five classes of  $SSF_{66,67}$ ,  $SSF_{40}$ ,  $SSF_{51,52,53}$ , and  $SSF_{41,42,43,44}$ , while frequency numbers of the maps indicate the event occurrences

## 5. SUMMARY AND CONCLUSION

The first steps towards the extensive evaluation of climatological events occurred in the Carpathian Basin during the past two millennia was made by Réthly who collected and published all the available climate related historical documents in his book series (Réthly, 1962; 1970; Réthly and Simon, 1999). In order to facilitate further climatological research on this enormous archive collection, it was essential to digitalize as much from the written information as possible. The applied code system presented in this paper served as a basic tool to analyze temperature and precipitation related time series reconstructions, as well, as extreme climatological episodes reported in historical sources. Based on the results presented in this paper the following conclusions can be drawn.

1. The applied code system classifies the 14159 historical records from the Réthly collection into a three level hierarchical class system involving 3 main

categories, 10 second-level classes, and 61 third-level classes. The main advantage of this three level classification is that characteristics of complex phenomena or several parameters may be analysed simultaneously.

2. Spatial information is compiled from reports on scales of settlements and subregions by using 28 geographical subregions defined according to the geographical characteristics and the historical background.

3. In case of temperature for the entire Carpathian Basin 65% and 35% of the reports refer to cold and warm conditions, respectively. The stands for the five main regions of the Carpathian Basin (ratios of the cold and warm conditions are about 63-69% and 31-37%, respectively).

4. A reliability index has been defined that was used to represent the total number of archive data in a given period.

5. A warm episode between 1500 and 1600, and a warming trend after 1700 are detected on the seasonal scale analysis of 50 year distribution of warm and cold conditions.

6. After the 1500s, floods, snow and wet conditions are reported in 81-89% of the precipitation related documents; which imply wetter climatic conditions during these centuries until 1900 than in the previous periods. The decreasing tendency in the relative occurrence of droughts is the most pronounced in winter, while it almost disappears in summer.

7. In order to analyze the spatial distribution of precipitation related event occurrences, a subregional event frequency factor (SFF) is defined using the ratio of the numbers of documents and the spatial extension of subregions. Large SFF values for the precipitation code composites dominate in the subregions with agricultural activity. This can be explained by the high sensitivity of these subregions.

8. Results of the temporal analysis of two special events suggest that historical records on hails in June increased from 24% in the 1600s to 32% in the 1800s, while in August the relative frequency decreased from 18% to 10% during the same period. Similarly, in case of thunderstorms, the relative occurrence in June increased from 14% to 19%, while it decreased in July from 23% to 15% between the 17<sup>th</sup> and 19<sup>th</sup> century.

**Acknowledgements.** Research leading to this paper has been supported by the *Hungarian National Science Research Foundation* under grants T-026629, T-034867, and T-038423, also by the AEROCARB and CHIOTTO projects of the *European Union* Nr. 5 program under grants EVK2-CT-1999/0013, EVK2-CT-2002/0163, and the *Hungarian National Research Development Program* under grant NKFP-3A/0006/2002. ESRI software has been used to create maps. Furthermore, supports of the Bolyai Janos Research Fellowship of the Hungarian Academy of Sciences is appreciated.

## REFERENCES

- Bartholy, J., Pongracz, R., Molnar, Zs., 2001: Extreme climatic events occurred in the last millennium in the Carpathian Basin based on Réthly's documentary research. *Geophysical Research Abstracts*, **3**, GRA3 5586, 9146.
- Bartholy, J., Pongracz, R., Molnar, Zs., 2003: Millennial trends, extremes in the Carpathian Basin based on the Réthly documentary collection. *Geophysical Research Abstracts*, **5**, 12896.
- Bradley, R.S., Jones, P.D., 1992: *Climate since A.D. 1500*. Routledge, London.
- IPCC. 2001: *Climate Change 2001: Third Assessment Report. The Scientific Basis*. Cambridge University Press, Cambridge.
- Lamb, H.H., 1969: *Climatic fluctuations. World Survey of Climatology, Vol.2.* (ed: Flohn, H.), Elsevier, Amsterdam, 173-249.
- Lamb, H.H., 1995: *Climate, History and the Modern World*. Routledge, London.
- Luterbacher, J., 2001: The Late Maunder Minimum (1675-1715) – climax of the "Little Ice Age" in Europe. In: *History and Climate: Memories of the Future?* (eds: Jones, P.D., Ogilvie, A.E.J., Davies, T.D., Briffa, K.R.), Kluwer, New York. 29-54.
- Pfister, C., 1993: *Historical Weather Indices from Switzerland. IGBP PAGES/World Data Center-A for Paleoclimatology Data Contribution Series # 93-027*. NOAA/NGDC Paleoclimatology Program, Boulder CO, USA.
- Pfister, C., Kington, J., Kleinlogel, G., Schuele, H., Siffert, E., 1994: High resolution spatio-temporal reconstructions of past climate from direct meteorological observations and proxy data. In: *Climatic Trends and Anomalies in Europe 1675-1715.* (eds: Frenzel, B., Pfister, C., Glaeser, B.), Fischer, Stuttgart, 329-376.
- Pongracz, R., Bartholy, J., 2002.: Study of extreme climatological events occurred in the Carpathian Basin during the last millennium. *Geophysical Research Abstracts*, **4**, OA26-1FR5P-122.
- Racz, L., 1999: *Climate History of Hungary Since 16th Century: Past Present and Future*. Centre for Regional Studies of Hungarian Academy of Sciences, Pécs.
- Réthly, A., 1962: *Meteorological Events and Natural Disasters in Hungary until 1700.* (in Hungarian) Academic Press, Budapest.
- Réthly, A., 1970: *Meteorological Events and Natural Disasters in Hungary between 1701-1800.* (in Hungarian) Academic Press, Budapest.
- Réthly, A., Simon, A., 1999: *Meteorological Events and Natural Disasters in Hungary between 1801-1900. Vol.I-II.* (in Hungarian) Hungarian Meteorol. Service, Budapest.