# USE OF CLUSTER ANALYSIS FOR ASSASSMENT OF RAINFALL REGIME

*M. KOLÁŘ, T. LITSCHMANN, J. SKLENÁŘOVÁ*

Department of Geography, Masaryk University, Kotlářská 2,
611 37 Brno, Czech Republic

Received for publication: September 1997

## SUMMARY

The contribution deals with the issue of climatological regionalisation on the basis of the study of the precipitation zone in the wide surroundings of the Nové Mlýny Basin in South Moravia. After compilation of correlation matrixes covering the relationships between annual rainfall totals in selected stations, and formation of correlation chains, potential differences between rainfall modes of individual areas of the investigated territory were defined, using cluster analysis aided with the non-hierarchical procedure of "typical representatives". The gained regionalisation is relatively stable even when data from different time intervals are used. Another significant fact is that the conclusions resulting from the mathematical and statistical analyses correspond to their logical, geographically accountable determination.

## KEY WORDS

Cluster analysis – Rainfall – South Moravia

## INTRODUCTION

The limestone cliffs of the Pavlovské vrchy (Mts.), quite suddenly protruding from the Dyjsko-svratecký úval (Graben), represent a strong orographic dominant of the Břeclav district. The highest point of the hill range, Děvín Mt., 554.4 m above sea level, stands out over 350 m above the surrounding meadows. This sudden obstacle in a comparatively flat landscape is therefore quite likely to be deforming the domains of individual meteorological elements, in both the close and the wider surroundings. While thermal conditions of the Pavlovské vrchy (Mts.) have been dealt with for example by Prošek (1978), a study of their effect on the precipitation domain has not yet been published. Regarding the fact that the nearest relevant synoptic station is only as far as in Malacky, there is no wonder that potential anomalies of the precipitation regimes, most likely caused by orographic effects, are little known by our synoptic forecast centres, although the inhabitants of certain regions of the Břeclav district know them as a common phenomenon. In most cases they are represented by an absence of rainfall in the particular areas, under certain climatic conditions, while there is rain over the rest of the territory. This fact is not to be neglected and should be analysed from both climatological and synoptic viewpoints.

The present study aims at primary processing of the available climatic data from the wide surroundings of the Pavlovské vrchy (Mts.), for the purpose of defining the zones with different precipitation regimes, potentially including definitions of the differences between the individual regimes, observed in the area. With a view to the fact that the authors only had monthly totals of precipitation data at hand, the analysis cannot be as deep as the suggested problem would require, also regarding the potential use of the analysis for more accurate weather forecast for this area.

## THE LOCATION OF THE STUDIED TERRITORY

The area under investigation is the landscape within the frontiers defined by the line drawn between Retz (Austria), Znojmo and Džbánice, and between Židlochovice, Těšany and Klobouky in the north. The eastern frontier is identical with the line drawn between Dubňany, Prušánky and Břeclav, and the frontier to the south is marked with the towns of Valtice, Laa a. d. T. and Jaroslavice.

From the orographical point of view the greatest proportion of this area belongs to the province of the Western Carpathians, going into two systems. The middle part of the territory, a belt going from north-east to south-west, belongs to the Outer Carpathian Depression, here represented by the depression of the Dyjskosvratecký úval (Graben) valley. The area east of this belt (roughly from Mikulov to Klobouky), belongs to the system of the Outer Western Carpathians, with their edges represented by the Mikulovská vrchovina (Highland) to the south and the Ždánický les (Highland) and the Kyjovská vrchovina to the north-east. The depression along the Dyje (the "Věstonice Gate"), and the small south-eastern section of the investigated area, are part of another system, the West-Panonian Basin, consisting of the Dolnomoravský úval (Graben) in the system of the Viennese Basin. The western part of the territory lies in the province of the Bohemian Highland and in the geomorphological units of the Jevišovická vrchovina (Upland), the Boskovická brázda (Furrow) and the Bobravská vrchovina (Upland). Over 50% of the area is represented by a flat landscape relief, the rest is a typical highland relief, especially the Ždánický les Mts. and the Kyjovská vrchovina (Highland), i.e. east of the limestone cliffs of the Pavlovské vrchy (Mts.).

## DATA AND METHODOLOGY

To achieve the aim of the study, data of monthly precipitation totals from 1961-1994, provided by 22 stations situated in the investigated area and administered by the Czech Hydrometeorological Institute and the Austrian Central Institute of Meteorology and Geodynamics, have been used. The basic data are shown in Table 1.

| No. | Station | Latitude [° '] N | Longitude [° '] E | Altitude [m] |
|---|---|---|---|---|
| 1. | Božice | 48 50 | 16 17 | 216 |
| 2. | Branišovice | 48 58 | 16 26 | 200 |
| 3. | Břeclav | 48 46 | 16 54 | 152 |
| 4. | Bulhary | 48 50 | 16 45 | 165 |
| 5. | Drnholec | 48 51 | 16 29 | 185 |
| 6. | Dubňany | 48 55 | 17 04 | 175 |
| 7. | Džbánice | 49 00 | 16 13 | 342 |
| 8. | Hustopeče | 48 57 | 16 44 | 193 |
| 9. | Jaroslavice | 48 46 | 16 14 | 189 |
| 10. | Klobouky | 49 00 | 16 52 | 248 |
| 11. | Laa a. d. T. | 48 43 | 16 23 | 187 |
| 12. | Lednice | 48 48 | 16 48 | 164 |
| 13. | Miroslav | 48 57 | 16 19 | 270 |
| 14. | Velké Pavlovice | 48 54 | 16 50 | 215 |
| 15. | Pohořelice | 48 59 | 16 31 | 184 |
| 16. | Prušánky | 48 50 | 16 59 | 185 |
| 17. | Retz | 48 46 | 15 57 | 242 |
| 18. | Těšany | 49 03 | 16 46 | 203 |
| 19. | Valtice | 48 45 | 16 46 | 205 |
| 20. | Dolní Věstonice | 48 53 | 16 39 | 172 |
| 21. | Židlochovice | 49 02 | 16 37 | 185 |
| 22. | Znojmo | 48 52 | 16 03 | 306 |

**Table 1.** List of used stations

Most of the stations have provided a complete series of observations, in the case of any measurements being absent these have been calculated with the help of the multiple linear regression (Brázdil, Litschmann, 1984) based on data from some of the surrounding stations. The greatest number of measurement absences have been found with the Dolní Věstonice station, therefore the relations of this station to the others have to be taken with the greatest care.

The question of climatological regionalization has been studied quite profoundly in domestic as well as world literature, for the question of determination of natural climatic regions is relevant for the solutions of numerous practical problems connected either with the weather forecast or the needs of agriculture or many others. The advance of computer technology, including the relevant software, provides the climatologists with ever more powerful and more effective means of solutions of various problems needing extensive data files processing.

The problem similar to that dealt with in this article has been solved by Sumner et al. (1993), for the region of Mallorca, or by Puvaneswaran and Smithson (1993) for the region of Srí Lanka. Although the regions might seem very remote, at first sight, there is an element unifying them all, and that is the very occurrence of orographical obstacles, whose effect is even stronger in the island areas. For illustration: in the case of Mallorca the principal orographical obstacle is the Serra de Tramontana mountain range with the highest mountain of Puig Major (1443 m above sea level), with Srí Lanka the highest peak of Pidurutalagala reaches 2524 m above sea level. In both cases the natural precipitation zones were defined with the help of cluster analysis, aided, in the case of Mallorca, with the correlation chains method, which, at the beginning of the project solution, helped to identify the relations between individual rainfall regimes of the particular stations. Cluster analysis has also been used by Fovell and Fovell (1993) for the United States.

In the course of the very solution we have only used tested procedures which had yielded reliable results before in similar problems solutions. The first step was to construct a correlation matrix of the links between annual precipitation totals for individual stations. Assuming that the sets of annual rainfall totals need not always show normal distribution, we have calculated the correlation coefficients by means of a non-parametric procedure according to Sperman. The results are summed up in Table 2.

It is obvious from this table that the links between the rainfall totals of individual stations can be either very loose (there is an especially loose link between Hustopeče and Retz of Austria, $r = 0.33$), or very tight (the value of $r = 0.91$ was achieved between Pohořelice and Branišovice and between Pohořelice and Drnholec). Table 2 shows in bold letters pairs of stations whose mutual correlation coefficient has reached the value of 0.85 or higher. These correlations are presented graphically in Fig. 1, where individual stations are connected by their correlation chains to provide a spatial image of the relationships between the rainfall totals of individual stations. The greatest number of close correlations with the neighbouring stations can be found with the stations of Pohořelice, Branišovice, Velké Pavlovice and Lednice. As stated above, for the Dolní Věstonice station a great proportion of the data was calculated on the basis of data provided by the neighbouring stations, and therefore the relatively great number of close correlations cannot be considered objective.

Another step towards finding potential differences between the precipitation regimes of individual parts of the studied area can be the use of cluster analysis. As stated by Gong and Richman (1995), the use of this methodology in climatology has been more and more frequent in the course of the last decade, although the results have largely been affected

Station No.

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1.0 | **.89** | .64 | .74 | .83 | .55 | **.85** | .63 | .82 | .68 | .80 | .74 | **.90** | .72 | **.86** | .63 | .41 | .60 | .82 | .76 | .76 | .76 |
| 2 | | 1.0 | .66 | .76 | **.87** | .70 | .72 | .70 | **.88** | .76 | **.87** | .81 | **.85** | .75 | **.91** | .71 | .52 | .69 | .80 | .77 | .82 | .71 |
| 3 | | | 1.0 | .66 | .69 | .55 | .45 | .63 | .61 | .62 | .62 | .76 | .56 | .72 | .70 | .73 | .42 | .59 | .74 | .68 | .67 | .60 |
| 4 | | | | 1.0 | .81 | .66 | .66 | .74 | .71 | .76 | .84 | **.86** | .71 | .76 | .83 | .64 | .45 | .72 | .80 | .84 | .74 | .69 |
| 5 | | | | | 1.0 | .74 | .74 | .73 | .82 | .79 | .84 | **.85** | **.86** | .84 | **.91** | .77 | .49 | .77 | .75 | .83 | **.87** | .73 |
| 6 | | | | | | 1.0 | .55 | .70 | .66 | .75 | .72 | .75 | .57 | .79 | .75 | .78 | .50 | .80 | .68 | .75 | .76 | .50 |
| 7 | | | | | | | 1.0 | .55 | .73 | .60 | .69 | .66 | .84 | .64 | .76 | .54 | .39 | .60 | .75 | .72 | .68 | .73 |
| 8 | | | | | | | | 1.0 | .61 | .73 | .75 | .71 | .55 | .77 | .77 | .78 | .33 | .81 | .61 | **.88** | .77 | .44 |
| 9 | | | | | | | | | 1.0 | .64 | .72 | .75 | **.89** | .68 | **.86** | .59 | .52 | .58 | .68 | .70 | .76 | .70 |
| 10 | | | | | | | | | | 1.0 | .77 | .77 | .66 | .79 | .78 | .71 | .49 | .84 | .75 | .78 | .84 | .45 |
| 11 | | | | | | | | | | | 1.0 | **.86** | .80 | .77 | .84 | .69 | .56 | .73 | .79 | **.87** | .81 | .59 |
| 12 | | | | | | | | | | | | 1.0 | .73 | **.88** | **.88** | .82 | .53 | .77 | .84 | **.88** | **.85** | .70 |
| 13 | | | | | | | | | | | | | 1.0 | .70 | **.85** | .59 | .47 | .61 | .70 | .72 | .75 | .77 |
| 14 | | | | | | | | | | | | | | 1.0 | **.86** | **.89** | .39 | **.85** | .77 | **.90** | **.85** | .63 |
| 15 | | | | | | | | | | | | | | | 1.0 | .80 | .49 | .80 | .76 | **.85** | **.90** | .79 |
| 16 | | | | | | | | | | | | | | | | 1.0 | .45 | .83 | .68 | .80 | .80 | .60 |
| 17 | | | | | | | | | | | | | | | | | 1.0 | .41 | .41 | .38 | .48 | .43 |
| 18 | | | | | | | | | | | | | | | | | | 1.0 | .67 | **.87** | **.87** | .47 |
| 19 | | | | | | | | | | | | | | | | | | | 1.0 | .80 | .70 | .67 |
| 20 | | | | | | | | | | | | | | | | | | | | 1.0 | **.87** | .61 |
| 21 | | | | | | | | | | | | | | | | | | | | | 1.0 | .56 |
| 22 | | | | | | | | | | | | | | | | | | | | | | 1.0 |

**Table 2.** Correlation coefficients of annual rainfall totals between individual stations in the period 1961-1994

by the procedures and data selection useds. For small-scale selections (including this study) non-hierarchical procedures are recommended, involving a selection of representatives of future clusters typical in one way or another, and their co-ordination with others on the basis of input data resemblances. We have selected data files of stations with the greatest numbers of pair correlation coefficients above 0.85 as the typical representatives of the individual precipitation zones. These stations include the above-mentioned Branišovice, Velké Pavlovice, Pohořelice and Lednice. Since the stations of Branišovice and Pohořelice are close neighbours and therefore when the results of cluster analysis for four areas were processed the areas belonging to these two stations overlapped to a great extent, for further processing only the stations of Pohořelice, Velké Pavlovice and Lednice have been used.

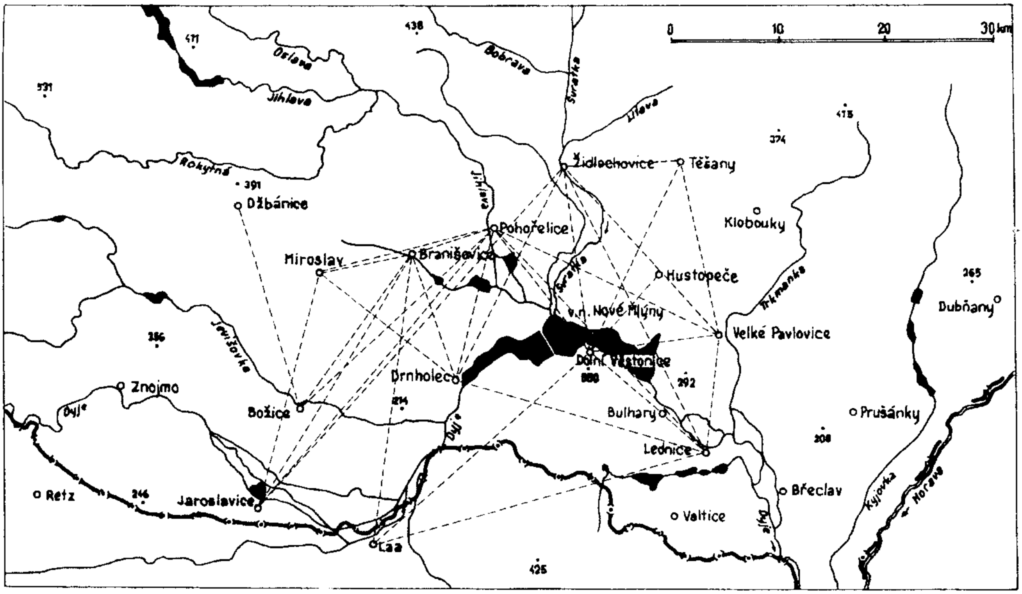


**Fig. 1.** Correlation chains for r > 0.85

## CONCLUSION

The results of cluster analysis using the non-hierarchic procedure of typical points can be assessed from Fig. 2. In accordance with our initial ideas the individual stations got sorted into clusters with clear spatial frontiers, quite corresponding with the geomorphologic division of the landscape, for while one of the clusters can be found in the Dyjskosvratecký úval (Graben) and is formed by the stations of Božice, Branišovice, Drnholec, Džbánice, Jaroslavice, Miroslav, Pohořelice, Retz, Židlochovice and Znojmo, another covers the area of the Ždánický les Mts. and the Kyjovská vrchovina (Highland) (including the stations of Dubňany, Hustopeče, Klobouky, Velké Pavlovice, Prušánky, Těšany, and Dolní Věstonice) and the third one is situated inside the Dolnomoravský úval-Graben (the stations of Břeclav, Bulhary, Laa a. d. T., Lednice and Valtice).

These areas are quite stable and there is not much difference in whether the input data used for the cluster analysis are monthly totals for a selected decade or annual totals for the whole of the three decades. The only difference is the allocation of the Břeclav station, as when the annual totals are used the station is allocated to the second of the three areas.

Note that the dividing line between the stations of the first and the third areas is formed by the Mikulovská vrchovina (Highland), with a substantial part formed by the cliffs of Pavlovské vrchy (Mts.).

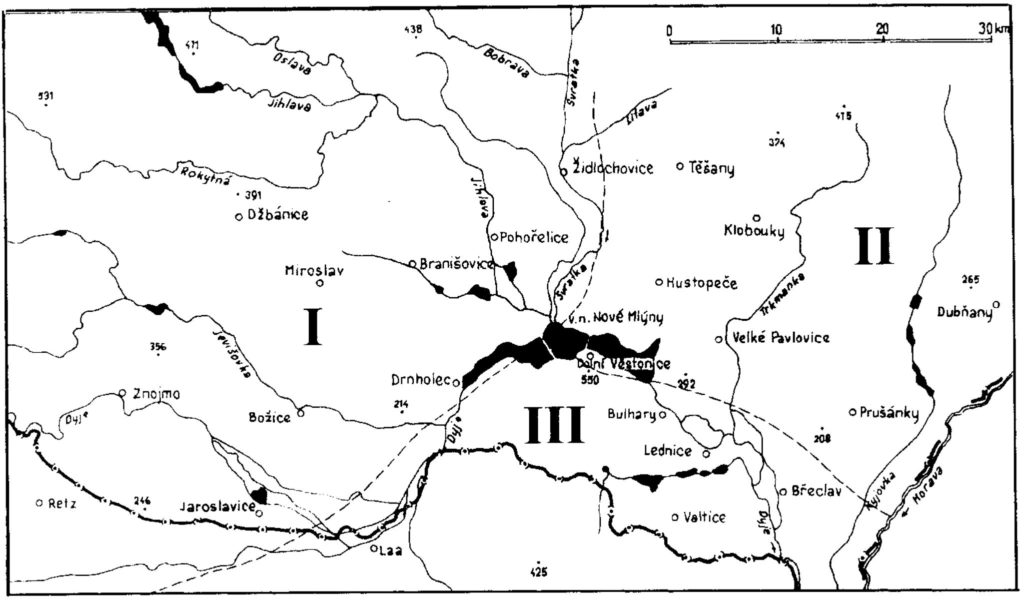


**Fig. 2.** Regions (I–III) defined on the basis of precipitation zone cluster analysis

## REFERENCES

BRÁZDIL, R., LITSCHMANN, T. (1984): Comparing different methods for completing the missing data on precipitation. Scripta Fac. Sci. Nat. Univ. Purk. Brun., 14, 6, 289-300.

GONG, X., RICHMAN, M. B. (1995): On the application of cluster analysis to growing season precipitation data in North America east of the Rockies. J. Climate, 8, 4, 897-931.

FOVELL, R. G., FOVELL, M. C. (1993): Climate zones of the conterminous United States defined using cluster analysis. J. Climate, 6, 11, 2103-2135.

JOHNSON, G. L., HANSON, C. L. (1995): Topographic and atmospheric influences on precipitation variability over a mountainous watershed. J. Appl. Meteorol., 34, 1, 68-87.

MILLS, G. F., LANA, X., SERRA, C. (1994): Catalonian precipitation patterns: Principal component analysis and automated regionalization. Theor. Appl. Climatol, 49, 201-212.

MILLS, G. F. (1995): Principal component analysis of precipitation and rainfall regionalization in Spain. Theor. Appl. Climatol, 50, 169-183.

PROŠEK, P. (1978): Vliv reliéfu na teplotní poměry Pavlovských vrchů. Univerzita J. E. Purkyně, Brno, 99 pp.

PUVANESWARAN, K. M., SMITHSON, P. A. (1993) An objective classification of homogenous rainfall regimes in Sri Lanka. Theor. Appl. Climatol, 48, 133-145.

SUMNER, G., RAMIS, C., GUIJARRO, J. A. (1993): The spatial organization of daily rainfall over Mallorca, Spain. Int. J. Climatol., 13, 1, 89-109.

SUMNER, G., RAMIS, C., GUIJARRRO, J. A. (1995): The impact of surface circulation on significant daily rainfall patterns over Mallorca. Int. J. Climatol., 15, 6, 673-696.