IDŐJÁRÁS
Quarterly Journal of the Hungarian Meteorological Service
Vol. 115, No. 1–2, January – June 2011, pp. 51–70

Time variation of the effect of geographical factors on spatial distribution of summer precipitation over the Czech Republic

Jacques Célestin Moliba Bankanza

Institute of Atmospheric Physics,
Boční II 1401, 141 31 Prague, Czech Republic; E-mail: moliba@ufa.cas.cz

(Manuscript received in final form March 26, 2011)

Abstract—This study deals with modeling of spatial distribution of summer (JJA) precipitation over the Czech Republic. The aim is to analyze the time variation of the relationships between geographical factors and precipitation during summer. Various candidates geographical predictors are evaluated in the stepwise regression models for summer precipitation, namely: (1) a set of omnidirectional parameters of the elevation that characterize an area of 3×3 km around meteorological stations, (2) various cross products calculated on the basis of geographical coordinates and elevation or topographic parameters, (3) slope and four facets of slope aspect characterizing the orographic regimes in the Czech Republic, (4) land cover parameters describing an area of 10×10 km around meteorological stations, and (5) geographical coordinates. The orographic parameters are derived from the 1 km resolution digital elevation model (DEM); the land cover parameters are derived from the 1 km resolution CORINE (COoRdination of INformation on the Environment) land cover data. Daily precipitation data for the period 1971–2003 have been used. The precipitations were collected from 203 stations throughout the country. Stepwise regression models of summer precipitation are generated for each year, and each overlapping decade from 1971 to 2003. To ensure the stability of the regression equations and comparability of regression models in time, similar suitable and stable independent variables in time should be selected. Therefore, orthogonally rotated principal component analysis (PCA) and frequency of significant predictors entering models are used to select them. Multivariate regression precipitation models are generated using definitive (PCA or stepwise based) selected predictors. Ten independent geographical variables have been selected as the most important predictors for precipitation regression models. They consist of latitude, longitude, slope aspect of the grid westward from the central grid, slope aspect of the grid northward from the central grid, slope of the grid northeastward from the central grid, slope of the grid eastward from the central grid, slope of the grid northward from the central grid, maximum value of elevation (percentile 95%) of northwestern grid from the central grid, minimum value of elevation (percentile 5%) of the central grid, and vegetation. The relationships between these significant predictors and precipitation are stable in time. No significant trend in regression coefficients has been found during 1971–2003.

Key-words: multivariate regression, precipitation models, summer precipitation, geographic influence, environmental variables, temporal variability, Czech Republic

1. Introduction

Precipitation is a climatic variable that has a high spatial and temporal variability. Its spatial distribution is influenced by various factors such as terrain height, slope, etc. For mapping purposes, it is practical to estimate the effect of such factors on precipitation distribution. Spatial modeling is a suitable tool that allows to explore the relationship between the target variable and predictors, and to get continuous information on precipitation over a targeted area. Many studies have been undertaken recently to assess and model the relationship between the climatic variables, and independent factors. Several geographical variables, including land cover (*Joly et al*., 2003), proximity to the water bodies (*Weisse* and *Bois*, 2001; *Vicente-Serrano et al*., 2003; *Marquinez et al*., 2003; *Daly et al.* 2002), atmospheric circulation (*Johnson* and *Hanson*, 1995; *Basher* and *Zheng*, 1998; *Courault* and *Monestiez*, 1999), and topography (*Johnson* and *Hanson*, 1995; *Goodale et al*., 1998; *Daly et al*., 2002) have been frequently used as relevant independent variables to model spatial patterns of precipitation. The latter has a significant influence on spatial variability of precipitation (*Joly et al*., 2003; *Weisse* and *Bois*, 2001, etc.). Therefore, numbers of these studies have been focused on modeling the influence of topographic features on the spatial variability of climate variables (*Prudhomme* and *Reed*, 1999; *Johnson* and *Hanson*, 1995; *Drogue et al*., 2002; *Weisse* and *Bois*, 2001; *Diodatto*, 2005, etc.).

According to its geographical position in Central Europe, the Czech Republic is subject to both oceanic and continental influences. Topographically, the inner part of the country is dominated by lowlands and surrounded by highlands. Such topographic feature contributes to modifying airflow over the country and can induce a strong convective precipitation, especially in the mountains (Moravsko-slezské Beskydy, Jeseníky, Krkonoš, Jizerské hory, and Krušné hory). Extreme precipitation events are more frequent and intense over these highlands due, among other factors, to the influence of exposition to airflow (*Kakos*, 2001).

In this study, the relationships between geographical factors and summer precipitation are examined through a stepwise regression model. Summer precipitation is analyzed instead of other seasons for several reasons. First, the annual cycle of precipitation in the Czech Republic is characterized by a tendency for maximum rainfall during summer. Therefore, summer precipitation contributes significantly on the character of the precipitation fluctuation (*Tolasz et al*., 2007). Second, summer precipitation, usually of shorter duration and greater intensity (*Tolasz et al*., 2007), is characterized by the high frequency of occurrence of extreme precipitation events (*Kaspar* and *Muller*, 2008), which are often connected with several natural hazards including hydrological flood and soil erosion. The need of precipitation information during summer time is crucial for the risk management. Moreover, spatial models examine spatial dependence of climate variables using a single time realization of the variable, i.e., they widely use the mean values for a given period, as input. However, the performance of a model depends not only on the density of the station network and the choice of methods, but also on the temporal variability (*Hulme et al.*, 1997). Therefore, the choice of the period of study can bias the results of interpolation (*Hulme et al.*, 1997). The spatial variability of environmental variables is commonly a result of complex processes working at the same time and over long periods of time, rather than an effect of a single realization of a single factor (*Hengl*, 2007). Geostatistics are less powerful than the statistical climatology based on sample in time, because they are based on single realization in time *(Szentimrey* and *Bihari*, 2007). The temporal variability seems to be an important task in modeling spatial variation of climate variables. This aspect has received substantial attention in several studies: *Basher* and *Zheng* (1998) take into account seasonal behavior of precipitation (ENSO variations) for mapping precipitation patterns of a data-sparse tropical south-west Pacific Ocean region. *Brown* and *Comrie* (2002) created the 39-year time series of maps and datasets of winter temperature and precipitation for the southwest US by comparing 30 years (1961–1990) modeled means with 39 observed winter temperature and precipitation values. *Johnson* and *Hanson* (1995) modeled the relative contribution of topographical and meteorological variability to regional precipitation variability. In order to improve interpolation of spatially generated weather data, *Baigorria et al.* (2007) analyzed changes in spatial correlations and compared spatial correlation on daily and monthly basis. Therefore, using seasonal rainfall amounts, temporal analysis is needed to find and determine: (1) how the relationships between independent variables and precipitation vary within years and decades; (2) how the model is affected by temporal changes. The aims of this study are: (1) to model spatial pattern of summer (JJA) precipitation in the Czech Republic at year and ten-year time steps from 1971 to 2003 using geographical variables *as independent variables;* (2) to analyze the time variation of the relationships between geographic variables, and the summer precipitation during 1971–2003.

2. Datasets

Digital Elevation Models (DEM)

DEM with the resolutions of 100 m and 1 km have been used. The fine spatial resolution of topographic and elevation variables have been retained in this study, because large-scale topographic features at a resolution of 1–15 km yield a high correlation with precipitation (*Daly et al.*, 1994; *Daly*, 2006). The 100 m DEM resolution data have been used to calculate the following smoothed elevation parameters: (1) the upper (percentile 95%) and (2) lower (percentile 5%) percentiles of elevation for a grid of 1 km resolution, and (3) mean elevation for each grid with 1×1 km resolution. On the other hand, slope, and slope aspect are obtained directly from the 1 km resolution DEM data.

Land cover data

Land cover data are obtained from the CORINE (COoRdinate INformation on the Environment) land cover dataset. These data are available in the following link: http://www.dataservice.eea.europa.eu/dataservice. They describe the land cover units in Europe. According to the CORINE land cover classification, four main types of landscape characterizing the Czech Republic were identified, and used as candidate geographical independent variables. They are related to vegetation, agricultural area, water bodies, and artificial areas.

Precipitation data

Daily precipitation data for the period 1971–2003 have been used. The dataset consists of 203 stations distributed over the whole country (*Fig. 1*). Meteorological stations are unevenly distributed across these different land cover units and topographic patterns. Most of them are distributed across urban (towns, small cities, villages) and agricultural areas. Only few stations are located in vegetation-covered areas. Considering ground elevation, about 80% of meteorological stations are located below 600 m. Only 12% of them are located above 600 m on the highlands or mountainous regions that have a significant influence on precipitation distribution (*Table 1*). The lack of observation in forested and mountainous areas shows how much it is important to model the relationships between rainfall, and elevation and/or other geographical variables.

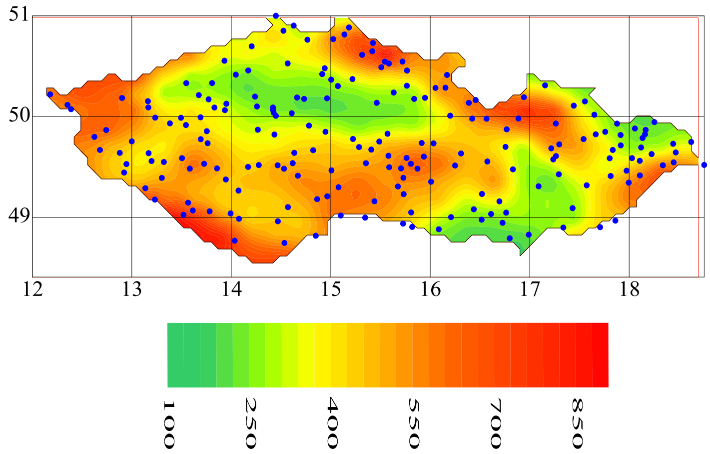


Fig. 1. Spatial distribution of meteorological stations with regard to DEM data (in m). Dots represent geographic position of meteorological stations.

Table 1. Elevation of meteorological stations

Elevation (m)	Number of stations	% of total area
< 200	12	6
200 – 400	85	42
400 – 600	82	40
> 600	24	12

In order to perform time variation analysis (see Section 4), two subsets of summer precipitation series are derived from the summer precipitation dataset: (I) yearly summer precipitation amounts and (II) the overlapping decade precipitation mean, with a shift of one year from 1971–1980 to 1994 –2003. The lengths of both subsets of summer precipitation series are 33 and 24, respectively.

3. Methods: model development

3.1. Independent variables

Spatial fields of precipitation are correlated with many environmental or geographic factors especially elevation and geographic coordinates. In this study, 54 candidate independent variables, which can explain spatial variability in the climate data, have been evaluated, and then selected (*Table 2*). A large number of geographical variables are evaluated, because none is a priori the most important. They are related to:

- Omnidirectional variables describing elevation (27 variables), topographic features (slope and slope aspect: 13 variables). Those morpho-topographic variables represent the values from the grids omnidirectionally oriented around central grids. The central grids can be defined as grids in which stations are located. Eight directions around the central grid have been defined as: (1) north, (2) east, (3) south, (4) west, (5) north-east, (6) south-east, (7) south-west, and (8) northwest.
- Cross products involving (eight variables): (a) geographical coordinates and elevation variables (maximum: percentile 95%, minimum: percentile 5%, and average elevation to north and east) and (b) geographical coordinates and topographic features (slope). Cross products were calculated (as indicated in *Table 2*) to obtain west-east or south-north gradient of elevation and topography (*Brown* and *Comrie*, 2002; *Vicente-Serrano et al*., 2003).
- Land cover parameters were selected from a grid data with spatial resolution of 1 km resolution. An area of 10×10 km around meteorological stations was delimited and four indexes (Iwat, Iveg, Iagr, Iurb) characterizing the main units of landscapes (water bodies,

vegetation cover, agricultural area, and urban and artificial area) in this area were calculated. Indexes are calculated as a ratio (%) of the area covered by land cover units to the total area (grid) of 10×10 km resolution.

- Geographic coordinates: latitude (Lat) and longitude (Long).

Table 2. Candidate independent variables for stepwise regression models. They are related to the Elevation variables, topographic variables (slope and slope aspect), cross products involving geographical position and morpho-topographic variables (elevation and topographic variables), geographical coordinates, and land use or cover variables. Abbreviations Emax, Emin, Eavg are related to maximum, minimum, and average elevation. Slope is related to slope values in %, while Oreg abbreviates slope aspect. The numbers after letters indicate orientation of grid from which slope or elevation values have been taken out, inside a bound of 3×3 km around the central grid. Central grid is defined as a grid in which meteorological stations are located. For elevation and slope eight directions have been defined (see section 3.1), while for slope aspect 4 have been taken into account (1 for east, 2 for west, 3 for north, and 4 for south). Abbreviations ending with "-grad" are related to cross product involving longitude (with number 1 on the end of letters) or latitude (with number 2 on the end of letters) and elevation variables (Egrad, Exgrad, Engrad) or slope (Sgrad). Iagr, Iveg, Iurb, and Iwat abbreviate four Indexes of landscape units (agriculture, vegetation, urban area, water bodies)

Candidate independent variables	**Abbre-viations**	**Candidate independent variables**	**Abbre-viations**
Central grid average elevation	Eavg0	Central grid slope values	Slope0
Central grid minimum elevation (percentile 5%)	Emin0	Slope in the north	Slope1
Central grid maximum elevation (percentile 95%)	Emax0	Slope in the east grid	Slope2
Average elevation in the north from the central grid	Eavg1	Slope in the south	Slope3
Minimum elevation (percentile 5%) in the north	Emin1	Slope in the west	Slope4
Maximum elevation (percentile 95%) in the north	Emax1	Slope in the north-east	Slope5
Average elevation in the east from the central grid	Eavg2	Slope in the south-east	Slope6
Minimum elevation (percentile 5%) in the east	Emin2	Slope in the south-west	Slope7
Maximum elevation (percentile 95%) in the east	Emax2	Slope in the north-west	Slope8
Average elevation in the south from the central grid	Eavg3	Slope facet east	Oreg1
Minimum elevation (percentile 5%) in the south	Emin3	Slope facet west	Oreg2
Maximum elevation (percentile 95%) in the south	Emax3	Slope facet north	Oreg3
Average elevation in the west from the central grid	Eavg4	Slope facet south	Oreg4
Minimum elevation (percentile 5%) in the west	Emin4	Cross product Long × average elevation	Egrad1
Maximum elevation (percentile 95%) in the west	Emax4	Cross product Lat × average elevation	Egrad2
Average elevation in the north-east from the central grid	Eavg5	Cross product Long × slope	Sgrad1
Minimum elevation (percentile 5%) in the north-east	Emin5	Cross product Lat × slope	Sgrad2
Maximum elevation (percentile 95%) in the north-east	Emax5	Cross product Long × 95% percentile of elevation	Exgrad1
Average elevation in the south-east from the central grid	Eavg6	Cross product Long × 5% percentile of elevation	Engrad1
Minimum elevation (percentile 5%) in the south-east	Emin6	Cross product Lat × 95% percentile of elevation	Exgrad2
Maximum elevation (percentile 95%) in the south-east	Emax6	Cross product Lat × 5% percentile of elevation	Engrad2
Average elevation in the south-west from the central grid	Eavg7	Index for the ratio of Agricultural area	Iagr
Minimum elevation (percentile 5%) in the south-west	Emin7	Index for the ratio of the Vegetation covered area	Iveg
Maximum elevation (percentile 95%) in the south-west	Emax7	Index for the ratio of Urban area	Iurb
Average elevation in the north-west from the central grid	Eavg8	Index for the ratio of the Water bodies	Iwat
Minimum elevation (percentile 5%) in the north-west	Emin8	Longitude	Long
Maximum elevation (percentile 95%) in the north-west	Emax8	Latitude	Lat

3.2. *Selection of suitable predictors for regression model and temporal analysis*

Analysis of time variation of relationship between precipitation and geographical factors mentioned in *Table 2* was the objective of this study. Relationships are analyzed through regression models that are performed at various time steps. In order to assure stability of the regression equations and comparability of models in time, it was necessary to select similar suitable and stable independent variables. Selection of suitable independent variables was based on two approaches: stepwise-based approach (STW), and principal component analysis (PCA)-based approach, both for yearly-based precipitation models (STW I / PCA I) and overlapping decade-based precipitation models (STW II / PCA II).

3.2.1. *Stepwise regression based models (STW)*

Stepwise selection of suitable predictors has been made in two steps. At the first step, the significant candidate independent variables have been selected for each model at various time steps. At the second step, only the most frequently selected significant predictors have been taken into account.

(a) It is important to remind that the set of geographical variables used as predictors (*Table 2*), particularly topographic and elevation parameters, are collinear. The choice of suitable predictors from this set has a significant influence on the behavior of models. Hence, forward stepwise linear regression was used to model summer precipitation as function of the collinear geographical factors at time steps of annual (STW I) and overlapping decades (STW II) from 1971–2003. All predictors mentioned in *Table 2* are used. A p-value of 0.05 has been used to force out of the model any non-significant effects, and to select significant, and non-collinear independent variables. Stepwise regression has been used in many studies (*Ninyerola et al.*, 2000, 2007; *Marquinez et al.*, 2003; *Vicente-Serrano et al.*, 2007) as an accurate method in examining relations between precipitation and collinear independent variables.

(b) On the second step, the most frequently selected significant predictors by both STW I and STW II-based models was considered. A threshold frequency value of 20 – 40% was defined to select them. Geographical variables of which frequency value does not reach at least 20% are considered as improper for temporal analysis, and are discarded; while independent variables exceeding the defined threshold are retained. If the retained variables are collinear, the operation is repeated (using a higher threshold value, i.e., 30 – 40%) until no co-linearity is found among them. Using this procedure, six predictors have been retained for both STW I and STW II-based precipitation models (see *Table 3* and *Figs. 2* and *3*). For the STW I-based models, the selection was ended at the first step, where frequency of significant independent variables exceeded 30%. However, for the STW II-based models, 11 collinear predictors have been selected at the first step (*Fig. 3a*). The operation was repeated for the 11 predictors to select a final series of six non-collinear predictors exceeding 40% (*Fig. 3b*).

Table 3. Definitive selected independent variables for all models

I. Year by year time series			**II. Moving long-term precipitation mean**		
		Models			
STW I	**PCA-A I**	**PCA-B I**	**STW II**	**PCA-A II**	**PCA-B II**
Slope2	PCI	Emin0	Slope2	PCI	Emin0
Slope5	PCII	Slope1	Slope5	PCII	Slope1
Emax8	PCIII	Vegetation	Emin8	PCIII	Vegetation
Lon	Lon	Lon	Lon	Lon	Lon
Lat	Lat	Lat	Lat	Lat	Lat
Oreg2	Oreg2	Oreg2	Oreg2	Oreg2	Oreg2
–	–	–	–	Oreg3	Oreg3

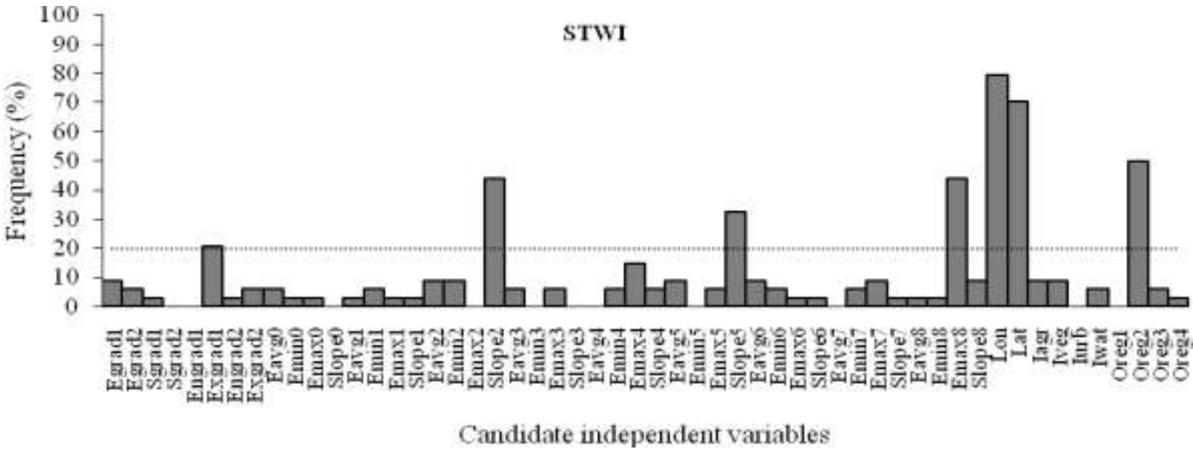


Fig. 2. Frequency of significant variables entering models for STWI-based models.

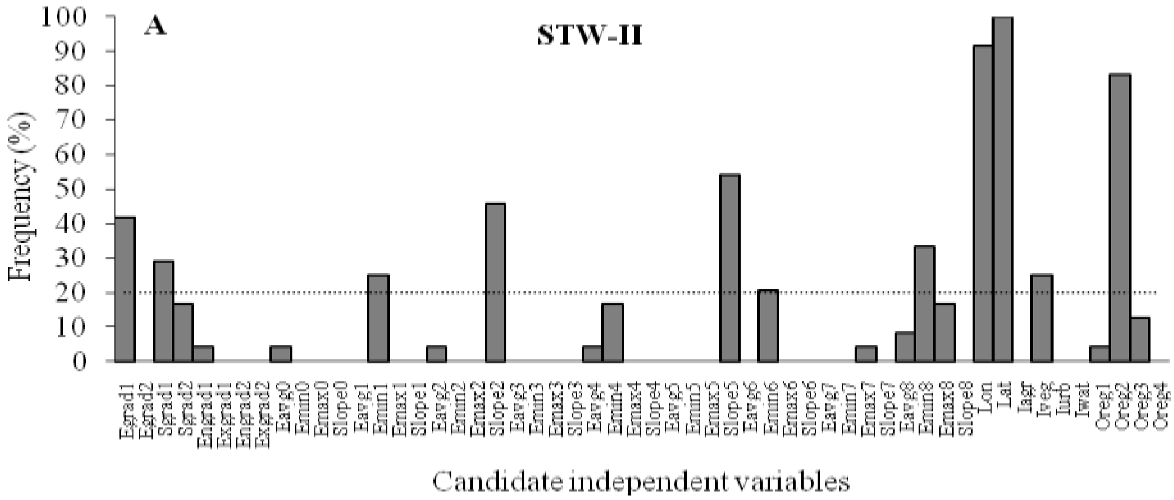


Fig. 3a. Frequency of significant independent variables for STW II-based models. Dotted line represents the threshold value for selecting more frequent predictors. First step (A) of selection.

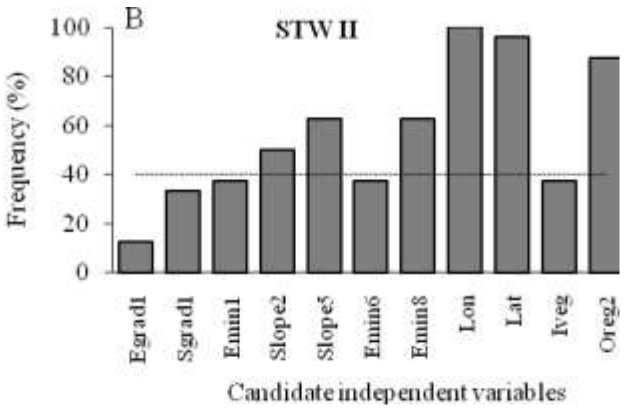


Fig. 3b. Frequency of significant independent variables for STW II-based models. Dotted line represents the threshold value for selecting more frequent predictors. Second step (B) of selection.

3.2.2. PCA – based models (PCA)

Principal component analysis (PCA) has been used to eliminate random variability in independent factors, and to generate stable models in time. The sets of independent variables have been checked for correlation before performing PCA. Eight independent variables, which were less correlated or uncorrelated with other independent variables, were discarded for the PCA. They were: geographical coordinates (longitude and latitude), two units of land cover (urban area and water bodies), and four orographic facets (slope aspect). PCA was performed for the remaining 46 variables. The number of principal components (PCs) to retain for rotation was given using screen test. Three PCs have been retained. They are related to: (1) characteristics of elevation, (2) topographic features, (3) land use and land cover parameters: agricultural and vegetation covers. The three PCs explain about 90.3% of total variability. Morpho-topographic variables that have the highest loadings with those retained PCs were then selected. For the further regression models, both PCs scores (three variables), and independent variables (three morpho-topographic variables), that were selected assuming highest loading, were considered as candidate independent variables. These candidate predictors selected using PCA were recombined with the 8 discarded variables before performing PCA. Then the stepwise regression has been performed to select significant and noncollinear variables. The frequencies of variables (*Fig. 4*) have been analyzed (considered as in Section 3.2.1.b). The frequent variables have been considered as stable and, therefore, suitable for multivariate regression model. This approach helped to avoid the problem encountered during interpretation of the PCs that involve independent uncorrelated variables. The final selected independent variables are shown in *Table 3*. Using this approach, six significant independent variables were selected for yearly-based precipitation models (PCA I) and seven independent variables were selected for decade-based precipitation models (PCA II) (see *Table 3*). PCs scores (PCA-A) or selected variables assuming the highest loadings (PCA-B) have been used to compare their effect on the models. *Table 3* shows important geographical variables that influence significantly the spatial patterns of precipitation in the Czech Republic. All approaches selected the geographical coordinates (including continentality) and the westward slope aspect.

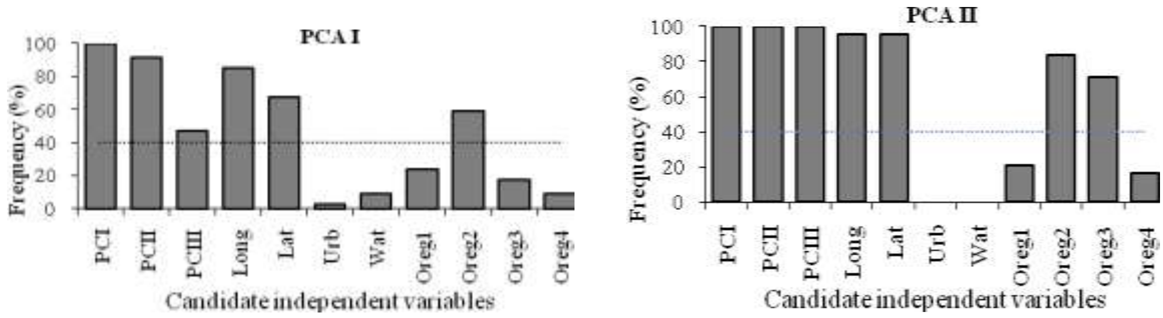


Fig. 4. Frequency of significant variables entering models for PCA I and II-based models. Dotted lines represent the threshold value for selecting more frequent predictors.

3.3. Regression models and trend detection

Once stable significant variables were selected, multivariate regression precipitation models were then performed using them as predictors. The multiple regression relationship is obtained through the following equations:

$$\begin{cases} P_{est}(t_1) = b_0 + b_1(x_1) + \ldots + b_n(x_n) \\ P_{est}(t_2) = b_0 + b_1(x_1) + \ldots + b_n(x_n) \\ \ldots \\ P_{est}(t_n) = b_0 + b_1(x_1) + \ldots + b_n(x_n) \end{cases}, \tag{1}$$

where b_{1-n} is the multiple regression coefficient adjusted for each retained independent variable x_n; P_{est} represents the predicted rainfall, and t_{1-n} is the time step (i.e., year or decade).

In order to carry out temporal analysis of the relationship between significant geographical factors and precipitation, time series of regression coefficients from each time resolution and approaches-based models were built. The linear trend in those series was estimated using a least-squares regression. The significance of trends was determined using the confidence interval (*CI*) given by the following equation:

$$CI = t +/- \frac{2.042 \cdot \sigma e}{\sqrt{n} \cdot \sqrt{\sigma x^2}}, \tag{2}$$

where t is the trend value, n is the length of the time series of the regression coefficients, σe is the standard deviation of the residuals, and σx is the standard deviation of the independent variable.

4. Results

4.1. Model performance

Several standard statistical measures of models performance and accuracy were calculated. The goodness of fit of the model and the proportion of the variation of summer precipitation explained by the model are measured by the coefficient determination (R^2). The magnitude and sign of errors of the regression model are given by mean absolute error (MAE) and root mean square error (RMSE).

The time variation of the coefficients of determination for all approaches and time resolution-based precipitation models is displayed in *Fig. 5*. The magnitude of errors is measured by the rootmeansquare error (*Fig. 6*) and mean absolute error (*Fig. 7*).

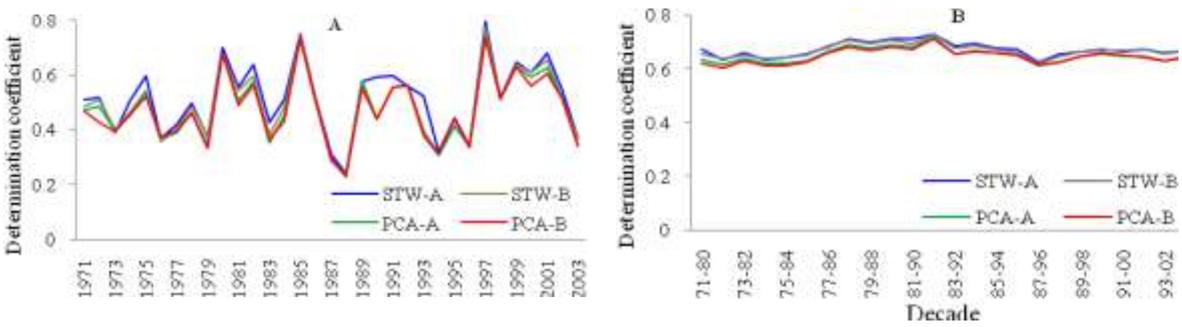


Fig. 5. Time variation of determination coefficients of STW and PCA-based models for year (A) and decade (B) time resolution.

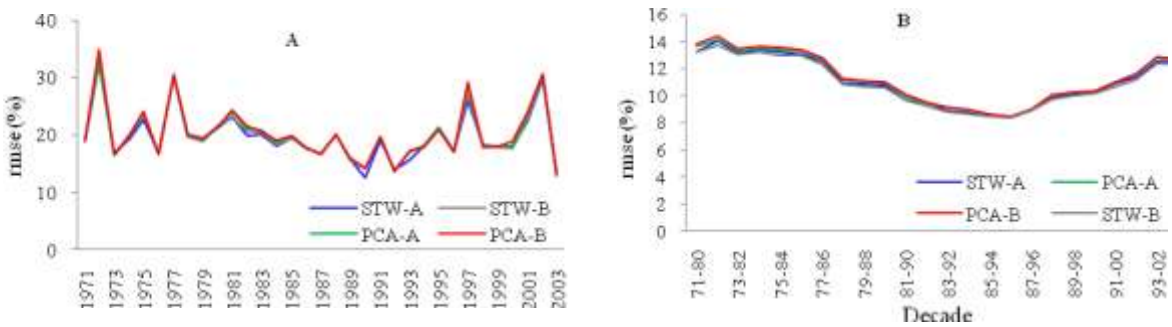


Fig. 6. Time variation of the root-mean-square errors of STW and PCA-based models for year (A) and decade (B) time resolution.

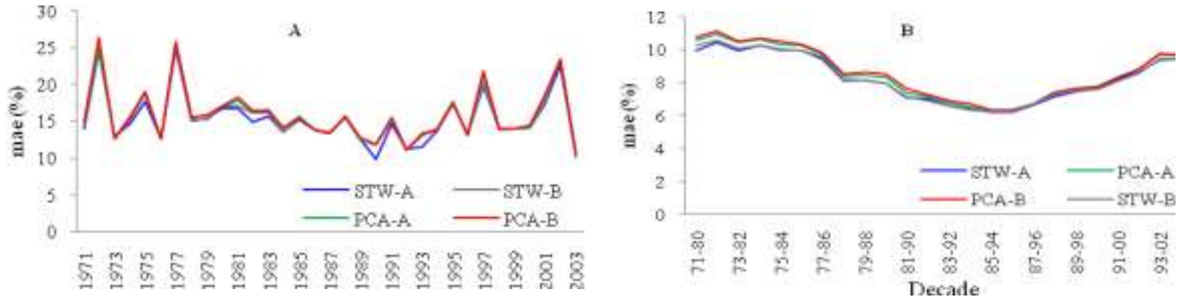


Fig. 7. Time variation of the mean absolute errors of STW and PCA-based models for year (A) and decade (B) time resolution.

The coefficients of determination and errors for model based on different approaches vary consistently in time, whereas their time variations are inconsistent for time resolution-based models. In fact, the STW-based models are more efficient than PCA-based models. Nevertheless, the difference between both approaches-based models is minor. Model accuracy varies widely among models generated at various time resolutions. Models based on long-term time resolution are more accurate and explain more variation than those based on shorter time resolution. An important variability in summer precipitation (> 0.61 for both approaches-based models) is consistently captured by ten-year models for both PCA and STW-based approaches. The magnitude of errors (RMSE, MAE) does not reach 15% (37 mm) of summer precipitation for ten-year models. On the contrary, terrain influences did not always account for an important variability of summer precipitation for the annual precipitation models. The coefficients of determination of these models describe an important inter-annual variability; they vary from 0.23 (1988) to 0.7 (1997). Similarly, the magnitude of errors is fluctuant. The larger (35% – 86 mm) and the smaller RMS errors (13% – 32 mm) have occurred in 1972 and 2003, respectively. Unlike other years, where model errors are inversely proportional to the explained variance, both variance explained and error yielded by the models are large in 1997 and 2002.

Several models were unable to capture more than 50% of variability in precipitation and to generate small error. Considering that a good precipitation model must capture at least 50% of precipitation variability (*Ninyerola et al.*, 2000) and yield very small prediction error, we can conclude that numbers of these models failed. The temporal variability of summer precipitation combined with the ability of the set of predictors to capture variation in data can explain it. *Fig. 8* displays the relationships between rainfall departure (from the long-term precipitation average) and coefficient of determination. It reveals that the stronger the negative rainfall anomalies for a given year, the smaller the variation explained in precipitation data. Indeed, the coefficients of determination fall under 0.5 or rarely (once) overtake this value during dry years. Therefore, additional predictors (different from the used set of predictors) or additional analysis on removing temporal variability were needed to improve models. If it is necessary to add other auxiliary variables, then it should be important to consider that not only geographical factors but also other factors such as atmospheric circulation affect the spatial distribution of precipitation. However, their mechanisms are more complex and not easy to evaluate their statistical relationships.

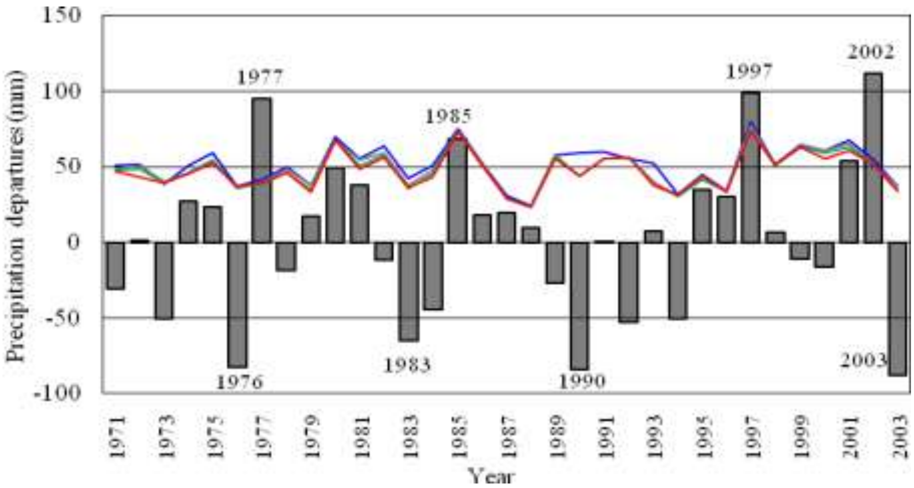


Fig. 8. Comparison of determination coefficients (color lines) with precipitation departures (histogram).

4.2. Time variation of effect of significant geographical factors

The relationships between summer precipitation and geographic variables, selected using STW approach and PCA, have been examined through regression model. Resulting coefficients of regression for each approach and time resolution-based model are plotted in *Figs. 9–12*. Six independent variables have been selected using STW approach: slopes of grids eastward and northeastward from the locations, maximum elevation northwestward from locations, geographic coordinates (longitude and latitude), and slope aspect westward.

At the annual time resolution, spatial patterns of summer precipitation show an increase of precipitation with a growing value of elevation and slope aspect. However, they show an increase and a decrease of precipitation to increasing latitude, longitude, and slope. Considering the inter-annual variability of the relationships between precipitation and each significant geographical variable, it can be pointed out in some extreme cases of strong relationships. They are found between precipitation and elevation (emax8), latitude, slope (slope5), and slope aspect (oreg2), respectively, during 1980, 2002, 1997, and 1972. The precipitation models based on STW approach reveal that the spatial pattern of precipitation during the heavy precipitation events of 1997 and 2002 were strongly related to the slope northeastward and latitude, respectively. The fields of intense precipitation during such years have a significant influence on the spatial pattern of precipitation across the whole country. For example, the decrease of precipitation with an increasing longitude during 1997 is mainly due to the fields of intense precipitation in the northeastern part of the country. Spatial patterns of summer precipitation and precipitation amount fluctuate in time according to different factors such as atmospheric conditions.

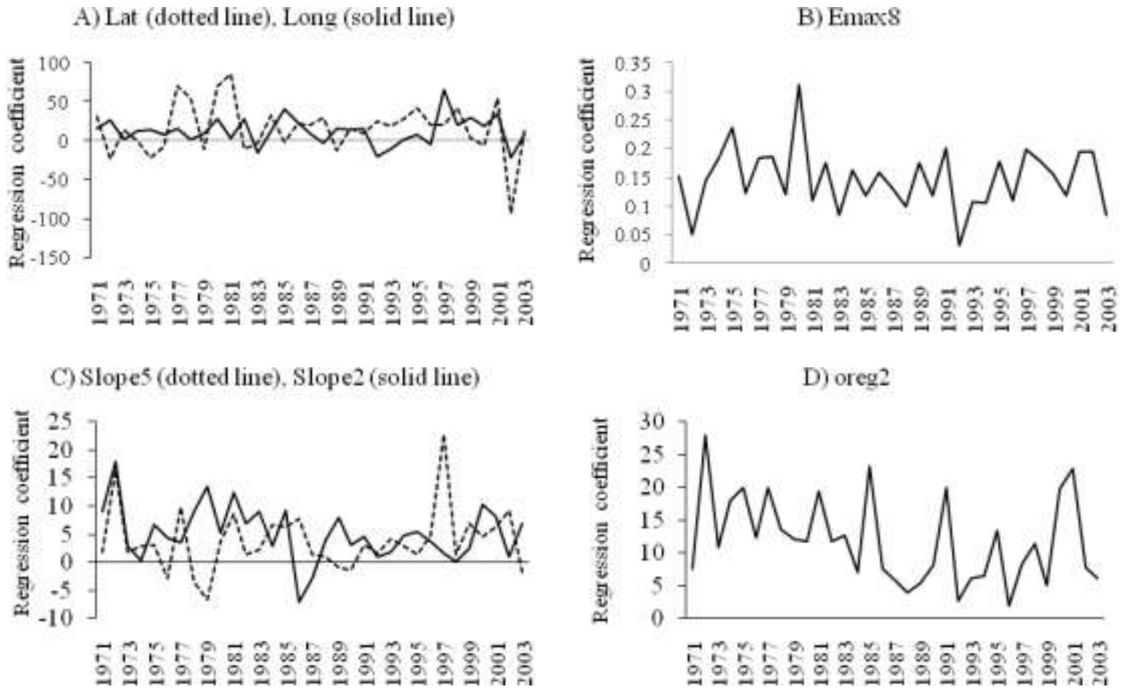


Fig. 9. Time variation of the regression coefficients of significant predictors (as indicated in A, B, C, D) for STW I-based models.

During the whole period under study, the precipitations have had a positive relationship with elevation variables. The well-known effect of elevation on the spatial pattern of summer precipitation, which is related to an increase of precipitation with growing elevation, is observed along the entire period. The highlands are seemingly rainier during summer than lowlands all over the country. However, during some years (1976, 1986, 2003, etc.) spatial pattern of precipitation is characterized by a decrease of rainfall with a growing value of slope (slope2, slope5). Thus, if multiple factors of mountain regions are taken into account, their relationships with precipitation become more complex than a simple increase with growing elevation. During the period 1971–2003, the precipitations increased strongly in 1980 and slightly in 1992 with maximum elevation northwestward from the locations. The slope orientated eastward from the locations has influenced considerably the summer precipitation during 1972.

For the decade-based models, minimum elevation northwestward (Emin8) has been selected as predictor instead of Emax8. Unlike the yearly-based models, no geographical variable was related to any decreases of summer precipitation. Although the effect of minimum elevation (Emin8) on precipitation is less variable in time, the influences of the topographic and geographical position fluctuate largely in time. The variation of the effect of topographic features (slope5, slope2, and oreg2) shows two peaks at the beginning and at the end of the period under study. The influence falls in the middle of this study period (i.e., during 1981–1997). This can be related to drought that occurred in the Czech Republic in this period (*Kaspar* and *Muller*, 2008).

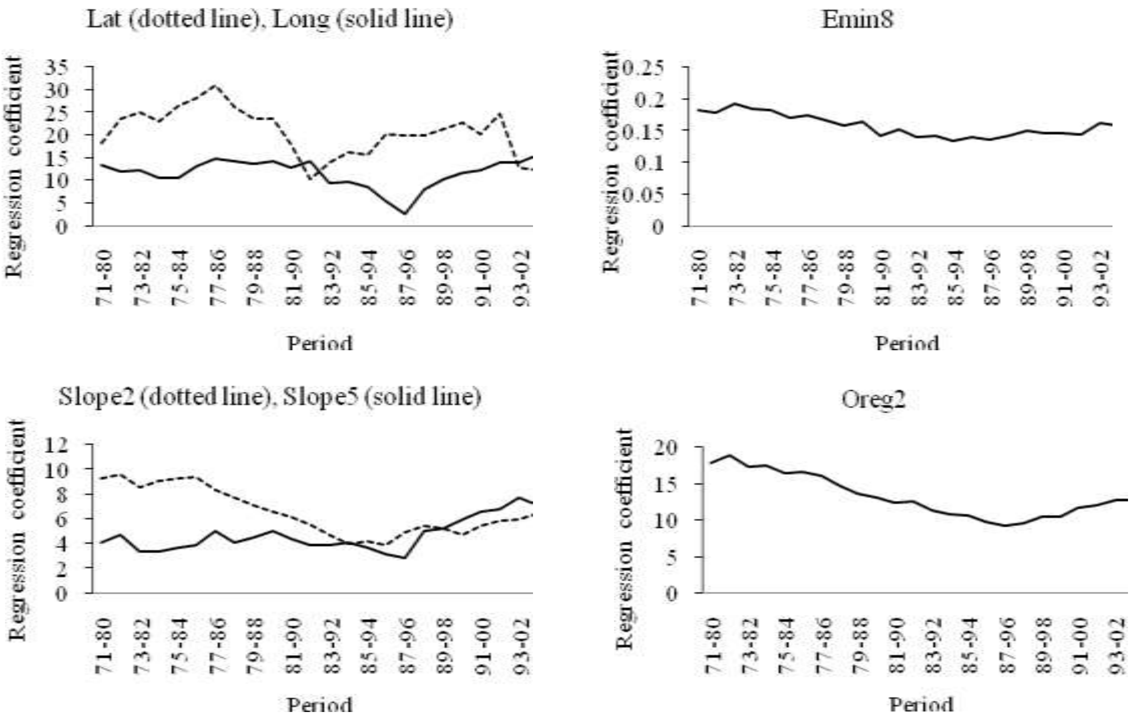


Fig. 10. Time variation of the regression coefficients of the significant predictors for STW II-based models.

Time variation of the relationships between spatial patterns of summer precipitation and significant geographical variables selected using PCA-based approach has been also analyzed (*Figs. 11* and *12*). New variables were selected using the loading of the components: slope northward from the locations (Slope1), vegetation (Veg), and minimum elevation at the locations. An additional geographical variable was specifically selected for the ten-year models: slope aspect oriented northward. These additional predictors influence the spatial patterns of precipitation as well. They are related to the increase of precipitation, except the vegetation during some years (1984, 1992, and 2002). A strong relationship with precipitation is observed during 1972 for slope2 and in 1995, 2002 for the vegetation.

The PCA-based model show an increase of summer precipitation for the two PCs scores during the period considered by this study (*Figs. 11* and *12*). In particular, spatial patterns of summer precipitation in 1972 and 1980 have been related to PC1 and PC2, respectively. Similarly to STW-based models, PCA-based models reveal that spatial distribution of precipitation during the summer of 1997 is related to topographic features and geographical coordinates, while spatial distribution of summer precipitation in 2002 is influenced by urban effect, elevation, and especially latitude.

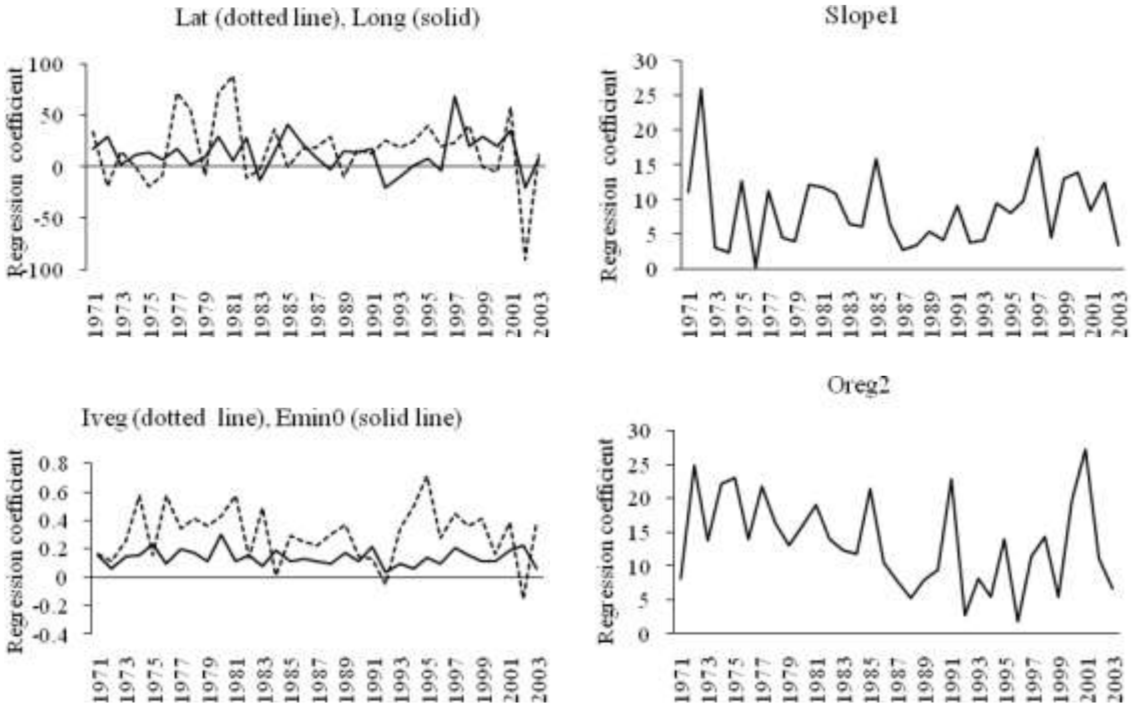


Fig. 11. Time variation of the regression coefficients of the significant predictors for PCA I-based models.

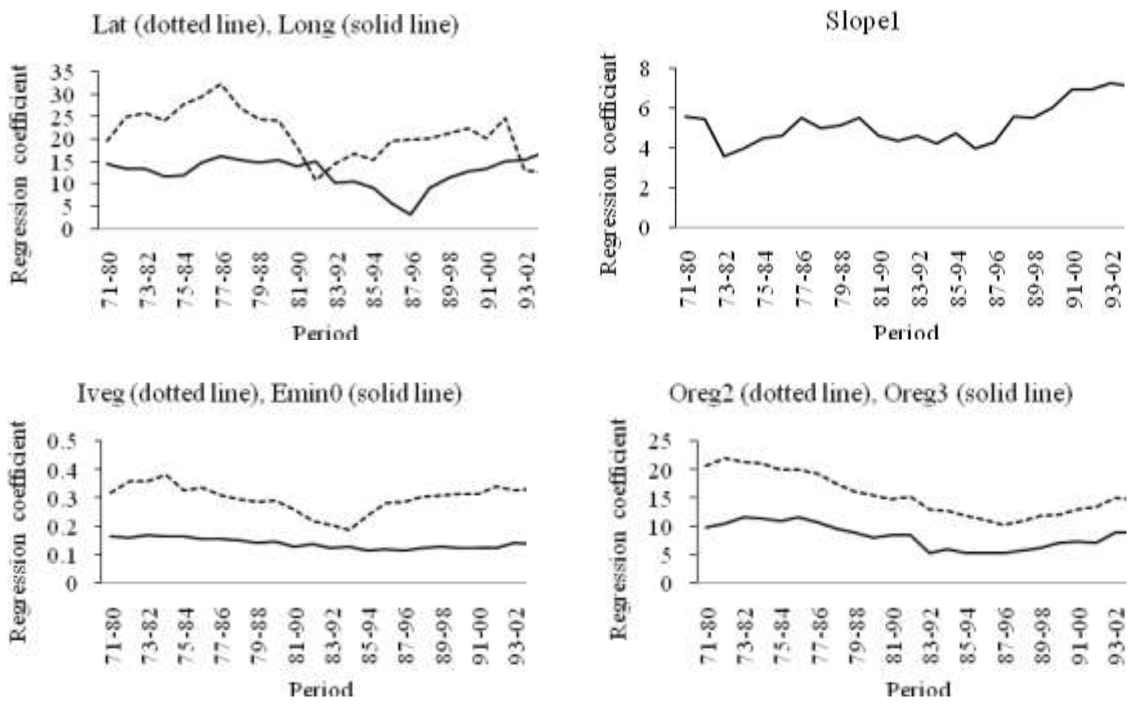
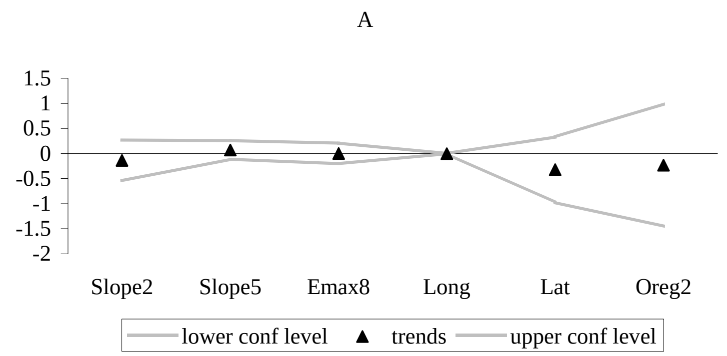


Fig. 12. Time variation of the regression coefficients of the significant predictors for PCA II-based models.

The extremeness of precipitation during the summer of 1997, especially during the two heavy precipitation episodes in July, is associated with the atmospheric advection of moist air stream (*Rezacova et al.*, 2005). Meteorological features were specifically characterized by an intensive influx of moisture into Central Europe and intensive upward motions in the precipitation area. Finally, the regression coefficients of precipitation models reveal that topographic feature and geographic coordinates have stronger influence on spatial distribution of summer precipitation over the Czech Republic than other geographical factors.

4.3. Trends in coefficients of regression of significant predictors

The trends in regression coefficients for the yearly-based models during the period 1971–2003 are displayed in *Fig. 13* (*a,b,c*). It is obvious that the trends in regression coefficients are negative for almost all selected independent variables. The only exception concerns the slope of grids northeastward from the locations, which has positive trend. The magnitude of the trends is higher for latitude and west slope (reaching $-0.37\ \mathrm{yr}^{-1}$ and $-0.23\ \mathrm{yr}^{-1}$) than for other independent variables, especially elevation, vegetation, and longitude. The sign and magnitude of trends are similar for the same predictors independently selected from different model-based approaches. The positive and negative trends detected are insignificant. Therefore, the relationships between the spatial patterns of summer precipitation and geographical variables, during the relatively short period considered in this study, are stable in time.



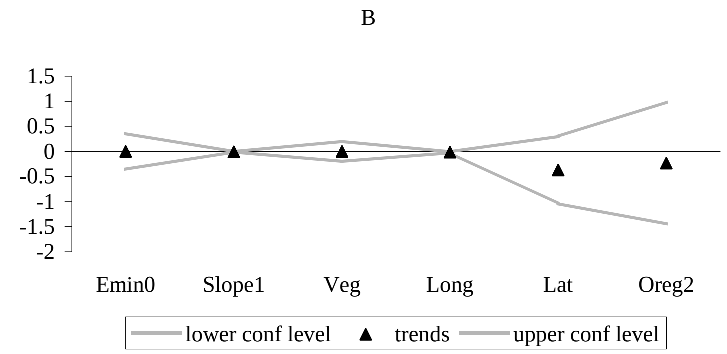


Fig. 13a,b. Trends in coefficient of regression estimated by STW-based (A) and PCA-B (B) precipitation models. Models are generated at the annual time step.

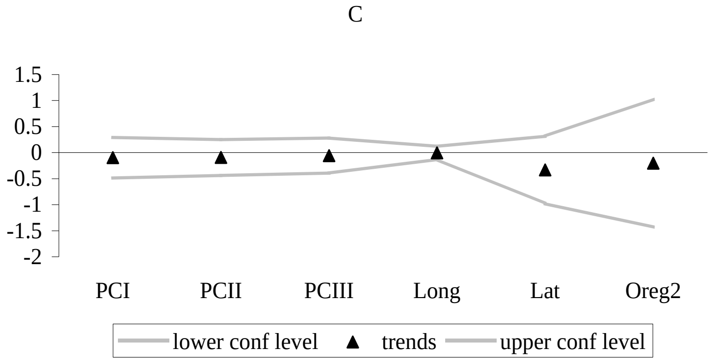


Fig. 13c. Trends in coefficient of regression estimated by PCA A-based precipitation models. Models are generated at the annual time step.

5. Conclusions

Relationships between geographical variables (land cover, geographical positions, morpho-topographical features) and rainfall spatial pattern were analyzed in this study using regression models. Stepwise regression and rotated PCA identified several significant geographical variables: slope2, slope5, emax8, emin0, vegetation, oreg3, oreg2, and geographical coordinates. Geographical coordinates (including continentality) and slope orientation westward are the most significant predictors for modeling the spatial distribution of summer precipitation over the Czech Republic. Except the minimum elevation (Emin0), all important topographical factors correspond to the outside grids at the distance of 1 km from the location. Therefore, the terrain characteristics had more significant influence when a larger area than a station location is taken into account for selecting independent variables. Similarly, the models working with independent variables selected from stations are slightly inaccurate in comparison with models working with independent geographical variables describing a large area around the location. Furthermore, PCA and STW-based approaches for selecting significant geographical variables to model spatial patterns of precipitation over the Czech Republic are consistent. Nevertheless, PCA-based models seem to be less powerful than STW-based models. They yield a small explained variance and a large prediction error. In the same way, models based on a one-year time resolution are notably less powerful than models based on long-term time resolution.

The time variation of explained variance indicated that precipitation variability has an influence on the variance accounted by models and on its performance. Thus, smaller explained variance is accounted by models during dry years. A temporal analysis of regression coefficients from ten-year precipitation models showed positive relationships between precipitation and geographical factors. Spatial patterns of averaged summer precipitation at a decade time resolution are modeled as increasing with continentality, morpho-topographic features, and latitude. On the other side, at the smaller time-resolution, negative relationships were found, especially between precipitation and topographical features, vegetation, and geographical positions. The spatial pattern of precipitation during the summer of 2002, for example, is strongly related to decreasing latitude.

Trend analysis of regression coefficients revealed that relationships between summer precipitation patterns and morpho-topographical features, land cover, and geographical positions are stable in time. No significant trend in the model parameters (effect of geographical factors) has been found during 1971–2003. Model parameters are stable in time. Therefore, spatial prediction of precipitation based on single realization in time (i.e., long term average) is not biased by the length of the sample period.

Spatial precipitation patterns vary in time according to the effect of geographical factors, as well as performance of models. The models failed to capture the relationships between the precipitation patterns and the geographical factors during dry years, namely 1987 and 1988. These years have been dominated by intensive drought (*Kaspar* and *Muller*, 2008). Precipitation patterns during these years could be well modeled using other auxiliary independent variables such as the dominant mode of atmospheric circulation that are linked with both spatial and temporal variability of precipitation. Although this was not the aim of this analysis, the conclusions of this study show the necessity to investigate further on the relationships between precipitation pattern and atmospheric circulation.

References

Baigorria, G.A., Jone, J.W. and *O'Brien, J.J.*, 2007: Understanding rainfall spatial variability in southeast USA at different timescales. *Int. J. Climatol. 27,* 749-760.

Basher, R.E., Zheng, X., 1998: Mapping rainfall fields and their ENSO variation in data-sparce tropical south-west Pacific Ocean region. *Int. J. Climatol. 18,* 237-251.

Brown, D.P., Comrie, A.C., 2002: Spatial modeling of winter temperature and precipitation in Arizona and New Mexico, USA. *Clim. Res. 22,* 115-128.

Courault, D., Monestiez, P., 1999: Spatial interpolation of air temperature according to atmospheric circulation patterns in southeast France. *Int. J. Climatol. 19,* 365-378.

Daly, C., 2006: Guidelines for assessing the suitability of spatial climate data sets. *Int. J. Climatol. 26,* 707-721.

Daly, C., Neilson, R.P., Philip,s D.L., 1994: A statistical-topographic model for mapping climatological precipitation in mountainous terrain. *J. Appl. Meteorol 33,* 140-158.

Daly, C., Gibson, W.P., Taylor, G.H., Johnson, G.L., Pasteris, P., 2002: A knowledge-based approach to the statistical mapping of climate. *Clim. Res. 22,* 99-113.

Diodatto, N., 2005: The influence of topographic co-variables on the spatial variability of precipitation over small regions of complex terrain. *Int. J. Climatol. 25,* 351-363.

Drogue, G., Humbert, J., Deraisme, J., Mahr, N., Freslon, N., 2002: A statistical – topographic model using an omnidirectional parametrization of the relief for mapping orographic rainfall. *Int. J. Climatol. 22,* 599-613.

Goodale, C.L., Aber, J.D., Ollinger, S.V., 1998: Mapping monthly precipitation, temperature, and solar radiation for Ireland with polynomial regression and a digital elevation model. *Clim. Res. 10,* 35-49.

Hengl, T., 2007: A practical guide to geostatistical mapping of environmental variables. *JRC Scientific and Technical Report* - European Communities, Luxembourg, 143 pp.
Hulme, M., *New, M.*, 1997: Dependence of large-scale precipitation climatologies on temporal and spatial sampling. *J. Climate 10*, 1099-1113.
Johnson, G.L., Hanson, C.L., 1995: Topographic and Atmospheric influences on precipitation variability over a mountainous watershed. *J. Appl. Meteorol. 34,* 68-87.
Joly, D., Nilsen, L., Fury, R., Elvebakk, A., and *Brossard, T.*, 2003: Temperature interpolation at a large scale: test on a small area in Svalbard. *Int. J. Climatol. 23*, 1637–1654.
Kakos, V., 2001: Maximum precipitation in the Czech Republic in terms of synoptic meteorology (in Czech). In *Sbornik přednášek ze semináře výsledkům grantového projektu VaV/510/97*, 46-60.
Kaspar, M., Muller, M., 2008: Selection of historic heavy large-scale rainfall events in the Czech Republic. *Nat. Hazards Earth Syst. Sci. 8*, 1359–1367.
Marquinez, J., Lastra, J., Garcia, P., 2003: Estimation models for precipitation in mountainous regions: the use of GIS and multivariate analysis. *J. Hydrol. 270*, 1-11.
Ninyerola, M., Pons, X., Roure, J.M., 2000: A methodological approach of climatological modelling of air temperature and precipitation through GIS techniques. *Int. J. Climatol. 20*, 1823-1841.
Ninyerola, M., Pons, X., Roure, J.M., 2007: Objective air temperature mapping for Iberian Peninsula using spatial interpolation and GIS. *Int. J. Climatol. 27,* 1231-1242.
Prudhomme, C., Reed, D.C., 1999: Mapping extreme rainfall in mountainous region using geostatistical techniques: a case study in Scotland. *Int. J. Climatol. 19,* 1337-1356.
Rezacova, D., Kaspar, M., Muller, M., Sokol, Z., Kakos, V., Hanslian, D., Pesice, P., 2005: A comparison of the flood precipitation episode in August 2002 with historic extreme precipitation events on the Czech territory. *Atmos. Res. 77*, 354-366.
Szentimrey, T., Bihari, Z., 2007: Mathematical background of the spatial interpolation methods and the software MISH (Meteorological Interpolation based on Surface Homogenized Data Basis). In *Proceedings from the Conference on Spatial Interpolation in Climatology and Meteorology,* Budapest, Hungary, 2004. *COST Action 719*, COST Office, 2007, 17-27.
Tolasz, R. et al. (eds.), 2007: *Climate Atlas of Czechia.* Czech Hydrometeorological Institute, Prague, Hardback ISBN 978-80-86690-1.
Vicente-Serrano, S.M., Saz, M.A., Cuadrat, J.M., 2003: Comparative analysis of interpolation methods in the middle Ebro valley (Spain): application to annual precipitation and temperature. *Clim. Res. 24*, 161-180.
Weisse, A.K., Bois, P., 2001: Topographic effects on statistical characteristics of heavy rainfall and mapping in the French Alps. *J. Appl. Meteorol. 40*, 720-740.