Developments in Statistics
Andrej Mrvar and Anuška Ferligoj (Editors)
Metodološki zvezki, 17, Ljubljana: FDV, 2002

Spatial Interpolation of Mean Yearly Precipitation using Universal Kriging

Damijana Kastelec[1] and Katarina Košmelj[2]

Abstract

The objective of our work was to find an appropriate method for the spatial interpolation of mean yearly precipitation (MYP) into a regular grid with 1 km resolution, for Slovenia. A geostatistical approach and the universal kriging method was chosen. In the analysis, the dependence of MYP on geographical variables like altitude, latitude and longitude was considered.

In the first step, we used exploratory methods to examine spatial variability of MYP and its dependence on geographical variables. The positive trend of MYP with altitude and the negative trend of MYP with longitude were taken into account in spatial interpolation.

The spherical variogram model was chosen, considering variogram anisotropy. On the basis of cross-validation technique we chose for the influential surrounding the ellipse with the longer axis 80 km. The universal kriging results were encouraging and comparable with the subjectively obtained map (Zupančič, 1996; Zupančič, 1998). The analysis showed that the results for the eastern part of Slovenia are better than for the western mountainous part. This is due to the lack of observational stations in the mountainous part of Slovenia where the spatial variability of MYP is high.

[1] Biotehnical Faculty, University of Ljubljana, Slovenia; Damijana.Kastelec@bf.uni-lj.si.
[2] Biotehnical Faculty, University of Ljubljana, Slovenia; Katarina.Kosmelj@bf.uni-lj.si.

1 Introduction

The geographical position of Slovenia and its climate induce large spatial variability of precipitation amounts. The maximum of mean yearly precipitation (MYP) is around 3000 mm in the western mountainous part of Slovenia, and the minimum of MYP is around 800 mm in the north-eastern part of Slovenia. The precipitation is measured daily on irregularly spatially distributed meteorological stations (Fig. 1). On the basis of daily data the yearly sums of precipitation are calculated. Usually, thirty-year means of the yearly sums of precipitation (MYP) are considered in climatology.

In practise, we often need to predict the value of MYP for a location where no measurements exist. The predictions are made on the basis of the data from the neighbourhood locations. In general, a map of MYP is required, where the predictions for several locations have to be made. Different maps of MYP for Slovenia were hand-drawn by Zupančič and by Pristov according to their subjective knowledge (Zupančič, 1996; Kolbezen and Pristov, 1997).

The objective of our work was to find an appropriate method for spatial interpolation of MYP into a regular grid with 1 km resolution, for Slovenia. A geostatistical approach was chosen. The universal kriging method was used and the dependence of MYP on geographical variables was considered.

In the paper, first some plots to examine the dependence of MYP on geographical variables is presented. In the next step, the universal kriging theory is summarised. Then the results of variogram estimation is presented. A cross-validation technique is used for the verification of the universal kriging model. The results of cross-validation using different influential surroundings are presented. Then, our map of MYP is compared with the map drawn by Zupančič (1996). The quality of the universal kriging predictions is analysed on the basis of kriging variances. Additionally, kriging predictions of MYP also taking into account some Croatian and Italian meteorological stations near the Slovenian border is presented.

For geostatistical calculations, the program Gstat 2.0g (Pbesma, 1998) was used. For mapping, we used IDRISI for Windows program (Eastman, 1997).

2 Exploratory analysis

We deal with mean yearly sums of precipitation (MYP) for 362 meteorological stations in Slovenia for the period 1961-1990. The data were collected by the Hidrometeorological Institute of Slovenia. In one case of spatial interpolation we also took into account the data from 25 meteorological stations in Croatia and 20 meteorological stations in Italia (Fig.1).

First, graphs to explore the dependence of the MYP on some geographical variables were plotted. In general, in Slovenia the amount of precipitation increases with altitude and decreases with longitude from West to Est. Figures 2, 3 and 4 show this dependence. The data from the meteorological stations in the north-western part of Slovenia (the drainage basin of the river Soča, Fig.5, left) are marked with black triangle. This region is specific: there is the maximum of MYP in Slovenia which is determined by relief and by weather dynamics. If we exclude the data from this region, the negative linear trend of MYP with longitude and positive linear trend of MYP with altitude are evident. There is no significant trend of MYP on latitude.

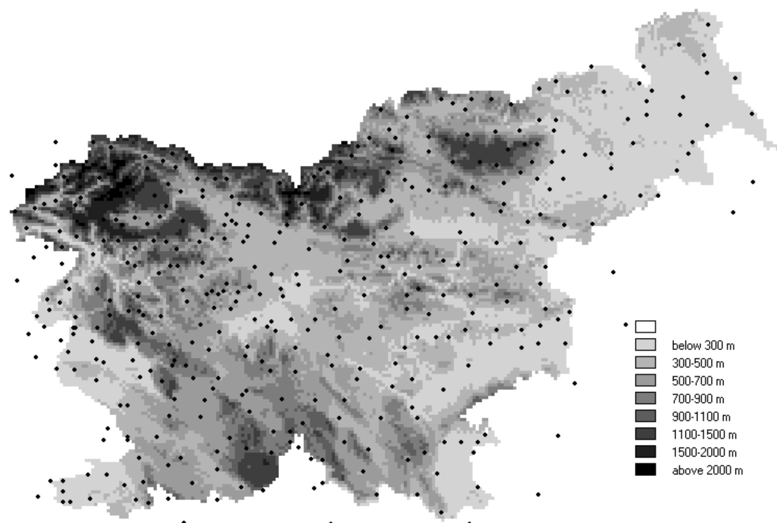

Figure 1: The relief of Slovenia and the locations of meteorological stations in the period 1961-1990, and some locations of precipitation measurements in Croatia and Italia.

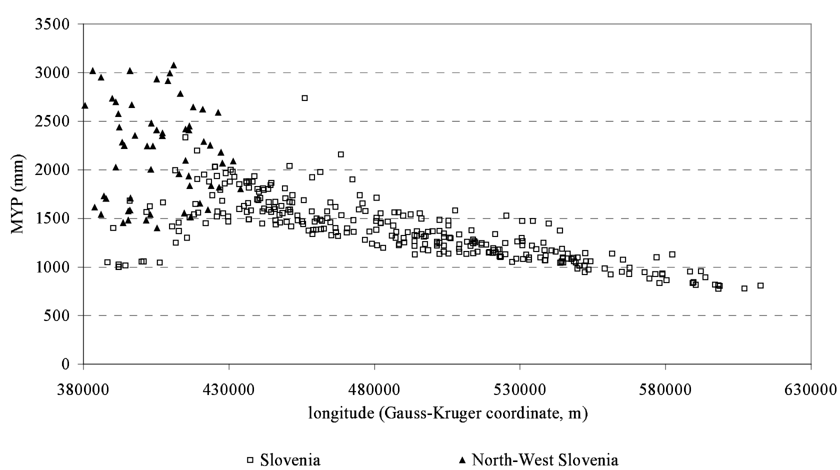

Figure 2: Dependence of mean yearly precipitation (MYP) on longitude for 362 locations in Slovenia (period 1961-1990).

There are very few locations with an altitude higher than 1000 m. The Fig. 5 (right) shows the distribution of the topography for Slovenia in comparison with the distribution of the meteorological station altitude. The discrepancy above 1000 m is evident: 11.5% of the topography in comparison with 5.2% of the stations.

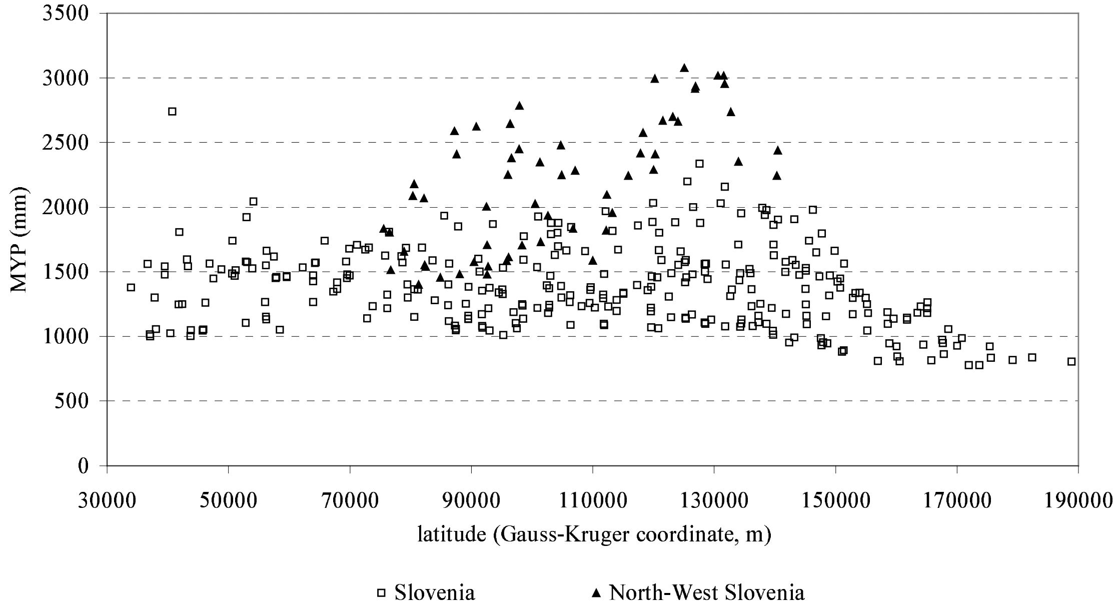

Figure 3: Dependence of mean yearly precipitation (MYP) on latitude for 362 locations in Slovenia (period 1961-1990).

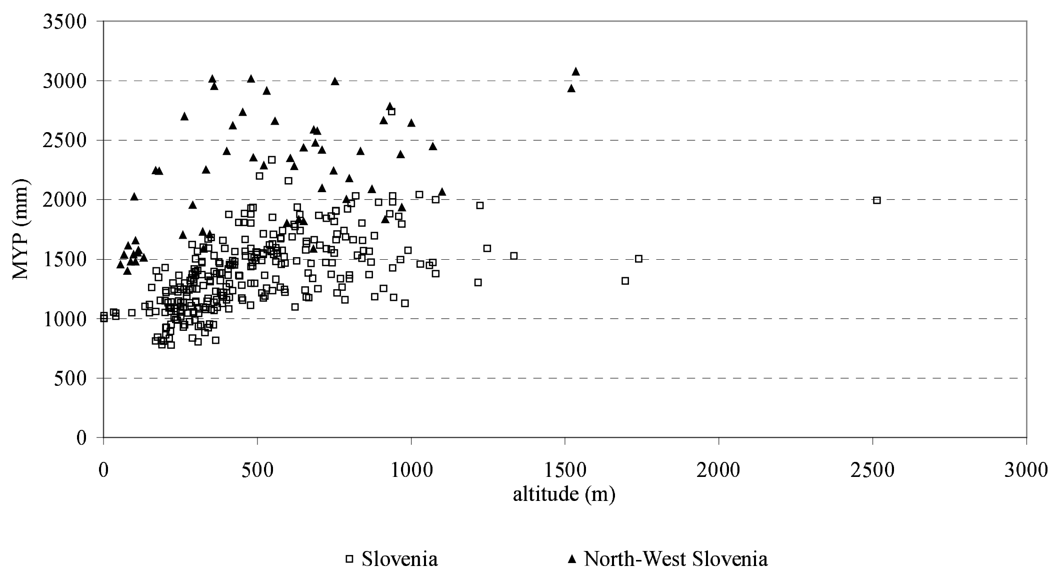

Figure 4: Dependence of mean yearly precipitation (MYP) on altitude for 362 locations in Slovenia (period 1961-1990).

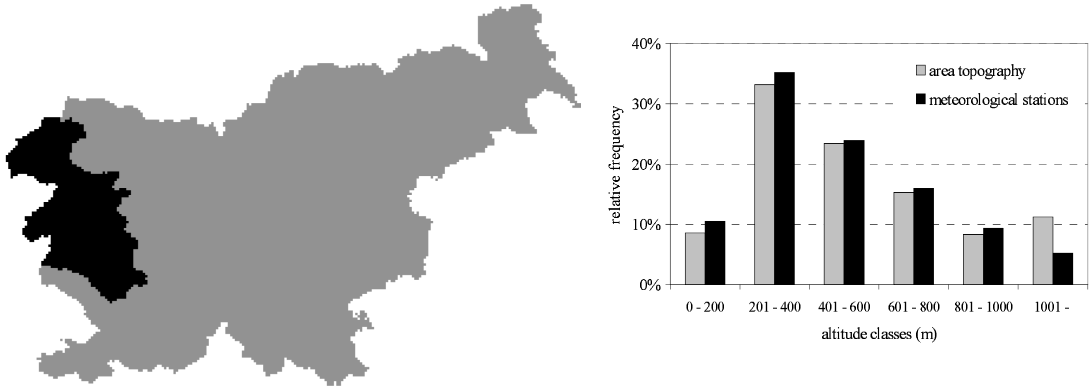

Figure 5: The North-West Slovenia, drainage basin of the river Soča (left). The distribution of the topography (grid resolution 100 m) and the meteorological stations with altitude for Slovenia (right).

3 The spatial interpolation

3.1 The universal kriging theory

The universal kriging, the optimal linear unbiased predictor, considering linear trend with altitude and longitude was used for MYP spatial interpolation. The mathematical model is:

$$U(\mathbf{s}) = \beta_0 + \beta_1 x_1(\mathbf{s}) + \beta_2 x_2(\mathbf{s}) + \delta(\mathbf{s}), \tag{1}$$

$U(\mathbf{s})$ is a random variable representing MYP at the location $\mathbf{s}$, $x_1(\mathbf{s})$ is longitude and $x_2(\mathbf{s})$ is the altitude at location $\mathbf{s}$, β_0, β_1 and β_2 are coefficients of the linear trend, $\delta(\mathbf{s})$ is an intrinsic stationary random process with existing variogram $2\gamma(\mathbf{h})$, where $\mathbf{h} = \mathbf{s}_i - \mathbf{s}_j$ is a distance vector between locations $\mathbf{s}_i$ and $\mathbf{s}_j$.

The random proces value on location $\mathbf{s}_0$, $U(\mathbf{s}_0)$ is:

$$U(\mathbf{s}_0) = \mathbf{x}^T(\mathbf{s}_0)\boldsymbol{\beta} + \delta(\mathbf{s}_0). \tag{2}$$

According to the universal kriging theory, the predicted value $\hat{U}(\mathbf{s}_0)$ can be expressed as a linear combination of the measured values $U(\mathbf{s}_i)$ on n locations:

$$\hat{U}(\mathbf{s_0}) = \sum_{i=1}^{n} \lambda_i U(\mathbf{s}_i). \tag{3}$$

Linear coefficients λ_i, $i = 1,\ldots,n$ are calculated under the condition for uniformly unbiased predictor:

$$\boldsymbol{\lambda}^T \mathbf{X} = \mathbf{x}^T(\mathbf{s}_0). \tag{4}$$

and under the constraint of minimal kriging variance $\sigma^2(\mathbf{s}_0)$ at location $\mathbf{s}_0$:

$$\sigma^2(\mathbf{s}_0) = E\left(U(\mathbf{s}_0) - \hat{U}(\mathbf{s}_0)\right)^2. \tag{5}$$

It can be shown that for calculation of λ-s, a variogram which is a measure of spatial continuity of the data, is required. The details of this theory can be found in Cressie (1993).

3.2 Variogram model estimation

The spatial continuity of the data was examined on the basis of our variogram analysis. The variogram $2\gamma(\mathbf{h})$ is defined as:

$$\mathrm{var}(U(\mathbf{s}+\mathbf{h})-U(\mathbf{s}))=2\gamma(\mathbf{h}). \quad (6)$$

The $\gamma(\mathbf{h})$ is often called a semivariogram. The estimates for the variogram $2\hat{\gamma}(\mathbf{h})$ are calculated by the equation:

$$2\hat{\gamma}(\mathbf{h})=\frac{1}{|N(\mathbf{h})|}\sum_{N(\mathbf{h})}(U(\mathbf{s}_i)-U(\mathbf{s}_j))^2, \quad (7)$$

$N(\mathbf{h})$ denotes the set of pairs of locations at distance $\mathbf{h}$ and $|N(\mathbf{h})|$ denotes the number of corresponding pairs of locations.

Firstly, the semivariogram values for the detrended MYP were calculated (Fig. 6) and the spherical variogram model with nugget effect was fitted:

$$\gamma(\mathbf{s}_i-\mathbf{s}_j)=c_0 Nug(\mathbf{0})+c_1 Sph(a) \quad (8)$$

$$Sph(a)=\begin{cases}\left(\frac{3(\mathbf{h})}{2a}-\frac{1}{2}\left(\frac{(\mathbf{h})}{a}\right)^3\right); 0\le\mathbf{h}\le a\\ 1; h>a\end{cases} \quad (9)$$

$$Nug(\mathbf{0})=\begin{cases}0, \mathbf{h}=0\\ 1, \mathbf{h}>0\end{cases} \quad (10)$$

The results indicate that the variogram range for Slovenia is around 40 km (Fig. 6). This value represents the longest distance with correlated values of the random process $U(\mathbf{s})$. The range 40 km seems acceptable for Slovenia according to its relief and climate diversity. The nugget effect presents the semivariance of MYP on microscale.

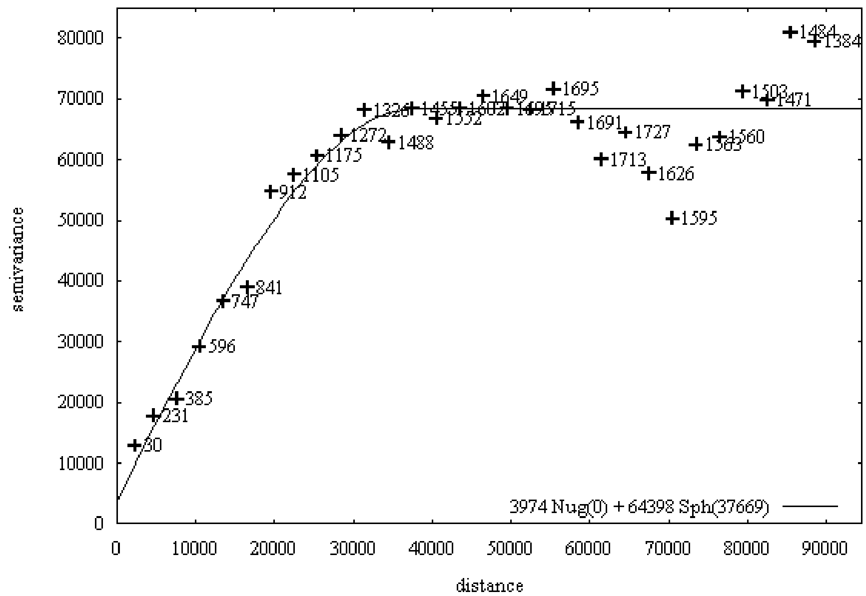

Figure 6: The isotropic variogram model for $\delta(\mathbf{s})$ made by Gstat 2.0g program. Numbers at crosses present $|N(\mathbf{h})|$. The distances are expressed in meters and the semivariances (semivariogram values) in mm^2.

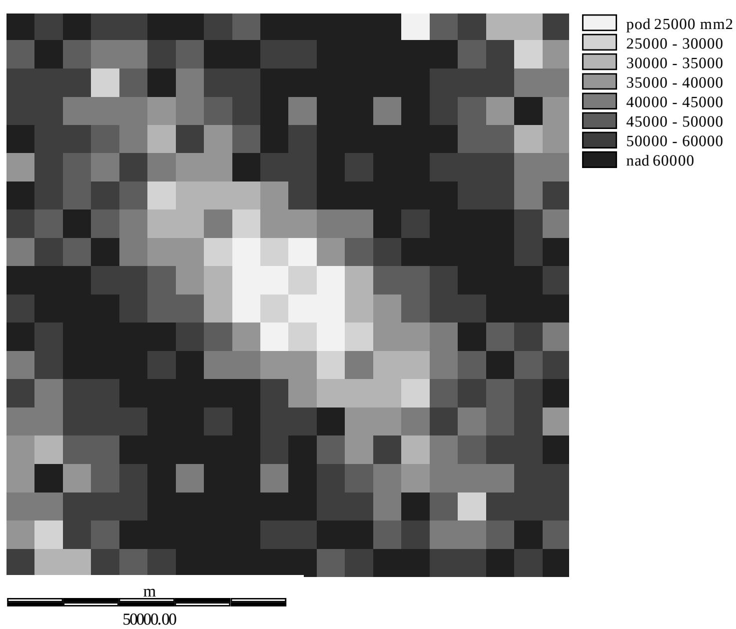

Figure 7: The variogram map for MYP.

The semivariogram on Fig. 6 is based on the assumption that the spatial continuity of $U(\mathbf{s})$ is the same in all directions. This assumption is not acceptable for MYP in Slovenia. The anisotropy of the random process was examined through the variogram map (Fig. 7), which graphically presents the semivariogram values for different distances and for different directions. The variogram map shows that the spatial continuity in the direction North-West to South-East (NW-SE) is much stronger than in the opposite direction (NE-SW). Therefore, we calculated two semivariograms, for each direction separately. The variogram range for the NW-SE direction is around 80 km, and for the perpendicular direction around 30 km. The range ratio is 0.36. This geometric anisotropy was taken into account in further analyses.

3.3 The influential surrounding

The influential surrounding is the area around the prediction location, from which the measured values are taken into account by kriging. In our case, the influential surrounding is the ellipse with the longer axis in the direction NW-SE. The variograms indicate that the upper limit for the longer axis is 160 km, and for the shorter axis 60 km.

We carried out several universal kriging calculations to find the appropriate dimension of the ellipse (Fig. 8 and 9). The cross-validation technique was used for the evaluation of the kriging results. By cross-validation, each meteorological station is eliminated out of the kriging calculation separately, and the kriging prediction is made for the eliminated location on the basis of the remaining locations.

The size of the influential surrounding has an impact on the kriging predictions and on their kriging variances. Fig. 8 presents the relative residuals (difference between the measured and the predicted MYP divided by the measured MYP) for different sizes of ellipse, and Fig. 9 the corresponding kriging variances. In our case, the relative residuals and the kriging variances do not depend on the ellipse size, except for some locations. Further analysis showed that these meteorological stations are located near the Slovenian border. It turns out that the reasonable value for the ellipse longer axis is 80 km.

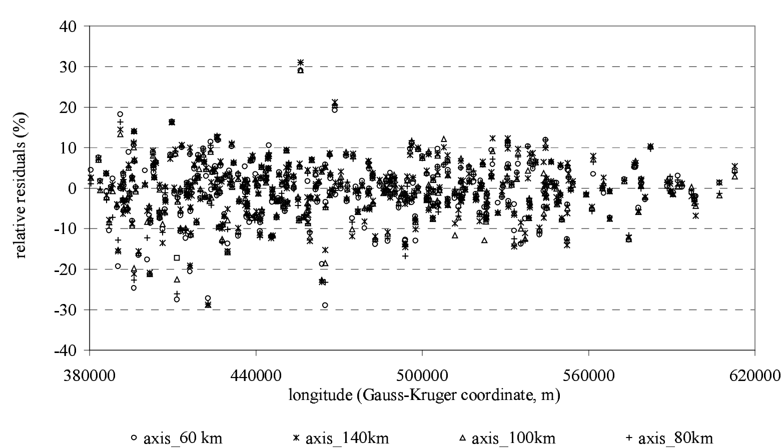

Figure 8: The relative residuals given by cross-validation using different influential surroundings (ellipses with the longer axis 60 km, 140 km, 100 km and 80 km).

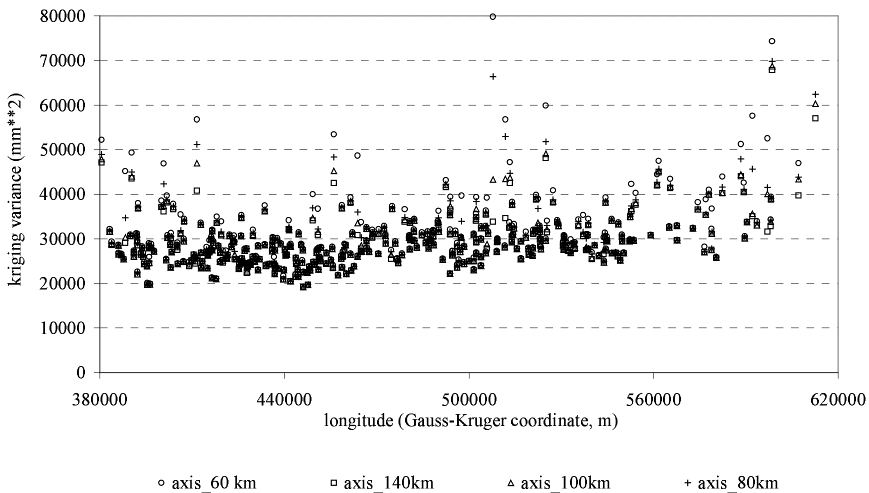

Figure 9: The kriging variances given by cross-validation using different influential surroundings (ellipses with the longer axis 60 km, 140 km, 100 km and 80 km).

3.4 Mapping MYP

Fig. 10 presents the map of the kriging predictions of MYP using anisotropic variogram and the ellipse with longer axis 80 km as influential surrounding. The kriging predictions were calculated for a regular grid with 1 km resolution. The universal kriging results were encouraging and similar to the subjectively determined map (Zupančič, 1996).

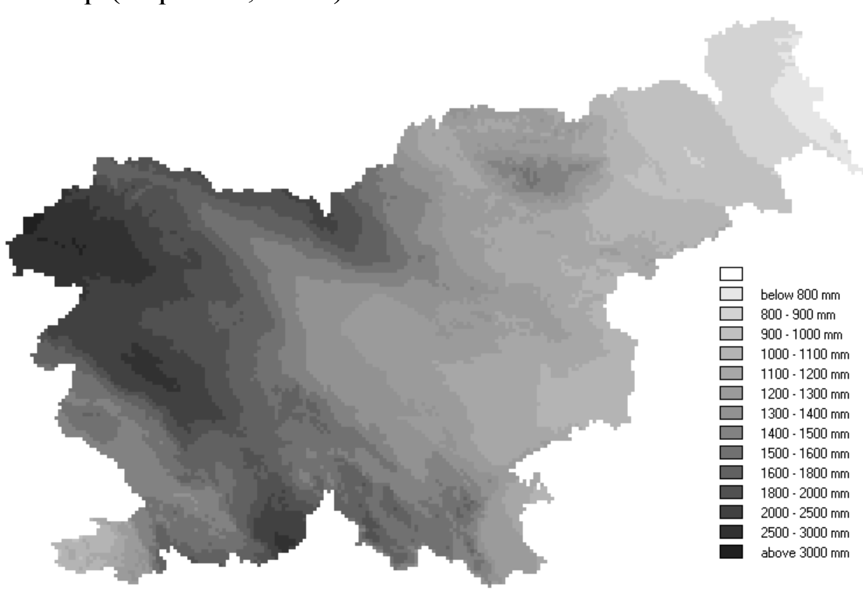

Figure 10: The map of MYP for Slovenia.

3.5 The quality of the kriging predictions

The quality of the kriging predictions in Fig. 10 was examined on the basis of cross-validation results and on the basis of a kriging variances analysis. Fig. 11 presents the histogram of the residuals obtained by cross-validation. Around 88% of the relative residuals are in the range -10% to 10%. Fig. 12 shows the locations which have relative residuals smaller than -15% or greater than 15%, and the locations which have kriging variances greater than 36000 mm^2 (90-th percentile). The relative residuals and the kriging variances are the greatest for the locations near the Slovenian border and for the locations on the specific mountainous regions. In general, the relative residuals for the eastern part of Slovenia are smaller then for the western mountainous part (Fig. 8). This is due to the lack of meteorological stations in the mountainous region where the spatial variability of MYP is high.

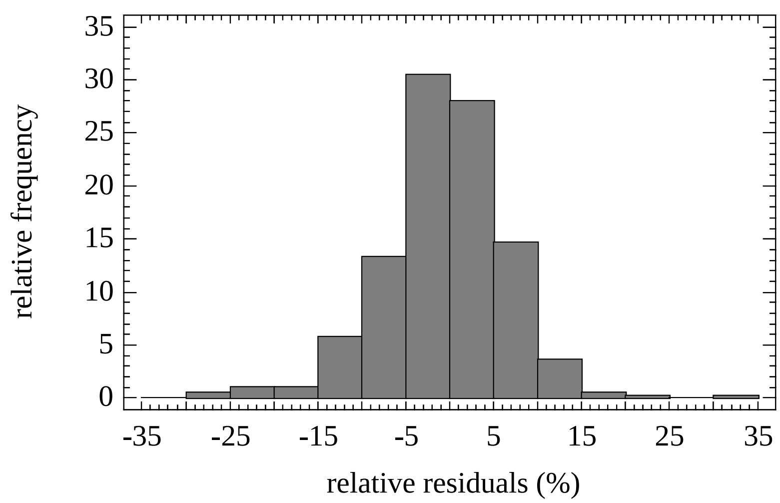

Figure 11: The distribution of relative residuals.

Fig. 13 presents the map of kriging variances for the predictions on Fig. 10. It is evident that the kriging variance is small in the neighbourhood of the measurement locations and large for the regions with small number of measurement locations as well as for some regions near Slovenian border.

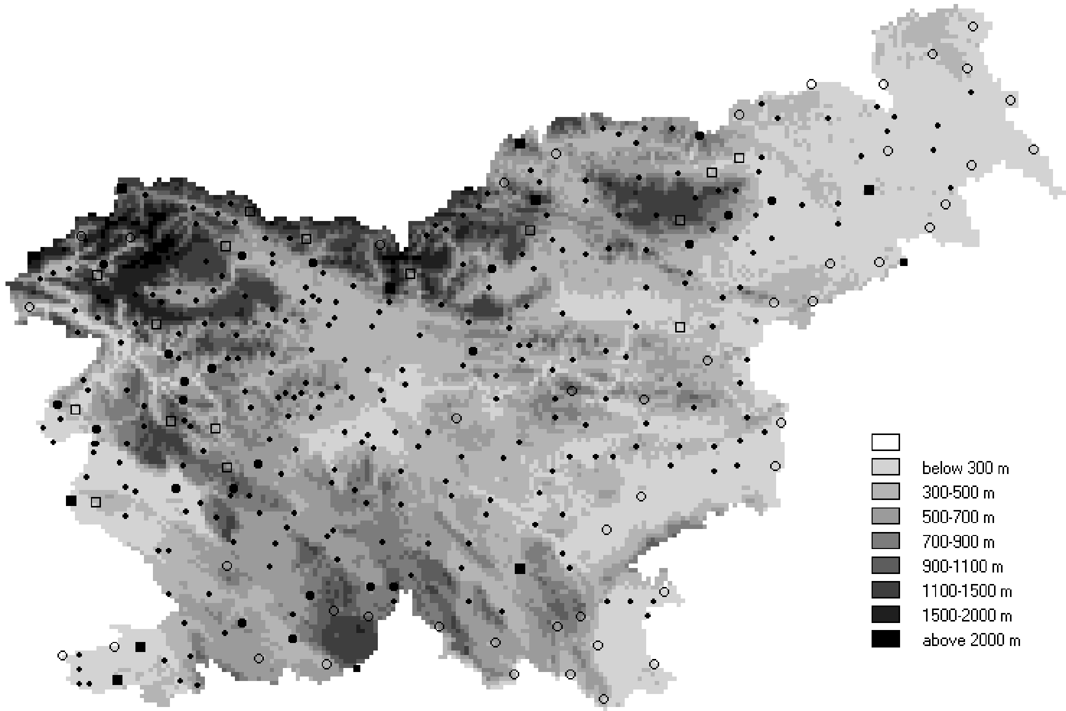

Figure 12: The type of locations according to the value of the relative residual and kriging variance: **large black circle -** residuals smaller than -15%; **transparent square -** residuals larger than +15%; **transparent circle -** kriging variance larger than 90-th percentile; **large black square** - residuals smaller than -15% and kriging variance larger than 90-th percentile at the same time; **small black square** - - residuals larger than +15% and kriging variance larger than 90-th percentile at the same time; **small black circle** - the others.

.

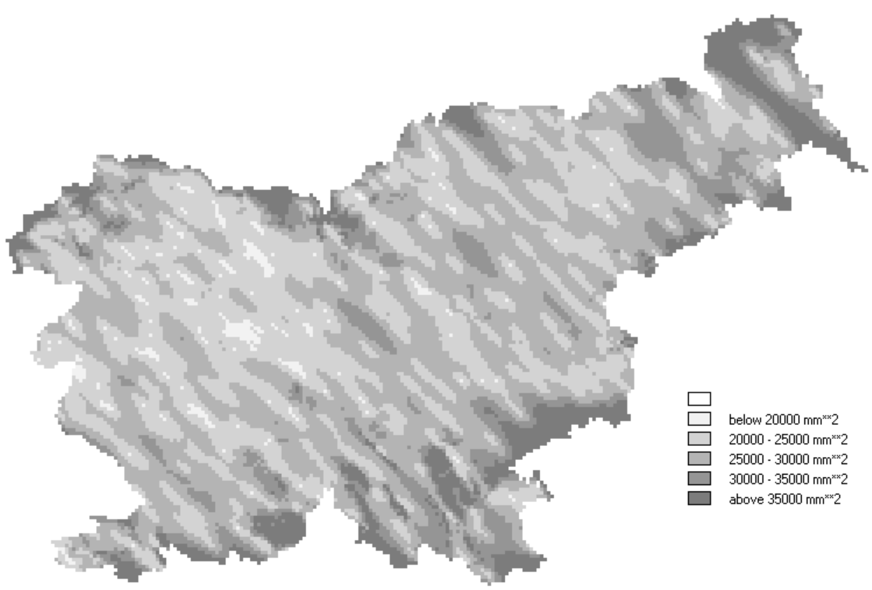

Figure 13: The map of kriging variances for map of MYP (Fig.12).

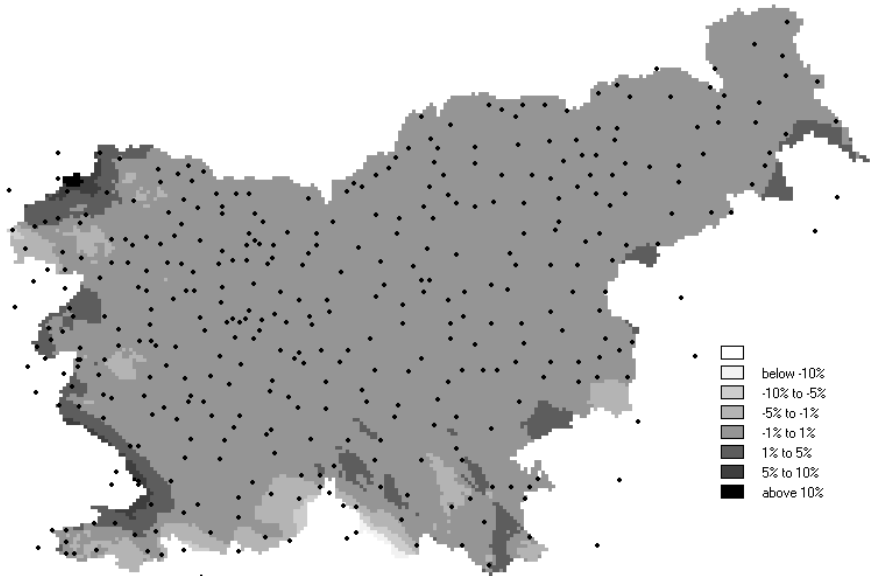

Figure 14: Relative differences between kriging predictions of MYP if only Slovenian data and if also some Croatian and Italian data were used.

In further work, we examined the quality of predictions near the Slovenian border. Therefore, some Croatian and Italian data were also included into the kriging calculations (Fig. 1). Fig. 14 and 15 present the relative differences between kriging predictions of MYP and kriging variances if only Slovenian data were used and if also the additional data were used. The relative differences in predictions are for some areas near the Slovenian border even greater than 10%, and the kriging variances are even smaller for more than 20%.

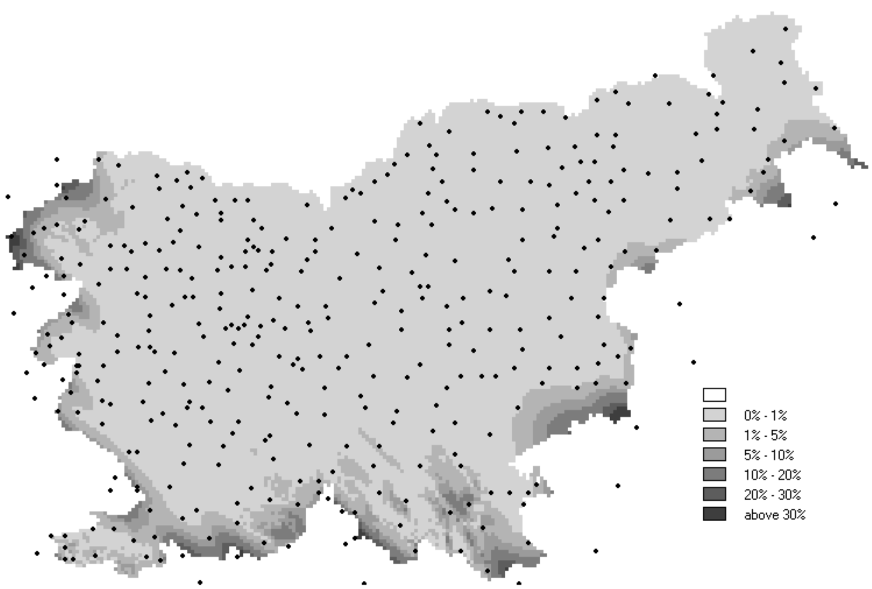

Figure 15: Relative differences between kriging variances if Slovenian data were used and if also some Croatian and Italian data were used.

4 Conclusions

The universal kriging results were encouraging and comparable with the subjectively obtained map (Zupančič, 1996).

The analysis of residuals calculated by cross-validation, showed that the results for the eastern part of Slovenia are better than for the western mountainous part. This is due to the lack of observational stations in the mountainous part of Slovenia where the spatial variability of MYP is high.

The analysis of kriging variances showed that the quality of predictions near the Slovenian border is considerably lower than in the inner region.

Some data from Italian and Croatian meteorological stations were added and the change in predictions and kriging variances on the locations near the Slovenian border were analysed. The results showed that we should include in spatial interpolation as many measurements across the Slovenian border as possible.

References

[1] Bergant, K. and Kajfež-Bogataj, L. (1998): Influence of different data densities on the mapping of agroclimatological parameters - case of precipitation. *Research Reports*, **71**, 91-97. Biotechnical Faculty, University of Ljubljana.

[2] Cressie, N.A.C. (1993): *Statistics for Spatial Data*, Revised Edition, Wiley Series in Probability and Mathematical Statistics. New York: Wiley.

[3] Eastman, J.R. (1997): *Idrisi for Windows, User's Guide.* Version 2.0, Revision 5, Clark labs for Cartographic Technology and Geographics Analysis.

[4] Gandin, L.S. (1963): *Objective Analysis of Meteorological Fields*. Gidrometeorologicheskoe Izdatel'stvo (GIMIZ), Leningrad (translated by Israel Program for Scientific translations, Jerusalem, 1965).

[5] Geodetska uprava RS. *Digitalni model reliefa*.

[6] Kastelec, D. (1999): *Analiza višine novozapadlega snega v Sloveniji*. Hidrometeorološki zavod republike Slovenije.

[7] Kolbezen, M. and Pristov, J. (1997): *Površinski vodotoki in vodna bilanca Slovenije*. Hidrometeorološki zavod RS, Ljubljana.

[8] Matheron, G. (1963): Principles of geostatistics. *Economic Geology*, **58**, 1246-1266.

[9] Pebesma, E.J. (1998): *Gstat User's Manual*. http://www.geog.uu.nl/gstat/

[10] Vrhovec, T. and Sluga G. (1998): La percentuale in acqua del manto nevoso sulle Alpi Giulie. *Neve e Valanghe*, **35**, 6-11.

[11] Zupančič, B. (1996): Mean yearly precipitation. In T. Cegnar (Ed.): *Climate of Slovenia*. Hidrometeorološki zavod RS, Ljubljana, 51.

[12] Zupančič, B. (1998): Padavine. In *Geografski atlas Slovenije. Država v prostoru in času*. DZS, Inštitut za gegrafijo, Geografski inštitut Antona Melika ZRC SAZU, Ljubljana, 98-99.