



CZECH TECHNICAL UNIVERSITY IN PRAGUE

**Faculty of Civil Engineering
Department of Mapping and Cartography**

Spatial Modeling of Climate

Master Thesis

Study Programme: Geodesy and Cartography
Branch of study: Geodesy and Cartography

Thesis advisor: Ing. Jiří Cajthaml, Ph.D.

Arnošt Müller

Prague 2010

Declaration

I declare that I have completed this Master thesis called ‘Spatial Modeling of Climate’ by myself and have cited all used references and literature according to the ‘Methodical Directive About Ethical Preparation of Final Academic Works’ issued by the Czech Technical University in Prague.

Prohlašuji, že jsem předloženou práci vypracoval samostatně a že jsem uvedl veškeré použité informační zdroje v souladu s Metodickým pokynem o etické přípravě vysokoškolských závěrečných prací.

In Prague on

.....
Arnošt Müller

Acknowledgement

I would like to express my appreciation to my thesis advisor Ing. Jiří Cajthaml, Ph.D. from the Department of Mapping and Cartography at the Czech Technical University, for his advice, help with the MapServer, and friendly communication. I would like to thank my research project advisor Dr. Bertram Ostendorf from the School of Earth and Environmental Sciences at the University of Adelaide for his support, advice, and useful comments. I am very grateful to editors Aisha Salazar and Maureen Robinson for the Abstract and Conclusion, Kelly Arbon for chapter 5, and Katherine Widomski for the remainder. I am also thankful to Tomas Zejdlik from the radar department as well as Mgr. Michal Zak, Ph.D. from the Department of Climatology of the CHMI for providing me annual climatological and monthly radar data.

Abstract

This thesis presents mapping of two climate variables: mean air temperature and precipitation over the Czech Republic using spatial interpolation of climatological data. Geographical information is used to predict the climatological variables through a regression relationship. Geographic Information System ArcGIS provides the means to produce climatological surfaces for the entire Czech Republic between the years 1998–2009. The temporal resolution of the maps is annual and the spatial resolution is 90 m.

Climatological data were obtained from a small subset of only 22 meteorological stations available online at the Czech Hydrometeorological Institute websites. More detailed data from 132 stations – were obtained only for the year 2008 and were used for independent tests of accuracy. Monthly radar data were also obtained for the year 2008 for testing purposes. Chapter 5 presents a separate study about using radar in precipitation modeling.

Independent variables were derived from a 90 m digital elevation model (DEM) acquired by the Space Shuttle Radar Topographic Mapping Mission (SRTM). Significant independent variables for predicting mean temperature were altitude and latitude. Four out of 28 independent variables were selected for precipitation: altitude, longitude, latitude and variable ZxW25 depicting topographical barriers in a specific westerly direction and radius of 25 km. The final regression models show a high degree of explained variance for mean air temperature ($R^2 = 0.90\text{--}0.97$, and root-mean-squared error RMSE = 0.40 °C) and a moderate degree of explained variance for precipitation ($R^2 = 0.78\text{--}0.92$, RMSE = 106 mm). This study provides a deeper analysis of the influence of topography on mean temperature and precipitation in the Czech Republic.

Annual variation in mean air temperature and precipitation as well as deviation from 40-year normals was visualized in an animated map. Final climatological surfaces were published interactively using UMN MapServer and ka-Map template.

Abstrakt

Předkládaná diplomová práce se zabývá mapováním dvou klimatických prvků v České republice: roční průměrné teploty vzduchu a průměrného ročního úhrnu srážek. Geografický informační systém ArcGIS slouží jako nástroj pro vytvoření klimatických vrstev pro celou Českou Republiku v prostorovém rozlišení 90 m v období 1998–2009.

Klimatické veličiny jsou měřeny na meteorologických stanicích, které jsou volně dostupné na webu Českého hydrometeorologického úřadu (ČHMÚ), a kterých má tato práce k dispozici pouhých 22. Úřadem byla poskytnuta data ze 132 stanic, avšak pouze pro rok 2008. Tato data byla použita k nezávislému určení přesnosti výsledných klimatických veličin. Dále byly od ČHMÚ získány měsíční sumy srážkových úhrnů měřené radarově také pro rok 2008. Využitím radarových měření k mapování ročního úhrnu srážek se zabývá samostatná kapitola 5.

Bodová klimatická měření byla interpolována do prostoru pomocí vícenásobné regresní analýzy a korigována lokální interpolací reziduí. Nezávislé veličiny byly odvozeny z globálního digitálního modelu terénu (DEM) Space Shuttle Radar Topographic Mapping Mission (SRTM). Jako nezávislé veličiny v teplotním regresním modelu byly na základě statistických testů vybrány dvě veličiny, nadmořská výška a zeměpisná šířka. V případě srážek byly z celkového počtu 28 nezávislých veličin vybrány čtyři: nadmořská výška, zeměpisná šířka a délka, a veličina $ZxW25$ charakterizující terénní překážky ve směru na západ do vzdálenosti 25 km. Použitý regresní model průměrné teploty vykazuje silný vztah s nezávislými veličinami ($R^2 = 0.90\text{--}0.97$ a střední chyba $RMSE = 0.40$ °C) a model srážek uspokojivý vztah s nezávislými veličinami ($R^2 = 0.78\text{--}0.92$, $RMSE = 106$ mm). Díky použité metodě umožňuje tato práce hlubší pohled na závislost klimatu na topografii v České Republice.

Rozdíly průměrných teplot vzduchu a srážek mezi jednotlivými roky, a také jejich odchylek od čtyřicetiletého normálu, jsou nejlépe patrné z animovaných map. Výsledné klimatické vrstvy byly zveřejněny interaktivně pomocí UMN MapServer a šablony ka-Map.

Key Words

Annual mean air temperature, annual precipitation, rainfall, radar, climate, Czech Republic, GIS, ArcGIS, regression analysis, multivariate regression, DEM, SRTM, animated map, UMN MapServer, ka-Map.

Klíčová slova

Průměrná roční teplota vzduchu, průměrný roční úhrn srážek, radar, klima, Česká republika, GIS, ArcGIS, regresní analýza, víceproměnná regrese, DMT, SRTM, animovaná mapa, UMN MapServer, ka-Map.

Table of Contents

1	Introduction	1
1.1	Objectives.....	2
2	Literature Review	3
2.1	Mean Air Temperatures.....	5
2.2	Precipitation	6
2.3	Remote Sensing.....	7
3	Materials.....	8
3.1	Study Area.....	8
3.2	Climate Characteristics of the Czech Republic.....	8
3.3	Data Sources.....	9
3.3.1	Climatological Data.....	9
3.3.1.1	Missing Data	11
3.3.2	DEM	12
3.3.3	Variables Derived from DEM	14
3.3.3.1	Altitude (ALT).....	14
3.3.3.2	Slope (SLP)	14
3.3.3.3	Aspect (ASP).....	16
3.3.3.4	Curvature (CRV)	16
3.3.3.5	Solar Radiation (SOLRAD).....	16
3.3.3.6	Maximum Altitude in Wedge (Zx)	18
3.3.3.7	X0, Y0.....	20
3.3.4	DBF files.....	20
3.3.5	ArcČR 500 v. 2.0a	20
4	Methods.....	21
4.1	Global Methods	21
4.2	Local Methods.....	22
4.3	Geostatistical Methods.....	23
4.4	Combined Methods.....	23
4.4.1	Multiple Regression with Residual Correction	24
4.4.2	R ² and Adjusted R ²	25
4.4.3	Fisher's F-test.....	25
4.4.4	Student's t-test	26
4.4.5	Protocols of Regression from Geoda.....	26
4.5	Limitations	27
4.6	Delimitations	27
5	Using Radar-derived Rainfall in Precipitation Modeling in CR.....	28
5.1	Introduction.....	28
5.1.1	Aims and Objectives.....	29
5.2	Theoretical Framework.....	29
5.2.1	Methodology	30
5.2.2	Data Sources.....	30

5.2.2.1	Radar Data.....	30
5.2.2.2	Specific Independent Variables	34
5.2.3	Atmospheric Data	37
5.3	Analysis.....	37
5.3.1	Regression Analysis.....	37
5.3.2	Residual Analysis	39
5.3.2.1	Residuals	39
5.3.2.2	Residual Plots.....	40
5.4	Findings.....	46
5.4.1	Discussion	47
5.5	Significance.....	48
5.5.1	Suggestions for Future	48
6	Analysis	49
6.1	Linear Relationships	49
6.2	Backward Stepwise Approach.....	52
6.3	Regression Models	54
6.4	Corrector Maps.....	54
6.5	MT and P Maps	57
6.6	Deviation from Normal.....	57
6.7	Visualization.....	57
7	Accuracy.....	58
7.1	Statistical Criteria	58
7.2	Precision of DEM.....	59
7.3	Numerical Assessment.....	59
7.4	Spatial Assessment	61
8	Results	65
8.1	Spatial Pattern of MT and P	67
8.2	Deviations from Normal	67
9	MapServer.....	69
9.1	ka-Map.....	70
9.2	Mapfile.....	72
10	Conclusions	73
10.1	Future Work.....	74
	Bibliography	I
	List of Abbreviations	V
	List of Figures	VI
	List of Tables.....	VIII
	List of Appendices.....	IX

1 Introduction

Spatial modeling of climate variables is of wide interest since many other environmental variables depend on climate. Climate maps are needed in disciplines related to Earth Sciences such as hydrology, forest management, agriculture, ecology, urban environments, energy, and climate change. [31, 37] These environmental disciplines use spatial information of climate as a basis for understanding the processes they study. Climate does not only influence the environmental disciplines, but also affects whole economies of many countries and the lives of their citizens, mainly in planning and management of all logistics involved in transportation of people and goods. [16] Therefore, the monitoring and analysis of climate and weather conditions can be beneficial to many areas.

There is a need for detailed continuous climate surfaces in digital form. [38] Daly et al. introduce the discipline called geospatial climatology, ‘the study of the spatial patterns of climate and their relationships with geographic features’. [15] As the climate measurements are typically point source in nature, one of the challenges facing meteorology is the interpolation of point climatological phenomena across a wide spatial domain. [8] Accurate climatological data are collected at meteorological stations (in further text often referred to as ‘stations’), which are discrete point locations in space. Values at any other point must be derived from neighboring stations or from relationships with other variables. [31] The method of spreading discrete measurements over an area is called spatial interpolation. Various interpolation methods have been taking into account geographic information by using Geographic Information Systems (GIS).

Over the last decade, the use of GIS in a variety of applications involving the processing of climatological and meteorological data has rapidly increased. Climatological datasets can be displayed in GIS in a variety of ways, for example: rainfall measurements as point features, rain radar as rasters, and isolines as vectors. In weather forecasting, GIS has become a key management component in weather processing systems allowing instantaneous plotting, interpolation and animation of weather data. [8] GIS are no more used only as spatial visualization facility, but have evolved into powerful management tool used for capturing, modeling, and analyzing spatial data. Climatological and meteorological phenomena are naturally spatially variable and hence GIS represent a useful solution for the management of vast climate datasets.

As the climate and geography are closely related, one can study this relationship and then use it to model the climate. The geographic information often comes from digital elevation model (DEM). DEMs have enabled good estimates of an area's climatology without the need of extensive weather records. [8] DEMs were traditionally derived using land surveying techniques but are now determined remotely using radar, e.g. the Shuttle Radar Topography Mission (SRTM). SRTM captured topographic data for 80% of the Earth's land surface at a resolution of up to 30m. [8] Remotely sensed DEM are manipulated into GIS and can then provide a baseline for climatological datasets. Remote sensing has also provided other inputs in climate modeling, such as NDVI (normalized vegetation index) and LST (land surface temperature), which have been used as independent variables in regression models to enhance local differences. [13] Remote sensing enables the acquisition and calibration of comprehensive datasets whereas GIS provides a standard means to display, overlay, and combine the data for analysis, which makes remote sensing and GIS intimately related. For example, studies regarding the urban heat island phenomenon integrating remote sensing data with GIS assess how temperature is spatially influenced by land use. [8]

1.1 Objectives

This study has several objectives: Firstly, to select the most suitable mapping and spatial interpolation method that takes into account geographical variables, and uses GIS techniques to obtain maps of different climatological variables over a relatively large area of the whole Czech Republic (CR). This study intends to explore the utility of regression analysis, which allows the studying of topographic influence on spatial climate patterns. The next sub-objective is derived from the interpolation technique applied to create the cartography. It is to investigate and select meaningful geographic factors, which play an important role in annual mean air temperature and precipitation modeling in the area of CR. The second objective is to evaluate the use of radar data in precipitation modeling. The third and last is to publish the results in an interactive manner on the Internet using a MapServer.

Resulting maps will show the annual change in mean air temperature and rainfall patterns over the past twelve years and more importantly, the recent years will be compared to the long-term 30-year mean (normal) for both climate variables, which can provide an insight into the problem of the climate change.

2 Literature Review

Chapman and Thornes [8] review the role of GIS in climatology and meteorology by discussing methods used to derive and refine factual climatological applications in various disciplines: In agriculture, there is massive potential for ‘agroclimatic’ modeling in order to predict yield by combining maps of soils, nutrients, climate, weather stress, fertility, etc. In hydrology, measurements of precipitation are an obvious starting point for many hydrometeorological models including floods. The opposite extreme to flooding is modeling the distribution of droughts, which can be used for fire alert systems. In forestry, GIS is used to model and monitor the spread of forest fires via the combination of climatological and remotely sensed imagery. In ecology, biodiversity modeling has been successfully applied in studies such as distribution of plants with respect to rainfall. In urban environments, climatological data can provide information regarding pollution, but is mostly directed at studying the urban heat island phenomenon. In energy, temperature and humidity are the primary factors controlling energy demand. For example, GIS has been used to aid the locations of wind farms by modeling wind energy potential. [8]

All of the environmental disciplines above are potentially subjects to the impact of climate change. Hence, the assessment of the effects of climate change is truly a multidisciplinary exercise of which GIS provides a fundamental unifying role. GIS has become a processing and visualization tool for climate models. Such models can predict the global impacts of hypothesized climate change scenarios. [8]

The broadest group of studies regarding temperature and precipitation modeling starts from the premise that altitude and elevation explain most of the spatial variability but proceed to evaluate other factors such as terrain attributes (aspect and morphology of the relief), atmospheric factors (humidity and wind) and maritime factors (distance from coast and effects of sea currents). [11] The second group of studies addresses the maximum efficiency of statistical and geostatistical interpolation techniques. For example, Hutchinson [26] develops a method called thin-plate smoothing splines which is then used in works of Lapen and Hayhoe [29] or Hijmans et al. [25].

Many studies dealing with modeling of climate are focusing on mountainous regions [29, 31, 46, 50], where climate variables are difficult to predict due to the topographic complexity which generates microclimate environments. [15] Squires and McNab [45] as well as Marquinez [31] claim that interpolation using traditional techniques (IDW, kriging, etc.) is not accurate enough especially in mountainous regions.

Vincente-Serrano et al. [50] provided a comparative analysis of interpolation methods applied to annual precipitation and temperature. The study was carried out in the north east of Spain in a valley where geographic and spatial climate diversity is significant. Interpolation methods tested include: global interpolators (trend surfaces and regression models), local interpolators (Thiessen polygons, inverse distance weighting, splines), geostatistical methods (simple kriging, ordinary kriging, co-kriging, etc.) and mixed methods (combined global, local and geostatistical methods). Vincente-Serrano et al. obtained the best results using geostatistical methods and a regression-based model. The authors concluded their study by stressing out the importance of testing various interpolation methods before the most appropriate scheme for a given area and climatic variable is selected.

Goodale et al. [20] interpolated temperature and precipitation in Ireland by means of polynomial regression, using latitude, longitude, and altitude derived from a 1km DEM as predictors. Guler et al. [22] used a simple linear regression to assess spatial climatic layers at finer spatial resolution of 250m in Samsun, Turkey. Gyalistras [23] developed a 50-year (1951-2000) dataset for Switzerland. His approach for modeling mean monthly temperature and precipitation is a method called AURELHY (analysis using relief for hydrometeorological applications) and is based on principal component analysis and 5 km DEM. Unlike other studies, he did not consider mean values, but rather presents the amplitudes of climatic variables. Brown and Comrie modeled winter temperature and precipitation in Arizona and New Mexico, USA for the period 1961–1990. They used regression models at 1 km resolution and kriging or IDW interpolation to account for model residuals. [6]

While most of the studies focus on one country or one region within one country, Agnew and Palutikof [1] constructed maps of mean seasonal temperature and precipitation for the whole Mediterranean on a grid of 1km spatial resolution. Their regression-based approach included longitude, latitude, elevation, distance and direction to the nearest coast, slope, aspect, and the ratio of land to sea. Hijmans et al. [25] created global monthly precipitation and temperature surfaces at 1 km resolution.

In the CR, the Climate Atlas of Czechia was published in 2007 by the Czech Hydrometeorological Institute (CHMI) in cooperation with experts from the Department of Geomatics at the Palacky University in Olomouc. The Atlas is a paper book compilation comprising over 300 maps and presents results of long-term (1961–2000) climatological

measurements divided into 11 sections. Horizontal resolution of the maps is 500 m. The authors use a regression approach to create so called ‘fictive stations’ in order to have sufficient spatial density of observations, which they then interpolate with the IDW or kriging interpolation. They also consider linear regression for variables significantly dependent on elevation. Mean temperature for example was calculated from 311 stations. [28] More information about data processing in the Climate Atlas of Czechia can be found in [10]. Authors of the Climate Atlas of Czechia mention a warming trend of about 0.03 °C per year based on annual average temperature between 1961–2000. [48] Although the Atlas is so complex and well presented, it is only available in hard copies and at large scales from 1:1 000 000 to 1:5 000 000. One of this thesis’ objectives is to publish resulting digital maps on the internet, in a similar way as in the Digital Climatological Atlas of Spain at <http://opengis.uab.es/wms/iberia/mms/>. Another objective is to produce up to date maps for the recent years, which can be compared to the long-term normals.

The work of Ninyerola et al. [36-38] is seen as the key study regarding this thesis. The reasons why are further described in sections 2.1.1 and 2.1.2.

2.1 Mean Air Temperatures

An exhaustive comparison of eight different interpolation methods for temperature estimation can be found in a conference paper by F. C. Collins. [12]

Ninyerola et al. present a mapping technique of monthly and annual mean temperature over the entire Iberian Peninsula. The spatial interpolation method is based on multiple regression (statistical global analysis) with residual correction (local interpolation). This method allows exploring the relationship between climatic elements (such as mean air temperature) and independent geographical variables. As the most accurate interpolation technique they find the multiple regression with residuals interpolated using IDW. The authors not only test various interpolation techniques, but also compare global versus regional models. Ninyerola et al. claim that the best results for the entire peninsula are obtained when the model uses all stations together and not only regional subsets. The authors use altitude, latitude, continentality, solar radiation and a cloudiness factor as the independent geographical variables, which are elaborated from a 180 m resolution DEM. Ninyerola et al. describe this method as both empirical and statistical. Empirical, because it uses data obtained from stations for building and for validating the model and statistical, because it is based on a multiple regression analysis and its corresponding validation. The

authors believe they created a ‘very simple, useful and realistic model’.

Luis Rodríguez-Lado et al. [41] proved that multiple regression with residual correction for mapping air temperature of the State of São Paulo in Brazil at 0.5 km spatial resolution is an accurate method. Claps et al. [11] modeled monthly and annual mean temperatures based on 1 km DEM using linear multivariate regression analysis, a technique very similar to Ninyerola’s. Significant variables were altitude, latitude, distance from the sea and terrain concavity. For the interpolation of residuals, ordinary kriging was used.

2.2 Precipitation

Daly et al. [15] introduced in USA well established regression model called PRISM (parameter-elevation regression on independent slopes model), which had been developed at the Oregon State University. It is a knowledge-based system, which combines the strengths of both human expert and computer based statistical methods. The model uses a weighted climate-elevation regression function which stresses out the dominant influence of elevation on climate. Aspect and topographic exposure are also accounted for in the PRISM model. The maps including gridded datasets are available online at <http://www.prism.oregonstate.edu/>.

Basist et al. [3] developed statistical relationships between topography and mean annual precipitation for ten distinct mountainous regions around the world. Among the topographic variables studied, exposure to the prevailing winds was the single most important feature. Sun et al. [46] carried out a multivariate regression analysis combined with residuals correction to predict precipitation in the Daqing Mountains, Inner Mongolia in China. For the multivariate regression analysis, they used five topographic factors: altitude, slope, aspect, longitude and latitude in a third-order polynomial regression equation. The topographic factors for the area of about 9300 m² were derived from a 100 m DEM. To obtain residuals for correction, they used ordinary kriging interpolation method, which, according to the authors, did not significantly improve the prediction model. Their prediction model explained approximately 73% of spatial variability of precipitation.

Ninyerola et al. [37] developed monthly precipitation maps in the Iberian Peninsula at 200 m spatial resolution with the use of multiple regression combined with residual correction method.

Throughout the literature the following independent variables are found significant for precipitation prediction:

- Ninyerola [37] uses 5 variables to map monthly precipitation of the Iberian Peninsula: altitude, latitude, continentality (distance to the nearest coast), terrain curvature, and solar radiation.
- Vincente-Serrano [50] studies a large number of variables (many of them they created by calculating the mean value within radii of 2.5, 5, 10, and 25 km) in his comparative study of in the middle Ebro Valley in Spain, four of which are found significant predictors: maximum altitude within a wedge (in the North direction), latitude, latitude \times continentality, and altitude smoothed to 25 km.
- Basist [3] introduces exposure to the prevailing winds, exposure \times elevation and slope \times orientation.
- In addition to the variables already mentioned, Agnew and Palutikof [1] use direction of the nearest coast, a land to sea ratio, aspect, and slope.

2.3 Remote Sensing

Since imagery from weather satellites such as NOAA or MODIS as well as satellites such as Landsat is easy accessible, Cristobal et al. [13] investigated the combination of remote sensing and GIS data in air temperature modeling. They were working at regional scale (looking at Catalonia, Spain) at various temporal resolutions (daily instantaneous and mean, monthly, and annual). The authors found out that the most powerful remote sensing predictors for air temperature modeling are land surface temperature (LST) and normalized difference vegetation index (NDVI). Although the authors proved the significance of remote sensing predictors, the overall improvement with regard to the geographical model was modest, only 0.1 °C. Florio et al. [19] also reported an improvement of only 0.06 °C when remote sensing variables are introduced in the analysis. In respect to these small improvements, the idea of using remote sensing in this work was rejected.

3 Materials

3.1 Study Area

The Czech Republic (CR) spreads out in the temperate climate zone of the northern hemisphere. It is located in the center of Europe, between approximate longitude coordinates 12.09 °W and 18.86 °E (~500 km), and between latitude coordinates 48.55 °S and 51.06 °N (~285 km). CR shares borders with Germany, Austria, Slovakia, and Poland and occupies the area of almost 79 000 km². The altitudes range from 115 m to 1602 m above the Baltic Sea level. Official terrestrial cartography is in Krovak's conic projection, however all cartography has been projected into the UTM cylindrical projection. The area falls into the 33N zone. [35]

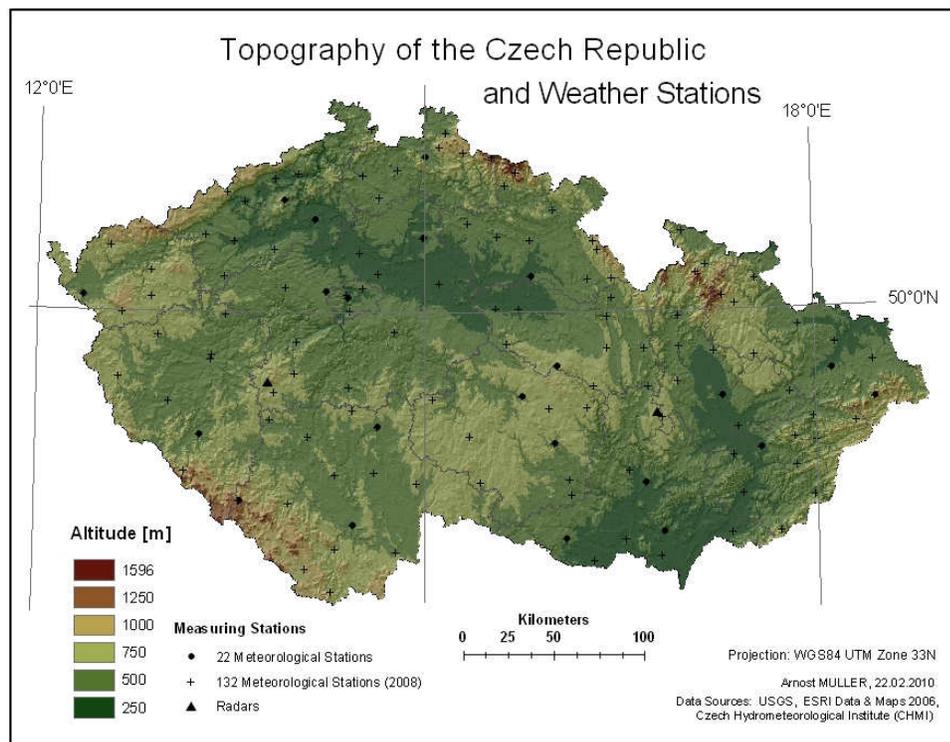


Figure 3.1: Study area with classified DEM and meteorological stations including two radars

3.2 Climate Characteristics of the Czech Republic

The natural environment of CR is characterized by a moderate, humid climate and four altering seasons. The complexity of the climate of the CR is related to its orography. Mean altitude is 430 m [47], however mean altitude obtained from the SRTM DEM is 450.7 m. The higher value derived from the DEM is probably caused by vegetation cover and urban settlement, which are reaching above the Earth's surface. Nevertheless, variation in altitude

has larger influence on both climate and weather in the Czech Republic. There are approximately 67% of total area in altitudes lower than 500 m, 32% between 500 and 1 000 m and only 1% higher than 1 000 m. [47]

The climate of the Czech Republic is influenced by both the continental and the ocean climate. The elongated shape along the 50th latitude results in a slight increase in continentality towards the East. Higher continental influence with east or north-east winds causes warmer, dry summers and stronger, colder winters. [47]

Precipitation is characterized by substantial spatial and temporal variability. During the winter precipitation is mainly linked to passing frontal systems and pressure lows, and is characterized by lower intensity and longer duration. During the summer, precipitation is usually shorter duration and higher intensity. The spatial differences of precipitation are amplified by orographic effects, such as increasing precipitation total with increasing elevation and the effects of exposure, where wind facing sides of the mountains receive higher rainfalls than leeward slopes. [48] According to the Climate Atlas of Czechia, south-westerly and westerly winds are the most frequent. In Moravia, the most frequent wind direction is northwest and in some cases north-south, because it is modified by the terrain. [48]

3.3 Data Sources

Data come from several sources: Czech Hydrometeorological Institute (CHMI), United States Geological Survey (USGS), Environmental Systems Research Institute (ESRI), and ArcDATA Praha.

3.3.1 Climatological Data

Temperatures, precipitation, and sunshine duration are three basic climatological variables measured at meteorological stations. Most of those stations in the Czech Republic are operated by employees of the CHMI or volunteers, who provide their data to this institute. The CHMI belongs to the Ministry of the Environment of the Czech Republic. CHMI operates 209 climatological stations (including 38 professional synoptic weather stations) and 585 precipitation stations (situation in January 2008). See a map of the climatological stations at http://www.chmi.cz/meteo/ok/images/st_cz.gif, where dark blue squares stand for professional stations, light blue for automatic, red for basic and khaki green for military operated stations.

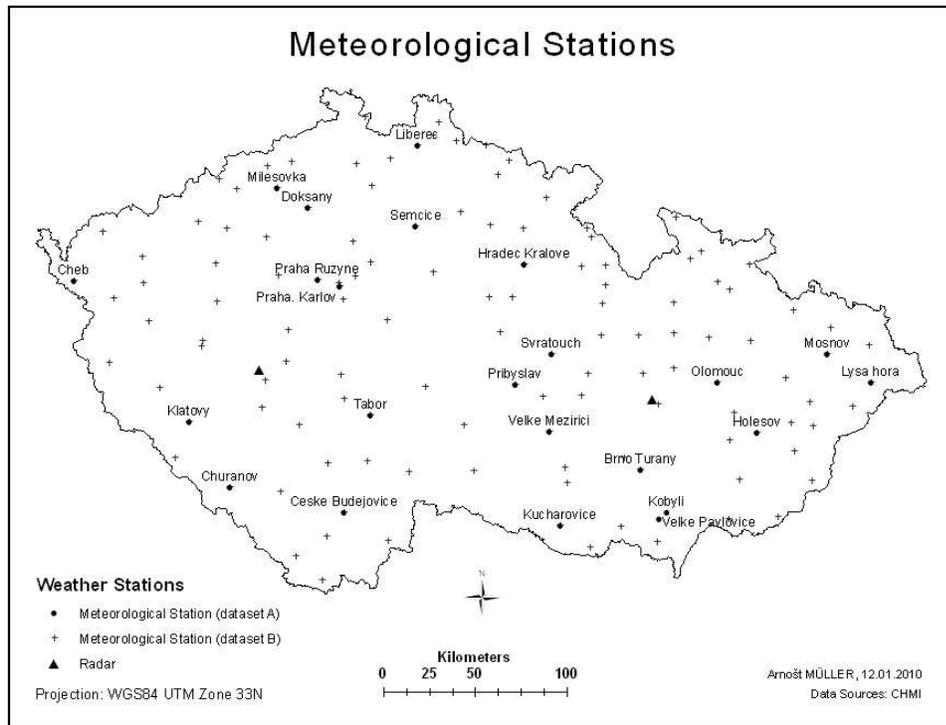


Figure 3.2: Meteorological stations (22 in dataset A and 132 in dataset B) and two radars

Three datasets were obtained from CHMI:

A) 22 meteorological stations, whose monthly and annual data for the years 1998–2009 and one long-term 30-year period (1961–1990) are published online at <http://www.chmi.cz/meteo/ok/infklima.html>. In fig. 3.3 below there are 23 stations because of the fact that station ‘Velke Pavlovice’ was replaced by ‘Kobyli’ in 2009. This fact is further discussed in the subsection 3.3.1.1.

Annual means for mean temperatures (MT) and precipitation (P) were copied into a spreadsheet. Tab 3.2 shows the distribution of the 22 stations with altitude. Just 22 stations for the area of almost 79 000km² and for about 1500 m variation in altitude are extremely sparse coverage (see tab. 3.2 and fig. 3.2 for location of the 22 stations). This dataset comprising 22 stations is further referred to as ‘dataset A’.

B) 135 meteorological stations with annual data for the year 2008. Although CHMI claims to provide data for academic purposes free of charge, after negotiations in September 2009 only the year 2008 with 135 stations was obtained in a form of a spreadsheet. Free of charge, but with no guarantee. This ‘dataset B’ is therefore used for the purposes of testing, evaluation and accuracy assessment.

C) Monthly radar rainfall sums for 2008. Refer to chapter 5 for detailed data description.

Table 3.2: Characteristics of 22 met. st

Mean ALT	449 m
Median of ALT	379 m
Maximum ALT	1 322 m
Nr. of stations higher than 500m	5
Nr. of stations higher than 1000m	2
Area per station	3 591 km ²

Table 3.3: Characteristics of 132 met. st.

Mean ALT	453 m
Median of ALT	401 m
Maximum ALT	1 328 m
Nr. of stations higher than 500m	48
Nr. of stations higher than 1000m	5
Area per station	598 km ²

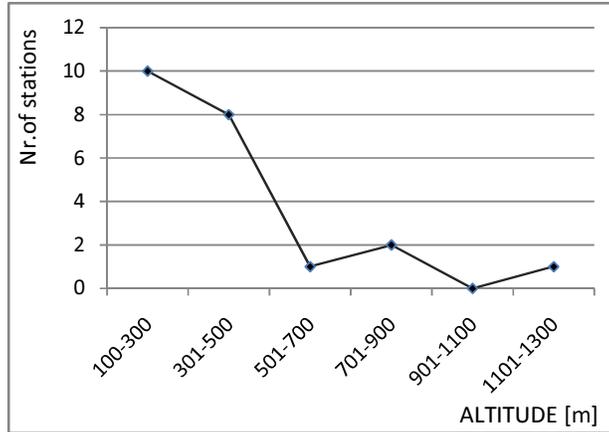


Table 3.4: Nr. of stations in altitude categories

Altitude Zone [m]	Nr. of Stations
100-300	9
301-500	8
501-700	1
701-900	2
901-1100	0
1101-1300	1
1301-1500	1

Figure 3.3: Station distribution with altitude (22 stations)

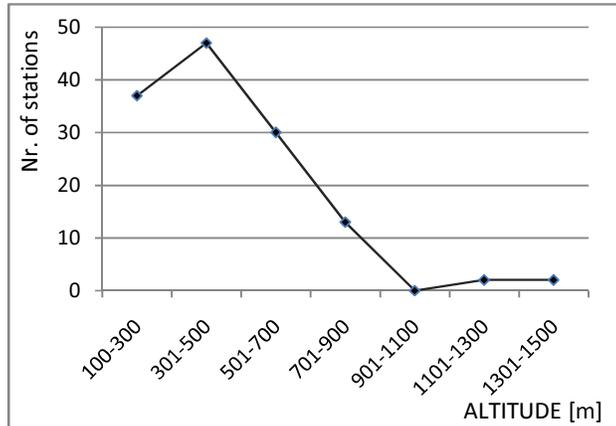


Table 3.5: Nr. of stations in altitude categories

Altitude Zone [m]	Nr. of Stations
100-300	37
301-500	47
501-700	30
701-900	13
901-1100	0
1101-1300	2
1301-1500	2

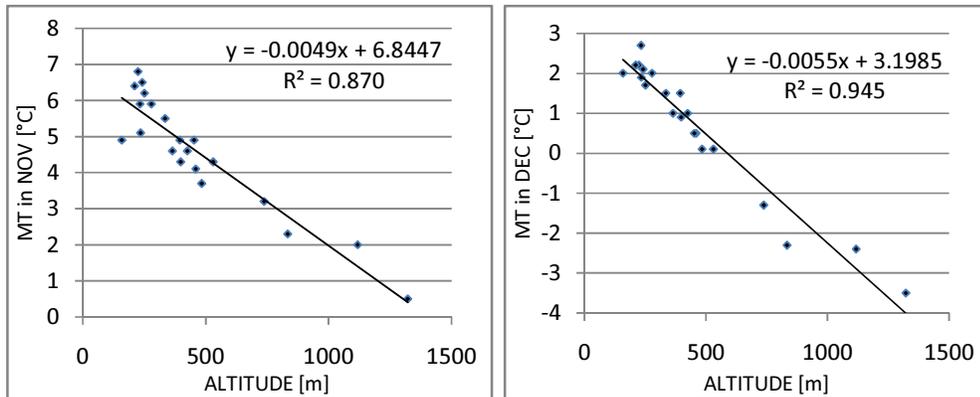
Figure 3.4 : Station distribution with altitude (132 stations)

3.3.1.1 Missing Data

In dataset B, there were two stations with no MT records and one station with no P record. Those three stations (stations H4SLAT01 Slatiny, O3HOST01 Hostalkova, and H3CHTU01 Chotusice, all highlighted in red in \input_data\2008.xls on the DVD attached) were therefore removed and the number of stations was reduced from 135 to 132 stations.

Regarding dataset A, station ‘Velke Pavlovice’ was shut down in November 2008 and a new station ‘Kobyli’ started observations in January 2009. In order to preserve consistency and to have the same number of monitoring station for each year, two MT and two P monthly records needed to be interpolated for the last two months of 2008 to obtain annual values/sums.

MT records were interpolated using linear regression (‘TREND’ function in MS Excel) with ALT as the independent variable. Data table is included in Appendix 1.



Figures 3.5a, b: Linear regression between MT in November and December and altitude.

Missing P records were interpolated using multivariate linear regression in Geoda software with ALT, CURV and values of other months as independent variables. Data table, protocols, and regression equations are included in Appendix 1.

3.3.2 DEM

Digital Elevation Model (DEM), displayed in fig. 3.6, originates from the NASA Shuttle Radar Topography Mission (SRTM), launched in year 2000. The mission captured topographic data for 80% of the Earth’s land surface at a horizontal resolution of 1 arc second for USA (~30 m on the equator). Horizontal resolution of SRTM global datasets is 3 arc seconds (~90 m along the equator). For latitudes of central Europe each pixel represents a rectangle of 60 × 90 m. The horizontal datum of SRTM datasets is WGS84 (geographic coordinate system using latitude and longitude). SRTM is vertically referenced to the WGS84/EGM96 geoid. [35] ‘The SRTM global datasets comprise an annulus between 60°N and 56°S latitude. The model is global, because GTOPO30 data were used to fill in latitudes beyond 60°N and 56°S, as well as void areas within the SRTM data. GTOPO30 is another coarser DEM of 30 arc seconds (~1 km) horizontal resolution developed by the USGS EROS Data Center in 1996 from a variety of data sources.’ [35]

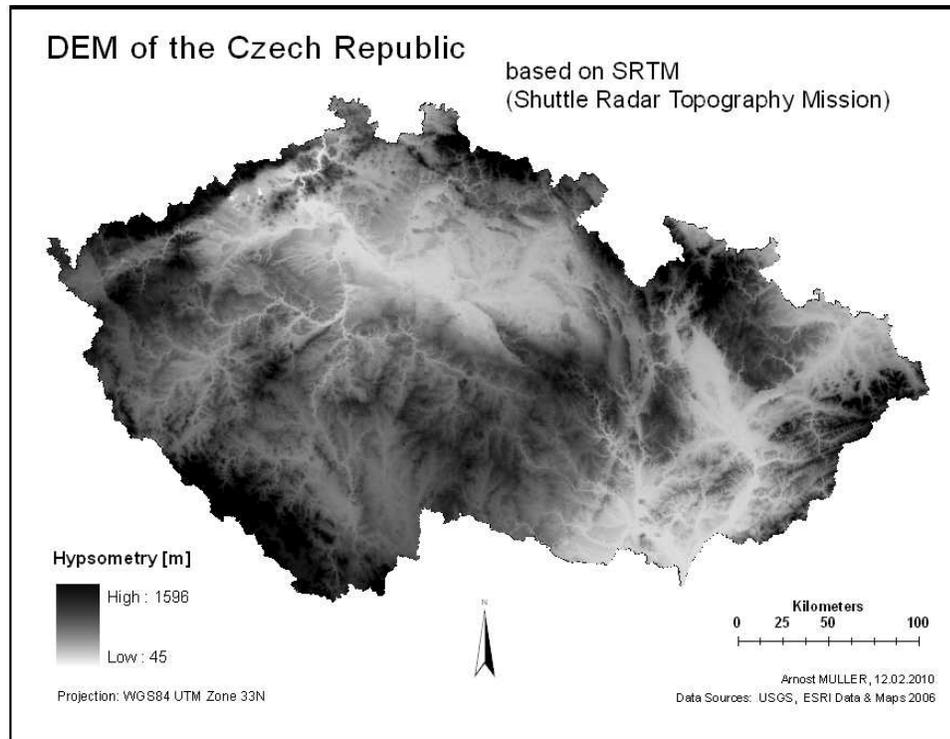


Figure 3.6: SRTM DEM, each pixel contains one value representing altitude in meters.

Although NASA has already published a ‘finished’ version of SRTM (version 2 with interpolated void areas), DEM used in this work is the same as in author’s previous work [35]. Void areas contained in SRTM version 1 were filled in by ESRI and the DEM used in this thesis was downloaded from ESRI Data and Maps 2006. The DEM came in decimal degrees in WGS84 datum. It needed to be converted to UTM coordinates in meters to allow some calculations such as curvature or solar radiation. The conversion was performed using the ‘Project Raster’ tool in ArcGIS toolbox by selecting bilinear interpolation. Bilinear interpolation is suitable for continuous data (such as elevation surfaces) and determines the new value of a cell based on a weighted distance average of surrounding cells [17]. The default output cell size was kept on the default value of 61.94 m. This value was the most similar to the size of the original grid in that the number of cells in rows and columns of the new raster remained more or less the same. All calculations including interpolation are then calculated at 61 m grid. The resulting climate surfaces, however, are resampled to 90 m when a low pass filter is applied.

3.3.3 Variables Derived from DEM

In complex terrain, climatic patterns are delineated by topographic slopes and barriers. [15] Several variables which appear throughout the scientific literature to be significant predictors of climate were derived from the DEM in ArcGIS: Altitude, slope, aspect, curvature, solar radiation, maximum altitude in a wedge and UTM coordinates X and Y.

3.3.3.1 Altitude (ALT)

Climate varies strongly with altitude. Temperature typically decreases with altitude. [15] While there is a known empirical relationship between temperature and elevation (6.5 °C per 1 km increase in elevation) [20], precipitation also generally increases with elevation, but the relationship is not that straightforward. [31] Altitude is an excellent statistical predictor, because it is sampled at greater spatial density than climate variables in a form of a regular grid – DEM, in our case 61 m (refer to section 3.3.2). [15] The strong relationship between altitude and climate has been utilized in many scientific studies, e.g. [1, 35, 36, 50].

Daly et al. [14] describe the orographic effects on altitude: ‘The relationship between precipitation and elevation varies from one slope face to another, depending on location and orientation. Relationships between measured precipitation and elevation are sometimes strengthened when the elevation of each data point is given in terms of its height on a smoothed terrain.’ As suggested Daly, smoothed ALT to 1, 5, and 10 km was therefore also tested in the precipitation model selection process.

3.3.3.2 Slope (SLP)

Slope (fig. 3.7) characterizes the steepness of terrain. The ‘Slope’ tool of ArcGIS calculates the ‘maximum rate of change from a cell to its neighbors’. [17] The effects of slope on climate are often combined with aspect and are therefore explained in the next section.

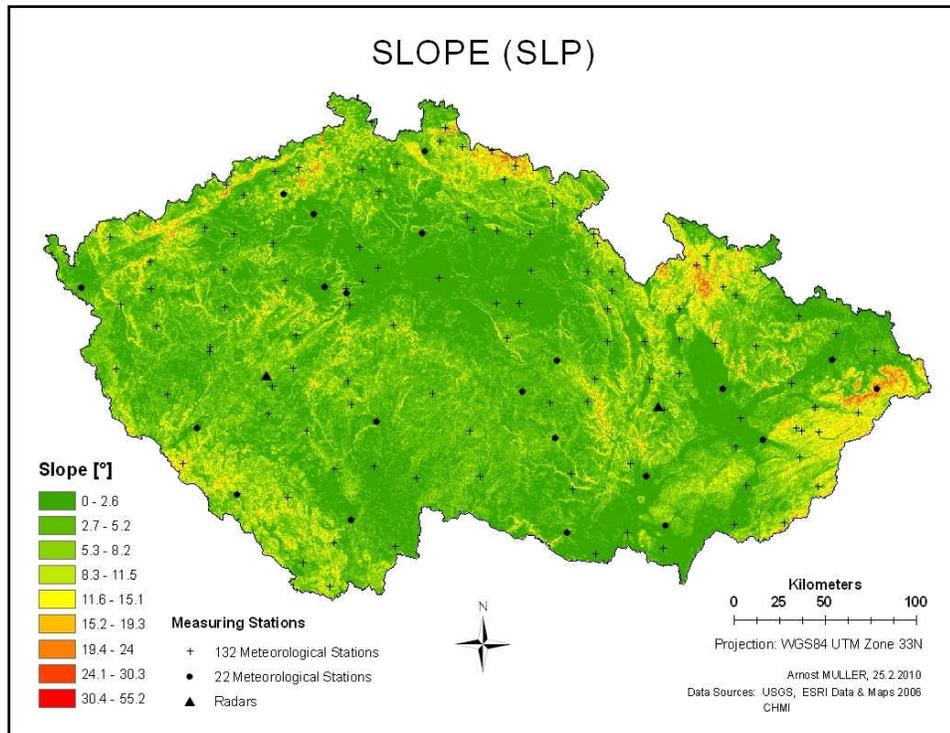


Figure 3.7: Slope

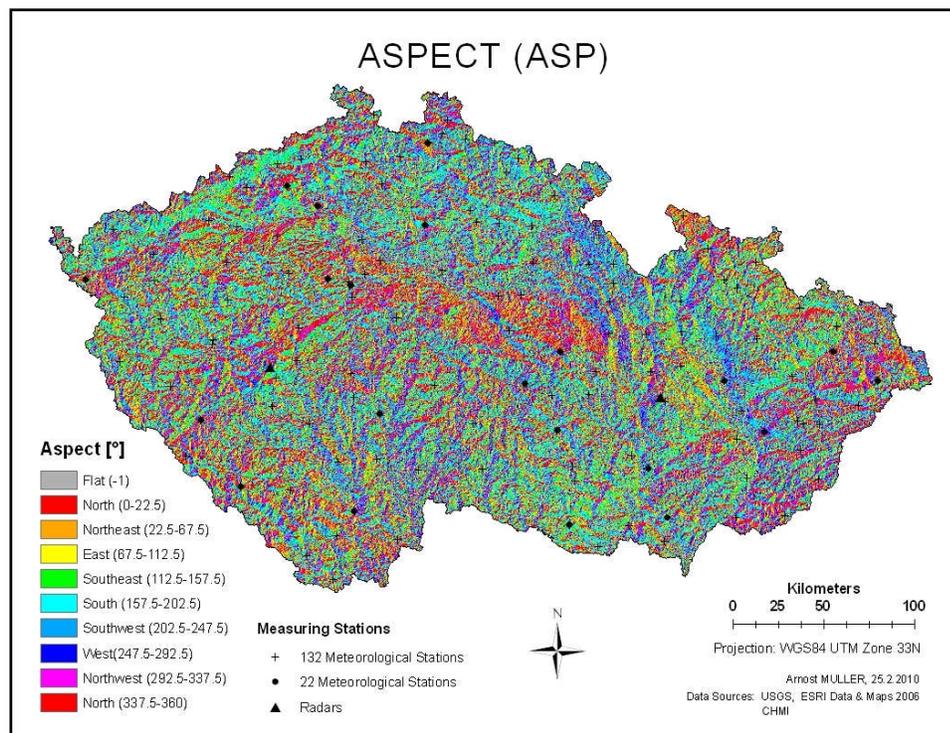


Figure 3.8: Aspect

3.3.3.3 Aspect (ASP)

Aspect (fig. 3.8), the direction of slope, was calculated in ArcGIS using the corresponding function of the Spatial Analyst Extension. Values of 0° face to the North, values of 90° to the East. Slope orientations typically affect the amount of sunlight received, which directly influences MT. In the northern hemisphere places with southerly aspects tend to be warmer and drier than places with northerly aspect. [15, 17] As humid air sweeps up the slopes of a mountain range, the air cools, and clouds form. Eventually, rain or snow falls from the clouds. The rainiest places are usually those facing the wind (called windward slopes). As winds blow down the opposite slopes (leeward slopes), the air warms, and clouds tend to vanish. Leeward slopes of mountain ranges are therefore dry. [34]

3.3.3.4 Curvature (CRV)

Curvature (fig. 3.9) is the second derivative of the surface, in other words the slope of the slope. A value of 0 indicates the surface is flat. A positive curvature indicates the surface is upwardly convex at that pixel. Convex parts of surfaces are typically ridges, generally exposed and drain. A negative curvature indicates the surface is upwardly concave at that pixel. Concave parts of surfaces are typically channels or valleys. The values between ± 0.5 represent moderate relief, while values over ± 4 represent extremely steep relief. [17]

3.3.3.5 Solar Radiation (SOLRAD)

SOLRAD was used by Ninyerola as a predictor in MT modeling. [36] Solar radiation model contains topographic information (slope and aspect) that determines amounts of incident solar radiation and can influence cloud formation or wind circulation. [17] SOLRAD can be calculated for certain year in ArcGIS with the 'Area Solar Radiation' tool (example in fig. 3.10), but the calculation is very demanding and can take several hours. SOLRAD further refers to the theoretical value derived from the model.

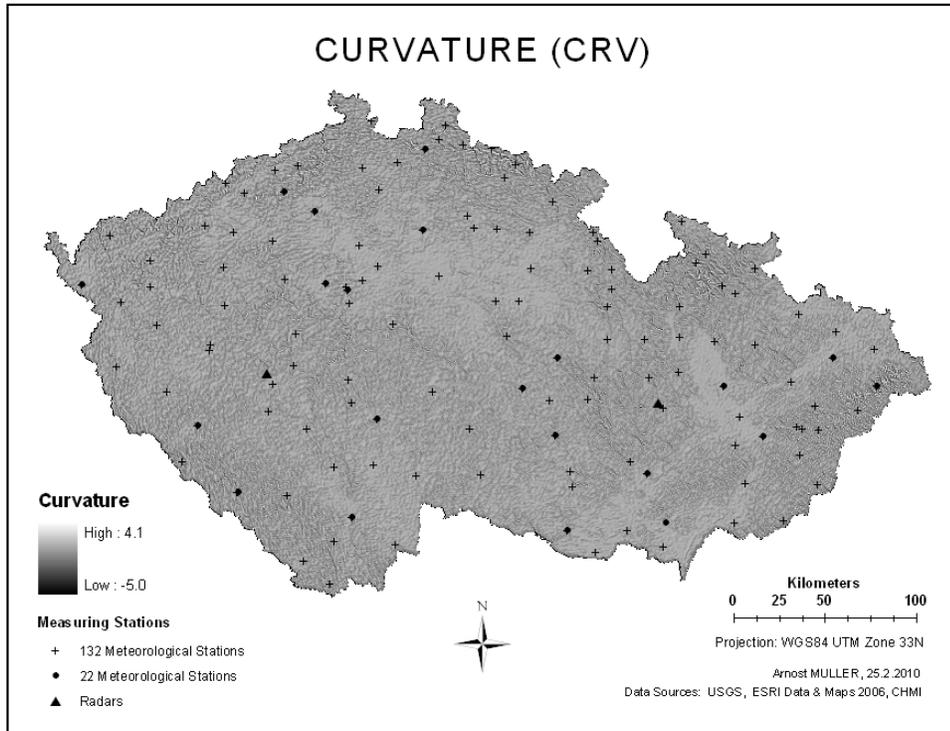


Figure 3.9: Curvature with hillshade effect ($Z = 2$)

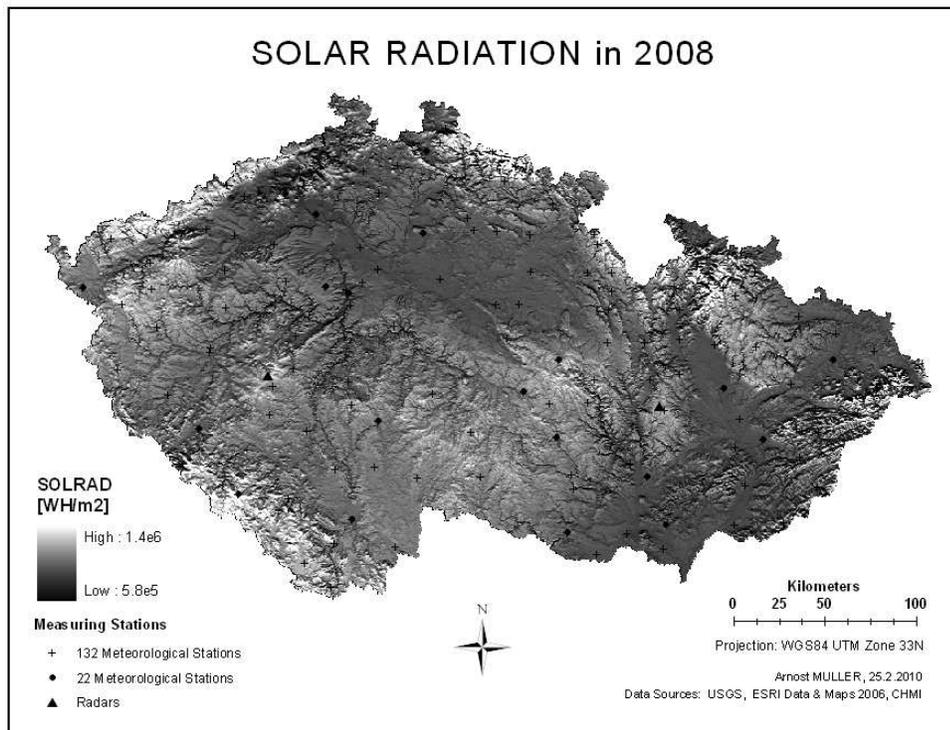


Figure 3.10: Solar Radiation in 2008

3.3.3.6 Maximum Altitude in Wedge (Zx)

The right full term describing this variable should be ‘the maximum elevation in a wedge of given aspect and radius’ (ZxDIR&RADIUS), according to Agnew and Palutikof. They emphasize that this variable is very useful in areas with dominant wind directions, because the leeward slopes are drier and warmer than windward slopes. They include Zx variable in all of the models for predicting seasonal precipitation.

In this thesis, there are 16 Zx variables. Zx in 4 directions (W, SW, NW, N, see fig. 3.11.) multiplied by 4 radiuses of 1, 5, 10, and 25 km. For example, ZxW25 is the maximum altitude in W direction within a 45° wedge with the radius of 25 km (see fig. 3.12). The 4 directions were selected with respect to the direction of prevailing winds in the CR. According to the ‘wind roses map’ in the Climate Atlas of Czechia, westerly and south-westerly winds are the most frequent. In the eastern part of the CR, the prevailing direction is modified by the terrain so that the northwest and in some cases north direction is often the most frequent. According to Daly, orographic effects influencing precipitation may operate at relatively large spatial scales, responding to smoothed topographic features rather than detailed variations in terrain, therefore the 4 radiuses were selected. Zx with greater radiuses of 5, 10, and 25 km were calculated on a resampled 1 km DEM in order to save the time needed for calculation.

In ArcGIS, Zx is calculated using the ‘Focal Statistics’ tool, selecting the ‘wedge’ option, and specifying the starting and ending angle of the wedge, its radius, and ‘statistics type’ as ‘maximum’. For SW direction for example, the starting angle would be 22° and ending 67° to get the wedge of 45°. This tool looks into specific direction and distance and finds the highest altitude of the DEM. The Zx variable can simulate orographic barriers in prevailing wind directions.

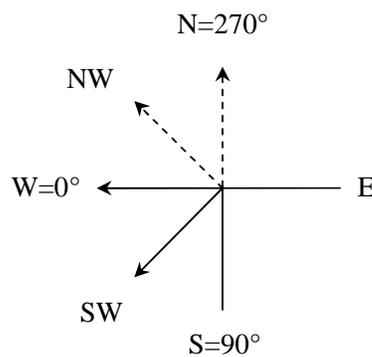


Figure 3.11: Wedge angles with 0 to the West, angles increase counter-clockwise

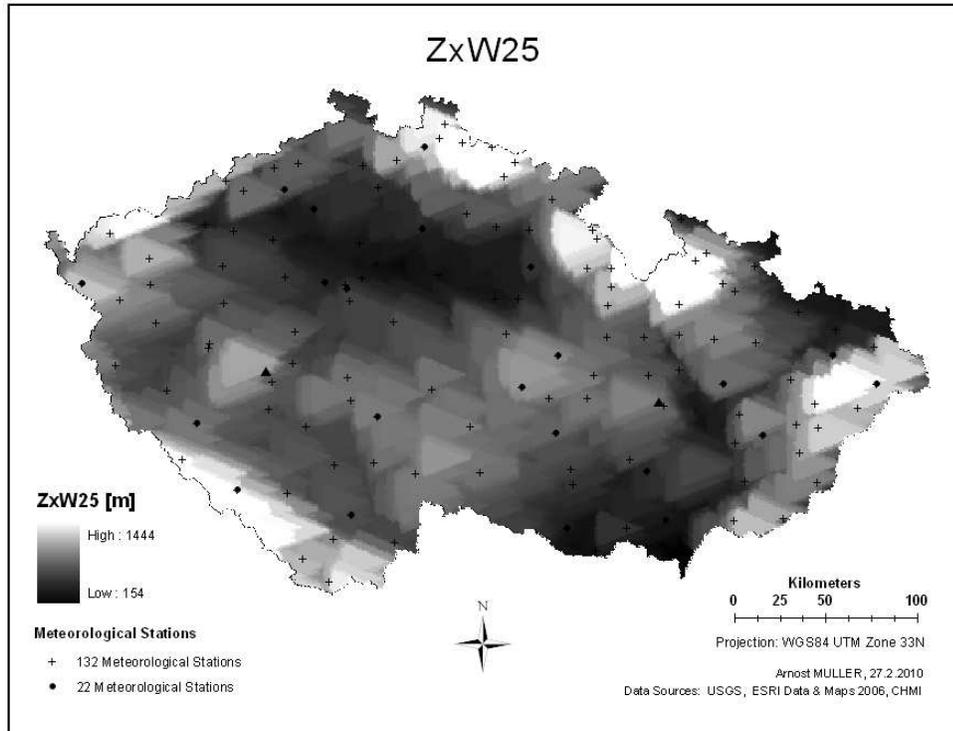


Figure 3.12: Maximum altitude within 45° wedge in the W direction and 25km radius

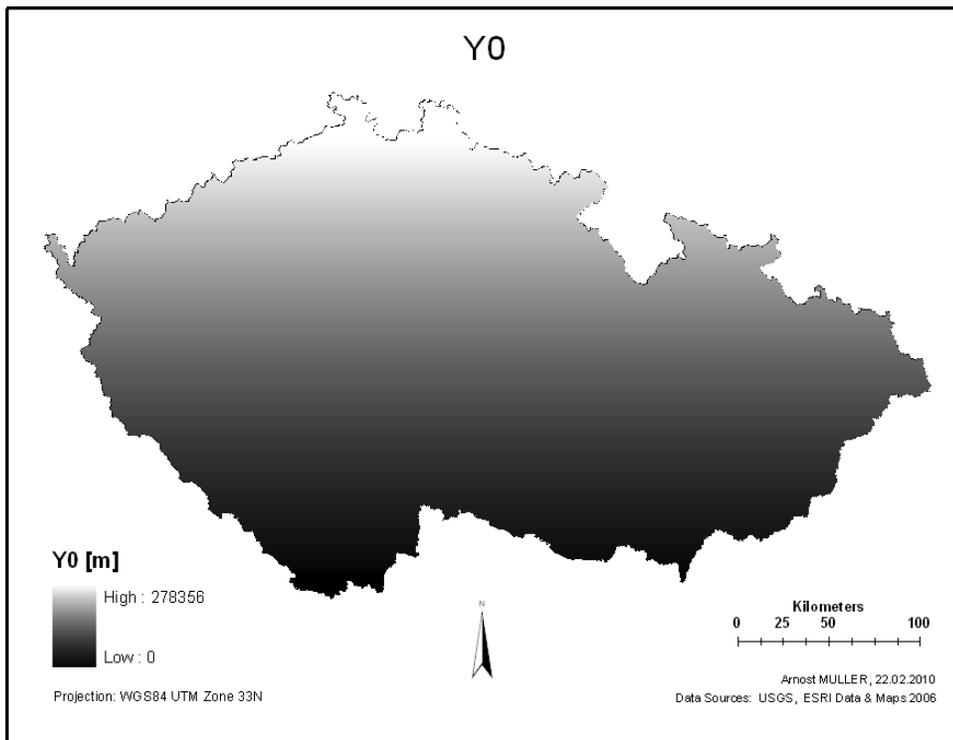


Figure 3.13: Y0 is the UTM coordinate representing latitude

3.3.3.7 X0, Y0

X and Y are location variables. Variable X represents latitude and Y longitude. Both are UTM coordinates expressed in meters. New raster matrix needed to be created for both X and Y. This was done using the 'Trend' tool. At first, a new shapefile with 3 random points outside the CR borders was created, because 3 points are needed to determine a surface. In the attribute table, each of the three points was assigned UTM X and Y coordinates using the 'Calculate Geometry' option. Linear trend surface was then fitted through the X or Y values with specified output cell size matching the DEM (61.940m). In order to obtain reasonable coefficient numbers in calculations of regressions, minimum values of X and Y were subtracted from each cell using the 'Raster Calculator' and so were X0 (eq.3.1) and Y0 (eq.3.2) created.

$$X0 = X - 292684.937(Xmin) \quad (3.1)$$

$$Y0 = Y - 5377820.5(Ymin) \quad (3.2)$$

3.3.4 DBF files

Input data tables were obtained from CHMI and formatted in MS Excel spreadsheets (\input_data\2008.xls on the DVD). The problem was that MS Excel does not support DBF database files which can be imported into ArcGIS. XLS files were opened in Open Office Calc and saved as DBF files with the Eastern European encoding (both Windows-1250 and ISO-8859-2) or UTF-8 but neither of them worked correctly in ArcGIS. Displaying the imported DBF file as a point shapefile in ArcGIS is done with the 'Display XY Data' function, by selecting the appropriate X coordinate (longitude), Y coordinate (latitude) and the geographical datum WGS84. Point shapefiles with attribute tables filled in with independent variables information are also included on the DVD in the geodatabase.

3.3.5 ArcČR 500 v. 2.0a

The dataset ArcČR 500 version 2.0a was obtained from ArcData Praha, s.r.o. free of charge for academic purposes in 2007 for the author's Bachelor Thesis. Shapefiles containing administrative regions of the Czech Republic, rivers and streams, water bodies, and all cities were used to extend the spatial context of resulting climate maps. The dataset came in WGS84 and was also converted to UTM Zone 33N.

4 Methods

Accurate climatological data are collected at meteorological stations, which are discrete point locations in space. Values at any other point must be derived from neighboring stations or from relationships with other variables. [31] The method of spreading discrete measurements over a continuous surface represented by a regularly spaced grid is called spatial interpolation.

Methods of mapping climate from point data fall into two categories: human expertise and statistical. The first group is based on human experience and knowledge and involves manual preparation of climate maps often related to topographic analysis. Statistical methods use a numerical function, calculated or prescribed, to spread irregularly spaced point data to a regularly spaced grid. [15] Various statistical methods have been developed to predict the spatial distribution of climate variables:

4.1 Global Methods

Global methods include all sampling points, in this case all weather stations, in calculations. Global methods use external information, e.g. topographic data, to create dependence models between the external (independent) variable and the modeled (dependent) variable, which is typically done by a polynomial function or by the means of simple or multiple regression models. The underlying hypothesis is that climate at any location is influenced by the environmental attributes of the surroundings. The geographical variables of the weather station (typically coordinates - latitude and longitude, distance to water bodies) or topographic variables (e.g. elevation, aspect, and slope) are used as independent variables. The value at an unsampled point is predicted by the following function [50]:

$$z(x) = \beta_0 + \beta_1 P_1 + \beta_2 P_2 + \dots + \beta_n P_n \quad (4.1)$$

where z is the predicted value at point x , β_0 - β_n are the regression coefficient and P_1 - P_n are the values of independent variables at point x . Simple linear regression (with one independent variable) fits a regression line (example fig. 3.5 or fig. 5.9). Multiple regression determines a regression surface using least-square estimation. Global methods are inexact interpolators in that the values predicted do not coincide with the real observations of weather stations.

The relationships between climatic and independent geographic or topographic variables have been extensively studied in the scientific literature. Linear regression was used in the following works [22, 35] , and multiple regression was used in [1, 20, 36]. All of the above studies reveal elevation as the most important predictor in the distribution of weather data.

4.2 Local Methods

In contrast to global methods, local interpolators only use the data from the nearest sampling points (weather stations). First, a number of weather stations are selected. Prediction at a given point is done by a mathematical function which takes data from the selected stations. Local interpolators are exact methods, as predictions coincide with the measured values. [26] Local interpolators such as inverse distance weighting (IDW), Thiessen polygons or thin-plate splines [25, 26] have also been used for climate mapping, but are nowadays used as additional interpolators in combination with a geostatistical or global interpolator, as described in sections 2.1.1 and 4.4.

IDW is a simple statistical method based on the assumption that the value at unsampled point $z(x)$ is a distance weighted average of the climatic values at selected sampling points $z(x_1), z(x_2), \dots, z(x_n)$. The distance d_{ij} is the weighting factor, because climate values are in general more alike between closer points than between distant points. [50] The IDW formula looks as following:

$$z(x) = \frac{\sum_{i=1}^n z(x_i) \times d_{ij}^{-r}}{\sum_{j=1}^n d_{ij}^{-r}} \quad (4.2)$$

where r is the power parameter (positive, real number, $r = 2$ by default in ArcGIS). The choice of power parameter can significantly affect the interpolation results. As r increases, 'IDW approaches the nearest neighbor interpolation method where the interpolated value simply takes on the value of the closest sample point', says Collins. [12] In other words, the r exponent provides the possibility for the user to control the significance of known points on the interpolated values. [18]

The next local interpolation method is splines. Splines algorithms are mathematically complex, but are standard in current GIS. For each $z(x)$, a new function is created according to the number of sampling points available in radius r . The predicted value $z(x)$ is determined by two terms [50]:

$$z(x) = T(x) + \sum_{i=1}^n \lambda_i \times \psi(r_i), \quad (4.3)$$

where $T(x)$ is a polynomial smoothing term and the sum contains a group of radial functions. Splines assume smoothness of variation. Splines have the advantage of creating contour lines which are visually appealing, but may mask uncertainty present in the data. [12]

Thiessen polygons, also called nearest neighbors, simply take the value of the nearest point where climatic information is available. The result is a polygon network with abrupt spatial discontinuities in the values when passing from one polygon to the other. [50]. Therefore, it is not an appropriate interpolation for variables with gradual spatial variation, such as climatological variables which are the subject of this thesis. [5]

4.3 Geostatistical Methods

Kriging is a group of geostatistical methods which assume that the spatial variation of continuous climatic variable is too irregular to be modeled by a mathematical function, so prediction by a probabilistic surface is proposed instead. [50] The climatic value z at point x can be expressed by the following formula [5]:

$$z(x) = m(x) + \varepsilon'(x) + \varepsilon'' \quad (4.4)$$

where $m(x)$ is the drift component which represents the structural variation of the climatic variable, $\varepsilon'(x)$ are the residuals, the difference between the drift and the original sampling data values, while $\varepsilon''(x)$ are spatially independent residuals. Kriging methods are based on a weighted average of the data available in the n neighboring stations. The weights are chosen so that the calculation is not biased and variance is minimal. [50] This is done through a semivariogram model that best fits the data (see fig. 6.3). [14] Examples of works where kriging is used to interpolate residuals are [1, 6, 46].

4.4 Combined Methods

Global methods are not exact interpolators, since the predicted value does not equal the value recorded at the station. The known error or difference between the two values is called a residual [50]:

$$\text{residual} = \text{observed data} - \text{predicted data} \quad (4.5)$$

Combined methods generally use a global method for prediction and a local method for interpolation of residuals. The later method is in fact a correction used to obtain exact climatic data at the measured locations [50]:

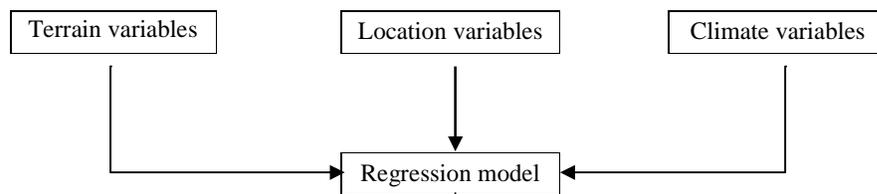
$$\text{observed data} = \text{predicted data} + \text{residual interpolation} \quad (4.6)$$

The correction interpolation increases the precision because it accounts for some of the remaining spatial variation hidden in the residuals. Combined methods were successfully applied in the following studies [1, 38, 50].

4.4.1 Multiple Regression with Residual Correction

The combination of statistical (multivariate regression) and spatial interpolation (IDW, spline, or kriging) has been demonstrated to be effective in MT and P modeling in the scientific literature [1, 36-38, 50]. The main purpose of this work is to investigate the role of multiple regression analysis with the residual correction method in climate modeling in CR. According to Burrough and McDonnell [5], this method is scientifically interesting because, in addition to the interpolation process, it gives information about the relationship between the geographic reality of the land and climate. Fig. 4.2 illustrates the whole process:

Stage I: Developing the regression model



Stage II: Estimating and refining the climate surface

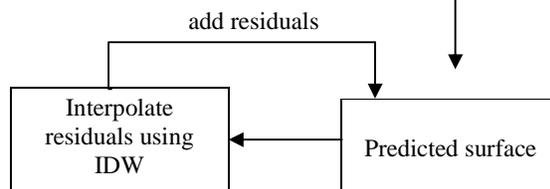


Figure 4.1: Two-stage methodology for constructing baseline climatologies [1]

The input data has following characteristics: topographical data as independent variables are in the form of rasters while climatological data as dependent variables are point features. Multivariate regression models are calculated in Geoda software. Backward stepwise approach and t-statistics is used for choosing significant independent variables. Backward stepwise selection begins with the examination of the combined effect of all of the independent variables on the dependent variable. One by one, independent variables (usually starting with the weakest predictor) are removed based on some outset criteria (t-statistics), and a new analysis is performed. [30] One should avoid the possibility of any of the variables being a linear combination of another variable. The goal of the backward stepwise approach is to find the combination of independent variables when adj-R^2 is maximized. On the other hand, since there are only 22 observations (dataset A), 4 to 5 independent variables will be the maximum, because the more variables are included, the lower is the degree of freedom of the model.

One final model is then selected for MT and one for P. Putting in the regression coefficients (β_n) into ‘Raster Calculator’, prediction maps called ‘potential maps’ [36] are created. Predicted values were extracted from prediction maps at weather station locations, then subtracted from observed values and so were residual values obtained for each station. Residual values are interpolated into a raster using IDW interpolation in ‘Spatial Analyst’. Residual surface is then pixel by pixel added to the potential maps using ‘Raster Calculator’. Ninyerola calls these residual surfaces ‘corrector maps’ [36]. ‘These corrector maps will not be uniform, but they will show maximum variability in the more unpredictable areas, and minimum variability in the predictable sites. In this sense, corrector maps can be seen as anomaly maps, of great interest to reveal the singularities of the climate at the local scale. The most unpredictable areas are usually correlated with rugged zones.’ states Ninyerola.

4.4.2 R^2 and Adjusted R^2

How well the model predicts is described by the correlation coefficient R^2 , also called coefficient of determination. It determines the goodness of fit of the model. In the case of linear regression this would be how well the line approximates the points (see fig. 3.4 or 5.9 for illustration). In case of multivariate regression, R^2 is calculated using the determinants $|R|$ and $|R_{yy}|$ of the correlation matrix R . $|R|$ is the determinant of the whole matrix, while $|R_{yy}|$ is the determinant of the matrix without the first row and column. [24]

$$R^2 = 1 - \frac{|R|}{|R_{yy}|}$$

4.4.3 Fisher’s F-test

Fisher’s F-test can test the overall effectiveness of the regression model, but it does not advise whether all of the independent variables are significant, if more variables should be included or some excluded from the analysis. [24] F-test is calculated as the ratio of two variances. In Geoda, F-statistic tests the null hypothesis that all regression coefficients are jointly 0 and gives the associated probability. F-test will typically reject the null hypothesis and is therefore not that useful. [2]

4.4.4 Student's t-test

Student's t-test in case of multivariate regression provides a measure of significance of an independent variable. It helps to analyze the within-group variation and to decide which variable are to be excluded from the analysis. The higher the t-statistics value, the more significant is the variable and vice versa. For the quality of the model it is better to include a variable which is not significant (is useless) rather than not including a variable which is significant. [24]

4.4.5 Protocols of Regression from Geoda

Geoda software developed at the Arizona State University, USA, is used as the primary software for calculation of multivariate regression. Geoda reads shapefiles as input. Geoda's 'Regress' tool offers the option to include the full covariance matrix of the regression coefficient estimates in the output protocol, and the predicted values and residuals for each observation. The output protocols (in appendix 1 and on the DVD in \regression_protocols\ folder) contain several statistical characteristics of both dependent and independent variables. RMSE (see table 7.1) is in Geoda referred to as S.E. of regression (standard error of regression). The explanation of the various statistical measures based on [2] is given below:

- The first few lines characterize the dependent variable (by providing the mean value and standard deviation) and the model (number of observations, number of independent variables and degrees of freedom).
- R^2 and $\text{adj-}R^2$ provide the measure of the goodness of fit of the model.
- F-statistics and probability of F-statistics can tell whether the model is not effective more than whether it is.
- The sum of squared residual enters the calculation of Sigma-square (which is a standard error estimate or RMSE in tab. 7.1) and Sigma-square ML. The last one in the column is S.E. of regression ML (standard error estimate or RMSE_{ML} in eq. X+1). The ML index means that the statistical measure does not compensate for the loss of freedom.
- The area between dashed lines displays the independent variables and their significance characterized by the t-statistics (Student's t-test) and its probability.

- The three statistical measures on the right (Log likelihood, Akaike Information Criterion (AIC) and Schwarz Criterion (SC)) are used for comparisons across the various spatial (and non-spatial) regression models. The higher (or less negative) the log-likelihood, the better the fit. For the information criteria, the direction is opposite, and the lower the measure, the better the fit.
- Below the dashed line there is the multicollinearity condition number. This diagnostic suggests problems with the stability of the regression results. Values over 30 are in general problematic.
- The Jarque-Bera test is used to examine the normality of the distribution of the errors. The low probability of the test indicates non-normal distribution of the error term.

4.5 Limitations

The major limitation of this work is the lack of high spatial density of meteorological stations, as discussed in Data Sources in section 3.3.1.

The terrain variables are all derived from a DEM which may contain errors. The accuracy of the DEM is further analyzed in section 6.

4.6 Delimitations

The author selects to study only two basic climate variables – mean temperatures (MT) and precipitation (P).

5 Using Radar-derived Rainfall in Precipitation Modeling in the Czech Republic

This chapter presents a separate study about using radar data in annual rainfall modeling in CR. This analysis was done prior to the whole thesis, therefore 21 meteorological stations are used (missing data were not filled) and latitude (LAT) is introduced as an independent variable instead of Y0.

5.1 Introduction

Precipitation, as one of the basic climatic variables, is essential for life, however, it can also transform into destructive power. It is used as an input in various models not only in hydrologic modeling, e.g. flood prediction, but also in agriculture applications for estimating yields, land management, forestry or in atmospheric simulation models.

Rainfall data are traditionally collected at meteorological stations, which are discrete point locations in space. These measurements are called rain gauges. Rain gauges observations, although represent only points, are still considered as close to true rainfall as we can get at present state of art technologies. [27] Values at any other point must be derived from neighboring stations or can be remotely sensed, e.g. by ground-based radar. Weather radar has advantages in contrast to other methods. First, it encompasses large spatial domains of up to 260 km. Second, radar images are acquired at fine temporal resolution, for example every 10 min. Third, radar can ‘see’ much larger atmospheric space than rain gauges located on the ground. Radar can detect the areal distribution of precipitation at more detailed spatial scale than rain gauge network and therefore, the final rain field pattern should be determined by radar as recommended by Krajewski [27]. Conversely, precipitation obtained only from radar data cannot be directly used because radar measurements are affected by various types of errors and the transformation of measured radar reflectivity into rain rates is far from accurate. [27, 44] The solution presented in the literature e.g. [27, 42, 43] include merging radar and rain gauges estimates.

Krajewski [27] presents a brief discussion of the use of rain gauges in radar-rainfall estimation and concludes that both radar and rain gauge networks are equally important. Sokol and Bliznak [44] study heavy short term precipitation of the Czech Republic at 1h temporal and 1 km spatial resolution by merging radar and rain gauge precipitation [43]. Sokol and Bliznak [2] focus on the precipitation-altitude relationship and conclude that the

relationship depends on the accumulation period. 1-hour high rain rates are without apparent dependence on altitude, while longer 2 and 3 hours low and high rain rates are impacted by altitude. 6-hour precipitation events reveal influence of the mountains, where precipitation occurs at larger scale.

The focus of this study is therefore to investigate the use of radar data in modeling annual precipitation over the area of the Czech Republic (CR).

5.1.1 Aims and Objectives

The broad aim of this study is to explore the use of radar data in precipitation modeling. The secondary objective is to test the influence of following topographic, locational and atmospheric variables on residual errors: altitude (ALT), longitude (E_lon), latitude (N_lat), aspect (ASP), slope (SLP), curvature (CRV), distance from the radar antenna (DIST), aspect perpendicular to the radar beam referred to as directional difference (DIF), mean temperature (MT), and solar radiation (SOLRAD). Residual errors are calculated as the difference between radar-predicted rainfall and rain gauges observations, which are considered true. If some factors influencing residual errors are found significant, this would mean radar errors are not random and therefore can be removed by calibration. The variables DIST, ALT and DIF are expected to have the largest influence on radar measurement accuracy amongst other studied variables.

5.2 Theoretical Framework

Rain gauge observations are typically used in calibration and validation of radar derived estimates. The process can be summarized in three steps: Firstly, parameters of the basic relationship between radar-measured reflectivity and rainfall rate (Z-R relationship) are estimated. The second stage involves adjustment of the mean field bias, which is the ratio of the true areally averaged rainfall (approximated by rain gauge observations) to the corresponding radar-rainfall. The third stage may use rain gauges to locally adjust radar-rainfall patterns by merging the two sets of rainfall estimates according to some criterion, e.g. mean error variance. [27]

Radar images can be imported into a Geographic Information System (GIS), which provides a standard means to display, overlay, and combine the data with other layers, e.g. topographic or climatic, for analysis. [21]

5.2.1 Methodology

In order to examine the use of radar rainfall data for precipitation mapping, regression analysis is applied to determine the relationship between rain gauges and radar estimates. Residual analysis using regression approach is then applied to study factors influencing residuals, which can provide insight into radar measurement errors. Given the data provided by the Czech Hydrometeorological Institute (CHMI), the analysis is firstly performed for annual estimates in the year 2008 with 134 stations and secondly at a monthly scale using gauge data from only 21 stations.

5.2.2 Data Sources

Precipitation data were obtained from the CHMI for year 2008 in two forms:

- 1) Monthly radar sums
- 2) Rain gauges at two temporal scales (monthly and annual) with a different number of stations (21 and 134 respectively).

5.2.2.1 Radar Data

The Czech radar network (CZRAD), operated by the CHMI, consists of two polarization C-band Doppler radars. Optimal location of the radars with respect to topography causes reflectivity to be significantly influenced by terrain blockage of radar echo only in small areas of the CR. [44] More information about the radar network and instruments' specification is detailed in [9].

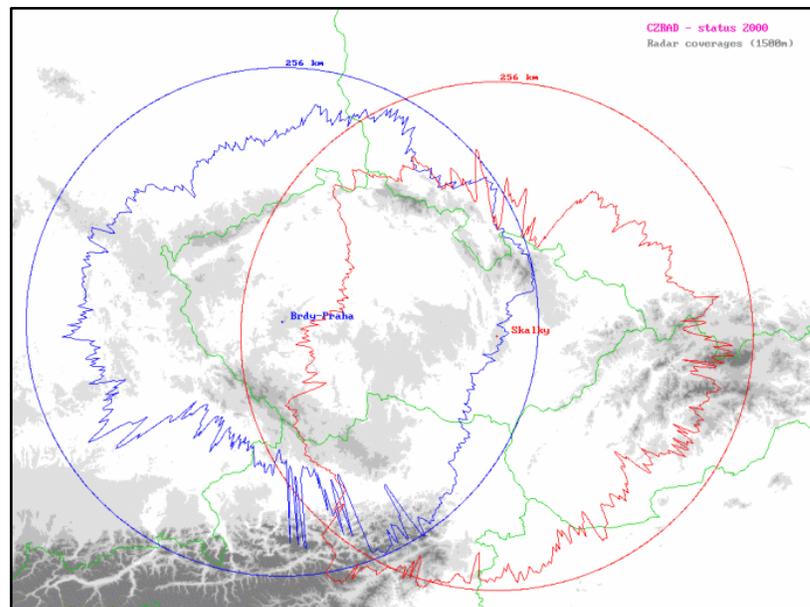


Figure 5.1: Map downloaded from <http://www.chmi.cz/meteo/rad/eindex.html> showing the maximum coverage of the CHMI weather radars (circles) and the effective coverage for P estimation.

Monthly raw radar rainfall sums (already combined from both radars) for the year 2008 were provided by the CHMI in two dimensional RPD format. The data are stored in Cartesian coordinates. The RPD file consists of two parts: header and binary data, which can be decoded in a Linux system environment using simple commands:

```
dd if=file.rpd ibs=512 count=1 | tr -d '\000' >header
dd if=file.rpd ibs=512 skip=1 | gzip -d >image
```

In order to import the image into a raster ArcGIS compatible format, the image file was renamed to `image.BIL` and new header `image.HDR` needed to be created (based on the information provided in the original header):

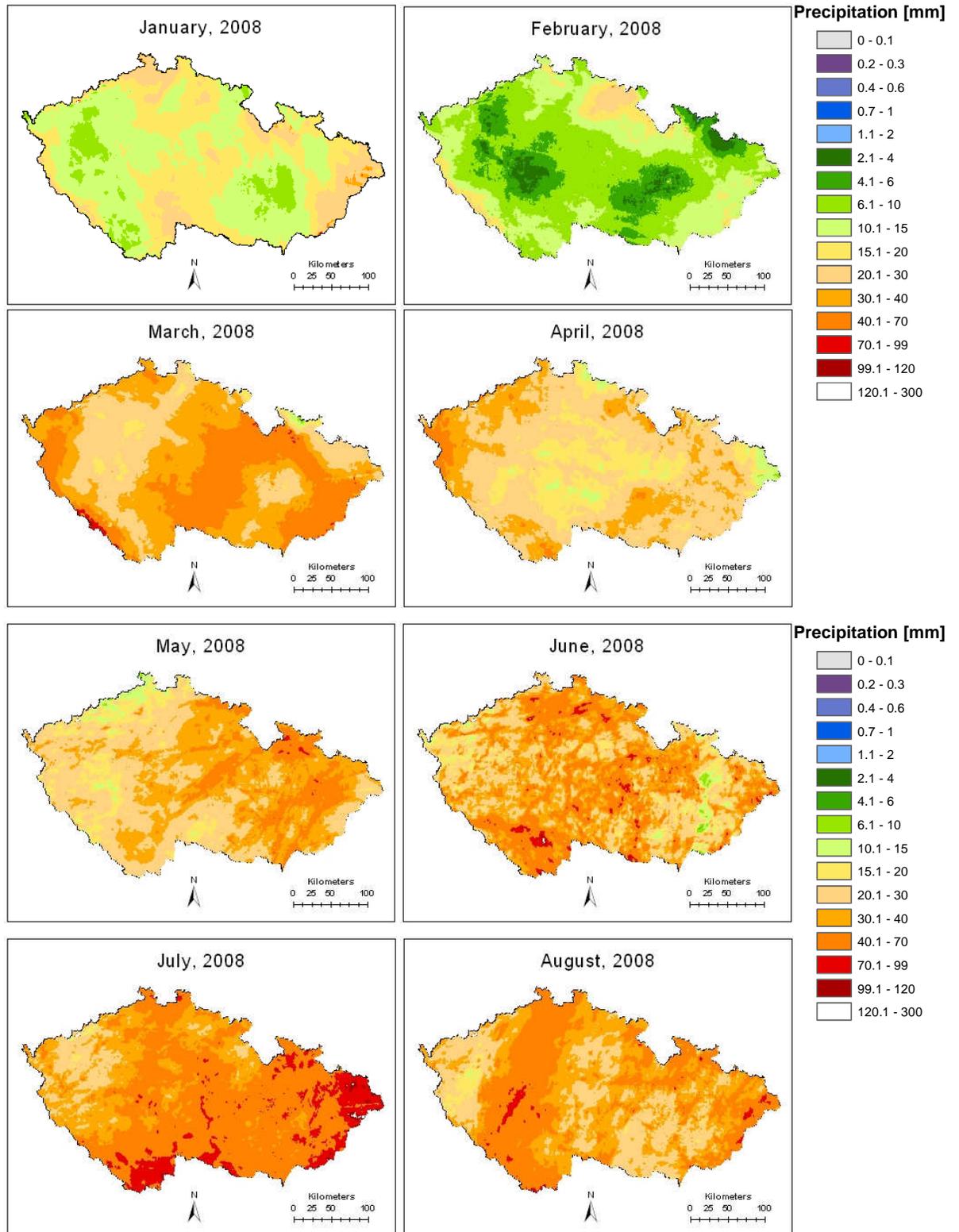
```
BYTEORDER  I
LAYOUT      BIP
NROWS       528
NCOLS       728
NBANDS      1
NBITS       16
NODATA      -9999
XDIM        1000.0
YDIM        1000.0
```

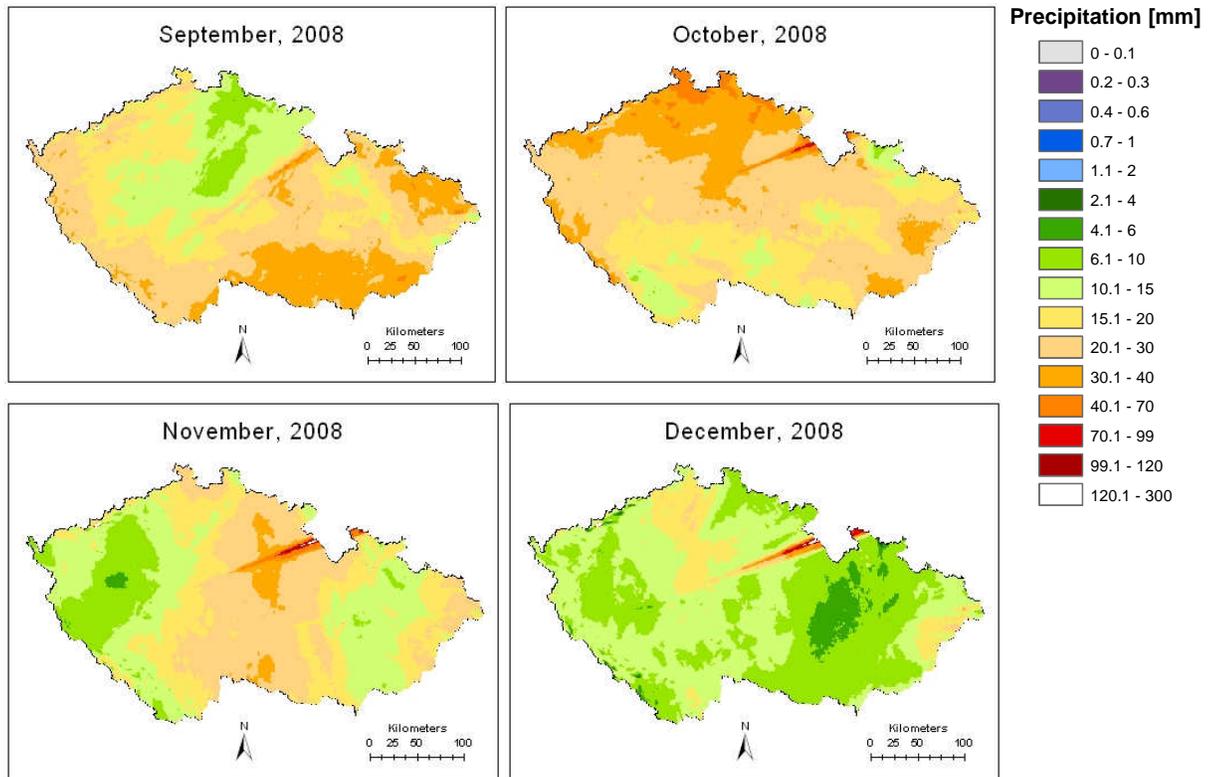
Using the ‘Raster to Other Format’ conversion tool 12 images were imported into ArcMap as GRIDs. Each image was then assigned gnomonic projection using ‘Define Projection’ tool. The gnomonic projection is defined by CHMI [39] by the following parameters:

```
<pacz23> # ID of projection
# combined information with 1x1km resolution (x_0, y_0 are in [m])
proj=gnom lat_0=50.008 lon_0=14.447 a=6379000.
x_0=301500 y_0=-217500 es=0. no_defs <>
x_res = 728 # nr. of cols
y_res = 528 # nr. of rows
pix_res = 1.0 # size of pixel in the centre of projected image in [km]
```

Radar sum rasters were then transformed from gnomonic to UTM projection (based on WGS84 ellipsoid, zone 33N) in order to be able to ‘extract values to points’ from radar rasters. The horizontal spatial resolution is 1×1 km.

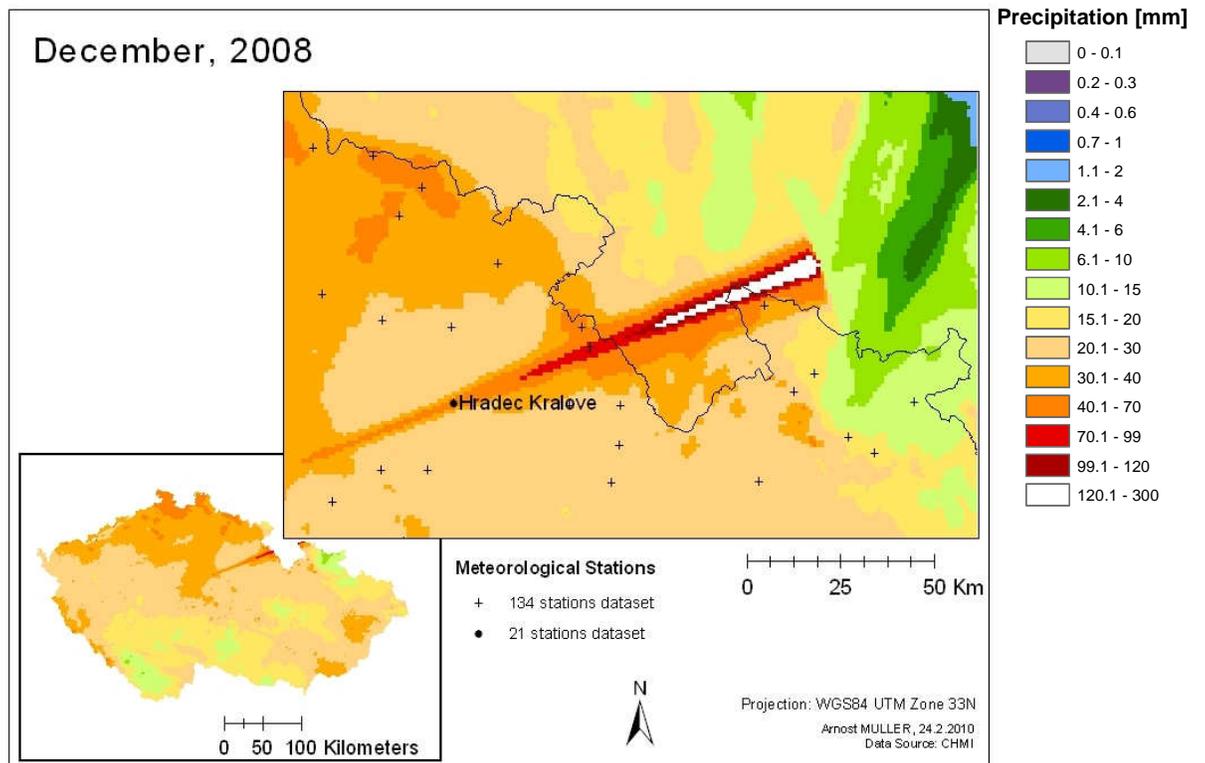
Monthly Radar Sums of Precipitaion in 2008





Set of figures 5.2 above: Radar sums by months for 2008

Figure 5.3 below: Zoom in to the area where radar beam casts a 'shadow' caused probably by different signal of similar frequency



5.2.2.2 Specific Independent Variables

Apart from variables described in section 3.3.3, two new variables are introduced.

- DIST (Fig. 5.4)

Distance to the Radar Antennas was calculated as the shortest Euclidean distance, using the function corresponding name.

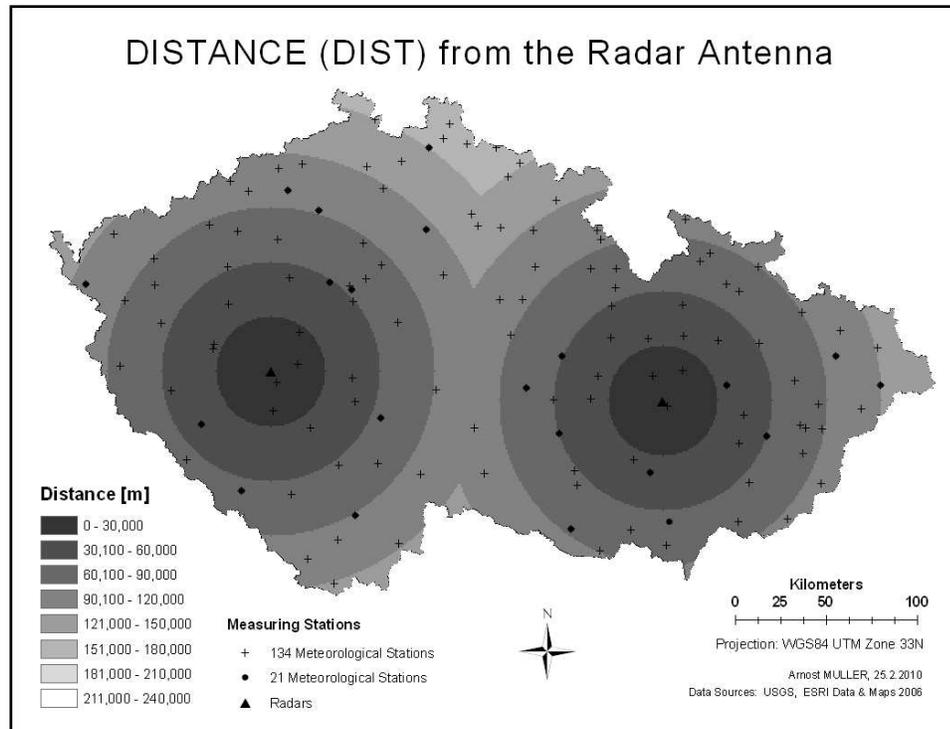


Figure 5.4: Distance

- DIF, DIF5 and DIF10 (Figures 5.6–5.8)

Directional difference is derived from ASP using GIS techniques. Difference between direction to the radar antenna (DIR) and aspect (ASP) is calculated as

$$DIF = | \cos(ASP - DIR) | \quad (5.1)$$

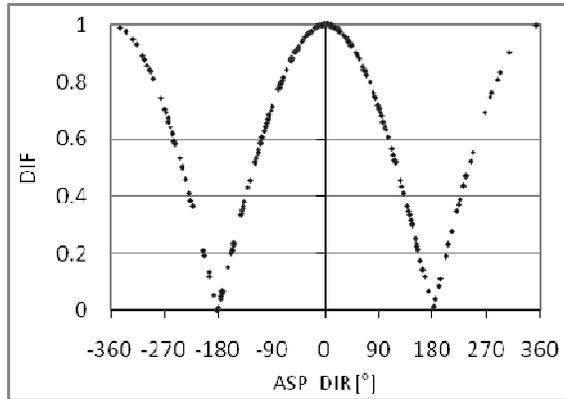


Figure 5.5: Chart explaining the meaning of the variable DIF

When $ASP-DIR = \pm 90$ or ± 270 , the aspect is perpendicular to the radar beam. When $ASP-DIR = 0$ or ± 180 , the aspect is facing the same direction as the radar beam. By averaging neighboring pixels two additional variables were created: DIF5 and DIF10, which represent directional difference for area of 5 and 10 km, respectively.

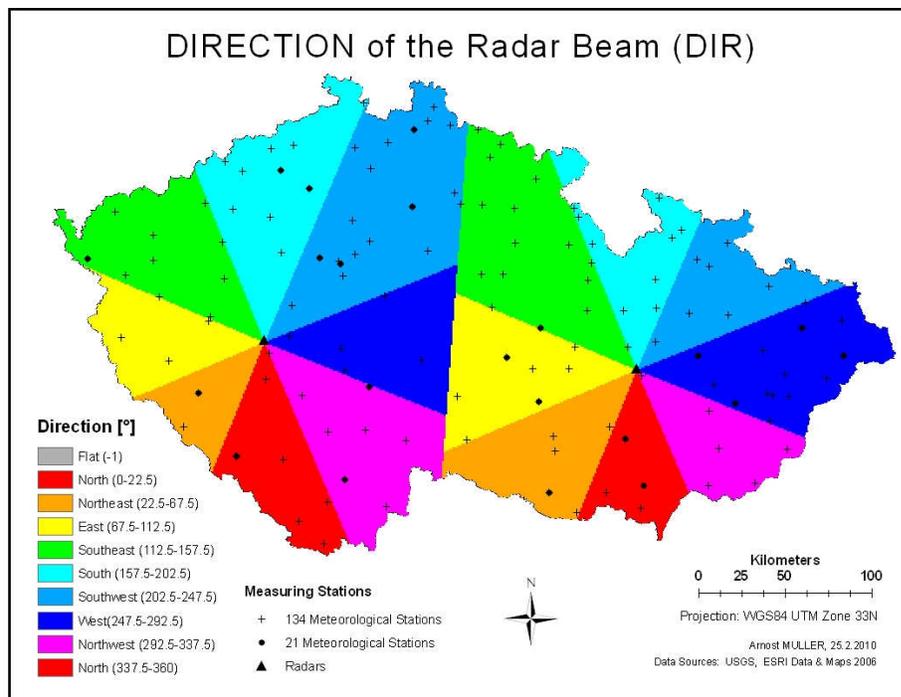


Figure 5.6: Direction

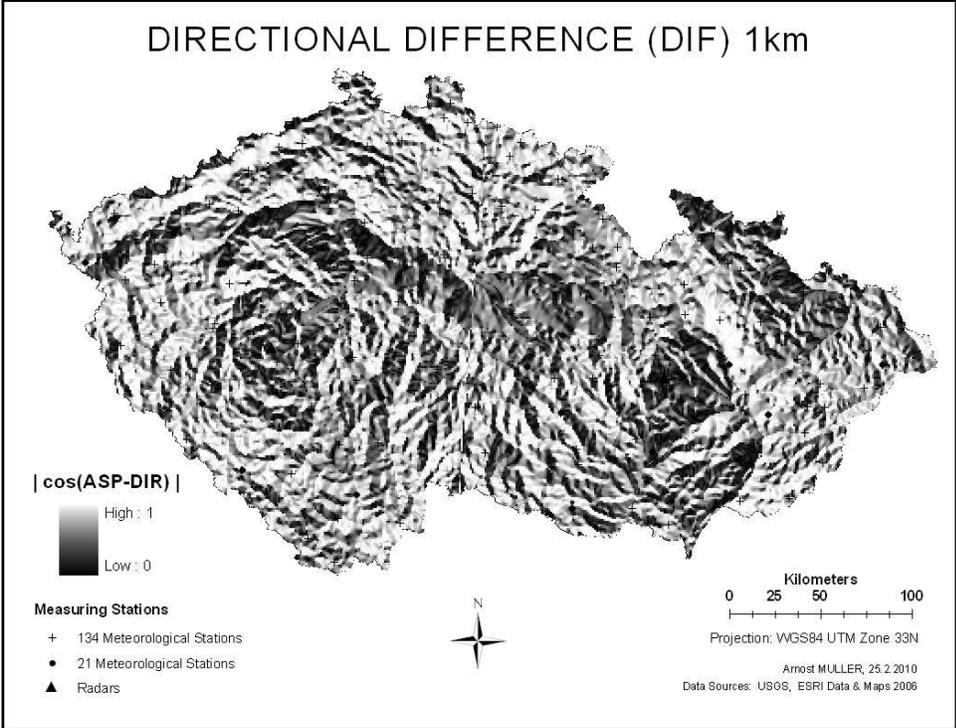


Figure 5.7: Directional Diference

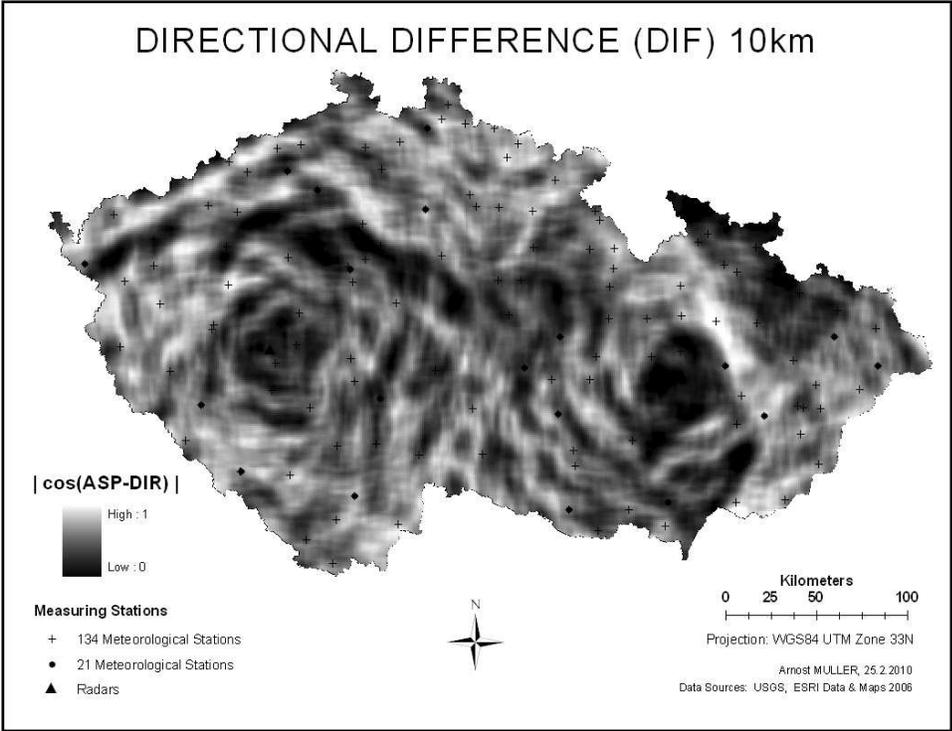


Figure 5.8: Directional Difference smoothed to 10km resolution

5.2.3 Atmospheric Data

Since mean air temperature (MT) can be interpolated over large areas with sufficient accuracy (standard deviation less than 0.1 °C) [35], it could also be used as an independent variable in rainfall modeling. However, it should not be used in models together with altitude with which it is highly correlated.

Meteorological stations recording rain gauges do not only record MT, but also sunshine duration (SD). SD, in addition to solar radiation (SOLRAD) is also included in the studied variables. While SD is a real measured variable, SOLRAD is a theoretical variable because it is derived from DEM. SOLRAD is calculated using ASP, SLP, and solar angle and does not include actual information about cloudiness.

5.3 Analysis

Since the aim of this study is to model precipitation by using radar data, gauge measurements enter the regression analysis as dependent variable and rainfall sums as independent variable. The analysis is carried out with two datasets:

- 1) annual 2008 data of 134 monitoring stations
- 2) monthly data for the same year of only 21 available stations.

5.3.1 Regression Analysis

Using MS Excel or Geoda software, simple linear regression model was fitted between radar (independent) and rain gauge data (dependent) variables as displayed in scatter plots in figures 12–14. Geoda software creates a protocol of the regression calculation with various statistical measures and includes the residuals between the predicted and the real gauge data, refer to appendix 2.

The scatter plot in fig. 5.9 reveals some relationship between rain gauges and radar measurements. The slope coefficient of 1.74 indicates that radar underestimates the gauge rainfalls by nearly twofold. When forcing the linear curve to intercept the [0,0] point, it increases to 1.94. The relationship is slightly linear, but not very strong. The correlation coefficient for annual data $R^2 = 0.18$ is low, which means that there is low dependence between gauge and radar precipitation estimates. F-statistics reaches 29 with the probability of 2.7×10^{-7} (refer to appendix 2) which indicates that the relationship is true and is not the consequence of chance.

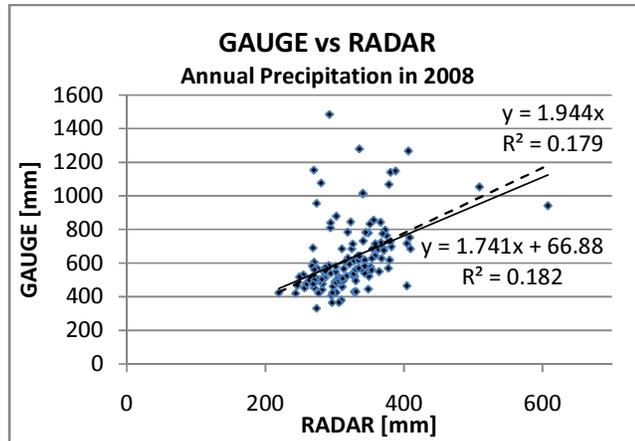


Figure 5.9: Gauge precipitation measurements of 134 meteorological stations plotted against annual radar rainfall sums over the same location.

One would expect a stronger relationship between the radar and the gauge data since both methods measure the same variable – precipitation. Radar rainfall integrated in time to represent rainfall accumulations are typically adjusted to rain gauge-based areal average of the corresponding rainfall. [27] The regression function implies that the radar data are raw and were probably not calibrated with the gauge measurements, otherwise one would see a random noise pattern in the plot, slope coefficient approximating 1 and the regression line intercepting close to the [0,0] point on the axis. The slope coefficient is almost 2, indicating that radar generally underestimates rainfall by almost a half.

If stations with residuals higher than 300mm (arbitrary selected threshold) were removed, the R^2 would increase from 0.18 to 0.41 (refer to fig. 5.10 below).

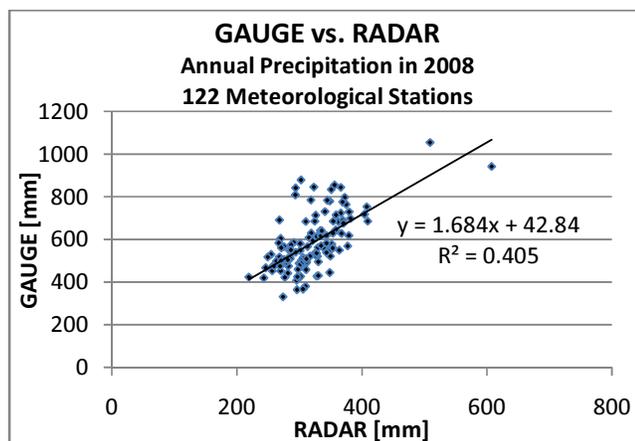


Figure 5.10: Gauge precipitation measurements of 122 meteorological stations (12 stations with residuals over 300mm were removed) plotted against annual radar rainfall sums.

The scatter plot in fig. 5.11 represents the monthly dataset. The slope coefficient of 1.4 is lower than 1.7 in annual data. It is pulled down by 2 stations with the highest residuals lying in the highest altitudes ('Churanov' and 'Lysa Hora'). However, the correlation ($R^2 = 0.50$) is better than in the annual case, where there is more stations situated in altitudes over 1000 m above the sea level.

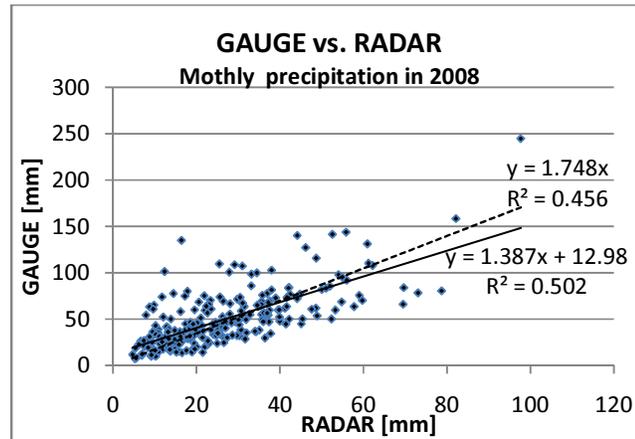


Figure 5.11: Gauge precipitation measurements of 21 meteorological stations plotted against monthly radar rainfall sums over the same location.

5.3.2 Residual Analysis

The following analysis seeks to determine if long-term annual and monthly precipitation can be predicted and modeled using radar rainfall. The residual analysis is based on the assumption that rain gauge data are true [27] and not biased, while annual radar sums of rainfall are not accurate. Further analysis focuses on the source and nature of errors of radar measurement by looking at the residuals of the two regression models, monthly and annual, depicted in figures 12 and 14. The analysis is carried out primarily on the annual dataset with 134 gauge stations which is considered more representative than monthly dataset with only 21 stations. The monthly scatterplots are included for comparison and confirmation of the annual results. Radar-derived precipitation generally underestimates gauge measurements and the underestimation increases with increasing distance from the radar. [43]

5.3.2.1 Residuals

Residuals (RES) are calculated as:

$$RES = PRED - P \quad (5.1)$$

where $PRED$ are predicted values by the regression model and P (monthly) or $P08$

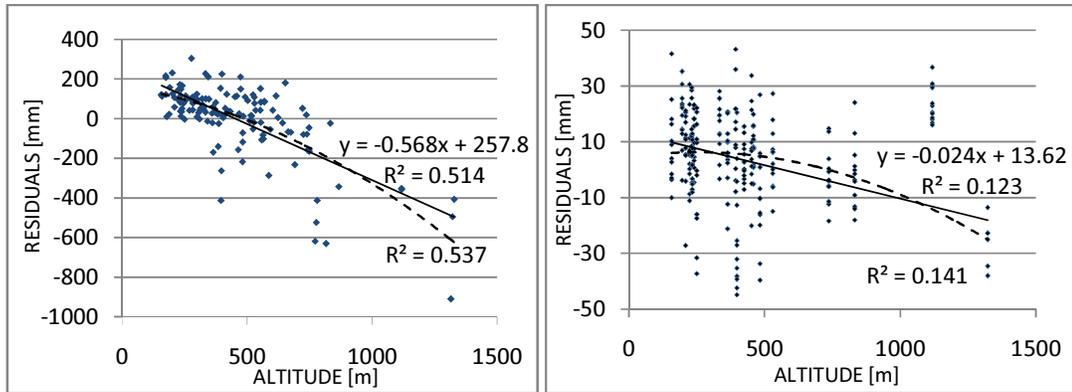
(annual) are gauge values. When the residual value is positive, the model is over-predicting the actual precipitation. If the residual is negative, the model based on radar rainfall is under-predicting the real value of precipitation. Residuals for all stations are shown in input data table, Appendix B. The highest residual values overreaching the arbitrary selected threshold of $\pm 300\text{mm}$ were highlighted in red color. Note that these stations are situated in high altitudes where there is generally greater precipitation than in lower altitudes and where radar measurement is less accurate, primarily due to the radar beam being obstructed by mountain ridges.

All of the residual values resulting from linear regression (figures 5.9 and 5.11) are negative. One exception is the station 'Hradec Kralove', where the residual value is positive, which means the model is over-predicting the real value. Meteorological station Hradec Kralove lies in a 'radar beam shadow' behind an obstacle for last three months of the year 2008 (refer to fig. 5.3). Therefore, this station was excluded from the analysis of monthly datasets. Stations 'Destne v Orli.H.' and 'Javornik' from the annual dataset lying in the same 'radar beam shadow' do not show any outlying residuals.

5.3.2.2 Residual Plots

Residuals (refer to eq. 5.1) were plotted against several variables. A linear (represented by solid line) as well as second order polynomial (dashed line) regression curve was fitted through the data to determine any underlying relationships and dependencies between residuals and other variables. This method helps to explain the variation in residuals and reveals factors which affect radar measurements.

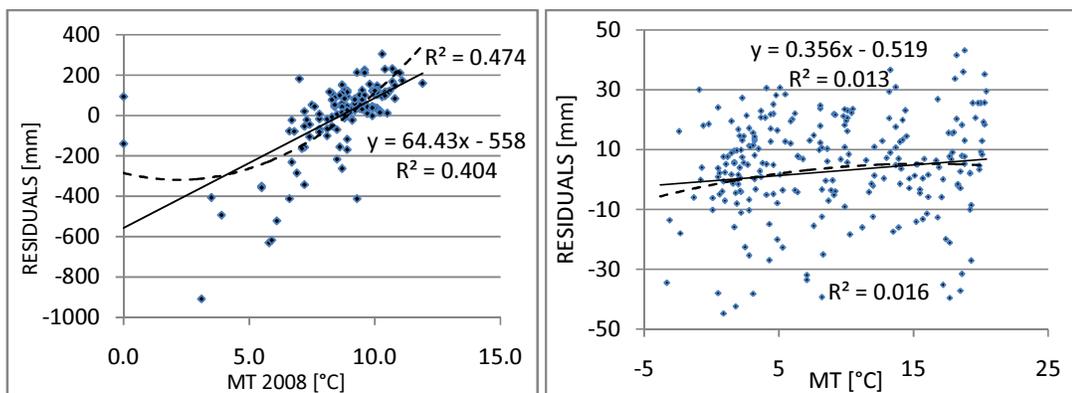
The following text contains pairs of scatter plots, representing one variable plotted against annual (on the left hand side) and monthly (on the right) residuals. The annual residual plots are further analyzed in the text, while the monthly plots are included as further reference for comparison and confirmation of annual results. This is because the annual dataset containing 134 stations is considered more representative.



Figures 5.12a and b: Altitude vs. annual and monthly residuals

It is generally accepted that altitude significantly influences the spatial distribution of precipitation. The main reason for increasing precipitation with altitude is the orographic lift, which occurs on windward slopes, where the arising air mass expands and cools adiabatically which results in increasing relative humidity, creating clouds and precipitation. [44]

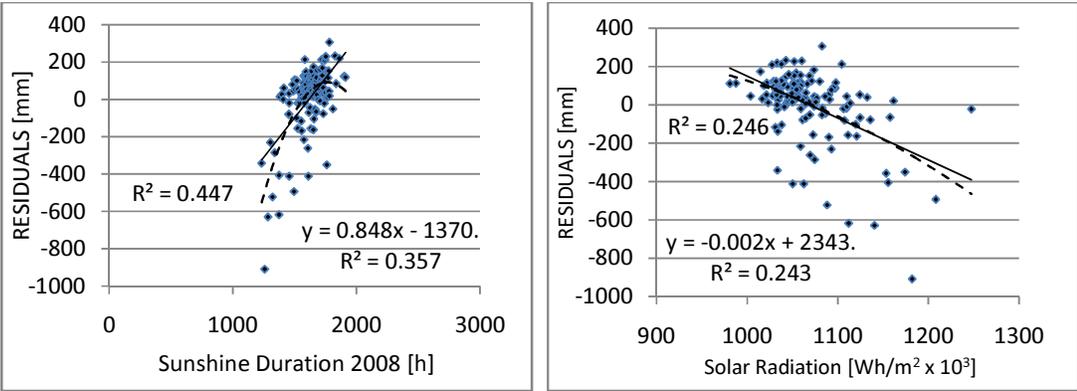
As shown in fig. 5.12a, altitude has a strong influence on residuals. Altitude itself can already explain 51% of variation in residuals. The lowest residuals around 0 are in altitudes between 400 m and 500 m. Altitudes lower than 400 m show generally positive residual values, while most of the altitudes above 500–m have negative residual values. The highest residuals appear at the highest altitudes.



Figures 5.13a and b: Mean Temperature vs. annual and monthly residuals

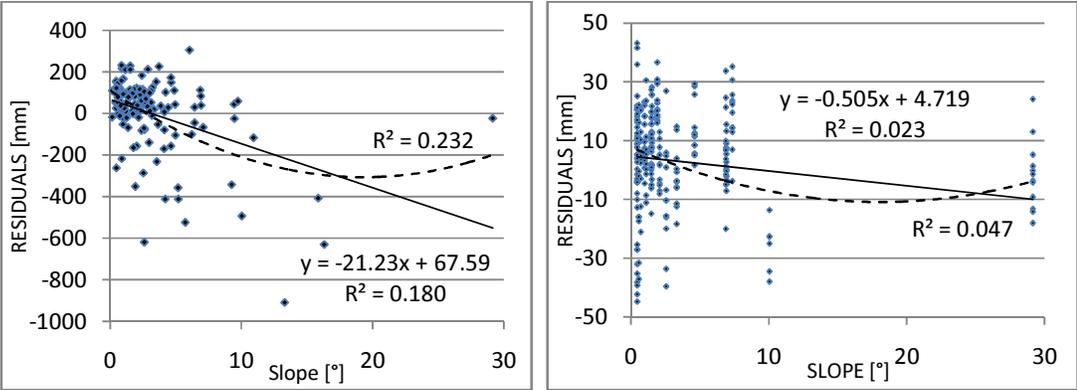
Fig. 5.13a reveals a relatively strong linear relationship ($R^2 = 0.40$) between mean annual temperature and residuals. Higher residual values emerge at colder stations, which again, lie in higher altitudes. Mean temperature and altitude are highly correlated variables ($R^2 = 0.90$ and higher) [35] and therefore its inclusion in regression models together with

altitude should be considered. However, only one of the two variables should be used as predictor in one model. The monthly dataset on the other hand shows no correlation of residuals and monthly mean temperature. This is not a direct effect of temperature but is caused by altitude. There is a high correlation with altitude for annual averages and a much weaker relationship of temperature and altitude for monthly averages, because one can get high temperatures at high altitudes during summer.



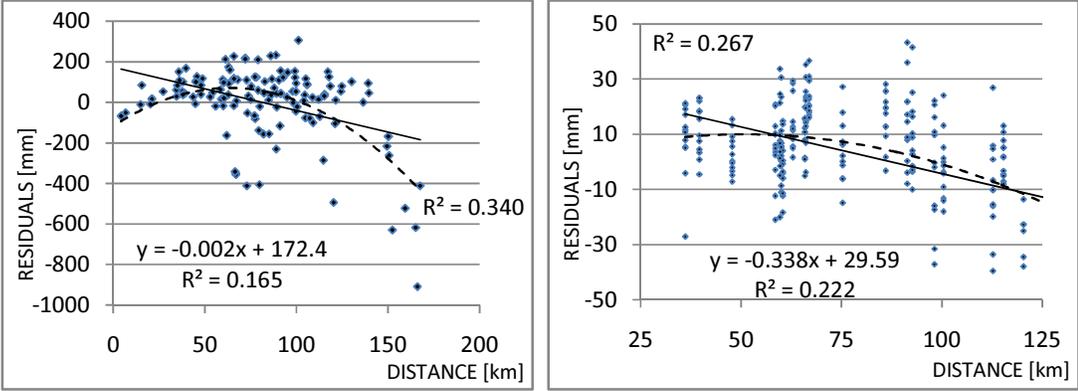
Figures 5.14a and b: Sunshine Duration and solar radiation vs. annual residuals

Figures 5.14a and b are both from the annual dataset and show solar influence on residuals. It is interesting to note that the trend is negative in the case of solar radiation and positive in the case of sunshine duration. Sunshine duration is a physical measure monitored at the meteorological stations (takes into account cloud cover), while solar radiation is a theoretical measure derived from DEM and calculated using ArcGIS. Solar radiation reveals some linear trend which considering the R^2 of 0.24 could explain ¼ of variation of residuals. Solar radiation contains information about slope and aspect from the DEM. Both slope and aspect can affect radar measurements as shown below in figures 5.15 and 5.16.



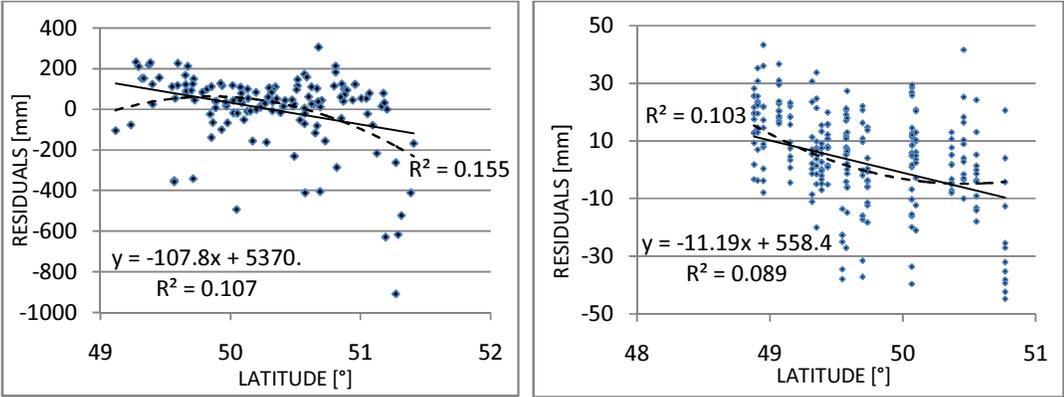
Figures 5.15a and b: Slope vs. annual and monthly residuals

Regarding slope in fig. 5.15a, the majority of stations lie on flat surface of slope 5° or lower. Generally speaking, the higher the slope, the higher the residuals (negative, which means radar is under-predicting).



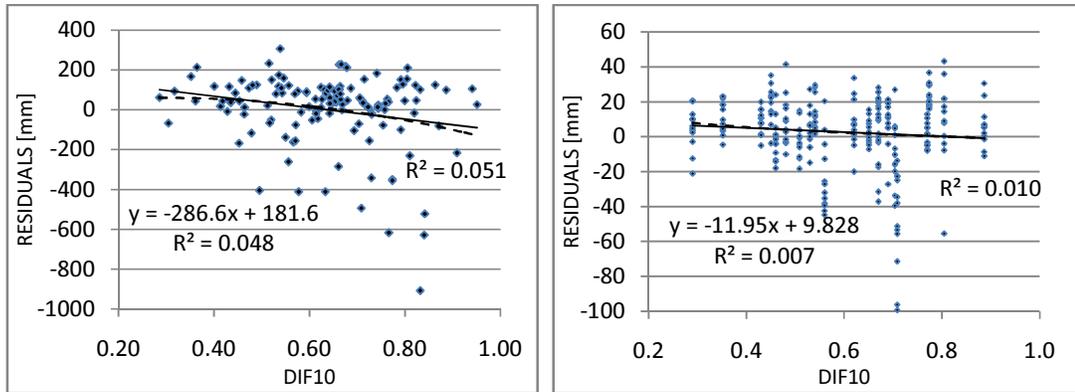
Figures 5.16a and b: Distance from the radar antenna vs. annual and monthly residuals

Fig. 5.16a focuses on the distance from the radar antenna. The relationship between the distance and residuals is rather quadratic than linear. This can be explained by the stations situated too close or too far from the antenna resulting in higher residuals. Highest residuals (also negative) are associated with those stations farthest away from the radar antenna. Lowest residual values are at distances ranging from 50 to 100 km.



Figures 5.17a and b: Latitude vs. annual and monthly residuals

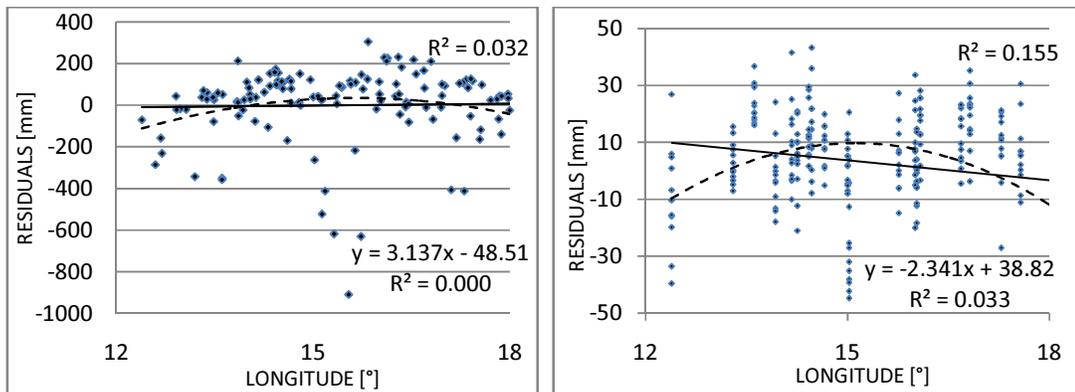
Latitude in fig. 5.17a does not reveal any strong relationship with residuals, but the outliers lie in latitudes above 51°, which is around the northern border, where mountains spread out.



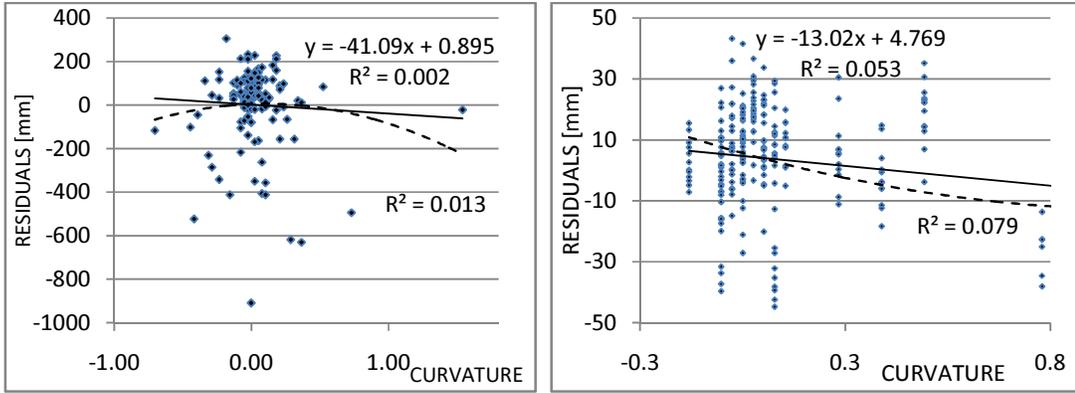
Figures 5.18a and b: Directional Difference smoothed to 10km vs. annual and monthly residuals

Directional difference (DIF) between the direction towards the radar antenna and aspect is displayed in fig. 21a. In other words, this variable describes the horizontal angle of the radar beam and the reflectance area. Higher correlation, but still very low ($R^2 = 0.048$), between directional difference and residuals was found for DIF smoothed to 10 km compared to 1 km or 90 m.

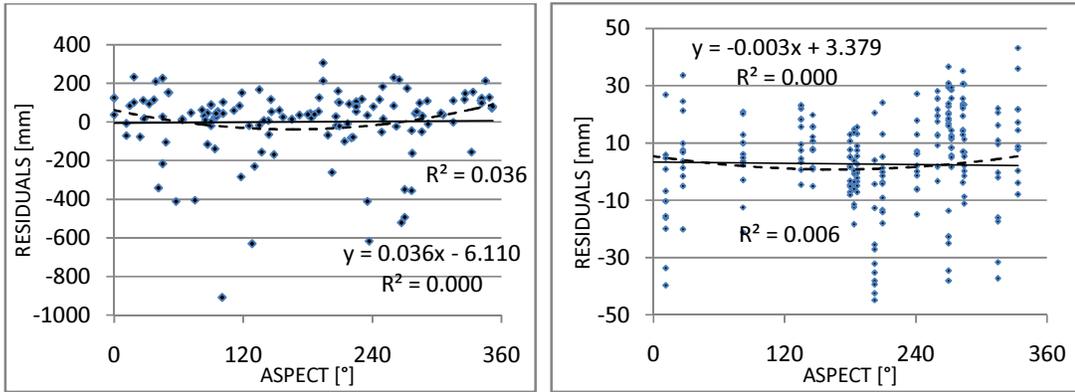
Longitude in fig. 5.19a nor curvature in fig. 5.20a nor aspect in fig. 5.21a does not have any significant influence on residuals. In the case of curvature, residual values around 0 are spread out around curvature of 0, where the terrain is flat.



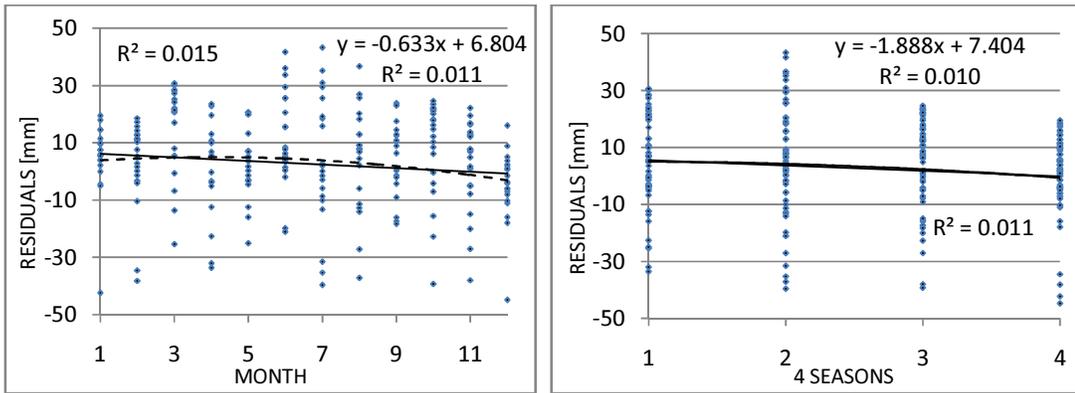
Figures 5.19a and b: Longitude vs. annual and monthly residuals



Figures 5.20a and b: Curvature vs. annual and monthly residuals



Figures 5.21a and b: Aspect vs. annual and monthly residuals



Figures 5.2a and b: Residuals vs. months and seasons

Table 5.1: Legend for 4 seasons (fig. 5.22b)

4 SEASONS:

- 1 - Spring
- 2 - Summer
- 3 - Fall
- 4 - Winter

Neither months nor 4 seasons show significant correlation with residuals, concluding that there is no seasonal variation in radar-rainfall prediction.

5.4 Findings

The correlation coefficients (R^2) obtained through the annual residual plots show how certain factors influence the residuals. Monthly residuals from a small number of stations are less representative and are therefore included mainly for comparison with annual residuals. Tab. 3 summarizes the annual R^2 values which are sorted in descending order according to the linear regression fit. ‘*’ in tab. 3 means that SD08 was excluded from the analysis, because not all of the gauge stations measure sunshine duration (SD08). An additional reason for this is that sunshine duration would have to be interpolated over the whole area with certain precision in order to be used as areal prediction factor.

Table 5.2: Regression coefficients between annual residuals and various factors

	R^2	
	<i>linear</i>	<i>polynomial</i>
ALTITUDE	0.514	0.537
MT08	0.404	0.474
*SD08	0.357	0.447
SOLARRAD	0.243	0.246
SLOPE	0.180	0.232
DIST	0.156	0.338
N_LAT	0.107	0.155
DIRDIFF10	0.048	0.051
DIRDIFF5	0.041	0.043
DIRDIFF	0.019	0.041
CURVAT	0.002	0.013
E_LON	0.000	0.032
ASPECT	0.000	0.036

Factors with the highest R^2 values have the highest influence on residuals and hence can bias radar measurements. Multiple regression analysis was performed with the factors in table 5.2 lying above the dashed line (the line was drawn arbitrary). Significant factors were identified using backward stepwise approach, which is documented in table A-2.1 in appendix 2.

Significant factors are found ALT, N_LAT and DIST. Relationship between distance and residuals is better described by a polynomial relationship rather than linear as shown in fig. 19a. The polynomial relationship is also underlined in tab. 5.2 where the R^2 reaches the

value of 0.338. Distance squared (DIST2) was therefore introduced to the regression model to simulate 2nd order polynomial (quadratic) fit. The results are stronger than in the linear case, since R^2 increases from 0.66 to 0.74 (see tables A-2.1 and 2.2. in appendix 2)

The three factors included in the multivariate polynomial regression model can explain 74% of variance of residuals. This finding is important because it means that the residuals are predictable from topographic and locational variables and not a consequence of random variation.

5.4.1 Discussion

The findings above are important because it means that the residuals are predictable from topographic and locational variables and are not only a consequence of random variation. The regression analysis was performed including the station suggested for exclusion in section 5.3.2.1. The exclusion of the 3 stations mentioned earlier would lead to slightly better fittings of the regression curves and slightly higher regression coefficients. Because such an improvement would not be significant, the three stations were not left out from the analysis.

Second, stations with residuals greater than ± 300 mm per annum should not be excluded from this analysis, because all of these stations lie in high altitudes and represent the mountainous regions where precipitation is hard to model and always under-predicted by the radar. This can be caused by the mountains acting as natural obstacles blocking the radar beam.

Using radar data in precipitation modeling can be easier in low to mid altitudes, as evident in figures 5.9–5.12, but it can be difficult, if not impossible, in mountainous regions.

Findings are the result of analyzing residuals between radar and gauge rainfall sums at annual and monthly scale. Looking at finer temporal resolution could introduce some other topographic, locational or atmospheric variables.

5.5 Significance

Findings confirm the hypothesis that radar measurement errors are not only a cause of random variation but are significantly affected by the following topographic and locational variables: altitude, distance from the radar beam and less significantly by latitude. Including these three geographic variables in radar measurements in some form of location specific calibration could improve the radars accuracy and remove up to 74% variance of residual errors.

5.5.1 Suggestions for Future

Having monthly rain gauge data for all of the meteorological stations in the Czech Republic one could increase the number of statistical samples for analysis from 134 to

$$12(\text{months}) \times 134(\text{stations}) = 1608(\text{samples}). \quad (5.2)$$

Having wind direction observations it would be possible to develop and test new interactive variables such as the product of slope and orientation (orientation of the prevailing winds at some specific height) or exposure of a slope with regard to wind directions.

6 Analysis

The magnitude of the multivariate regression analysis is in selecting the best fitting model by choosing significant independent variables. For the residual correction, an appropriate local interpolation technique needs to be selected.

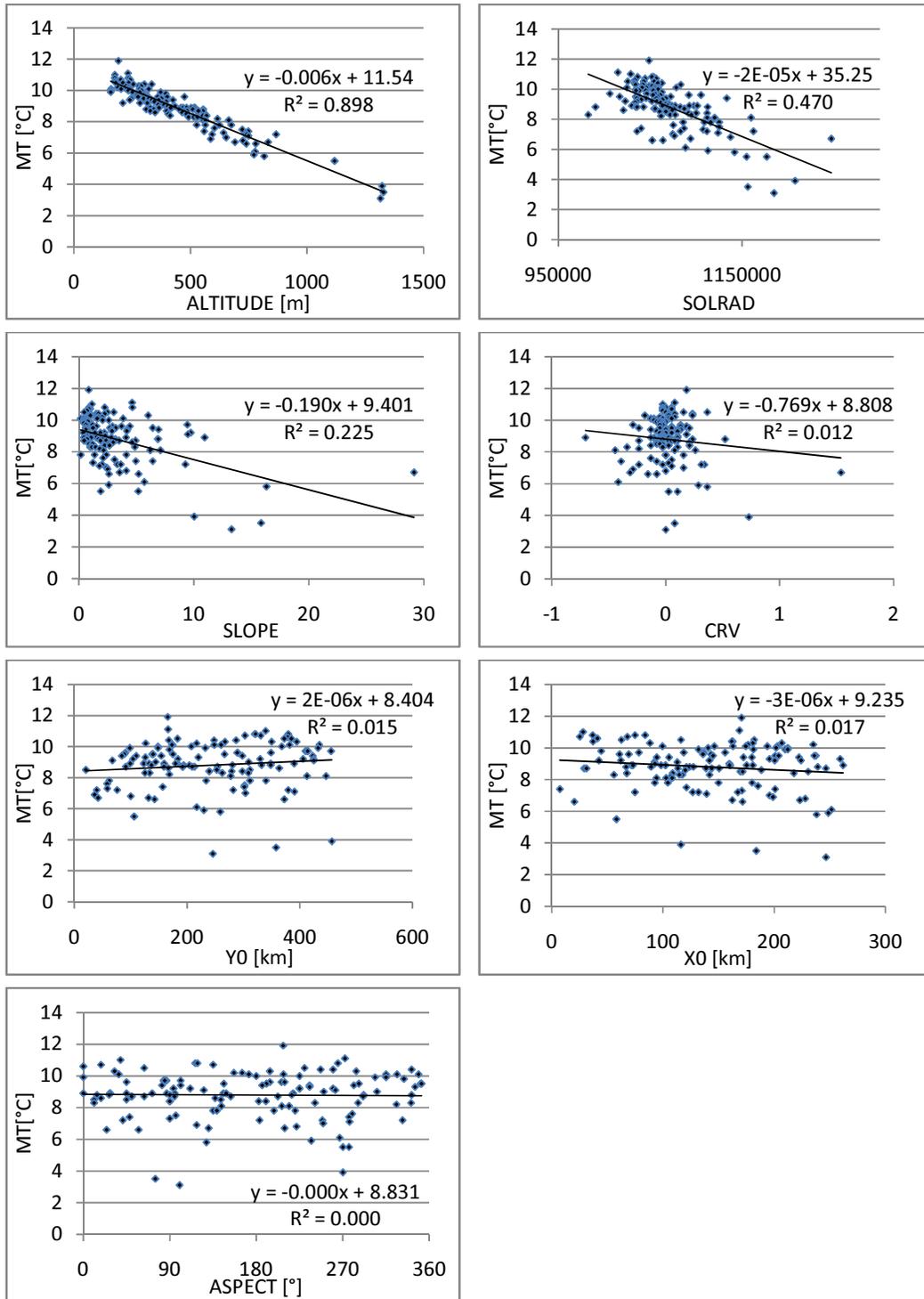
6.1 Linear Relationships

The first reasonable step in multivariate regression analysis is finding linear relationships by looking at scatter plots and correlation coefficients of the independent variables with the climate variables. The relationship between climate and independent variables is expressed by the correlation coefficient R^2 . Tab. 6.1 shows the most significant linear predictors followed by scatter plots with linear regression equations with R^2 produced in MS Excel. The dashed line in tab. 6.1 and 6.2 separates the significant from non-significant variables based on the probability of F-statistics.

Table 6.1: Correlations of various independent variables with MT (dataset A, 132 observations)

MT 2008			
Variable	R^2	F	prob F
ALT	0.898	1146.7	0.00
SOLRAD	0.470	115.5	0.00
SLP	0.230	37.8	0.00
Y	0.017	2.3	0.13
X	0.015	2.0	0.16
CRV	0.013	1.7	0.19
ASP	0.000	0.3	0.85

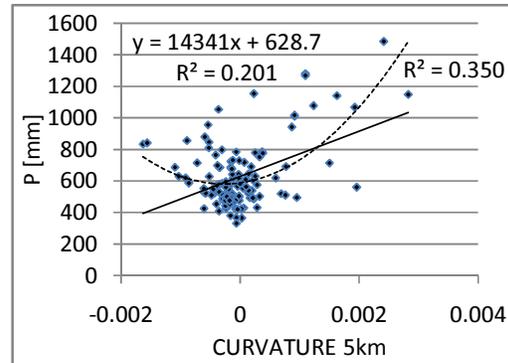
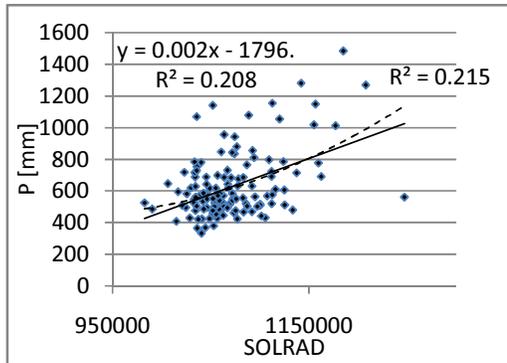
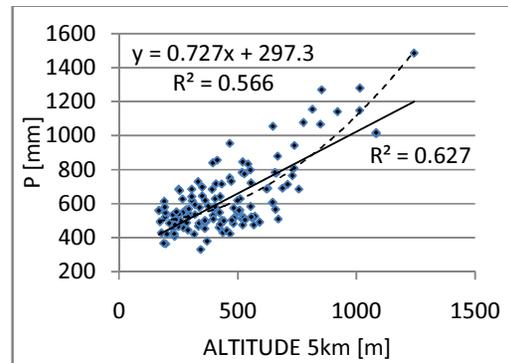
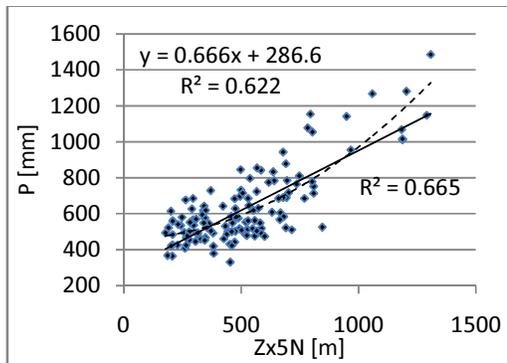
The best predictor of MT is undoubtedly altitude (see fig. 6.1, tab. 6.1). Aspect on the other hand shows no correlation with MT, although it was expected to be a useful predictor. Notice the cloud of points in slope as well as in curvature scatter plots in fig. 6.1 and 6.2. Neither of them should be used as predictors, since the slope of the curve (or shape in case of quadratic regression) is determined by only a few outlying points.

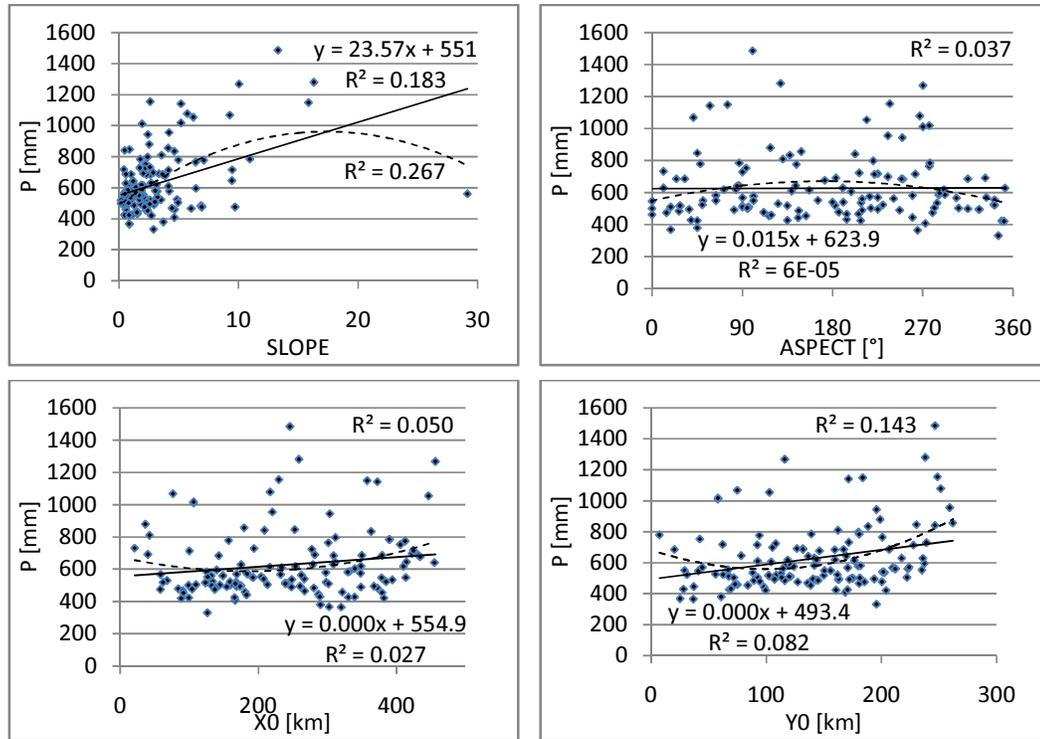


Set of figures 6.1: Scatter plots of various independent variables with MT (dataset A, 132 observations)

Table 6.2: Correlations of various independent variables with P (dataset A, 132 observations)

variable	R ²		F	prob F
	linear	poly 2nd		
ZxN5	0.622	0.698	214	0.00
ALT5km	0.578	0.627	178	0.00
ZxW25	0.549	0.628	158	0.00
ALT1km	0.539	0.592	152	0.00
ALT	0.497	0.526	128	0.00
SOLRAD	0.209	0.215	34	0.00
CRV5km	0.202	0.350	32	0.00
SLP	0.183	0.267	29	0.00
X0	0.079	0.141	11	0.00
Y0	0.028	0.051	4	0.05
CRV1k	0.000	0.309	0	0.81
CRV	0.000	0.012	0	0.96
ASP	0.000	0.037	0	0.93





Set of figures 6.2: Scatter plots of various independent variables with P (dataset A, 132 observations)

6.2 Backward Stepwise Approach

Selecting significant predictors is in fact building the regression model. The backward stepwise approach was performed manually in Geoda software. In the case of MT, there were 7 independent variables tested: ALT, X0, Y0, SLP, ASP, CRV, and SOLRAD. Below is a table showing the backward elimination of independent variables for dataset B, tables for dataset A are in Appendix 2. ‘*’ denotes the best fitting (but still reasonable) model, final model selected with respect to the results of backward elimination of each year in dataset A are marked in bold.

Table 6.3: Backward elimination of ind. variables in MT prediction (dataset A, 132 stations)

MT2008													
R ²	adj.R ²	ALT		Y		CRV		ASP		SOLRAD		SLP	
		t	prob.	t	prob.	t	prob.	t	prob.	t	prob.	t	prob.
0.921	0.918	-23.57	0.000	-5.74	0.000	0.89	0.378	1.21	0.229	1.08	0.281	0.62	0.539
0.921	0.918	-25.10	0.000	-5.75	0.000	1.13	0.259	1.15	0.251	1.07	0.288		
0.921	0.918	-37.75	0.000	-5.67	0.000	1.89	0.062	1.27	0.205				
*0.920	0.918	-37.64	0.000	-5.56	0.000	1.95	0.053						
0.917	0.916	-37.43	0.000	-5.43	0.000								
0.898	0.897	-33.86	0.000										

$$MT = 12.010 - 0.00603761 * [DEM-srtm] - 0.000003320479 * [Y0]$$

In the case of P, there were all together 26 independent variables. For dataset B the backward stepwise approach is shown in tab. 6.4 below, but only the variables eliminated without the t-statistics, due to space constrains.

Table 6.4: Backward elimination of independent variables in P prediction (dataset A, 132 stations)

R²	adjR²	F	MK	P2008: Backward Stepwise Approach
0.86	0.82	21	413	ALT_1k_5k_10k,X,Y,SOLRAD,ASP,SLP,CRV_1k_5k,16 x Zx
0.86	0.82	28	363	ALT_5k_10k,X,Y,SOLRAD,ASP,SLP,CRV,14 x Zx (-ZxSW5 - ZxW10)
0.85	0.83	37	299	ALT_5k_10k,X,Y,SOLRAD,ASP,SLP,10 x Zx (-ZxNW5 - ZxN25 - ZxSW25 -ZxW5)
0.85	0.83	45	255	ALT_10k,X,Y,SOLRAD,ASP,SLP,8 x Zx (-ZxN10 -ZxSW10)
0.85	0.83	51	231	ALT_10k,X,Y,SOLRAD,ASP,SLP,6 x Zx (-ZxW1 - ZxN1)
0.84	0.83	59	97	ALT_10k,X,Y,ASP,6 x Zx (-ZxW1 - ZxN1)
0.84	0.83	70	85	ALT, X, Y, ZxW25, ZxSW1, ZxNW1_10_25, ZxN5
0.82	0.82	98	68	ALT, X, Y, ZxW25, ZxNW1, ZxN5
*0.82	0.81	143	11	X, Y, ZxW25, ZxN5
0.80	0.79	166	10	X, Y, ZxN5
0.77	0.76	106	11	ALT1k, X, Y, ZxW25

$$P = 0.5065762 * [ALT1k] + 0.0004720109 * [X0] + 0.001042431 * [Y0] + 0.2504948 * [ZxW25] - 27.013$$

Tables of dataset A are included in Appendix 2. Having to test about 26 independent variables it would be time demanding to carry out the selection approach manually, excluding one by one variable, for each year and each climate variable. In order to cut the time needed for the selection process for dataset A, two shortcuts had to be ‘taken’. Since altitude is acknowledged as a good predictor, 4 altitude variables (1 from the DEM and 3 smoothed to different scales) were tested separately and only the one most significant entered the backward stepwise selection with all other independent variables. Second, only those variables proven to be significant in the dataset B in year 2008 were included in the backward stepwise selection, because dataset B with 132 observations is more representative. Third, all 16 Zx variables were included, but only 3-5 Zx variables with the student t-test greater than 1 were further taken into account. This way the number of independent variables was reduced and the backward stepwise approach was performed from this point the same way as in the case of MT, excluding one by one variable, which is documented on the DVD attached in folder \regression_protocols\ (containing about 110 protocols from Geoda).

6.3 Regression Models

Based on the backward selection process, final regression models look as following:

- Regression model for MT has two independent variables:

$$MT = b_o + b_1ALT + b_2Y0 \quad (6.1)$$

- Regression model for P has four independent variables:

$$P = b_o + b_1ALT1km + b_2Y0 + b_3X0 + b_4ZxW25 \quad (6.2)$$

Values of regression coefficients b_0 - b_4 are in equations below the tables of backward selection in Appendix 2. Both linear models provide satisfying predictions and therefore higher-order polynomial models were not tested. Altitude and latitude can explain more than 90% of the spatial variability of MT. Precipitation is in general more difficult to predict and therefore more variables are needed in order to explain more than 78% of the variability of P.

Regression equations were put into the ‘Raster Calculator’ in order to create ‘potential maps’.

6.4 Corrector Maps

Corrector maps are calculated by interpolating of residuals. For both MT and P, IDW was used with the power of 2 and the maximum of 10 possible neighboring points that can be used to calculate the value at each cell. This method was suggested by Ninyerola [38] for MT because it yielded better results than splines. Corrector maps are classified into 9 equal interval classes for visualization purposes.

In the case of P, IDW with the power of 1, 2, and 3, spline with the tension of 400 and kriging was tested using independent observations of dataset B. Among IDW techniques, the power of 2 was slightly better than 1 or 3. Splines yielded the same results as IDW with the power of 2. For the kriging method, a semivariogram was created (fig. 6.3), which did not show any characteristic autocorrelation curve fitting the residuals. In addition, there was a long time needed for computation of ordinary kriging, therefore the IDW was selected.

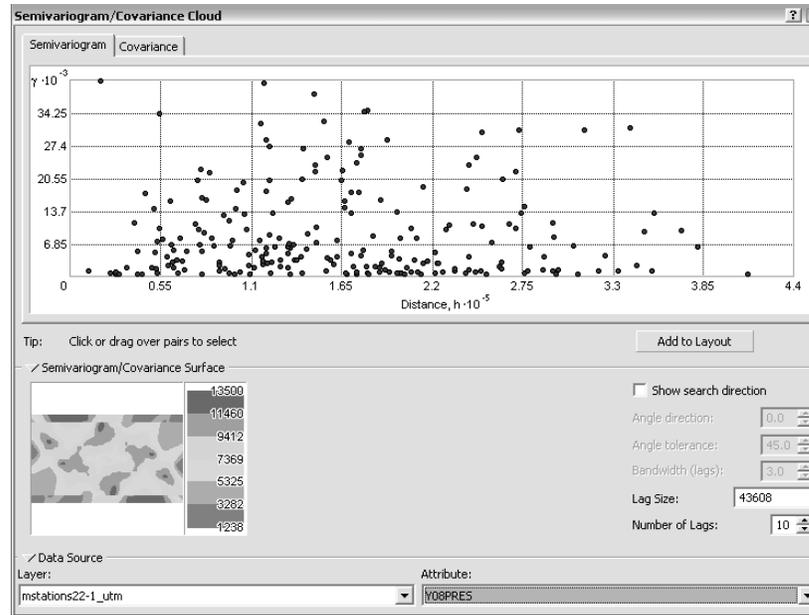


Figure 6.3: Semivariogram of residuals of P in 2008

In figures 6.4 and 6.5 on the next page are two examples of a residual correction map for the year 2008 given for comparison and illustration. Areas in green with negative residuals are locations, where the model over-predicts the reality, while areas in red and white colors are where the model under-predicts the reality. Corrector maps are added to 'potential maps' (see eq. 4.6) using 'Raster Calculator' in order to obtain the final climate surface. Corrector maps for MT and P are stored on the DVD attached in the \maps\MT\corrector_map and \maps\P\corrector_map.

The overall pattern among the twelve MT corrector maps is under-prediction in the eastern and sometimes central part of the CR and over-prediction in the western part of the CR. The area around Prague is where the lowest values appear, where temperature is always under-predicted. This has most likely something to do with the heat island problem. Peaks of over-prediction where temperatures are usually lower are located around the stations 'Tabor' and 'Pribyslav' in the centre of CR, 'Liberec' in the North and stations onwards from 'Olomouc' in the east.

The overall pattern of P corrector maps is under-prediction along the borders in the south-west, north and the west and over-prediction in the central part and western part. Also, station 'Milesovka' in the north-west receives less precipitation than predicted.

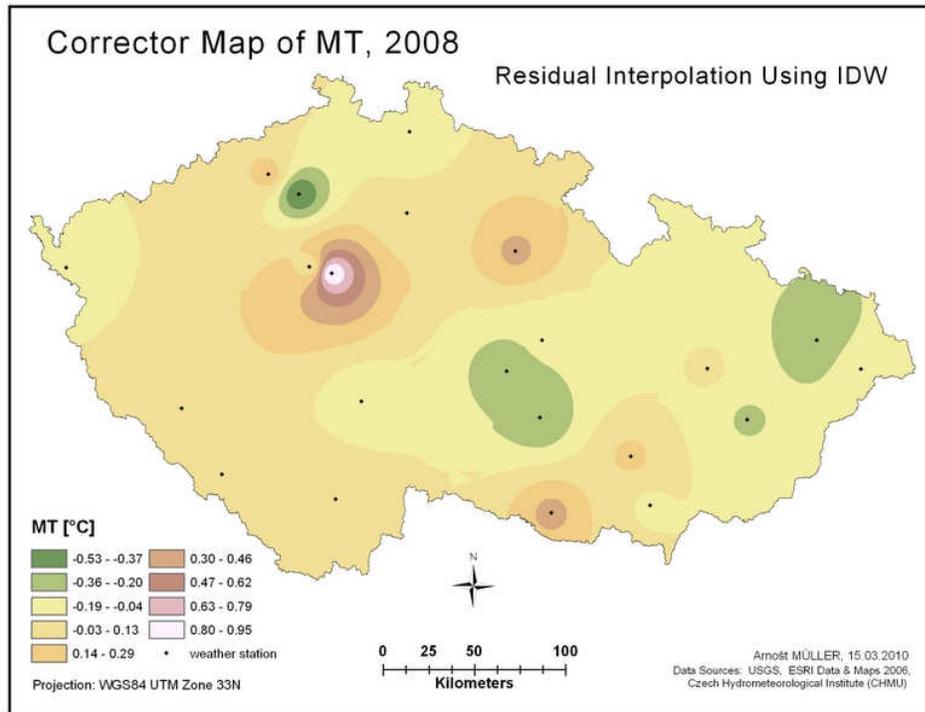


Figure 6.4: Corrector Map created by interpolating residuals of MT in 2008 (dataset A, 22 stations)

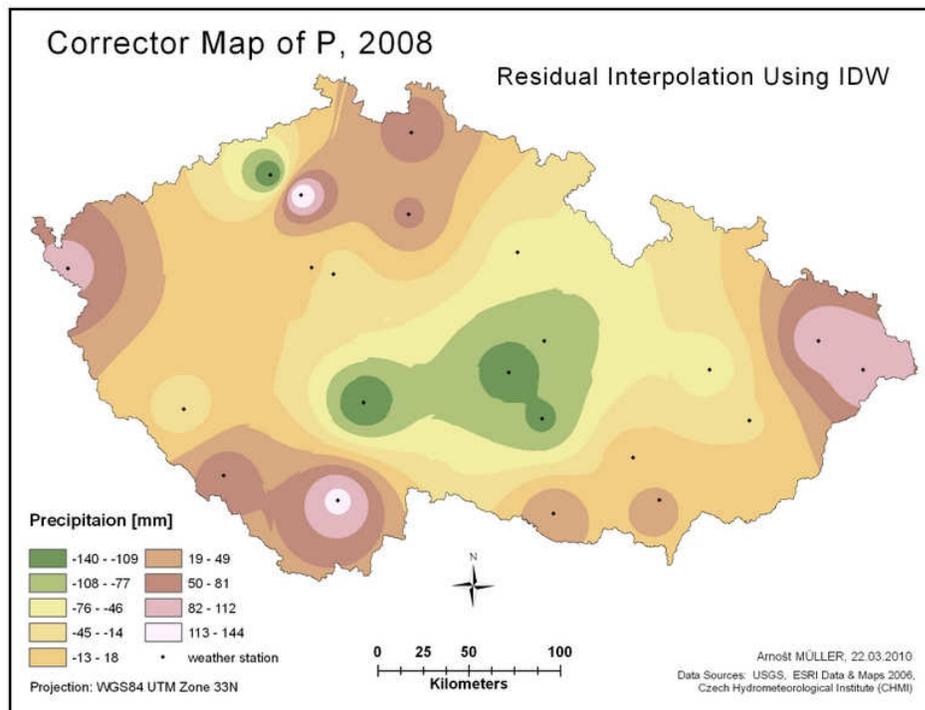


Figure 6.5: Corrector Map created by interpolating residuals of P in 2008 (dataset A, 22 stations)

6.5 MT and P Maps

Once the final climate surface is created, low-pass filter (focal mean of 3 surrounding cells in case of MT and 10 surrounding cells in case of P) is then applied to smooth the surface, particularly the edges, and to get rid of ‘lonely’ pixels. Final maps, 12 for each year and each climate variable (MT and P) are stored in IMG format and exported into JPG format. Four examples of final maps are in section 8.1, all from the year 2008. For comparison, there are maps created from 22 stations (dataset A) as well as from 132 stations (dataset B).

6.6 Deviation from Normal

The World Meteorological Organization (WMO) defines the temperature normal as the ‘period average, computed for a uniform and relatively long period comprising at least three consecutive ten-year periods’. In CR, CHMI publishes 30-years normals for the period 1961–1990. In author’s previous work, deviations from normals were for some computational problem calculated for the 22 stations and then interpolated using IDW. However it is much more effective to subtract the whole rasters, pixel by pixel, to calculate an anomaly map, according to the following equations, where $i = 1998, 1999, \dots, 2009$.

$$D_{MT} = MT_i - MT_{1961-90} \quad (6.3)$$

$$D_P = P_i - P_{1961-90} \quad (6.4)$$

Deviation rasters are smoothed with a low pass filter (focal mean of 20 neighboring cells) and stored at 180 m spatial resolution.

6.7 Visualization

For the animation using a GIF file and for the MapServer, final climate surfaces were smoothed (with the focal mean statistics, P with 10 and MT with 3 neighboring cells). MT surfaces were smoothed with 3 neighboring cells only in order to smooth the edges and keep the precision of measured values at meteorological stations.

It was crucial for the animation to maintain the same classes and layout, so that the only features changing is the map itself and the year. The same classification is essential, because it makes comparison among the twelve years possible. In addition, P layers were laid over a shaded relief with 10% and MT with 15% transparency in order to emphasize the climate-topography relationship. Animated GIF files were created with the free software GIMP with 3 seconds rate and the resolution of 1600px and 800px.

7 Accuracy

The dataset's accuracy should be well documented (in metadata) in order to enable users to estimate the reliability of any derived results. The accuracy of the resulting maps was assessed using statistical criteria mentioned in section 7.1. Nevertheless, not only statistical criteria were used to determine the validity of the interpolated climatic maps. Daly et al. emphasize that subjective evaluation of the reasonableness of the maps is worthwhile. The climatic maps were compared with the Climate Atlas of Czechia in section 7.4.

7.1 Statistical Criteria

Table 7.1: Statistical criteria used to assess the agreement of the models

Definitions	N ...number of observations F ...degrees of freedom O ...observed value \bar{O} ...mean of observed values P ...predicted value
Least-square regression	R^2 ...correlation coefficient $adj-R^2$... adjusted R^2
Mean bias error (MBE)	$MBE = \frac{1}{N} \sum_{i=1}^N (P_i - O_i)$
Mean absolute error (MAE)	$MAE = \frac{1}{N} \sum_{i=1}^N P_i - O_i $
Root mean square error (RMSE)	$RMSE = \sqrt{\frac{1}{F} \sum_{i=1}^N (P_i - O_i)^2}$ with an adjustment for a loss in degrees of freedom F
Root mean square error (RMSE _{ML})	$RMSE_{ML} = \sqrt{\frac{1}{N} \sum_{i=1}^N (P_i - O_i)^2}$ without an adjustment for a loss in F

Statistical criteria above serve to determine the error between model predictions and the real data recorded at the weather stations. Correlation coefficient (also known as coefficient of determination) R^2 is the first measure of the goodness of fit of the model. [31] R^2 varies between +1 and -1, both bracket values indicating a perfect fit. [23]

Since the magnitudes of R^2 are not consistently related to the accuracy of predictions and tend to over-estimate the goodness of fit of the models, more objective statistics is needed. Adjusted correlation coefficient ($\text{adj-}R^2$) compensates for this optimistic trait in R^2 by taking into account the size of the sample and the number of prediction variables. Unlike R^2 , $\text{adj-}R^2$ does not necessarily increase when additional variables are added to the model. [31]

RMSE and MAE are good overall measures of model performance, because they summarize the mean difference in the units of predicted variable. RMSE puts a lot of weight on high errors, while MAE is less sensitive to extreme values. [50]

7.2 Precision of DEM

The SRTM data meet the absolute vertical accuracy of 16 m (linear error at 90% confidence), respectively, as it was specified for the mission. The vertical accuracy is actually significantly better than 16 m, closer to 10 m, according to USGS. [49] Tab. A-3.1 in appendix 3 shows the differences for all 132 stations between real altitudes of meteorological stations (ALT) and altitudes derived from

- 1) the original DEM (ALT_{DEM})
- 2) the DEM resampled by bilinear transformation when converting to UTM ($\text{ALT}_{\text{DEM_utm}}$)

The resulting statistical characteristics are below in tab. 7.2.

Table 7.2: DEM's accuracy

1) DEM	2) DEM_{utm}
MBE= 2.3 m	MBE= 1.9 m
MAE= 5.0 m	MAE= 4.7 m
RMSE= 7.8 m	RMSE= 7.0 m

Tab. 7.2 confirms the overall good agreement (less than 7 m) in real altitudes and altitudes derived from the DEM as stated by the USGS. The DEM transformed to UTM shows even better agreement with reality than the original one.

7.3 Numerical Assessment

When there are many observations available, typically 60% of them are used for estimating the regression model and the remaining 40%, often randomly selected from the whole dataset, are used for independent validation. If the model is considered reliable and the results are satisfying, the remaining 40% often enter the regression as well to improve the model by including all of the stations.

The case of this thesis is somehow the opposite. Since there is only a small subset of observations available (22 stations) for each year, the detailed dataset B with 111 observations (132 minus 21 stations which are the same as in dataset A) is used for validation of the one year (2008). So there are only 18% of observations used in prediction and 72% for validation. This way only one year is independently validated, nevertheless it can provide a good measure of the overall performance of the regression models. The validation of the regression model for the year 2008 is in tab. A-3.2 in appendix 3. In tab. 7.3 below are the results for both MT and P.

Table 7.3: Validation of MT and P using dataset B (132 stations)

MT08		P08	
MBE =	-22.2 °C	MBE =	-196 mm
MAE =	0.3 °C	MAE =	67 mm
RMS =	0.4 °C	RMS =	89 mm
MAX =	1.6 °C	MAX =	280 mm
MED =	0.3 °C	MED =	53 mm

According to the results of validation for the year 2008, we can assume that the regression models will predict with the same efficiency in the remaining years as well. The overall accuracy of MT given by RMSE (standard deviation) is 0.4°C. In case of P the accuracy given by RMSE is approximately 90mm.

The overall accuracy of the final climate maps is characterized by RMS, calculated as a quadratic mean of standard deviations from Geoda’s protocols of regression (tab. 7.4). The accuracy of MT is 0.4 °C and 106 mm of P.

The fit of the MT model is excellent (R^2 between 0.90 and 0.97), and the fit of P model is better than expected (R^2 between 0.78 and 0.92). The following text compares R^2 and RMSE values to results from works of other authors. The final regression models in Arizona and New Mexico showed a higher degree of variance for temperature ($R^2 = 0.98$), but a higher root mean-squared error $RMSE = 0.74^\circ C$. [6] The relationship between the average annual air temperature and four terrain variables in Italy explained 92% of the variance and produced a standard error of 0.89 °C. [11] In the middle Ebro Valley in Spain, R^2 of annual mean temperature reached only 0.74 and RMS 0.62 °C, while precipitation reached 0.95 and RMSE 28.2 mm using the same combined regression method. [50] Ninyerola, the author of the combined residual method, obtained $R^2 = 0.84$ and $RMSE = 137.8$ mm for mapping annual precipitation of the Iberian peninsula. [37] In China, precipitation in the whole year was predicted with 72.6% and RMSE 8.4%. [46]

Table 7.4: R^2 , standard deviations from protocols of regression & quadratic mean (RMS)

Year	R^2		RMSE	
	MT	P	MT [°C]	P [mm]
1998	0.95	0.90	0.40	94.5
1999	0.96	0.92	0.37	72.3
2000	0.97	0.86	0.33	109.6
2001	0.96	0.82	0.36	141.2
2002	0.96	0.78	0.36	145.4
2003	0.95	0.87	0.41	72.7
2004	0.95	0.92	0.41	70.4
2005	0.93	0.89	0.49	102.4
2006	0.90	0.85	0.52	105.4
2007	0.97	0.85	0.34	105.8
2008	0.97	0.82	0.34	101.0
2009	0.95	0.83	0.40	120.5

RMSE = **0.40** **106.1**

7.4 Spatial Assessment

Spatial assessment is here called a visual comparison of the resulting maps with some other, ideally more precise, map. One comparison was already given in figures 6.5–6.8, with maps interpolated using the spatially denser dataset B with 132 observations. 22 observations, being very sparse for the whole area of CR, do not capture local anomalies, but are overall sufficient for predicting annual MT and P.

Maps in figures 7.1 and 7.2 were simply calculated by subtracting two maps of the same year (2008), one created from dataset A (22 observations) minus the one interpolated from dataset B (132 observations). The maps clearly show that the map created from dataset B is able to depict local anomalies (thanks to the residual correction method) while the map from dataset A is more general. Blue colors are areas that are colder/wetter than they should be, while areas in red/pink colors are warmer/drier than they should be.

Figures 7.3–7.6 provide an independent assessment of spatial patterns of MT and P through a comparison with a map from different source – the Climate Atlas of Czechia. Annual MT and P for the long-term period between 1961–1990 were interpolated using the same regression models (eq. 6.1 and 6.2) with 22 stations, and displayed using the same classification. Similar color scheme to the Climate Atlas was selected to allow for visual comparison, see figures 7.3–7.6. The confrontation in figures 7.3–7.6 confirms that the combined regression method yields excellent results of MT and reasonable areal distribution of P.

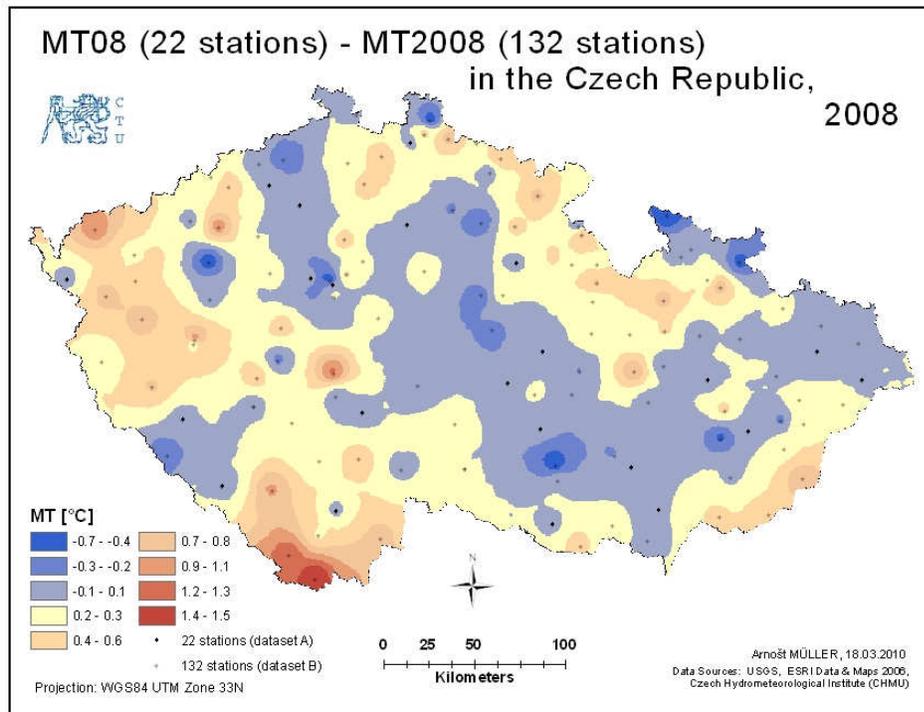


Figure 7.1: Deviation of MT (dataset A – dataset B)

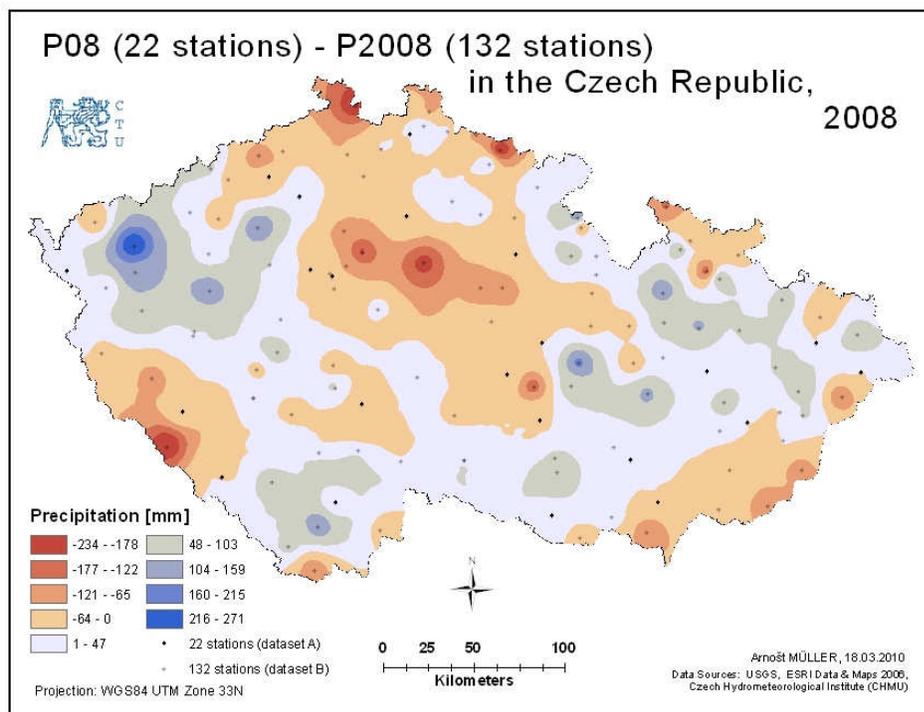


Figure 7.2: Deviation of P (dataset A – dataset B)

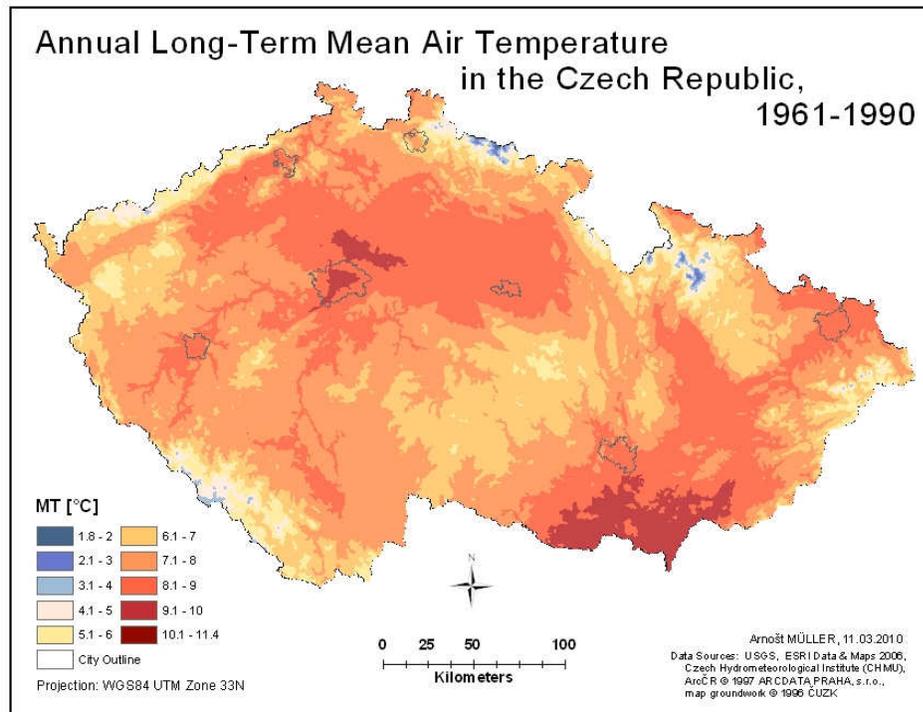


Figure 7.3: 30-year mean of mean air temperature

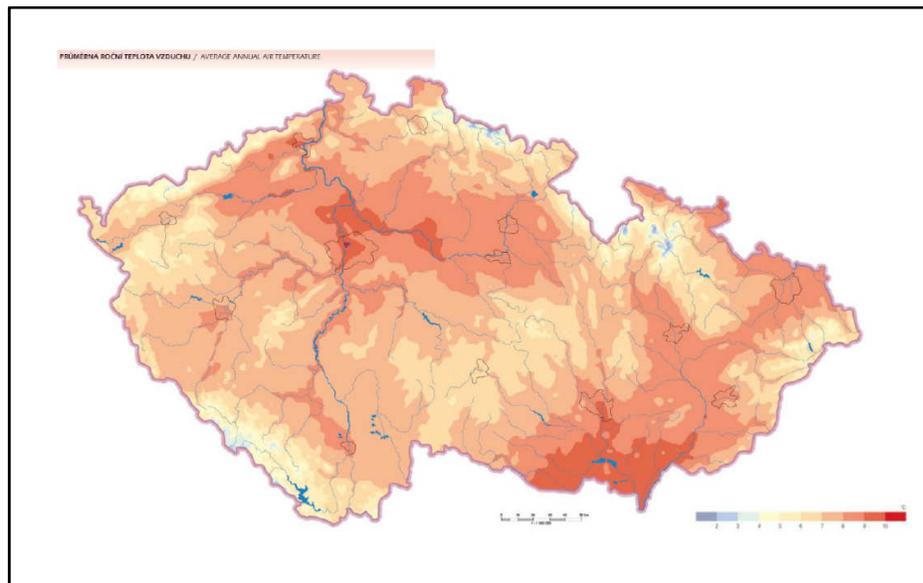


Figure 7.4: 40-year mean (1961-2000) of annual mean air temperature (Source: Climate Atlas of Czechia) [48]

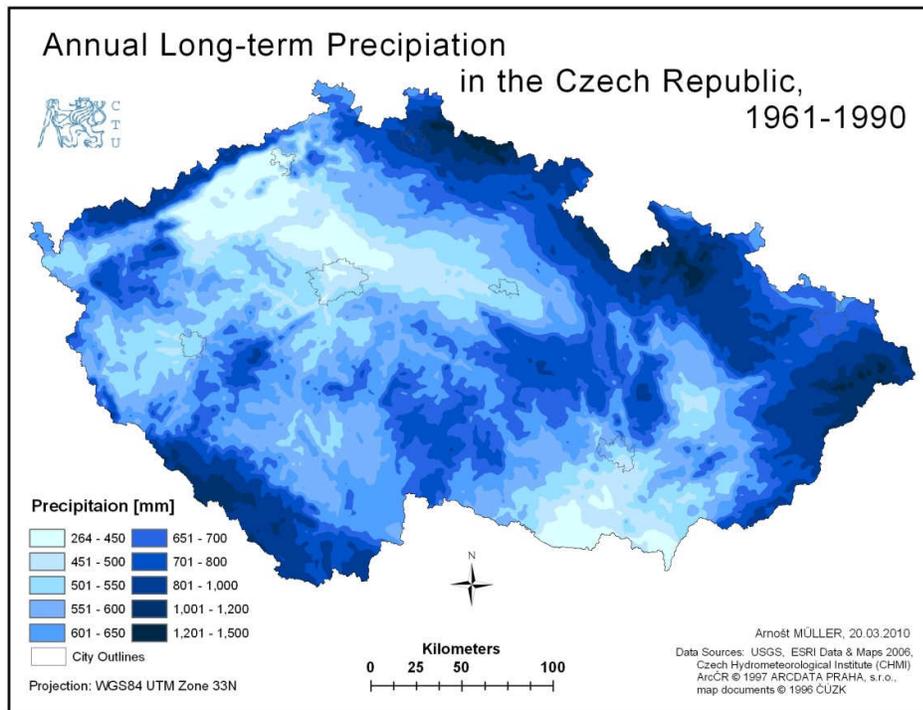


Figure 7.5: 30-year mean of annual precipitation

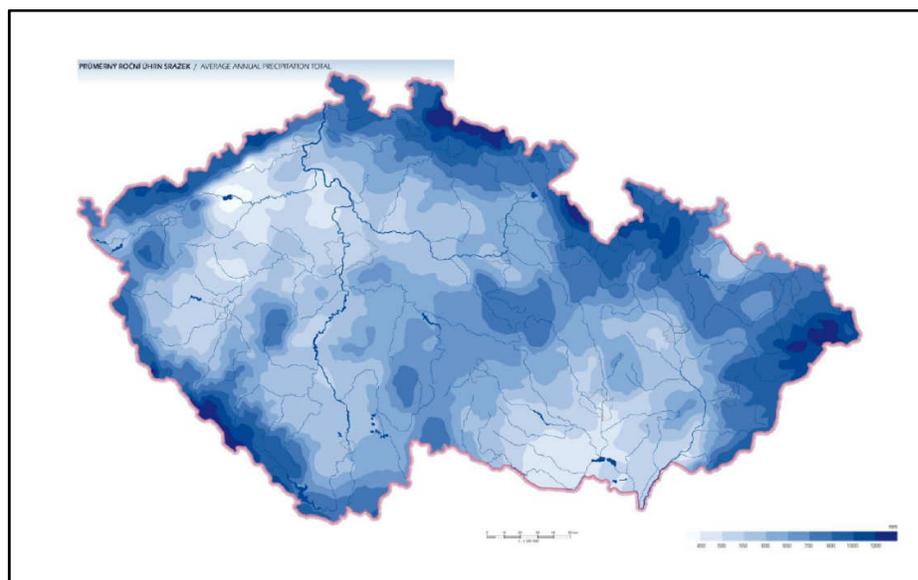


Figure 7.6: 40-year mean (1961-2000) of annual precipitation (Source: Climate Atlas of Czechia) [48]

8 Results

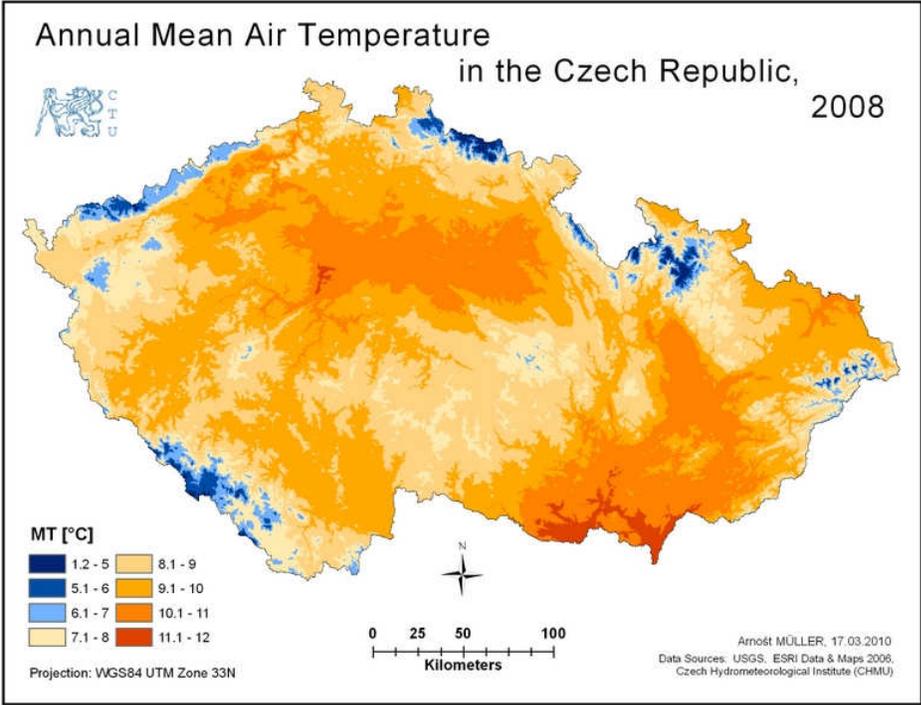


Figure 8.1: One example of final MT map

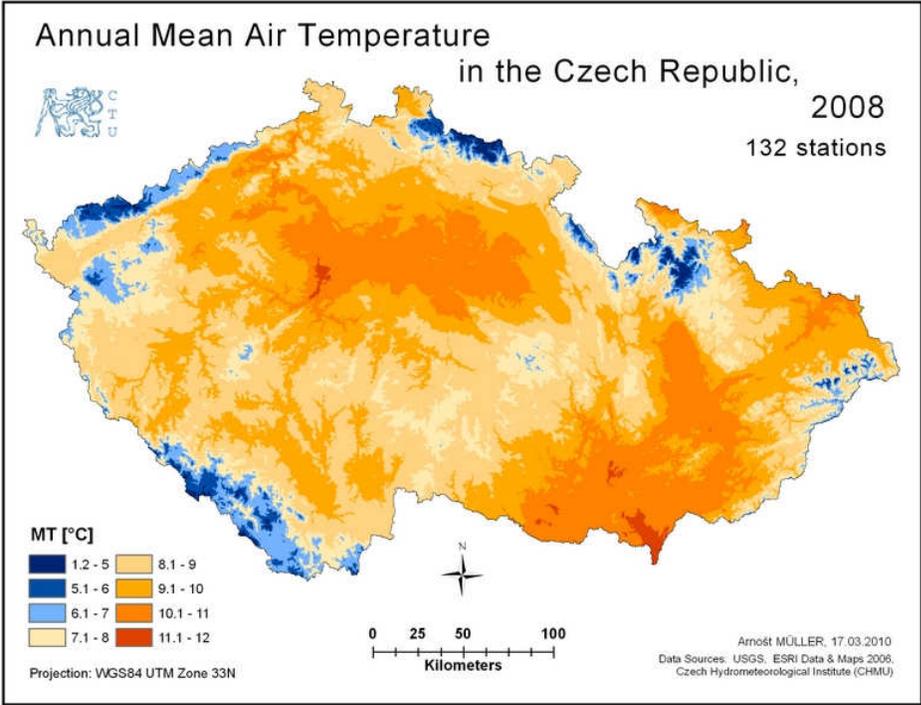


Figure 8.2: The same map as in fig. 8.1 but interpolated from 132 stations, included for comparison.

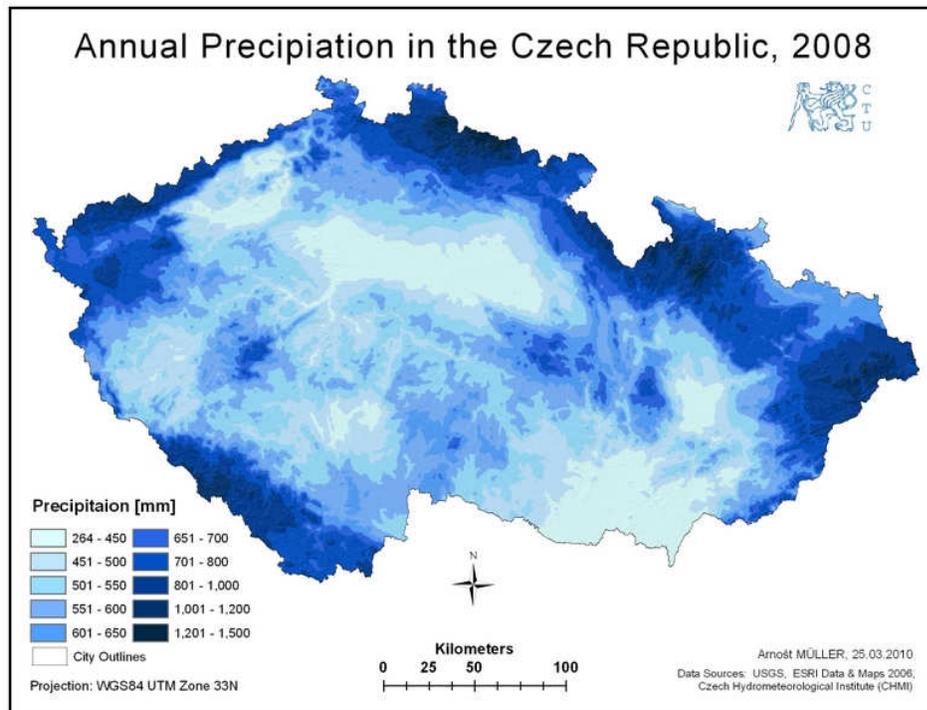


Figure 8.3: One example of final P map

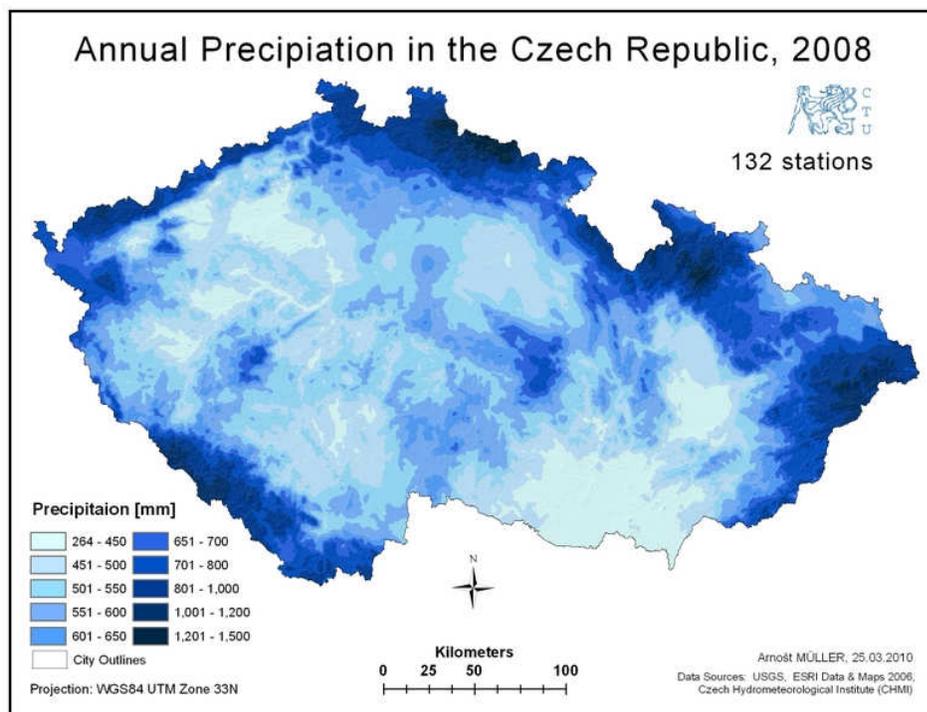


Figure 8.4: The same map as in fig. 8.3 interpolated from 132 stations and included for comparison.

8.1 Spatial Pattern of MT and P

Resulting MT, P maps and their deviations from normal are stored on the DVD attached, at <http://maps.fsv.cvut.cz/~muller/> as both static and animated maps, and also published interactively at <http://maps.fsv.cvut.cz/ka-map/>.

Spatial pattern of MT copies the topography. Since temperature is highly dependent on altitude, the warmest areas are river valleys in central Bohemia and south Moravia, while the coldest areas are situated along the borders in the mountains. There are only two focus points around Prague and in southern Moravia where annual MT reaches over 11 °C (only in the years 2000, 2007, and 2008). The spatial pattern of precipitation is similar to the MT. The rainiest places are along the borders in the mountains, while the driest areas lie in the mountain shadow in the west Bohemia and in the south Moravia. The year 2002 when CR was hit by a 100-year flood on the Vltava River is notably wetter compared to other years.

The agreement of the year 2008 MT map interpolated from 22 stations (fig. 8.5) with the more accurate map interpolated from 132 stations (fig. 8.6) is very good, although in general the map in fig. 8.5 is a little warmer. This is due to the residual correction, which allows taking into account local anomalies. Spatial pattern of P is also satisfying although the map created from dataset A is generally slightly over-predicted. The changing pattern of MT and P throughout the last twelve years is clearly to see in the animated maps.

8.2 Deviations from Normal

Maps of deviations from normal are published in the same way as MT and P maps, but are not included in the MapServer application. According to the deviations of MT the warmest years were 2000, 2002, 2007, and 2008. The only focal points colder than normal appear in years 1998, 2004, and 2005. Maps were smoothed with the 'Focal Mean' function with 20 neighboring cells. Using 'Raster Calculator', the arithmetic mean of the 12 deviation maps of both MT and P was calculated for each pixel. MT (fig. 8.5) in the last 12 years was higher than the 30-year normal, overall by 1.0 °C (mean value of the whole raster in fig. 8.5). In the case of P (fig. 8.6), less precipitation than normal fell in central Bohemia and northern Moravia by approximately 40 mm, while on the rest of the CR fell more P than normal ranging from 20 to 160 mm (mean value of the whole raster was 39.7 mm).

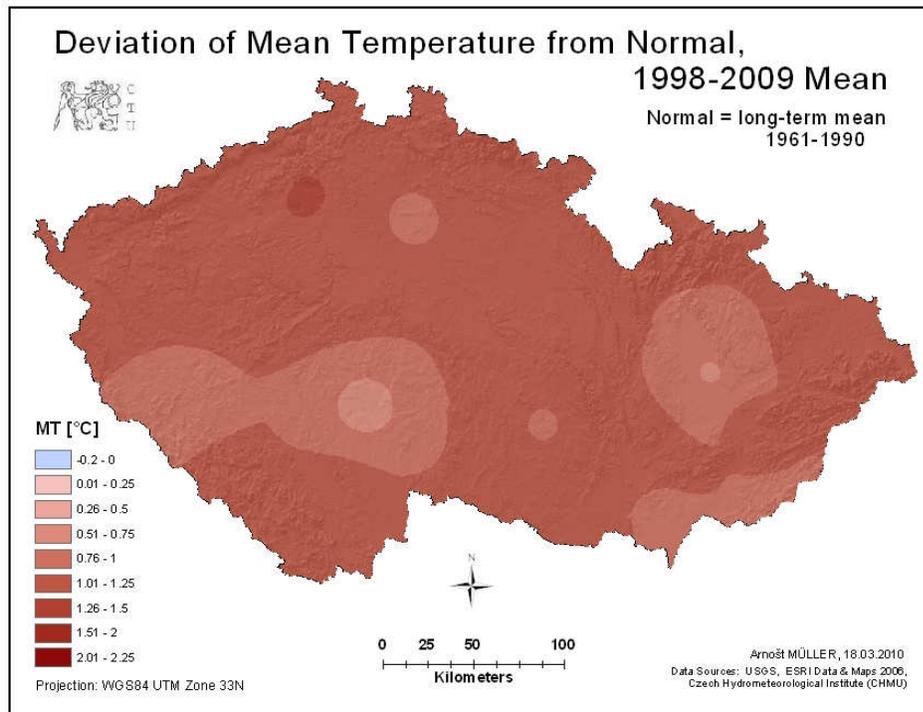


Figure 8.5: 12-year mean of deviation from normal of MT

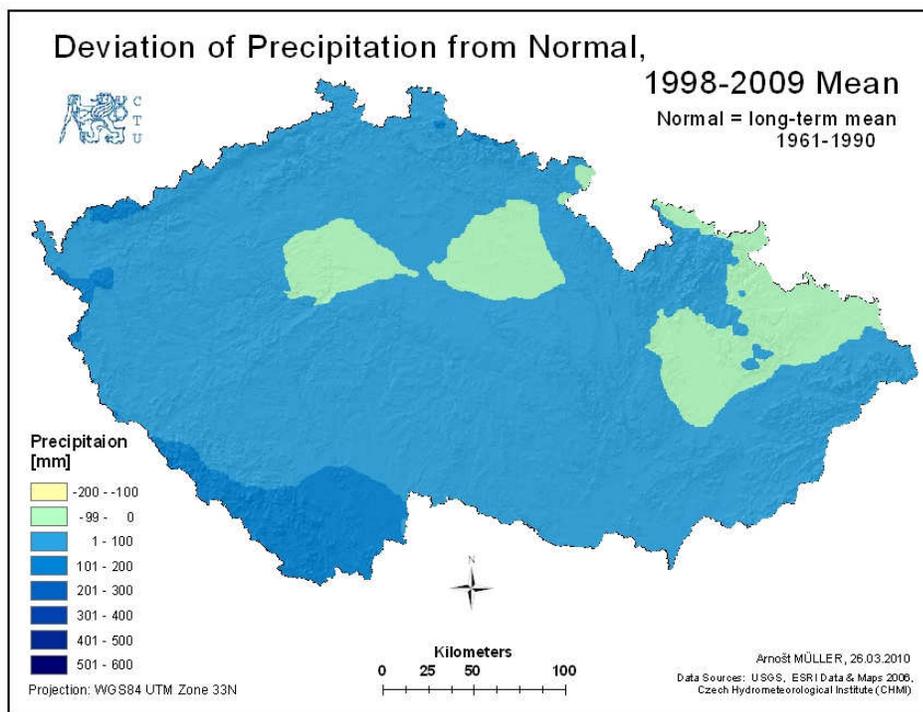


Figure 8.5: 12-year mean of deviation from normal of P

9 MapServer

The popularization of web mapping technologies by Google has encouraged the development of more interactive web mapping techniques. [33] For an ordinary user looking at a map on the Internet it is now much more than a simple look at a static map, which shows isolated information about a specific theme. By using a GIS it is possible to combine information and visualize them in form of layers. The Internet can be used in an effective manner to visualize as well as provide access to information for a wide range of users. [16]

A map server is in fact a GIS, which is operated by text parameters. Map servers are running 'above' a web server such as Apache, which makes a request by handing over the parameters. Map server uses information passed in the request and the mapfile (discussed later in section 9.2) to create an image of the requested map. [7] A variety of map servers exist. The commercial branch is certainly represented by ArcGIS Server from ESRI, or TopoL Internet Server and T-MapServer from Czech firms, while the Open Source branch is quickly approaching the commercial one. The most known Open Source map servers are GeoServer and the MapServer from the University of Minnesota (UMN). UMN MapServer is running at maps.fsv.cvut.cz at CTU and is therefore used in this thesis.

UMN MapServer is commonly referred to as 'MapServer'. MapServer is an Open Source geographic data rendering engine written in C. Developed at UMN in cooperation with NASA and the Minnesota Department of Natural Resources, MapServer is now a project of OSGeo, and is maintained by developers from around the world. The purpose of the MapServer is to dynamically display spatial data (maps, rasters and vector data) over the internet.

A simple MapServer application consists of:

- **Map File** - a structured text configuration file which tells the MapServer how to access data and draw the map, more in section 9.2.
- **Data** - MapServer can utilize many geographic data source types (vector formats such as ESRI shapefile, PostGIS, KML, or DGN, and raster formats such as TIFF/GeoTIFF, PNG and many others via GDAL).
- **HTML Pages** - are the interface between the user and the MapServer.
- **Template File** - controls how the maps and legends output by MapServer will

appear in the browser and also determines how the user can interact with the MapServer application (browse, zoom, pan, query).

- **MapServer CGI** - The binary or executable file that receives requests and returns images, data, etc. By default, this program is called 'mapserv'.
- **HTTP Server** - serves up the html pages for the user's browser. You need a working HTTP (Web) server, such as Apache, on the same machine as you have the MapServer.

Some of the features of the MapServer are [32]:

- Advanced cartographic output (scale dependent feature drawing, feature labeling including label collision mediation, fully customizable, template driven output, map element automation – scalebar, reference map, and legend)
- Support for popular scripting and development environments (PHP, Java, Perl, etc.)
- Cross-platform support (Linux, Windows, Mac, etc.)
- Support of numerous Open Geospatial Consortium (OGC) standards (WMS, WFS, etc.)
- Map projection support (on-the-fly map projection using the Proj.4 library)

9.1 ka-Map

MapServer alone does not provide the high level of interactivity, pre-rendering, caching of tile images, smooth panning, etc. [33] ka-Map is an open source template that uses a java script API for developing highly interactive web-mapping interfaces. [40] ka-Map coupled with MapServer is a powerful combination of open source web-mapping technologies. MapServer prepares the map images, and ka-Map serves them to the web browser. [33] ka-Map uses AJAX and the PHP MapScript to render maps. [33]It supports the usual array of user GIS interface elements such as: continuous panning without reloading the page, zooming to pre-set scales, scalebar, legend, and keymap support. [40] In summary, ka-map has 4 requirements, which need to be running together: Apache, MapServer, PHP and Mapscript.

The ka-map package is in fact a folder structure consisting of many HTML, PHP and CSS files. The whole structure occupies less than 4MB memory. ka-Map is not installed, you can start 'out of the box' with minimal configuration. In fact, you need to do 6 steps:

1. Decompress the package downloaded into a folder.
2. It is recommended to set a web server alias by adding the following script into web server's configuration file:

```
Alias /ka-map/ "[full path to ka-map folder]"
<Directory "[full path to ka-map folder]/htdocs/">
    Options Indexes
    AllowOverride None
    Order allow,deny
    Allow from all
</Directory>
```

This allows entering a simple URL (<http://maps.fsv.cvut.cz/ka-map/>) and having it pointed to the file path where ka-Map content is stored. It also sets the rights of access to the 'htdocs' folder for the client.

3. Rename the main configuration file `\ka-map\include\config.dist.php` to `config.php`.
4. Forth step is setting up two library pointers in the `config.php` file:

```
$szPHPMapScriptModule='php_mapscript.'.PHP_SHLIB_SUFFIX;
$szPHPGDModule = 'gd.'.PHP_SHLIB_SUFFIX;
```

5. Tell ka-Map where the mapfile is, as well as some other map-specific settings. This is done in the `config.php` file in the `$aszMapFiles` array:

```
$aszMapFiles = array(
    'mt' => array(
        'title' => 'Mean Temperatures',
        'path' => '/data/_projekty/amuller/mt.map',
        'scales' => array( 2000000, 1000000,500000,250000),
        'format' =>'PNG'),
    'P' => array(
        'title' => 'Precipitation',
        'path' => '/data/_projekty/amuller/p.map',
        'scales' => array( 2000000, 1000000,500000,250000),
        'format' =>'PNG')
);
```

ka-Map allows including multiple mapfiles, which can be selected in a drop down menu in the upper left corner of the application (fig. 9.1).

6. Set up the temporary folder, where kaMap creates its tile cache. The directory does not have to be web-accessible, but it must be writable by the web-server-user and allow creation of both directories and files.

```
$szBaseCacheDir = "/var/www/tmp/";
```

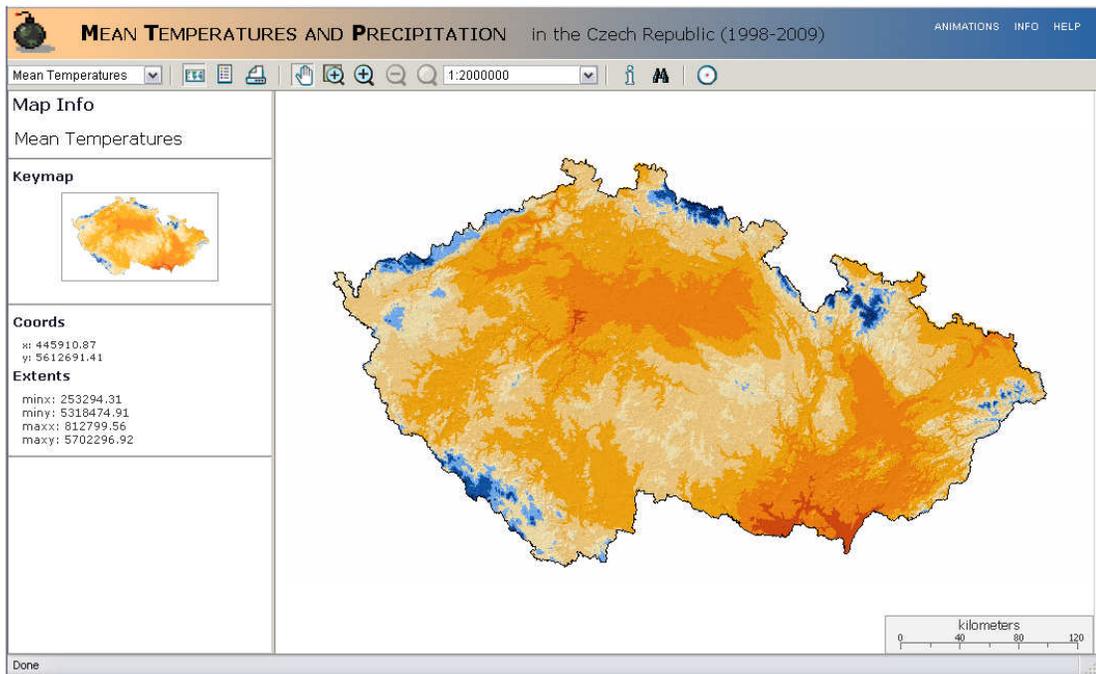


Figure 9.1: Screenshot of the MapServer with ka-Map

9.2 Mapfile

Mapfile is a structured text configuration file for data access and styling for MapServer. It defines the area of the map, tells the MapServer where the data is stored and where to output images. It also defines map layers, including projections and symbology. It must have a MAP extension otherwise MapServer will not recognize it. Mapfile is made up of different objects. Each object has a variety of parameters available for it. All mapfile parameters are documented in the mapfile reference [32]. A part of the mapfile source code displaying three layers including comments is included in Appendix 4.

10 Conclusions

A GIS-based technique has been applied for mapping spatial distribution at high spatial resolution of two climatological variables: annual mean temperature and annual precipitation for the period 1998–2009. The resulting maps suggest that the combined regression approach is a useful tool for interpolating from sparse point data. The combined approach involves regression using geographical and terrain variables as predictors followed by local interpolation of residuals. Altitude has proven to be overall the best predictor of both climatological variables. Altitude and latitude have been found to be the most powerful predictors of annual mean temperature. In case of annual precipitation, significant variables are ZxW25, altitude, longitude and latitude. Aspect was expected to be a significant predictor of MT, but this did not happen.

The final regression model of mean temperature enables us to describe 90–97% of spatial variability with the standard deviation RMSE = 0.4 °C. Final precipitation model shows a moderate degree of explained spatial variance 78–92% with RMSE = 106 mm. In comparison with the Climate Atlas of Czechia published by the CHMI, where the authors consider linear models dependent on elevation, this study provides a deeper analysis of the influence of topography on MT and P.

The focus of the radar study in chapter 5 was the use of radar data in precipitation modeling over the area of the Czech Republic. The first finding was that radar rainfall sums do not coincide nor significantly correlate (correlation coefficient $R^2 = 0.18$) with rain gauge observations due to high residual errors especially in mountainous regions as shown in figures 5.9–5.12. By the means of regression analysis residuals between the radar predicted rainfalls and rain gauge observations were calculated and then studied using residual regression analysis. A multivariate second order polynomial regression model was developed with three topographic and locational variables as the best predictors: altitude, distance from the radar antenna and latitude, which can explain up to 74% of variance of the residual errors. Such finding is important in regards to radar residual errors which are not random, but can be partially predicted which may allow for calibration and improvement in a radar's accuracy. The regression analysis was firstly conducted with an annual dataset containing 134 rain gauge stations and secondly for comparison with a monthly dataset including 21 rainfall recording gauge stations. The latter dataset is less representative but confirms the results of annual datasets and allows us to look at seasonal

variation. Nevertheless, as shown in fig. 5.22, there is no significant seasonal variation in radar-rainfall prediction.

Animated maps of deviations of MT from normals show that the last 12 years have been significantly warmer than the 30-year normal, approximately by 1 °C, which suggests that the hypothesis of global warming is true.

The key success of this thesis is in obtaining good results with only a subset of meteorological stations by using the topography-climate relationship. The resulting climatological maps available at <http://maps.fsv.cvut.cz/ka-map/> or <http://maps.fsv.cvut.cz/~muller/> have potential applications in many disciplines related to Earth Sciences.

10.1 Future Work

The largest limitation of this work is having only a subset of 22 meteorological stations available. If data from all stations had been provided, such as dataset B in year 2008, final maps would have had better accuracy and would have underlined local anomalies such as heat islands, as cities reflect and emit more heat than surrounding natural areas.

Annual means were studied, although for example mean maximum and mean minimum temperature is important in bioclimatological and agricultural research. It would certainly be interesting to look at the amplitudes of both climatological variables.

Details of how to use radar data (particularly the pattern of annual radar sums) in precipitation modeling remain unexplored, due to the high errors between radar and gauge measurements, which would need to be removed through some calibration.

Bibliography

- [1] AGNEW, M.D., PALUTIKOF, J.P. GIS-based Construction of Baseline Climatologies for the Mediterranean Using Terrain Variables. *Climate Research*, 2000, vol. 14, no. 2, p. 115-127. ISSN: 0936-577X.
- [2] ANSELIN, L. *Exploring Spatial Data with GeoDaTM : A Workbook*. 2005, University of Illinois: Urbana, IL, USA. p. 226.
- [3] BASIST, A., et al. Statistical Relationships between Topography and Precipitation Patterns. *Journal of Climate*, 1994, vol. 7, no., p. 1305-1315.
- [4] BOLDIŠ, P. *Bibliografické citace dokumentů podle ČSN ISO 690 a ČSN ISO 690-2: Část 2 - Modely a příklady citací u jednotlivých typů dokumentů*. v. 3.0. 2004. Available from: <<http://www.boldis.cz/citace/citace2.ps>>.
- [5] BORROUGH, P.A., MCDONELL, R.A., *Principles of Geographical Information Systems*: Oxford University Press. 1998.
- [6] BROWN, D.P., COMRIE, A.C. Spatial modeling of winter temperature and precipitation in Arizona and New Mexico, USA. *Climate Research*, 2002, vol. 22, no. 2, p. 115-128. ISSN: 0936-577X.
- [7] ČEPICKÝ, J. *Mapový server snadno a rychle* [online]. 3.11.2005. [cited 1.3.2010] <<http://www.root.cz/clanky/mapovy-server-snadno-a-rychle-1/>>.
- [8] CHAPMAN, L., THORNES, J.E. The use of geographical information systems in climatology and meteorology. *Progress in Physical Geography*, 2003, vol. 27, no. 3, p. 313-330. ISSN: 0309-1333.
- [9] CHMI. *CHMI Radar Network* [online]. 2002. [cited 24.2. 2010]. Available from: <<http://www.chmi.cz/meteo/rad/eindex.html>>.
- [10] CHMI. *GIS Methods and Data Processing. Climate Atlas of Czechia* [online]. 7.11.2005. [cited 25.11. 2009]. Available from: <<http://www.chmi.cz/meteo/ok/atlas/en/menu.html>>.
- [11] CLAPS, P., et al. Spatial distribution of the average air temperatures in Italy: Quantitative analysis. *Journal of Hydrologic Engineering*, 2008, vol. 13, no. 4, p. 242-249. ISSN: 1084-0699.
- [12] COLLINS, F.C.B., P.V. *A Comparison of Spatial Interpolation Techniques in Temperature Estimation*. 2008: Santa Fe.
- [13] CRISTOBAL, J., et al. Modeling air temperature through a combination of remote sensing and GIS data. *Journal of Geophysical Research-Atmospheres*, 2008, vol. 113, no. D13. ISSN: 0148-0227.

- [14] DALY, C. *Overview of the PRISM Model* [online]. 7.11.1996. [cited 12.01.2010]. Available from: <<http://www.prism.oregonstate.edu/docs/overview.html>>.
- [15] DALY, C., et al. A knowledge-based approach to the statistical mapping of climate. *Climate Research*, 2002, vol. 22, no. 2, p. 99-113. ISSN: 0936-577X.
- [16] DOBESH, H., et al., *Spatial interpolation for climate data: the use of GIS in climatology and meteorology*. London: Antony Rowe Ltd, Chippenham, Wiltshire. 2007.
- [17] ESRI. *ArcGIS 9.2 Desktop Help* [online]. 2008. [cited 24.2.2010]. Available from: <<http://webhelp.esri.com/arcgisdesktop/9.2>>.
- [18] ESRI. *ArcGIS Desktop Help* [online]. 2008. [cited 2.2.2010] <<http://webhelp.esri.com/arcgisdesktop/>>.
- [19] FLORIO, E.N., et al. Integrating AVHRR satellite data and NOAA ground observations to predict surface air temperature: a statistical approach. *International Journal of Remote Sensing*, 2004, vol. 25, no. 15, p. 2979-2994. ISSN: 0143-1161.
- [20] GOODALE, C.L., et al. Mapping monthly precipitation, temperature, and solar radiation for Ireland with polynomial regression and a digital elevation model. *Climate Research*, 1998, vol. 10, no. 1, p. 35-49. ISSN: 0936-577X.
- [21] GOROKHOVICH, Y., VILLARINI, G. Application of GIS for processing and establishing the correlation between weather radar reflectivity and precipitation data. *Meteorological Applications*, 2005, vol. 12, no. 1, p. 91-99. ISSN: 1350-4827.
- [22] GULER, M., et al. Assessment of some spatial climatic layers through GIS and statistical analysis techniques in Samsun Turkey. *Meteorological Applications*, 2007, vol. 14, no. 2, p. 163-169. ISSN: 1350-4827.
- [23] GYALISTRAS, D. Development and validation of a high-resolution monthly gridded temperature and precipitation data set for Switzerland (1951-2000). *Climate Research*, 2003, vol. 25, no. 1, p. 55-83. ISSN: 0936-577X.
- [24] HAMPACHER, M., RADOUCH, V., *Teorie chyb a vyrovnávací počet 10, 20: Příklady a návody na cvičení*. Praha: Vydavatelství ČVUT. 2000.
- [25] HIJMANS, R.J., AL., E. Very High Resolution Interpolated Climate Surfaces for Global Land Areas. *International Journal of Climatology*, 2005, vol. 25.
- [26] HUTCHINSON, M.F. Interpolating Mean Rainfall Using Thin-plate Smoothing Splines. *International Journal of Geographical Information Systems*, 1995, vol. 9, no. 4, p. 385-403. ISSN: 0269-3798.
- [27] KRAJEWSKI, W.F. *Rainfall Estimation Using Weather Radar and Ground Stations*, in *III International Symposium on Weather Radars*. 1995: São Paulo, Brazil.

- [28] KVÉTOŇ, V., ŽÁK, M. New climate Atlas of Czechia. *Studia Geophysica Et Geodaetica*, 2007, vol. 51, no. 2, p. 345-349. ISSN: 0039-3169.
- [29] LAPEN, D.R., HAYHOE, H.N. Spatial analysis of seasonal and annual temperature and precipitation normals in southern Ontario, Canada. *Journal of Great Lakes Research*, 2003, vol. 29, no. 4, p. 529-544. ISSN: 0380-1330.
- [30] LUDFORD, P.J. *Multiple Regression* [online]. 2003 [cited 10.01.2009 2009]. Available from: <http://www-users.cs.umn.edu/~ludford/Stat_Guide/Multiple_Regression.htm>.
- [31] MARQUINEZ, J., et al. Estimation models for precipitation in mountainous regions: the use of GIS and multivariate analysis. *Journal of Hydrology*, 2003, vol. 270, no. 1-2, p. 1-11. ISSN: 0022-1694.
- [32] MCKENNA, J., et al. *MapServer 5.6.3 Documentation* [online]. 2010. [cited 1.3.2010] <<http://www.mapserver.org/documentation.html>>.
- [33] MITCHEL, T. *Build AJAX-Based Web Maps Using ka-Map* [online]. 10.8.2005. [cited 1.3.2010] <<http://www.xml.com/pub/a/2005/08/10/ka-map.html?page=1>>.
- [34] MORAN, J.M. *Climate World Book 2005* [online]. 2005 [cited 13.11.2009] <<http://www.irthebest.com/Climate.html>>.
- [35] MÜLLER, A. *Climate Characteristics of the Czech Republic using ArcGIS*, in Faculty of Civil Engineering, Department of Mapping and Cartography. 2008, The Czech Technical University: Prague. p. 42.
- [36] NINYEROLA, M., et al. A methodological approach of climatological modelling of air temperature and precipitation through GIS techniques. *International Journal of Climatology*, 2000, vol. 20, no. 14, p. 1823-1841. ISSN: 0899-8418.
- [37] NINYEROLA, M., et al. Monthly precipitation mapping of the Iberian Peninsula using spatial interpolation tools implemented in a Geographic Information System. *Theoretical and Applied Climatology*, 2007, vol. 89, no. 3-4, p. 195-209. ISSN: 0177-798X.
- [38] NINYEROLA, M., et al. Objective air temperature mapping for the Iberian Peninsula using spatial interpolation and GIS. *International Journal of Climatology*, 2007, vol. 27, no. 9, p. 1231-1242. ISSN: 0899-8418.
- [39] NOVÁK, P. *Definice formátu radarových dat ČHMÚ-OR*. CHMI. 2006.
- [40] OMINIVERDI. *ka-Map! Documentation Wiki* [online]. 2007. [cited 1.3.2010] <<http://ka-map.ominiverdi.org/wiki/>>.
- [41] RODRÍGUEZ-LADOI, L.S., et al. Modelling air temperature for the state of São Paulo, Brazil. *Scientia Agricola*, 2007, vol. vol.64, no. 5, p. 460-467. ISSN: 0103-9016.

- [42] SALEK, M., NOVAK, P. AND SEO, J. *Operational Application of Combined Radar and Raingauges precipitation estimation at CHMI*, in *ERAD*. 2004.
- [43] SOKOL, Z. The use of radar and gauge measurements to estimate areal precipitation for several Czech River basins. *Studia Geophysica Et Geodaetica*, 2003, vol. 47, no. 3, p. 587-604. ISSN: 0039-3169.
- [44] SOKOL, Z., BLIZNAK, V. Areal distribution and precipitation-altitude relationship of heavy short-term precipitation in the Czech Republic in the warm part of the year. *Atmospheric Research*, 2009, vol. 94, no. 4, p. 652-662. ISSN: 0169-8095.
- [45] SQUIRES, M., et al. *Combining Deterministic and Probabilistic Methods to Produce Gridded Climatologies*, in *Twenty-fourth Annual ESRI User Conference*. 2004, ESRI: San Diego, USA.
- [46] SUN, R., et al. A Multivariate Regression Model for Predicting Precipitation in the Daqing Mountains. *Mountain Research and Development*, 2008, vol. 28, no. 3-4, p. 318-325. ISSN: 0276-4741.
- [47] The Ministry of Interior of the Czech Republic. *Česká republika* [online]. [cited 20.11.2009] <http://portal.gov.cz/wps/portal/_s.155/713/_s.155/7205?docid=3>.
- [48] TOLASZ, R., et al., *Climate Atlas of Czechia*. 1st ed. Praha, Olomouc: CHMI, Palacky University 2007. ISBN: 9788086690261, 9788024416267.
- [49] USGS. *Shuttle Radar Topography Mission (SRTM) "Finished" FAQ* [online]. 2007. [cited 02.11.2009 2010]. Available from: <http://seamless.usgs.gov/faq/srtm_faq.php>.
- [50] VICENTE-SERRANO, S.M., et al. Comparative analysis of interpolation methods in the middle Ebro Valley (Spain): application to annual precipitation and temperature. *Climate Research*, 2003, vol. 24, no. 2, p. 161-180. ISSN: 0936-577X.

List of Abbreviations

ArcGIS	GIS software package from ESRI
CHMI	Czech Hydrometeorological Institute
CR	Czech Republic
DEM	Digital elevation model
ESRI	Environmental Systems Research Institute, Inc.
GEOTOPO30	Global DEM at 30 arc seconds (1-km) resolution
GIS	Geographic information system
IDW	Inverse distance weighting
MT	Mean Temperature
NASA	National Aeronautics and Space Administration
P	Precipitation
PRISM	Parameter-elevation Regressions on Independent Slopes Model
RMSE	Root mean square error
SRTM	Space Shuttle Radar Topography Mission
USGS	U.S. Geological Survey

List of Figures

Figure 3.1: Study area.....	8
Figure 3.2: Meteorological stations.....	10
Figure 3.3: Station distribution with altitude (22 stations).....	11
Figure 3.4 : Station distribution with altitude (132 stations).....	11
Figures 3.5a, b: Linear regression between MT in Nov and Dec and altitude.	12
Figure 3.6: SRTM DEM	13
Figure 3.7: Slope	15
Figure 3.8: Aspect	15
Figure 3.9: Curvature	17
Figure 3.10: Solar Radiation	17
Figure 3.11: Wedge angles.....	18
Figure 3.12: ZxW25	19
Figure 3.13: Y0	19
Figure 4.1: Two-stage methodology.....	24
Figure 5.1: Coverage of the CHMI weather radars.....	30
Set of figures 5.2 above: Radar sums by months for 2008.....	33
Figure 5.3 below: Zoom in to the area where radar beam casts a ‘shadow’	33
Figure 5.4: Distance	34
Figure 5.5: Chart of DIF	35
Figure 5.6: Direction	35
Figure 5.7: Directional Diference	36
Figure 5.8: Directional Difference smoothed to 10km.....	36
Figure 5.9: Gauge precipitation measurements of 134 stations.....	38
Figure 5.10: Gauge precipitation measurements of 122 stations.....	38
Figure 5.11: Gauge precipitation measurements of 21 stations.....	39
Figures 5.12a and b: Altitude vs. annual and monthly residuals	41
Figures 5.13a and b: Mean Temperature vs. annual and monthly residuals.....	41
Figures 5.14a and b: Sunshine Duration and solar radiation vs. annual residuals.....	42
Figures 5.15a and b: Slope vs. annual and monthly residuals	42
Figures 5.16a and b: Distance from the radar antenna vs. annual and monthly residuals ..	43
Figures 5.17a and b:Latitude vs. annual and monthly residuals	43
Figures 5.18a and b: Directional Difference vs. annual and monthly residuals	44

Figures 5.19a and b: Longitude vs. annual and monthly residuals	44
Figures 5.20a and b: Curvature vs. annual and monthly residuals	45
Figures 5.21a and b: Aspect vs. annual and monthly residuals	45
Figures 5.2a and b: Residuals vs. months and seasons	45
Set of figures 6.1: Independent variables vs. MT (dataset A).....	50
Set of figures 6.2: Independent variables vs. P (dataset A)	52
Figure 6.3: Semivariogram of residuals of P in 2008.....	55
Figure 6.4: Corrector Map of MT in 2008 (dataset A).....	56
Figure 6.5: Corrector Map of P in 2008 (dataset A).....	56
Figure 7.1: Deviation of MT.....	62
Figure 7.2: Deviation of P	62
Figure 7.3: 40-year mean of MT	63
Figure 7.4: 40-year mean of MT (Source: Climate Atlas of Czechia)	63
Figure 7.5: 40-year mean of P.....	64
Figure 7.6: 40-year mean of P (Source: Climate Atlas of Czechia)	64
Figure 8.1: One example of final MT map	65
Figure 8.2: MT2008.	65
Figure 8.3: One example of final P map.....	66
Figure 8.4: P2008	66
Figure 8.5: 12-year mean of deviation from normal of MT	68
Figure 8.5: 12-year mean of deviation from normal of P.....	68
Figure 9.1: Screenshot of the MapServer with ka-Map	72

List of Tables

Table 3.2: Characteristics of 22 meteorological stations.....	11
Table 3.3: Characteristics of 132 meteorological stations.....	11
Table 3.4: Nr. of stations in altitude categories.....	11
Table 3.5: Nr. of stations in altitude categories.....	11
Table 5.1: Legend for 4 seasons.....	45
Table 5.2: Regression coefficients between annual residuals and various factors.....	46
Table 6.1: Correlations of independent variables with MT (dataset A)	49
Table 6.2: Correlations of independent variables with P (dataset A)	51
Table 6.3: Backward elimination in MT prediction (dataset A)	52
Table 6.4: Backward elimination in P prediction (dataset A)	53
Table 7.1: Statistical criteria	58
Table 7.2: DEM's accuracy.....	59
Table 7.3: Validation of MT and P using dataset B	60
Table 7.4: R^2 , RMSE	61
Table A.1: Dataset A, MT in year 2008.....	IX
Table A.1: Dataset A, P in year 2008.....	X
Table A-2.1: Backward Stepwise Approach	XII
Table A-2.2: 2 nd Order Polynomial Multivariate Regression	XII
Table A-3.1: Vertical Accuracy of the DEM.....	XV
Table A-3.2: Validation of MT and P in 2008.....	XVII

List of Appendices

Appendix 1: Missing Data.....	IX
Appendix 2: Backward Stepwise Approach.....	XII
Appendix 3: Accuracy.....	XV
Appendix 4: Mapfile.....	XIX

Appendix 1: Missing Data

Table A.1: Dataset A, MT in year 2008 and missing values in red

NR	MSTATION	ALT	NOV	DEC	Y08 MT
1	Brno Turany	241.0	6.5	2.1	10.7
2	Ceske Budejovice	394.1	4.9	1.5	9.8
3	Doksany	158.0	4.9	2.0	9.9
4	Holesov	223.6	6.8	2.2	10.3
5	Hradec Kralove	278.0	5.9	2.0	10.3
6	Cheb	483.0	3.7	0.1	8.5
7	Churanov	1117.8	2.0	-2.4	5.5
8	Klatovy	425.0	4.6	1.0	9.4
9	Kucharovice	334.0	5.5	1.5	10.4
10	Liberec	397.7	4.3	0.9	8.7
11	Lysa hora	1321.8	0.5	-3.5	3.9
12	Milesovka	833.0	2.3	-2.3	6.7
13	Mosnov	250.4	6.2	1.7	9.9
14	Olomouc	210.0	6.4	2.2	10.5
15	Praha Karlov	232.0	5.9	2.7	11.1
16	Praha Ruzyne	364.0	4.6	1.0	9.4
17	Pribyslav	530.0	4.3	0.1	8.3
18	Semcice	234.0	5.1	1.9	10.0
19	Svratouch	737.0	3.2	-1.3	7.2
20	Tabor	459.0	4.1	0.5	8.9
21	Velke Mezirici	452.0	4.9	0.5	8.8
22	Velke Pavlovice	196.0	5.89	2.13	10.81

Regression equations from MS Excel (fig. 3.5):

$$MT_{NOV08} = -0.004 * ALT + 6.844, \quad R^2 = 0.869$$

$$MT_{DEC08} = -0.005 * ALT + 3.198, \quad R^2 = 0.945$$

Table A.1: Dataset A, P in year 2008 and missing values in red

NR	MSTATION	ALT	CURV	JAN	FEB	MAR	JUN	JUL	AUG	NOV	DEC	Y08P
1	Brno Turany	241.0	0.05	17	10	32	35.9	62	44.6	29.6	23.5	426
2	Ceske Budejovice	394.1	-0.03	18.9	10	32.4	78.4	66.2	60	45	24.7	569.3
3	Doksany	158.0	0.00	29	21	24.7	80.6	98	60.8	15.7	38.7	560.8
4	Holesov	223.6	0.23	35.4	18	45	25.4	108	47.2	27.9	39.1	534.5
5	Hradec Kralove	278.0	-0.18	26.3	25	47	45.9	64.6	46.5	46.8	20.3	465.7
6	Cheb	483.0	-0.05	28.9	41	69.2	69.5	98.6	60	25.1	36.4	731.4
7	Churanov	1117.8	0.03	41.9	78	140	76.9	127	110	65.3	54.4	1011
8	Klatovy	425.0	-0.13	21.5	29	47.2	33.3	55.6	54.9	29.6	35.2	478.8
9	Kucharovice	334.0	0.03	15.4	7.4	28.7	84.1	70.3	38.6	30.3	21.5	445.4
10	Liberec	397.7	0.08	80.1	63	74	51.5	116	84.2	68.6	74.1	841.2
11	Lysa hora	1321.8	0.73	107	61	77.8	142	245	144	74.9	102	1269
12	Milesovka	833.0	1.54	30.4	20	36.4	44.9	75.9	69.4	15.1	50.2	560.8
13	Mosnov	250.4	-0.05	28.9	12	30.3	77.1	159	103	13.8	43.1	686.3
14	Olomouc	210.0	0.00	25.9	11	38.5	47.8	75.7	86.1	22.9	26.3	484.8
15	Praha Karlov	232.0	0.08	20.4	8	14.3	61.8	63.7	50.3	17.6	28.5	408.1
16	Praha Ruzyne	364.0	0.00	22.1	13	20	66	73.7	68.7	23.7	29.1	492.1
17	Pribyslav	530.0	-0.03	33.7	23	53.6	56.9	74.4	73	66.8	33.5	563.2
18	Semcice	234.0	-0.05	38	32	49.5	49.5	72.2	49.6	38.8	39.5	540.1
19	Svratouch	737.0	0.34	31.4	29	85.7	83.4	95.1	67.5	34.6	31.2	690.4
20	Tabor	459.0	0.10	32.2	19	60.2	47.4	52.5	56.6	44.7	28.3	442.2
21	Velke Mezirici	452.0	0.05	36.3	12	50.2	29.8	92.2	31.8	64.1	25.6	484.5
22	Velke Pavlovice	196.0	0.44	17.3	12	34.7	44.9	50.4	34.9	14.5	27.0	378.1

REGRESSION

Dependent Variable : **NOV08** Number of Observations: 20
Mean dependent var : 37.705 Number of Variables : 6
S.D. dependent var : 19.4719 Degrees of Freedom : 14

R-squared : 0.801272 F-statistic : 11.2896
Adjusted R-squared : 0.730297 Prob(F-statistic) : 0.000163695
Sum squared residual: 1506.97 Log likelihood : -71.6
Sigma-square : 107.641

Variable	Coefficient	Std.Error	t-Statistic	Probability
CONSTANT	24.74683	6.879042	3.597425	0.0029133
JAN	0.6771153	0.1561442	4.336473	0.0006836
JUN	-0.1873656	0.1254143	-1.493973	0.1573764
AUG	-0.289728	0.1524117	-1.900956	0.0780942
ALT	0.05403731	0.01276454	4.233392	0.0008346
CURV	-31.60545	7.778165	-4.063355	0.0011625

$P_{NOV08} = 24.75 + 0.6771 * JAN - 0.1874 * JUN - 0.2897 * AUG + 0.05404 * ALT - 31.61 * CURV$

REGRESSION

Dependent Variable : **DEC08** Number of Observations: 20
Mean dependent var : 39.22 Number of Variables : 5
S.D. dependent var : 18.7669 Degrees of Freedom : 15

R-squared : 0.961063 F-statistic : 92.5582
Adjusted R-squared : 0.950679 Prob(F-statistic) : 2.19784e-010
Sum squared residual: 274.273 Log likelihood : -54.5627
Sigma-square : 18.2849

Variable	Coefficient	Std.Error	t-Statistic	Probability
CONSTANT	13.71841	2.809142	4.883486	0.0001986
FEB	0.9121725	0.1077099	8.468792	0.0000004
MAR	-0.3633209	0.0680562	-5.338543	0.0000828
JUL	0.1972553	0.02817413	7.001291	0.0000043
CURV	12.40645	2.701437	4.592538	0.0003522

$P_{DEC08} = 13.72 + 0.9122 * FEB - 0.3633 * MAR + 0.1973 * JUL + 12.41 * CURV$

REGRESSION

Dependent Variable : **NOV05** Number of Observations: 21
Mean dependent var : 31.8952 Number of Variables : 4
S.D. dependent var : 20.9225 Degrees of Freedom : 17

R-squared : 0.918745 F-statistic : 64.0723
Adjusted R-squared : 0.904405 Prob(F-statistic) : 1.79042e-009
Sum squared residual: 746.96 Log likelihood : -67.2984
Sigma-square : 43.9389

Variable	Coefficient	Std.Error	t-Statistic	Probability
CONSTANT	12.79761	2.842123	4.502833	0.0003138
JAN	-0.5820471	0.1308637	-4.447736	0.0003531
DEC	0.53418	0.06947958	7.688302	0.0000006
MAR	0.9119744	0.1813468	5.028897	0.0001032

$P_{NOV05} = 12.80 - 0.5820 * JAN + 0.5342 * DEC + 0.9120 * MAR$

Appendix 2: Backward Stepwise Approach

Radar Study (Chapter 5)

Table A-2.1: Backward Stepwise Approach in Multivariate Linear Regression

R²	adj. R²	ALTITUDE		N_LAT		DIST		SOLARRAD		DIRDIFF10		MT2008		SLOPE	
		T	prob.	t	prob.	T	prob.	t	prob.	t	prob.	T	prob.	t	prob.
0.69	0.67	6.99	0.000	4.42	0.000	2.71	0.008	-1.70	0.092	1.77	0.079	-1.43	0.155	0.40	0.687
0.68	0.67	7.35	0.000	4.51	0.000	2.77	0.006	-1.66	0.099	1.76	0.081	-1.42	0.159		
0.68	0.67	10.3	0.000	4.66	0.000	3.00	0.003	-1.81	0.073	1.79	0.076				
0.67	0.66	10.52	0.000	4.45	0.000	3.57	0.000	-1.99	0.048						
0.66	0.65	13.55	0.000	4.18	0.000	3.64	0.000								
0.63	0.62	13.52	0.000	6.30	0.000										

Table A-2.2: 2nd Order Polynomial Multivariate Regression

R²	adj. R²	ALTITUDE		DIST		DIST2		N_LAT	
		T	prob.	t	prob.	t	prob.	t	prob.
0.74	0.73	13.88	0.000	-4.52	0.000	5.99	0.000	2.57	0.011
0.72	0.72	13.38	0.000	-4.94	0.000	7.02	0.000		

Mean Temperatures (Chapter 6, section 6.2)

Below are tables showing the backward elimination of independent variables for dataset A.

* denotes the best fitting (but still reasonable) model, final model selected is in bold.

MT09

R²	adj. R²	ALTITUDE		Y		ASP		SLP		CURV		SOLRAD	
		t	prob.	t	prob.	t	prob.	t	prob.	t	prob.	T	prob.
0.963	0.948	-10.24	0.000	-2.29	0.037	1.57	0.137	1.32	0.206	-0.89	0.386	-0.15	0.881
0.961	0.952	-17.91	0.000	-2.62	0.017	1.52	0.146	1.03	0.320				
*0.959	0.952	-19.88	0.000	-2.41	0.027	1.56	0.014						
0.953	0.948	-19.60	0.000	-2.67	0.015								
0.935	0.932	-17.03	0.000										

$$y = 11.656 - 0.00563341 * [\text{DEM-srtm}] - 0.000003786632 * [Y0]$$

MT08

R²	adj. R²	ALTITUDE		Y		ASP		SLP		CURV		SOLRAD	
		t	prob.	t	prob.	t	prob.	t	prob.	t	prob.	T	prob.
0.978	0.969	-13.79	0.000	-3.06	0.008	1.65	0.120	1.79	0.094	-1.37	0.191	0.11	0.910
0.975	0.969	-23.42	0.000	-3.11	0.006	1.53	0.145	1.34	0.199				
*0.973	0.968	-25.07	0.000	-2.76	0.013	1.64	0.118						
0.969	0.966	-24.26	0.000	-3.37	0.003								
0.951	0.948	-19.63	0.000										

$$y = 12.197 - 0.00588921 * [\text{DEM-srtm}] - 0.000003992299 * [Y0]$$

MT07

R ²	adj. R ²	ALTITUDE		Y		SLP		CURV		SOLRAD		ASP	
		t	prob.	t	prob.	t	prob.	t	prob.	t	prob.	T	prob.
0.976	0.966	-13.59	0.000	-2.58	0.021	1.72	0.106	-1.34	0.201	0.73	0.474	0.65	0.524
0.975	0.967	-14.00	0.000	-3.08	0.007	1.69	0.111	-1.28	0.218	0.91	0.374		
*0.974	0.968	-21.73	0.000	-2.99	0.008	1.97	0.065	-1.22	0.241				
0.972	0.967	-22.54	0.000	-2.87	0.010	1.67	0.112						
0.967	0.964	-23.65	0.000	-2.35	0.030								
0.958	0.956	-21.27	0.000										

$$y = 12.211 - 0.00584728 * [\text{dem-srtm}] - 0.0000028924 * [Y0]$$

MT06

R ²	adj. R ²	ALTITUDE		ASP		Y		SLP		CURV		SOLRAD	
		t	prob.	t	prob.	t	prob.	t	prob.	t	prob.	T	prob.
0.917	0.884	-6.74	0.000	1.51	0.151	0.20	0.844	0.97	0.349	-0.88	0.395	-0.037	0.971
0.912	0.897	-13.54	0.000	1.57	0.134	0.42	0.677						
*0.911	0.902	-13.84	0.000	1.54	0.139								
0.900	0.895	-13.40	0.000										
0.900	0.889	-13.03	0.000			-0.08	0.935						

$$y = 10.671 - 0.00490266 * [\text{dem-srtm}] - 0.000000202995 * [Y0]$$

MT05

R ²	adj. R ²	ALTITUDE		ASP		Y		SLP		CURV		SOLRAD	
		t	prob.	t	prob.	t	prob.	t	prob.	t	prob.	T	prob.
0.938	0.914	-6.14	0.000	1.32	0.206	0.13	0.902	0.86	0.401	-0.67	0.510	-0.503	0.622
0.934	0.924	-15.85	0.000	1.25	0.228	0.13	0.900						
*0.934	0.928	-16.28	0.000	1.31	0.206								
0.929	0.925	-16.13	0.000										
0.929	0.921	-15.75	0.000			-0.32	0.756						

$$y = 10.447 - 0.00554196 * [\text{dem-srtm}] - 0.0000005441739 * [Y0]$$

MT04

R ²	adj. R ²	ALTITUDE		ASP		N_LAT		CURV		SLP		SOLRAD	
		t	prob.	t	prob.	t	prob.	t	prob.	t	prob.	T	prob.
0.966	0.952	-10.78	0.000	2.07	0.056	-0.55	0.593	-1.09	0.293	0.97	0.348	0.08	0.931
0.963	0.957	-21.43	0.000	2.14	0.046	-0.49	0.627						
*0.962	0.958	-21.87	0.000	2.48	0.023								
0.950	0.948	-19.51	0.000										
0.954	0.949	-19.75	0.000			-1.19	0.248						

$$y = 10.923 - 0.00583272 * [\text{dem-srtm}] - 0.000001723783 * [Y0]$$

MT03

R ²	adj. R ²	ALTITUDE		Y0		ASP		SLP		CURV		SOLRAD	
		t	prob.	t	prob.	t	prob.	t	prob.	t	prob.	T	prob.
0.957	0.940	-0.90	0.000	-1.88	0.079	1.46	0.165	1.06	0.306	-0.81	0.427	0.25	0.806
*0.953	0.945	-18.88	0.000	-1.74	0.098	1.61	0.125						
0.946	0.940	-18.23	0.000	-2.29	0.033								
0.931	0.928	-16.44	0.000										

$$y = 11.372 - 0.00538561 * [\text{dem-srtm}] - 0.000003392663 * [Y0]$$

MT02

R ²	adj. R ²	ALTITUDE		Y0		ASP		SLP		CURV		SOLRAD	
		t	prob.	t	prob.	t	prob.	t	prob.	t	prob.	T	prob.
0.964	0.949	-9.88	0.000	-1.62	0.126	1.58	0.136	1.22	0.240	-1.26	0.228	-0.52	0.609
*0.959	0.952	-20.29	0.000	-1.19	0.074	1.38	0.186						
0.955	0.950	-20.02	0.000	-2.41	0.026								
0.941	0.938	-17.86	0.000										

$$y = 11.528 - 0.0052164 * [\text{DEM-srtm}] - 0.000003110357 * [\text{Y0}]$$

MT01

R ²	adj. R ²	ALTITUDE		Y0		SLP		ASP		CURV		SOLRAD	
		t	prob.	t	prob.	t	prob.	t	prob.	t	prob.	T	prob.
0.968	0.956	-10.91	0.000	-1.49	0.158	1.77	0.097	1.67	0.116	-1.59	0.133	-0.045	0.656
0.968	0.958	-18.67	0.000	-1.68	0.113	1.76	0.098	1.65	0.118	-1.65	0.118		
0.958	0.952	-18.31	0.000	-2.11	0.049	0.80	0.434						
*0.957	0.953	-20.59	0.000	-1.97	0.063								
0.948	0.946	-19.17	0.000										

$$y = 10.618 - 0.00538076 * [\text{DEM-srtm}] - 0.000002560351 * [\text{Y0}]$$

MT00

R ²	adj. R ²	ALTITUDE		ASP		Y0		SLP		CURV		SOLRAD	
		t	prob.	t	prob.	t	prob.	t	prob.	t	prob.	T	prob.
0.973	0.962	-12.10	0.000	1.47	0.160	-1.21	0.245	1.25	0.231	-1.07	0.301	-0.076	0.941
*0.970	0.965	-23.79	0.000	1.48	0.156	-1.12	0.278						
0.968	0.964	-23.64	0.000	1.93	0.068								
0.961	0.960	-22.35	0.000										
0.966	0.963	-23.37	0.000			-1.66	0.113						

$$y = 11.968 - 0.00556648 * [\text{DEM-srtm}] - 0.000001940305 * [\text{Y0}]$$

MT99

R ²	adj. R ²	ALTITUDE		ASP		Y0		SLP		SOLRAD		CURV	
		t	prob.	t	prob.	t	prob.	t	prob.	t	prob.	T	prob.
0.969	0.956	-10.90	0.000	1.82	0.089	-1.05	0.311	1.13	0.270	-0.70	0.492	-0.55	0.590
*0.965	0.960	-22.20	0.000	1.81	0.088	-1.04	0.313						
0.963	0.959	-22.16	0.000	2.26	0.036								
0.954	0.951	-20.26	0.000										
0.959	0.955	-21.17	0.000			-1.64	0.116						

$$y = 11.230 - 0.00559034 * [\text{DEM-srtm}] - 0.000002129976 * [\text{Y0}]$$

MT98

R ²	adj. R ²	ALTITUDE		Y0		ASP		SLP		CURV		SOLRAD	
		t	prob.	t	prob.	t	prob.	t	prob.	t	prob.	T	prob.
0.961	0.946	-9.34	0.000	-1.32	0.208	1.56	0.140	1.35	0.195	-1.35	0.197	-0.72	0.480
*0.955	0.947	-19.28	0.000	-1.576	0.132	1.27	0.219						
0.951	0.946	-19.16	0.000	-2.07	0.053								
0.940	0.937	-17.67	0.000										

$$y = 11.071 - 0.00550756 * [\text{DEM-srtm}] - 0.000002955612 * [\text{Y0}]$$

Appendix 3: Accuracy

Table A-3.1: Vertical Accuracy of the DEM

d = ALT-ALTDDEM

NR	NAME	ALT	ALT DEM _{utm}	d	ALT _{DEM}	d	NR	NAME	ALT	ALT DEM _{utm}	d	ALT _{DEM}	d
1	Holesov	223.6	225	-1	226	-2	41	Pec pod Snezkou	816	834	-18	818	-2
2	Kromeriz	233	232	1	233	0	42	Upice	413	407	6	407	6
3	Protivanov	675	678	-3	676	-1	43	Velichovsky	299	291	8	293	6
4	Stitna nad Vlari – Popov	315	311	4	311	4	44	Vrchlabi	482	485	-3	484	-2
5	Stare Mesto	235	219	16	219	16	45	Destne v Orlic. Horach	653	657	-4	657	-4
6	Strani	383	381	2	383	0	46	Polom	748	726	22	724	24
7	Straznice	176	170	6	170	6	47	Rokytnice v Orlic.horach	564	577	-13	577	-
8	Vizovice	313	301	12	309	4	48	Rychnov nad Kneznou	335	327	8	323	12
9	Brod nad Dyji	175	175	0	174	1	49	Usti nad Orlicí	402	399	3	400	2
10	Brno	241	237	4	237	4	50	Zamberk	405	404	1	404	1
11	Bystrice nad Pernstejnem	553	549	4	549	4	52	Gajer	515	514	1	515	0
12	Brno	236	234	2	233	3	53	Hradec Kralove	278	255	23	254	24
13	Dukovany	400	397	3	396	4	54	Mokosin	255	255	0	258	-3
14	Dyjakovice	201	197	4	197	4	55	Pardubice	225	227	-2	227	-2
15	Kostelni Myslova	569	569	0	571	-2	56	Sez	529	520	9	523	6
16	Kostelni Myslova	569	569	0	571	-2	57	Svratouch	737	730	7	728	9
17	Kucharovice	334	337	-3	337	-3	58	Holovousy	321	299	22	299	22
18	Lednice	176	173	3	173	3	59	Jicin	283	282	1	282	1
19	Nedvezi	722	718	4	718	4	60	Podebrady	189	192	-3	189	0
20	Sedlec	474	472	2	472	2	62	Zelezna Ruda	867	869	-2	866	1
21	Vatin	555	554	1	553	2	63	Klatovy	425	422	3	414	11
22	Velke Mezirici	452	455	-3	455	-3	64	Plzen	360	355	5	358	2
23	Cerna v Posumavi	739	739	0	737	2	65	Plzen	328	335	-7	335	-7
24	Churanov	1117.8	1117	1	1117	1	66	Stankov	362	361	1	361	1
25	Churanov	1118	1111	7	1109	9	67	Konstantinovy Lazne	527	529	-2	529	-2
26	Husinec	492	492	0	490	2	68	Kralovice	468	467	1	468	0
27	Kocelovice	515	512	3	512	3	69	Krasne Udoli	642	649	-7	649	-7
28	Kocelovice	519	512	7	512	7	70	Marianske Lazne	691	690	1	687	4
29	Rozmital pod Tremsinem	524	524	0	524	0	71	Primda	742	742	0	743	-1
30	Temelin	503	501	2	499	4	72	Cheb	483	480	3	480	3
31	Vraz	433	434	-1	434	-1	73	Karlovy Vary	603	598	5	597	6
32	Borkovice	419	416	3	416	3	74	Sindelova	587	585	2	586	1
33	Bynov	475	475	0	476	-1	75	Belotin	306	304	2	304	2
34	Ceske Budejovice	394.06	393	1	393	1	76	Cervena	749	748	1	747	2
35	Cesky Krumlov	554	540	14	541	13	77	Javornik	289	284	5	285	4
36	Jindrichuv Hradec	524	524	0	524	0	78	Jesenik	465	465	0	466	-1
37	Nadejkov	616	613	3	612	4	79	Karlova Studanka	780	788	-8	798	-
38	Tabor	459	462	-3	462	-3	80	Lucina	300	295	5	295	5
39	Vyssi Brod	559	557	2	557	2	81	Lysa hora	1321.8	1314	8	1314	8
40	Labska bouda	1315	1319	-4	1320	-5	82	Mosnov	250.4	245	5	245	5

NR	NAME	ALT	ALT _{DEM_utm}		ALT _{DEM}		NR	NAME	ALT	ALT _{DEM_utm}		ALT _{DEM}	
			d	d	d	d				d	d		
83	Opava	270	258	12	259	11	133	Liberec	397.7	400	-2	399	-1
84	Ostrava	238.6	236	3	235	4	134	Straz pod Ralskem	310	308	2	308	2
85	Serak	1328	1328	0	1327	1	135	Varnsdorf	365	372	-7	374	-9
86	Svetla Hora	593	591	2	589	4							
87	Mesto Albrechtice	483	502	- 19	500	- 17							
88	Dubicko	275	271	4	269	6							
89	Jevicko	342	343	-1	344	-2							
90	Luka	510	508	2	503	7							
91	Olomouc	210	207	3	207	3							
92	Paseka	290	288	2	288	2							
93	Sumperk	328	324	4	324	4							
94	Trebarov	375	386	- 11	387	- 12							
95	Horni Becva	565	558	7	556	9							
97	Maruska	664	639	25	645	19							
98	Prerov	202.7	201	2	201	2							
99	Valasske Mezirici	334	333	1	333	1							
100	Vsetin	387	380	7	379	8							
101	Lany	415	421	-6	421	-6							
102	Neumetely	322	318	4	318	4							
103	Praha	232	234	-2	218	14							
104	Praha	282	280	2	281	1							
105	Praha	191	204	- 13	204	- 13							
106	Praha	302.04	302	0	302	0							
107	Pribram	555	553	2	557	-2							
108	Praha	364	362	2	362	2							
109	Brandys n. Labem	179	179	0	179	0							
110	Desna	772	768	4	770	2							
111	Semcice	234	231	3	231	3							
112	Tuhan	160	161	-1	161	-1							
113	Kosetice	534	535	-1	534	0							
114	Nedrahovice	348	350	-2	350	-2							
115	Novy Rychnov	624	615	9	619	5							
116	Ondrejov	485	494	-9	493	-8							
117	Pribyslav	530	525	5	525	5							
118	Doksany	158	156	2	156	2							
119	Doksany	158	156	2	156	2							
120	Tusimice	322.4	323	-1	323	-1							
121	Kopisty	240	241	-1	241	-1							
122	Milesovka	833	816	17	792	41							
123	Nova Ves v Hor.	725	730	-5	730	-5							
124	Smolnice	345	343	2	342	3							
125	Strojetice	372	370	2	370	2							
126	Teplice	236	217	19	218	18							
127	Ustí nad Labem	375	364	11	364	11							
128	Zatec	210	214	-4	213	-3							
129	Bedrichov	777	771	6	771	6							
130	Ceska Lipa	246	253	-7	253	-7							
131	Doksy	284	287	-3	287	-3							
132	Hejnice	396	397	-1	396	0							

Table A-3.2: Validation of MT and P in 2008

NR	NAME	MT 2008	MT08	d	P 2008	P08	d	NR	NAME	MT 2008	MT08	d	P 2008	P08	d
2	Kromeriz	10.8	10.3	0.5	454	480	-27	47	Rokytnice v Orlic.horach	7.8	8.1	-0.3	798	832	-34
3	Protivanov	7.8	7.7	0.1	605	722	-117	48	Rychnov nad Kneznou	9.4	9.6	-0.2	578	658	-80
4	Stitna nad Vlari – Popov	9.2	10.0	-0.8	718	624	94	49	Usti nad Orlicí	9.0	9.2	-0.2	684	630	53
5	Stare Mesto	10.5	10.5	0.0	522	457	66	50	Zamberk	8.5	9.1	-0.6	643	684	-41
6	Strani	9.2	9.7	-0.5	753	658	95	52	Gajer	8.3	8.5	-0.2	630	610	20
7	Straznice	10.6	11.0	-0.4	546	545	1	54	Mokosin	10.4	10.1	0.3	490	390	99
8	Vizovice	9.7	9.9	-0.2	645	647	-2	55	Pardubice	10.1	10.3	-0.2	495	399	97
9	Brod nad Dyji	10.8	11.1	-0.3	365	383	-19	56	Sez	8.7	8.5	0.2	618	590	28
10	Brno	10.7	10.7	0.0	426	426	0	58	Holovousy	10.1	9.7	0.4	536	553	-17
11	Bystrice nad Pernstejnem	8.4	8.3	0.1	519	600	-81	59	Jicin	9.9	9.7	0.2	567	610	-43
13	Dukovany	9.6	9.6	0.0	380	469	-89	60	Podebrady	10.2	10.4	-0.2	615	413	201
14	Dyjakovice	10.7	11.1	-0.4	368	345	22	62	Zelezna Ruda	7.2	6.9	0.3	1068	830	238
15	Kostelni Myslova	8.4	8.5	-0.1	515	554	-39	64	Plzen	9.6	9.6	0.0	467	496	-29
16	Kostelni Myslova	8.4	8.5	-0.1	510	554	-44	65	Plzen	8.7	9.7	-1.0	453	509	-56
18	Lednice	11.0	11.0	0.0	429	332	97	66	Stankov	9.1	9.7	-0.6	531	441	90
19	Nedvezi	7.5	7.3	0.2	511	691	-180	67	Konstantinovy Lazne	7.8	8.5	-0.7	518	620	102
20	Sedlec	9.6	9.0	0.6	431	495	-64	68	Kralovice	8.9	8.9	0.0	423	567	144
21	Vatin	8.1	8.2	-0.1	724	568	155	69	Krasne Udoli	7.3	7.7	-0.4	567	709	142
23	Cerna v Posumavi	6.6	7.8	-1.2	685	746	-61	70	Marianske Lazne	6.7	7.4	-0.7	810	805	4
25	Churanov	5.5	5.5	0.0	1017	1015	2	71	Primda	7.2	7.3	-0.1	691	677	14
26	Husinec	8.3	9.2	-0.9	525	596	-71	73	Karlovy Vary	7.6	7.9	-0.3	475	755	280
27	Kocelovice	8.9	8.9	0.0	516	529	-13	74	Sindelova	6.9	7.9	-1.0	879	836	43
28	Kocelovice	8.9	8.9	0.0	518	529	-11	75	Belotin	9.7	9.7	0.0	618	674	-56
29	Rozmital pod Tremsinem	8.3	8.7	-0.4	584	569	15	76	Cervena	7.1	7.1	0.0	785	838	-53
30	Temelin	8.9	9.0	-0.1	502	525	-23	77	Javornik	10.3	9.6	0.7	686	558	128
31	Vraz	9.0	9.3	-0.3	503	482	21	78	Jesenik	8.6	8.6	0.0	834	793	41
32	Borkovice	8.9	9.4	-0.5	461	493	-32	79	Karlova Studanka	6.6	6.8	-0.2	1141	980	161
33	Bynov	8.7	9.3	-0.6	550	532	18	80	Lucina	9.7	9.7	0.0	642	733	-91
35	Cesky Krumlov	8.7	9.0	-0.3	523	666	-144	83	Opava	9.7	9.8	-0.1	549	615	-67
36	Jindrichuv Hradec	8.8	8.8	0.0	555	561	-6	84	Ostrava	10.1	9.9	0.2	677	624	52
37	Nadejkov	8.2	8.1	0.1	493	546	-53	85	Serak	3.5	3.6	-0.1	1149	1232	-83
39	Vyssi Brod	7.4	9.0	-1.6	779	669	110	86	Svetla Hora	7.2	7.9	-0.7	685	784	-99
40	Labska bouda	3.1	3.4	-0.3	1485	1407	78	87	Mesto Albrechtice	8.9	8.4	0.5	784	752	32
41	Pec pod Snezkou	5.8	6.3	-0.5	1281	1049	232	88	Dubicko	9.6	9.9	-0.3	489	566	-78
42	Upice	8.4	9.0	-0.6	630	678	-48	89	Jevicko	8.8	9.6	-0.8	586	540	46
43	Velichovky	9.5	9.9	-0.4	539	527	13	90	Luka	8.8	8.7	0.1	576	599	-23
44	Vrchlabi	8.5	8.4	0.1	846	902	-56	92	Paseka	10.0	9.9	0.1	495	606	111
45	Destne v Orlic. Horach	7.0	7.6	-0.6	943	888	54	93	Sumperk	8.9	9.6	-0.7	619	754	135
46	Polom	7.4	7.2	0.2	765	916	-152	94	Trebarov	9.2	9.3	-0.1	582	563	19

NR	NAME	MT 2008	MT08	d	P 2008	P08	d
97	Maruska	8.1	7.9	0.2	776	776	-1
98	Prerov	10.1	10.5	-0.4	422	510	-88
99	Valasske Mezirici	9.4	9.6	-0.2	698	789	-91
100	Vsetin	9.1	9.4	-0.3	715	765	-50
101	Lany	8.9	9.1	-0.2	500	547	-48
102	Neumetely	9.5	10.0	-0.5	485	494	-9
104	Praha	10.1	10.5	-0.4	500	421	78
105	Praha	11.9	11.2	0.7	425	395	29
106	Praha	10.4	10.5	-0.1	522	460	62
107	Príbram	8.8	8.6	0.2	475	562	-87
109	Brandys nad Labem	10.5	10.7	-0.2	580	387	193
110	Desna	5.9	6.6	-0.7	1155	1159	-5
112	Tuhan	10.1	10.6	-0.5	495	412	83
113	Kosetice	8.7	8.5	0.2	505	540	-35
114	Nedrahovice	8.6	9.7	-1.1	509	478	31
115	Novy Rychnov	7.8	8.1	-0.3	609	602	8
116	Ondrejov	8.7	8.9	-0.2	549	569	-20
119	Doksany	10.0	9.9	0.1	560	561	-1
120	Tusimice	9.5	9.5	0.0	420	446	-27
121	Kopisty	9.4	10.0	-0.6	505	482	23
123	Nova Ves v Horách	6.8	7.0	-0.2	713	774	-61
124	Smolnice	9.3	9.4	-0.1	331	493	-162
125	Strojetice	9.9	9.3	0.6	501	519	-19
126	Teplice	10.2	10.1	0.1	552	444	108
127	Ustí nad Labem	9.5	9.1	0.4	595	555	41
128	Zatec	9.2	10.2	-1.0	476	429	47
129	Bedrichov	6.1	6.5	-0.4	1078	1117	-39
130	Ceska Lipa	9.5	9.7	-0.2	628	564	64
131	Doksy	9.2	9.6	-0.4	570	560	10
132	Hejnice	9.3	8.7	0.6	956	873	82
134	Straz pod Ralskem	8.8	9.3	-0.5	729	732	-3
135	Varnsdorf	8.9	8.9	0.0	856	632	224

MBE = -22.2 °C MBE = -196 mm
 MAE = 0.3 °C MAE = 67 mm
 RMSE = 0.4 °C RMSE = 89 mm
 MAX = 1.6 °C MAX = 280 mm
 MED = 0.3 °C MED = 53 Mm

Appendix 4: Mapfile

```
MAP
  NAME MT          # each object should have a name
  PROJECTION
    'init=epsg:32633' #WGS84 UTM Zone 33N
  END
  SIZE 450 300    # default size of a map in pixels
  EXTENT 261955 5361927 805940 5665141
  UNITS meters    # map units
  SHAPEPATH "/data/amuller/MT/" # path to data
  IMAGECOLOR 255 255 255 # background color
  IMAGETYPE png   # type of the resulting image
  TRANSPARENT ON #allow transparency
  FONTSET '/data/amuller/fonts/font.list'

LEGEND
  KEYSIZE 12 15
  LABEL
    TYPE TRUETYPE
    FONT arial
    SIZE 10
    COLOR 0 0 0
    ALIGN RIGHT
  END
  STATUS ON
END

REFERENCE
  STATUS ON
  IMAGE '/data/amuller/MT/RefMap.jpg'
  SIZE 150 84
  EXTENT 261955.53631154 5361927.43289661 805939.957613716 5665140.53932282
  COLOR -1 -1 -1
  OUTLINECOLOR 255 0 0
END #REFERENCE

SYMBOL
  NAME "square"
  TYPE VECTOR
  FILLED true
  POINTS
    0 0
    0 1
    1 1
    1 0
    0 0
  END
END

### Beginning of Layers ###

LAYER
  GROUP 'Relief'
  NAME 'Shaded Relief'
  DATA '/data/amuller/base/shaded_relief.jpg'
  TYPE RASTER
  STATUS OFF
  PROJECTION
    'init=epsg:32633'
  END
END

LAYER
  GROUP 'MT'
  NAME mt2008
  DATA MT08_fm3.img
```

```

TYPE RASTER
STATUS ON
PROJECTION
  'init=epsg:32633'
END
METADATA
  'OPACITY' '85'
END
CLASSITEM "[pixel]" # class using an EXPRESSION using only [pixel].
CLASS
  EXPRESSION ([pixel] > 0 AND [pixel] <= 5)
  STYLE
    COLOR 0 38 115
  END
END
CLASS
  EXPRESSION ([pixel] > 5 AND [pixel] <= 6)
  STYLE
    COLOR 0 77 168
  END
END
#
#
#
# more classes

LAYER
  GROUP 'Cities'
  NAME '9 Largest Cities'
  DATA '/data/amuller/base/cities_large'
  TYPE POINT
  STATUS OFF
  PROJECTION
    'init=epsg:32633'
  END
  LABELITEM 'NAZEV_ENG'
  CLASS
    NAME '9 Largest Cities'
    STYLE
      SYMBOL 'square'
      COLOR 0 0 0
      SIZE 5
    END
    LABEL
      TYPE TRUETYPE
      SIZE 12
      FONT Arial
      COLOR 0 0 0
      POSITION UC
      OFFSET 0 0
      PRIORITY 10
      # MAXSCALEDENOM 1100000
    END #LABEL
  END #CLASS
END

END

### END of ALL Layers ###
END # End of mapfile

```