ANALYZING RADAR-MEASURED RAINFALL VS. RAIN GAUGES IN GIS

Arnošt, MÜLLER

Department of Mapping and Cartography, Faculty of Civil Engineering, Czech Technical University,
Thákurova 7, 166 29 Praha 6 – Dejvice, *arnost.muller@fsv.cvut.cz*

ABSTRACT

Rainfall data are traditionally collected at discrete point locations in space, at meteorological stations (rain gauges). Values at any other point must be interpolated or can be remotely sensed by ground-based radar, which can detect the areal distribution of precipitation at more detailed spatial scale. Nevertheless, radar measurements are affected by various types of errors and the transformation of measured radar reflectivity into rain rates is far from accurate. This study provides a deeper analysis of the influence of topography on radar measured precipitation.

By the means of linear regression analysis residuals between 134 rain gauges and corresponding radar estimated rainfalls were calculated, and then studied using residual regression analysis with the following independent variables: altitude, longitude, latitude, aspect, slope, curvature, distance from the radar antenna, aspect perpendicular to the radar beam referred to as directional difference, mean air temperature, and solar radiation. The independent variables were derived from the 90 m SRTM DEM in ArcGIS. A multivariate second order polynomial regression model was developed with three topographic and locational variables as the best predictors: altitude, distance, and latitude, which can explain up to 74% of variance of the residual errors. This means that radar measurement errors are not only a cause of random variation, but can be partially predicted, which may allow for some type of correction and improvement in radar's accuracy.

Keywords: rainfall, rain gauge, radar, GIS, regression, terrain analysis

INTRODUCTION

Precipitation, as one of the basic climatological factors, is used as an input in various models in hydrologic modeling, e.g. flood prediction, in agriculture applications for estimating yields, in land management, or in atmospheric simulation models. Rainfall data are traditionally collected at meteorological stations (rain gauges), which are discrete point locations in space. Values at any other point must be derived from neighboring meteorological stations or can be remotely sensed, e.g. by ground-based radar. The main advantage of rain gauges is a fine temporal resolution. In fact, gauges record continuously and are able to detect even short (minute) rainfalls [1]. Rain gauge observations are still considered as close to true rainfall as we can get at present state of art technologies [4].

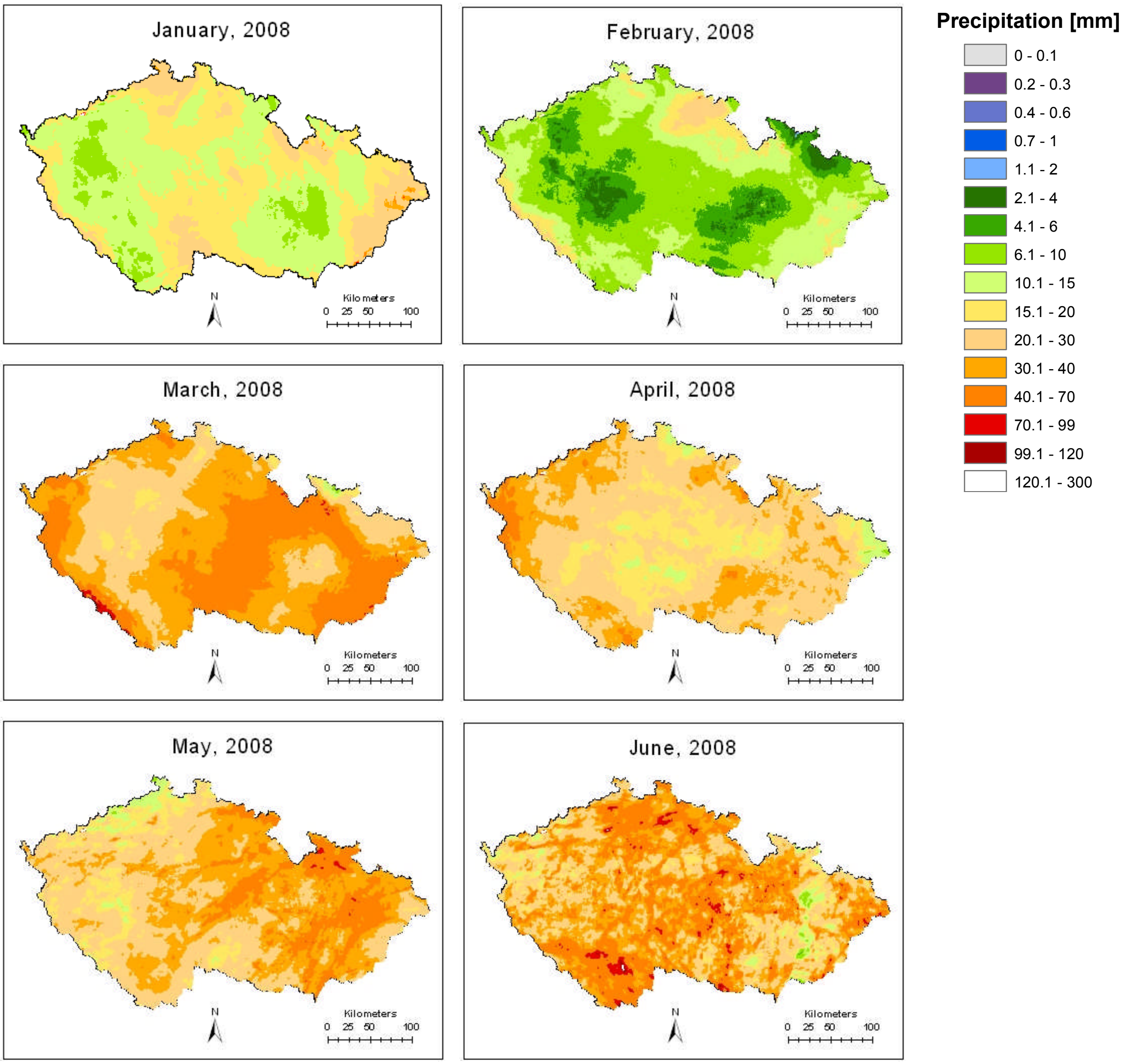
While rain gauges measure at discrete locations, weather radar samples at discrete time instances (e.g. every 10 minutes). Radar's main advantage is that it can 'see' much larger atmospheric space than rain gauges located on the ground. Radar can detect the areal distribution of precipitation at more detailed spatial scale than rain gauge network and therefore, the final rain field pattern should be determined by radar, as recommended by Krajewski [4]. However, 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 [4, 8].

The objective of this study is to test the influence of following topographic, locational and atmospheric variables on residual errors: altitude, longitude, latitude, aspect, slope, curvature, distance from the radar antenna (DIST), aspect perpendicular to the radar beam referred to as directional difference (DIF), mean air temperature and solar radiation. Residual errors are calculated as the difference between radar-estimated rainfall and rain gauges observations, which are considered as a good approximation of the true ground rainfall. If some factors influencing residual errors are found significant, this would mean radar errors are not random and therefore can be removed by some type of correction.

DATA AND METHODS

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, for analysis [3]. 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 provides insight into radar measurement errors. According to the data provided by the Czech Hydrometeorological Institute (CHMI), the analysis is performed with annual estimates of 134 meteorological stations in the year 2008.

Study area is the area of the whole Czech Republic. The Czech radar network (CZRAD), operated by the CHMI, consists of two polarization C-band Doppler radars [2]. 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 [8]. Monthly raw radar rainfall sums (already combined from both radars) for the year 2008 were obtained in two-dimensional binary format (.RPD), were imported into ArcGIS and visualized (fig. 1.) using the same color schema as CHMI uses for radar estimated rainfall representations.



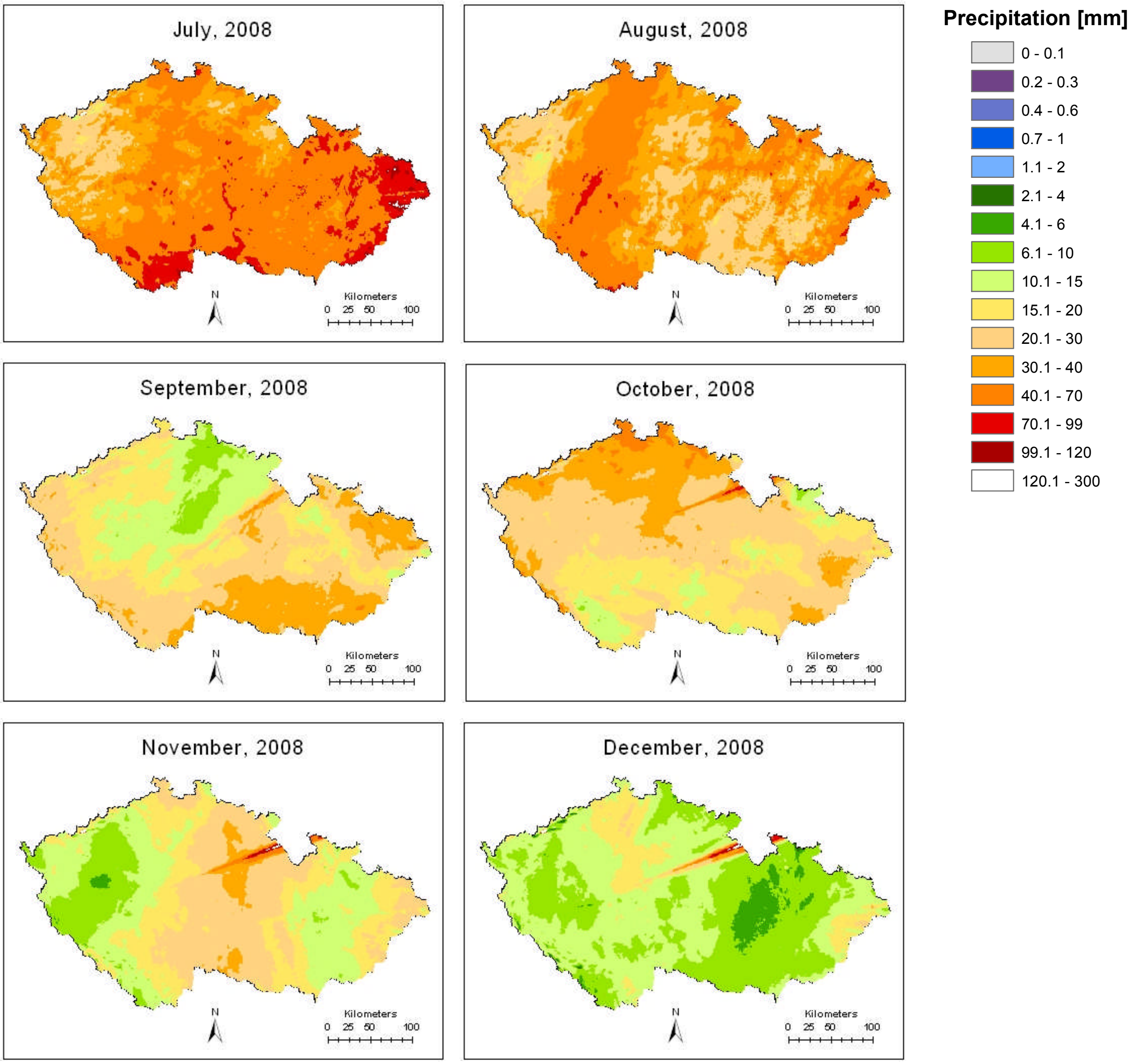


Fig. 1. Vizualisation of input data (monthly radar sums of precipitation in 2008)

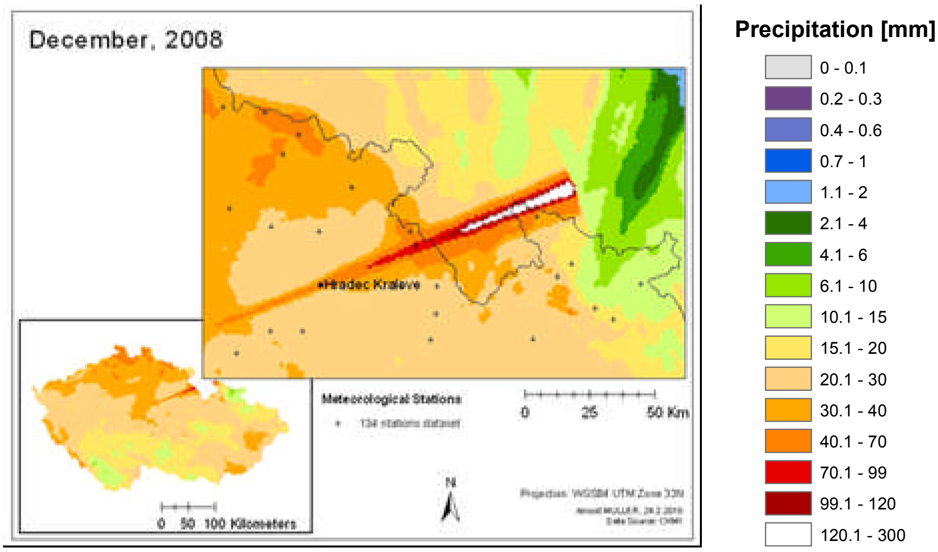


Fig. 2. The artificial ‘precipitation beam’ pattern caused by a microwave telecommunication link interfering with radar signal.

Fig. 2. displays a ‘precipitation artifact’, which appears in the data from September till November 2008. It is caused most likely by a microwave telecommunication link (probably by Wireless Local Area Network, WLAN) interfering with radar signal. Stations lying within this artifact are discussed later in the Analysis section.

Topographical data come from the digital elevation model (DEM) from the NASA/NGA Shuttle Radar Topography Mission (SRTM). Horizontal resolution of SRTM global datasets is 3 arc seconds, for latitudes in the Czech Republic each pixel represents 90 x 60 m. Variables derived from the DEM have the same spatial resolution. The vertical accuracy of the DEM characterized by RMSE (calculated from differences in real altitudes of 134 meteorological stations and altitudes derived from the DEM) is 7 m. Derived topographic variables are: slope, aspect, curvature, solar radiation, distance from the radar antenna and the variable named directional difference (DIF).

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

$$DIF=| \cos(ASPECT-DIR) |$$

The mathematical meaning of the DIF variable is illustrated in fig. 4 on the left. When ASPECT-DIR = ±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.

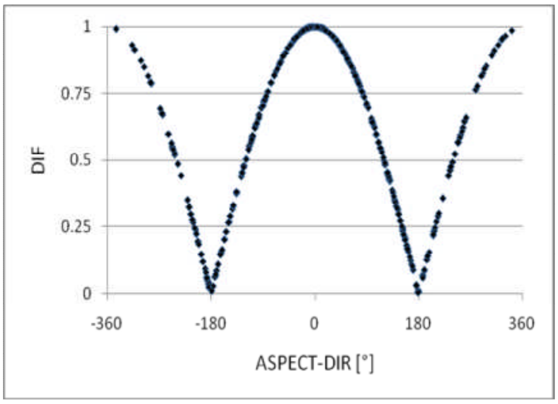


Fig. 4.: Chart explaining the meaning of the variable DIF

Fig. 5. illustrates the direction of the radar beam, which is then subtracted from aspect. Fig. 6. shows the variable DIF resampled to 1 km grid. Variable DIF was resampled from 90 m spatial resolution to 1 km as well as 10 km.

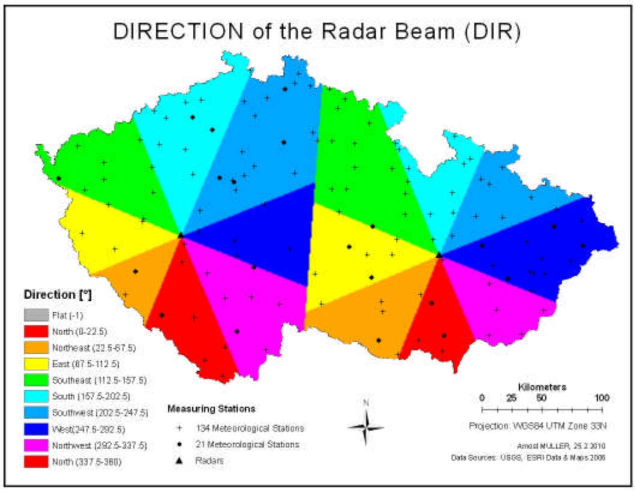


Fig. 5.: Direction of the radar beam (DIR)

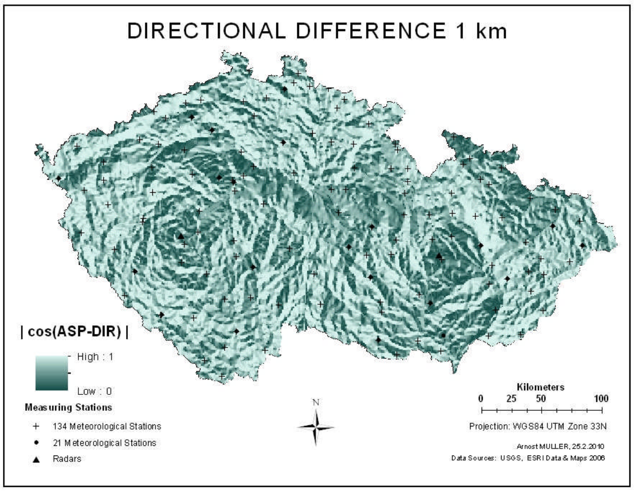


Fig. 6: Directional Difference (DIF)

ANALYSIS

Gauge measurements enter the regression analysis as dependent variable and radar rainfall sums as independent variable. The regression relationship is based on the assumption that rain gauge data are true and not biased, while annual radar sums of rainfall are not accurate [4]. The scatter plot in fig. 7. reveals some relationship between rain gauges and radar measurements. The slope coefficient of 1.74 indicates that radar underestimates the gauge rainfall by nearly twofold. The coefficient of determination for annual data R^2=0.18 is low, which means that there is low dependence between gauge and radar precipitation estimates.

One would expect a stronger relationship between the radar and the gauge data since both methods measure the same variable – precipitation. But as noted by Austin, radar samples almost instantaneously (at intervals of several minutes) a volume of atmosphere which has a surface projection of 1 square kilometer. The gauge accumulates continuously rain falling on an area which is smaller than 1 square meter. This error is known as nonrepresentative sampling. Rainfall often varies significantly over distances of less than a kilometer, while it may also change during time intervals of a minute or less. Therefore, the gauge measurements may not be representative of that in the entire area sampled by radar and similarly, radar-estimated rain rates observed instantaneously at any given measurement cell may not be representative of intensities during the intervals between observations [1].

Radar rainfall integrated in time to represent rainfall accumulations are typically adjusted to rain gauge-based areal average of the corresponding rainfall [4, 6, 7]. The radar data used in this analysis represent raw rainfall accumulations. If those had been calibrated with the gauge measurements, we would see a random noise pattern in the plot.

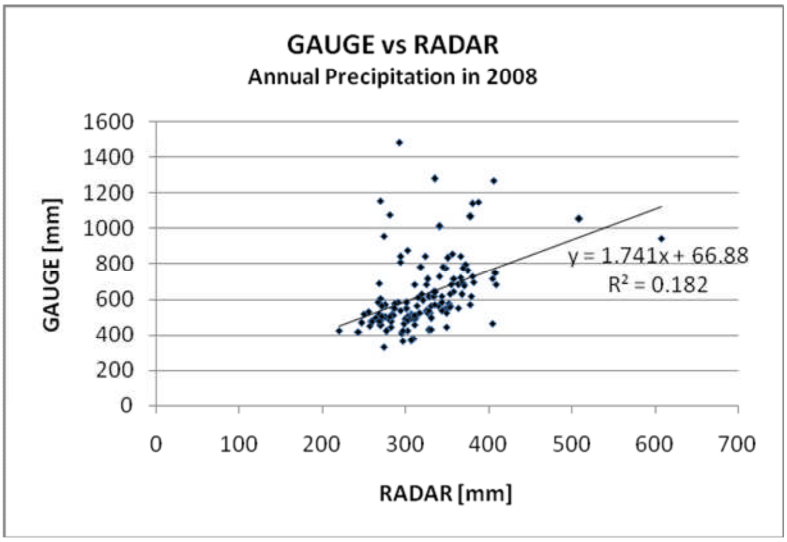


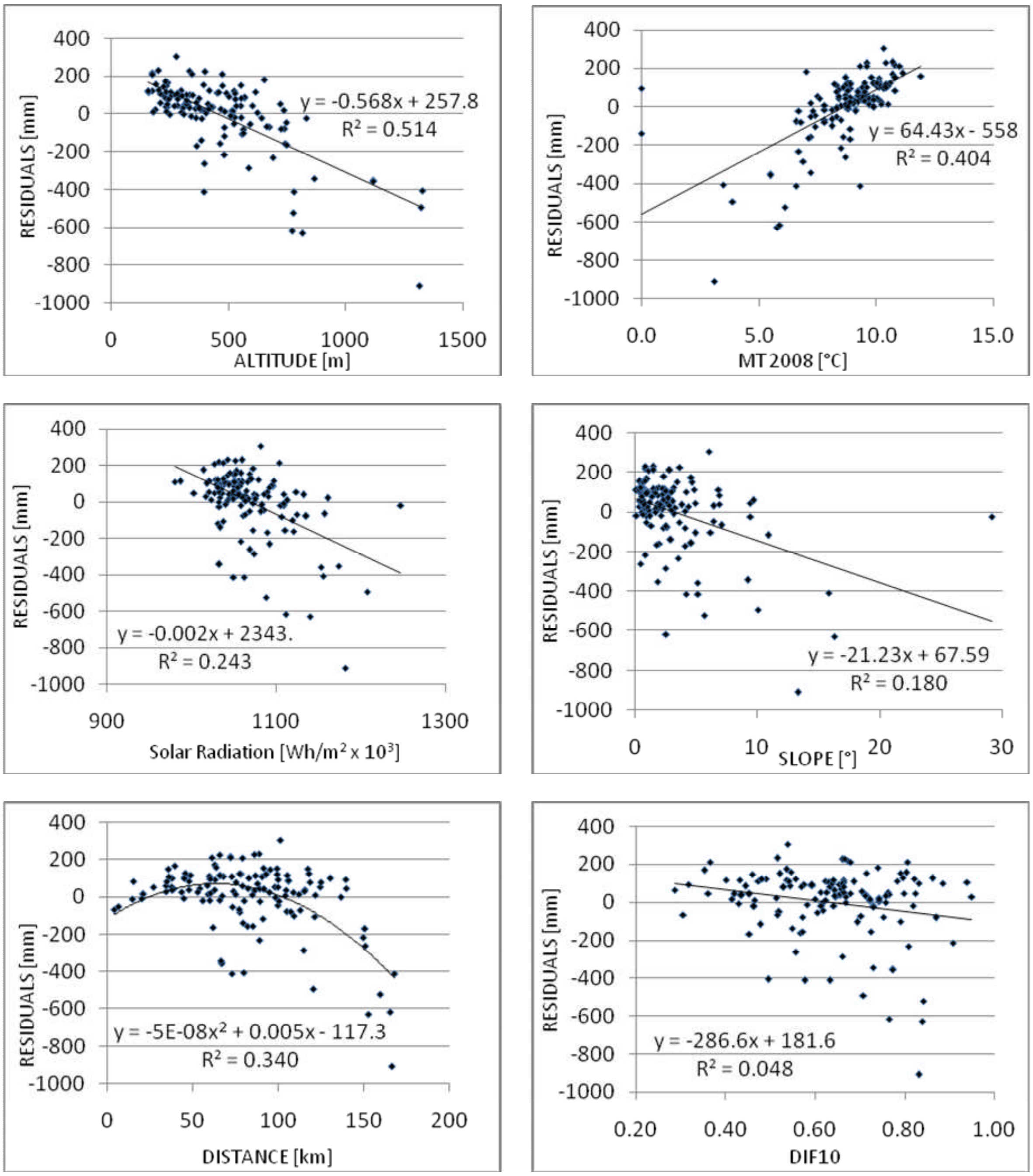
Fig. 7.: Gauge precipitation measurements of 134 meteorological stations plotted against annual radar rainfall sums over the same location

The residual analysis investigates the residuals of the regression model (illustrated in fig. 7.). Residuals (RES) are calculated as

$$RES = PRED - P$$

where PRED are predicted values by the regression model and P are rain gauge values. When the residual value is positive, the model (radar) is over-predicting the actual precipitation and vice versa. The highest residual values were found at stations 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 were negative, meaning that radar generally underestimates gauge rainfall. One exception was found at the station 'Hradec Kralove', where the residual value was positive. Meteorological station 'Hradec Kralove' lies in a 'radar beam artifact' caused by WLAN (refer to fig. 2.). Therefore, this station was excluded from the analysis. Stations 'Destne v Orl.H.' and 'Javornik' lying in the same 'WLAN artifact' did not show any outlying residuals and were not excluded.

Residuals were plotted against several variables (fig. 8.). A linear or a second order polynomial 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.



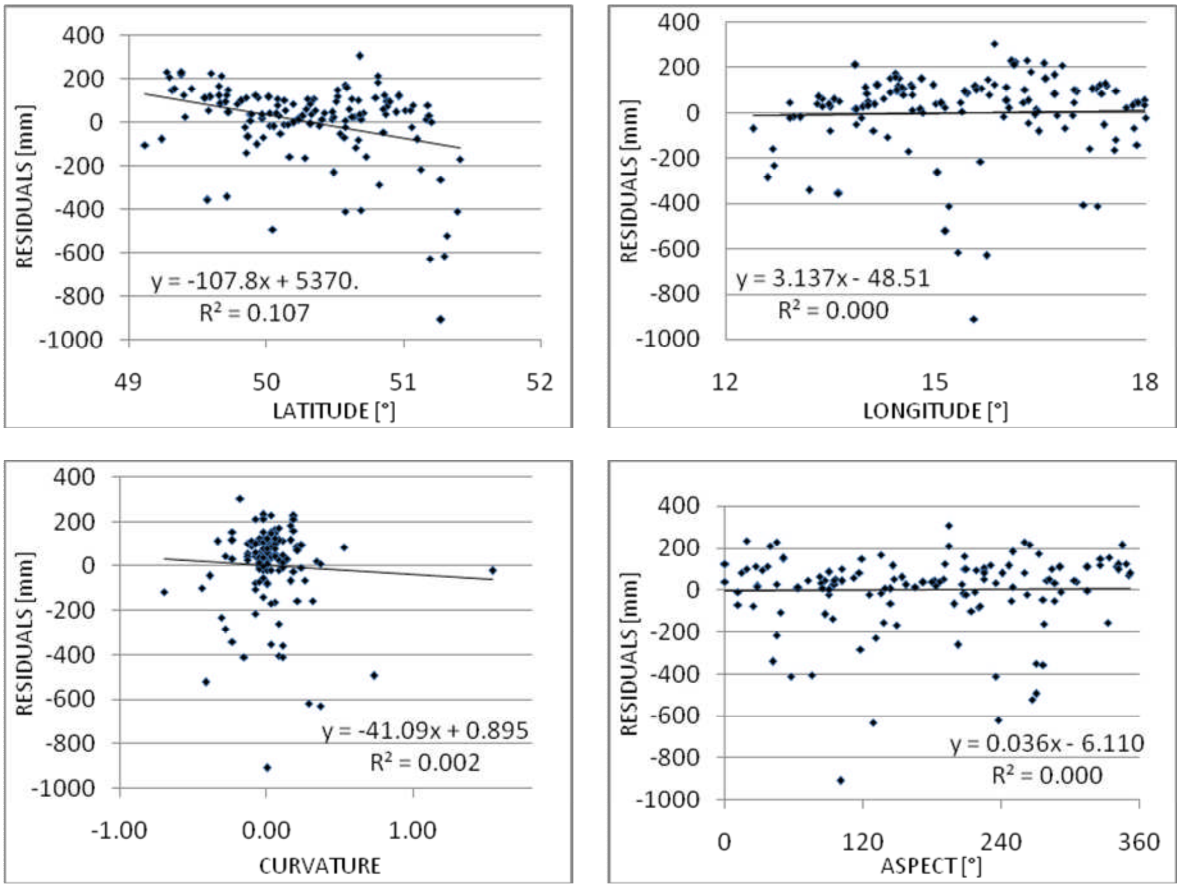


Fig. 8. Residual plots

RESULTS

Radar-derived precipitation generally underestimates gauge measurements and the underestimation increases with increasing distance from the radar. The main causes are the dispersing radar beam and its height above the curved Earths surface [1, 7].

The relationship between the distance from the radar antenna 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 (in absolute numbers negative) are associated with those stations farthest away from the radar antenna. Lowest residual values are at distances ranging from 50 to 100 km.

It is generally accepted that attitude 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 humidity, creating clouds and precipitation [8]. As shown in the corresponding fig. 8., 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 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.

Since mean air temperature (MT) can be interpolated over large areas with sufficient accuracy (standard deviation less than 0.5°C) [5], it could also be used as an independent variable in rainfall modeling. However, it should not be used together with altitude, with which it is highly correlated.

Solar radiation was calculated as a theoretical value for the year 2008 in ArcGIS using aspect and slope, and the solar angle. It does not include any actual information about cloudiness. Solar radiation in fig. 8. reveals some linear trend which considering the R^2 of 0.24 could explain ¼ of variation of residuals.

Regarding slope (fig. 8.), the majority of stations lie on flat surface of slope 5° or lower. Generally speaking, the higher the slope, the higher the residuals (in absolute numbers; when negative, it means radar is underestimating).

Directional difference (DIF) beteween the direction towards the radar antenna and aspect 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.

Latitude does not reveal any strong relationship with residuals, but the outliers lie in latitudes above 51°, which is around the norhtern borders, where the mountains spread out. Longitude nor curvature nor aspect does not have any significant influence on residuals (R^2=0.0?). In the case of curvature, residual values around 0 spread out around curvature of 0, where the terrain is flat. In other words, higher residuals appear where the terrain is not flat, but curved.

The R^2 values obtained through the residual plots show how certain factors influence the residuals. Tab.1. summarizes the R^2 values which are sorted in descending order according to the linear regression fit.

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 tab. 1. lying above the dashed line, which was selected arbitrary. Significant factors were identified using backward stepwise approach, which is documented in tab. 2.

Tab. 1.: Coefficients of determination between annual residuals and various factors

	R^2	
	linear	*Polynomial*
ALTITUDE	0.514	0.537
SOLARRAD	0.243	0.246
SLOPE	0.180	0.232
DISTANCE	0.156	0.338
LATITUDE	0.107	0.155
DIF10	0.048	0.051
DIF5	0.041	0.043
DIF	0.019	0.041
CURVATURE	0.002	0.013
LONGITUDE	0.000	0.032
ASPECT	0.000	0.036

Tab. 2.: Backward Stepwise Approach in Multivariate Linear Regression

			ALTITUDE		N_LAT		DIST		SOLARRAD		DIRDIFF10		MT2008		SLOPE	
R^2	**adj. R^2**	**F**	t	*prob.*	t	*prob.*	T	*prob.*	t	*prob.*	T	*prob.*	T	*prob.*	t	*prob.*
0.69	0.67	39	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	46	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	54	10.3	*0.000*	4.66	*0.000*	3.00	*0.003*	-1.81	*0.073*	1.79	*0.076*				
0.67	0.66	66	10.52	*0.000*	4.45	*0.000*	3.57	*0.000*	-1.99	*0.048*						
0.66	**0.65**	**85**	**13.55**	***0.000***	**4.18**	***0.000***	**3.64**	***0.000***								
0.63	0.62	110	13.52	*0.000*	6.30	0.000										

Tab. 3.: 2nd Order Polynomial Multivariate Regression

			ALTITUDE		DISTANCE		DISTANCE^2		LATITUDE	
R^2	**adj. R^2**	**F**	T	*prob.*	t	*prob.*	t	*prob.*	t	*prob.*
0.74	**0.73**	**90**	**13.88**	***0.000***	**-4.52**	***0.000***	**5.99**	***0.000***	**2.57**	***0.011***
0.72	0.72	112	13.38	*0.000*	-4.94	*0.000*	7.02	*0.000*		

The best fitting model (in tab. 2. and 3.) is highlighted bold. Significant factors according to tab. 3. are altitude, latitude and DIST. Relationship between DIST and residuals is better described by a polynomial relationship rather than linear. Distance squared was therefore introduced to the regression model to simulate 2nd order polynomial (quadratic) fit, refer to tab. 3.. The results are stronger than in the linear case, since R^2 increases from 0.66 to 0.74.

The three factors included in the multivariate polynomial regression model (tab. 3.) 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 are not a consequence of random variation.

CONCLUSIONS AND FUTURE PLANS

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. The first finding was that radar rainfall sums do not coincide nor significantly correlate (coefficient of determination R^2=0.18) with rain gauge observations due to high residual errors especially in mountainous regions. Radar underestimates annual rainfall at all gauges included in the 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 findings are important in regards to radar residual errors are not random, but can be partially predicted which may allow for some type of correction and improvement in radar’s accuracy.

The residual analysis was carried out at annual scale, given the data provided by CHMI. Having finer data e.g. at monthly scale could introduce some other topographic, locational or atmospheric variables and would allow us to look at seasonal variation. The main sources of radar’s inaccuracy discussed here come from topography, but there is the influence of the atmosphere as well, such as the attenuation of radar’s signal passing through near clouds, which blocks detection of further clouds. Such atmospheric effects are variable in time and space. Having wind direction observations it would be possible to develop and test new interactive variables mentioned in the scientific literature 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. Least but not last, the radar’s ‘visibility’ could be modeled and included as one of the independent factors in the analysis.

The ‘nonrepresentative sampling’ error could be reduced by interpolating rain gauge values (prior to analysis) into a grid (mean areal precipitation) matching the size of radar’s pixels, as described by Sokol [7].

ACKNOWLEDGEMENTS

This work has been supported by grant SGS10/151/OHK1/2T/11 of the Czech Technical University (CTU).

I would like to thank 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 also thankful to Tomáš Žejdlík from the Radar Department as well as Mgr. Michal Žák, Ph.D. from the Department of Climatology of the CHMI for providing me annual climatological and monthly radar data. Least but not last, I would like to thank Milan Šálek from the Regional Forecasting Office Brno, CHMI, for his corrections.

REFERENCES

[1] Austin, P.M. (1987) Relation Between Measured Radar Reflectivity and Surface Rainfall. Monthly Weather Review, 115, 5, 1053-1071.

[2] CHMI. http://www.chmi.cz/meteo/rad/eindex.html, CHMI Radar Network, 24.2. 2010.

[3] Gorokhovich, Y., Villarini, G. (2005) Application of GIS for processing and establishing the correlation between weather radar reflectivity and precipitation data. Meteorological Applications, 12, 1, 91-99.

[4] Krajewski, W.F. (1995) Rainfall Estimation Using Weather Radar and Ground Stations, International Symposium on Weather Radars, São Paulo, Brazil.

[5] Müller, A. (2010) Spatial Modeling of Climate, Faculty of Civil Engineering, Department of Mapping and Cartography, The Czech Technical University in Prague.

[6] Řezáčová, D., Sokol, Z., Kráčmar, J. , Novák, P. (2001) Statistical adjustment of radar-based daily precipitation to ground data from the Czech territory, Proceedings of 30th International Conf. on Radar Meteorology, 19-24 July 2001. Amer. Meteorol. Soc., Munich, 570-572.

[7] Sokol, Z. (2003) The use of radar and gauge measurements to estimate areal precipitation for several Czech river basins. Studia Geophysica Et Geodaetica, 47, 3, 587-604.

[8] Sokol, Z., Bližňák, V. (2009) Areal distribution and precipitation-altitude relationship of heavy short-term precipitation in the Czech Republic in the warm part of the year, Atmos. Res., 94, 652-662.