EVALUATION OF THE RADAR PRECIPITATION MEASUREMENT ACCURACY USING RAIN GAUGE DATA

Aurel Apostu[1], Mariana Bogdan[1], Coralia Dreve[1], Silvia Radulescu[1]

Administratia Nationala de Meteorologie, CNTMST – GRNRM, Bucuresti

Abstract

A weather radar system can provide spatial and temporal information on the clouds and precipitation over several hundreds square kilometres. Using raw data and specific processing algorithms, radar products are obtained. Among these products, a special attention is paid to rain rate product and precipitation accumulation products over predefined time intervals.

The facts that the rain rate is not a directly measured parameter by the radar and there are several limitations of the radar equipment that induce some times big errors in measuring the quantity of precipitation, opened a large opportunity for researches and studies, all these with the goal of obtaining, from radar measured data, the best estimates of the precipitation accumulation. The conventional algorithms for precipitation accumulation actually integrate the rain rate over a user selectable time interval.

Another source of information for the ground level precipitation is the rain gauges network. But, even these sensors cannot be considered "truth" for precipitation accumulation measurements, being affected by several limitations. It was demonstrated, both scientifically and experimentally, the best results are obtained when both data sources (radar and rain gauges) are used together.

This paper presents the results of a statistic analysis of the precipitation data measured during 2001 – 2003 period for the Muntenia region (southern Romania). The two data rows (measured with the Bucharest weather radar and with rain gauges) were processed using statistic – mathematic techniques in order to determine the correlation intensity and type. Two techniques for radar data adjustment using rain gauge data were applied and the results were studied and commented.

Key words: correlation, quantity of precipitation, radar, rain gauge

1. Introduction

The weather radar is, among other sensing instruments, one of the most modern and precise methods for atmosphere investigation.

The estimation of the precipitation quantity that will reach the ground, the estimation of the precipitation distribution inside hydrological basins and supplying the necessary data for hydrological forecast numerical models initialisation in order to prevent the effects of the flash floods, these are just few of the most important application fields for the weather radar.

Comparing to a rain gauge network, the weather radar as a precipitation sensor has couple of major advantages:

- the spatial continuity of the measurements (generally, the density of the weather stations provided with rain gauge is small comparing to the scale of the developed convective cells producing heavy rain, so that the convective cells could be "missed" by the rain gauge network or errors related to the maximum intensities caught by rain gauges could appear);
- possibility for real-time surveillance, from a single spot, over a large surface (for example, a radar with 10 centimetres wavelength and one degree beam width can perform precipitation measurements over 100 km range);
- digital acquisition, processing and storage for the radar data.

Typical radar "deficiencies" are ground clutter, beam blocking, vertical profile of reflectivity, different Z–R relations, etc.

The oldest instrument delivering a point measurement value of precipitation amount is the in-situ can-type precipitation gauge. The rain gauge measures the average precipitation quantity during a rain event, the highest error being induced by the airflow around the sensor (the drift of precipitation particles due to wind field deformation around the gauge). These instruments are also subject to a few another systematic errors, the most important sources of which are:

- lost water by wetting of the inner walls of the gauge;

- evaporation of water accumulated in the gauge “vessel”;
- splashing of raindrops or blowing of snow flakes out or into the gauge.

Nevertheless, data from precipitation gauges are still required for calibration of remote sensing techniques (radar, even satellite).

The relative error in measuring at the ground level the spatial and temporal distribution of a rain event using a rain gauge network depends on the observed precipitation structure and the network characteristics:

- increases with the distance between measuring points

and

- decreases when precipitation intensifies.

The rain gauge network used for radar measurements testing should have an average density of one measurement point at each 10 – 20 square kilometres. In such conditions (and if local scale storms are not present and air flow speeds are not high), the error for measuring precipitation quantities with a rain gauge network is less than 5%. A high-density rain gauge network is a satisfactory way to measure the precipitation at the ground level, but this is not enough. Radar information can be considered either supplementary data to be used for interpolation between measuring point on the ground or a way to extend the measurements over the terrestrial network.

2. Data and methods

The authors performed in this paper a first evaluation of the radar liquid precipitation measurement accuracy. The data used were provided by the Doppler weather radar system in Bucharest (EEC DWSR-2500C type, commissioned in October 2000). The precipitation data were collected for the 2001 – 2003 period in the Muntenia region, without being categorized by the season and precipitation type.

The authors used radar accumulation products (ACC) to be compared to the rain gauge data. ACC data was generated from:

a) maximum column reflectivity (CMAX)
   and
b) first elevation (0.5 degrees) radar reflectivity (PPIZ).

The two data rows are not simultaneous because the EDGE software (the software designed by manufacturer used for radar control, data acquisition, product generation and display) does not allow the simultaneous generation of the ACC products from two different reflectivity products. The ACC products are raster format with 1 km resolution.

The relationship used for converting reflectivity data (dBZ) into rain rate (mm/h) is Marshall – Palmer standard:

$$Z = 200\ R^{1.6}$$

The obtained rain rate is then integrated over specific time intervals (like one, three, six, 24 hours or any user selected time interval) resulting the ACC products for these intervals.

The correction factor automatically entered for all ACC products was F=1.

The raw radar data was acquired in a 24/7 manner, with one volume scan every 10 minutes, for a range of 240 km (figure 1). Beam blockage correction, rainfall attenuation, range correction and smoothing were applied.

For the analysis presented in this paper, 24 hours ACC products were used (the start moment for the ACC was 00:00 hours). The radar data were compared to the rain gauge measured precipitation accumulation, for the same 24 hours interval, from 35 meteorological stations located in the Bucharest radar surveillance area. Most of these weather stations are placed in the plain, in the range of 120 – 150 km relative to the radar location, and only few of them are placed in hilly and mountain areas.

In order to eliminate, as much as possible, the errors caused by the localization precision of each station and those caused by air movement near the rain gauge, the radar data were read in two different ways:

- the precipitation accumulation value exactly on the station spot (for all 35 stations);
- maximum value in a 5 km range relative to each station, or the average accumulation value when the maximum value was exactly on the station spot.

In this way, two rows of radar data (accumulation) were obtained for each type of radar ACC product (CMAX and PPIZ).

In this way, two rows of radar data (accumulation) were obtained for each type of radar ACC product.

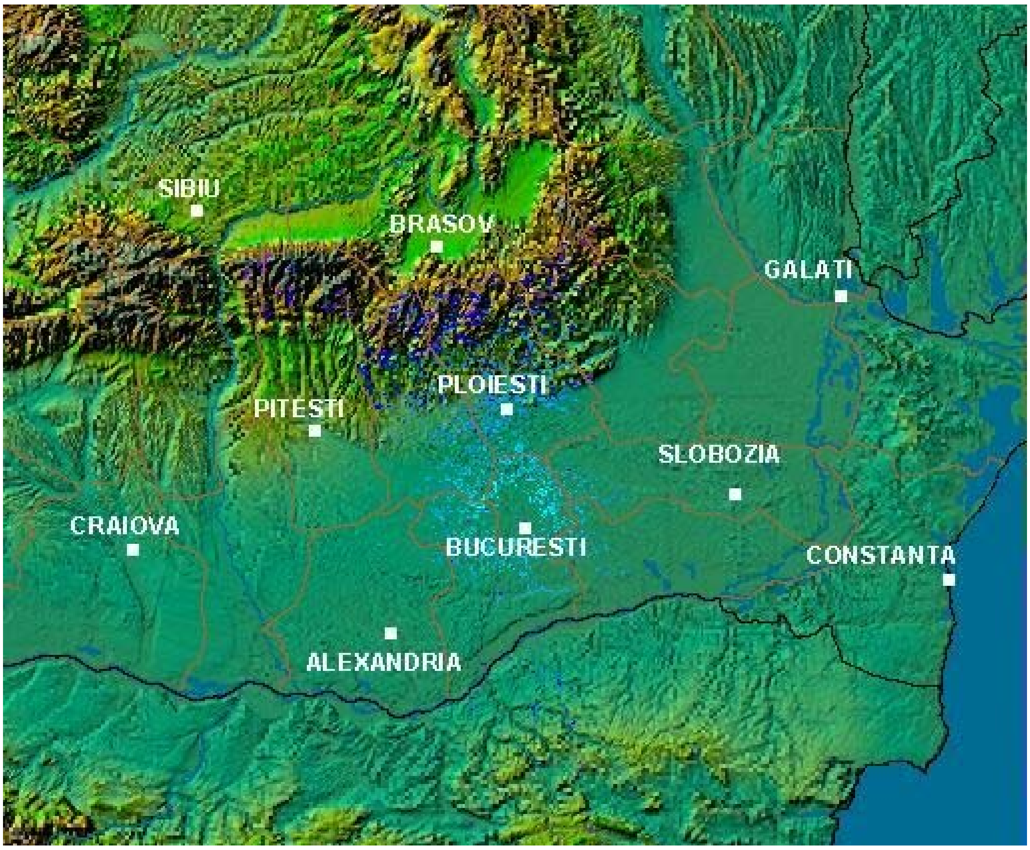


Figure 1

These four pairs of random variables, associated to the evaluation of the precipitation accumulation using the radar and, respectively, the rain gauge, were introduced in a statistical study based on co-relational theory. Considering the large drawn samples size (2934 for ACC generated from PPIZ and 2500 for ACC generated from CMAX) the normality assumptions necessary for parametric correlation are no longer a matter of concern, due to the Central limit theorem. However, they were tested with normal probability plots. For the correlation coefficients, Pearson's r formula was used since numerical variables, with no significant departures from normality are involved. Then three significance tests were applied to the obtained values (T test, classic Z test using Fisher transformation and ANOVA analysis). Using Fisher transformation, confidence intervals for the populations' correlation coefficients (ρ) were computed.

3. Results

The numerical values obtained for **r** indicates a good correlation between variables (table 1). The best correlation was obtained for the radar product generated from PPIZ and read on the station spot, while the worse correlation was for radar products obtained from CMAX in a 5 km range.

Table 1

r (PPIZ)	**r_5 (PPIZ)**	**r (CMAX)**	**r_5 (CMAX)**
0.68	0.668	0.625	0.595

According to the values obtained for the significance tests statistics, (much bigger than the critical values at the significance level α=0.01) the null hypothesis is rejected. There is indeed a good correlation between the variables.

Moreover, the performed statistical tests gave as well information about the regression slope; this parameter is significantly different from zero and therefore the regression lines were computed.

$R_{PPIZ} = 0.337*P + 0.852$; (1) $P = 1.375*R_{PPIZ} + 1.417$; (2) equation (2) - figure 2 a

$R_{PPIZ5} = 0.457*P + 1.874$; (3) $P = 0.977*R_{PPIZ5} + 0.839$; (4) equation (4) - figure 2 b
$R_{CMAX} = 0.44*P + 1.329$; (5) $P = 0.888*R_{CMAX} + 0.902$; (6) equation (6) - figure 2 c
$R_{CMAX5} = 0.482*P + 1.634$;(7) $P = 0.736*R_{CMAX5} + 1.003$.(8) equation (8) - figure 2 d

Figure 2a

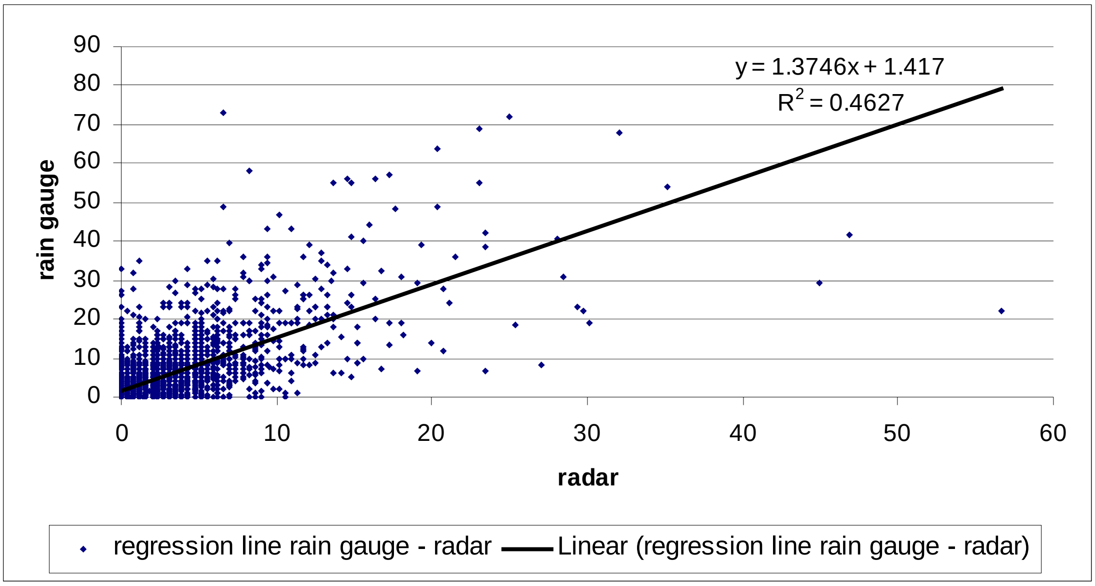
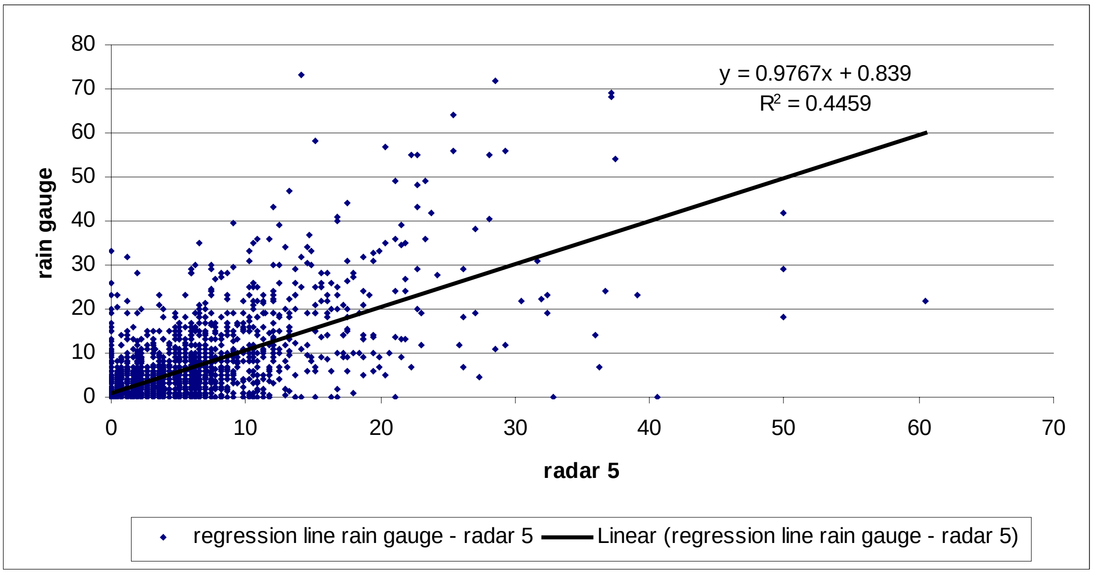


Figure 2b



In the next step, the samples were divided into 35 repeated selections for each random variable, the selection being done using the geographical criterion (for each rain gauge in the network). The presented co-relational analysis was repeated in order to determine a weight for each rain gauge in the network. Figures 3, 4, 5 and 6 are graphic representations for the 4 types of correlation coefficients.

Considering that, from statistical point of view, a coefficient between 0 and 0.4 means no correlation between variables, values between 0.4 and 0.7 are equivalent to a good correlation and values between 0.7 and 1.0 suggests a strong correlation, the cases with a correlation coefficient less than 0.4 were carefully considered.

For Braila station, located at a distance over 160 km from the radar site, the correlation is weak for PPIZ (radar beam, even at the first elevation, is approximately 3 km high), but it gets a lot better for CMAX. About Penteleu, Voinesti and Intorsura Buzaului stations (all being mountain stations), the correlation can be considered satisfactory, even good for at least two pairs of data rows. It can be noted the fact that, for all these stations, the best correlation is obtained for ACC products generated using PPIZ in 5 km range, while for ACC from CMAX the correlation is weak or even absent.

Figure 2c

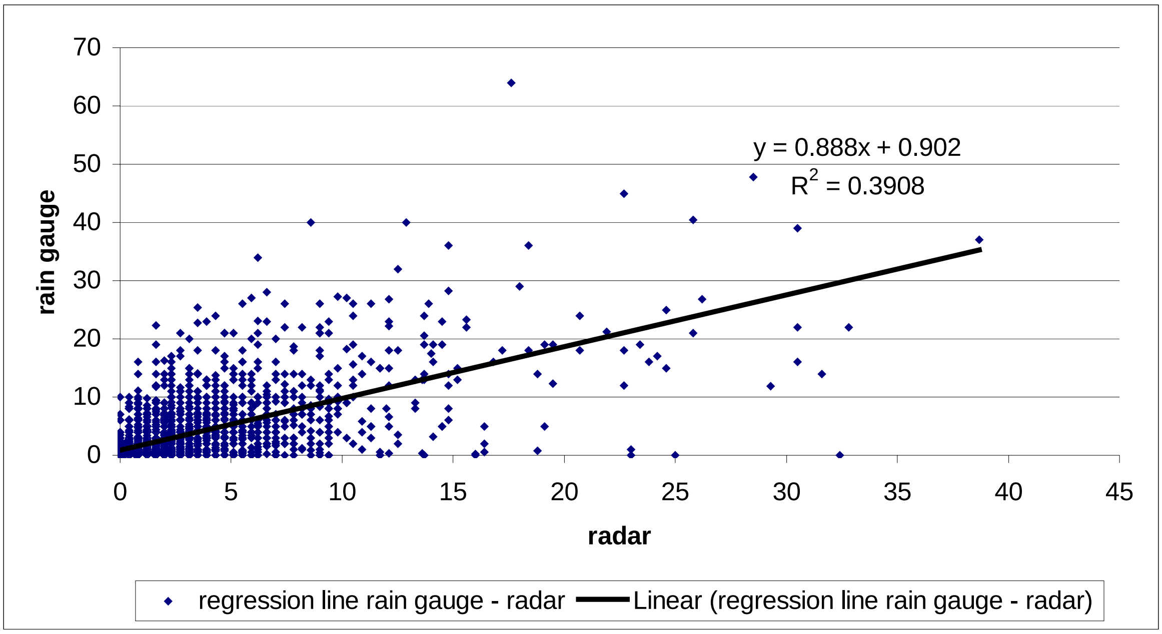


Figure 2d

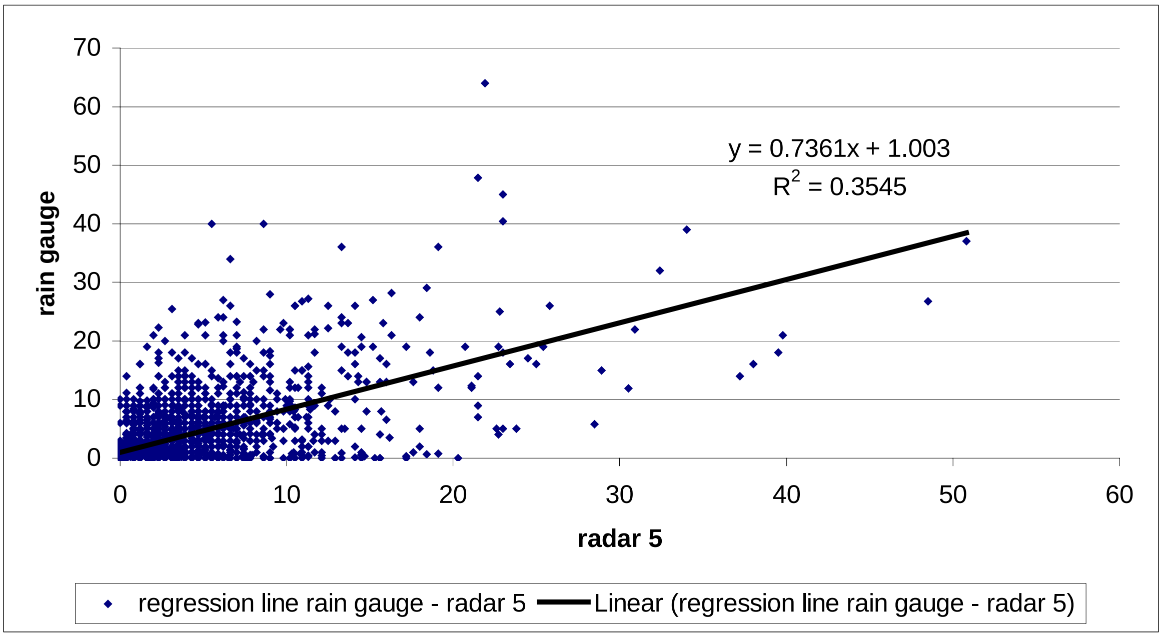


Figure 3

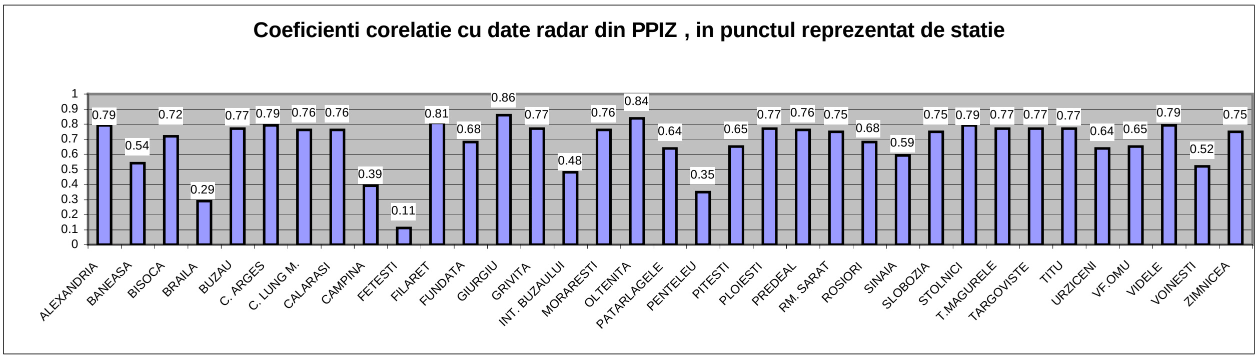


Figure 4

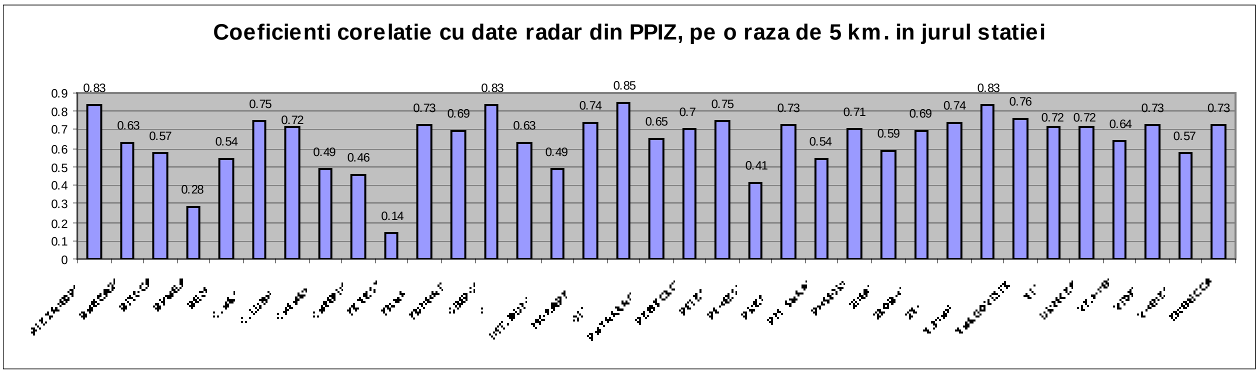


Figure 5

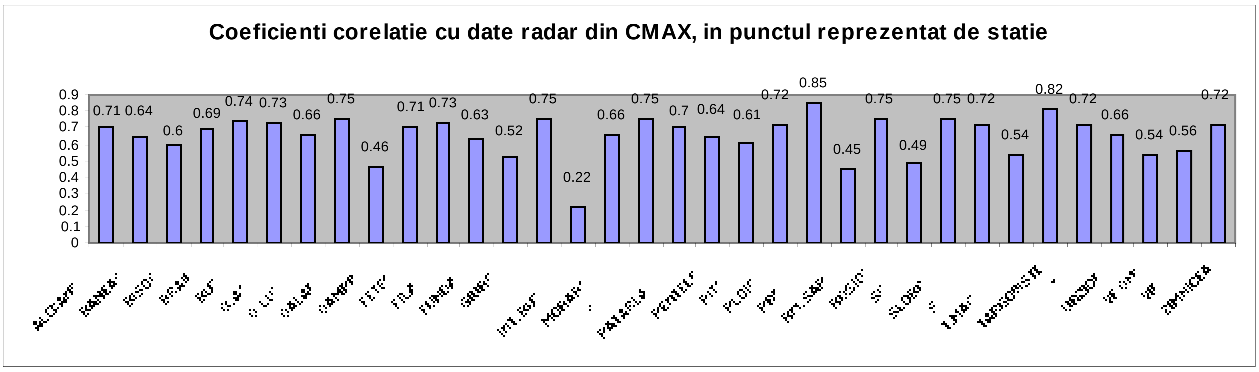
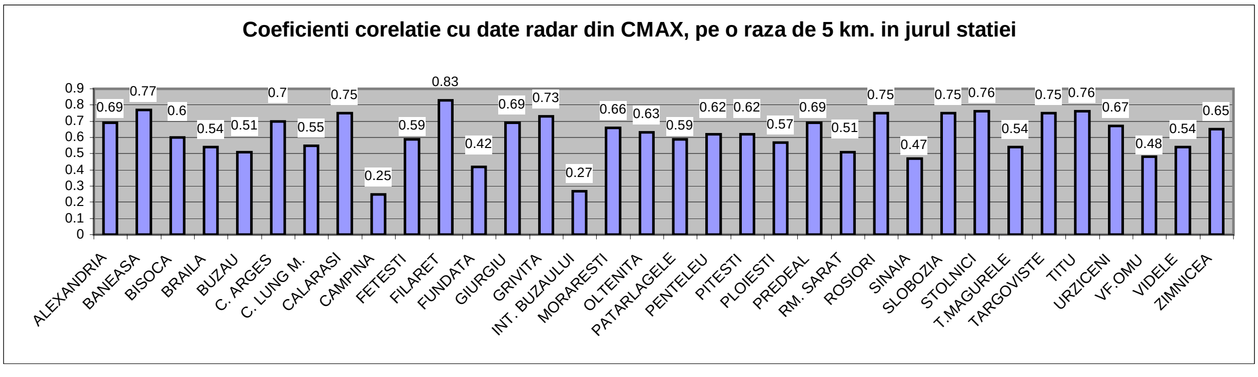
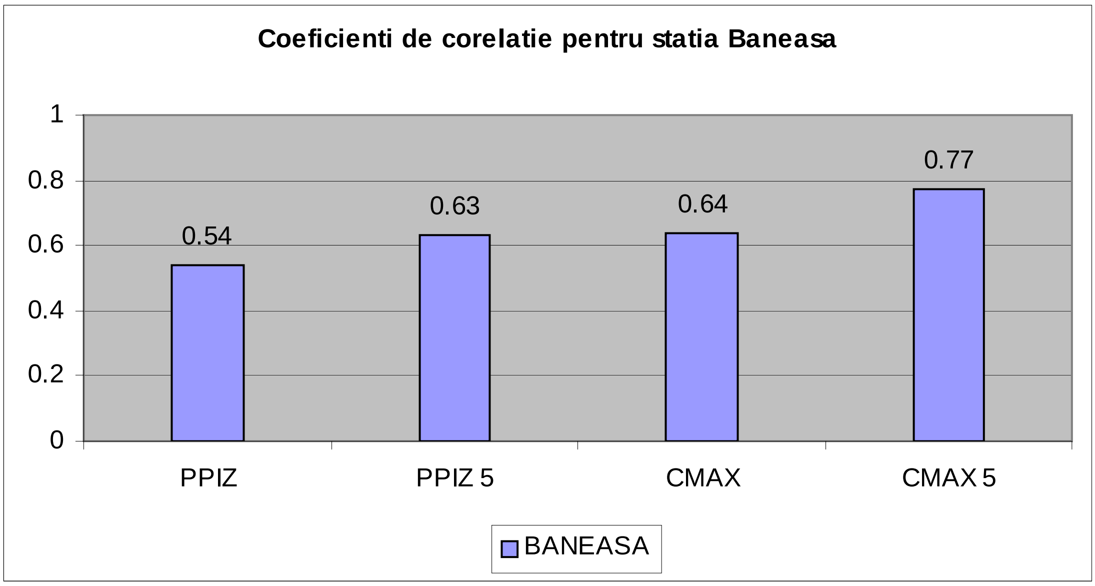


Figure 6



For Baneasa station (Figure 7), where the radar site is, the correlation for the CMAX 5 km (0.77) is the best from all 4 types of data, fact explainable by the "cone of silence" specific to all radar sites. The value of the correlation coefficients increasing with the distance from the radar site and with the elevation of the antenna, demonstrates this fact:

Figure 7



Another reason for comparing radar data with rain gauge data is to find a way to improve the accuracy and quality of the radar data. This can be achieved either

- modifying the coefficients of the Z-R relationship

or

- uniformly applying to the radar data of a correction factor obtained using the rain gauge data.

In order to adjust the radar data using the second method presented above, two multiplicative adjusting factors were applied. These factors were obtained using the following formulas:

$$F_1 = \sum_{i=1}^{n} P_i \Big/ \sum_{i=1}^{n} R_i \,; \qquad F_2 = (1/n) * \sum P_i / R_i \,.$$

When F1 is used, the data is weighted by the rain quantity, while when applying F2 all radar – rain gauge data pairs have equal weights. The corrections refer also to the errors due to the radar calibration and Z-R relationship. The following values were obtained for the factors F1 and F2:

$F_{1\,PPIZ} = 1{,}95$; $F_{1\,PPIZ\,5} = 1{,}18$; $F_{2\,PPIZ} = 2.18$; $F_{2\,PPIZ\,5} = 1.24$;
$F_{1\,CMAX} = 1{,}21$; $F_{1\,CMAX\,5} = 1.04$; $F_{2\,CMAX} = 1.41$; $F_{2\,CMAX\,5} = 1.33$.

As we are dealing, in both situations, with multiplicative factors, meaning linear transformations of radar data, the linear model performance remains the same (mean square error over sample variance).
Only the parameters of the regression lines (slope and intercept) are changing as can be seen from the equations.
However, another aspect ignored by the linear regression (but important) was analysed. It is about the basic question of whether the two methods of measurement agree sufficiently closely. The quantities that best answer this question are the differences between the pairs of data. If the measurements are comparable, the differences should be small, cantered around 0, and show no systematic variation with the mean of the measurement pairs. In terms of distribution parameters, the matter is best summarized by the standard deviation of the differences. If this number is small enough from a practical standpoint, then it can be said that the measurements are comparable.

From the analysis on this paper, it can be noticed that the radar data underestimate the precipitation values (the underestimation is bigger when the reading is taken right on the station spot). This can be corrected using the first adjusting factor: the average of the difference variable becomes zero and the standard deviation remains approximately the same.

Figure 8a

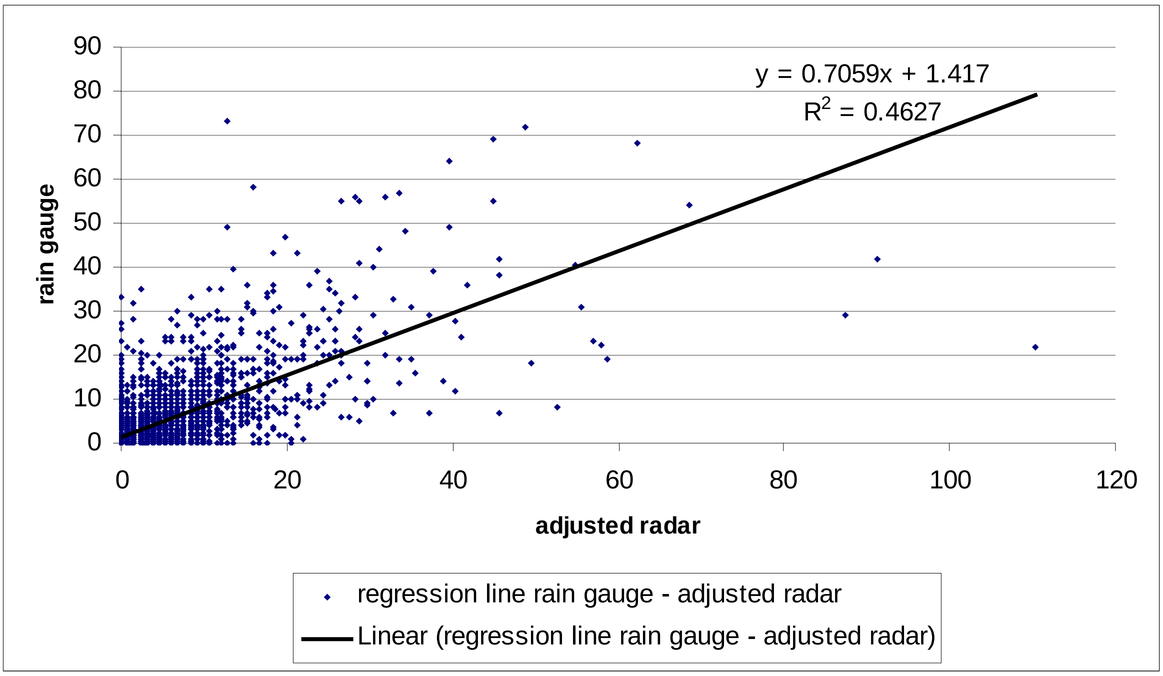
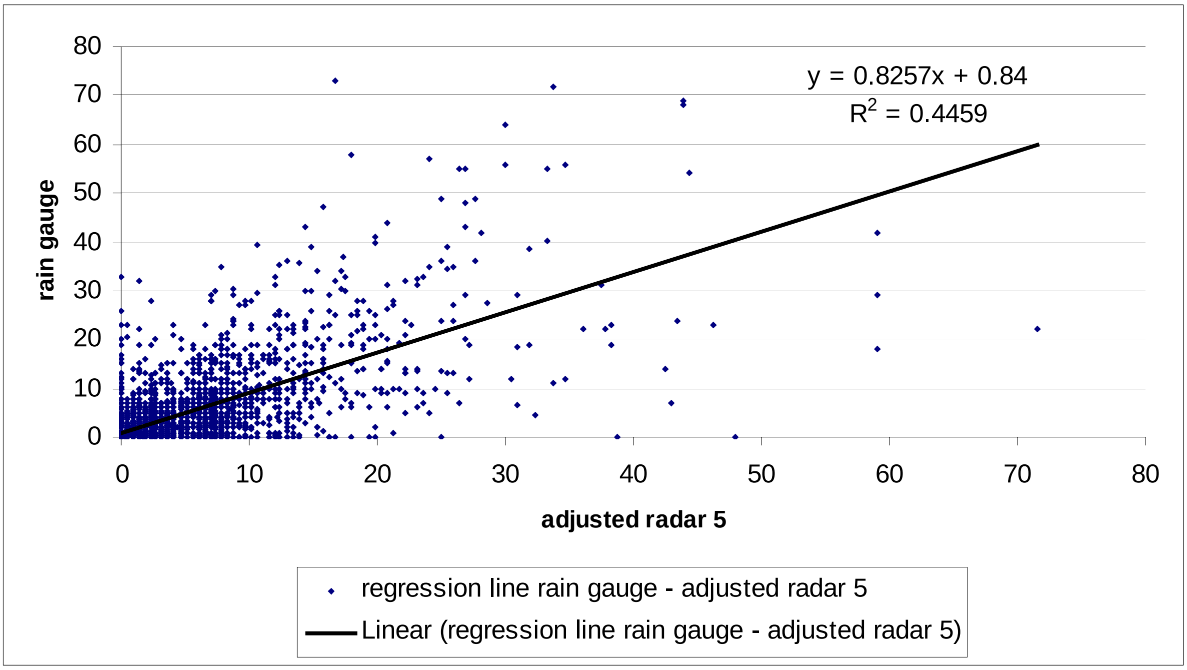


Figure 8b



In the case of the second adjusting factor, the values of the radar data increase too much and an overestimation of the precipitation is obtained; moreover, the standard deviation increases as well.
Consequently, the first adjustment factor (F_1) was preferred and the new coefficients for the regression lines were computed:

$R_{PPIZ} = 0.655*P + 1.66$; (1) $P = 0.706*R_{PPIZ} + 1.417$; (2) equation (2) - figure 8 a
$R_{PPIZ5} = 0.54*P + 2.216$; (3) $P = 0.826*R_{PPIZ5} + 0.839$. (4) equation (4) - figure 8 b
$R_{CMAX} = 0.531*P + 1.603$; (5) $P = 0.736*R_{CMAX} + 0.902$; (6) equation (6) - figure 8 c

$R_{CMAX5} = 0.502 * P + 1.703$; (7) $P = 0.706 * R_{CMAX5} + 1.003$; (8) equation (8) - figure 8 d

Figure 8c

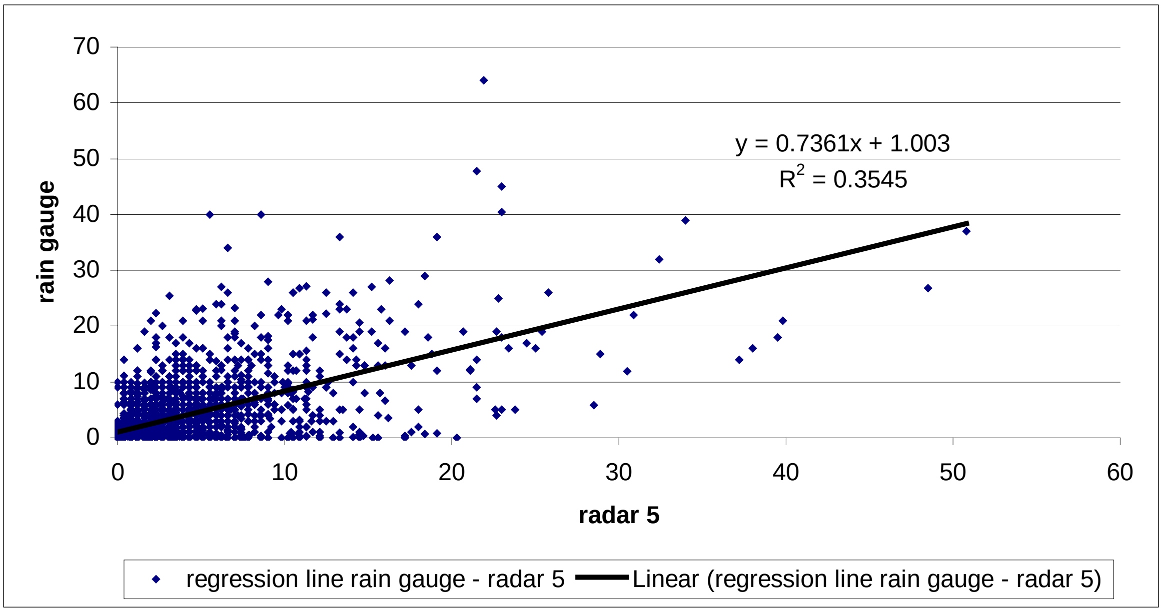
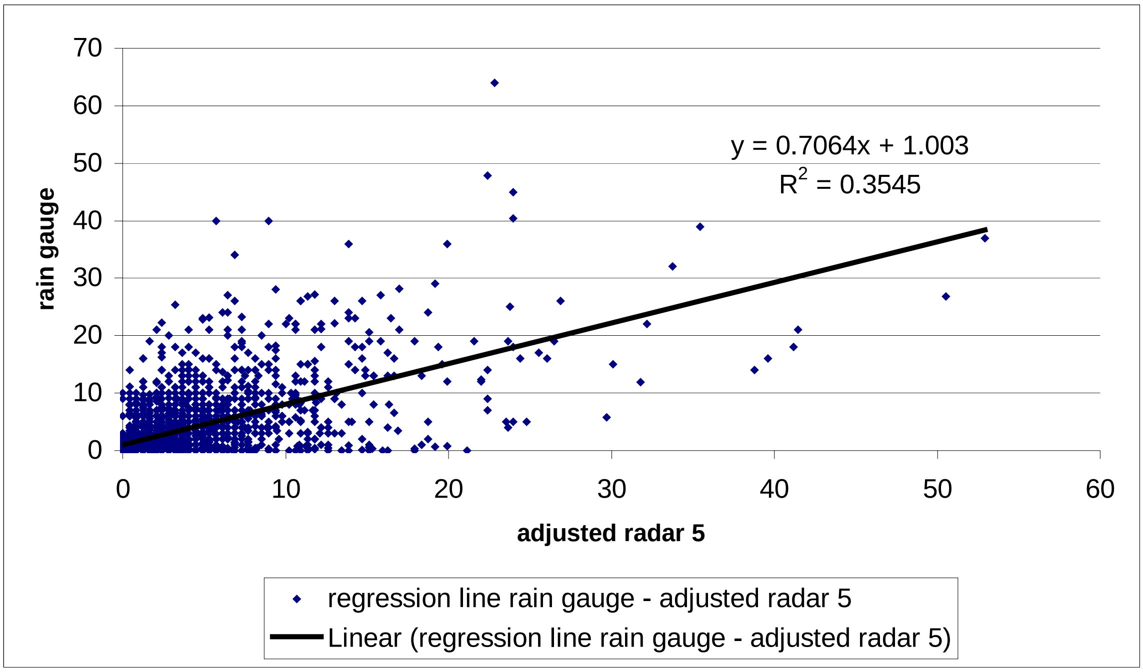


Figure 8d



Results of testing the factor F1

As a result of the above analysis, starting with 01 November 2004, the ACC product was generated using PPIZ and an automatic correction factor F1 = 1.95 was applied.

The consequences of adjusting the radar data by F1 were analysed using a sample of 497 data collected from the moment of its application. As shown above, the purpose of this correction concerns only the differences between the rain gauge precipitation values and the radar values taken exactly on the station spot. Therefore, the best way to verify the consequences is to compare the distribution parameters of the differences before and after adjustment. These are as can be seen below:

Before: $m_1 = -2.34$; $\sigma_1 = 6$
After: $m_2 = -0.7$ $\sigma_2 = 5.43$.

It can be noted that the average value had significantly decreased, which proves the applied correction to be a useful way of getting the two measured values much closer than they were. The standard deviation had also decreased even though not as significantly as the mean.

As a new sample is always a good opportunity of testing estimates homogeneity, Pearson's **r** has been computed again. The obtained values are:

r(PPIZ)	**r_5(PPIZ)**
0.76	0.738

It can easily be noticed the difference between the values; these coefficients do not even belong to the computed confidence intervals. Considering that very large samples estimates results are highly reliable, a sampling error may have occurred in this case. It was found testing the homogeneity of means and variances across the two samples for each variable, that the rain gauge precipitation data had a greater average in the smaller sample.

This study we may say, emphasized again the importance of the sample size in parametric analysis; very large samples are required especially when a variable phenomenon like precipitation is involved.

4. Conclusions

1. It was noticed a good correlation between precipitation data measured with DWSR-2500 C weather radar and precipitation data measured with rain gauge; the radar measured quantities are, in general, underestimated; the underestimation is more evident when the reading is taken in the station spot. However, there are cases when the radar quantities are higher than those measured with the rain gauge at the weather stations.
2. For the majority of the weather stations, the correlation between radar data and rain gauge data is better when the radar ACC products are generated using PPIZ first elevation products.
   For the mountain stations, the correlation is satisfactory for PPIZ 5 km and is missing for ACC products generated using CMAX.
   For the stations located at a great distance from the radar site (120 – 150 km), the correlation is better for the accumulations generated using CMAX.
3. The multiplicative factor that best adjusts the 24 hours radar accumulation data for Muntenia region (using DWSR-2500C equipment) is F1 (that weights proportional with the rain quantity) applied to radar accumulation using PPIZ products. From the data analysed and processed in this study, for the F1 factor was found the value of 1.95.
4. For other "reference" stations (like Baneasa), where the correlation is better when radar accumulation is generated using other product than for the majority of the weather stations, a specific adjustment factor will be computed and tested. This factor will be applied in a non-automatic manner.
5. In parametric analysis, when a variable phenomenon like precipitation is involved, very large samples are required.

Bibliography

[1] Rinehart, Ronald E, 1997, “Radar for Meteorologists”, Third edition

[2] Henri Sauvageot, 1982, “RADARMÉTÉOROLOGIE. Télédétection active de l’atmosphere”

[3] INMH, 1997, “Elemente de Meteorologie Radar “

[4] van Gorp, Jacques J., 6-6-2004, EU-action COST-717, WG-2, “A warning to the adjusters of radar precipitation measurements by means of rain gauges.”

[5] Sevruk, B. (2000), “Correction of the wind-induced error of Tipping-Bucket precipitation gauges in Switzerland using numerical simulation”, (presented at TECO-2000), in Instrument and Observation Methods report No. 74 (WMO/TD-No. 1028,144-147, 2000.