Hydrol. Earth Syst. Sci. Discuss., 10, 2453–2480, 2013
www.hydrol-earth-syst-sci-discuss.net/10/2453/2013/
doi:10.5194/hessd-10-2453-2013

Hydrology and Earth System Sciences
Open Access
Discussions

Precipitation accumulation analysis – assimilation of radar-gauge measurements and validation of different methods

E. Gregow[1], **E. Saltikoff**[1], **S. Albers**[2,3], **and H. Hohti**[1]

[1]Finnish Meteorological Institute, P.O. Box 503, 00101 Helsinki, Finland
[2]NOAA/ESRL/Global Systems Division, Boulder, Colorado, USA
[3]Cooperative Institute for Research in the Atmosphere, Fort Collins, Colorado, USA

Received: 25 January 2013 – Accepted: 6 February 2013 – Published: 28 February 2013

Correspondence to: E. Gregow (erik.gregow@fmi.fi)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

We investigate the appropriateness of four different methods used for combining radar data with precipitation gauge data to produce precipitation accumulation fields. These methods were validated for high-latitudes weather conditions of Finland. The reference method uses radar reflectivity only, while three assimilation methods are used to blend radar and surface observations together, namely: the linear analysis regression, the Barnes objective analysis and a new method based on a combination of the regression and Barnes techniques (RandB). The Local Analysis and Prediction System (LAPS) is used as platform to calculate the four different hourly accumulation products over a 6-months period covering summer 2011. The performance of each method is verified against both dependent and independent observations (i.e. observations that are or are not included, respectively, into the precipitation accumulation analysis). The new developed RandB-method performs best according to our results. Applying the regression- or Barnes assimilation analysis separately still yields better results for the accumulation products compared to precipitation accumulation derived from radar data alone.

1 Introduction

The concept of precipitation accumulation is of great importance for various applications in meteorology and hydrology. Climate projections under possible climate change scenarios point out to likely higher frequency of storms, with intensified precipitation over Europe. This will most probably have a significant effect on the surface water balance, therefore having a large impact on society and its economical aspects. Hydrological models, which are based on analyzed precipitation accumulation, do need a very high accuracy of the precipitated water amount in order to issue warnings, e.g. for sudden flooding. Fire-weather warnings are another example of products where

HESSD
10, 2453–2480, 2013

end-users require high quality data of precipitation accumulation during the summer period.

Radar-derived precipitation products are generated at high spatial resolution but embed measurement uncertainties. On the other hand, surface precipitation observations, such as standard gauge observations and road-weather measurements, have usually higher accuracy and are essential when used for correcting radar-based precipitation accumulation fields, but have limited spatial representativeness. The literature provides many studies on the benefits one can gain from the combination of radar measurements and surface observations to derive the final accumulated precipitation product (Goudenhoofdt and Delobbe, 2009). Radar reflectivity generates a good first guess for the accumulated precipitation, with the advantage of high spatial resolution, though there are certain inherent inaccuracies when deriving this product from radars (Koistinen and Michelson, 2002). Measurements of precipitation at ground-level are performed at point location and the errors associated with the observations are well characterized (Steiner et al., 1999). Different more or less sophisticated assimilation methods exist whereby surface point observations are blended together with radar data in order to establish a corrected precipitation accumulation, e.g.: co-kriging (Sun et al., 2000), the statistical objective analysis method (Pereira et al., 1998) and bias-adjustments using Kalman-Filter (Chumchean et al., 2006; Anagnostou and Krajewski, 1999). A summary of the methods and operational usage in different countries is compiled in the COST-717 report (Gjertsen et al., 2003). Problems linked to radar-gauge bias correction methods have been discussed in, e.g. Seo et al. (2002).

In this study, we use the Local Area and Prediction System – LAPS (McGinley et al., 1991, 1992) as a platform for testing and validating 4 different precipitation accumulation analyses: the radar only (hereafter LAPS_radar) and 3 assimilation methods, namely: the linear analysis regression, the Barnes objective analysis and a combination of those two methods (hereafter Regression, Barnes and RandB, respectively). The interest and incentive for using LAPS is that methods, such as Regression and Barnes, already existed in the system and could thereby be developed further, and that

HESSD
10, 2453–2480, 2013

LAPS is applicable for operational usage (Albers et al., 1996; Amy 2003), which is of critical interest for end-users who demand as close to real-time products as possible.

The aim of this article is to test and validate these methods for typical high latitudes summer weather conditions encountered in Finland (extending between 60 and 70° N) and to provide some guidance in the use of these methods. Section 2 introduces the LAPS model (Sect. 2.1), the radar data (Sect. 2.2) and the gauge network data (Sect. 2.3). The different analysis methods for estimating precipitation accumulation are introduced in Sect. 3. The results are presented and analysed in Sect. 4, while Sect. 5 provide some conclusions and outlook.

2 Methods and material

We describe here the model and data used to determine the gridded background fields involved in the estimation of the precipitation accumulation.

2.1 The Local Analysis and Prediction System (LAPS)

The Finnish Meteorological Institute (FMI) operates the Local Analysis and Prediction System (LAPS) for production of 3-D analysis fields of different weather parameters (Albers et al., 1996). LAPS uses a data fusion method, in which a high-resolution spatial analysis, using statistical methods, is performed on top of the coarser resolution background fields. Observations are fitted to the coarser first-guess analysis mainly by successive correction method, while high resolution topographical datasets are taken into account when creating the final high resolution analysis fields. Those analysis products are mainly used for now-casting purposes; i.e. what is currently happening and what will happen in the next few hours.

The coarser background first-guess field is the latest available forecast from the European Centre for Medium-Range Weather Forecasts (ECMWF) model, with a current horizontal grid spacing of approximately 16 km (ECMWF, 2011). The following ECMWF

HESSD
10, 2453–2480, 2013

parameters are used at 16 vertical pressure levels: vertical velocity, specific humidity, temperature, geopotential, vectorized winds, surface geopotential, surface pressure, pressure at mean sea level, 2-m temperature and dew-point temperature, vectorized wind at 10 meter, sea surface temperature, skin temperature and land-sea mask.

The FMI LAPS setup uses a pressure coordinate system including 44 vertical levels distributed with a higher resolution (e.g. 10 hPa) at lower altitudes and decreasing with height. The horizontal resolution is 3 km and the domain used in this article covers the whole Finland and some parts of the neighbouring countries (see Fig. 1a).

The fine-scale structures in the resulting 3-D-analysis are extracted from the observations. Therefore, LAPS highly relies on the existence of high-resolution, both spatial and temporal, observational network and especially on remote sensing data. At present, the LAPS suite implemented at FMI is able to process several types of in-situ and remotely sensed observations such as: radar radial winds/reflectivity (C-band), weighting gauges, road weather observations, soundings, Synop, Metar, air-traffic observations, lidars and Meteosat9 satellite data. The first three of these listed measurements are used for calculating the precipitation accumulation within LAPS.

2.2 The radar network

FMI operates eight C-band Doppler radars, which nearly cover the whole country. In southern Finland, the distance between radars is 140–200 km and measurements are made in bins that are 500 m long and 1° wide, up to 250 km in range. Thus, data from two or three radars are available over most of the study area. The location of the radars and their coverage is shown in Fig. 1a. As Finland has no high mountains, the horizon of all the radars is near 0° elevation with no major beam blockage, and, in general, the radar coverage is excellent up to 68° N latitude.

HESSD
10, 2453–2480, 2013

The effective radar reflectivity factor Z_e (usually called reflectivity) is derived from the expression:

$$Z_e = \frac{P_r \cdot r^2}{L \cdot C \cdot K^2}. \quad (1)$$

where P_r is the average received microwave power, r is the measurement range, L is the two-way attenuation in the propagation path (antenna–scatterers–antenna), C is a radar constant (including parameters of the radar hardware) and K is the dielectric factor (depending on the relative fraction of ice and water in the hydrometeors).

The reflectivity uses dBZ as unit, which is expressed as:

$$\mathrm{dBZ} = 10 \cdot \log_{10} Z_e. \quad (2)$$

The uncertainty factors affecting radar reflectivity are the electronic mis-calibration, calibration differences between radars, beam blocking, attenuation due to both precipitation (Battan, 1973) and wet radome (Germann, 1999). At mid-latitudes, the main source of uncertainty of radar-based rainfall estimates is the vertical profile of reflectivity, which causes a range-dependent error (Zawadski, 1984). At large distances, the radar probes the upper parts of the cloud, where reflectivity is weaker. This is compensated in the FMI system with the VPR- (Vertical Profile of Reflectivity) correction, which also compensates for overestimation in a melting layer when appropriate (Koistinen et al., 2003). Clutter is removed with Doppler-filtering, and any residual clutter with a post-processing procedure based on fuzzy logics (Peura, 2002).

In this study, the radar data was used as volume measurements, repeated every 5 min and consisting of 5 elevation angles, typically between 0.4 and 45°. Details of the FMI radar network and processing routines are described in Saltikoff et al. (2010).

2.3 Surface observations

For this study, a total of 447 rain gauges, both weighting gauges and optical sensors, provide detailed point information, which is used to correct the radar first-guess field

HESSD
10, 2453–2480, 2013

(introduced in Sect. 2.2). The verification period ranges from 11 April and 14 October, 2011, i.e. by and large the non-winter season (no snow-phase precipitation).

The surface precipitation observations are from standard weighting gauges and optical sensors mounted on road-weather masts. Weighting gauges are subject to different sources of errors such as mechanical malfunction, wind-drift (Hanna, 1995) and icing, which all affect the accuracy of measurements. FMI manages 77 stations instrumented with the weighting gauge Vaisala model VRG101. Measurements with this instrument have high cumulative accuracy (0.2 mm) provided that the precipitation event exceeds 0.5 mm. Depending on the station, the gauges measure the accumulated precipitation in intervals of 10 to 60 min. Summing these measurements over a 60 min period yields 1 h accumulation data.

The Finnish Transport Agency (FTA) runs 370 road-weather stations with optical sensor measurements (Vaisala Present Weather Detectors models PWD11 and PWD22), which have a precipitation detection sensitivity of 0.05 $\mathrm{mm\,h^{-1}}$ or less, within 10 min. Observation uncertainties from this type of stations are connected to their location in the immediate vicinity of roads with heavy traffic, where splash-effects and wind eddies, generated by big vehicles, occasionally affect the resulting accumulation. The precipitation intensity is measured in intervals ranging between 10 s to 5 min and finally summed up to 1 h precipitation accumulation information.

Another source of uncertainties in surface accumulation observations results from the limited spatial representativeness of many stations with respect to their surroundings, due to the insufficient density of measuring stations for certain areas (Cherubini et al., 2002). Note that if measurements consistently indicate poor data quality, those stations are blacklisted within LAPS and do not contribute to the precipitation accumulation analysis. Hereafter in this article, the weighting gauges and road-weather measurements are indistinctly called gauges and their distribution is shown in Fig. 1b.

HESSD
10, 2453–2480, 2013

3 Description of the four analysis methods

Thanks to its high resolution reflectivity pattern, weather radar data provides the best first-guess to calculate precipitation accumulation. The radar-based accumulation is calculated in the LAPS routine with the standard Z–R equation formula. On the other hand, gauges usually measure the accumulation with higher quality and are consequently used to correct the radar field. In this study, three different assimilation methods have been tested as to their capacity to perform the best radar-gauge correction: the Regression-, the Barnes- and a new RandB-method. These methods use the quotient between gauge and radar (hereafter G/R) for their corrections.

It has been noted, both in the literature and in our experiments, that the G/R-quotient is highly variable depending on the type of precipitation, a problem related to the used default drop size distribution (Battan, 1973). During light precipitation (drizzle) for instance, the gauges give usually larger values compared to radar and therefore the resulting G/R-quotient gets very large, often exceeding values of 30. On the contrary, in heavy precipitation cases (rain showers with embedded Cumulonimbus-clouds), the G/R-quotient usually leads to low values, less than 0.25. This discrepancy is related to the use of the standard Z–R equation formula for all liquid precipitation cases, even though we know that drop size distributions vary from one precipitation case to another. Another factor affecting the G/R-quotient may be the radar beam overshooting in shallow drizzle events. These circumstances could breed a substantial impact on the analysis and therefore the G/R-quotient has to be controlled when used within the different methods (see Sects. 3.2 and 3.3).

3.1 LAPS_radar based accumulation

The reflectivity Z parameter measured by the radar is converted to precipitation intenssity R (mm h^{-1}) using a pre-selected Z–R equation (Marshall and Palmer, 1948) as of the type:

$$R = A \cdot Z^b. \quad (3)$$

HESSD
10, 2453–2480, 2013

Where A and b are empirical factors describing the shape and size distribution of the hydro-meteors. Two sets of values are used in FMI's implementation of LAPS: for liquid precipitation, which is relevant in this study, $A = 315$ and $b = 1.5$ and for snow $A = 200$ and $b = 1.5$. This is a gross simplification since the drop size and particle shapes vary according to weather situation (drizzle/convective, wet snow/snow grain). Problematic situations include both convective showers with heavy rainfall and the opposite case of drizzle with little precipitation. Although such situations contribute only for a fraction of the annual precipitation amount, they might be important during, e.g. flooding events. On the other hand, looking at long-term averages, the radar accumulation data match the gauge accumulation values within reasonable accuracy (Aaltonen et al., 2008).

The intensity field is calculated every 5 min after each new set of radar observations has arrived. The 1 h accumulation is thereafter obtained by summing up over the 5 min intervals. The LAPS routines calculate the rain field every hour during summer as well as the snow-field during winter. In this study, only the rain-fields are considered.

3.2 The linear regression analysis method

As described above, accumulation estimates based only on radar data usually differs from gauge observation values due to radar errors (see Sect. 2.2) or problems with the gauges (Sect. 2.3). This is why various statistical methods have been used to address and reduce these differences. In the linear regression analysis method (hereafter Regression method), as a first step, the gauge-radar pairs from a given grid-point undergo a quality check to prohibit dubious differences between gauge and radar values. The aim is to avoid comparisons involving uncertain radar measurements and spurious surface observations. The selection is performed by discarding gauge-radar pairs exceeding specific thresholds based on the G/R-quotient. The thresholds are based on approximately 2 times standard deviation, STDEV(R/G), from LAPS radar dependent dataset (see Table 1). The thresholds used in the Regression method are:

If $G/R > 2.0$ then the gauge-radar pair is discarded.

If $G/R < 0.5$ then the gauge-radar pair is discarded.

HESSD
10, 2453–2480, 2013

The first threshold handles surface observations that are suspected to be false. The second criteria attempt to avoid cases where the radar gives too high reflectivity, for example in strong convective precipitation (including hail). Once these criteria are enforced, the remaining data form a dataset of representative gauge-radar pairs from which a linear regression can be established, calculated with the least-square method, which minimized the errors between the measurement pairs:

$$Y = k \cdot X + c. \quad (4)$$

In Eq. (4), Y is the corrected radar estimate, X is the first-guess accumulation from radar, while the analysis derives the regression coefficients k (the slope) and c (the interception point with the y-axis).

The Regression method benefits from having many gauge-radar pairs, since it will then create a more robust statistical relationship between the measurements. The behaviour of the linear curve has to be constrained since the shape of the curve is strongly influenced by the amount of gauge-radar pairs. Criteria for this have been set so to constrain k values between 0.2 and 5.0, and c values between −5 and +5 mm, in Eq. (4). These constraints were based on average vertical profile adjustments of reflectivity and relates to ranges of up to 200 km from radar station, during summer period (Koistinen et al., 2003). The linear function is applied to the whole radar accumulation field, i.e. corresponds to a regional scale correction.

3.3 Barnes objective analysis method

The Barnes interpolation forces the radar field to converge towards gauge accumulation measurements, using an objective multi-pass telescoping strategy (Barnes, 1964; Heimstra et al., 2006). The G/R-quotient is used to interpolate the first-guess radar field closer to the observation value and in order to optimize the result, several iteration steps are performed within the Barnes analysis at successively finer scales. For grid-points distant from any G/R observations, the G/R field tends smoothly towards a value of 1.

HESSD
10, 2453–2480, 2013

Depending on the precipitation pattern, this method can potentially result in a highly overestimated or underestimated reflectivity field being spread to the surroundings. For example, if there is one ground station situated at the border of a convective rain-shower (Cumulonimbus-cloud), where only light precipitation occur, the G/R-quotient would probably in this case exceed the value of 30, as described in Sect. 3. For the station point itself, this quotient gives an adequate correction but spreading this large quotient to the surrounding precipitation pattern could potentially give very large over-estimates of the accumulation within, for example in this case, the nearby core of rain-shower with heavy precipitation. Quality checks and thresholds have been set to avoid situations where such over- or underestimations of nearby precipitation areas are likely. If the G/R-quotient gives very large (> 30) or very small (< 0.25) values, this might still give a signal of an adequate trend, even though the signal is over-amplified. This trend has to be maintained and adapted but is given less weight in the resulting accumulation. Consequently, the chosen criteria must incorporate these aspects. The thresholds for the Barnes G/R-quotient are based on approximately 2 times standard deviation, STDEV(R/G), from LAPS radar dependent dataset (see Table 1). The following thresholds were used:

If $0.25 < G/R < 2.0$ then allow the derived quotient.

If $0.25 < G/R$ and $G/R > 0.0$ then reset $G/R = 0.25$

If $2.0 < G/R$ and $G/R < 100.0$ then reset $G/R = 2.0$

The modified Barnes scheme allows weighting (w_0) with distance (d) from the gauge station point with respect to the radius of influence (r), normalized by the instrument error (err_0), which is here set to be 1.5 in Eq. (5). The G/R-increment gives the initial increment (p_0) at the first iteration step, and the background weight (w_b) adjusts the output to be closer to radar value further away from observation point in Eq. (6).

$$w_0 = \frac{e^{-\left(\frac{d}{r}\right)^2}}{\mathrm{err}_0^2}, \tag{5}$$

HESSD
10, 2453–2480, 2013

$$p_{ij} = \frac{\sum(p_0 \cdot w_0)}{\sum p_0 + w_b}. \quad (6)$$

After the first iteration step, the p_{ij} output becomes the new G/R-increment (p_0) for the next iteration step in Eq. (6). The iterations continue with successively decreasing values of r in Eq. (5) until the observation increments have been diminished to tolerable values, in this case RMSE = 0.13 mm or alternatively after 10 iteration steps, in order to minimize the calculation time.

3.4 New method, combination of Regression and Barnes methods

This new method combines the above described Regression and Barnes analyses. First, the Regression method is used to correct the overall radar estimate, i.e. a regional scale correction. The resulting accumulation field is thereafter used as a new first guess, initializing the Barnes analysis which rectifies the radar field on local scales. Assuming that the new first-guess field from the Regression analysis is closer to the real precipitation accumulation, the Barnes correction method will not need to be too aggressive in its correction, thus minimizing the risk of exaggerating the surrounding precipitation with too low, alternatively too high, G/R-quotients.

4 Results and verification

The performance of the different methods has been verified against surface gauge observations of precipitation accumulation data. The verification period spans from 11 April to 14 October 2011, therefore assuming precipitation is in the form of liquid water, and the time sampling interval is one hour. The observations have been divided into two subsets: (i) one set including observations of all stations (but 7 of them) and (ii) a group of 7 SYNOP stations (excluded from the former set) used as an objective dataset for verification (Figs. 2–3 and 4–5, respectively). Accordingly, in the calculation

HESSD
10, 2453–2480, 2013

of the 1 h precipitation accumulation, the analysis depends on the station information from the first subset (i), hereafter called "dependent" stations. While the accumulation analysis is independent of the 7 stations in the second subset (ii), hereafter called "independent" stations. Each of these seven stations is representative of a characteristic Finnish climatological or physiographical areas such as coast-line, inland, lake district, etc.

The statistical quantification of the validation of the different analysis methods are based on the Root Mean Square Error (RMSE Eq. 7), and the Mean Absolute Error (MAE Eq. 8), calculated with these datasets, i.e.:

$$\mathrm{RMSE} = \sqrt{\frac{\sum(\mathrm{Analysis} - \mathrm{Gauge})^2}{N}}, \quad (7)$$

$$\mathrm{MAE} = \frac{\sum(\mathrm{Analysis} - \mathrm{Gauge})}{N}. \quad (8)$$

RMSE is a quadratic scoring rule, which measures the average magnitude of the error. Since the errors are squared before they are averaged, RMSE gives a relatively high weight to large errors. MAE measures the average magnitude of the errors in a set of analyzes, without considering their direction. It measures the accuracy for continuous variables. MAE is a linear score, which means that all the individual differences are weighted equally in the average. MAE and RMSE can be used together to diagnose the variation in the errors in a set of analyzes. RMSE will always be larger or equal to MAE. The greater the difference between them (RMSE-MAE), the greater is the variance in the individual errors in the sample (see Tables 1 and 2). If RMSE = MAE, then all the errors are of the same magnitude.

In Fig. 2 we show, separately for the four different methods, the relationship between the analyzed accumulation data at the LAPS grid-point closest to a gauge station and the corresponding gauge observations for the dependent stations. The correlation calculated from the datasets and the statistics of the comparisons are compiled in Table 1. It appears from these comparisons that the new RandB-method yields the best

HESSD
10, 2453–2480, 2013

agreement for accumulation precipitation compared to gauge observations, though the Barnes method also provides reasonable results. On the other hand, the regression method alone is not very successful but still improves the accumulation analysis to some extent. The LAPS_radar method, which is based on radar information only, gives the poorest results in our study.

In order to investigate the error dependencies between radar and gauges, we use an indicator that describes the hydrological aspects of the errors (Szturc et al., 2011), namely, the absolute difference between observed and analyzed precipitation accumulation as a function of the magnitude of the observed value (i.e. gauge data). Figure 3 shows that the linear fit has a smaller angle coefficient as one passes from the LAPS_radar, to Regression, Barnes and RandB analysis methods. This shows that the departure between analyzed and observed values decreases and again the RandB analysis performs best of the different methods.

We next investigate the agreement between the analyzed precipitation accumulation values and observations (gauge values) for the independent stations (Table 2). Note that for independent stations, there is much less data available. Through the independent stations we want to prove that the methods also work for areas where there are no observing stations available. Thus, verifying that there are no over- or under-amplified accumulation patterns devolving from especially the Barnes method (see Sect. 3.3), but also from the Regression method. The scatter plots (Fig. 4) indicate less scatter and slightly better agreement, i.e. smaller RMSE, MAE and higher correlation coefficient, compared to the dependent stations analysis (Fig. 2). The linear fitted curves in Fig. 4 are strongly influenced by the small amount of observation points, hence the distribution of high accumulation values (i.e. $> 10\,\mathrm{mm\,h^{-1}}$) have a large impact on the fitted curve. Although the comparison between the linear fitted curves in Fig. 4a–d gives a clear indication of how the different methods compare to each other, we plotted also for the independent stations the absolute difference between analyzed precipitation accumulation and observation as a function of gauge observations (Fig. 5). The same trend as with dependent station data is observed: less dependence of the RandB

HESSD
10, 2453–2480, 2013

method with increasing precipitation tending nearly to a systematic overestimation of ca. 5 mm h^{-1}.

In Sects. 2.2 and 3.1 we gave an explanation to the errors that are attributed to radar measurements, such as the range-dependent error and Z–R inaccuracies. These errors are related to the prevailing weather situation (e.g. thunderstorms or warm-fronts) and, hence, the type of precipitating hydro-meteors occurring at that time. Such influence was further investigated by dividing the different weather situations into two categories describing their air-mass stability: strong convection (hereafter convective) and light-moderate convection (hereafter non-convective), which relates to thunderstorms and warm-fronts, respectively. Each category includes 10 cases of a full 24-h day, also selected from the period 11 April to 14 October, 2011. The convective cases were determined by using FMI's lightning location system (Tuomi et. al., 2008) together with FMI radar archive, while the non-convective (warm-front) cases were selected from analyzed frontal passages over southern Finland as tagged by the duty forecaster at FMI.

The dataset representing the convective weather situations have fewer data values, compared to warm-front cases (see # values in Fig. 6). This is expected since convective precipitation is less likely to hit a gauge measuring device and generally last for shorter time, while large-scale precipitation events occurring during warm-fronts, have a much higher probability to come across a gauge station and have a larger temporal and spatial dimension. The results (Fig. 6) clearly show that the convective cases gives larger RMSE and MAE values, compared to non-convective cases. This is expected as convective precipitation situations display more spatial heterogeneity and thus a stronger decoupling from the gauge observations. This categorisation also indicates that the RandB-method performs best out of the four different methods, though only slightly better than the Barnes method.

HESSD
10, 2453–2480, 2013

5 Discussions and conclusions

In this article we compare the results from 4 different analysis methods on how to calculate the hourly precipitation accumulation: LAPS_radar, Regression, Barnes and a newly developed method RandB (combination of Regression and Barnes). The LAPS_radar serves as the reference method and since it is based on the common Z–R formula, this method is also similar to what is used at many meteorological services. The LAPS_radar is further used as the first-guess field when merging gauges data into the analysis routine of the three other methods.

The Regression method has the limitation of requiring a large number of valid gauge-radar pairs in order to fulfil the least-square calculations and thereby creating a sufficient linear curve fit between the gauge network and radar observations. If there are not enough valid pairs, or if the criteria for a linear dependency are not fulfilled, then the regression method will not be used and the analysis will fall back to the original LAPS_radar based initial precipitation accumulation field. Therefore, there will not always be any impact on the accumulation field when using the Regression analysis. The Barnes method will in the same way fall back to the original LAPS_radar based accumulation field if there are no observations available, or if the radar-gauge pairs does not fulfil the thresholds stipulated for the G/R-quotient. The new RandB-method encounters the same restrictions as described above, since it is a combination of the Regression and Barnes methods.

In order to be meaningful for operational purposes, the studied merging methods should therefore show at least as good a result as the LAPS_radar precipitation accumulation analysis. Figures 2, 4 and 6 confirm that applying an assimilation method improves the overall results. In Figs. 3b–d and 5b–d one can see that dots congregate closer to the zero value along the x-axis, indicating a better match between analyzed and observed value. The calculated statistics, including both the dependent, independent, convective and non-convective datasets, also state that agreement is improved by applying a merging method. The error values of RMSE and MAE are

HESSD
10, 2453–2480, 2013

decreasing, compared to LAPS_radar values, and for the RandB-method with the dependent dataset the corresponding reduction in RMSE and MAE are 29 % and 47 %, respectively. The correlation, for RandB dependent dataset, is increasing (41 %) accordingly and the variance (RMSE-MAE) is decreasing when applying the different assimilation methods. Similar results are seen in the independent, convective and non-convective datasets.

When studying the results from two different stability weather situations, i.e. convective and non-convective, the main findings are that the RMSE and MAE are considerably higher in convective cases. This indicates that the four accumulation methods adopted in this study is more sensitive to convective situations. We interpret that this is related to the larger spatial variability of convective precipitation as well as different drop size distributions. In convective situations, the real intensity is variable within each radar measurement bin (typically representing several cubic kilometres), and it is a random process, which is only partly captured at a single gauge (orifice diameter of order of 15 cm). Also the Z–R equation used in Finland has been optimized for total rainfall, which in areas of extra-tropical cyclones consists largely of frontal precipitation, e.g. warm-fronts. As a consequence, when the discrepancy between radar and gauge observations (i.e. large G/R-quotients) is significant for the convective cases, the thresholds (see Sects. 3.2 and 3.3) are more frequently exceeded within the Regression-, Barnes- and RandB- analyses. This leads to fewer corrections being done from the gauge measurements and the resulting accumulation analysis is worse for convective weather situations, compared to non-convective cases.

On the other hand, optimising the Z–R equation for some specific types of precipitation should lead to a more faithful merging, which should be reflected in the agreement between analysed and observed precipitation. When such approach would be performed, using a much larger dataset basis, the RMSE and MAE value of the agreement for specific precipitation types should naturally tends towards better performance than without any differentiation between precipitation types, and could be used thus as a test.

HESSD
10, 2453–2480, 2013

The conclusive results from this study is that the new developed RandB-method, i.e. the combination of Regression and Barnes analysis methods, generates the best estimate of 1 h precipitation accumulation. Also, applying either Regression or Barnes method separately, still yields a better result than solely using radar accumulation, i.e. LAPS_radar method.

Acknowledgements. We want to thank Victoria Sinclair for her help in finding the warm front cases and Sylvain Joffre for advice and encouragement.

References

Aaltonen, J., Hohti, H., Jylhä, K., Karvonen, T., Kilpeläinen, T., Koistinen, J., Kotro, J., Kuitunen, T., Ollila, M., Parvio, A., Pulkkinen, S., Silander, J., Tiihonen, T., Tuomenvirta, H., and Vajda, A.: Strong Precipitation and Urban Floods (Rankkasateet ja taajamatulvat RATU), Finnish Inst. of the Environement (Suomen ympäristökeskus), Finland, 80 pp., 2008.

Albers, S. C., McGinley, J. A., Birkenheuer, D. L., and Smart, J. R.: The local analysis and prediction system (LAPS): analyses of clouds, precipitation, and temperature, Weather Forecast., 11, 273–287, 1996.

Amy, H.: Precipitation Verification of the Local Analysis and Prediction System (LAPS) Storm Total Precipitation Estimates, St. Cloud State University and NOAA/NWS Weather Forecast Office, Minnesota, 2003.

Anagnostou, E. N. and Krajewski, W. F.: Real-time radar rainfall estimation – Part I: Algorithm formulation, J. Atmos. Ocean. Tech., 16, 189–197, 1999.

Barnes, S. L.: A technique for maximizing details in numerical weather map analysis, J. Appl. Meteorol., 3, 396–409, 1964.

Battan, L. J.: Radar Observation of the Atmosphere, University of Chicago Press, Chicago, 1973.

Cherubini, T., Ghelli, A., and Lalaurette, F.: Verification of precipitation forecasts over the Alpine regions using a high-density observing network, Weather Forecast., 17, 238–249, 2002.

Chumchean, S., Sharma, A., and Seed, A.: An integrated approach to error correction for real-time radar-rainfall estimation, J. Atmos. Ocean. Tech., 23, 67–79, 2006.

HESSD
10, 2453–2480, 2013

ECMWF: http://www.ecmwf.int/research/ifsdocs/CY37r2/index.html, last access: 30 December 2011.
Germann, U.: Radome attenuation – a serious limiting factor for quantitative radar measurements?, Meteorol. Z., 8, 85–90, 1999.
Gjertsen, U., Salek, M., and Michelson, D. B.: Gauge-adjustment of radar-based precipitation estimates – a review, COST-717 working document No. WDD 02 200310 1, 33 pp., 2003.
Goudenhoofdt, E. and Delobbe, L.: Evaluation of radar-gauge merging methods for quantitative precipitation estimates, Hydrol. Earth Syst. Sci., 13, 195–203, doi:10.5194/hess-13-195-2009, 2009.
Hanna, E.: How effective are tipping-bucket raingauges? A review, Weather, 50, 336–342, 1995.
Koistinen, J. and Michelson, D. B.: BALTEX weather radar-based precipitation products and their accuracies, Boreal Environ. Res., 7, 253–263, 2002.
Koistinen, J., Michelson, D. B., Hohti, H., and Peura, M.: Operational measurement of precipitation in cold climates, in: Weather Radar Principles and Advanced Applications, edited by: Meischner, P., Springer, Germany, 337 pp., 2003.
Marshall, J. S. and Palmer, W. M. K.: The distribution of raindrops with size, J. Atmos. Sci., 5, 165–166, 1948.
McGinley, J. A., Albers, S. C., and Stamus, P. A.: Validation of a composite convective index as defined by a real-time local analysis system, Weather Forecast., 6, 337–356, 1991.
McGinley, J. A., Albers, S. C., and Stamus, P. A.: Local data assimilation and analysis for nowcasting, Adv. Space Res., 12, 179–188, 1992.
Pereira Fo, A. J., Crawford, K. C., and Hartzell, C. L.: Improving WSR-88D hourly rainfall estimates, Weather Forecast., 13, 1016–1028, 1998.
Peura, M.: Computer vision methods for anomaly removal, in: Second European Conference on Radar Meteorology, ERAD, Delft, Netherlands, 312–317, 2002.
Saltikoff, E., Huuskonen, A., Hohti, H., Koistinen, J., and Järvinen, H.: Quality assurance in the FMI Doppler Weather radar network, Boreal Environ. Res., 15, 579–594, 2010.
Seo, D. J. and Breidenbach, J. P.: Real-time correction of spatially nonuniform bias in radar rainfall data using rain gauge measurements, J. Hydrometeorol., 3, 93–111, 2002.
Steiner, M., Smith, J. A., Burges, S. J., Alonso, C. V., and Darden, R. W.: Effect of bias adjustment and rain gauge data quality control on radar rainfall estimation, Water Resour. Res., 35, 2487–2503, 1999.

HESSD
10, 2453–2480, 2013

Sun, X., Mein, R. G., Keenan, T. D., and Elliott, J. F.: Flood estimation using radar and raingauge data, J. Hydrol., 239, 4–18, 2000.
Szturc, J., Ośródka, K., and Jurczyk, A.: Quality index scheme for quantitative uncertainty characterization of radar-based precipitation, Meteorol. Appl., 18, 407–420, 2011.
Tuomi, T. J. and Mäkelä, A.: Thunderstorm climate of Finland 1998–2007, Geophysica, 44, 29–42, 2008.
Zawadzki, I.: Factors affecting the precision of radar measurement of rain, in: 22nd Conference on Radar Meteorology, Zurich, Switzerland, 10–13 September 1984, 251–256, 1984.

HESSD
10, 2453–2480, 2013

Table 1. Statistical verification results of the different methods for the dependent stations dataset.

	LAPS_radar	Regression	Barnes	RandB
Number of observations	111 821	102 016	111 821	111 821
STDEV(R/G)	1.11	1.23	0.53	0.55
RMSE	1.38	1.32	1.03	0.98
MAE	0.73	0.69	0.43	0.39
RMSE-MAE	0.85	0.63	0.60	0.59
CORR	0.51	0.56	0.69	0.72

HESSD
10, 2453–2480, 2013

Table 2. Statistical verification results of the different methods for the independent stations dataset.

	LAPS_radar	Regression	Barnes	RandB
Number of observations	2648	2436	2648	2648
STDEV(R/G)	1.67	1.47	1.41	1.19
RMSE	1.29	1.23	0.95	0.91
MAE	0.72	0.68	0.44	0.40
RMSE-MAE	0.57	0.55	0.51	0.51
CORR	0.60	0.65	0.80	0.81

HESSD
10, 2453–2480, 2013

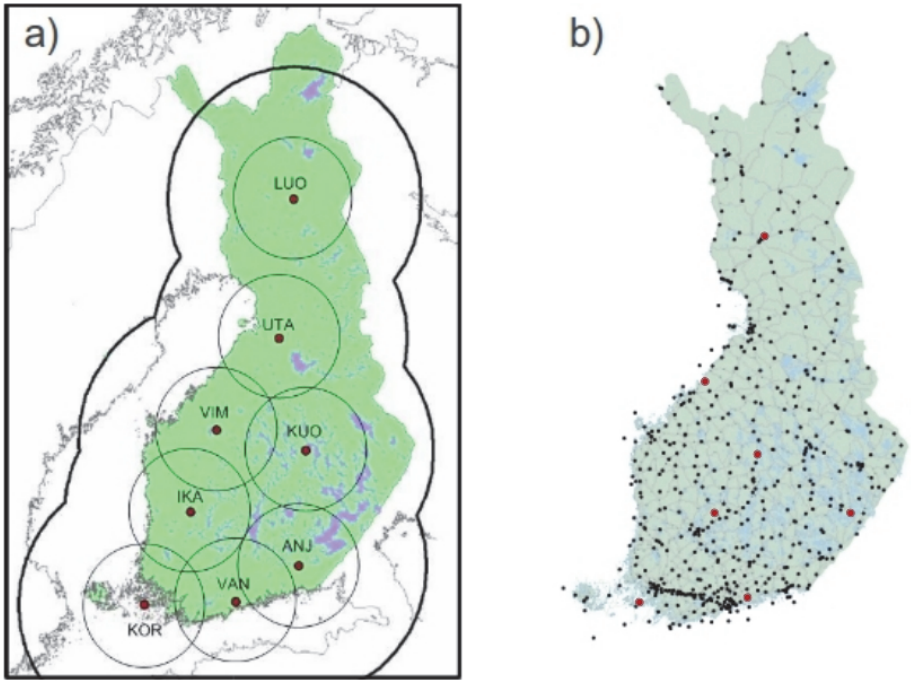

Fig. 1. (a) The rectangular frame of the map depicts the LAPS analysis domain. The red dots represent the 8 Finnish radar stations and the thick, black curved lines display their coverage. The thin circles surrounding each radars represent the areas where measurements are performed below 2 km height. **(b)** The Finnish surface gauge network (dots on the map) used to measure precipitation accumulation. The red dots indicate the position of the seven "independent" stations used for the verification.

HESSD
10, 2453–2480, 2013

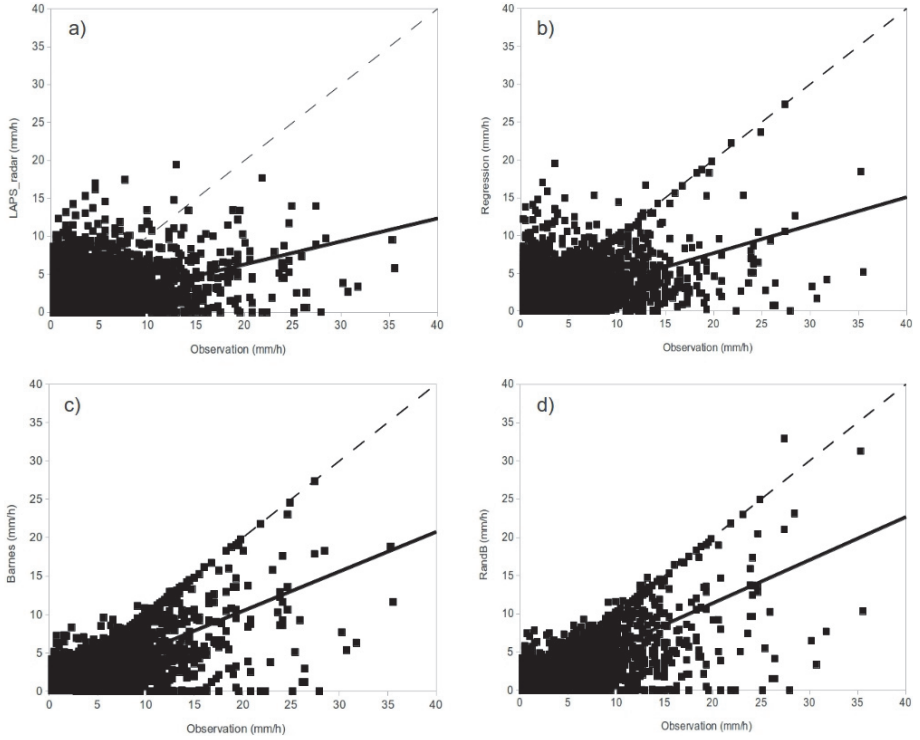

Fig. 2. Scatter plots of analyzed precipitation accumulation (y-axis) against observed rain-gauge values (x-axis) for the dependent stations: **(a)** LAPS_radar versus observations; **(b)** regression vs. observations; **(c)** Barnes vs. observations, and **(d)** RandB vs. observations. The continuous line is a linear fit to the dataset and the dashed line represents the perfect 1 : 1 fit in the plots.

HESSD
10, 2453–2480, 2013

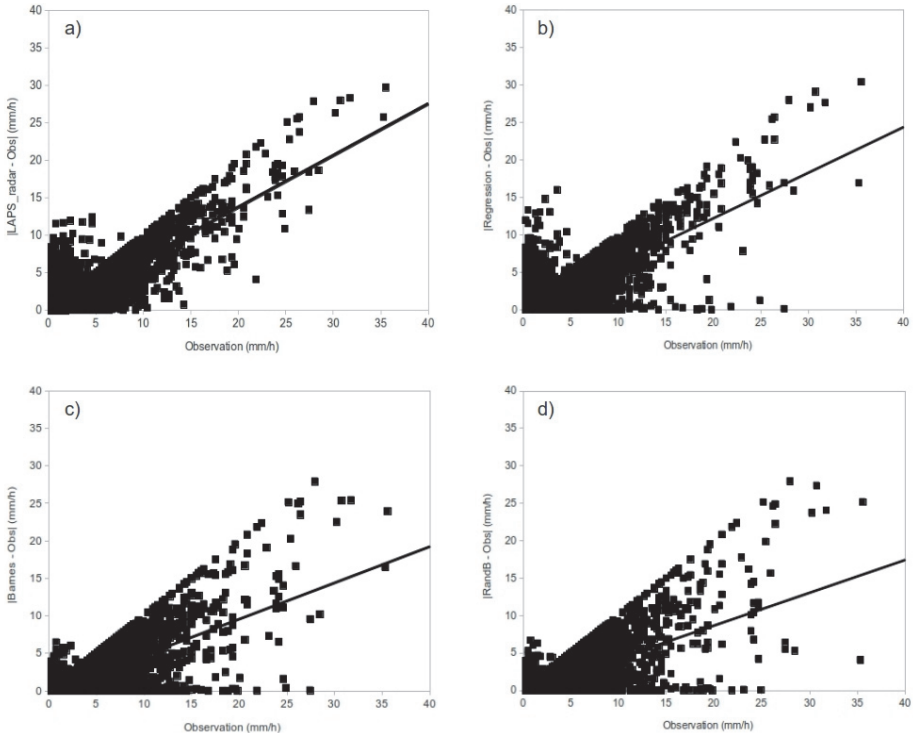

Fig. 3. Absolute value of the difference between observed and analyzed precipitation accumulation (y-axis) plotted against observed rain-gauge values (x-axis) for the dependent stations: **(a)** |LAPS_radar-Obs| versus observations; **(b)** |Regression-Obs| vs. observations; **(c)** |Barnes-Obs| vs. observations, and **(d)** |RandB-Obs| vs. observations. The continuous line is a linear fit to the dataset.

HESSD
10, 2453–2480, 2013

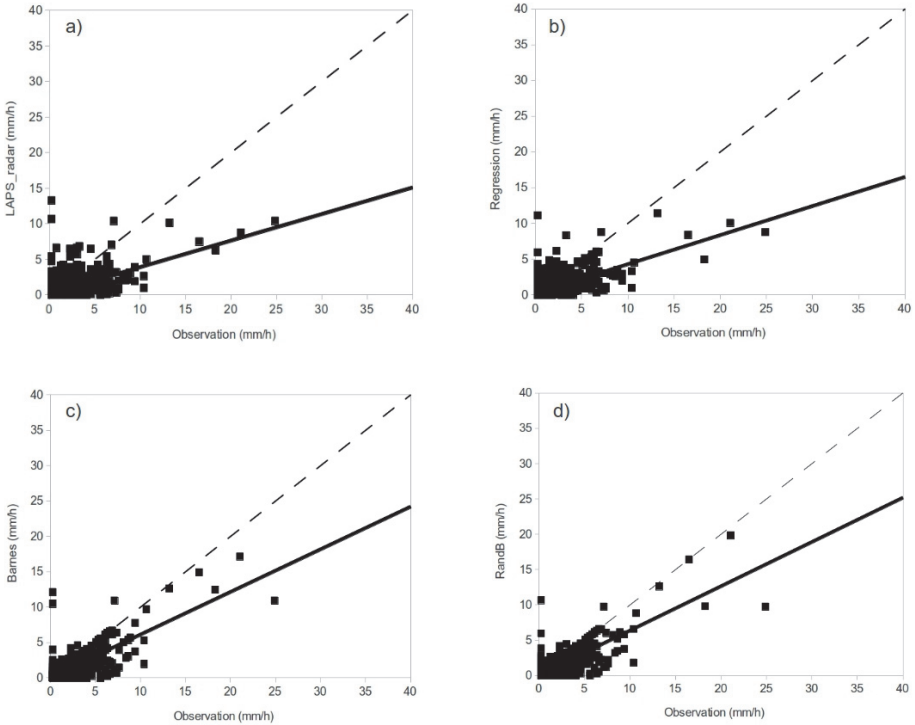

Fig. 4. Scatter plots of analyzed precipitation accumulation (y-axis) against observations (x-axis) for 7 independent stations: **(a)** LAPS_radar versus observations; **(b)** regression vs. observations; **(c)** Barnes vs. observations, and **(d)** RandB vs. observations. The continuous line is a linear fit to the dataset and the dashed line represents the perfect 1 : 1 fit in the plots.

HESSD
10, 2453–2480, 2013

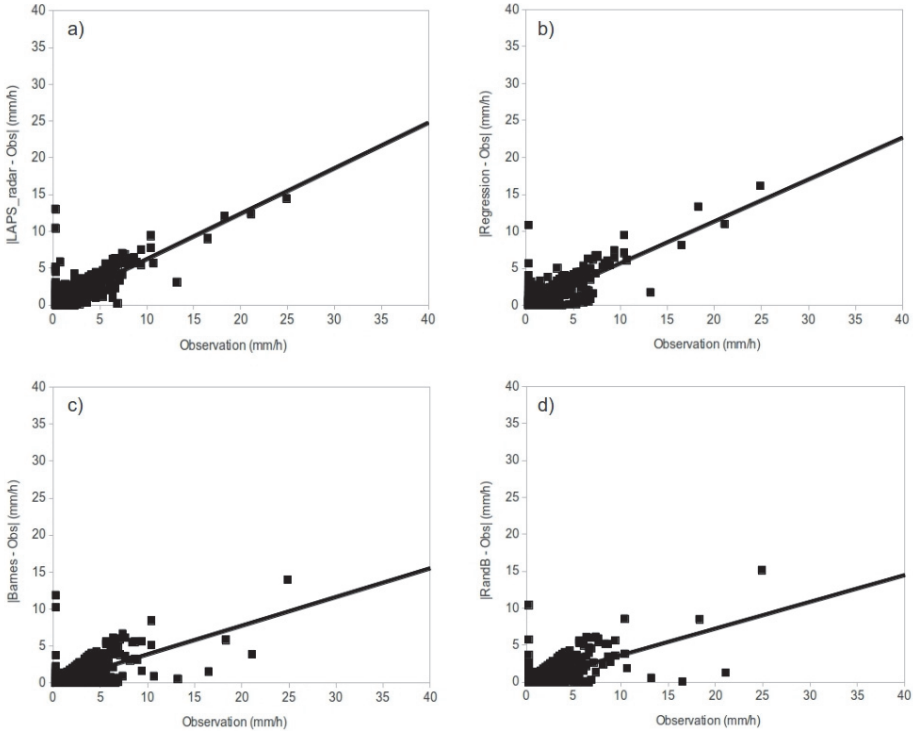

Fig. 5. Absolute value of the difference between observed and analyzed precipitation accumulation (y-axis) plotted against observations (x-axis) for 7 independent stations: **(a)** |LAPS_radar-Obs| versus observations; **(b)** |Regression-Obs| vs. observations; **(c)** |Barnes-Obs| vs. observations, and **(d)** |RandB-Obs| vs. observations. The continuous line is a linear fit to the dataset.

HESSD
10, 2453–2480, 2013

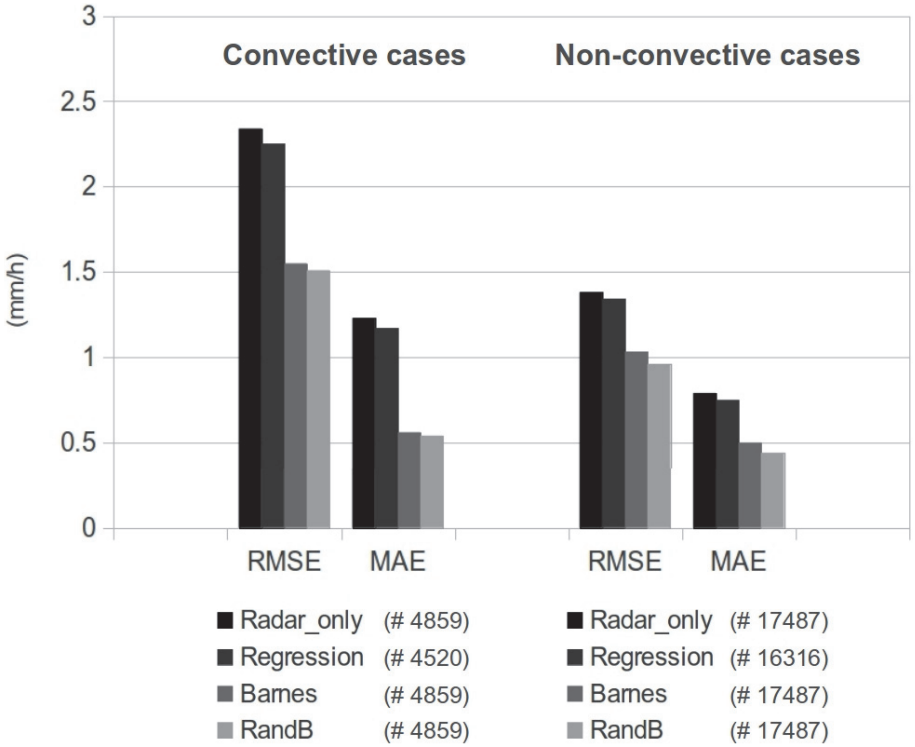

Fig. 6. Statistical verification results for the four different accumulation methods split into two different air-mass stability situations; left panel: convective cases (i.e. thunderstorms) and right panel: non-convective cases (i.e. warm-fronts). The symbol # indicates the number of observations used in the calculations.

HESSD
10, 2453–2480, 2013