

The New French Operational Radar Rainfall Product. Part II: Validation

P. TABARY, J. DESPLATS, K. DO KHAC, F. EIDELMAN, C. GUEGUEN, AND J.-C. HEINRICH

Centre de Météorologie Radar, Direction des Systèmes d'Observation, Météo-France, Trappes, France

(Manuscript received 26 October 2005, in final form 27 September 2006)

ABSTRACT

A new operational radar-based rainfall product has been developed at Météo-France and is currently being deployed within the French operational network. The new quantitative precipitation estimation (QPE) product is based entirely on radar data and includes a series of modules aimed at correcting for ground clutter, partial beam blocking, and vertical profile of reflectivity (VPR) effects, as well as the nonsimultaneity of radar measurements. The surface rainfall estimation is computed as a weighted mean of the corrected tilts. In addition to the final QPE, a map of quality indexes is systematically generated. This paper is devoted to the validation of the new radar QPE. The VPR identification module has been specifically validated by analyzing 489 precipitation events observed over 1 yr by a representative eight-radar subset of the network. The conceptual model of VPR used in the QPE processing chain is shown to be relevant. A climatology of the three shape parameters of the conceptual VPR (brightband peak, brightband thickness, and upper-level decreasing rate) is established and the radar-derived freezing-level heights are shown to be in good agreement with radiosonde data. A total of 27 precipitation events observed by three S-band radars of the network during the winter of 2005 and the autumns of 2002 and 2003 are used to compare the new radar QPE to the old one. Results are stratified according to the distance to the radar and according to the height of the freezing level. The Nash criterion is increased from 0.23 to 0.62 at close range (below 50 km) and from 0.35 to 0.42 at long range (between 100 and 150 km). The relevance of the proposed quality indexes is assessed by examining their statistical relationship with long-term radar–rain gauge statistics. Mosaics of QPE and quality indexes are also illustrated.

1. Introduction

In a companion paper, Tabary (2007, hereafter Part I) proposed a new conventional radar quantitative precipitation estimation (QPE) product for the French operational radar network. The new radar QPE is entirely based on radar data. Following many works published previously (e.g., Joss and Lee 1995; Fulton et al. 1998; Anagnostou and Krajewski 1999; Dinku et al. 2002), the values are obtained as a weighted linear combination of all available measurements at the vertical of any surface pixel:

$$R_{\text{best}} = \sum \omega_i R_i^{\text{corrected}} / \sum \omega_i, \quad (1)$$

where $R_i^{\text{corrected}}$ is the rainfall rate of the i th tilt that has been corrected for VPR effects, partial beam blocking,

and advection. The various modules are commented upon in Part I. As the new QPE product is intended to be deployed operationally within the network, special emphasis is paid to the robustness and simplicity of the correction procedures [as argued by Joss and Lee (1995) and Fulton et al. (1998)]. The quality index of each contributing measurement is expressed empirically as

$$\omega_i = \delta(GC_i) \exp[-(h_i - h_{\text{terrain}})/h_0] (1 - T_i/100)^{1/b} \delta(T_i < 70\%), \quad (2)$$

where $\delta(GC_i)$ is equal to zero if the pixel was classified as ground clutter (1 otherwise), h_i is the altitude of the beam, h_{terrain} is the altitude of the underlying terrain, h_0 is equal to 1 km, T_i is the reflectivity partial beam blocking (in %), b is the exponent of the Z – R relationship ($b = 1.6$), and $\delta(T_i < 70\%)$ is equal to 0 if the occultation rate is beyond 70% (1 otherwise). This empirical formulation is justified in Part I. In the current version, the only factor that can make the quality index

Corresponding author address: Pierre Tabary, Centre de Météorologie Radar, Direction des Systèmes d'Observation, Météo-France, 7, rue Teisserenc-de-Bort, 78195 Trappes, France.
E-mail: pierre.tabary@meteo.fr

vary with time is the occurrence of ground clutter. Indeed, depending on the propagation conditions, any pixel can be dynamically classified either as ground clutter or rain. The quality index is set to zero when the pixel is identified as ground clutter, which means that it will not be taken into account in the surface rainfall estimation.

The best surface estimation of the rainfall accumulation R_{best} goes along naturally with a quality index:

$$\omega_{\text{best}} = \text{MAX}_i(\omega_i), \quad (3)$$

that is, the maximum weight that entered into the weighted linear combination.

The present paper is entirely devoted to the evaluation of this new radar QPE. Because VPR effects are recognized as a major source of error in radar rainfall estimation, especially during the winter season (Zawadzki 1984), section 2 is entirely devoted to the specific evaluation of the VPR identification module. The old and new QPEs are compared qualitatively and quantitatively in section 3. Section 3 also includes a test of the relevance of the proposed empirical quality indexes as well as examples of mosaics of QPEs and quality indexes. Conclusions and perspectives are presented in section 4.

2. Validation of the VPR identification module

a. Database and methodology

The VPR identification module described in Part I relies on VPR candidates and hourly rainfall accumulation ratio curves, as initially proposed by Andrieu and Creutin (1995) and Andrieu et al. (1995). The VPR candidates are obtained by using a conceptual model with four parameters (see Part I, Fig. 5): brightband peak (BBP, expressed in terms of a ratio of rainfall rates), brightband thickness (BBT, expressed in m), freezing-level height (FLH, expressed in m), and upper-level reflectivity decreasing rate (DR, expressed in dB km⁻¹). This modelization is close to what Kitchen et al. (1994) and Kitchen (1997) proposed. The principle of the method consists of letting those four parameters take a limited number of predetermined, climatology-based values. The size of the family, and hence the computation time, increases very rapidly with the number of values tested for each parameter. As the method is used in an operational environment, it is necessary to limit as much as possible the ranges of variation of the four parameters. The purpose of this section is basically to address two questions: Is the proposed four-parameter conceptual VPR model relevant when accounting for the variability of the vertical structure of

TABLE 1. Description of the database used to validate the VPR module.

Station	No. of days processed	First day	Last day
Arcis/Aube	73	26 Nov 2003	10 Oct 2004
Bordeaux	57	14 Nov 2003	13 Oct 2004
Collobrières	21	28 Dec 2003	13 Oct 2004
Grèzes	53	19 Dec 2003	13 Oct 2004
Opoul	50	26 Oct 2003	13 Oct 2004
Plabennec	90	6 Nov 2003	13 Oct 2004
Sembadel	66	14 Nov 2003	11 Oct 2004
Toulouse	79	22 Oct 2003	12 Oct 2004
All radars	489		

precipitation? What is the midlatitude climatological range of variations of each shape parameter? To answer those two questions, 489 rainfall episodes observed by eight radars of the network over a 1-yr period have been analyzed. Table 1 describes the VPR validation database. The technical characteristics and locations of the eight radars are given in Table 1 and also in Part I, Fig. 1) The selection of those eight radars was dictated by several considerations: that the scan strategy of the radar should at least comprise two elevation angles and that it be stable over the study period (October 2003–October 2004), that it have no particular problem regarding the antenna pointing precision has been reported, etc. On the other hand, the VPR identification module had to be tested on radars with different antennas, wavelengths, and physical environments. Overall, from a geographical and technical point of view, the French radar network is quite well represented with this eight-radar subset.

For each radar, the list of rainy days in the study period has been obtained subjectively by examining carefully the 24-h rainfall accumulation plan position indicators (PPIs) produced routinely for all radars within the network. Only the most significant episodes were kept for each radar. Some episodes were removed because of the presence of evident artifacts such as sea clutter, clear air, abnormal propagation, etc. This was especially important for the Collobrières and Opoul sites, which are located on the coast of the Mediterranean Sea and are frequently strongly affected by sea clutter (hence the reduced number of days for those two radars).

The VPR identification used in this paper differs by different aspects with respect to the operational method described in Part I. First of all, unlike the process in operations, no model-derived first guess of the freezing-level height (FLH) is used here. FLH is here allowed to take any value between 0.1 and 4.3 km with a 0.2-km spacing, which yields a total of 22 possible

values. For obvious computational time reasons, it would not be possible to proceed in this way in operations. Another reason for doing this is to allow for a fair comparison between radar-derived FLH and radiosounding data (section 2b). Second, the number of possible values of the three shape parameters of the conceptual VPR is much higher in the present study than in the operational version. Indeed, here are their ranges of variations:

- BBP: 1, 1.5, 2., 2.5, 3., 3.5, 4, 4.5, and 5;
- BBT: 20, 200, 400, 600, 800, and 1000 m; and
- DR: 0, -1.5, -3, -4.5, -6, -7.5, -9, -10.5, and -12 dB km⁻¹.

As one aim of this paper is to document the climatology of the shape parameters, the range of variation of each parameter should be as large as possible in order to cover extreme values. Notice in particular the possibility of an ultrathin bright band (20 m), which has already been reported in the literature (Fabry and Zawadzki 1995). Recall also that BBP is expressed in this method in terms of a ratio of millimeters per hour. A BBP equal to 5 corresponds to a reflectivity enhancement of 11 dB, which is about the maximum documented in the literature (see Part I, Table 2).

The last difference between the VPR identification method used here and the operational one lies in the data that are used as inputs of the algorithm: 24-h rainfall accumulation PPIs are used here, instead of running hourly accumulation PPIs. This was solely motivated by practical considerations: performing the operational version over 1 yr of data for eight radars would have meant processing a huge number of 5-min images. The main drawback of that simplification is that the derived conceptual VPR only represents the whole 24-h period in an average sense. Evolutions of the VPR during the 24-h period are not accounted for. If there is a mixture of stratiform and convective rain within the 24-h period, then the identified VPR will be a composite of a convective and a stratiform profile. As convective profiles are usually close to constant with height, one can expect the composite profile to be a stratiform profile with an attenuated brightband peak. Similarly, if the FLH changes during the day, which happens in the case of frontal systems, then the identified profile will correspond to an average FLH, or, to be more precise, to the FLH of the subperiod when most precipitation fell. An additional expected consequence of any change of FLH during the 24-h period is that the brightband thickness of the 24-h VPR is likely to be positively biased with respect to all hourly VPRs. We recognize these limitations but it should be noted that the same arguments also apply—from a spatial point of view—to

the operational method, as the VPR is assumed to be spatially uniform. Furthermore, 0°C isotherm heights inferred from the 0000 and 1200 UTC radiosoundings were systematically compared and little difference was found on average between the two values.

Overall, given the above-mentioned possible values retained for the FLH, DR, BBP, and BBT parameters, 10 692 VPRs are tested for each rainy day. Depending on the radar scanning strategy, one or three simulated ratio curves can be computed for each rainy day. The criteria and thresholds used to obtain the observed ratio curves are exactly the same as the ones given in Part I. Likewise, the optimal VPR is chosen as the one leading to the smallest difference between the observed and simulated ratio curves.

Figure 1 shows a sample of observed (thin lines) and simulated (thick lines) ratio curves and identified VPRs for 5 days from the Sembadel radar, which is located in the eastern-central France. As the VCP of Sembadel comprises three elevation angles, three ratio curves can be computed. The identified VPR and its characteristics are indicated on the right-hand side of each set of ratio curves. The agreement between observed and simulated ratio curves is quite good and demonstrates, at least qualitatively, the ability of the conceptual model to account for the vertical structure of precipitation. All identified profiles for those 5 days present a bright band, which is not surprising considering the time of the year (November–December). However, it has to be noted that the brightband characteristics are quite variable over time: the brightband thickness varies between 20 m (labeled as 0.0 km in Fig. 1) and 1 km, and the brightband peak between 2.5 and 5. The tremendous variability of stratiform profiles has been evidenced in numerous climatologies (Novak and Kracmar 2001; Vignal and Krajewski 2001; Fabry and Zawadzki 1995).

A careful examination of the ratio curves displayed in Fig. 1 reveals a slight undulation of the observed curve around the simulated one beyond the brightband bump. This oscillation is probably due to the fact that the upper part of the actual profile, above the freezing-level height, is made of two sections: one that is \approx 1000 m thick (i.e., between 0° and -6°C) and has a moderate decreasing rate and another one above that has a larger decreasing rate. This is consistent with the findings of Fabry and Zawadzki (1995) and the modelization of Kitchen (1997). The simple conceptual model of VPR that is used in the present paper assumes a constant reflectivity decreasing rate above the freezing-level height, so its upper part should be considered as a best fit for the upper part of the actual profile.

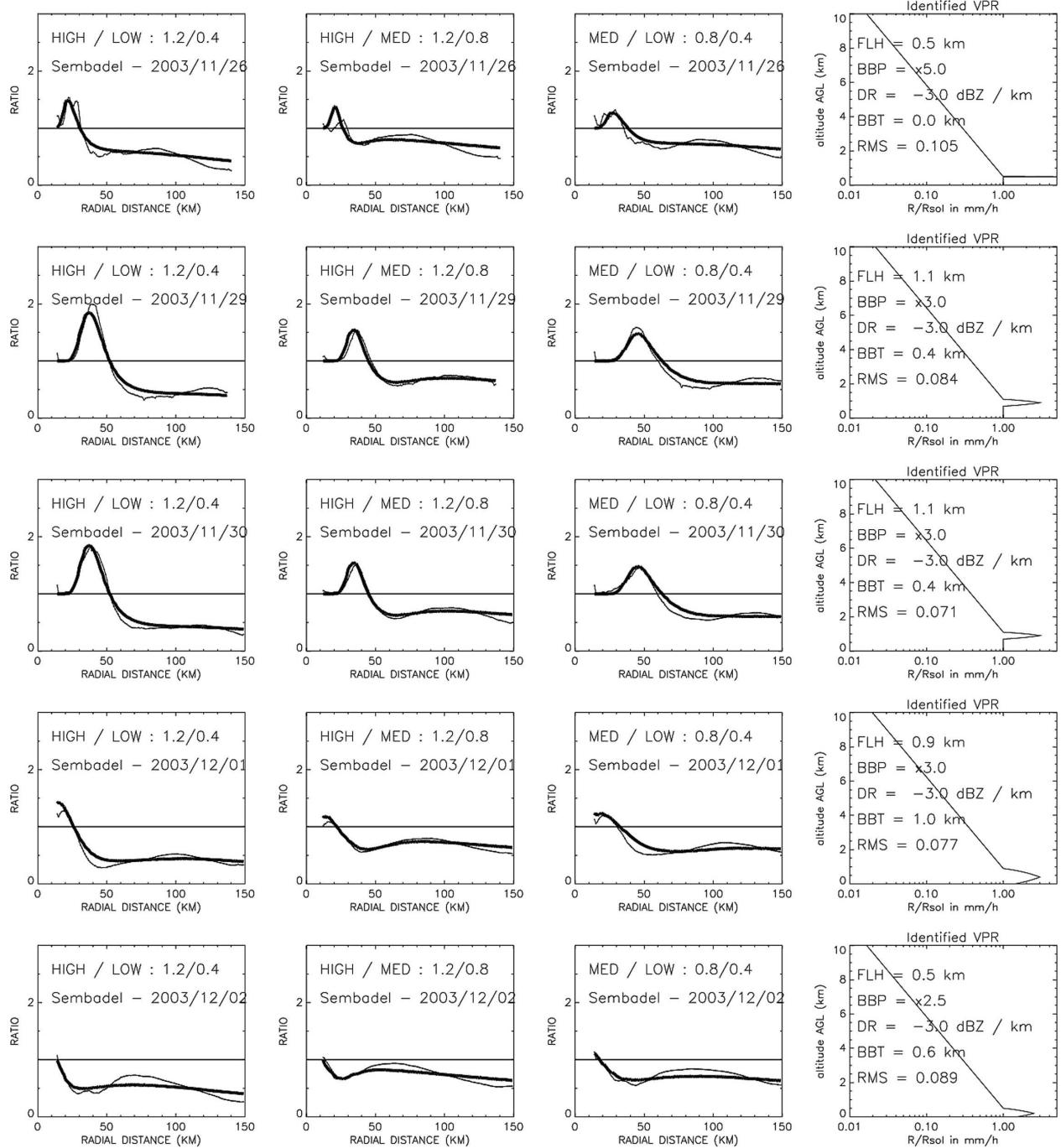


FIG. 1. Examples of observed and simulated ratio curves and identified VPR for five rainy days from the Sembadel radar. The thick line (thin line) corresponds to the simulated (observed) ratio curve. The three elevation angles from the Sembadel radar and the dates are indicated in the ratio curves plots. The identified VPR together with the values of the four parameters (DR, FLH, BBT, and BBP) are shown to the right of the ratio curves plots. Notice that the x axis is logarithmic for the VPR plots.

b. Results

A key question that needs to be answered is, Is the simple conceptual model of VPR able to reproduce the actual variations of reflectivity along the vertical? In

the present approach, the actual vertical variability of the reflectivity field is given by the rainfall accumulation ratio curves. The question then becomes, Are the simulated ratios in good agreement with the observed ratios? Figure 2 allows us to answer that question in a

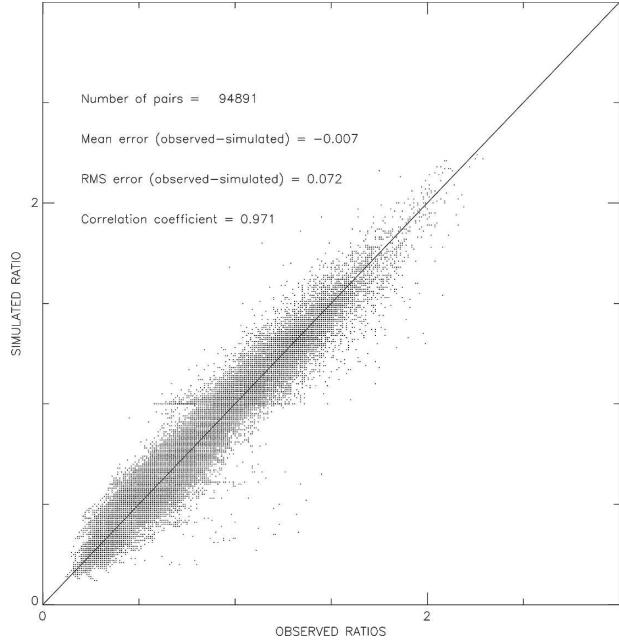


FIG. 2. Overall scatterplot of the whole set of simulated vs observed rain-rate ratios for the eight radars of the analysis. All available ratios of all days of the database have been put together. The total number of points, correlation coefficient, RMSE, mean error, and std dev are given in the graph.

more quantitative fashion than does Fig. 1. Figure 2 is a scatterplot of all of the simulated versus observed ratios. All (observed, simulated) ratios of all ratio curves of all rainy days of all radars have been plotted. The total number of points is about 100 000. The global scores, indicated in Fig. 2, are excellent: the correlation is equal to 0.97, the root-mean-square error (RMSE) is 0.07, and there is no bias. Considering the fact that the huge majority of the ratios are between 0.5 and 2, an RMSE of 0.07 means that the relative error is very small. As a conclusion, despite its simplicity and the fact that the decreasing rate above the FLH is assumed to be constant, the proposed conceptual VPR appears to be quite able to reproduce a large variety of observed vertical reflectivity profiles at the daily time step. Even though the variability of VPR is expected to be slightly higher at the hourly time step, the conceptual VPR family should still be well suited to fitting the actual profiles.

The histograms of the three shape parameters (BBP, BBT, and DR) are presented for all radars in Fig. 3. Each histogram is normalized by the total number of observations. The number of days that entered into the computation of the histogram is indicated in each graph. To create the brightband thickness histogram, only stratiform days were considered, that is, days with a brightband peak strictly larger than 1. The BBP his-

togram (Fig. 3a) has a rather large distribution, its mean being around 2.5. The statistical occurrence of convection (BBP = 1) is around 10%. The brightband thickness histogram, presented in Fig. 3b, reaches its maximum at around 200–400 m. There is however a non-negligible percentage of thick (15%) and ultrathin (7%) bright bands. While such very thin bright bands have already been inferred from vertically pointing radars (Fabry and Zawadzki 1995), they are almost impossible to observe with a scanning radar unless the antenna pattern is taken into account in the VPR retrieval. The decreasing rate histogram (Fig. 3c) is much more Gaussian than the previous ones. There is a well-defined maximum located between -1.5 and -4.5 dB km^{-1} . No attempt was made to stratify the DR histograms according to the type of situation (convective versus stratiform) but it is foreseen that stratiform situations would lead to a steeper decrease of the reflectivity profile above the freezing-level height. Unlike Koistinen (1991), we observe a very small percentage of decreasing rate smaller than -7.5 dB km^{-1} . This may be due to the fact that the freezing-level height rarely decreases below 1000 m in France. Overall, the range of variations of the three shape parameters of the VPR is in quite good agreement with what has been previously reported in the literature (Koistinen 1991; Fabry and Zawadzki 1995; Joss and Lee 1995; Kitchen 1997; Novak and Kracmar 2001; Vignal and Krajewski 2001) and the choice made in the operational version (Part I) to restrict the tested values of BBP, BBT and DR, respectively, to 1, 2, 3, 4, and 5; 200, 400, 600, and 800 m; and -1.5 , -3 , -4.5 , and -6 dB km^{-1} is acceptable. The mean value of the decreasing rate is found to be around -3 dB km^{-1} . This finding questions the default VPR (-1.5 dB km^{-1}) of the operational version. At 250 km, which is the maximum range of operational radar images, the 0.5° elevation tilt is already at an altitude of 6 km. Using the Marshall–Palmer relationship, the rain-rate correction factor corresponding to a VPR decreasing steadily at a rate of -3 dB km^{-1} is then equal to 13. In an operational environment, it can be dangerous to apply such large corrections, even though they may be justified by long-term analyses. Consequently, if a default VPR with a decreasing rate of -3 dB km^{-1} were to be used in operations, then correction factors would have to be bounded.

A systematic comparison of the radar-derived FLH with radiosonde (RS) data has been carried out in order to further validate the VPR retrieval. Only stratiform days are included in the comparison. In addition, some days had to be removed due to the absence of RS data. This leads to a total of 433 days. For each radar, data from the closest radiosounding station were used. The

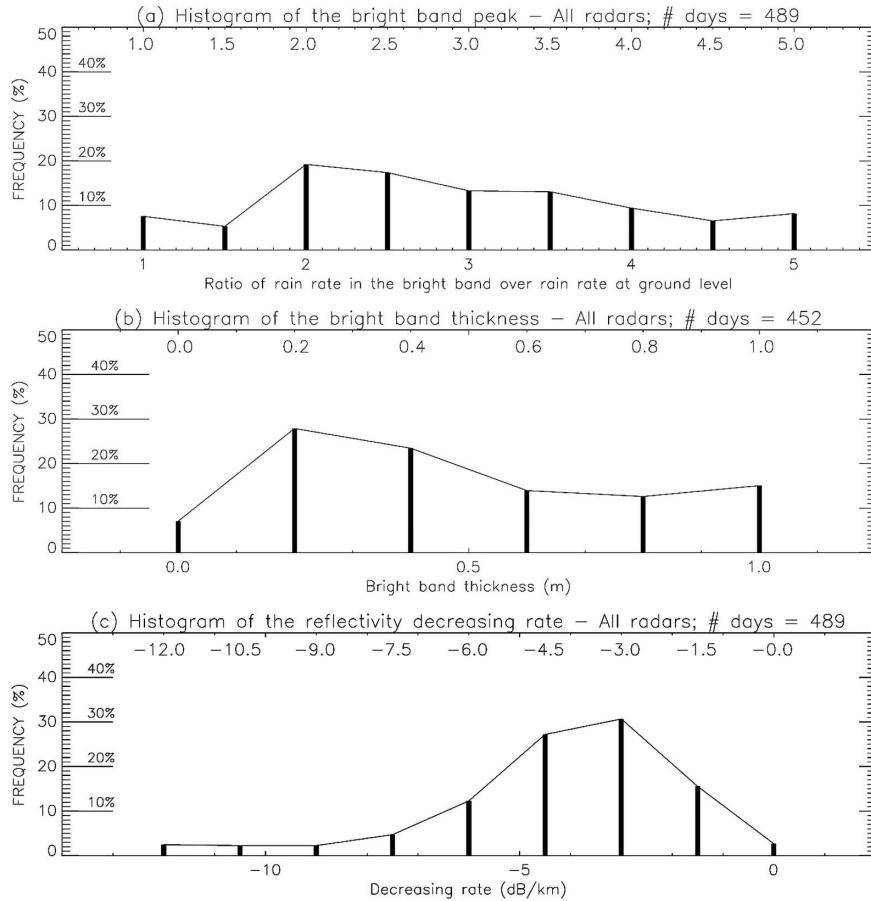


FIG. 3. Climatological distribution of the (a) brightband peak, (b) brightband thickness, and (c) decreasing rate for all radars. The number of days that were used to compute the histogram is indicated in each graph. Recall that the brightband peak stands for the ratio of the rain rate (mm h^{-1}) in the middle of the bright band to that at ground level. Only stratiform days (BBP > 1) were considered to build the BBT histogram. The label “0 km” in the brightband thickness histogram actually refers to an ultrathin bright band of 20-m thickness.

distance between the radar site and the radiosounding station is about 100 km on average over the eight radars but varies from 0 km (Plabennec and Bordeaux) up to 200 km (Opoul and Collobrières). Such large distances may impact the comparison in some situations (e.g., passage of a warm or cold front). However, inspection of model analyses suggests that in most cases the 0°C isotherm field is very smooth. The radar-derived FLHs were systematically compared to the “quasi-instantaneous” 1200 UTC radiosounding. This may be another source of error in some circumstances, as the retrieved VPR is actually an average over the day or, to be more precise, over the period when the rain fell around the radar (which is not necessarily always around noon). Each radar-derived FLH was compared to the dry 0°C isotherm but also, following Giangrande and Ryzhkov (2004), to the wet-bulb 0°C isotherm.

Figure 4 shows the time series of the radar-derived

FLH (thick, solid line) for all eight radars. The envelope of the RS 0°C dry- and wet-bulb temperatures has been superimposed in gray. All time series have a length of about 1-yr (October 2003–October 2004) even though the number of days that compose the series is variable. The annual cycle of the freezing-level height is evident in all graphs. The short-term variability may be also very important. Despite the fact that the eight radars of the study are located in different regions of France and are in some cases very far away from each other, the 0°C isotherms time series are quite comparable. Overall, for all radars put together, the mean 0°C isotherm height for the cold season is around 1500 m above sea level (ASL) and that of the warm season is around 3000 m ASL. The radar-derived FLH is generally very well correlated with the RS 0°C isotherms (both dry and wet bulb), whichever radar is considered. It is however difficult to decide, on the basis of Fig. 4

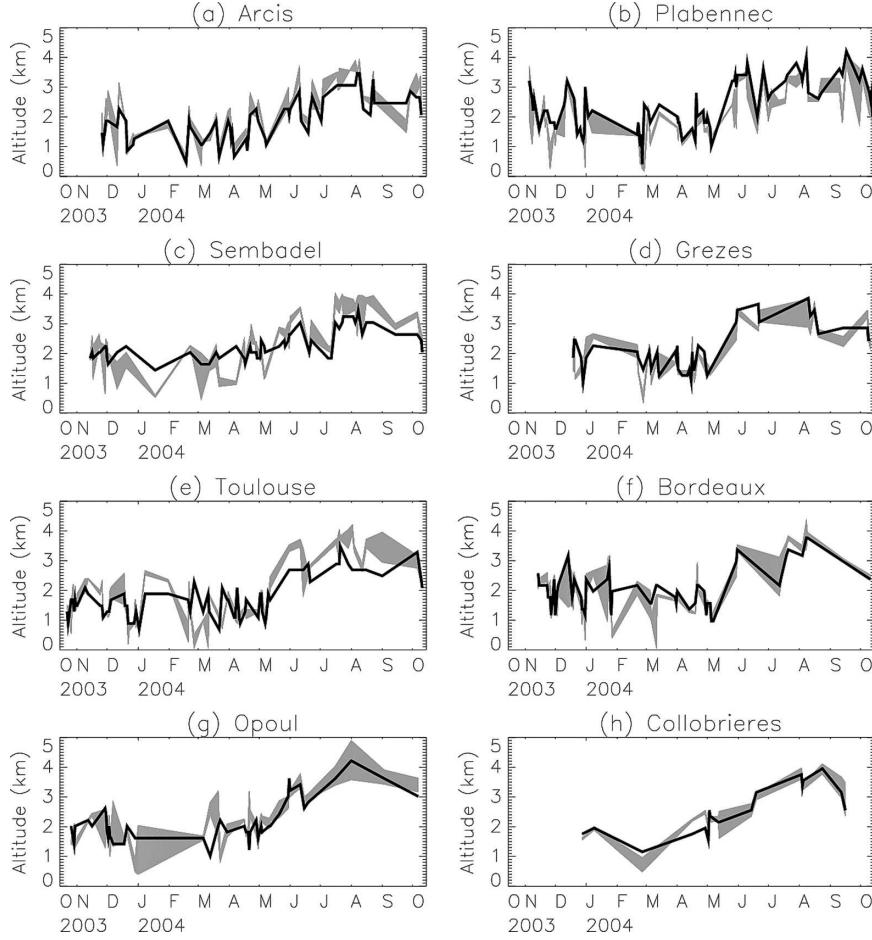


FIG. 4. Time series of the radar-derived FLH for the eight radars of the study. For each radar, the envelope of the 0°C dry- and wet-bulb temperatures deduced from the closest radiosounding station has been superimposed in gray. The radar-derived FLH is included in the comparison if a stratiform profile was identified ($BBP > 1$).

alone, which between the dry- and the wet-bulb 0°C isotherms is in better agreement with the radar observations. The mean error, RMSE, and correlation coefficient were computed for the dry- and wet-bulb cases. A slightly negative bias was found in the dry temperature case (-0.15 km, radar $<$ RS). That bias becomes slightly positive (0.19 km, radar $>$ RS) when wet-bulb temperatures are considered bias. The RMSE is around 500 m in both cases although, like the bias, it is a bit larger in the wet-bulb temperature case. This value of the RMS may seem to be large. One should however keep in mind the fact that the radar and the RS are spaced apart both in space and time. The space-time distance certainly accounts for a significant fraction of the RMSE. Finally, the correlation coefficient is excellent: it is equal to 0.77 in the wet-bulb temperature case and 0.83 in the dry-bulb temperature case. This confirms the qualitative impression given by Fig. 4. Using

model forecasts at a hourly time step and at the precise radar locations may help reducing the RMS error. In the operational version (Part I), only three possible values are tested for the FLH parameter: $FLH_0 - 200$ m, FLH_0 , and $FLH_0 + 200$ m, where FLH_0 is a mesoscale model prediction of the dry 0°C isotherm height at the radar location. We recognize that the range of variation could be increased a little bit (say to $FLH_0 - 400$ m, $FLH_0 + 400$ m) to better account for model errors. This, however, would mean adding more VPR candidates, which in turn would increase the computation time. Additionally, with the forthcoming advent of a new generation of operational nonhydrostatic numerical weather prediction models with frequent analyses (Ducrocq et al. 2005), one can expect the model errors to become smaller and smaller. Mittermaier and Illingworth (2003), using the Met Office Unified Model, found an RMS error between radar-observed and

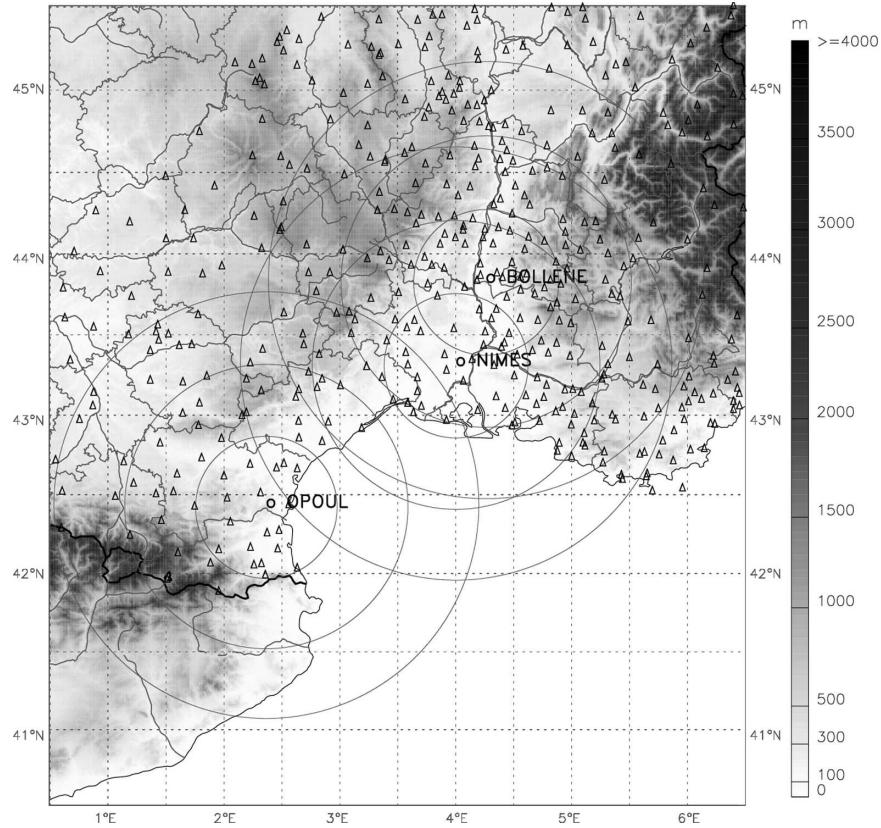


FIG. 5. Map of the validation area showing the orography, the locations of the three radars, and the rain gauge validation network. The 50-, 100-, and 150-km-range circles of the Bollène, Nîmes, and Opoul radars are also represented.

model-derived FLH of 150 m. It should finally be recalled that the data that are used to retrieve the VPR are corrected from partial beam blocking and ground clutter but not for attenuation occurring either in rain or in the bright band. Attenuation may in some particular cases severely impact the quality of the VPR retrieved in operations but the fact that the input data are averaged in time (1 h) and space (360° in azimuth) somehow contributes to mitigating that potential problem. Attenuation correction techniques based on dual polarization will undoubtedly benefit the retrieval quality of VPR.

3. Evaluation against rain gauges

a. Domain, database, and methodology

The benchmark is the old, pseudo-constant altitude plan position indicator (CAPPI) based, radar QPE (see Part I) and the reference rainfall field is provided by the Météo-France operational rain gauge network (Grégoire 2002), the density of which is about one gauge per 1000 km^2 (see Fig. 5). The evaluation is performed at an

hourly time step. Twenty-seven precipitation events taken from winter 2005 and autumns 2002 and 2003 have been selected. It is important to include warm and cold season episodes in the dataset to make the validation as significant as possible. Indeed, the cold season is characterized by more stratiform cases with lower bright bands (1000 m) whereas the warm season is more favorable to convective systems and higher 0°C isotherms (3000 m). The 27 episodes are described in terms of mean and maximum rain gauge accumulations in Fig. 6. They are ranked according to the freezing-level height, which is plotted at the top of the figure. The scale for rainfall accumulation is on the right-hand side of each graph which go up to 600 mm. The curves (with scales shown on the left-hand side of each graph) will be commented upon in section 3c. Three classes of distance have been distinguished: 0–50 (top), 50–100 (middle), and 100–150 km (bottom). Mean rainfall accumulations are represented by dark gray-shaded vertical bars and maximum accumulations by light gray-shaded bars. It is clear that the most significant episodes in terms of rainfall accumulations occurred during the

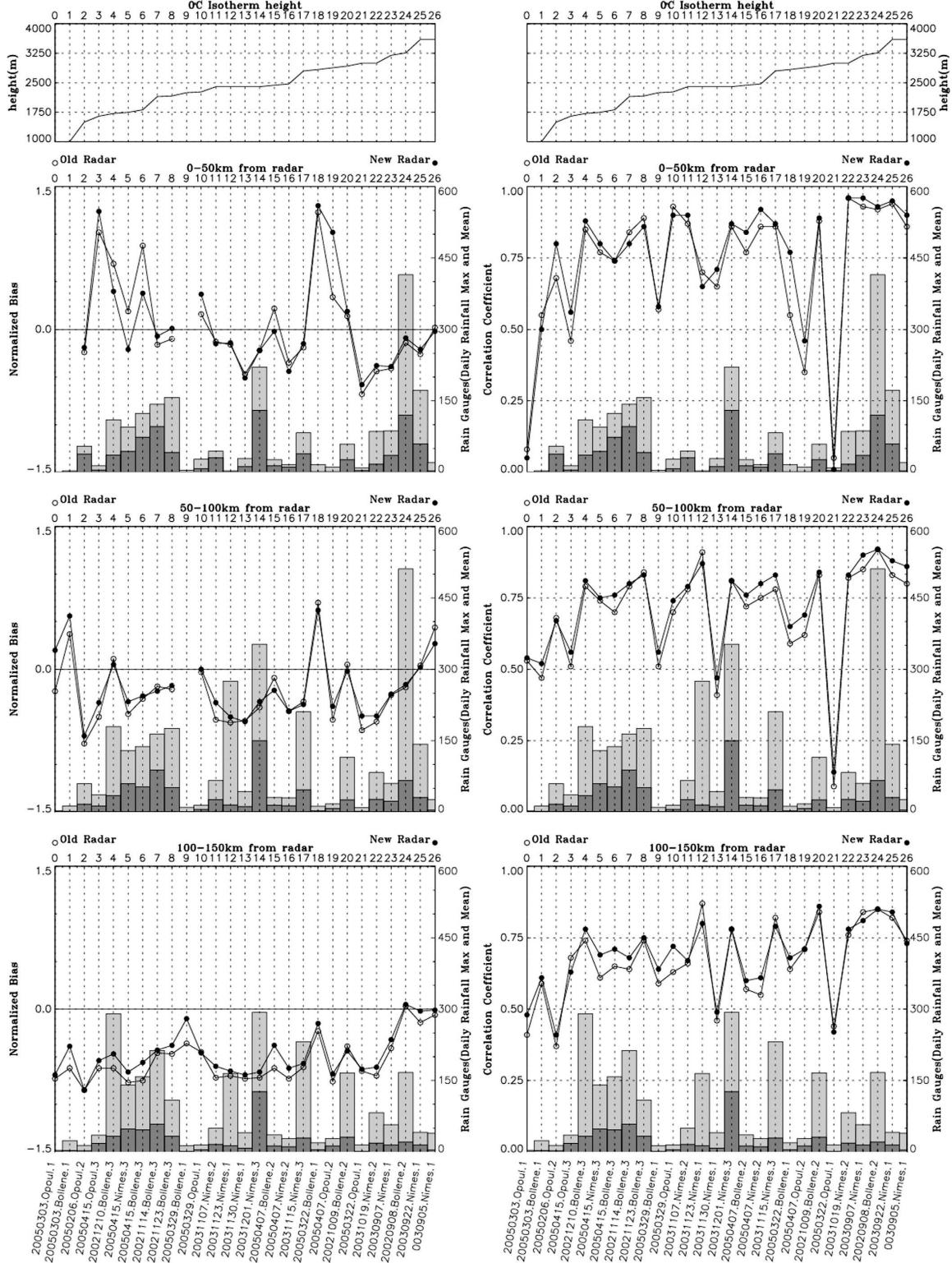


FIG. 6. Systematic comparison of the new and old radar QPEs over the 27 precipitation events. Results are stratified according to distance (0–50 km, top; 50–100 km, middle; 100–150 km, bottom). Two scores are shown: (left column) the Nash coefficient and (right column) the correlation coefficient. The new (old) radar QPE scores are represented by black (white) bullets. The maximum and mean rain gauge accumulations are superimposed on the scores. The FLH is plotted at the top of the two columns.

autumns. The radar data come from three S-band radars (Bollène, Nîmes, and Opoul) located in rather mountainous regions of southern France (see Fig. 5). The advantage of using S-band radars for the validation is that attenuation by rain can reasonably be neglected among the various possible sources of errors. The calibration of the all French operational radars is monitored continuously (Tabary 2003) and no calibration bias has been documented for the study periods.

b. A qualitative comparison between the old and new QPEs

Figure 7 shows the 48-h rainfall accumulation obtained with the old (Fig. 7a) and the new (Fig. 7b) radar QPEs for a catastrophic episode that occurred on 8–9 September 2002. To highlight the differences between the two algorithms, the ratio of accumulations was computed and is presented in Fig. 7d. More than 500 mm of rain were recorded in the area of maximum precipitation. Delrieu et al. (2005) give a detailed description of that case and several research QPE algorithms have been compared on that situation (Nicol et al. 2006, manuscript submitted to *J. Hydrometeor.*; Boudevillain et al. 2006, manuscript submitted to *J. Hydrometeor.*). The radar used to produce the accumulation is the Bollène radar. The range rings are spaced 50 km apart. The size and resolution of each image are 512 km × 512 km and 1 km², respectively. The 48-h accumulation has been obtained by adding up all 5-min QPE maps. The 0°C isotherm height was about 3 km throughout the episode. The black areas in the old QPE map indicate the cluttered region where no value is available. The rain gauge accumulations have been superimposed as small colored circles on the radar rainfall accumulation. The new radar QPE map appears to be much smoother than the old one. The obvious underestimations in the shielded sectors to the east and west of the radar are partially corrected for. The enhancement is clearly evidenced by the ratio map (Fig. 7d). The reason why the correction is not perfect is mainly due to unaccounted VPR effects on partial beam-blocking correction factors and nonorogenic masks such as the ones caused by trees and construction. The estimated rainfall accumulation at long range is also increased. However, because the 0°C isotherm was rather high (3000 m) during that episode, the effect of the VPR correction only becomes noticeable beyond 150 km. Rainfall estimation is available in cluttered areas in the case of the new QPE. Quite logically, at close range, in the absence of shielding or VPR contamination, the two QPEs give the same results. In particular, the maximum rainfall accumulation (500 mm) is measured at the same location and with the same magnitude by both QPE algorithms. Fig-

ure 7c represents the mean quality index map for the episode. Because the dynamic range of the quality indexes is very large (due to the exponential decrease with height), the parameter that is shown is the logarithm of the quality index that was rescaled between 0 (very good quality) and 255 (very bad quality). The quality is essentially distributed concentrically around the radar. This is due to the elevation of the beam with distance. This overall pattern is however modulated by a number of factors: in the close-range cluttered areas (indicated in black in Fig. 7a), the lowest-elevation angles were contaminated and the QPE was obtained from a higher tilt. This explains why the quality decreases in those areas. The signatures of the partial beam blocking to the east and west of the radar are also visible in Fig. 7c, even though they tend to be dominated by the strong range- (height-) dependent feature. Farther away to the east, when the beam passes over the French Alps, the crests and valleys become apparent. This is a direct consequence of the choice to let the quality indexes depend upon the height over the terrain [Eq. (2)]. In the general problem of estimating surface rainfall from radar data, it is clear that the closer the beam is to the ground, the better the surface rainfall estimation is likely to be. In mountainous areas then, it seems logical that the pixels located on the crests are assigned a better quality than those located deep down in the valleys. It is true, though, that the beam–terrain distance is not the only factor affecting the quality of rainfall estimation but at least it is a one that is easy to compute. Mountains, as opposed to plains, cause vertical motions and low-level microphysical processes, which impact the Z–R relationship and the VPR, and, hence, the quality of the radar rainfall estimation. More studies are needed though to quantify those effects and introduce them into the proposed quality indexes.

Figure 8 shows another case (1–3 December 2003) seen by the Nîmes radar. Like before, the old (Fig. 8a) and the new (Fig. 8b) QPEs are presented along with the mean quality index map (Fig. 8c) and their ratio (Fig. 8d). As in the previous case, the new QPE presents fewer discontinuities than the old QPE. In particular, the circular discontinuities caused by changes in the elevation angle have disappeared thanks to the pixel-by-pixel weighted linear combination of all available PPIs implemented in the new QPE algorithm. Mostly stratiform, this episode was characterized by brightband effects (strong circular feature at 50 km in the old QPE map) that are attenuated with the new QPE algorithm. This is evident in the ratio map (Fig. 8d). One must admit that there are still traces of the bright band that are visible. This may be related to the limitations of the VPR module (acknowledged in

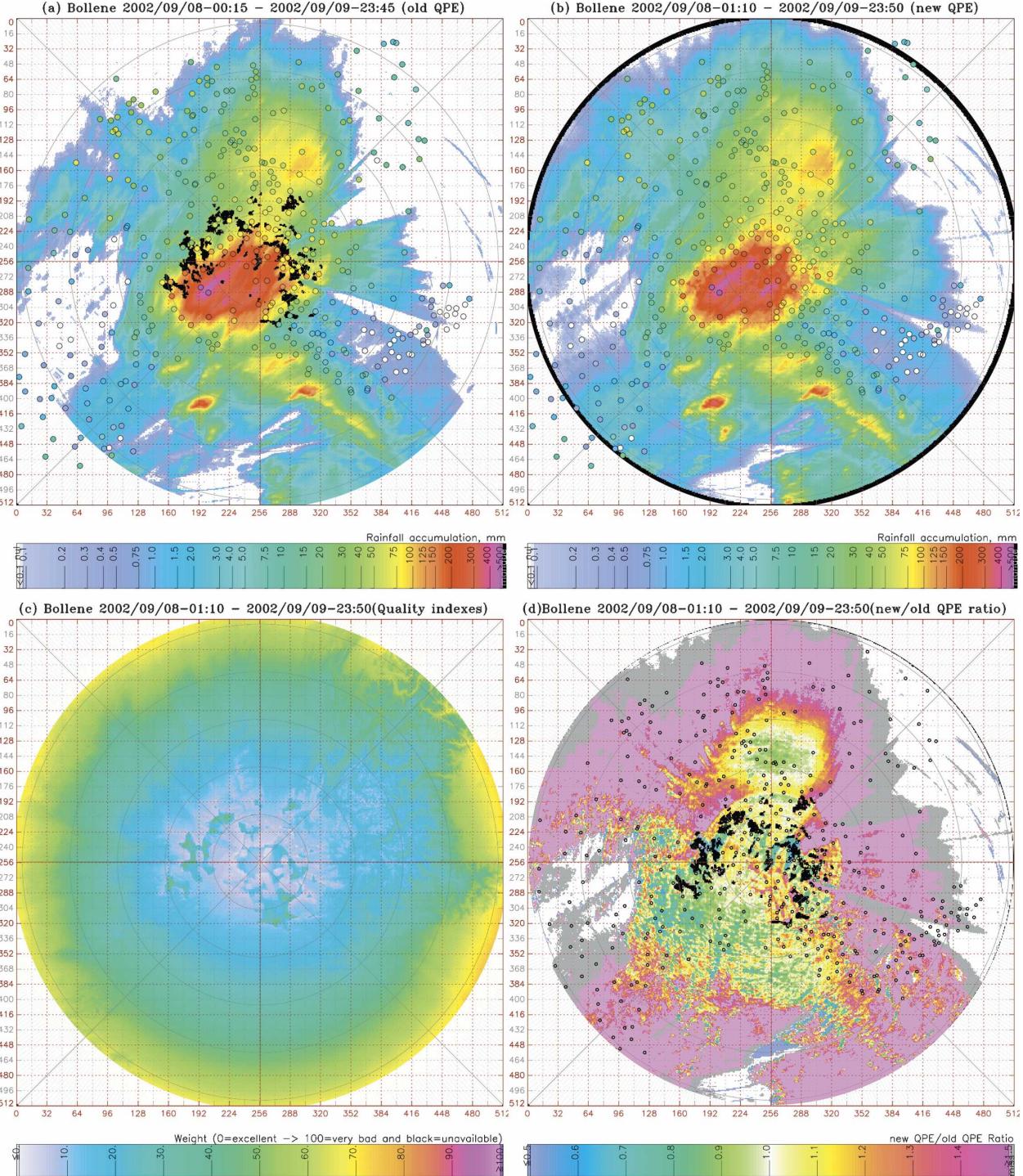


FIG. 7. The 48-h rainfall accumulation (8–9 Sep 2002) obtained with the S-band Bollène radar and with the (a) old and (b) new radar QPE algorithms. (c) The mean quality index map over the accumulation period. (d) The ratio of the rainfall accumulations obtained with the old and the new QPEs. Rain gauge measurements, represented by small colored circles, are overlaid on the QPE maps. The range circles are spaced 50 km apart. The black area in Fig. 6a is the ground-clutter area, where no value is available with the old QPE algorithm. The ratio is not computed if the old QPE accumulation is equal to zero.

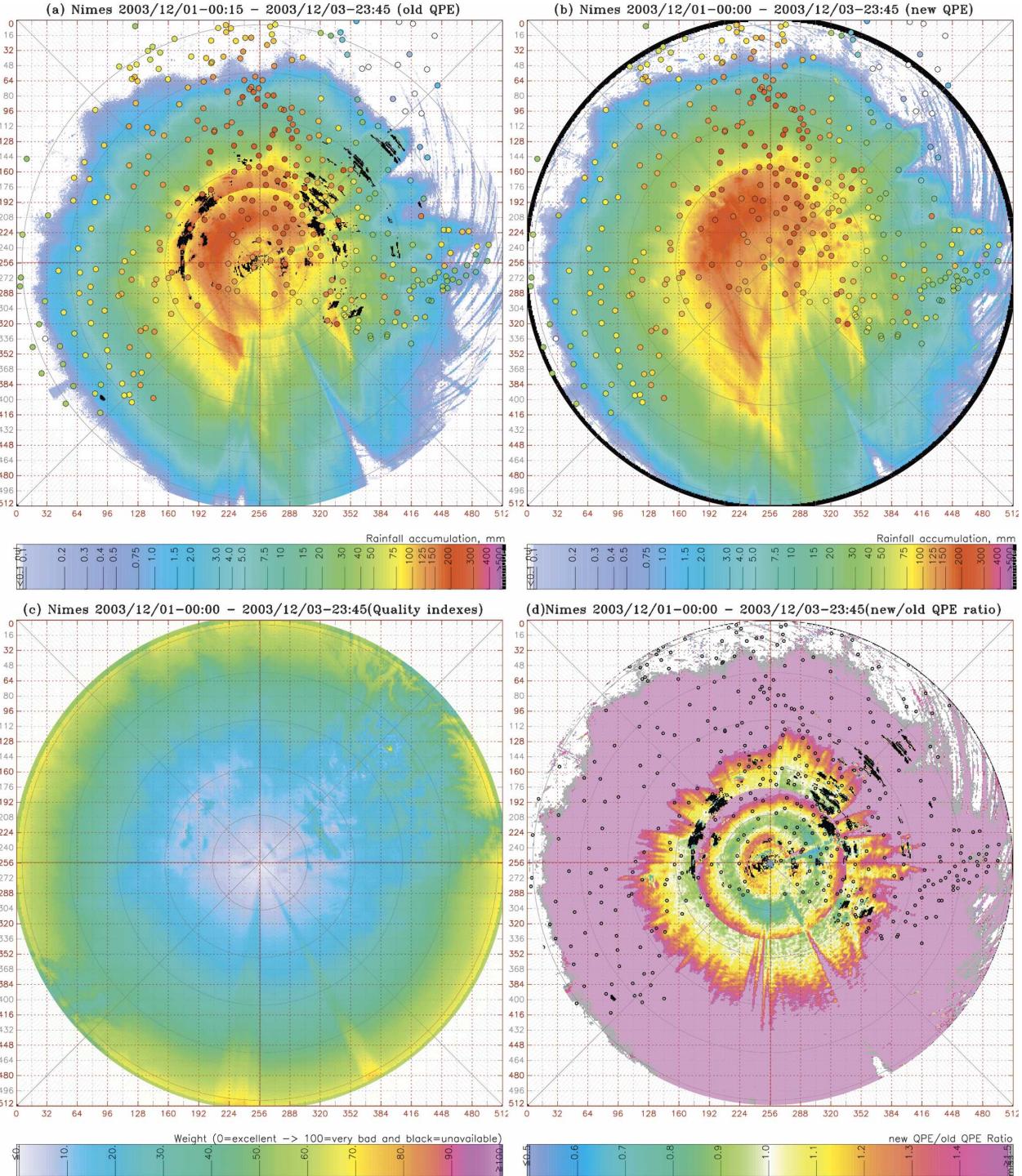


FIG. 8. Same as in Fig. 7 but for a 3-day event (1–3 Dec 2003) observed from the S-band Nîmes radar. Notice the discontinuities in the old QPE map that are due to the static, concentric way of creating the rainfall accumulation map.

Part I): 1) the VPR is assumed to be the same all around the radar and 2) a climatologically decreasing VPR is used when not enough precipitation is observed under the radar umbrella. The narrow shielded sector

to the south-southeast of the radar is caused by a series of trees. This is the reason why it is still apparent in the new QPE product. The new QPE algorithm provides a rainfall estimation in the former cluttered areas (in

black in Fig. 8a), which has been considered to be a major improvement by forecasters when they are presented with the new product during demonstration phases. The correction for underestimation at long range, achieved through the VPR correction, is also noticeable in the new QPE map (Fig. 8b) and in the ratio map (Fig. 8d). The quality index map has the same characteristics as in the previous situation: in addition to the strong-range dependency, the quality index decreases in cluttered areas, shielded sectors, and in the valleys.

The last case is an episode that occurred on 6 February 2005 over the Opoul radar (Fig. 9). The freezing-level height was particularly low (1000 m) throughout the episode and brightband effects caused a well-defined circular ring of rainfall overestimation in the vicinity of the radar. That effect has been corrected for in the new QPE product (Fig. 9b). The ratio map (Fig. 9d) shows that brightband effects were affecting all three elevation of the old pseudo-CAPPI QPE. In addition to reducing brightband effects, the VPR correction significantly enhances the rainfall accumulations beyond 30–40 km. With the Opoul radar being located between two mountain ranges, the Pyrénées to the southwest and the Massif Central to the northeast, the quality index map (Fig. 9c) presents a lot of fluctuations in those two regions.

Before moving on to the quantitative evaluation, it should be said that real-time demonstration phases of the new QPE have been organized during the autumns of 2002 and 2003 and the winter of 2005. The Météo-France forecasters who were presented with the new product unanimously appreciated the availability of rainfall estimation in cluttered areas, the extended coverage area, the smoothness of the product, and the removal of brightband signatures. They also strongly supported the systematic production of quality indexes.

c. Comparative analysis of the scores obtained with the old and new QPEs

Figure 10 shows the overall evaluation over the 27 episodes at an hourly time step. The old QPE scatterplots are displayed on the left and the new QPE scatterplots on the right. The assessment has been conducted separately for three classes of distance: 0–50, 50–100, and 100–150 km. The radar–rain gauge comparison has only been carried out in areas where the old QPE was available, that is, outside the cluttered areas in black in Figs. 7a, 8a, and 9a. The number of radar–rain gauge measurement pairs included in the analysis is indicated in each scatterplot. The normalized bias, normalized error, correlation coefficient, root-mean-square error, and the Nash criterion have been com-

puted (see Wilks 1995) and are indicated in each scatterplot.

The total number of radar–rain gauge pairs (N) is quite large and makes the statistical comparison quite significant. Here, N is close to 75 000 for the first class of distance (0–50 km), 170 000 for the second (50–100 km), and 210 000 for the third (100–150 km). Looking first at the first class of distance (Fig. 10, top), it can be seen that the new QPE overperforms significantly the old one whichever score is considered. For instance, the Nash criterion and the correlation coefficient are increased from 0.23 to 0.62 and from 0.72 to 0.81. The normalized bias is reduced from 0.25 (25% overestimation) to 0.04. The scores are closer in the second class of distance, though still in favor of the new QPE. The correlation and the Nash criterion are increased from 0.80 to 0.83 and from 0.63 to 0.66. The normalized bias is negative in both cases (−0.28). In the third class of distance, the correlation and the Nash criterion are also improved with the new QPE: 0.71 instead of 0.68 for the first case and 0.42 instead of 0.35 for the second. The normalized bias is less negative with the new QPE (−0.54 instead of −0.63) but still remains large. The reduction of the negative bias at long ranges is an expected consequence of VPR corrections (Kitchen et al. 1994; Vignal and Krajewski 2001; Germann and Joss 2002). In the present case, however, a negative bias still exists. There are several possible explanations for this result: one is that attenuation by gases is currently not accounted for. Doviak and Zrnić (1984) give the two-way attenuation by gases for propagation through a standard atmosphere at 10-cm wavelength. At an elevation of 1°, they report values of 0.6 dB at 50 km, 1.2 dB at 100 km, and 1.6 dB at 150 km. Those values would give, using the Marshall–Palmer Z – R relationship, normalized biases equal to −0.08, −0.15, and −0.20. Attenuation by rain or by the melting layer would also impact negatively the radar – rain gauge bias. However, those two potential causes are unlikely since only S-band data were used for the validation. Another explanation might be that the -1.5 dB km^{-1} climatological VPR profile that is used whenever the actual profile could not be obtained (see Part I) is not steep enough. Steeper profiles could be used in the future but that would mean accepting apply very large enhancement factors, which might be risky in an operational context.

The scores obtained at the hourly time step with the new and the old QPEs have been systematically computed for each episode to shed light on the temporal variability of the quality of the new QPE. The normalized bias and the correlation coefficient are plotted in Fig. 6. The 27 episodes have been ranked according to the freezing-level height, which is plotted at top of each

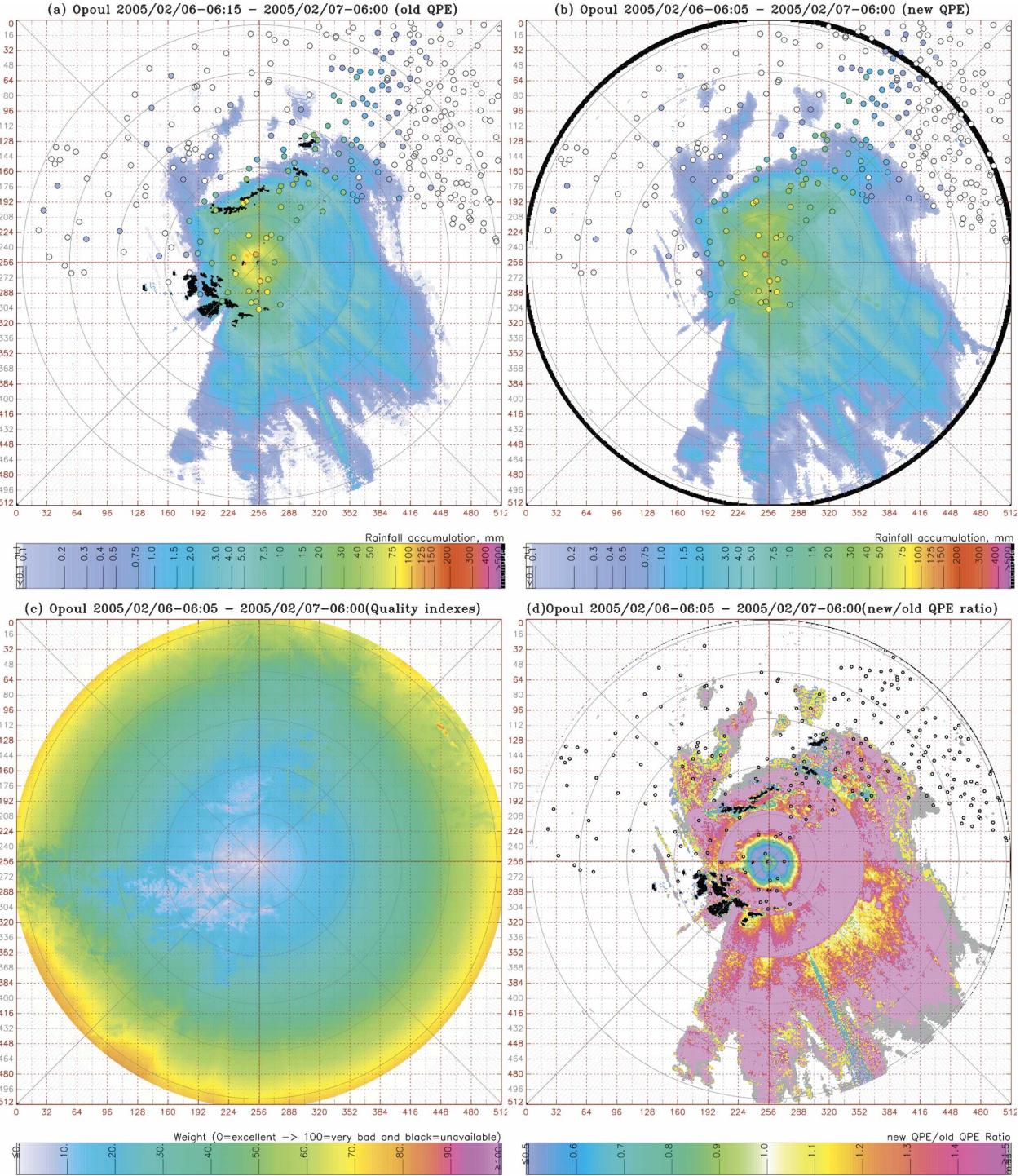


FIG. 9. Same as in Fig. 7 but for a pure brightband case (6 Feb 2005) observed from the S-band Opoul radar.

column in Fig. 6. The white and dark bullets represent the old and the new radar QPEs, respectively. The relevant scale for the scores is on the left-hand side of each graph. The scale for the normalized bias spans is $[-1.5, 1.5]$ and that for the correlation coefficient is $[0, 1]$. The

scores are not available in four cases (episodes 0, 1, and 9 in the first class of distance, and episode 9 in the second class of distance) because the number of radar–rain gauge pairs was too small. In the first class of distance, the scores are improved in the majority of the

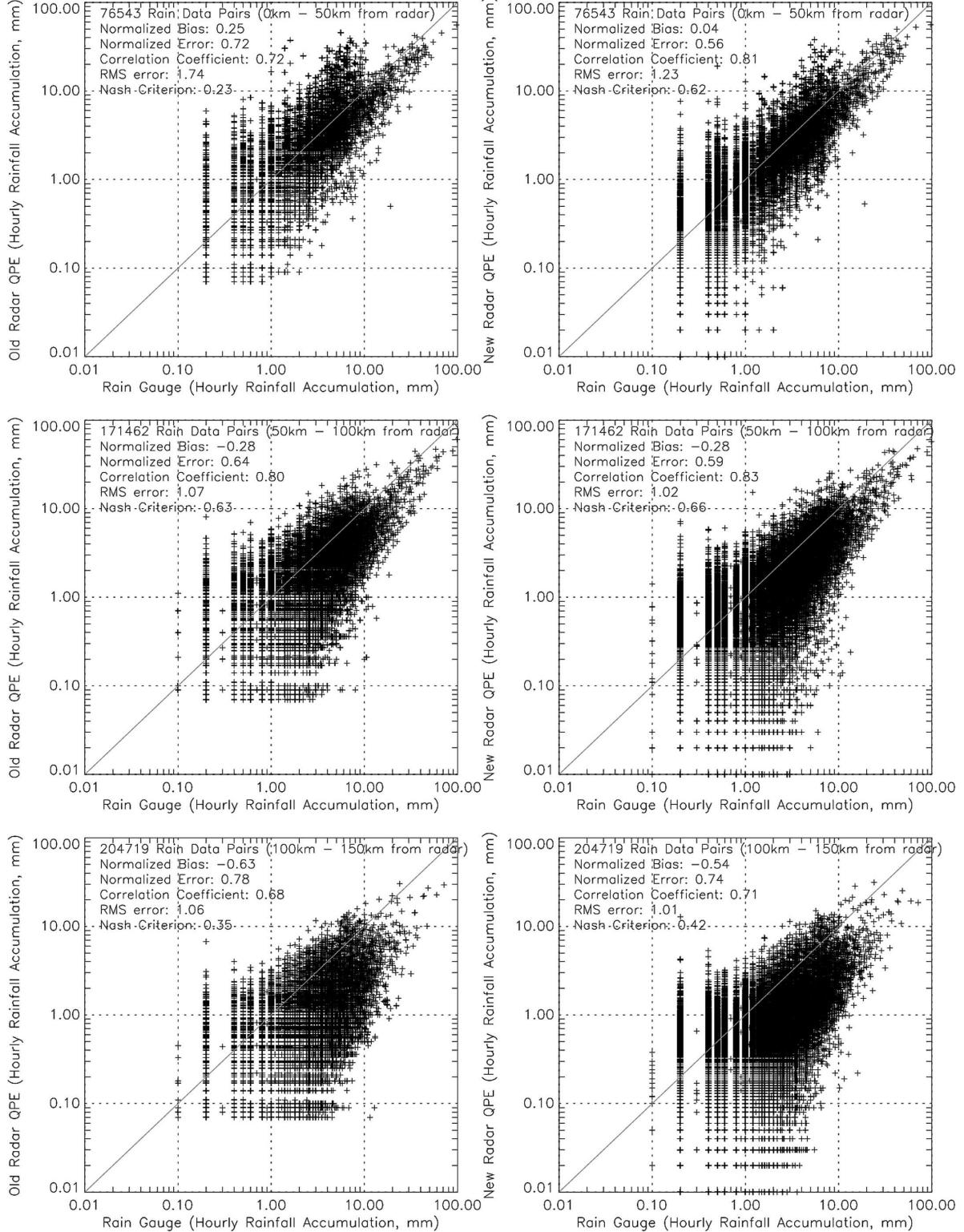


FIG. 10. (left) Hourly rainfall accumulations obtained with the old radar QPE vs hourly rain gauge measurements for distances between (top) 0 and 50 km, (middle) 50 and 100 km, and (bottom) 100 and 150 km. (right) Same as in the left column but with the new radar QPE. All 27 precipitation events of the validation database have been taken into account. The number of (radar–rain gauge) measurements pairs is indicated in each graph as are the values of the standard scores. All rain episodes have been included in this comparison.

TABLE 2. Illustration of the variations of Ω with the beam height and the occultation rate.

Value of the logarithm of the quality index coded over 1 byte (0–255)		Beam height over ground level (km)		
		0	0.5	1
Occultation rate (%)	0	0	13	25
	20	4	17	29
	50	11	24	36
	70	19	32	44
				70

cases. The correlation coefficient is increased in 20 cases out of 27. In several cases of low freezing-level heights (episodes 4, 5, and 6), the brightband-related positive bias is very much reduced with the new QPE. In the second class of distance, the correlation is improved in 24 cases out of 27 and the normalized bias tends, on average, to be closer to zero with the new QPE even though it is not systematic (see, e.g., episodes 0 and 1). The most striking feature of the last class of distance is the fact that the negative normalized bias of the old QPE is always reduced. This improvement is more pronounced when the freezing-level height is low. As in the other class of distances, the correlation coefficient is enhanced in most cases.

d. Evaluation of the proposed quality indexes

A quality index map accompanies each 5-min radar QPE. The quality index takes into account the presence of ground clutter, the beam height over the underlying terrain, and the occultation rate [Eqs. (2) and (3)]. In subsequent versions, the quality indexes might also depend upon the estimated attenuation by rain. There will still remain error sources such as $Z-R$ relationship fluctuations and calibration biases that will be impossible to predict in real time with conventional radars. The quality index is between 0 and 1 but it has a wide dynamic range due to its exponential dependency upon altitude. Consequently, it is the logarithmic of the quality index (Ω) coded on 1 byte that is displayed to users. Here, Ω is equal to 0 when the quality is maximal: this is the case for an unshielded measurement made at ground level. A value of 255 means that the quality is very poor, as is the case in strongly shielded or cluttered areas or when the beam is very high. Table 2 illustrates the variations of Ω with the occultation rate and with height (above ground level). The quality index formula is clearly empirical. In this section, we attempt to verify that it is however quite relevant to the characterization of the quality of the radar QPE. The methodology is close to the one developed by Pellarin et al. (2002) to validate the concept of hydrologic visibility. The radar–rain

gauge pairs of hourly accumulations of all 27 episodes have been stratified according to the corresponding quality index. The [0, 255] range of variation of Ω has been divided into 25 classes. Standard scores have then been computed for each class of quality index and each episode. Figure 11 shows the Nash criterion (top) and the correlation coefficient (bottom) for the autumn (left) and winter (right) episodes. There is one curve per episode. The thick dotted curve is the mean curve. Whichever score is considered, there is a lot of scatter among the various curves because the proposed formula for the quality indexes does not account for all possible discrepancies between radar and rain gauge measurements. In addition, the variance of the estimation of the scores is large for low and high values of Ω because the sample size of the radar and rain gauge measurements in those classes is very small. Overall, there is a very well defined trend in the mean curves. The Nash criterion and correlation coefficient gradually decrease with an increasing quality index. For autumn, the Nash criterion drops from 0.6 for Ω in [0, 10] down to 0 for Ω in [100, 110]. The correlation coefficient starts at 0.8 for Ω in [0, 10] and reaches 0 for Ω around [150, 160]. The scores are worse for winter, which is not surprising because the beam is more often in the snow region and unaccounted for VPR effects can cause large radar to rain gauge fluctuations.

e. Examples of mosaics of QPE and quality indexes

The French radar network consists of 24 radars in 2007. The 5-min QPEs product and their associated quality index maps are concentrated in real time in Toulouse. To this point, the QPE products have not been mosaicked, so forecasters in charge of the monitoring and surveillance of precipitation over the entire French territory have had to handle a lot of different radars and images in the case of synoptic-scale precipitation systems. In addition, in overlapping areas, it is often advantageous to combine the radar QPEs from neighboring radars so as to mitigate their errors and improve the quality of the rainfall estimation. For those two reasons, it was decided that a QPE mosaic should be developed and implemented. The QPE mosaic covers all of France. It has the same space–time resolutions as individual QPEs, namely 1 km^2 and 5 min. The composition rules for any pixel of the composite QPE are exactly the same as the ones given in Eqs. (1) and (3) for producing single-radar QPEs.

Given the fact that the average distance between radars is around 150 km, any pixel from the composite is usually covered by two or three radars. In most cases however, due to the height dependency of the quality

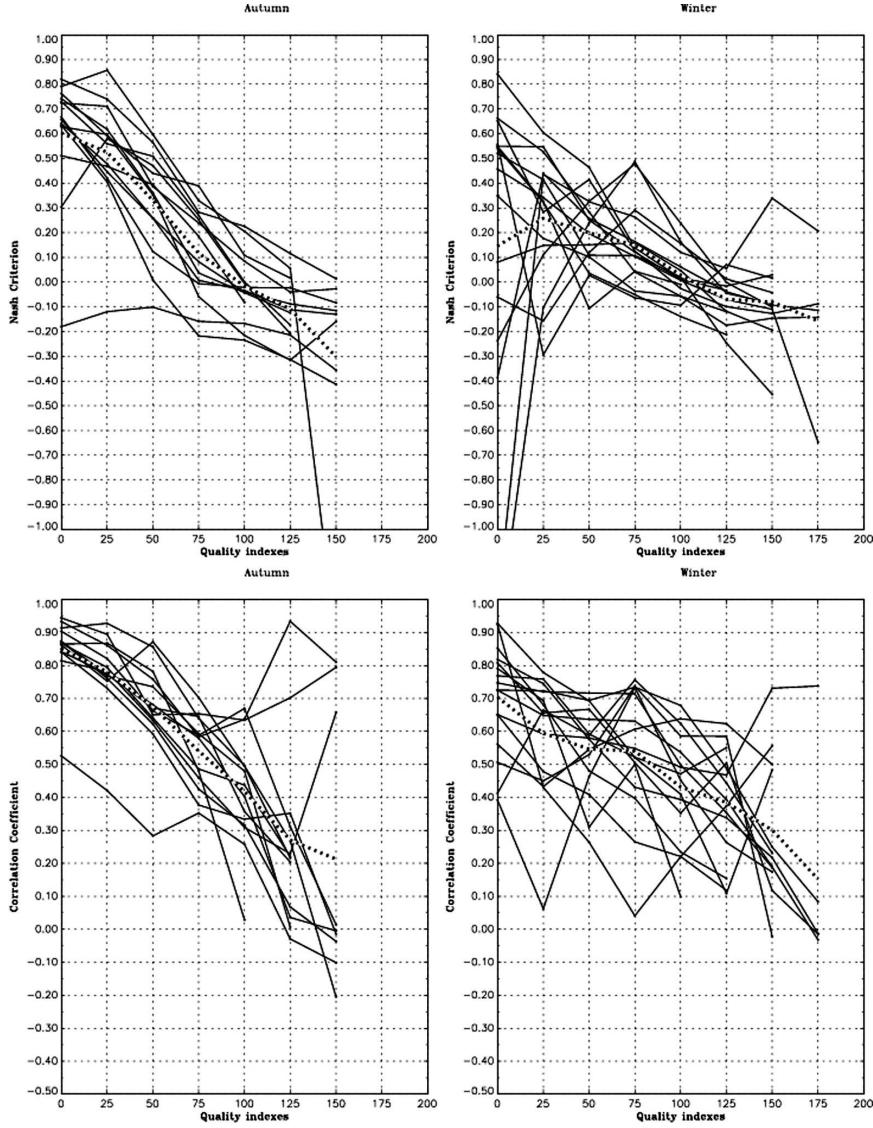


FIG. 11. Comparison of the quality indexes provided by the QPE algorithm to statistical radar–rain gauge scores. Two scores are used: (top) the Nash coefficient and (bottom) the correlation coefficient for (left) autumn and (right) winter. There is one curve per episode. The thick dotted line is the mean curve.

index, the closest radar has the largest weight in the weighted linear combination.

Figure 12 shows an example of a 24-h accumulation QPE composite product (Fig. 12a) along with its quality indexes (Fig. 12b). Twenty-one radars from the network have been used to create this composite. They are indicated by white crosses in Figs. 12a and 12b. The accumulation period started on 25 June 2006. An advantage of using a weighted linear combination in overlapping areas is that the resulting composite QPE presents no discontinuities, which are otherwise often obvious on long-term accumulations. Moreover, quantitative comparisons with rain gauge data (not

shown here) show that the QPE mosaic almost always outperforms any other single-radar QPE in overlapping areas. The mosaic of quality indexes (Fig. 12b) is very useful because it takes into account in a transparent way the geometry of the radar network. For instance, if a radar from the network has a problem and ceases transmitting images, then the rainfall estimation for its usual coverage area will still be obtained from its closest neighbors. The quality of the estimation will of course be worse than usual because the surface rainfall estimation is made from farther way, hence from farther aloft. This, however, will be reflected by the composite quality index map.

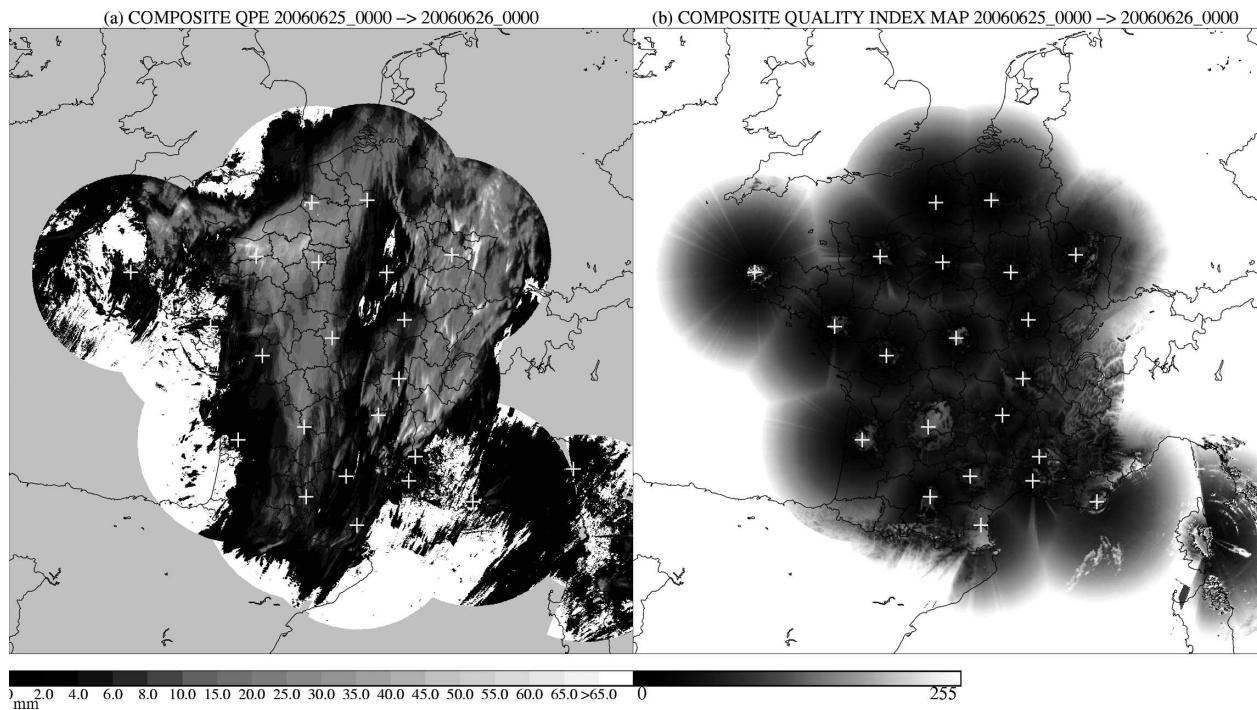


FIG. 12. Example of (a) a 24-h accumulation composite QPE product with (b) its quality indexes. Twenty-one radars from the network have been used to create this composite. They are indicated by white crosses in both maps. The accumulation period starts 0000 UTC 25 Jun 2006.

4. Conclusions

This paper is dedicated to the validation of the new French operational radar rainfall product (Part I). The new processing chain includes ground-clutter identification, partial beam-blocking correction, VPR correction, and a weighted linear combination of all available tilts.

A specific evaluation of the VPR identification module was carried out by analyzing 489 rainfall episodes gathered by a subset of eight radars from the French radar network. The conceptual model of VPR, which relies on the freezing-level height and three shape parameters (brightband peak, brightband thickness, and upper-level decreasing rate), was shown to be quite relevant in reproducing the actual profiles of reflectivity. A climatology of the three shape parameters was established in order to better define the family of VPR candidates. The radar-derived freezing-level heights were shown to be in quite good agreement with the radiosounding data. Better agreement was obtained with the dry 0°C isotherm height than with the analogous wet-bulb temperature.

The new QPE product has been compared to the old one for 27 episodes observed during the winter of 2005 and the autumns of 2002 and 2003 by three radars

within the network (Nîmes, Bollène, and Opoul). The new QPE product outperforms the old one for whichever distances (0–50, 50–100, and 100–150 km) and scores are considered. The correlation coefficient is almost systematically increased and the biases, whether positive or negative, are generally reduced. A negative bias still exists at long range. The benefits of the VPR, partial beam-blocking, and ground-clutter correction procedures have been assessed using S-band radars but they should be exactly the same for C-band radars, which constitute the majority of the French operational radars.

The proposed empirical quality indexes have been compared to long-term radar–rain gauge scores. Despite their simplicity, the quality indexes currently only depend upon partial beam blocking and the height over the terrain, they explain a large fraction of the radar–rain gauge variability. Such quality indexes can be very useful in applications such as mosaicking, assimilation in hydrological models, verification of numerical weather prediction models, and radar–rain gauge adjustment.

Acknowledgments. This work was done within the framework of the French Project Programme Aramis Nouvelles Technologies en Hydrométéorologie Exten-

sion et Renouvellement (PANTHERE) supported by Météo-France, the French Ministère de l'Ecologie et du Développement Durable, and the European Regional Development Fund (ERDF) of the European Union.

REFERENCES

- Anagnostou, E. N., and W. Krajewski, 1999: Real-time radar rainfall estimation. Part I: Algorithm formulation. *J. Atmos. Oceanic Technol.*, **16**, 189–197.
- Andrieu, H., and J. D. Creutin, 1995: Identification of vertical profiles of radar reflectivity for hydrological applications using an inverse method. Part I: Formulation. *J. Appl. Meteor.*, **34**, 225–239.
- , G. Delrieu, and J. D. Creutin, 1995: Identification of vertical profiles of radar reflectivity for hydrological applications using an inverse method. Part II: Formulation. *J. Appl. Meteor.*, **34**, 240–259.
- Delrieu, G., and Coauthors, 2005: The catastrophic flash-flood event of 8–9 September 2002 in the Gard region, France: A first case study for the Cévennes–Vivarais Mediterranean Hydrometeorological Observatory. *J. Hydrometeor.*, **6**, 34–52.
- Dinku, T., E. N. Anagnostou, and M. Borga, 2002: Improving radar-based estimation of rainfall over complex terrain. *J. Appl. Meteor.*, **41**, 1163–1178.
- Doviak, R. J., and D. S. Zrnić, 1984: *Doppler Radar and Weather Observations*. Academic Press, 458 pp.
- Ducrocq, V., F. Bouttier, S. Malardel, T. Montmerle, and Y. Seity, 2005: Le projet AROME, crues méditerranéennes: La réponse de l'état. *La Houille Blanche*, **2**, 39–43.
- Fabry, F., and I. Zawadzki, 1995: Long-term radar observations of the melting layer of precipitation and their interpretation. *J. Atmos. Sci.*, **52**, 832–851.
- Fulton, R. A., J. P. Breidenbach, D.-J. Seo, D. A. Miller, and T. O'Bannon, 1998: The WSR-88D rainfall algorithm. *Wea. Forecasting*, **13**, 377–395.
- Germann, U., and J. Joss, 2002: Mesobeta profiles to extrapolate radar precipitation measurements above the Alps to the ground level. *J. Appl. Meteor.*, **41**, 542–557.
- Giangrande, S. E., and A. V. Ryzhkov, 2004: Polarimetric method for bright band detection. Preprints, *11th Conf. on Aviation, Range, and Aerospace Meteorology*, Hyannis, MA, Amer. Meteor. Soc., CD-ROM, P5.8.
- Grégoire, P., 2002: Monitoring of the Radome network. WMO/TD-1123, Rep. 75 (TECO-2002), Bratislava, Slovakia, 2.2(1).
- Joss, J., and R. Lee, 1995: The application of radar-gauge comparisons to operational precipitation profile corrections. *J. Appl. Meteor.*, **34**, 2612–2630.
- Kitchen, M., 1997: Towards improved radar estimates of surface precipitation rate at long range. *Quart. J. Roy. Meteor. Soc.*, **123**, 145–163.
- , R. Brown, and A. G. Davies, 1994: Real-time correction of weather radar data for the effects of bright band, range and orographic growth in widespread precipitation. *Quart. J. Roy. Meteor. Soc.*, **120**, 1231–1254.
- Koistinen, J., 1991: Operational correction of radar rainfall errors due to the vertical reflectivity profile. Preprints, *25th Int. Conf. on Radar Meteorology*, Vol. I, Paris, France, Amer. Meteor. Soc., 91–94.
- Mittermaier, M. P., and A. J. Illingworth, 2003: Comparison of model-derived and radar-observed freezing level heights: Implications for vertical reflectivity profile correction schemes. *Quart. J. Roy. Meteor. Soc.*, **129**, 83–96.
- Novak, P., and J. Kracmar, 2001: Vertical reflectivity profiles in the Czech Weather Radar Network. Preprints, *30th Int. Conf. on Radar Meteorology*, Munich, Germany, Amer. Meteor. Soc., CD-ROM, P15.3.
- Pellarin, T., G. Delrieu, G. M. Saulnier, H. Andrieu, B. Vignal, and J. D. Creutin, 2002: Hydrologic visibility of weather radar systems operating in a mountainous region: A case study for the Ardèche catchment (France). *J. Hydrometeor.*, **3**, 539–555.
- Tabary, P., 2003: Efforts to improve the monitoring of the French radar network. Preprints, *31st Int. Conf. on Radar Meteorology*, Seattle, WA, Amer. Meteor. Soc., CD-ROM, P3C.9.
- , 2007: The new French operational radar rainfall product. Part I: Methodology. *Wea. Forecasting*, **22**, 393–408.
- Vignal, B., and F. W. Krajewski, 2001: Large-sample evaluation of two methods to correct range-dependent error for WSR-88D rainfall estimates. *J. Hydrometeor.*, **2**, 490–504.
- Wilks, D. S., 1995: *Statistical Methods in the Atmospheric Sciences: An Introduction*. Academic Press, 467 pp.
- Zawadzki, I., 1984: Factors affecting the precision of radar measurements of rain. Preprints, *22d Int. Conf. on Radar Meteorology*, Zurich, Switzerland, Amer. Meteor. Soc., 251–256.