New Algorithm for Rain Cell Identification and Tracking in Rainfall Event Analysis

Ting He [1,2,*], JianXin Zhang [1,*], JiYao Hua [1] and Yang Cai [1]

[1.] Information Center (Hydrology Monitor and Forecast Center), Ministry of Water Resources, P.R. China, No.2, Lane 2, BaiGuang Road, BeiJing 100053, China
[2.] Institute of Geographical Sciences, Free University of Berlin, Berlin, 12249, Germany
* Correspondence: heting@mwr.gov.cn, zhangjx@mwr.gov.cn

Abstract: This study proposes a new algorithm termed rain cell identification and tracking (RCIT) to identify and track rain cells from high resolution weather radar data. Previous algorithms have limitations when tracking non-consequent rain cells owing to their use of maximum correlation coefficient methods and their lack of an alternative way to handle the variation stages of rain cells during their life cycles. To address these deficiencies, various methods are implemented in the new algorithm. These include the particle image velocimetry (PIV) method for motion estimation and the rain cell matching rule to obtain the stage changes of rain cells. High resolution (5-min and 1-km) radar reflectivity data from three rainy days over the German federal state North Rhine Westphalia (NRW) are used to evaluate the proposed algorithm. The performance of the new algorithm is compared with a radar reflectivity map and verified by two object-oriented methods: structure–amplitude–location (SAL) and geometric index. The verification results suggest that the performance of the new algorithm is good. Application of the RCIT algorithm to the selected cases shows that the inner structure of rainfall events in the experimental region present extreme value distributions, with most rainfall events having a short duration with less intensity. The new algorithm can effectively capture the stage changes of rain cells during their life cycles. The proposed algorithm can serve as the basis for further hydro-meteorological applications such as spatial and temporal analysis of rainfall events and short-term flood forecasting.

Keywords: rain cell; tracking; PIV; feature-based verification

1. Introduction

Precipitation is a key process in Earth's water circle. Acquiring explicit knowledge about its inner behavior is critical to assisting us in understanding its interaction with hydrological processes. Rainfall events are characterized by several elements, such as duration, intensity, velocity, and spatial and temporal variability (Elena et al. 2017). The variability of rainfall events can be defined as "the variability derived from having multiple spatially-distributed rainfall fields for a given point in time" (Peleg et al. 2017). In hydro-meteorological applications, rainfall always varies over its life cycle; this variation also differs between different types of event (e.g., convective and stratiform rainfall events). As a consequence, the responses of hydrology models are sensitive to this variation. Modeling rainfall events and analyzing their spatial and temporal information is necessary.

For rainfall event monitoring, intensity and cumulative value are the two most common indexes and they are usually measured using a rain gage, which is the standard instrument for providing direct observations. Nevertheless, a rain gage cannot directly detect variability in rainfall events, it is also subject to errors owing to topography and wind effects. As a possible alternative, weather radar has played a major role in recent years owing to its high spatial and temporal resolution. This is advantageous in terms of (i) acquiring spatial and temporal patterns when modeling rainfall events and (ii) undertaking short-term rainfall forecasting at fine scales. Identifying and tracking rainfall events is a common task in radar-based meteorological and hydrological applications (Moseley et al. 2013; Novo et al. 2014; Guinard et al. 2015; Yeung et al. 2015).

Broadly speaking, the corresponding algorithms for radar-based rainfall event identification and tracking can be classified into pixel- and object-based approaches (Zahraei et al. 2013). The pixel-based approaches are also referred to as advection field tracking. These are pattern matching

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 August 2019 doi:10.20944/preprints201908.0011.v1

Peer-reviewed version available at Atmosphere 2019, 10, 532; doi:10.3390/atmos10090532

approaches for extracting motion vectors by searching for the maximum correlation coefficient of rain cells in two consecutive radar images. Pixel-based algorithms are highly effective when they are applied in convective rainfall nowcasting, which is usually found in frontal systems (Anna et al. 2018). Algorithms of this kind are capable of distinguishing between large scale convective and stratiform rainfall, although they are not so effective for individual RCIT. Herein, a rain cell is defined as the closed contours over which rainfall intensity exceeds a given threshold during one rainfall event (Féral et al. 2006). Representative algorithms can be listed as follows: TREC (Rinenart and Garvey 1978), TITAN (Dixon and Wiener 1993), COTREC (Li et al.1995), and SWIRLS (Li et al. 2014). Object-based approaches (also termed cell tracking) include (i) a detecting algorithm for searching a discrete rain cell's properties (e.g., centroid, area, echo-tops, and vertical-integrated) in consecutive radar images and (ii) a matching algorithm for tracking cell motion and shape changes (e.g., merging and splitting). The advantage of object-based approaches is in reflecting the dynamic of convective rainfall; they are suitable for convective rain storm analysis but not effective in straitiform rainfall identification. Representative algorithms can be listed as follows: SCOUT (Einfalt et al. 1990), SCIT (Johnson et al. 1998), Trace3D (Handwerker 2002), and PERsiann-ForeCAST (Zahrani et al. 2013).

Despite the thorough application of these rain event identification and tracking algorithms, they have the following deficiencies. For pixel-based approaches, motion estimates are mostly based on the maximum correlation coefficient, which may yield non-continuous results when fast decay of rainfall occurs. For object-based approaches, motion estimates obtained from the rain cell center of mass may lack accuracy owing to the random center of mass displacement problem (Han et al., 2009). This problem occurs when the shape of rain cells changes rapidly between successive radar images. As a consequence of these motion estimate inaccuracies, these algorithms may also encounter difficulties when handling merging and splitting scenarios (Muñoz et al., 2018), whereas a rain cell can begin its life cycle by simply emerging at a location with no rain. It is also possible that a rain cell can become separated from a large single rain cell or that several smaller rain cells can merge into one (Moseley et al. 2013). For such reasons, a new rain cell identification and tracking algorithm (RCIT) is proposed in this work. The algorithm is developed by combining the advantages of pixel- and object-based approaches and is able to handle the problem of detecting the stage changes of rain cells.

This paper is organized as follows. An introduction to the study area and radar data is given in Section 2. The RCIT algorithm is illustrated in Section 3. Section 4 presents structure, amplitude, and location (SAL) and geometric verification results and practical applications of the algorithm to North Rhine Westphalia (NRW) rainfall events. In Section 5, the main conclusions and further expectations of this work are given.

2. Rain Cell Identification and Tracking Algorithm - RCIT

The aim of the RCIT is to analyze rainfall events by fully utilizing the merits of weather radar. The inputs to the proposed algorithm are radar reflectivity maps and the outputs are the properties of rain cells such as area, intensity, and trajectory. The RCIT involves two modules: rain cell identification and tracking, as presented as in Figure 1.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 August 2019 doi:10.20944/preprints201908.0011.v1

Peer-reviewed version available at Atmosphere 2019, 10, 532; doi:10.3390/atmos10090532

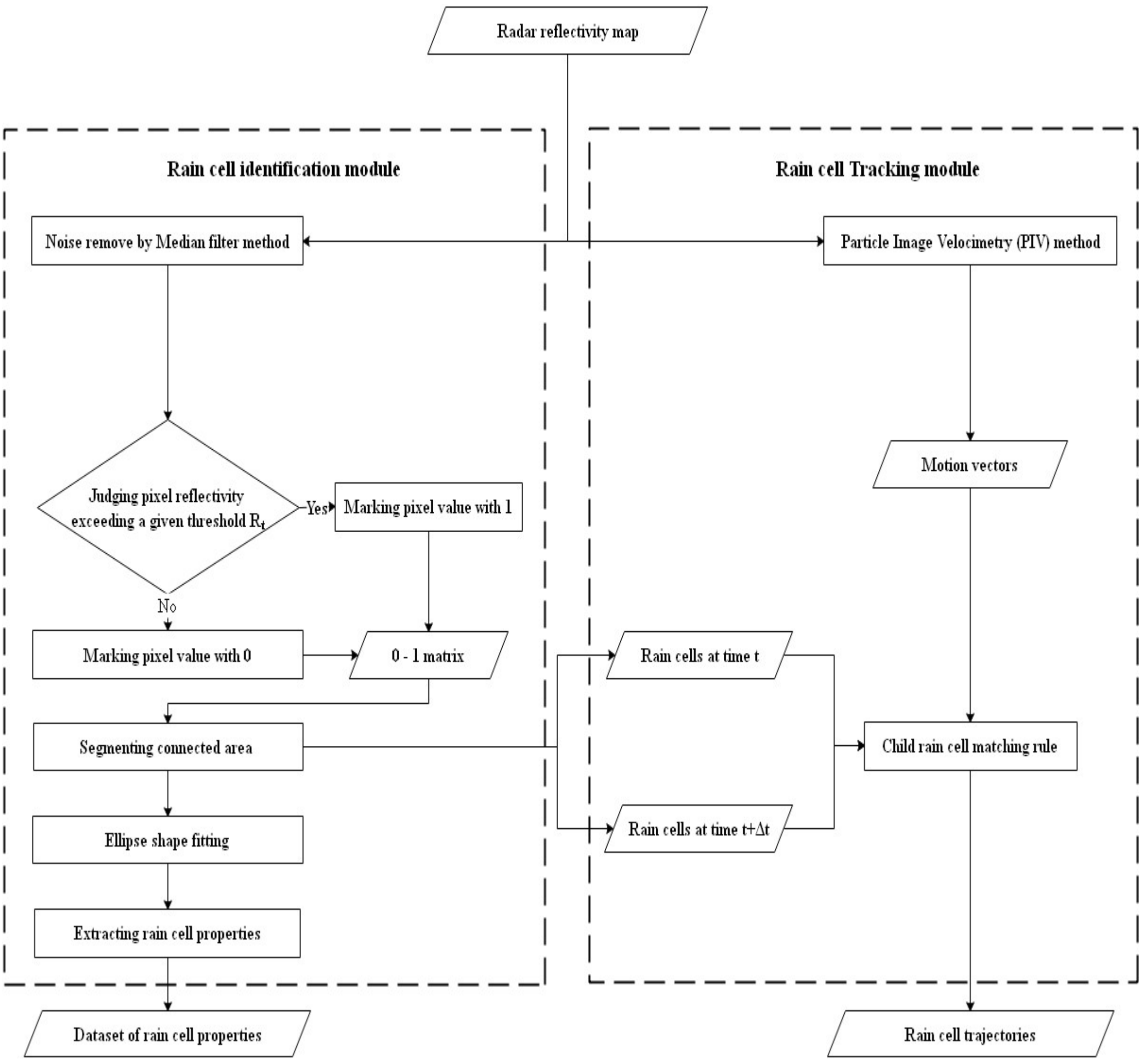


Figure 1. Step illustration of rain cell identification and tracking algorithm.

2.1. Rain Cell Identification Module

Similar to the object-based algorithms, the rain cell identification module of the RCIT is based on discerning a connected domain above a given threshold. As presented in the left portion of Figure 1, each selected radar reflectivity map in the Cartesian coordinates is initially filtered by the median filtering method (Anoraganingrum 1999) to remove noisy pixels (pixels with abnormally high reflectivity). Then, pixels above a given reflectivity threshold are assigned the value one, with the remainder assigned the value zero. A segmenting process is implemented to assemble and cluster pixels sharing the same reflectivity threshold into a connected area. In the segmenting process, the following rules suggested by Peleg and Morin (2012) are obeyed: (i) If the reflectivity of a rainy pixel is lower than a given threshold R_t, then it is set to null. (ii) For each rainy pixel and its eight neighbors, if more than five of them are null, then it is set to null. (iii) If the pixel is spurred, then it is set to null. Herein, spur pixels are those isolated pixels whose reflectivity is different to others along the horizontal and vertical directions in the labeled binary image. (iv) If the area of a connected region is smaller than 9 km^2, then it is ignored. All the segmented regions are then labeled and fitted with an ellipse shape. Their properties are extracted and stored in a relational database. The extracted properties are as follows:

i) Area [km^2] - Sum value for the number of pixels contained in one rain cell.

ii) Areal rainfall depth [mm] - Cumulative precipitation of one rain cell over a 5-minute interval.

iii) Maximum intensity [$mm.h^{-1}$] - Peak intensity of one rain cell.

iv) Areal mean rainfall depth [$mm.km^2$] - Ratio of the areal rainfall depth and area.

v) Eccentricity - Ratio of minor and major axes, which are acquired from the fitted eclipse. Used to describe the shape of one rain cell with a value range from 0 to 1.

vi) Center of mass [km] - Center of mass of a rain cell, which is weighted by the reflectivity of rainy pixels.

Property calculation: Areal rainfall depth, maximum intensity, and areal mean rainfall depth are based on the reflectivity (Z) and rain rate (R) converting function: $Z = aR^b$.

2.2. Rain Cell Tracking Module

The rain cell tracking module is established based on a hybrid approach, as illustrated in the right portion of Figure 1. In the first procedure, motion vectors are estimated by implementing the particle image velocimetry (PIV). This is an optical method of flow visualization that is used to obtain instantaneous velocity measurements and related properties of fluids (Merzkirch 2001; Adrian 2005; Westerweel et al. 2013). It consists of a class of flow measuring mechanisms that are characterized by recording the displacement of small particles embedded in a region of fluid. Figure 2 shows a PIV application in motion vector estimation. In the first step, a window box of size $r \times r$ is initially defined, which divides radar images into several sub-blocks. In the second step, a searching distance, $d = 2 \times v_{max} + 1$, is defined, where v_{max} is the preset maximum velocity. The minimum quadric difference (MQD), as suggested by Gui and Merzkirch (1996), is employed in searching the optimal grid points at time $t + \Delta t$, as in Equation (1):

$$MQD(\Delta x, \Delta y) = \sum_{i=1}^{N} \sum_{j=1}^{N} \left| R_1(X_i, Y_j) - R_2(X_i + \Delta x, Y_j + \Delta y) \right| \qquad (1)$$

where $R_1(X_i,Y_j)$ and $R_2(X_i,Y_j)$ are the reflectivity of grid points contained within the window boxes of radar images at time t and $t+\Delta t$, respectively; Δx and Δy ($\Delta x, \Delta y \in d$) are the locations of minimum reflectivity difference in the horizontal and vertical directions separately. The minimum reflectivity differences of grid points within the window boxes are reversed to simplify the calculation. In order to guarantee that the solitary peak locations can be calculated, Δx and Δy are corrected separately in the horizontal and vertical directions by fitting a second-order polynomial to the logarithm of the maximum reflectivity of the grid point and its three direct neighbors, as in Equation (2). In this way, the optimal grid points at time $t + \Delta t$ are identified, with their locations presented as $(x + \Delta x - \frac{d+1}{2}, y + \Delta y - \frac{d+1}{2})$. In the final step, the calculated motion vectors are smoothed by the median filter algorithm.

$$\Delta x = x - \frac{logR(x+1,y) - logR(x-1,y)}{2logR(x+1,y) - 4logR(x,y) + 2log\ (x-1,y)} \qquad (2a)$$

$$\Delta y = y - \frac{logR(x,y+1) - logR(x,y-1)}{2logR(x,y+1) - 4logR(x,y) + 2log\ (x,y-1)} \qquad (2b)$$

In the second procedure, rain cells at time t and $t + \Delta t$ are identified. Finally, a child rain cell matching rule is applied for identifying the most-matched rain cells. The child rain cell matching scheme considers the stage changes of rain cells between successive radar images (e.g., merge, split, growth, and decay), using certain indexes for determination such as overlap, area diversification, distance, and angle difference of center of mass. Before introduction of the rain cluster matching rule, certain definitions were identified: (i) for two radar images at time t and $t + \Delta t$, rain cells identified from radar images at t are termed parent cells and (ii) rain cells identified from radar images at $t + \Delta t$ are termed child cells. These definitions can be depicted as follows:

Peer-reviewed version available at Atmosphere 2019, 10, 532; doi:10.3390/atmos10090532

(1) A boundary box of a parent cell is defined, with a horizontal length of [10+max(pos_x), min(pos_x)−10] and vertical length of [10+max(pos_y), min(pos_y)−10], where pos_x and pos_y are Cartesian coordinates of pixels in the parent cell.

(2) The number of child cells falling into the boundary box is determined and their properties, e.g., area, areal rainfall depth, max intensity, areal mean rainfall depth, and center of mass, are selected.

(3) If only one child cell is searched in the boundary box and it overlaps with a parent cell, then it is the most-matched rain cell. If this child cell does not overlap with a parent cell and the distance and angle difference for the center of mass between it and the parent cell are less than 3 × mean (V_{motion_vector}) and 3 × θ_{motion_vector}, it is also the most-matched rain cell, where mean (V_{motion_vector}) and θ_{motion_vector} are the mean value of velocity and the prevailing direction of the motion vector, respectively.

(4) If two or more child cells fall into the boundary box without overlapping a parent cell, the matching rule is changeless; however one extra condition is included, i.e., child cells whose areas have minimum absolute differences with the parent cell are the most-matched rain cells.

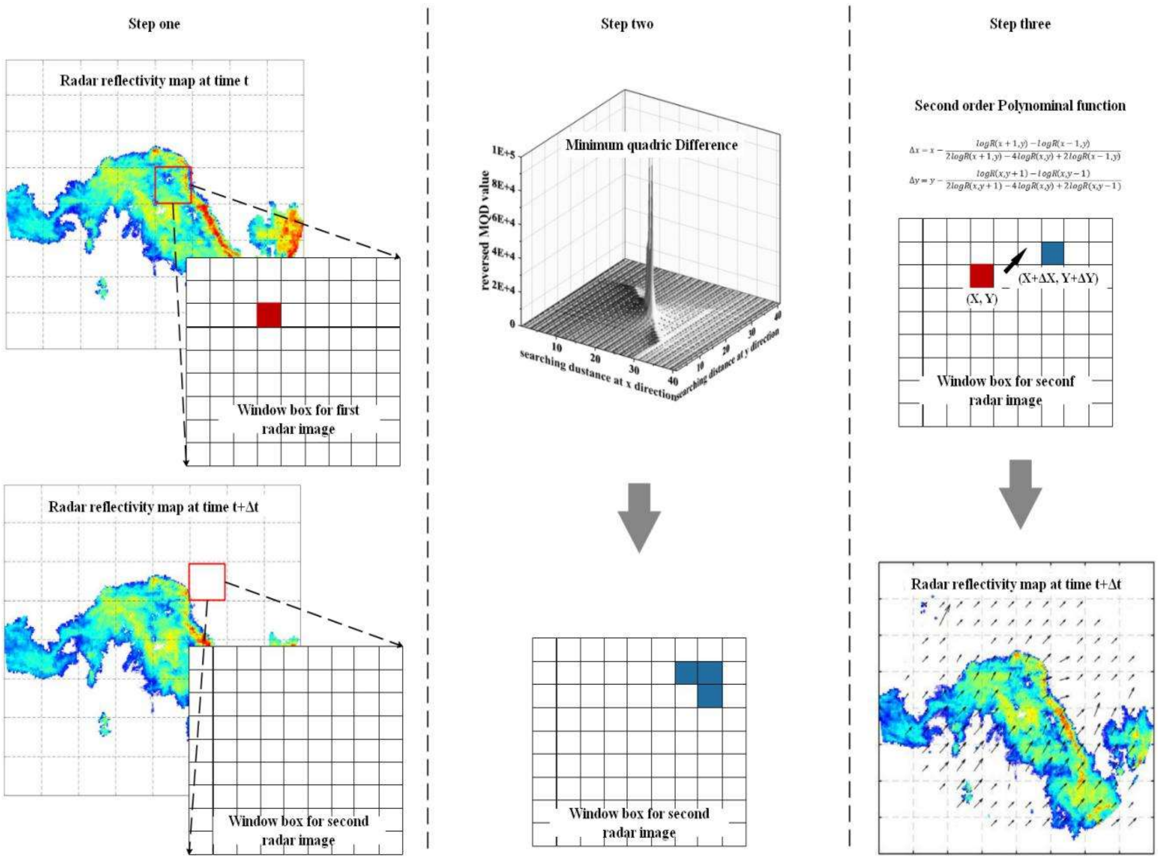


Figure 2. Illustration of PIV application in rain cell motion estimation. Step one: window boxes with the area of r × r are defined (rectangular with red color); step two: for any grid point in the window box at a previous timepoint (red block), the MQD algorithm is applied to deduce the minimum reflectivity differences, and grid points with minimum value in the next window box are identified (blue blocks); step three: the solitary locations of reversed MQD value, Δx and Δy, are corrected by applying Equation (2). The locations of the optimal grid point at the next timepoint are calculated using the second polynomial function, and the global motion vectors are extracted and smoothed by the median filter method.

3. Study Area and Data

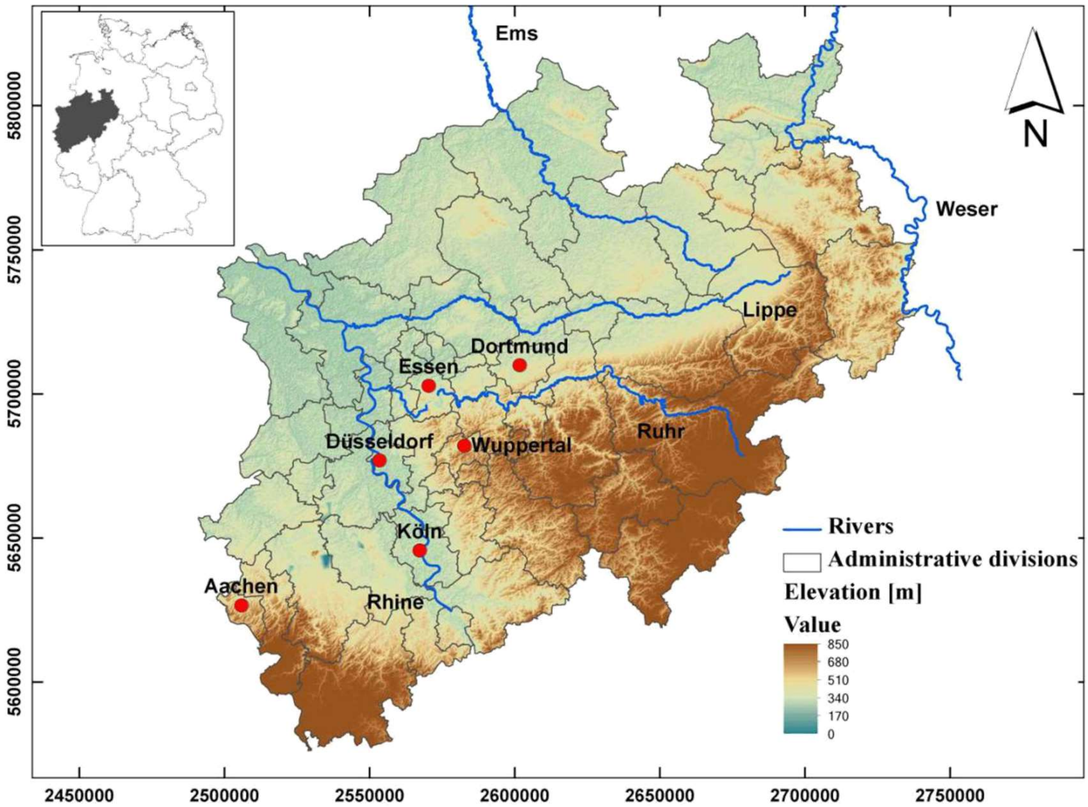
Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 August 2019 doi:10.20944/preprints201908.0011.v1

Peer-reviewed version available at Atmosphere 2019, 10, 532; doi:10.3390/atmos10090532

The study area is the federal state of NRW (Figure 3). It is bordered by the German federal states of Lower Saxony to the north and north-east, Hessen to the east, and Rhineland–Palatinate to the south, and by the countries of Belgium to the south-west and the Netherlands to the west. NRW includes upland regions of North Eifel in the south and mountains of the Sauerland in the south-east. There are five rivers in this study area: Rhine, Ruhr, Ems, Lippe, and Weser. Two main types of landscape can be found in NRW, namely the North German lowlands, with elevations just a few meters above sea level, and the North German low mountain range, with elevations of up to 850 m. The lowland areas comprise the Rhine–Ruhr area, which is one of the largest metropolitan areas, with a population of approximately 10 million. The circulation pattern of NRW is mainly affected by the air mass from the Atlantic Ocean along the direction toward the mainland. When arriving at the southern high mountain regions, the air mass stops and rises; this leads to more cloudiness and precipitation. On the eastern side of the mountain regions, drier air masses result in less cloudiness and less precipitation.

Radar data were obtained from the Essen radar deployed in Essen City, NRW. The Essen radar has been deployed over the study area and is a part of the radar network of the German weather service (DWD). The Essen radar is a dual-polarimetric C-band Doppler radar. It delivers radar volume scans (frequency: 800/1200 Hz, maximum range: 124 km) every 15 minutes for the Doppler velocity, together with intensity volume scans (frequency: 500 Hz, maximum range: 256 km) and precipitation scans (frequency: 600 Hz, maximum range: 150 km) every 5 minutes for the precipitation echo, with a high spatial resolution of 1 km in range and 1° in azimuth. For this study, the Essen radar provided precipitation scans with an elevation of 0.8° and a range of 128 km. The output reflectivity was selected with a plan position indicator (PPI) display type.

As radar measures precipitation in an indirect manner, the quality of radar data must be carefully checked. The sources that affect the quality of the Essen radar data include ground clutter and speckle, beam blockage, and attenuation. The corresponding quality correction methods in this work follow Golz et al. (2006). After quality checking, an open source package "Wradlib" (Heistermann et al. 2013) was applied to project the raw radar image onto a 256 × 256 km^2 Cartesian map with 1-km resolution. In total, 864 radar reflectivity images for three rainy days (May 26, 2007; July 19, 2008; and July 26, 2008), including some recorded convective rainfall events, were selected to evaluate the proposed algorithm. The daily rainfall distributions of these rainy days are shown in Figure 4.



204
205 **Figure 3.** Study area of North Rhine Westphalia (NRW) and its location in Germany. The main
206 administrative cities are marked with red dots.

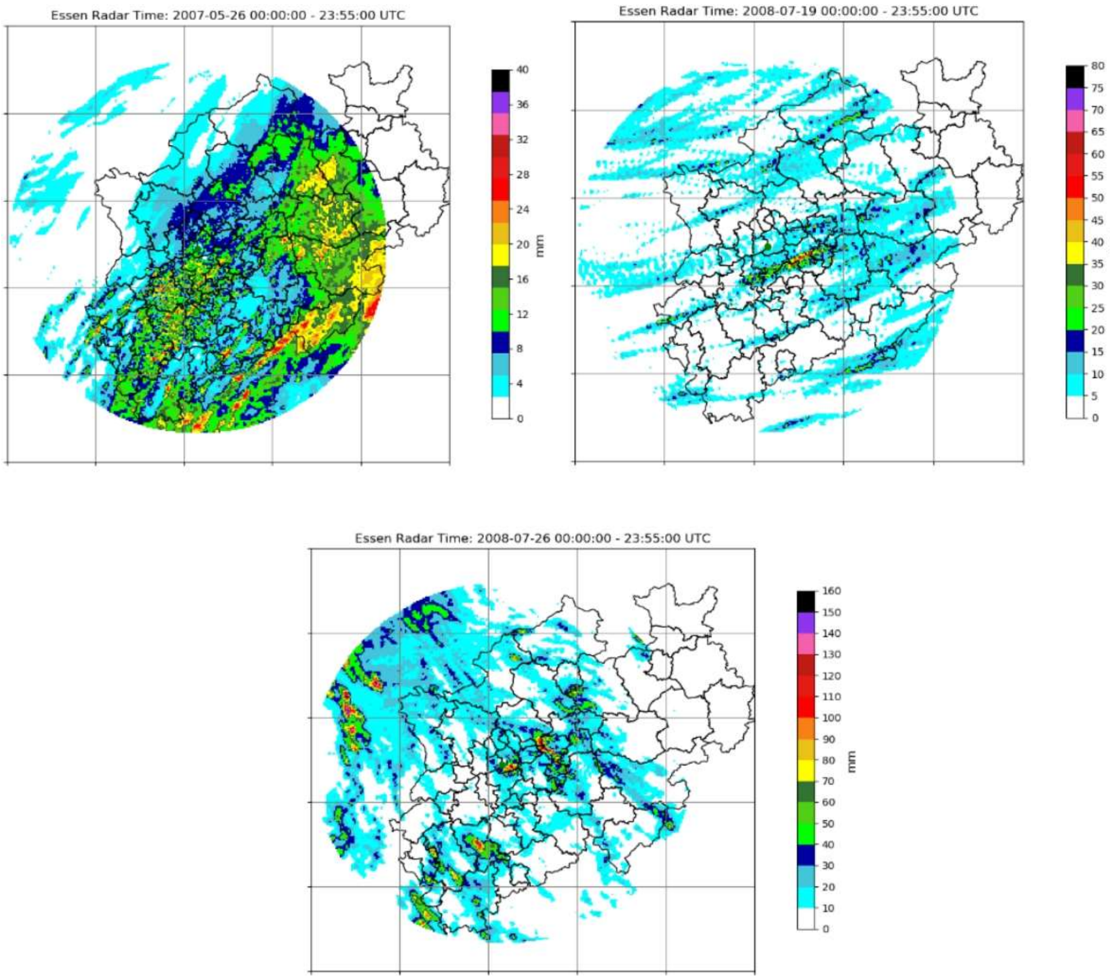


Figure 4. Daily rainfall distribution for selected rainy days. Radar reflectivity was converted into rain rate according to DWD standard Z–R relationship.

4. Results Analysis and Discussion

In the algorithm application process, the reflectivity threshold for rain cell identification was set to 19 dBZ, which was based on the classification of DWD presented by Weusthoff and Hauf (2008), as in Table 1. For the German weather radar system, two common Z–R relationships were used by Weusthoff and Hauf (2008): one was categorized for the RADOLAN product and the other uses constant a and b with values of 256 and 1.42, respectively. Although the DWD has stated that the categorized relationship statistically shows better results over long time periods, the standard relationship can be more compatible with local cases when a correction factor is added (Einfalt and Frerk 2012). Based on the above considerations, we applied the DWD standard Z–R relationship to radar reflectivity–rain rate conversion in the application cases.

Table 1. Conversion of radar reflectivity to rain rates using Z–R relationship (a = 256, b = 1.42) with thresholds according to the classification of the DWD.

Z [dBZ]	R [mm.h^{-1}]	Rain Rate [mm.5 min^{-1}]
> 55	> 150	> 12.5
46–55	35–150	2.92–12.5
37–46	8.1–35	0.68–2.92
28–37	1.9–8.1	0.16–0.68
19–28	0.4–1.9	0.03–0.16
7–19	0.06–0.4	0.005–0.03

4.1. Performance Assessment of RCIT Algorithm

Grid-point related error measurement is problematic for the rain cell tracking algorithm. A classic example illustrating this problem is the well-known "Double Penalty" problem, in which prediction of a precipitation object at the correct size and structure might yield very poor verification scores. For example, one rain cell is displaced slightly in space but the categorical verification scores heavily penalize such a situation. In traditional verification methods, a displacement simply leads to a false alarm, and it is also very poorly rated owing to its large root mean squared error (Davis et al. 2006). On the other hand, despite a great deal of effort in the statistical validation of grid-based rainfall estimated results, verification associated with the geometry patterns of rain cells has not been well researched or applied.

As an alternative, feature-based verification methods have been built upon the idea of identifying rainfall events as "objects". With this perspective, simulated and observed rainfalls are not compared directly at the same location; rather, objects of interest are extracted from simulated and observed data and then compared together so that verification statistics are obtained. A number of spatial verification methods have been proposed (Ebert 2008; Gilleland et al. 2009). In the present work two feature-based verification methods, SAL and geometric index, are implemented to verify the performance of the RCIT algorithm. A detailed introduction to the SAL and geometric index methods is presented in Appendix A. The data set for comparison was a simulated radar reflectivity map from the RCIT algorithm (termed sim_ref) and an observed radar reflectivity map (termed obs_ref).

Figure 5 shows the SAL verification results, which are arranged based on the three selected rainy days. For each SAL plot presented in Figure 5, the vertical axis denotes the A component, the horizontal axis denoted the S component, the dots represent values of the S and A components, and the color scale of the dots denotes the L component. Median values of the SAL components are presented as dashed lines. It can be seen that all S and A values are concentrated close to zero, as are most L component values. Table 2 gives the S, A, and L index values for the geometric index objects from the sim_ref and obs_ref data sets. All values were again organized based on the three selected rainy days. It is evident that the index differences between the geometric index objects from the sim_ref and obs_ref data sets were less than 0.05.

Table 2. Results of three geometric index components of geometric index verification objects for RCIT-simulated radar reflectivity maps (sim) and observed radar reflectivity maps (obs). Values are sorted at 5, 50, and 75 percentile levels.

Selected cases		C index			S index			A index		
		25%	50%	75%	25%	50%	75%	25%	50%	75%
May 26, 2007	obs	0.934	0.957	0.977	0.22	0.325	0.509	0.102	0.198	0.417
	sim	0.966	0.979	0.992	0.27	0.378	0.579	0.135	0.271	0.53
July 19, 2008	obs	0.847	0.895	0.938	0.143	0.217	0.29	0.031	0.071	0.118
	sim	0.907	0.943	0.969	0.154	0.233	0.297	0.043	0.086	0.134
July 26, 2008	obs	0.897	0.93	0.955	0.154	0.245	0.374	0.045	0.116	0.213
	sim	0.936	0.965	0.997	0.189	0.285	0.385	0.077	0.149	0.238

The SAL verification results suggest that the shape of most SAL objects from the sim_ref data set was the same as that for SAL objects from the obs_ref data set (except for a few cases that were slightly large and flat). The converted rainfall volume for some SAL objects from the sim_ref data set was less than that from the obs_ref data set; the origin that the rain cell area threshold used in the RCIT algorithm was 9 km^2, with rain cells less than this threshold ignored. However, the converted rainfall volume of most SAL objects from the sim_ref data set was close to that from the obs_ref data set. Location differences of SAL objects between the sim_ref and obs_ref data sets were not obvious. Geometric index verification results indicated that the geometric pattern of geometric index objects from the sim_ref data set was approximately the same as that for objects from the obs_ref data set (except for connectivity). Differences in the C index for geometric index objects from the two data sets were obvious, which may have been because the median filter method applied in the RCIT algorithm smoothed abnormal pixels in the radar reflectivity map. In general, the RCIT algorithm performed well based on the two feature-based verification methods.

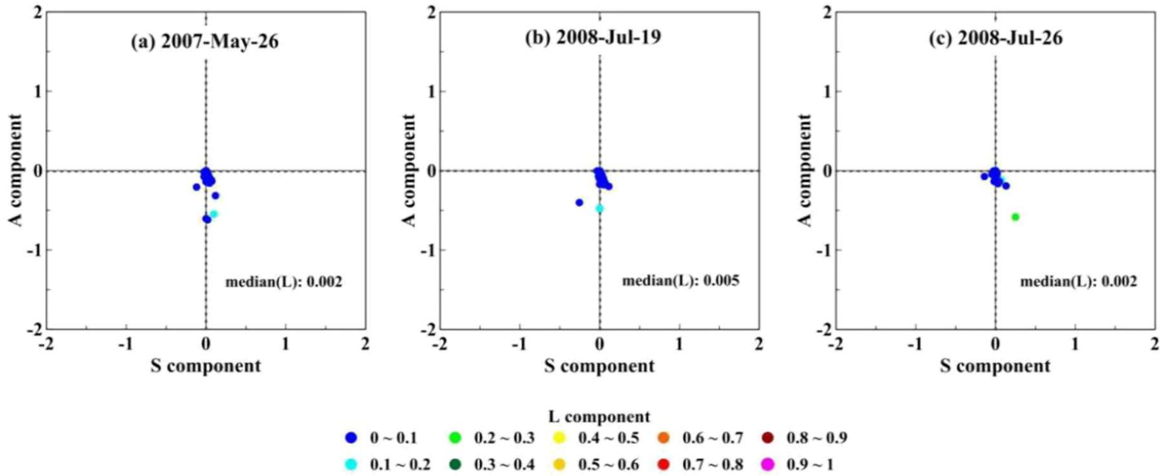


Figure 5. Value distributions of the three SAL components. Dashed lines in the vertical and horizontal directions in each sub-figure represent the median values of the S and A components, respectively, and dot color represents the of L component value. Results are sorted by selected rain days: (a) May 26, 2007; (b) July 19, 2008; (c) July 26, 2008.

4.2. Application of RCIT Algorithm in Rainfall Event Analysis

There were 10,346 rain cells identified from the radar reflectivity maps. Descriptive statistics of their properties are given in Table 3. It was found that a high standard deviation existed for the areas of these rain cells, with values ranging from 9 to 18,734 km^2 (most were less than 38 km^2). For areal rainfall depth, values ranged from 0.36 to 8861 mm and a high standard deviation again existed. For the maximum intensity property, a high standard deviation (34.08 mm.h^{-1}) was also found; the median value was 2.83 mm.h^{-1} and the range of values was 0.48 to 397.75 mm.h^{-1}. Areal mean rainfall depth was from 0.04 to 4.4 mm.km^2. Eccentricity ranged from 0 to 1, with a median value over 0.5.

Table 3. Descriptive statistics of rain cell properties. Indexes used for the statistics are minimum value, maximum value, standard deviation, and median value.

Property	Statistical properties			
	Minimum value	Maximum value	Standard deviation	Median value
Area	9	18734	1391	38
Areal rainfall depth	0.36	8861	559.9	4.4
Max intensity	0.04	33.2	2.8	0.24
Areal mean rainfall depth	0.04	4.4	0.3	0.1
Eccentricity	0	0.99	0.17	0.84

Inner structures of the selected events were described by statistically analyzing the physical and geometric properties of the rain cells. RCIT simulation results indicated that the properties (e.g., area, areal rainfall depth, max intensity, and areal mean rainfall depth) of the identified rain cells presented a wide range of values. The shape of the rain cells was somewhat elliptical, with a median value over 0.5. Histograms of the $\log_{10}$-transformed rain cell properties are given as in Figure 6. To determine the best theoretical distributions describing the empirical distributions, a multi-goodness of fit testing (GOF) approach combined with the Akaike information criterion (AIC), the Bayesian information criterion (BIC), and the Kolmogorov–Smirnov (K–S) test methods was applied (see Appendix B). Figure 7 shows their fitted cumulative distributions. Empirical distributions of the $\log_{10}$-transformed properties (area, areal rainfall depth, maximum intensity, and areal mean rainfall depth) could be fitted with the generalized Pareto distribution (GPD) presented in Equation (6), and the extreme value distribution (EVD) was found to fit the eccentricity property shown in Equation (7).

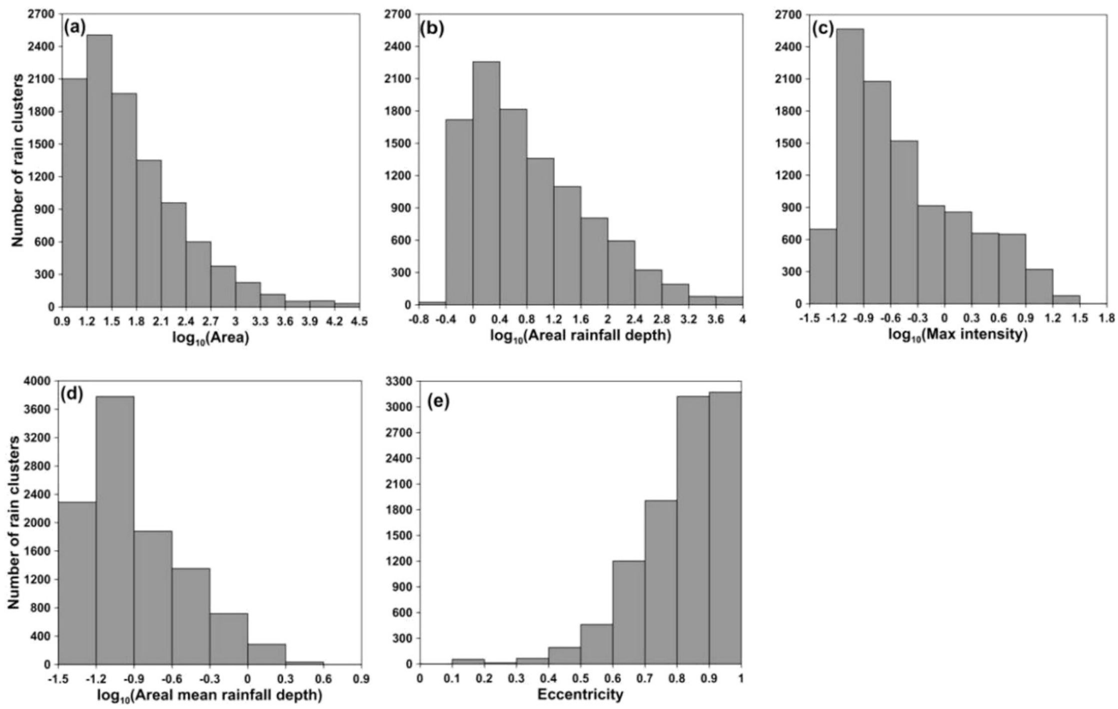
$$f(x|k,\mu,\alpha) = (\frac{1}{\alpha})(1 + k\frac{(x-\mu)}{\alpha})^{-1-\frac{1}{k}} \qquad (6)$$

where k is the shape parameter, and μ and α are location and scale parameters, respectively. For $\mu < 0$, k is above zero, and for $\mu < x < \alpha$, k is below zero. At the limit for k = 0, the GPD is the exponential distribution.

$$f(x|k,\mu,\alpha) = (\frac{1}{\alpha})e^{(-(1+k\frac{(x-\mu)}{\alpha})^{\frac{1}{k}})(1+k\frac{(x-\mu)}{\alpha})^{-1-\frac{1}{k}}} \qquad (7)$$

for $1 + k\frac{(x-\mu)}{\alpha}$, when k > 0, the generalized EVD is the Frchet distribution; k < 0 corresponds to the Weibull distribution; at the limit for k = 0, it is the Gumbel distribution.

A total of 1,107 rain cell trajectories were exploited. Histograms of their duration and motion vectors are presented in Figure 8. All rain cells held a mean duration of 40 minutes. For all the identified rain cells, the median value of their life cycles was 15 minutes, with an average moving speed of 11.59 m.s^{-1}. The moving directions of the rain cells were consistently toward the direction of motion observed in the radar images.



Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 August 2019 doi:10.20944/preprints201908.0011.v1

Peer-reviewed version available at Atmosphere 2019, 10, 532; doi:10.3390/atmos10090532

Figure 6. Histograms of $\log_{10}$-transformed rain cell properties: (a) area, (b) areal rainfall depth, (c) maximum intensity, (d) areal mean rainfall depth and property, (e) eccentricity for identified rain fields.

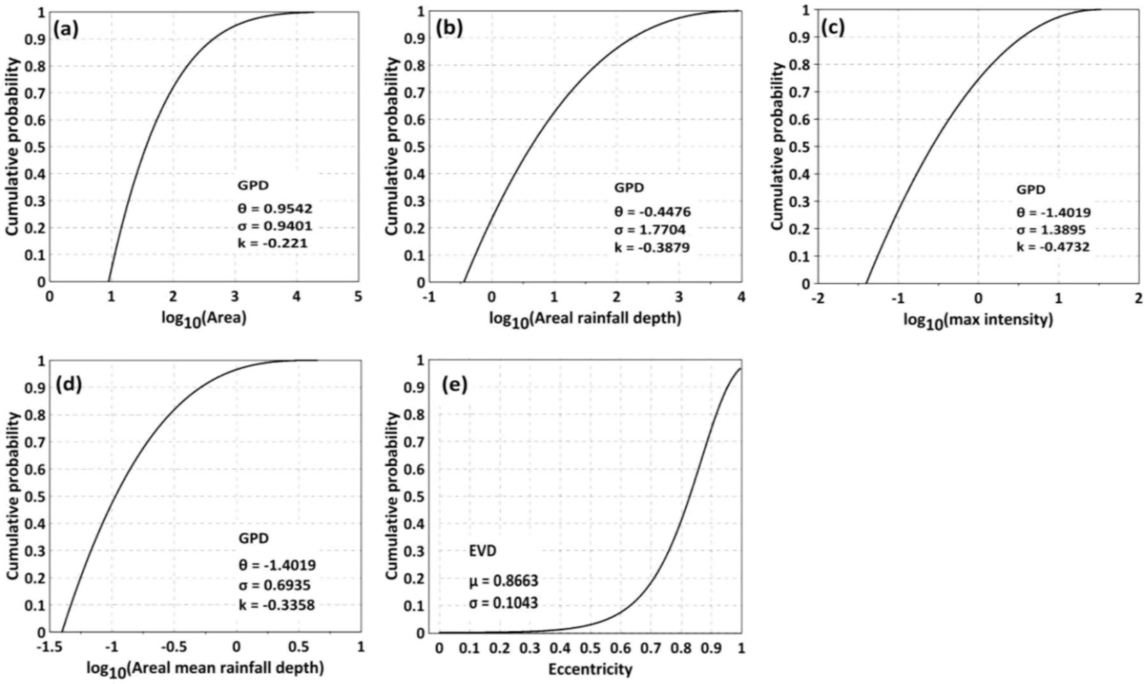


Figure 7. Cumulative curves of fitted probability density functions for $\log_{10}$-transformed rain cell properties: (a) area, (b) areal rainfall depth, (c) maximum intensity, (d) areal mean rainfall depth and property, (e) eccentricity.

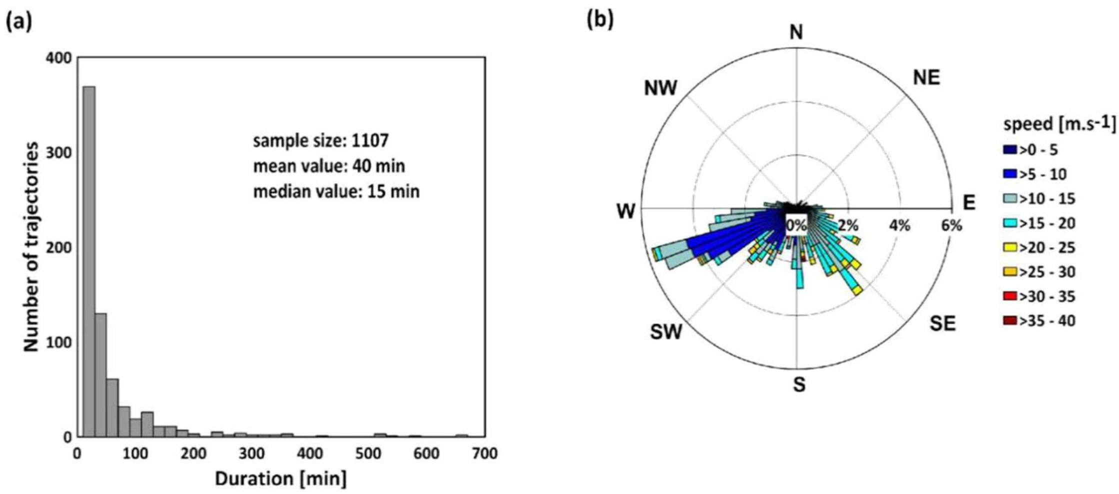


Figure 8. (a) Histograms of rain cell duration for identified rain fields, (b) wind rose plot of rain cell motion estimation result.

These results were in agreement with the study of Barnolas et al. (2010), in which the structures of heavy rainfall events recorded in Catalonia, Spain were analyzed. However, the results differed from those of Karklinsky and Morin (2006), in which the area of identified rain cells in southern Israel was better fitted to the log-normal distribution. Statistical analysis of the rain cell properties suggests that the inner structures of the selected rainy days can be expressed by the EVD. This suggests that most rainfall events had a limited covering area with less intensity and short duration; rainfall events with a long duration had a large covering area and high intensity.

During the life cycle of a rainfall event, the physical and geometrical features of rain cells continually change. Three common stages reflect these variations: cumulus, mature, and dissipating (Byers and Braham Jr 1948). In fact, the stage changes of rain cells are not only associated with its internal growth and decay but also with outer rain cells (e.g., merging or splitting). In this study, seven rain cell life stages were confirmed by the RCIT algorithm; their definitions can be listed as follows (Figure 9):

a. Initial: A rain cell having no parent cell is termed an initial rain cell.

b. Tracking: A rain cell with only one parent cell and having no interaction with other rain cells during its life cycle is termed a tracking rain cell.

c. Merge: A rain cell with at least two parent cells is termed a merged rain cell.

d. Split: A rain cell with only one parent cell but at least two child cells is termed a split rain cell.

e. Dissipate: A rain cell with at least one parent cell but no child cells is termed a dissipate rain cell.

f. 5-minute life cycle: A rain cell with a life cycle of only 5 minutes.

g. Complex stage: A rain cell for which merging and splitting simultaneously exist during its life cycle is termed complex stage.

The number of rain cells with different stages over their life cycles is summarized in Table 4. Rain cells with "5 minutes life cycle" and "tracking" stages were dominant. The "merging", "splitting", and "complex stage" cells occurred in isolated cases, indicating a stable inner structure of the identified cells. For the cases of July 19, 2008 and July 26, 2008, the number of rain cells in the "tracking" stage was even greater, indicating that the rainfall events occurring on these two days had long durations. The number of rain cells in the "merging" and "splitting" stages was greater in the July 26, 2008 case. This suggests that there were more convective rainfall events on that day since rain cells merge or split more frequently under such conditions.

Table 4. Number of rain cells with different life stages, sorted by selected rain day.

Stages	May 26, 2007	July 19, 2008	July 26, 2008
Initial	158	350	471
Tracking	608	1270	1787
Merge	7	6	39
Split	1	2	5
Dissipate	152	346	434
5 minute life cycle	632	3148	929
Complex stage	0	0	1

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 August 2019 doi:10.20944/preprints201908.0011.v1

Peer-reviewed version available at *Atmosphere* **2019**, *10*, 532; doi:10.3390/atmos10090532

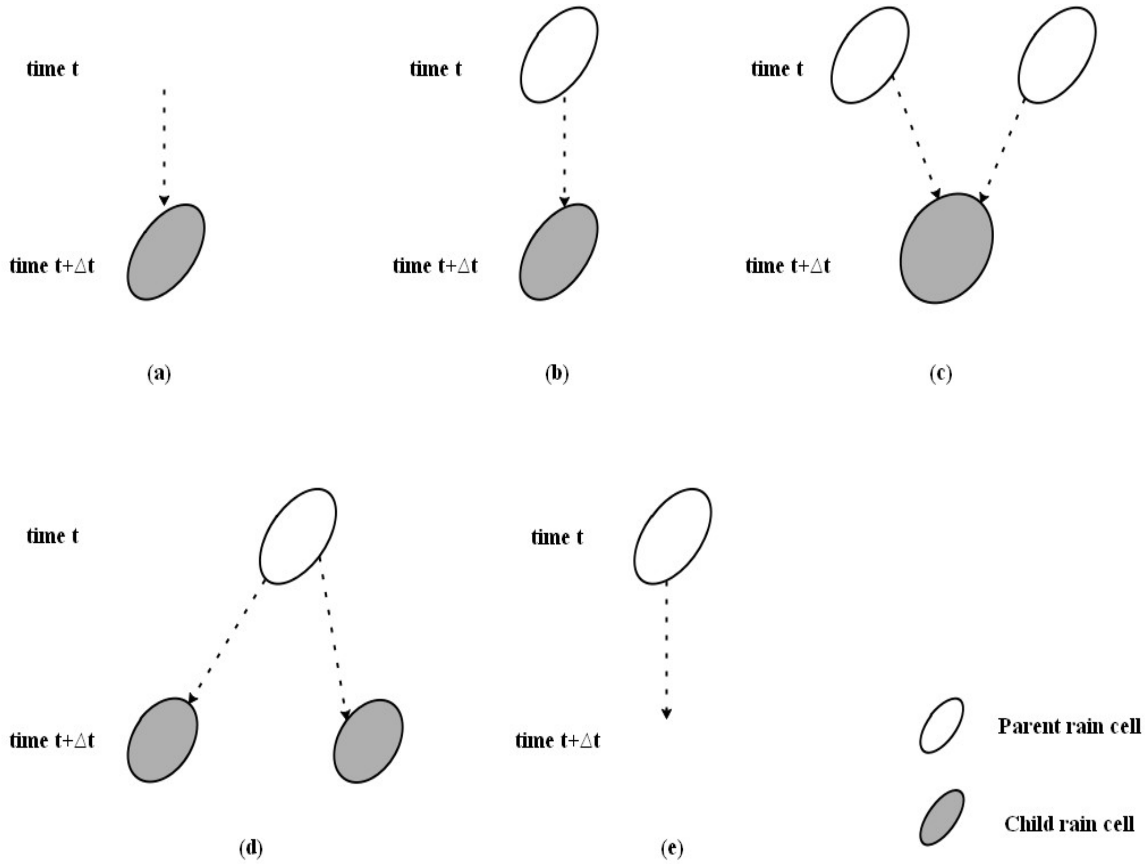


Figure 9. Stage definitions of rain cells: (a) initial, (b) tracking, (c) merging, (d) splitting, (e) dissipating.

5. Conclusion and Outlook

This study develops a new algorithm, RCIT, which utilizes the advantages of high resolution weather radar data. The proposed algorithm provides the following improvements:

1. It uses the PIV method in rain cell motion estimation. Rain cell motion estimation by past algorithms is mainly based on the maximum correlation coefficient method, which may lead to nonconsecutive motion when the shape and volume of a rain cell change rapidly. The PIV method avoids this situation.

2. A rain cell matching rule is proposed to discern the life cycle and stage change of rain cells. Past algorithms focus mainly on the tracking of rain cells without merging and splitting, when in fact rain cell stage variation is obvious over their life cycle, especially for convective rainfall events. The proposed rain cell matching rule implemented in the RCIT algorithm can easily and effectively discern the various stages of rain cells.

Two feature-based verification methods, SAL and geometric index, were used to test the performance of the RCIT algorithm. It is shown that all verification indexes fall within in a reasonable error range, confirming the good performance of the RCIT algorithm. Practical applications of the RCIT algorithm in analyzing the inner structure of historical rainfall events that occurred in the NRW are presented. This is the first time that the use of such a RCIT algorithm to depict the inner structures of rainfall events in an urban region with a high population density has been presented. The results show that the properties of rain cells in this region presented an EVD, indicating that the selected rainfall events had a short duration with low intensity. Long duration events with high intensity are rarely found and the stage changes of rain cells vary between events.

It should be noted that inputs for the proposed algorithm is not limited to radar data; other 2-D remote sensing data will also be used as the algorithm inputs, suggesting the versatility of the

Peer-reviewed version available at Atmosphere 2019, 10, 532; doi:10.3390/atmos10090532

proposed algorithm. In future application, it is intended that this algorithm will analyze the spatial–temporal variation of rainfall in small regions; this will lead to the determination of rainfall inputs with proper spatial and temporal scales for hydro-meteorological applications. The proposed algorithm will also be applied to rainfall nowcasting, which will improve the foresight period of flash floods in mountainous and urban regions. In addition, the features of the rain cell output from this algorithm can be used in sensitivity analyses of urban runoff in relation to short-term rainfall events, which will improve flood forecasting precision in small–medium catchments.

Acknowledgments: Authors would like to thank WUPPERVERBAND (water management in Wupper area) for radar data support and to thank Dr. Thomas Einfalt from Hydro&Meteoro Co Ltd for the guidance about radar data quality control. The financial support for this study was made available by the International Science & Technology Cooperation Program of China (Grant No: 2012DFG22140).

APPENDIX A

Structure–Amplitude–Location (SAL) and Geometric Index Verification Methods

Two feature-based verification methods, structure–amplitude–location (SAL) and geometric index, were applied to evaluate the performance of the RCIT algorithm and nowcasting methods. In SAL, term structure denotes the similarity in the shapes of modeled and observed rain cells; its value range is from −2 to 2. Amplitude denotes the similarity in the total precipitation values of modeled and observed rain cells; its value range is from −2 to 2. Location denotes the similarity of the center of mass for the modeled and observed rain cells; its value varies from 0 to 2. The accuracy of nowcasting methods can be evaluated based on the value of the three SAL components and a perfect nowcasting is confirmed by S, A, and L values of 0. More details on the SAL method can be found in Wernli et al. (2008).

Geometric index is a quantitative assessment method for the spatial patterns of rain cells (AghaKouchak et al. 2010). It compares the geometric features of modeled and observed rain cells via three indexes:

- Connectivity index: This is defined to compare simulated rain cells with respect to a reference object (e.g., observed rain cells). Its value is calculated based on the number of rain cells (NC) and the total number of non-zero pixels or pixels above a given threshold (NP), as in Equation (8):

$$C_{index} = 1 - \frac{NC-1}{\sqrt{NP}+N} \quad (8)$$

where C_{index} is the connectivity index, NP is the number of rainy pixels above a given threshold, and NC is the number of rain cells.

- Shape index: This index is introduced to quantitatively describe the shape discrepancy of rain cells, as in Equation (9):

$$S_{index} = 1 - \frac{P_{min}}{P} \quad (9)$$

where S_{index} is the single index, P_{min} is the theoretical minimum perimeter, and P is the actual perimeter of the rain cell.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 August 2019 doi:10.20944/preprints201908.0011.v1

Peer-reviewed version available at *Atmosphere* **2019**, *10*, 532; doi:10.3390/atmos10090532

- Area index: This is defined to depict the dispersiveness between the modeled and observed rain cells. Its value is the ratio of the area of its convex hull (the boundary of the minimal convex set containing a finite set of points in the rain cell), as in Equation (10).

$$A_{index} = 1 - \frac{A}{A_{Convex}} \quad (10)$$

where A is the area of the rain cell and A_{Convex} is the area of the convex hull.

APPENDIX B

Goodness of fit testing for fitted distributions of rain cell properties

The GOF test determines whether a data set is well fitted with a predefined distribution that gives the highest probability of producing the observed data. As such, a series of fit testing methods was developed, with the commonly applied tests as follows:

- The K–S test is based on the empirical cumulative distribution function (ECDF). Given N ordered data points Y_1, Y_2, …., Y_n, their ECDF is defined as:

$$E_N = \frac{n(i)}{N} \quad (11)$$

where n(i) is the number of points less than Y_i and Y_j, which are ordered from the smallest to largest value. This is a step function that increases by 1/N at the value of each ordered data point. The K–S test was developed according to the following hypotheses: H_0—the data follow a specified distribution; H_1—the data do not follow the specified distribution.

- AIC (Akaike, 1998) is based on the use of Kullback–Leible information as the discrepancy measure between the true distribution and the approximating distributions: Mi = gi(x,p_1,p_2,…,p_n). The AIC for the ith candidate distribution can be computed as:

$$AIC = -2\prod(\theta) + 2p \quad (12)$$

where $\prod(\theta)$ stands for the maximum log-likelihood of the sample of the dataset, p is the parameter's number of candidate distributions when the sample size n is small with respect to the number of the estimated parameter Pi. The smaller the value of AIC, the better fitting is the result for the candidate distribution.

- BIC (Schwarz, 1978) serves as an asymptotic approximation to a transformation of the Bayesian posterior probability of a candidate model. It is based on the empirical log-likelihood and does not require the specification of priors. BIC is defined as

$$BIC = -2\prod(\theta) + \ln(n)p \quad (13)$$

where the symbols are the same as those In Equation (12). A small value of BIC means that the candidate distribution fits well with the empirical distribution.

References

Adrian R. 2005. Twenty years of particle image velocimetry. Exp Fluids, 39: 159–169

AghaKouchak A, Nasrollahi N, Li J, Imam B, Sorooshian S. 2010. Geometrical characterization of precipitation patterns. J Hydrometeorol, 12: 274–285

Akaike H. 1998. Information theory and an extension of the maximum likelihood principle. In: Parzen E, Tanabe K, Kitagawa G, eds. Selected Papers of Hirotugu Akaike. Springer New York. Springer Series in Statistics, 199–213

Anna D M, Tomeu R, Carmen L M. 2018. A radar-based centroid tracking algorithm for severe weather surveillance: Identifying split/merge processes in convective systems. Atmos Res, 213: 110–120

Anoraganingrum D. 1999. Cell segmentation with median filter and mathematical morphology operation, in: Image Analysis and Processing. 1999. Proceedings. International Conference on Image Analysis & Processing. 1043–1046

Barnolas M, Rigo T, Llasat M C. 2010. Characteristics of 2-d convective structures in catalonia (ne spain): An analysis using radar data and gis. Hydrol Earth Syst Sci, 14: 129–139

Casati B, Wilson L J, Stephenson D B, Nurmi P, Ghelli A, Pocernich M, Damrath U, Ebert E E, Brown B G, Mason S. 2008. Forecast verification: Current status and future directions. Meteorol Appl, 15: 3–18

Davis C, Brown B, Bullock R. 2006. Object-based verification of precipitation forecasts. Part I: Methodology and application to mesoscale rain areas. Mon Weather Rev, 134: 1772–1784

Dixon M, Wiener G. 1993. Titan: Thunderstorm identification, tracking, analysis, and nowcasting—A radar-based methodology. J Atmospheric Ocean Technol, 10: 785–797

Ebert E, Wilson L, Weigel A, Mittermaier M, Nurmi P, Gill P, Gober M, Joslyn S, Brown B, Fowler T, Watkins A. 2013. Progress and challenges in forecast verification. Meteorol Appl, 20: 130–139

Ebert E E. 2008. Fuzzy verification of high-resolution gridded forecasts: A review and proposed framework. Meteorol Appl, 15: 51–64

Einfalt T, Denoeux T, Jacquet G. 1990. A radar rainfall forecasting method designed for hydrological purposes. J Hydrol, 114: 229–244

Elena C, Marie-Claire T V, Nick V D G. 2017. Spatial and temporal variability of rainfall and their effects on hydrological response in urban areas—a review. Hydrol Earth Syst Sci, 21: 3859–3878

Féral, Laurent, Sauvageot H, Castanet L, Lemorton J, Cornet, Frédéric, Leconte K. 2006. Large-scale modeling of rain fields from a rain cell deterministic model. Radio Sci, 41

Frerk I, Treis A, Einfalt T, Jessen M. 2012. Ten years of quality controlled and adjusted radar precipitation data for north rhine-westphalia–methods and objectives, in: 9th International workshop on precipitation in urban areas: Urban challenges in rainfall analysis. 6–9 December 2012, St. Moritz, Switzerland

Gilleland E, Ahijevych D, Brown B G, Casati B, Ebert E E. 2009. Intercomparison of spatial forecast verification methods. Weather Forecast, 24: 1416–1430

Gui L, Merzkirch W. 1996. A method of tracking ensembles of particle images. Exp Fluids, 21: 465–468

Guinard K, Mailhot A, Caya D. 2015. Projected changes in characteristics of precipitation spatial structures over North America. Int J Climatol, 35: 596–612

Handwerker J. 2002. Cell tracking with trace3d-a new algorithm. Atmos Res, 61: 15–34

Heistermann M, Jacobi S, Pfaff T. 2013. Technical note: An open source library for processing weather radar data (wradlib). Hydrol Earth Syst Sci, 17: 863–871

Johnson J T, MacKeen P L, Witt A, Mitchell E D W, Stumpf G J, Eilts M D, Thomas K W. 1998. The storm cell identification and tracking algorithm: An enhanced wsr-88d algorithm. Weather Forecast, 13: 263–276

Karklinsky M, Morin E. 2006. Spatial characteristics of radar-derived convective rain cells over southern Israel. Meteorol Z, 15: 513–520

Li L, Schmid W, Joss J. 1995. Nowcasting of motion and growth of precipitation with radar over a complex orography. J Appl Meteorol Climatol, 34: 1286–1300

Muñoz C, Wang L P, Willems P. 2018. Enhanced object-based tracking algorithm for convective rain storms and cells. Atmos Res, 201: 144–158

Merzkirch W. 2001. Particle image velocimetry, in: Mayinger, F., Feldmann, O. (Eds.), Optical Measurements. Springer Berlin Heidelberg. Heat Mass Transf., 341–357

Morin E, Yakir H. 2014. Hydrological impact and potential flooding of convective rain cells in a semi-arid environment. Hydrolog Sci J, 59: 1353–1362

Moseley C, Berg P, Haerter J O. 2013. Probing the precipitation life cycle by iterative rain cell tracking. J Geophys Res Atmos, 118: 13, 361–13, 370

Moseley C, Haerter J, Berg P, Eggert B. 2013. EGU General Assembly Conference Abstracts, 15, EGU2013–11380

Novo S, Mart´ınez D, Puentes O. 2014. Tracking, analysis, and nowcasting of Cuban convective cells as seen by radar. Meteorol Appl, 21: 585–595

Peleg N, Morin E. 2012. Convective rain cells: Radar-derived spatiotemporal characteristics and synoptic patterns over the eastern mediterranean. J Geophys Res Atmos, 117

Peleg N, Blumensaat F, Molnar P, Fatichi S, Burlando P. 2017. Partitioning the impacts of spatial and climatological rainfall variability in urban drainage modeling. Hydrol Earth Syst Sci, 21: 1559–1572

Rinenart R E, Garvey E T. 1978. Three-dimensional storm motion detection by conventional weather radar. Nature, 273: 287–289

Schwarz G. 1978. Estimating the dimension of a model. Ann Stat, 6: 461–464

Wernli H, Paulat M, Hagen M, Frei C. 2008. Sal-a novel quality measure for the verification of quantitative precipitation forecasts. Mon Weather Rev, 136: 4470–4487

Westerweel J, Elsinga G E, Adrian R J. 2013. Particle image velocimetry for complex and turbulent flows. Annu Rev Fluid Mech, 45: 409–436

Weusthoff T, Hauf T. 2008. The life cycle of convective-shower cells under post-frontal conditions. Q J Royal Meteorol Soc, 134: 841–857

Yeung J K, Smith J A, Baeck M L, Villarini G. 2015. Lagrangian analyses of rainfall structure and evolution for organized thunderstorm systems in the urban corridor of the northeastern United States. J Hydrometeorol, 16: 1575–595

Zahraei A, Lin Hsu K, Sorooshian S, Gourley J J, Hong Y, Behrangi A. 2013. Short-term quantitative precipitation forecasting using an object-based approach. J Hydrol, 483: 1–15

LIST OF TABLES

LIST OF FIGURES

Peer-reviewed version available at Atmosphere 2019, 10, 532; doi:10.3390/atmos10090532

Figure 6. Histograms of log_{10}-transformed rain cell properties: (a) area, (b) areal rainfall depth, (c) maximum intensity, (d) areal mean rainfall depth and property, (e) eccentricity for identified rain fields.

Figure 7. Cumulative curves of fitted probability density functions for log_{10}-transformed rain cell properties: (a) area, (b) areal rainfall depth, (c) maximum intensity, (d) areal mean rainfall depth and property, (e) eccentricity.

Figure 8. (a) Histograms of rain cell duration for identified rain fields, (b) wind rose plot of rain cell motion estimation result.

Figure 9. Stage definitions of rain cells: (a) initial, (b) tracking, (c) merging, (d) splitting, (e) dissipating.