Radar Based Quantitative Precipitation Estimation and Forecasting

a Case Study in North Rhine Westphalia, Germany

Ting He

Dissertation at the Department of Earth Sciences

Freie Universität Berlin

This dissertation is submitted for the degree of

Doctor of Nature Science

January 2016

Erstgutachter: Prof.Dr. Achim Schulte

Zweitgutachter: Prof.Dr. Karl Tilman Rost

Date of defense: December 11th, 2015

CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Background

Flood is a natural phenomenon with a prompt increase of water volumes from the drainage system, ponds or reservoirs. Anomaly weather situations, underlying surface conditions and human activities are the primary reasons for flood's generation. In urban catchments, flood usually occurs following the instantaneous rainfall events with high intensity. Rapid response of the stream flows in drainage networks can reach to peak level in a few minutes which leaves a short warning time for prevention operated. Main features of urban catchment's flood are concluded by Georgakakos (1986) and listed as follows :

- Heavy precipitation that persists over an area from minutes to few hours;
- Rapid runoff generating because of the steep slope in catchment;
- Impervious surface and saturated soil conditions limit streams infiltrate into subsurface;
- Sudden release of impounded water from sewer systems because of the fast increased water level which exceeds the warning line rapidly.

Urban regions with a density population and buildings have only a size within few hundreds of square kilometers or even smaller. Being compared to the rural ones, stream flows are more regularly in urban catchments because of more artificial channels, which also can lead the runoff to the peak value rapidly after one or a series of continuous and intensive rain events. Simultaneously, owed to steep slopes and saturated soil conditions in catchment, more streams spill over river channels and submerge facilities with more economic and social damages.

In the region of Europe, flood events occur more frequently and intensively in recent years , and these events show obviously seasonal and regional variations (Marchi et al., 2010):

- Floods occurred in the Mediterranean and Alpine-Mediterranean region implode in autumn season and in inner Europe (Central Europe and Eastern Europe), floods are prone to occur in summer season;
- Consistently with the seasonality effect, spatial extent and duration of the flood events is generally smaller for the continental events with respect to those occurring in the Mediterranean region.

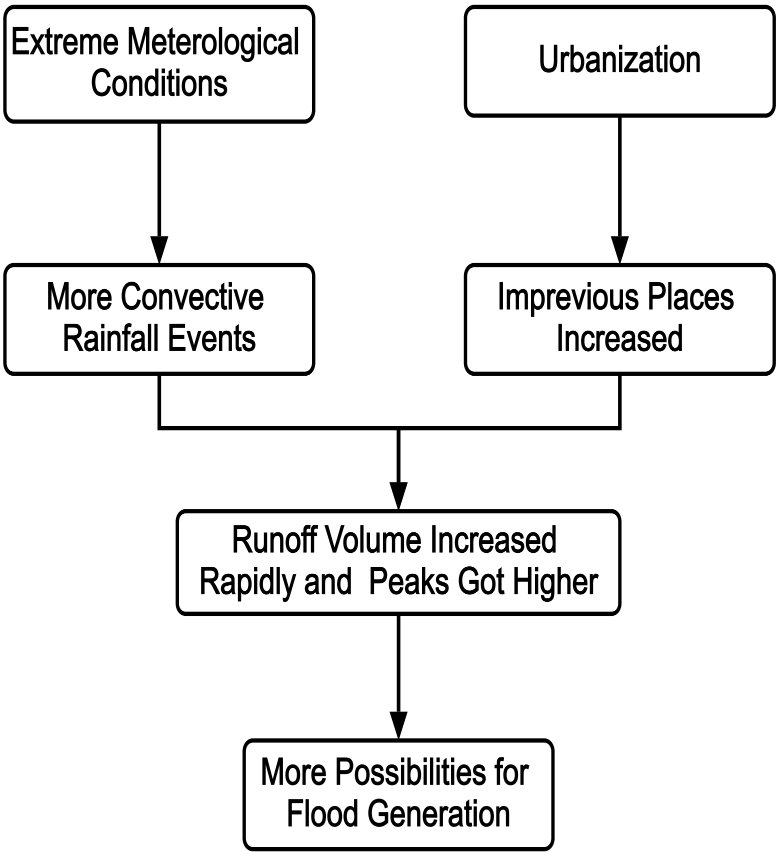


Figure 1.1 Main factors for leading to flood in urban catchment.

Reasons for resulting in more flood events occurred in urban regions are that more convective rainfall events happen as a result of local extreme meteorological phenomenons and catchment's hydrology factors changing which is affected by the increasing urbanization in recent years. Figure 1.1 presented the flood impacts from interactions among the extreme meteorological conditions, hydrology factors and urbanization.

Runoff generation in urban catchments is a complex process, which involves not only hydro-meteorological interaction but also hydro-hydraulic process and human effects. For the purpose of credible urban runoff simulation and flood forecasting, it is essential to provide a confidential approach for spatial and temporal precipitation estimation and forecast (e.g. Bruni et al., 2015; Gires et al., 2012; Wright et al., 2013). Considering the key factors for flood events generation in urban region, quantitatively estimating and forecasting of precipitation is the main researching direction of this thesis.

1.2 Precipitation - One Driving Phenomenon of Urban Hydrological Mechanisms

Among all the elements which can trigger urban floods, rainfall is the most direct and important one. As a driving phenomenon of runoff mechanisms, variability of rainfall constitutes a significant source of uncertainty in hydrological applications. Rainfall events can be described as a set of rain fields (or named as cell in convective conditions) that are organized by some spatial - temporal characteristics (e.g. shape, area, rain rate, movements). Singh (1997) observed that the effect of convective rain field's movements are significant for urban catchments, runoff can reach to a higher peak value if such rain fields move along the same direction with the stream flows and if the rain field's moving speed is the same with flow velocity, the discharge is more affected. Yakir and Morin (2011) investigated the hydrological response to the characteristics of convective rain fields in a small semi-arid watershed, the rain fields observed from weather radar was modeled for reproducing the spatial - temporal rainfall characteristics, and then were invited to a hydrological model for testing the runoff response to the inner structures of convective rainfall events. The authors found that the hydrological response in semi-arid catchment is mainly sensitive to the location and moving speed of convective rain fields which can produce higher runoff peaks.

For credible runoff modeling purpose, spatial - temporal precipitation information should be available to the hydrological model. These information needs to be quantified at various resolutions to be compatible with the model resolution. Mácaa and Torfs (2009) analyzed two flood events happened in a small catchment (less than 50 km^2) by a method which coupled a deterministic event based runoff model with a stochastic rainfall disaggregation model and proved that the runoffs in such a catchment is very sensitive to the temporal variability of precipitation. Liang et al. (2004) tested the effect of the spatial resolution of precipitation to the qualities of model calibration in a meso-scale watershed and found that a better calibration can be obtained with a critical spatial scale of 14 km, the modeling uncertainties start to be significant when this resolution is lower than 14 km. Vischel and Lebel (2007) compared the effects of different spatial resolutions of precipitation to the runoff response and found that the higher resolution scale for precipitation information can bring better predicting results for peak runoff.

For credible flood forecasting purpose, the runoff prediction by the hydrological model is a key. However, uncertainties from the rainfall inputs limit our abilities to extend the runoff's forecast leading time despite a good calibration for model itself (Singh and Woolhiser, 2002). Availability of rainfall forces at small scales as a requirement for runoff forecast is restricted by the limits of precipitation's forecasting ability. The most common and convenient way to obtain rainfall information is from gage observations. For many floods warning operations in meso-scale to macro-scale basins, a flood forecasting system based only on observed precipitation data might be all that is needed. For flood warnings in urban catchments, however, this may not give a sufficient leading time for operational warnings so that some prediction of future rainfalls will be required (e.g. Rossa et al., 2011; Zappa et al., 2010). Bárdossy and Das (2008) suggested that using too coarsely gage networks for estimating precipitation can lead to a poor simulating result and the hydrological model needs recalibration when different rain gage networks are used. Cooper and Fernando (2009) discussed that, if the rain gage's density is too low, confidence in the accuracy and ability of the model as a tool to predict peak runoff is diminished in the urban catchment. Decreased rain gage's density has a detrimental effect on the ability of a hydrological model for predicting the peak runoff as well as the shape of the runoff hydro-graph. Mishra (2013) tested the density effects of gages to the accuracies of rainfall estimating results and the author concluded that when the gage spacing intervals increased, errors of rainfall estimation were also obvious, and homogeneous distribution of rain gages which contain the sufficient number of equally spaced gages form a perfect network can monitor the rainfall more accurately over one region.

Another approach is the application of weather radar, which can acquire rainfall data at high spatial and temporal resolutions in hydro-meteorological applications. Rainfall inputs derived from weather radar can improve the accuracy of modeled peak runoff efficiently (e.g. Lopez et al., 2005; Méndez-Antonio et al., 2013). However, uncertainties from radar based rainfall estimation and forecast also exists, and most of these uncertainties are still unknown which are needed deeply investigated (Germann et al., 2009). Providing superior quality's spatial - temporal rainfall inputs is the key for improving the quality of runoff modeling and forecast in urban catchments.

1.3 Variability of Rainfall Events

Acquiring explicit knowledges about spatial - temporal characteristics of precipitation object is critical to assisting people to understand the inner behaviors of rainfall events and their interactions with the catchment's hydrological responses. Here, the precipitation object is presented as a connected cluster with intensity exceeding a given threshold. In hydro-meteorological applications (e.g. precipitation forecast, rainfall-runoff analysis and flood forecast), spatial rainfall inputs are generally calculated by applying some spatial interpolation methods (e.g. Krigin, Thiessen polygon, Co-Krigin and Inverse-Distance-Weight method) and are commonly treated as a uniform distribution. However, the precipitation object always varies over its life cycle (e.g. their size, integral rainfall volume or shape). Particularly, for those larger watersheds (e.g. above hundreds km^2), these variations are more obvious. Modeling precipitation object for getting the detail information of rainfall event's inner structure is necessary.

Importance of analyzing spatial - temporal characteristics of precipitation object for enhancing the quality of hydrological applications has been highlighted in some studies. Smith et al. (2004) used two spatial indexes of precipitation objects for testing the runoff's response to a trivial filtering watershed. Morin et al. (2006) applied a conceptual precipitation object tracking algorithm for representing spatial and temporal patterns of precipitation objects in thunderstorm events and used their characteristics as the input for runoff simulation, they also investigated that the high-peak runoff occurred in a semi-arid watershed is very sensitivity to the small changes of precipitation object's characteristics.

An essential problem cannot be ignored, as a micro physical phenomena, the circumstance of the cloud and its form of precipitation is still not fully understood, which makes it hard for an accurate description of precipitation objects. The common idea for precipitation object's spatial and temporal analysis can go back to the early algorithm which is proposed by Byers and Braham Jr (1948) who followed the 'Thunderstorm Project' in the late 1940s. They defined three stages for thunderstorm's cell over its life cycle: a towering cumulus stage, a mature stage, and a dissipating stage as in Figure 1.2.

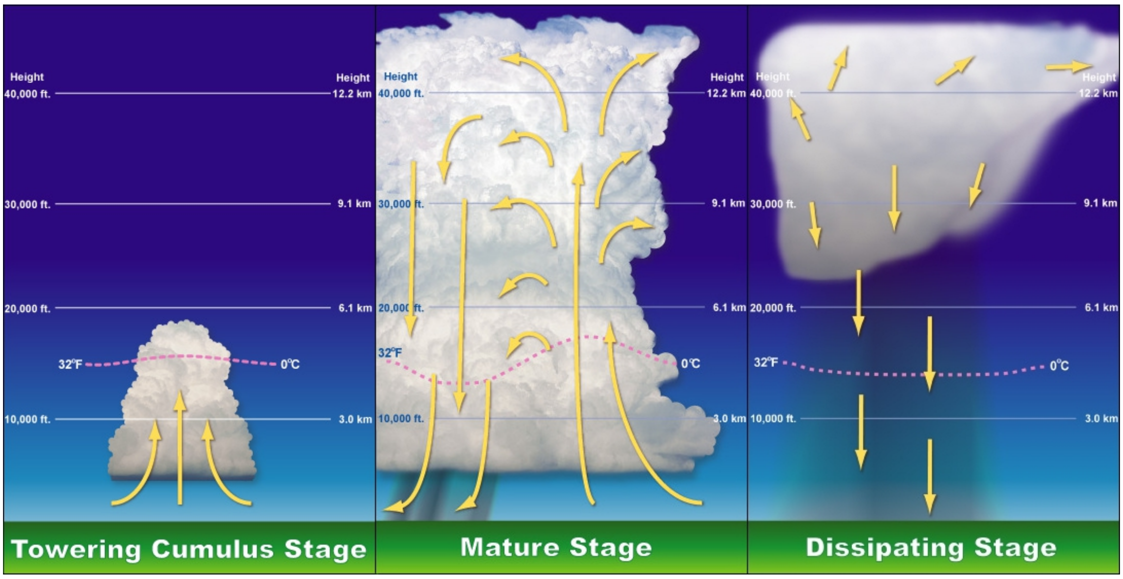


Figure 1.2 Life stages of convective cells, image is from Weather in National Weather Service (NWS) (2014).

The authors also recognized the formation of individual convective precipitation objects in the meso-scale system. Although not explicitly dealt with at that time, it is lately shown that this implied interactions among discrete precipitation objects that may have consequences for the convective evolution. Bluestein (1992) gave a historical overview of field programs dedicated to the study of severe convective storms. With the help of such field measurements and numerical modeling, the internal structures as well as the evolution of convective clouds and precipitation systems have been widely investigated over the past fifty years. Three stages of thunderstorm's cells over its life cycle are briefly introduced:

Cumulus stages: Characterized by vertical updraft within a precipitation object. The development of each cluster begins from the cumulus cloud stage. Cumulus cloud is a puffy cloud that can be readily recognized (left part of Figure 1.2). In this stage, the temperature inside the precipitation object is higher than in the external environment guaranteeing the uplifting of air. The greatest buoyancy forces are formed at upper levels of it, where the greatest temperature differences also occur. This is naturally since the buoyancy is proportional to the density differences in the fluid and thereby proportional to temperature differences.Precipitation can be observed inside the precipitation object, especially above the freezing level, which may happen in liquid, solid or both. However, as the updraft is carrying precipitation upwind, rain is not complied with on the ground at this stage.

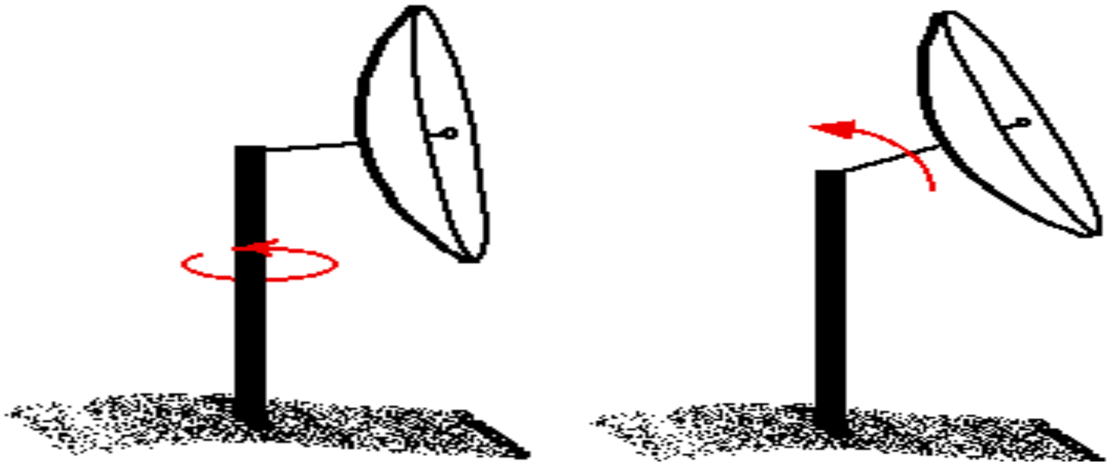
Mature stage: Both updrafts and downdrafts exist. While the air inside the convective precipitation object contains ascending, more and moister condensates forming visible water particles. This is followed by rain particles and above the freezing level snow and hail. Finally, the size of the mass and gravitation of particles exceeds the forces of flow dragging particles up, and these particles start falling down relative to the earth: the cell has reached the mature stage as illustrated in middle part of Figure 1.2. Strong updraft and hails amplifies the electrification of the convective precipitation object and therefore lighting is intense in the mature phase.

Dissipating stage: Precipitation object is featured by down-drafts. Finally causing the cluster to die out. In the dissipating stage, the ground precipitation diminishes until the final residual drops have fallen into the ground. Falling rain and evaporation cool air inside the precipitation objects and contributes to the dissipation. Finally, the vanishing updraft turns into a downdraft (right part of Figure 1.2), which spreads throughout the precipitation object's body. Since no updrafts occur in the dissipating phase, storm electrification and lighting decrease and disappear consequently. At the end, all that is left is an anvil-shaped cloud in the upper atmosphere consisting of crystallized ice.

1.4 Radar Fundamentals

Radar, the abbreviation of 'Radio Detection And Ranging', is a kind of active remote sensing equipment that is able to detect, track and imaging objectives in the atmosphere. It generates electromagnetic signals and emits them into the atmosphere in the form of beam, then receives the echoes from objects, after software's processing, users can get information such as distance between objects and signal emitting point, radial velocity, position and altitude.

Electromagnetic signals are emitted in the form of beam by antenna which is usually wrapped in a radome. Beam transfers along different directions with azimuth and vertical rotation of radar antenna. Three common scanning ways are used in weather radar system. Figure 1.3 presented two of them.



(a) PPI **(b)** RHI

Figure 1.3 Two scanning strategies of weather radar, image source:http://ww2010.atmos.uiuc.edu/.

- **PPI** (Plan Position Indicator), when elevation angle of antenna is constant, radar rotates at azimuth direction, it can rotate through 360^o (surveillance scan) or less than 360^o (sector scan). A conical area is presented by this scanning way, and is projected into a horizontal plane at specified elevation angle.
- **RHI** (Range Height Indicator), reversely to PPI, when the azimuth direction is constant, radar moves along vertical axis and then forms a vertical profile on screen at a specified horizontal position;
- **CAPPI** (Constant Altitude Plan Position Indicator) is a horizontal cut through the atmosphere objects, a PPI volume scans at multiple angles is required to produce it. This is achieved by scanning the atmosphere objects from the lowest elevation up to cloud top. The CAPPI algorithm constructs CAPPI by interpolating in height and range to the selected CAPPI surface.

Band selection is a function of the trade offs between the range of reflectivity and its cost which vary as a function of the physics of rain attenuations. Three types radar with different wavelength and frequency are commonly applied for rainfall observation:

- **S** band weather radar operates on a wavelength of 8 - 15 cm and a frequency of 2 - 4 GHz. Because of its long ranging wavelength and frequency, S-band radar is not easily attenuated. This makes it useful for near and far range's weather parameter's observation. The disadvantage of S-band weather radar is that it requires a large antenna dish and a large motor to power it.

- **C** band radar operates on a wavelength of 4 - 8 cm and a frequency of 4 - 8 GHz. Because of its wavelength and frequency, the dish size of it does not need to be very large. This makes C-band radar affordable for installation, but is more easily attenuated for the emitted signals, so it is commonly used for short range (up to 200 km) weather parameter's observation. Proper frequency of C-band radar allows it to create a smaller beam width using a smaller dish.

- **X** band weather radar operates on a wavelength of 2.5 - 4 cm and a frequency of 8 - 12 GHz. Because of the smaller wavelength, the X-band radar is more sensitive and can detect smaller particles. It is usually deployed for researching the cloud development of atmosphere because of their good performance in tiny water particle's detection and it is also used to detect strati-form precipitation. X-band weather radar also attenuates very easily, so it is used only in very short-range (less than 100 km). Due to the small size of it, X-band weather radar can be easily carried by mobile structures.

Apart from these commonly applied types, other kinds of weather radar with different frequency and wavelength were summarized by Rinehart (1990) as in Table 1.1.

Table 1.1 Performing parameters of weather radar with dfferent bands.

Band	Frequency [GHz]	Wavelength [cm]
L	1 - 2	30 - 15
Ku	12 - 1	2.5 - 1.7
K	18 - 27	1.7 - 1.2
Ka	27 - 40	1.2 - 0.75
W	40 - 300	0.75 - 0.01

Radar was early used in the military field for monitoring enemies situations in the last 30s. After the World War II, because of the excellence in long distance detecting, large area covering, strongly penetrating and less affection from weather elements (e.g. fog and rain), radar has been broadly applied in meteorological scope for detecting atmospheric objects that cannot be easily done by the traditional monitoring tools. Particularly, because radar can obtain rain fields at high temporal resolution and finer reflection for its distribution of horizontal and vertical direction, scientists have applied radar for QPE and QPF at different temporal and spatial scales since the last 50s. Experienced over half century's development, nowadays, radar can provide superior resolution's rainfall data to satisfy various meteorological and hydrological requirements.

1.5 Radar Based Precipitation Estimation and Forecast

Quantitative Precipitation Estimation (QPE) is a procedure for acquiring rain amounts that have fallen at a location or across a region. Correspondingly, Quantitative Precipitation Forecast (QPF) is a procedure for assessing the amounts of precipitation accumulated at a future time in a specified area. Accurate QPE and QPF can improve the knowledge of people in understanding the principle of rainfall event's occurrence, and their relations with catchment's response in the past and future. Even more, they can provide satisfactory rainfall inputs in hydrological model for flood warning and hazard's mitigation. During the last decades, as an efficient tool, weather radar has provided useful rainfall information for promoting the accuracy of QPE and QPF. Numerous researches have been done with the aim of improving QPE and QPF qualities by using weather radar.

1.5.1 Radar Based Precipitation Estimation

The direct way for rainfall estimation is using ground observing data. However, rainfall is not continuous in space and time as we expected. Gauge depended rainfall estimating results can arouse high bias. Radar has advantage of wide observing range and short scanning time, so combing it with gage observations can efficiently reduce such bias. Common procedures for radar based quantitative precipitation estimation go on as follows:

Reflectivity factor is sampled for an unlocked volume in space by scanning at selected elevation angles. The small fraction of short pulse emitted from the antenna by a predefined scanning strategy is scattered and reflected when it hits the targets, then reflecting pulse is captured by radar receiving equipment. Returning mean power p_r received by radar is transformed into the reflectivity Z proportionally, equation (1.1) describes this transformation:

$$\mathbf{Z} = \frac{\mathbf{P_r}^2}{\mathbf{C}\left|\mathbf{k}\right|^2} \tag{1.1}$$

whereas C is a constant value which can be presented as in equation (1.2):

$$\mathbf{C} = \frac{\pi^3}{\mathbf{1024}\ln\mathbf{2}} \frac{\Pr\mathbf{h}\mathbf{G^2}\theta\emptyset}{\lambda^2} \left| \frac{\mathbf{m^2}-\mathbf{1}}{\mathbf{m^2}+\mathbf{1}} \right| \tag{1.2}$$

where λ is the wavelength of radar, m is the complex fraction index, θ and $\emptyset$ is the horizontal and vertical width of radar beam, r is the distance from radar to the targets (meteorological or non-meteorological objects), while k is the imaginary part of the refraction index which depends on target surface: $|k|^2$ is about 0.93 for rainfall drops at $0^o c$ and 0.176 for ice particles.

Ground clutter contamination is identified. Ground clutter is a form of contamination for radar reflectivity, when unexpected objects such as trees, high buildings or hills are close to radar transmitter, this interfere normal propagation and returning of radar beams.

A relationship for translating measured reflectivity to rain rate is applied. For building the relationship between the radar measured reflectivity and rain rate, a hypothesis is proposed that targets from radar detection are rain drops in the atmosphere. Based on this assumption, the relation between reflectivity and raindrop size distribution is presented as in equation (1.3):

$$\mathbf{Z} = \int_{\mathbf{0}}^{\infty} \mathbf{D^6 N(D) dD} \tag{1.3}$$

whereas Z (unit: mm^6mm^3) is the reflectivity factor, $N(D)dD$ is the average numbers of raindrops with spherical diameters are between D and $D+dD$ (unit: mm) in per unit of air. For smoothing the wide range variations of reflectivity, the reflectivity factor is usually presented in a logarithmic scale which is called effective reflectivity (unit: dBZ) whereas $1dBZ = 10log_{10}Z$. Neglecting other meteorological parameter's impact, the relationship between rain rate R (unit: $mm.hr^{-1}$) and raindrop size distribution is also presented as in equation (1.4):

$$\mathbf{R} = \mathbf{6}\pi \times \mathbf{10^{-4}} \int_{\mathbf{0}}^{\infty} \mathbf{D^3 V(D) N(D) dD} \tag{1.4}$$

whereas $V(D)$ is the terminal falling speed of raindrop from cloud top to ground that is associated with raindrop's diameter. Combing equation (1.3) and (1.4), relationship between reflectivity and rain rate is given as in equation (1.5):

$$\mathbf{Z} = \mathbf{aR^b} \tag{1.5}$$

Correcting procedures are used for bias reducing combined with the measurements from ground gages. Rainfall measuring bias from weather radar are mainly caused by:

- **Hardware errors**. Radar equipment exists some instability after long working time (e.g. emitting components and receiving components). According to Collier and Hardaker (1996), when the error from radar receiver reaches to $\pm 1dBZ$, the bias of estimated rainfall can increases 16 $\%$.

- **Attenuation of radar electromagnetic signal**. In equation (1.1), P_r is associated with the wavelength of radar, antenna gain and beam width. When the antenna size is fixed, P_r is inversely to the 4th wavelength, the shorter of the wavelength, the larger of P_r and the more powerful detecting capability of radar. But the power of radar electromagnetic signal is attenuated by air and precipitation. The attenuation of radar signal is related to rainfall intensity R (unit: $mm.hr^{-1}$) and the radar wavelength λ (unit: cm) which can be presented by attenuation coefficient μ (unit: $dB.km^{-1}$) and their relationship can be presented as in equation (1.6):

$$\mu = \mathbf{c}\mathbf{R}^{\mathbf{d}} \tag{1.6}$$

  and default values for parameters in this equation were presented in Table 1.2 (Olsen et al., 1978).

Table 1.2 Default parameter values for equation (1.6).

λ	**c**	**d**
3.2 cm	0.01	1.21
5 cm	0.001	1.05
10 cm	0.000343	0.97

- **Radar reflectivity changes at vertical direction**. Owing to the evaporation of rain droplets, atmospheric motion and phase changing effects, radar reflectivity exists obvious variation at the vertical direction. Deviation between radar and gage observation is more obvious as horizontal propagating path extending, and the space scattering of radar beam is enlarged as the extension of range.

- **Rain drops changing and parameter variation of reflectivity and rain rate equation**. According to equation (1.2) and (1.3), radar reflectivity and rain rate have strong connection with rain drop size. Reflectivity of convective rainfall and warm cloud precipitation is different because of different rain drop size for these two types of precipitation, this results in hard decision for selecting right parameter a and b for converting reflectivity to rain rate in different types of rainfall events, even in the same one. Einfalt et al. (2004) presented an simple illustration for the influence from different a and b selection in resulting rain rate's variation as in Figure 1.4.

  In the figure, $Z = 200\,R^{1.6}$ is the Marshall–Palmer relation which is suitable for general strati-form events, $Z = 250\,R^{1.2}$ is adopted to tropical convective systems and $Z = 300\,R^{1.4}$ is recommended for summer deep convective rainfall events or other non-tropical convections.

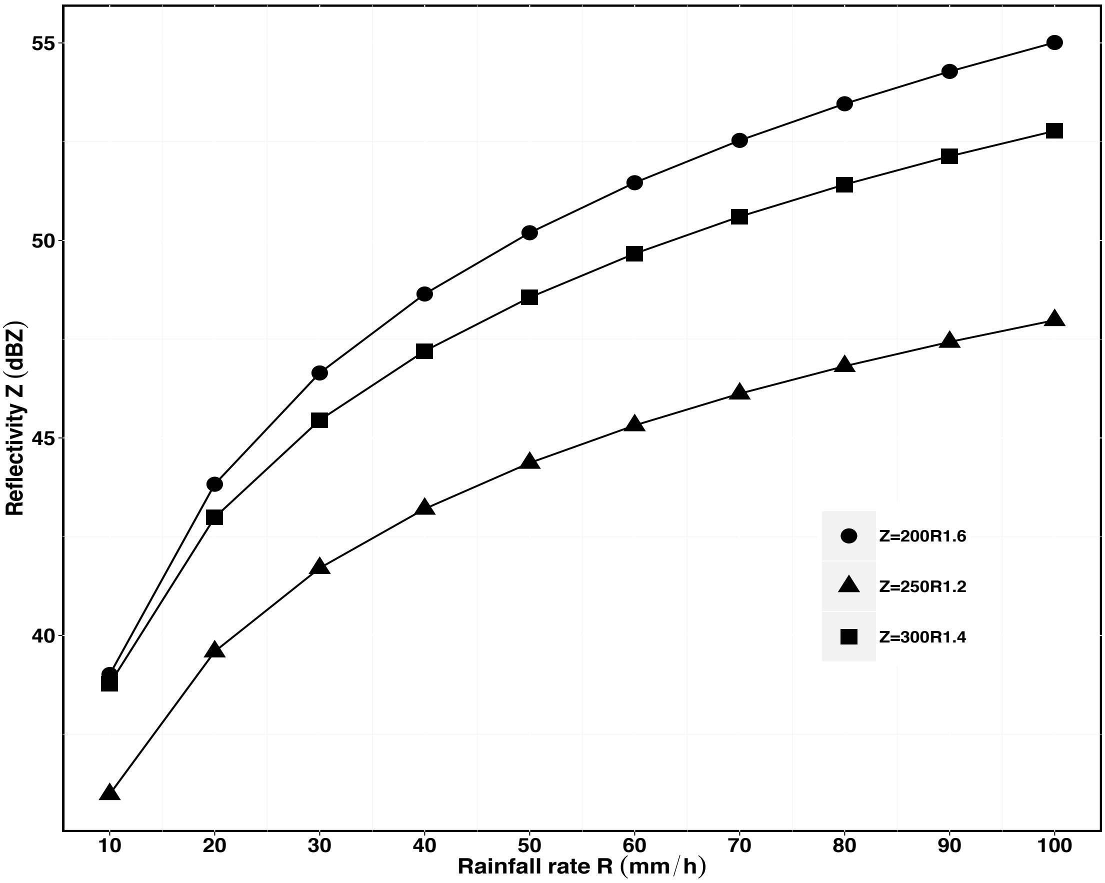


Figure 1.4 Influence from the different Z-R relationship selection to the rain rate estimation, after Einfalt et al. (2004).

- **Density of rain gages for radar-gage adjustment**. Fuyou et al. (2010) tested the effect of density of rain gages to the accuracy of radar-gage adjusting results. Authors applied the data from S-band radar observations and rain gage data in two meso-scale catchments located in An Hui province, China, the testing procedure was performed by reducing number of rain gages. They concluded that when using single rain gage for radar rainfall correcting, the results presented high bias as a result of rain field location, as increasing number of rain gage involved in the correcting procedure, the bias between true value and adjusted results was decreased and the mean squared error became more smoothly when the number of rain gages for adjustment was greater than or equal to 2.

1.5.2 Radar Based Precipitation Forecast

Radar based Quantitative Precipitation Forecast (RQPF) is one type of precipitation forecasting methods using weather radar with the benefits of high temporal and spatial resolution ('Quantitative precipitation forecasts based on radar data for hydrological models'). They are commonly applied in short-term precipitation for hydrological applications and thunderstorm and hail warnings. Grecu and Krajewski (2000) concluded that most RQPF methods can be statistically described as in equation (1.7):

$$\mathbf{\frac{\Delta Z^t}{\Delta t} + U_x \frac{\Delta Z^t}{\Delta x} + V_y \frac{\Delta Z^t}{\Delta y} = f(Z^t, ..., Z^{t-k\Delta t}, a) + w} \tag{1.7}$$

where Z^t is the radar measured reflectivity or translated rain rate which is estimated at time t. ΔZ^t is the reflectivity (rain rate)'s diversification from time $t - \Delta t$ to time t. Δt is the time lag between two radar image scanning moments. U_x and V_y are the horizontal and vertical velocity components separately. Δx and Δy are the horizontal and vertical velocity's variation from time $t - \Delta t$ to t, respectively. f is the function of parameter which is needed to be determined from radar observations at current time Z^t and previous times $(, ..., Z^t - k\Delta t)$. w is the noise element. According to the variation of advectiing fields is persistence or dynamic, equation (1.7) can be described as:

Eulerian persistent process: The Eulerian persistent process is consisted by a simple equation that the forecasted reflectivity (rain rates) equals to ones at previous time. This kind of process neglects the advection, reflectivity (rain rate)'s variability and noise components which reduces to:

$$\mathbf{\frac{\Delta Z^t}{\Delta t} = 0} \tag{1.8}$$

that is, the advectiing field is assumed to be stable in space and time. This forecasting method is simple but roughly and is surely not expected to give the satisfied forecasting results. In some actual applications, it is selected as the reference for evaluating the performance of other QPF methods.

Lagrangian persistent process: Austin and Bellon (1974) pointed out that in application of radar based rainfall forecast, advection of precipitation plays a fundamental role. The Lagrangian persistent process is a kind of process that takes rainfall advection into account, but this process still neglects the Lagrangian dynamic component and reputes that velocity fields is persistence. Therefore, it can be presented as:

$$\frac{\Delta \mathbf{Z^t}}{\Delta \mathbf{t}} + \mathbf{u_x} \frac{\Delta \mathbf{Z^t}}{\Delta \mathbf{x}} + \mathbf{u_y} \frac{\Delta \mathbf{Z^t}}{\Delta \mathbf{y}} = \mathbf{0} \tag{1.9}$$

In detail, this process requires the estimation of the velocity fields before the forecast starts. Subsequently, the forecasted reflectivity (or rain rate) is acquired by re-positioning downwind the current radar image accordingly to the storm shifts which occur at the next time steps, under this assumption, the estimated velocity fields remain unchanged during the forecasting process. common steps for operational application of RQPF are presented as follows:

- Quality control for the observed radar data;
- Estimating horizontal and vertical components U_x and V_y of identified reflectivity or transformed rain rate z^t from radar image;
- Estimation of parameters a in function f. The function f is usually treated as linearly with a being 0. When considering the growth or decay of the precipitation, f is a non-linearly process with $a \neq 0$;
- Forecasting future reflectivity (or rain rate) using equation (1.7);

1.6 Motivations and Objectives

From the year 2005 to 2008, the project 'Forecast and Management of Flash floods in Urban Areas' (URBAS) which aimed at presenting a better understanding of physical processes during urban flash floods was operated (Einfalt et al., 2009). Its main objectives included:

- Investigating the meteorological parameters, runoff and damage of urban floods in Germany.

- Advanced equipment and methods such as the German Radar Network from DWD and rainfall-runoff simulation models are applied to analyze rainfall and runoffs in urban catchments.
- Upgrading forecasting tools for improving flood early warning level, losing mitigation measures and disaster controlling.

After three year's efforts, databases for recording historical flood events occurred in Germany since last 90s have been established (Einfalt et al., 2009). Scientists concluded that most of the historical urban flood events were flashing ones generated by heavy rainfall events from these databases, and they were highly distributed in urban areas of west Germany. Figure 1.5 presented the spatial distribution of flash flood events occurred in Germany since the last 90s.

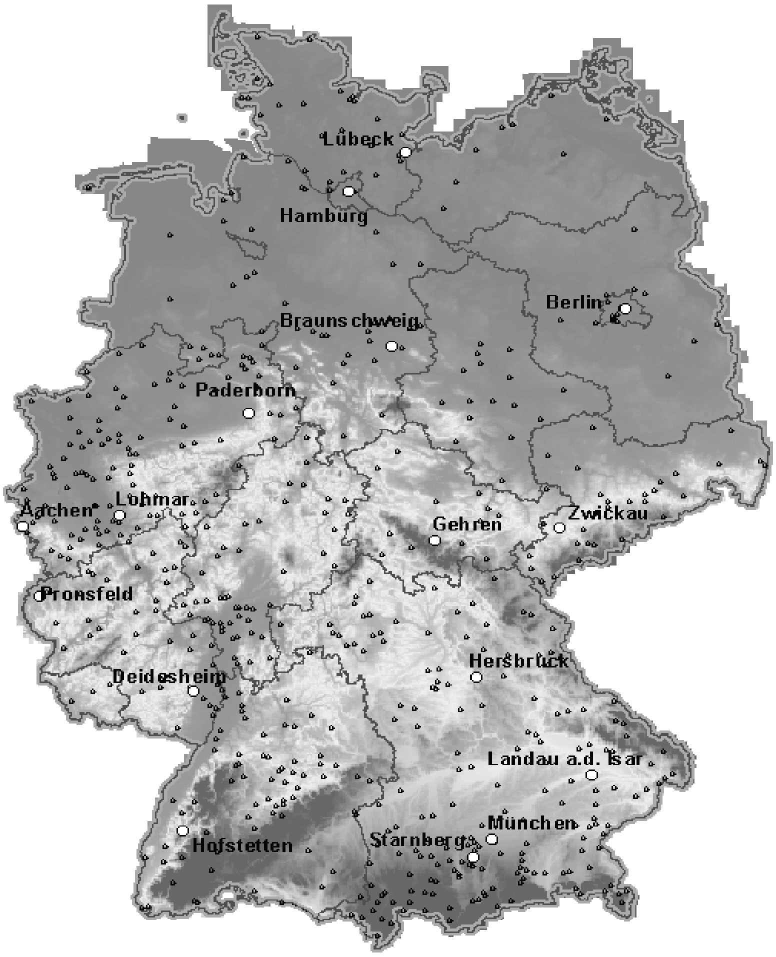


Figure 1.5 Spatial distribution of urban flash floods in Germany since 1990, after Einfalt et al. (2009).

Since the year of 2007, there has been a cooperation between German Weather Service

(DWD) and water associations for the state of North Rhine Westphalia (NRW) which were based on the exchange of data and experience, aiming at using radar data to optimize flood hazards protection. This included not only the management of historical events, but also good preventions for future events. Under this cooperation, facing to the small urban regions with sparse or no gage network distribution in the Wupper area of North Rhine Westphalia, historical heavy rain and flood events occurred in this region were analyzed by combing qualified rainfall data produced from weather radar and distributed hydrological model. Valuable information was also provided for aiding participants doing flood modeling in urban places (Einfalt et al., 2008). However, some weaknesses should not be hidden:

- Rainfall measurements produced from weather radar have been used for analyzing those historical events. Nevertheless, through the radar based rainfall analysis, inner structures of these historical events are still not fully understood. Quantitative estimation of their inner structures is essential for improving our understands for the origin of high peak runoff generation in this region.
- Radar based precipitation forecast for potential flood warning has been done in the earlier decades and indeed demonstrated its advantage. However, the uncertainties still exist. Finding the factors resulting those uncertainties are important for improving the quality of precipitation and flood forecasting.

Objects of this thesis were listed as follows:

- Developing a radar based precipitation objects identification and tracking algorithm for investigating the inner structures of rainfall events occurred in the state of North Rhine Westphalia, Germany.
- Developing the radar based short-term precipitation forecasting methods and evaluating their forecasting qualities by object based verification approaches.

Considering the researching objects of this thesis, following questions will be answered in this study:

- How are the inner structures of rain events occurred in North Rhine Westphalia?
- Can we find some simple mathematical functions to quantify the spatial - temporal variations for the characteristics of rainfall patterns in this region?
- Which is the most factor that affects the performance of the developed precipitation objects identification and tracking algorithm and the quality of precipitation forecasting

methods?

1.7 Outlines

This thesis is organized as follows:

- In **Chapter 2**, a literature review which is related to the radar based quantitative precipitation estimation and forecasting was presented.
- In **Chapter 3**, a introduction for the study area of this thesis was presented.
- In **Chapter 4**, a precipitation objects identification and tracking algorithm was developed. The Semi-Lagrange based advection field tracking method was introduced for short-term rainfall forecasting; a position tracking and predicting algorithm: Kalman Filter was introduced for extrapolation center of mass of precipitation object; a characteristic predicting method was proposed; two object-based spatial rainfall verification methods (SAL method and Geometric index method) were also introduced in this chapter.
- In **Chapter 5**, results were presented, and they were organized as: statistical description for the characteristics of modeled precipitation objects; mathematical relations among their characteristics; spatial-temporal analyzing for the modeled precipitation objects; verifications for the developed algorithm; uncertainties analyzing for the QPF results.
- In **Chapter 6**, discussions and outlooks of this thesis were presented, questions proposed in this study were also answered in this chapter.
- In **Chapter 7**, summary of this thesis were presented.

CHAPTER 2

LITERATURE REVIEW

Considering application requirements for radar based QPE and QPF, methods based on multi approaches have been proposed and developed during past two decades. Nakakita et al. (1996) concluded these methods into two main categories:

- Extrapolating advection fields based on the estimated motion vectors from radar or satellite images (e.g. advection field tracking approach, cell tracking approach).
- Using numerical models (conceptual or physical) or numerical weather prediction models for doing QPE and QPF at a high spatial resolution (e.g. spectral approach, numerical weather prediction model approach, artificial neural network approach).

A review for the related works about researching developments for QPE and QPF are presented according to their designing approaches.

2.1 Advection Field Tracking Approach

The advection field tracking is a kind of pattern matching approach for extracting background moving fields by searching the maximum correlation coefficient of reflectivity/intensity fields between two consecutive radar images. Radar images are first divided into a collection of boxes that can cover the whole echo fields, then the motion vectors of the echo fields are determined by the maximum correlation coefficient of two boxes which are from the current radar image and the previous one. Procedure of this approach is illustrated by Mecklenburg et al. (2000) as in Figure 2.1.

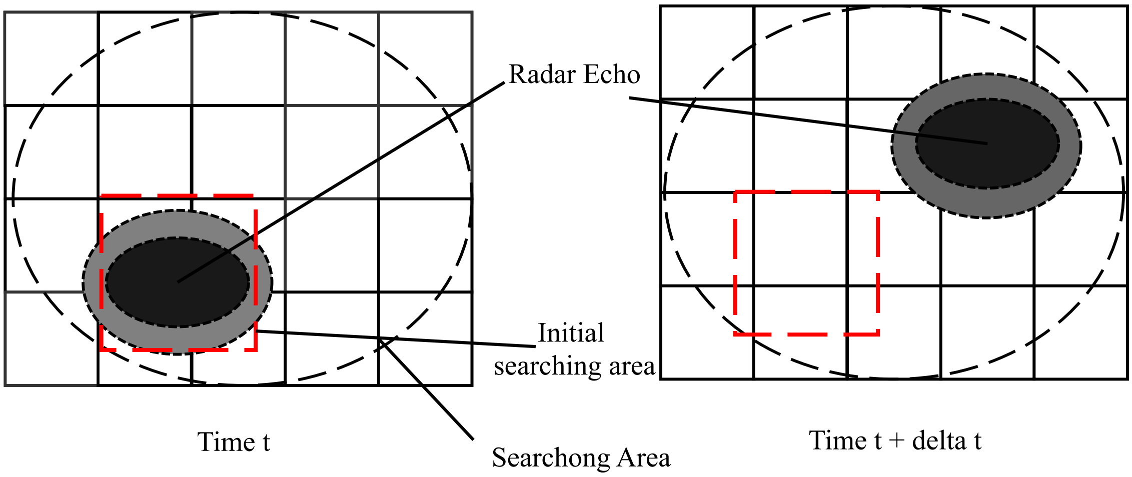


Figure 2.1 An illustration of advection field tracking approach, after Mecklenburg et al. (2000).

Rinehart and Garvey (1978) developed a radar based QPF algorithm named with Tracking Radar Echo by Correlation (TREC) which tries to find the best fits between arrays of reflectivity in two consecutive radar images by optimizing their correlation coefficients. The first radar image is subdivided into sets of regular boxes, which is used in determining the spatial resolution of motion vectors. An initial tracking area of a specified box size is defined around each grid point. The initial tracking area is compared to similar sized boxes within a searching range of the second radar-scanning image. The size of the searching range r is defined according to equation (2.1):

$$\mathbf{r} = \mathbf{v}_{\max} \times \boldsymbol{\Delta}\mathbf{t} \tag{2.1}$$

where v_{max} is the maximum velocity with its value obtained from atmospheric condition. Δt is the time interval between two successive radar scanning moments. For each box pair, in the first and second radar images, the correlation coefficient is computed and the best fit is found by choosing the pair of boxes with maximum computing result. Distance between the boxes and scanning time lag determines the motion vector of the tracking area. The correlation coefficient R for box pairs between two radar images is described in equation (2.2):

$$\mathrm{R} = \frac{\sum(\mathbf{Z_{t-1}(k)} \times \mathbf{Z_t(k)}) - \frac{\mathbf{1}}{\mathbf{N}}\sum \mathbf{Z_{t-1}(k)} \sum \mathbf{Z_t(k)}}{\sqrt{(\sum \mathbf{Z_{t-1}}^{\mathbf{2}}(\mathbf{k}) - \mathbf{N}\mu \mathbf{Z_{t-1}}^{\mathbf{2}})} \times \sqrt{(\sum \mathbf{Z_t}^{\mathbf{2}}(\mathbf{k}) - \mathbf{N}\mu \mathbf{Z_t}^{\mathbf{2}})}} \tag{2.2}$$

where Z_{t-1} and Z_t are the pixel's reflectivity in the defined box at time $t-1$ and t, respectively. N is the pixel number inner the box, k is the id of the box.

Dixon and Wiener (1993) added a geometry algorithm based on TREC for providing a possibility to detect merging and splitting of rain cells and the forecast was carried out by using a weighted linear regressive fitting methods from the historical storms for both position and size which is called Thunderstorm Identification, Tracking, Analysis and Nowcasting (TITAN). Li et al. (1995) applied the two-dimensional continuity equation to eliminate divergent components of extracted velocity fields which are called COTREC. Li et al. (2014) implemented TREC algorithm into Short range Warning of Intense Rainstorms in Localized Systems' (SWIRLS) for monitoring and extrapolating the motion vectors of radar echo right across the meso-scale spectrum and forecasting precipitation with leading time from zero to three hours. Liang et al. (2010) combined TREC algorithm with predicted winds from CHAF model for providing long term's precipitation forecast in Guang Dong province, China. Zahraei et al. (2012) developed a pixel based precipitation forecasting method which implemented a quadrilateral mesh based template-matching algorithm for tracking advection field at high resolution, the method proposed by the author was robust for sparse rain fields.

2.2 Cell Tracking Approach

Different with the advection field tracking approach, the cell tracking (or centroid tracking) methods include a detecting algorithm for identifying discrete cell's characteristics (e.g. centroid, area, echo-tops, vertical integrated liquid) in consecutive radar images and a matching algorithm for tracking cell's motion and shape changes (e.g. merging and splitting). Advantage of cell tracking methods is the fact that it can estimate cell's characteristics feasibly from cell trackers. Because of its superior performance in reflecting convective cell's dynamic, this approach is more suitable for convective rainfall events.

Einfalt et al. (1990) proposed a structured approach to determine the cell motion which contains cell's definition, characterization, matching and forecast which has been implemented into SCOUT rainfall processing system (Lempio et al., 2012). This algorithm tries to identify and recognize distinctive features of a cell in consecutive radar images. Rainfall event's characteristics derived by SCOUT include not only the size and centroid but also the orientation, elongation, intensity distribution, previous size, moment vectors and their root mean square variation and previous recognition of the cell. Johnson et al. (1998) developed Storm Cell Identification and tracking (SCIT) algorithm, in which the cell identification is done by predefined reflectivity thresholds (30, 35, 40, 45, 50, 55, 60 dBZ), and cell tracking is done through a combination optimization, which is solely upon a distance criterion. A concept called the cell based VIL was also present for considering storm volume titled in vertical direction. The SCIT algorithm performs well for convective rainfall events, but it fails in forecasting larger scale strati-form rainfall events. Handwerker (2002) invented an automated cell tracking algorithm - Trace3D by assembling the adjacent reflectivity regions which exceed some thresholds. The cell is identified and based on the cell's position in the radar image at previous time. The new position for the next time is tracked by an extrapolation procedure. Splitting and merging of rain cell is also dealt in Trace3D algorithm. Han et al. (2009) enhanced TITAN algorithm by combining TREC algorithm for tracking identified cell. Zahraei et al. (2013) built an object based storm cell tracking model called PERsiann-ForeCAST. Storm is identified from cloud and temperature data in this model, then an object-based tracking algorithm is done for tracking storm's motion in two or three consecutive rainfall images (PERSIANN-CCS data sets). Cell's growth and decay are also taken into consideration in this model.

2.3 Spectral Approach

Spectral based QPF methods can be described by equation (2.3), a radar image with $L \times L$ pixels is spatially decomposed into an additive cascade of n levels. X_k represents the kth scale of the decomposition.

$$\mathbf{dBZ_{i,j}} = \sum\nolimits_{\mathbf{k=1}}^{\mathbf{n}} \mathbf{X_{k,i,j}(t)} \tag{2.3}$$

where $dBZ_{i,j}(t)$ is the pixel's reflectivity in horizontal position i and vertical position j at time t. The translated rainfall intensity R has a multiplicative structure, for example, rain fields can be approximated by multiplying the independent component processes at different scale. This multiplicative structure converts into a summation when taking the logarithm transformation, which is the reason why dBZ is taken for the decomposition, not Z or R. X_k in the cascade represents the variability of the original field with structures of scales between $2^{-(k+1)L}$ and 2^{-kL} pixels.

Seed (2003) developed S-Prog algorithm which is based on this approach. In this algorithm, original reflectivity fields from radar images are decomposed, then the evolution of theses levels is considered separately: a measure for the lifetime of a certain level is obtained by studying the correlation of that level at time t_1 and at time $t_2 = (t_1 + \Delta t)$. The forecasted image is then composed by the summation of the different levels, taking into account their predicted lifetime as weight. The levels with the larger structures will have high correlation values with previous levels, and thus these levels will dominate for long leading times.

2.4 Numerical Approach

Numeric approach is a combination of advantages from weather radar and numerical weather prediction model (NWP) for providing high resolution and long leading time rainfall estimation and forecast. Usually, methods based on area tracking or cell tracking approach can only give a confidence rainfall forecast at short leading time, but NWP model is more skillful for providing longer leading time's forecast. However, the NWP model may perform well for med or large scale convective events. Purely application of NWP model may increase the precipitation forecasting uncertainties for the local area because of the coarse resolution of initial boundary conditions. QPE/QPF combined with NWP model can build a bridge for filling gaps between requirements of high resolution precipitation forecast and longer forecast leading time. The key point in doing QPF coupling NWP model is that how to assimilate radar observed data into numerical weather predicting model. According to Schlatter (2000), data assimilation can be traded as a process of estimating weather conditions on a regular grid from two main sources of information: Observations of atmospheric indices (e.g. wind, pressure, temperature, perceptible water) from different sources, whenever and wherever taken, and a numerical weather model, which couples through a series empirical mathematical equation about the atmosphere.

Assimilating radar data with NWP model can construct the best set of initial conditions for providing better rainfall forecasting results in local scale. Golding (1998) introduced nowcasting system called 'Nowcasting and Initialization for Modeling using Regional Observation Data Rainfall' (NIMROD), which is distributed at the Met Office headquarters in Britain since 1995. The system is intended to generate analysis and short-range forecasts of various meteorological variables especially precipitation. Bowler et al. (2006) developed a probabilistic precipitation forecasting scheme that blends an advection field extrapolation with a down scaled NWP forecast, known as STEPS. Performance evaluations showed that this scheme had a credible predictive skill at leading time over six hours. Sokol (2006) did one hour ahead precipitation forecasting combined with C-band weather radar and ALADIN/LACE model data in Czech Republic, the spatial resolution is 9×9 km. Milan et al. (2009) implemented a initial physical method in COSMO model of the German Weather Service for improving the precipitation forecasting quality. Liu et al. (2013) used five data assimilation modes in WRF model and concluded that when the radar data is combined with the numerical weather model, a good precipitation forecast exists.

2.5 Neural Network Approach

Artificial Neural Networks (ANN) is a technology by imitating how does the human brain process the real world objects for solving complex problems that can't be solved by simple mathematical methods. Thinking system of the human brain is a complicated network. This network is linked by numerous neurons, and information from the outside world is processed based on the strengths of connection between neurons. It is more simply comparing the real human neural system for structure of ANN, usually, ANN is constructed by an input layer with multi-parameters, hidden layer with some neurons and output layer with results.

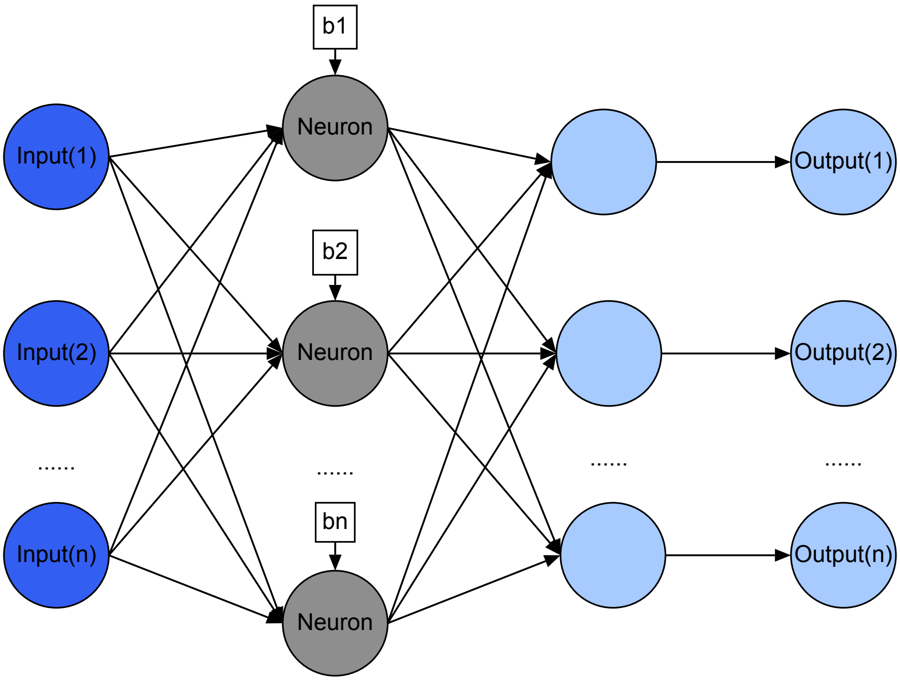


Figure 2.2 Structure of a three layer ANN. The blue, gray and cyan nodes presented input layer, hidden layer and output layer, respectively.

Figure 2.2 presented the structure of a three layer ANN. Neuron is the basic information process unit in ANN that is composed by weights, bias and exciting function. The ANN is trained by continuously changing weights and bias of neuron for having capability such as predicting and pattern recognition, then the trained ANN is used to predict future state of the system.

Denoeux and Rizand (1995) used consecutive radar images as input data for ANN training, then used the trained ANN for forecasting precipitation. Chiang et al. (2007) applied the dynamic ANN approach for doing QPF with a leading time from 0 to 1 hour in Taiwan province, China. Li et al. (2013) used ANN for doing one hour ahead rainfall forecasting in British based on NIMORD radar data, the input of ANN is pixel's reflectivity and its four neighbor's values.

CHAPTER 3

STUDY AREA

The study area is the Federal State of North Rhine Westphalia (NRW) which is situated in north-west of Germany and is with nearly 18 million inhabitants on an area approximately 34,000 km^2, thus 524 inhabitants per km^2. It is the most highly and most densely populated state in Germany.

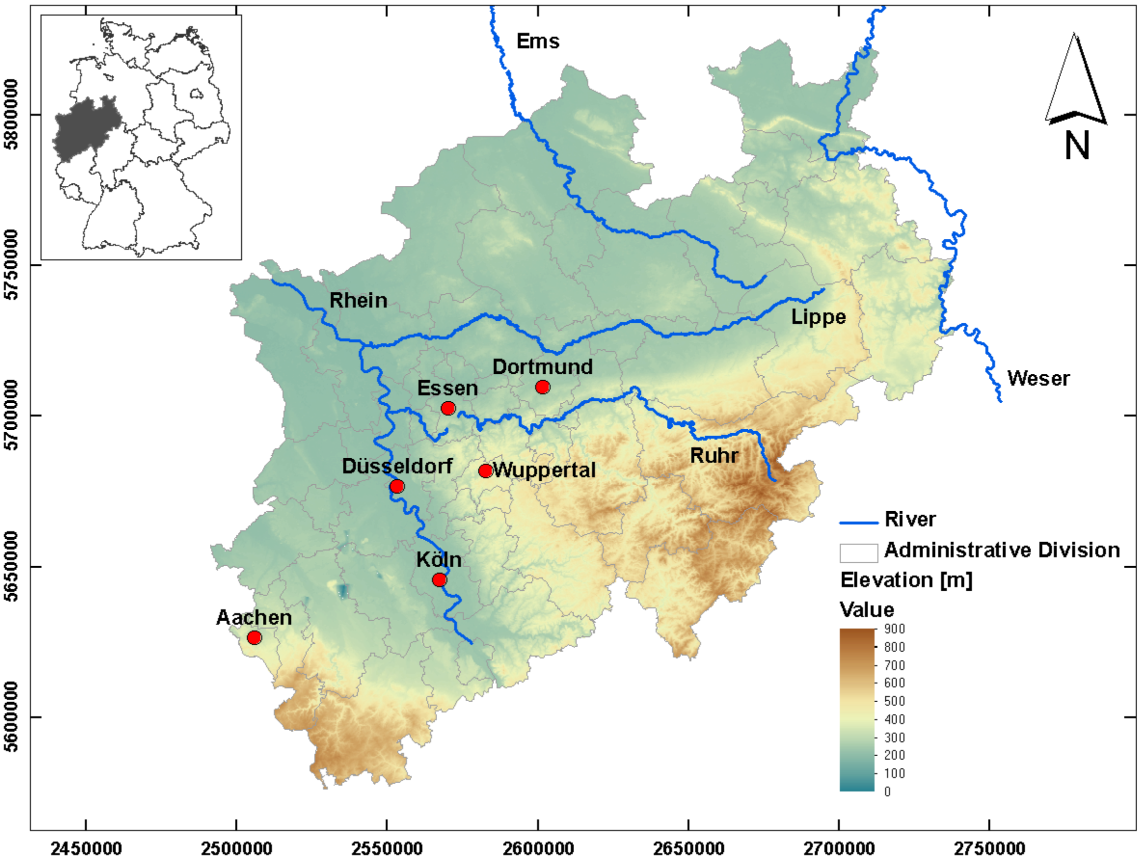


Figure 3.1 The study area North Rhine Westphalia and its location in Germany (inner plot). Main administrative cities are marked with red dots.

The state's area covers a maximum distance of 291 km from north to south, and 266 km from east to west and borders on the German states Lower Saxony, Hessen and Rhineland Palatinate as well as the countries Belgium and the Netherlands. The North Rhine Westphalia includes the upland regions of North Eifel in the southern part of the state and the mountains of the Sauer Land in the southeast. The most important rivers that run at least partially through North Rhine Westphalia include: Rhine, Ruhr, Ems, Lippe and Weser as in Figure 3.1.

The circulation pattern of North Rhine Westphalia is mainly affected by the air mass from Atlantic along the direction to south-west. When arriving to the southern high mountain regions, the air mass stops and arises. This leads at these places to a stronger cloudiness by which more precipitation falls. In the eastern sides of the mountains, there are more drier air dropping which result less cloudiness and precipitations.

3.1 Climate Background

North Rhine Westphalia belongs to the warm temperate rain climate, in which the mean temperature of the warmest month is below 22^oC and the coldest month above -3^oC (Kropp et al., 2006). North Rhine Westphalia is thus in a predominantly maritime sector, which is characterized in general with cool summers and mild winters. In the summer season (may - September), this region is influenced by longer periods of high air pressure, as a result, the weather in North Rhine Westphalia is hot and dry. Synoptic conditions in the winter season are often characterized with cold continental climate type. However, the forced lifting by the overflow of the mountain, results in a large-scale rainfall increase in precipitation at the windward side of the mountains (upwind) and over the mountains. In summer seasons, it comes through the different heating different inclined slopes on the mountain countries earlier and more often to convection, so that increasingly form of showers and thunderstorms occur. There is an increase trend for rainfall following altitude's shift, but only regionally valid and no longer affects smaller-scale structures below 1 to 2 km height.

3.1.1 Temperature Variation in NRW

The distribution air temperatures over the North Rhine Westphalia are strongly dependent on the height of the terrain as shown in Figure 3.2. In the southern mountain regions (Eifel, Weserberg Land, Bergisches Land, Sauer and Sieger Land), its value is approximate to 0.7^oC per 100 m height. The yearly averaged temperature value is shifted following the altitude of the terrain.

Through the figure 3.2, the northern low land areas such as the Rhine and Westphalian bay as well as other areas below 150 m, the yearly averaged air temperatures are about 9^{o}C. The highest mean annual temperature is up to 11.2^{o}C along the Rhine valley in the Rhine bay. In the region of Eifel, Sauer and Siger Land, yearly averaged temperatures are between 5 and 8^{o}C.

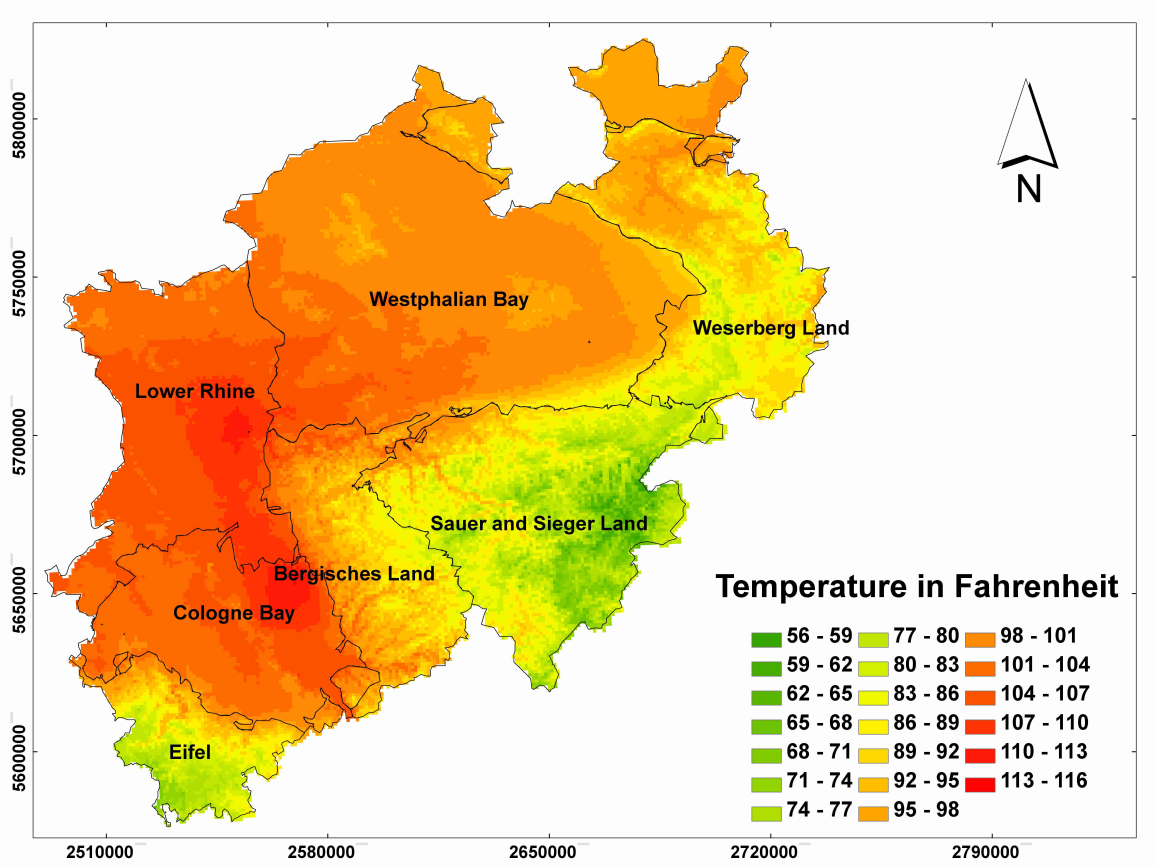


Figure 3.2 Yearly mean value of air temperature in North Rhine Westphalia during the period from 1981 to 2010, the unit of is displayed as fahrenheit. Data sources are from Climate Data Center (German weather service (DWD) (2014)) of DWD, ftp://ftp-cdc.dwd.de/pub/CDC/grids_germany/multi_annual/.

Multi-years particularly warm or cold periods are characterized by the days in which the temperature is above some given threshold. The characteristic days are the sum of all the days in a year which exceed a defined maximum temperature or a predetermined minimum temperature threshold, which are defined as follows:

- **Ice day**: A day on which the maximum air temperature is below 0^{o}C.
- **Frost Day**: A day on which the minimum of the air temperature is below the freezing

point.

- **Summer Day**: A day when the air temperature reaches 25^{o}C or above.
- **Hot Day**: A day on which the air temperature reaches 30^{o}C or above.

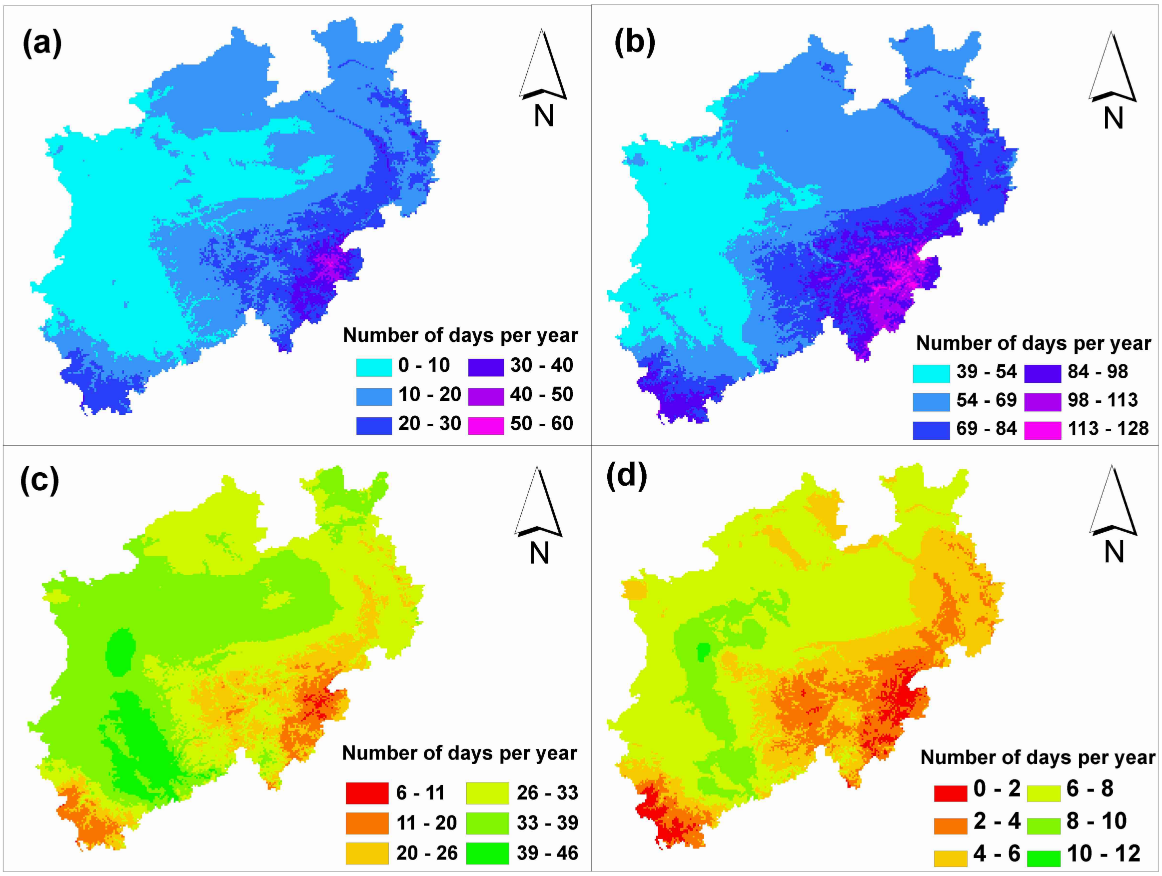


Figure 3.3 Distribution of identical day with different temperature threshold in North Rhine Westphalia during the period from 1981 to 2010, data sources are from Climate Data Center (German weather service (DWD) (2014)) of DWD, ftp://ftp-cdc.dwd.de/pub/CDC/grids_germany/multi_annual/. (a) for ice day; (b)for frost day; (c) for summer day; (d) for hot day.

Figure 3.3 showed the distribution of identical days with different temperature threshold over the state of North Rhine Westphalia. In the southern highlands of the Sauer and Sieger Land region, frost days occur almost a third of all days in the year. There is more in the long-term average for more than 40 ice days per year (figure 3.3a). Much of the northern low land areas such as Lower Rhine Basin, the Lower Rhine lowlands and the Westphalian bay are from frost least affected with less than 60 days of frost per year (figure 3.3b). The Rhine valley is particularly warm in the summer season with more than 33 summer days per year (figure 3.3c). The increasing height of the terrain reduces the number to a few summer days in the high altitudes of the Eifel, Sauer and Sieger Land. Most hot days are analogous to the summer days which can also be found in the Rhine Valley (figure 3.3d). The heights of the Sauer and Sieger Land have long-term average on any hot day.

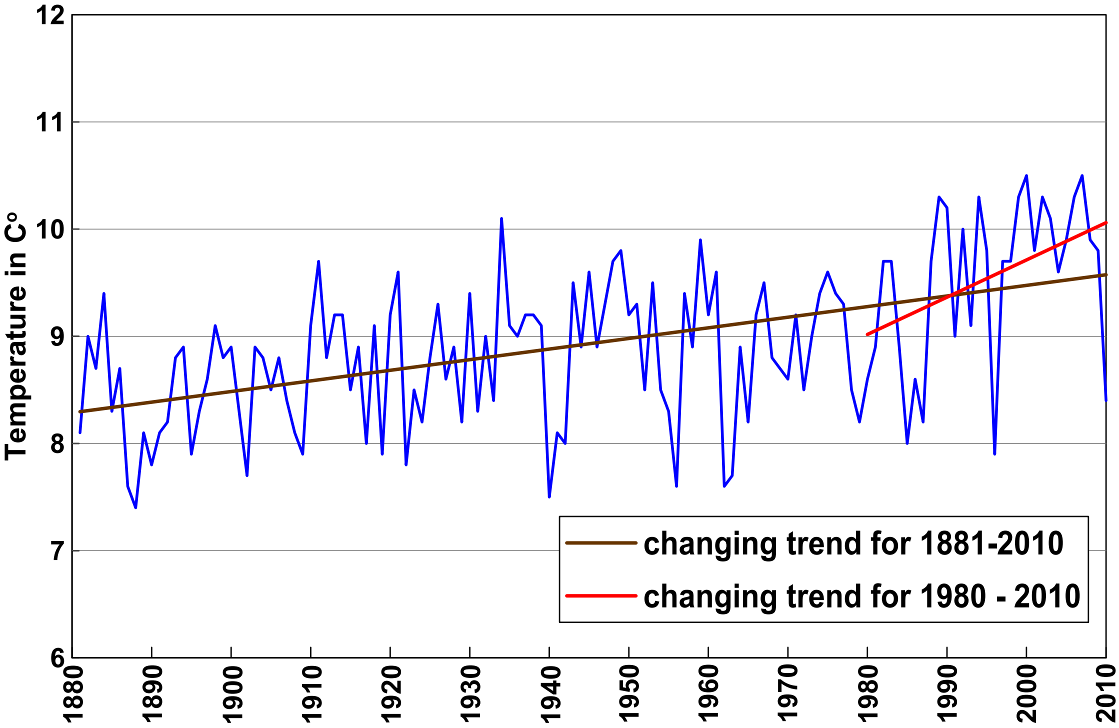


Figure 3.4 Multi-years averaged temperature over North Rhine Westphalia during the period from 1881 to 2010. Data sources are from Climate Data Center (German weather service (DWD) (2014)) of DWD, ftp://ftp-cdc.dwd.de/pub/CDC/regional_averages_DE/annual/air_temperature_mean/.

During the period from 1881 to 2010, the yearly averaged temperature over the North Rhine Westphalia is 8.9^oC (standard deviation is 0.7^oC). The lowest yearly averaged temperature occurred in 1888 (7.4^oC) and 1940 (7.5^oC) and 10.5^oC in 2000 and 2007. Generally, a highly significant increase in temperature was recorded in North Rhine Westphalia in the 130 -year period. In this period, the temperature is increased by 1.3^oC (from 8.3^oC to 9.6^oC, the difference between the beginning and end of the red trend line in figure 3.4).

3.1.2 Precipitation Variation in NRW

The yearly precipitation over the North Rhine Westphalia is 918 mm during the period 1981 to 2010. Figure 3.5 presented the spatial distribution of yearly precipitation in that region. Through the figure, it is seemed that precipitations of this region are also affected by the height of Teriann which is similar to the spatial distribution of yearly averaged temperature as shown in figure 3.2. However, considering the amount of precipitation in all sub-regions with different altitudes, clear differences still appear between the wind-turned mountain situations (Luv) and the leeward sector (Lee).

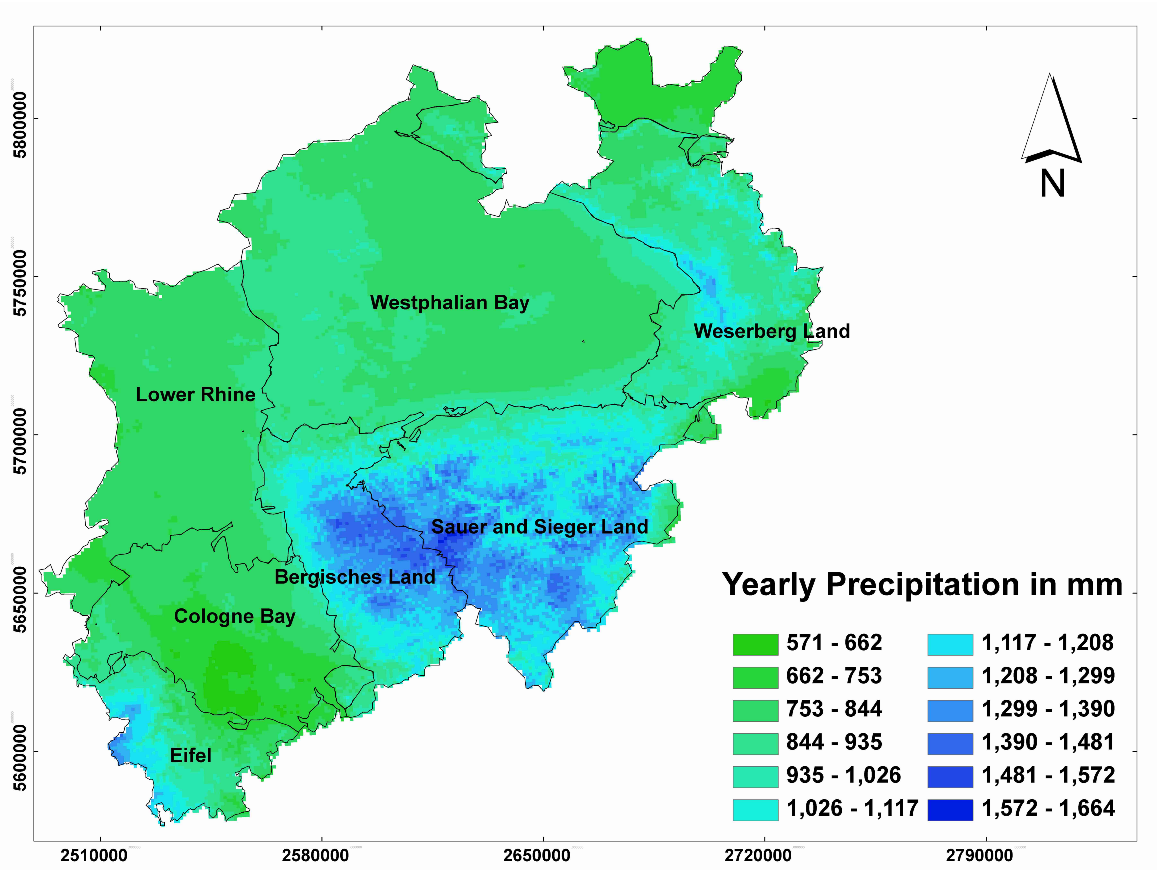


Figure 3.5 Multi-years precipitation distribution in North Rhine Westphalia during the period from 1981 to 2010. Data sources are from Climate Data Center (German weather service (DWD) (2014)) of DWD, ftp://ftp-cdc.dwd.de/pub/CDC/grids_germany/multi_annual/.

Through the figure 3.5, in the northern lowland regions such as Westphalian Bay and Lower Rhine lowlands, yearly precipitation is between 600 to 900 mm. In the southern mountain regions of Bergisches Land, Sauer and Sieger Land, the yearly total precipitation can reach to 1700 mm. It is particularly strong in the Bergisches Land region. Here, the effect of windward increases precipitation noticeable up to the Rhine area. In the region of Wuppertal/Remscheid in surroundings at about 250 m, a yearly precipitation of more than 1300 mm is observed. Yearly precipitation with value between 1400 to 1450 mm is observed in the Olpe district (460 m), as much as for Kahlen Asten (839 m). .The yearly precipitation is just over 600 mm between north-east perimeter of the Eifel and Ville. In the northern lowland region, strong sunlight causes frequently showers and thunderstorms which take more precipitations in its most part. However, In the central mountain region, large amount precipitations are observed when stronger westerlies introduce Atlantic air masses to frequent rainfall in the winter months.

The yearly averaged precipitation is 849 mm in North Rhine Westphalia during the period 1881 - 2010. Compared to the temperature, there is a stronger annual fluctuation as in figure 3.6.

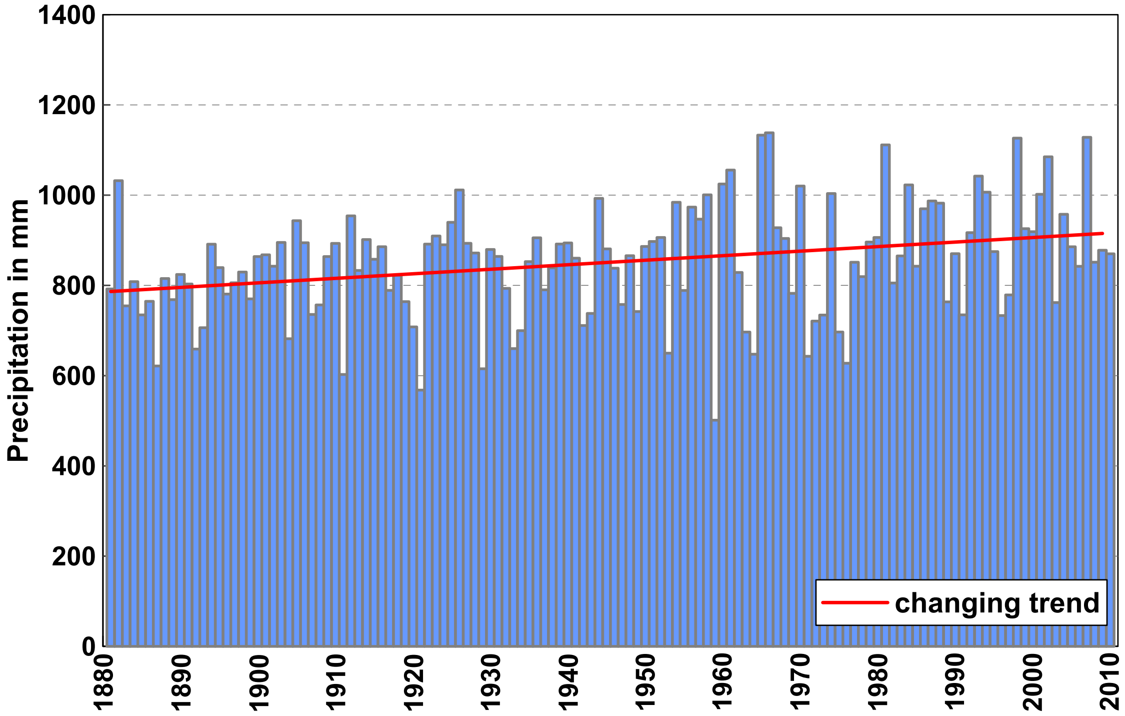


Figure 3.6 Average annual precipitation of North Rhine Westphalia during the period from 1881 to 2010. Data sources are from Climate Data Center (German weather service (DWD) (2014)) of DWD, ftp://ftp-cdc.dwd.de/pub/CDC/regional_averages_DE/annual/precipitation/.

To determine statistically significant trends which are suitable for long time series, the 'dry' 1970s Decade in the North Rhine Westphalia are taken into considerations. This is recognizable by means of decade moving average. The reason for such dry decades existed may be that the Atlantic climate anomaly with significantly cooler surface temperatures in the precipitation measurements resist because the weather patterns is often characterized by synoptic conditions with influx of Atlantic air masses in this region.

During the period from 1881 to 2010, an increasing trend was presented for yearly total precipitation as well as for the air temperature. During this period, the yearly averaged rainfall over North Rhine Westphalia rose by almost 130 mm (difference between start and end of the 110-year trend line in figure 3.6). This corresponds to about 10 mm per decade. Thus, the precipitation in 130 years has increased by 15 % over the long-term average.

3.1.3 Extreme rainfall events

The intensity change of rainfall may be a direct or indirect result of global warming. The temperature increasing affects, among others, directly on the water vapor content of the air and indirectly on the circulation systems of the earth. As a result, rainfall patterns can change regionally.

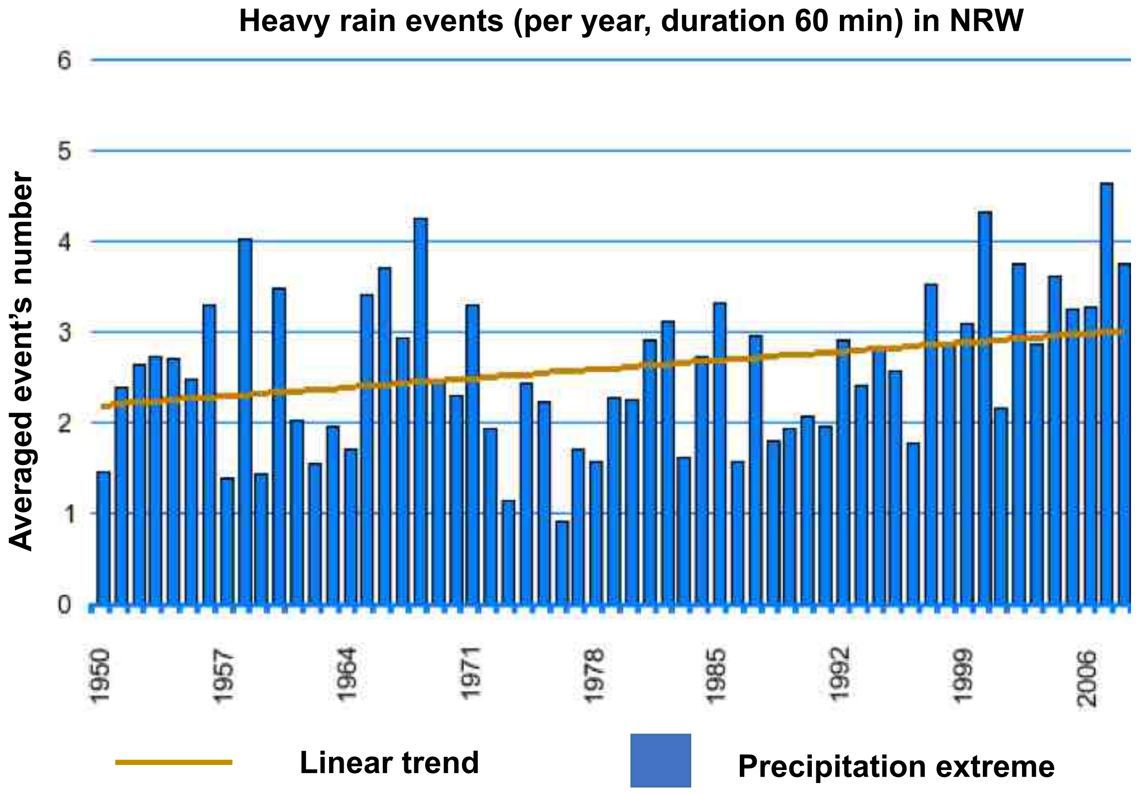
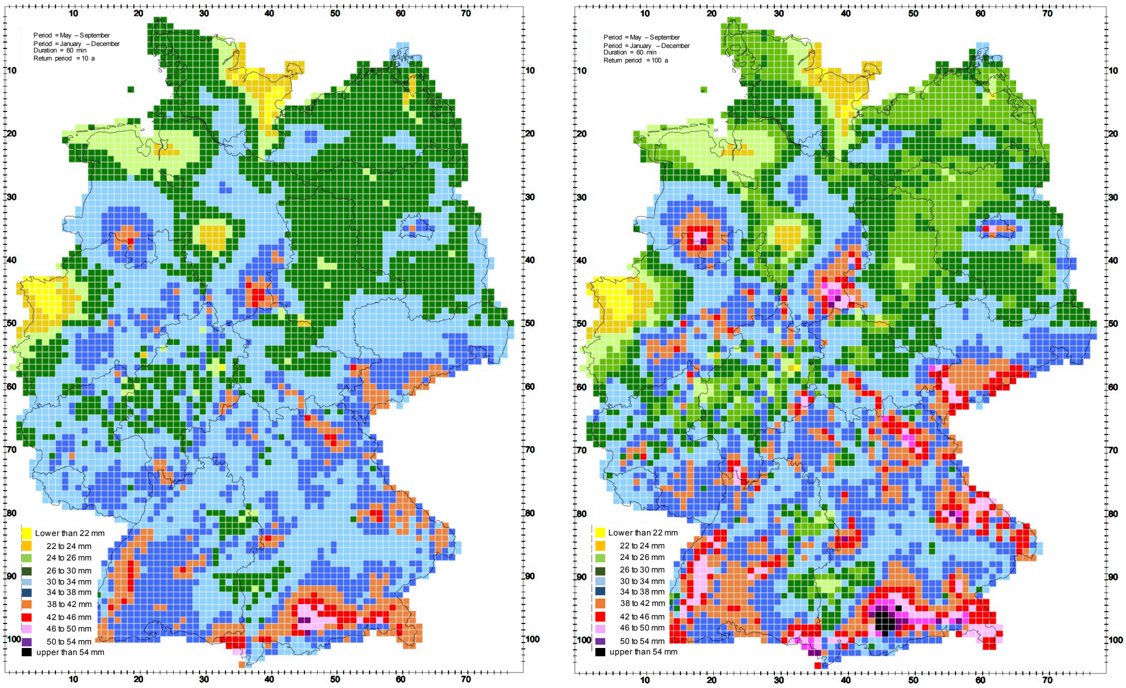


Figure 3.7 Averaged number of heavy rainfall events (threshold > 40 mm/day) per year for duration of 60 min (1950 - 2008). Image is from State Office for Nature, Environment and Consumer Protection of North Rhine Westphalia (2010), http://www.lanuv.nrw.de/kfm-indikatoren/index.php?indikator=6&mode=indi&aufzu=1.

The detection of these climate-induced changes is very difficult. The evaluation shows that for North Rhine Westphalia, a slight increase in the average number of heavy rainfall events per year existed during the period from 1950 to 2008. In the 2000s, short term local heavy rain events have multiplied, which took lots of damage in urban areas. Figure 3.7 gave the additional peaks of heavy rainfall with duration of 60 min between 1950 - 2008. The trend line presented in the figure is not statistically significant.

According to Quirmbach et al. (2012), during the past hydrological years (1950 - 2008), in many urban areas of North Rhine Westphalia, the number of dry days (daily precipitation less than 0.1 mm) had increased while at the same time this number had decreased in some mountainous areas. For winter periods, there had been a clear uniform reduction of dry days up to 20 days in the last 58 years, whereas they had increased only half number (10 days) in summer months. The number of days with daily precipitation above 10 or 20 mm had significantly increased with an average of 3.5 days in the last 58 years, which were observed for the whole state, and it was more clearly for winter months than for summer months.



(a) (b)

Figure 3.8 Spatial distribution of rainfall events with duration of 60 min of Germany with different recurrent intervals in summer season during the period from 1951 - 2000. (a) for return period of 10 year; (b) for return period of 100 year. Images are from German weather service (DWD) (2006).

Figure 3.8 showed the spatial distribution of rainfall events with duration of 60 min of Germany with different recurrent intervals in summer season during the period from 1951 to 2000, which is followed the DWD project KOSTRA-2000 (Bartels et al., 2005).

Through the Figure 3.8a, during the climate period 1951 - 2000, comparing to the reference period (January - December), rain events with duration of 60 min occurred in summer season (May - September) were more frequently distributed in southern Germany at the recurrent interval of 10a. For the range of North Rhine Westphalia, rain events with rainfall volume above 30 mm occurred more frequently in the southern mountain regions. In the northern low land region, rainfall volume of rain events were less than what occurred in the southern part, the value was between 22 to 30 mm.

Through the Figure 3.8b, spatial distribution of rain events with duration of 60 min at recurrent interval of 100a were more similar to what were presented in figure 3.8a but their rainfall volume were higher, and for North Rhine Westphalia, event's distribution at this recurrent interval was also consistent. It is suggested that the heavy rain events are more vulnerable to occur in the southern mountain region of North Rhine Westphalia compared to the northern low land region during the climate period 1950 - 2000, and the event's volume at recurrent interval of 100 years was higher than at recurrent interval of 10 years.

3.2 Geomorphology

There are two main types of landscapes in North Rhine Westphalia, one is the North German Lowlands with elevations a few meters above sea level. Another one is the North German Low Mountain Range with elevations of up to 850 m above sea level. Detailed landscape types are presented as in Figure 3.9.

The Sauer and Sieger Land is a rural, hilly region which spreads across most of the south-east part of North Rhine Westphalia and parts of federal state Hessen. The Sauer and Sieger Land is a heavily forested region with sparse inhabitation. It connects with the Bergisches Land to the West, and with the Sieger Land to the South, and continues into the Teutoburg Forest to the North-east. The major rivers running through the Sauer and Sieger Land region are the Ruhr and its tribute - Lenne. Several artificial lakes were created on the smaller rivers by building dams to store water for the adjacent area of Ruhr. The region includes several sub Mountains with highest elevations of 843 m above sea level in Rothaar mountain.

The Lower Rhine is a region around the Lower Rhine section of the river Rhine which locates in North Rhine Westphalia between approximately Oberhausen and Krefeld in the East and the Dutch border around Kleve in the West. The Lower Rhine region basically covers the districts of Cleves, Wesel, Viersen and Neuss as well as the independent cities of Duisburg, Mönchengladbach and Krefeld. On its disputed parts of Oberhausen as well as Düsseldorf can be seen as part of the Lower Rhine.The Lower Rhine region's landscape is mostly flat green grass land with wide views of the horizon.

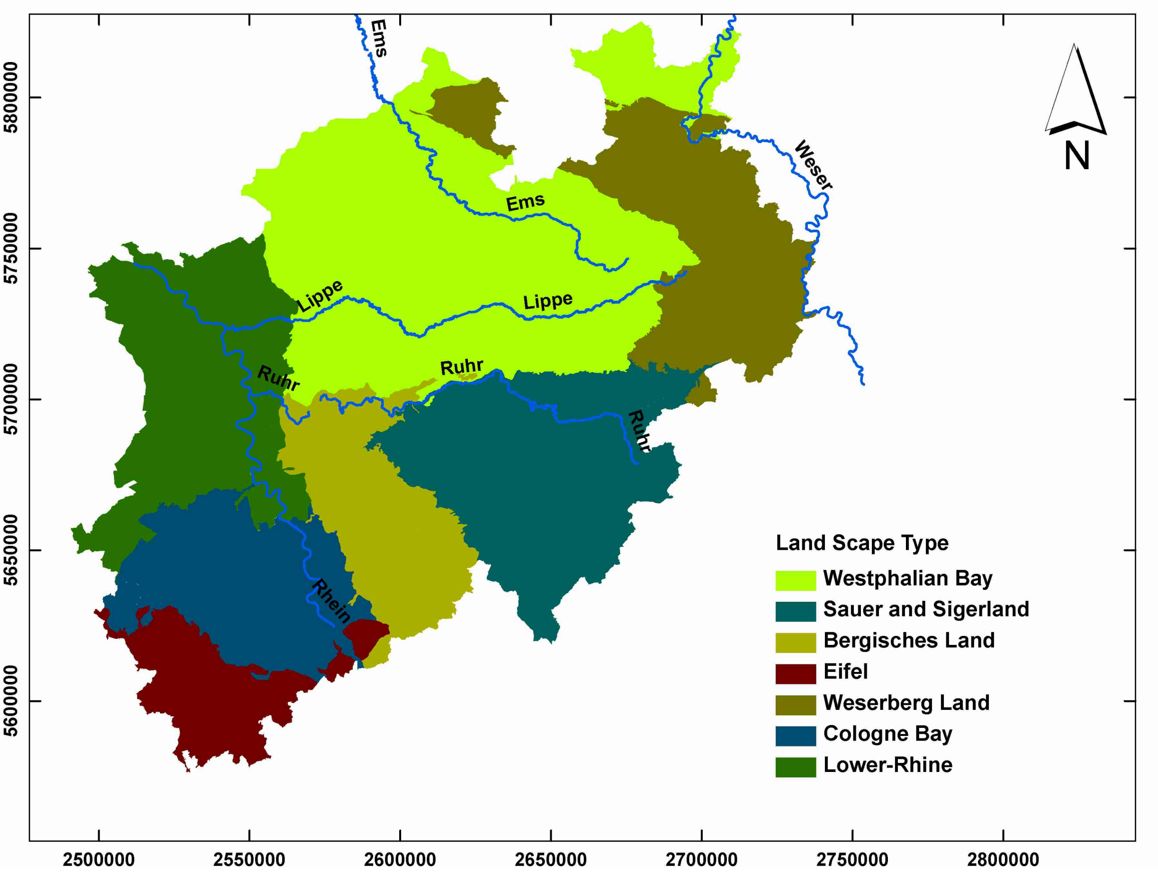


Figure 3.9 Main landscape types in North Rhine Westphalia. Data sources are from State Office for Nature, Environment and Consumer Protection of North Rhine Westphalia (2015), http://www.naturschutz-fachinformationssysteme-nrw.de/natura2000-meldedok/de/downloads.

The Eifel is a low mountain region which spreads across western Germany and eastern Belgium. It occupies parts of southwestern North Rhine Westphalia, northwestern Rhineland Palatinate and the south of Belgium. The Eifel region is bordered by the river Moselle in the South and the Rhine in the East. In the North, it is limited by the landscape of Jülich-Zülpicher Börde, and by the Ardennes of Belgium and Luxembourg in the West. The northern parts of Eifel include Rur Eifel, High Fens and the Limestone Eifel. The northeastern part of this region is called Ahr Hills. In the West, on the Belgian border, the hills of Eifel are known as Schneifel with a elevation of 698 m above sea level. The southern half of the Eifel is lower. It is cut by several rivers running from north to south direction towards the Moselle.

The Cologne Bay is a region with high densely population lying between the city of Bonn, Aachen, and Düsseldorf. The Cologne Bay is surrounded on the left bank of the Rhine by the High Fens and the Eifel, as well as by the Bergisches Land on the Rhine's right bank. In the South and Southeast, the rising Rhine Massif is visible from far off by the silhouette of the Siebengebirge, form the conclusion of the bay at city Königswinter. In the Northwest, the Cologne Bay opens out into the valleys of the Rhine and the Meuse. In the Northeast, it is bounded by the Münster Chalk Basin of the Westphalian Bay.

The Weserberger Land is a hill region along the Weser river. It spreads across federal states Lower Saxony, Hesse and North Rhine Westphalia, the elevation of this region is up to 527.8 m above sea level. Important cities within Weserberger Land region include Bad Karlshafen, Holzminden, Höxter, Bodenwerder, Hameln, Rinteln, and Vlotho. In addition to the whole of the Weser Valley between Hann. Münden und Porta Westfalica, several geologically associated, but clearly separate chains of uplands, ridges and individual hills are considered part of the Weserberger Land. This region roughly coincides with the natural region of the Lower Saxon Hills. The largest contiguous area of forest in the Weserberger Land lies on the Solling in the Solling-Vogler Nature Park.

The Bergisches Land is a low mountain range region within the North Rhine Westphalia with the highest elevation of 519.2 m above sea level.It lies in the east of Rhine river and south of city Ruhr. The landscape of this region is shaped by woods, meadows, rivers and creeks and 20 artificial lakes are contained. The biggest city in this region is Wuppertal, which is also the capital of this region.

The Westphalian Bay is a flat landscape whose most parts locating in Westphalia with only small southwestern parts in North Rhine and small northern parts in Lower Saxony. Together with its neighbor Lower Rhine Plain to the west, it represents the second most southerly region of the North German Plain after the Cologne Bay. The Westphalian Bay consists of the individual regions of Münster Land, the Emscher Land in the west and south, and regions even further south that flank the Sauer Land around the Hellweg. The Westphalian Bay is bounded by ridges of the Lower Saxon Hills from the northeast to east and by the northern part of the Süder berger Land to the south. These hills rise to heights of several hundred meters above the basin floor.

3.3 Land Covers

The land cover distribution of North Rhine Westphalia is presented as in Figure 3.10. Through the figure, about 50 % area of North Rhine Westphalia is identified as agricultural area which densely locate in the northern low land region, and 26 % area is forest region which distributes in southern Sauer and Sieger Land Region as well as Eifel region, and the others are artificial region (e.g. settlements, industrial area, traffic area) which distributes along the Rhine valley.

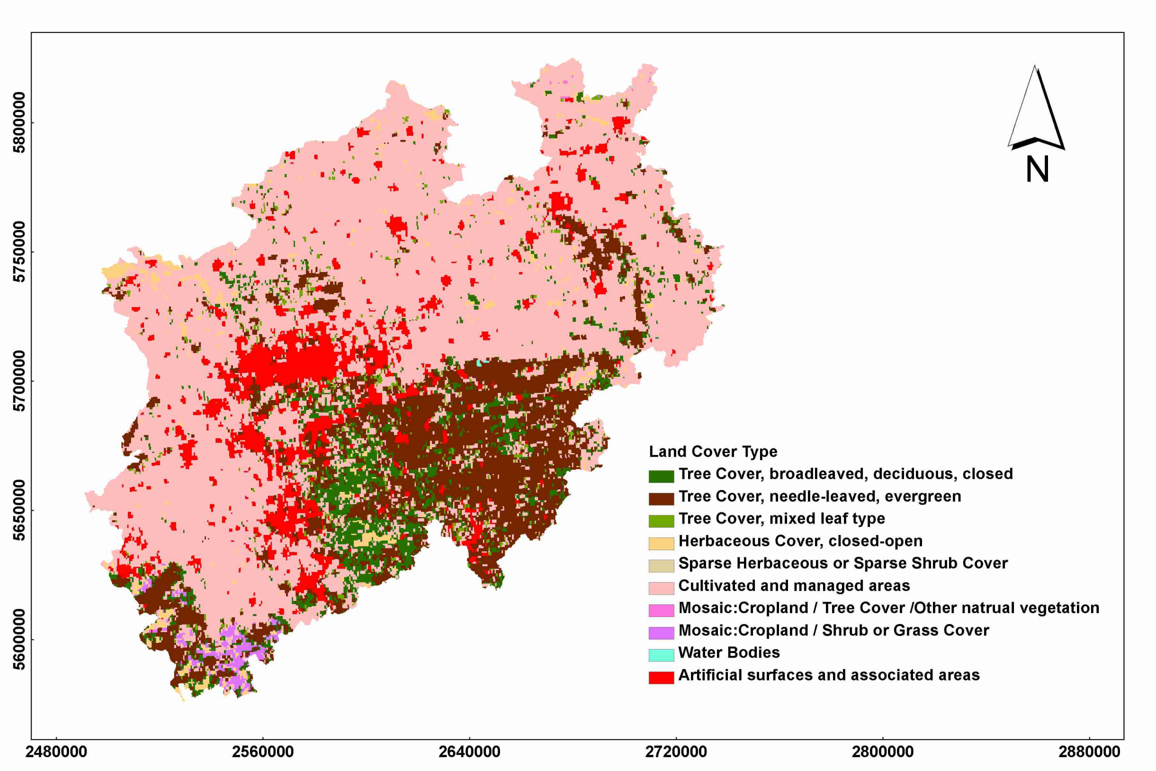


Figure 3.10 Land Cover types over the North Rhine Westphalia. Data sources are from European Environment Agency (EFA) (2015). http://www.eea.europa.eu/data-and-maps/data/nationally-designated-areas-national-cdda-10/.

Heavy industry is the main source of income of the state's economy. The relatively high costs and decreasing competitiveness of many of the Ruhr's heavy industries, however, including coal mining and metallurgy, have led to a concerted effort by the region to change both its economic structure and its image. By the early 21st century the state had succeeded in establishing itself as one of Germany's most important high-technology centers. The state's service industries are also increasingly highly developed.

Many commercial enterprises, trading houses, loan societies, and banks contribute to the state's economy. Outside of the Ruhr, much of the state's land is devoted to commercial farms, gardens, or orchards. Wheat and sugar beets are grown in the southern lowlands. In the north, fruits and vegetables are cultivated. In Münsterland and in the regions of the lower Rhine, cattle raising and pig breeding play considerable roles.

CHAPTER 4

METHODS

In this chapter, a precipitation object identification and tracking algorithm-RCIT was developed. The Semi-Lagrange based advection scheme was introduced which considers the diversity of velocity fields and then applied for doing radar based QPF. The position tracking and predicting algorithm- Kalman Filter was introduced and applied for the prediction of rain cluster's center of mass. Additionally, a characteristic predicting method which takes into account the growth & decay stages of precipitation objects over its life cycle was proposed. Two object based spatial rainfall verification methods- 'SAL' and 'Geometric Index' methods were introduced for evaluating the performance of RCIT algorithm and of QPF methods.

4.1 Data Sources

Radar images from three rainy days were employed in this study which were produced from the C-band operational Essen radar belonged the DWD radar network. The Essen radar works with a beam width approximate 1^o by precipitation scanning way with a range of 128 km and elevation angle of 0.8^o in five minute frequency. Radar reflectivity are indicated by $Z(r, \theta)$ where r is the radar covering range and θ is the azimuth. Their resolution are 1 km and 1^o separately. Figure 4.1 presented the daily accumulative rainfall over the NRW calculated from radar images sorted by rainy days (26-May-2007, 19-Jul-2008 and 26-Jul-2008). Then, a radar data process package - Wradlib developed by Heistermann et al. (2013) was applied, by which the radar images with polar coordinate are projected into a 256×256 km^2 Cartesian grid map with 1 km spatial resolution. The output of this procedure has been the reflectivity Z(i,j) in Cartesian map, whereas i and j are the horizontal and vertical coordinates, respectively. Finally, the power law relationship between radar reflectivity and rainfall intensity is used to transfer radar pixel's reflectivity to the rainfall intensity, the coefficients a and b are set to 256 and 1.42 defined by German Weather Service (DWD).

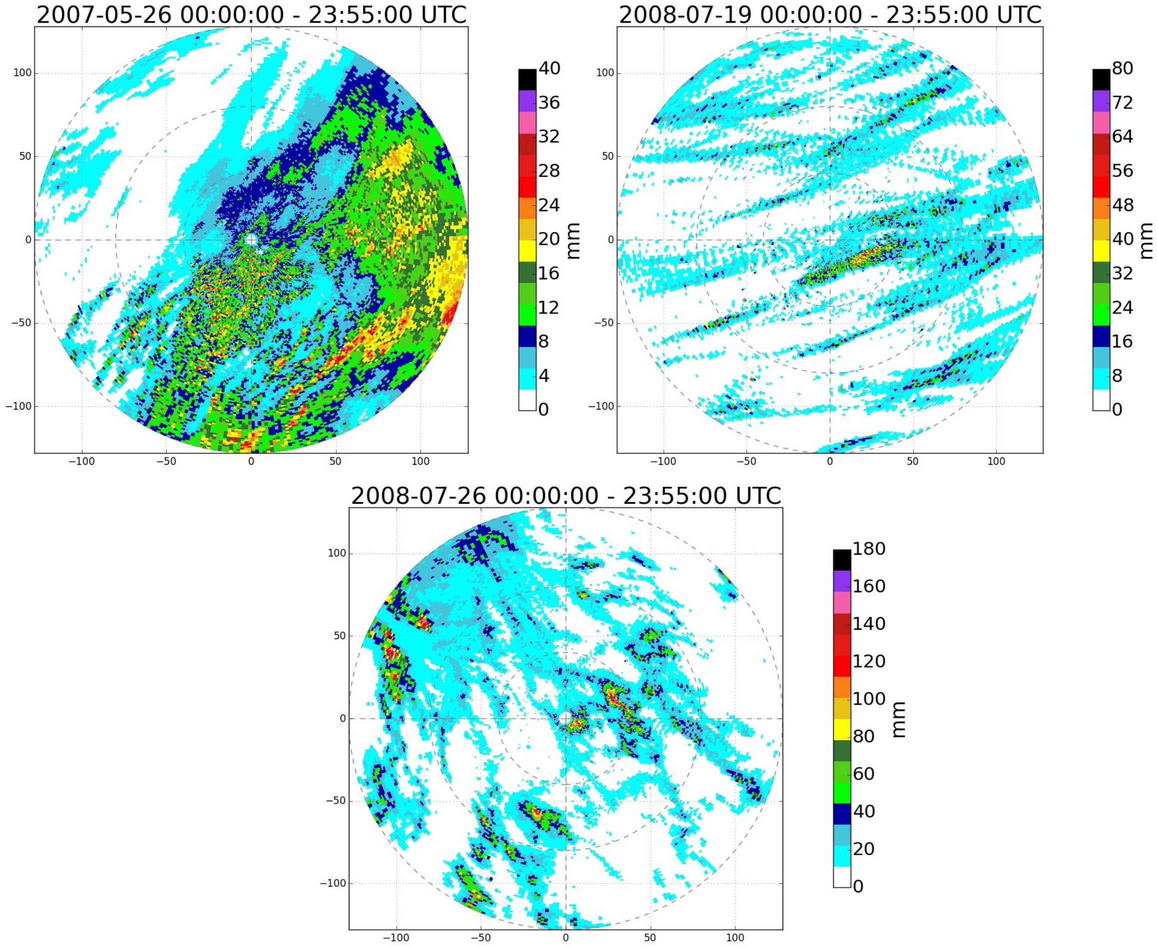


Figure 4.1 Calculated daily accumulated rainfall over the NRW. The upper left is for rainy day 26-May-2007 and the upper right is for rainy day 19-Jul-2008, the bottom is for rainy day 26-Jul-2008

4.2 Precipitation Object Identification and Tracking Algorithm - RCIT

Precise identification and tracking of precipitation object are the key requirement in radar based precipitation applications. Existed tracking algorithms are mainly developed for deep convective systems, primarily for purposes of extreme events' monitoring, nowcasting and warning. Examples of such algorithms include the KONRAD (Lang, 2001), the GANDOLF (Bowler et al., 2004) as well as the TITAN and the TREC. A broad overview has been put forward by researchers concerning various methods about storm tracking and nowcasting using radar, satellite or lighting data. Some of these methods use two dimensional radar images for pattern recognition (Bellon and Austin, 1984) or cross-correlation techniques (Tuttle and Foote, 1990). Others consider three-dimensional storm entities, for example to perform a so-called centroid tracking algorithm (Austin and Bellon, 1982). Within the researching range of 'World Weather Research Programmed Forecast Demonstration Project' (Sydney 2000), Pierce et al. (2004) performed a statistical and studied-oriented comparison of four different radar based precipitation forecasting schemes. They concluded that the centroid tracking and pattern-matching based extrapolation techniques are the most reliable in convective scenarios. Old techniques use a simple extrapolation to determine the new object position, recent techniques more concentrate on dealing with storm's initiation, growth and dissipation (e.g. Germann et al., 2006; Radhakrishna et al., 2012; Ruzanski and Chandrasekar, 2012). Some of these studies also consider the merging or splitting of rain clusters over their life cycles, which are critical for rain cluster's analysis (e.g. Moseley et al., 2013; Shimizu and Uyeda, 2012).

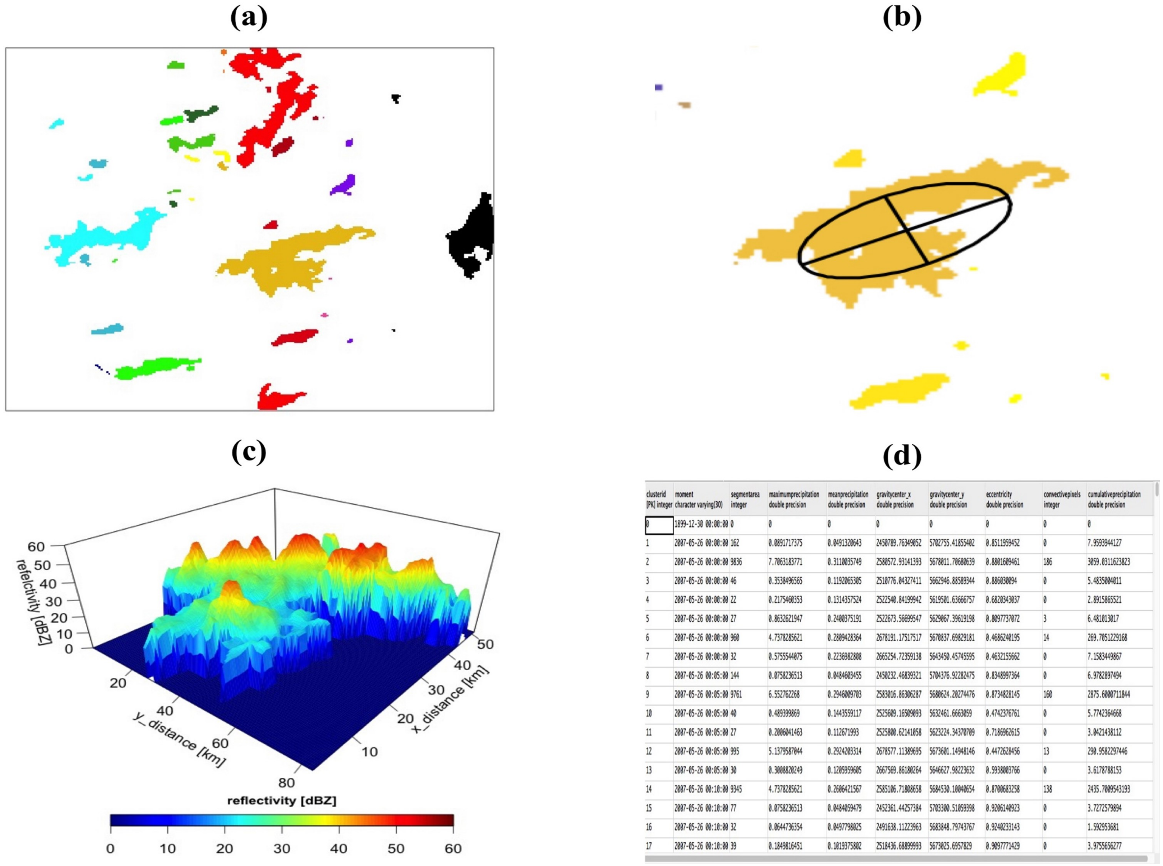
Under convective conditions, individual rain clusters are organized into meso-scale convective system(MCS) that are associated with severe precipitation, and exhibit organized structures. Under this knowledge, rain clusters can be modeled according to the known shape (e.g. circle or ellipse), and their characters can also be easily derived.

In this chapter, a precipitation object identification and tracking algorithm- RCIT (Rain Cluster Identification and Tracking) was proposed to accommodate the needs of radar based QPE and QPF. Processes for constructing the proposed algorithm were described in section 4.3 and 4.4 respectively.

4.3 Rain Cluster Identification Module

The objective of rain cluster identification is to identify and group connected rainy pixels from radar images, then to extract the physical and geometry characteristics of them. This procedure is not a simple task. The reason is that a rain cluster can be divided into a single group at various scales or embedded in a larger strati-form region. Commonly, identification of rain cluster can be done by two or three dimension radar data set.

In this study, the image processing techniques are applied for operating rain cluster's identification and their characteristic's extraction from radar images. Firstly, a median filter algorithm was employed for reducing noisy pixels containing in the radar image (subsection 4.3.1); secondly, a segmenting algorithm for identifying connected rainy pixels in radar image was developed, physical and geometric characteristics were derived from the identified rain clusters by ellipse fitting (subsection 4.3.2); finally, these characteristics were stored in a relation database for further analysis. Figure 4.2 presented an illustration for rain cluster identification module.



clusterid [PK] integer	moment character varying(30)	segmentarea integer	maximumprecipitation double precision	meanprecipitation double precision	gravitycenter_x double precision	gravitycenter_y double precision	eccentricity double precision	convectivepixels integer	cumulativeprecipitation double precision
0	1899-12-30 00:00:00	0	0	0	0	0	0	0	0
1	2007-05-26 00:00:00	162	0.0891717375	0.0491320643	2450789.76349052	5702755.41855402	0.8511959452	0	7.9593944127
2	2007-05-26 00:00:00	9836	7.7063183771	0.3110035749	2580572.93141393	5678011.70680639	0.8801609461	186	3059.0311623823
3	2007-05-26 00:00:00	46	0.3538496565	0.1192065305	2510776.04327411	5662946.88589344	0.886030094	0	5.4835004011
4	2007-05-26 00:00:00	22	0.2175460353	0.1314357524	2522540.84199942	5619501.63666757	0.6820343037	0	2.8915865521
5	2007-05-26 00:00:00	27	0.8632621947	0.2400375191	2522673.56699547	5629067.39619198	0.8097737072	3	6.481013017
6	2007-05-26 00:00:00	960	4.7378285621	0.2809428364	2678191.17517517	5670837.69829181	0.4686240195	14	269.7051229168
7	2007-05-26 00:00:00	32	0.5755544075	0.2236982808	2665254.72359138	5643450.45745595	0.4632155662	0	7.1583449867
8	2007-05-26 00:05:00	144	0.0758236513	0.0484603455	2450232.46839321	5704376.92282475	0.8348997364	0	6.9782897494
9	2007-05-26 00:05:00	9761	6.552762268	0.2946009703	2583016.86306287	5680624.20274476	0.8734828145	160	2875.6000711844
10	2007-05-26 00:05:00	40	0.489399869	0.1443559117	2525609.16509093	5632461.6663059	0.4742376761	0	5.7742364668
11	2007-05-26 00:05:00	27	0.2006041463	0.112671993	2525800.62141058	5623224.34370709	0.7186962615	0	3.0421438112
12	2007-05-26 00:05:00	995	5.1379587044	0.2924203314	2678577.11389695	5673601.14948146	0.4472628456	13	290.9582297446
13	2007-05-26 00:05:00	30	0.3008820249	0.1205959605	2667569.86180264	5646627.98223632	0.5938003766	0	3.6178788153
14	2007-05-26 00:10:00	9345	4.7378285621	0.2606421567	2585106.71808658	5684530.10040654	0.8700683258	138	2435.7009543193
15	2007-05-26 00:10:00	77	0.0758236513	0.0484059479	2452361.44257384	5703300.51059398	0.9206140923	0	3.7272579894
16	2007-05-26 00:10:00	32	0.0644736354	0.0497798025	2491638.11223963	5683848.79743767	0.9240233143	0	1.592953681
17	2007-05-26 00:10:00	39	0.1849816451	0.1019375802	2518436.68899993	5673025.6957829	0.9097771429	0	3.9755656277

Figure 4.2 The illustration of a rain cluster identification module for radar image at 16:55:00 on 19-July-2008. (a) the rainy pixels were segmented and labeled with different colors; (b) an ellipse shape was fitted to each segment; (c) the properties of the identified rain clusters were extracted; (d) the properties were stored in a relational database.

4.3.1 Median Filtering for Noise Smoothing

Through the observations from radar images, there were a number of isolated pixels existed with over high or low reflectivity. These pixels are far away from the connected pixel's center and a large blank area will present when grouping them into any connected pixels. For solving this problem, a median filter algorithm was applied. The median filtering is a nonlinear process which is useful in reducing impulsive, or salt and pepper noise (Ekstrom, 1984). For the median filter algorithm, a predefined size window slides along the image, and the median intensity value of the pixels within the window becomes the output intensity of the pixel being dealt with. Unlike low-pass filtering algorithm, the median filtering algorithm can preserve discontinuities in a step function and can smooth limited pixels whose value is modified significantly from their surroundings without affecting the true value of neighboring pixels.

Median Filter algorithm is operated on two dimensional radar images with Cartesian coordinates. Median Filter algorithm replaced each grid's value by a grid in its neighborhood that has the median total intensity, averaged over all channels, then the median value in r×r blocks centered on each pixel are found (in this study, r is chosen to 3 for convenient of calculation) and smaller neighborhoods are used at the edges of an image. Figure 4.3 presented error distributions between original radar images and filtered ones.

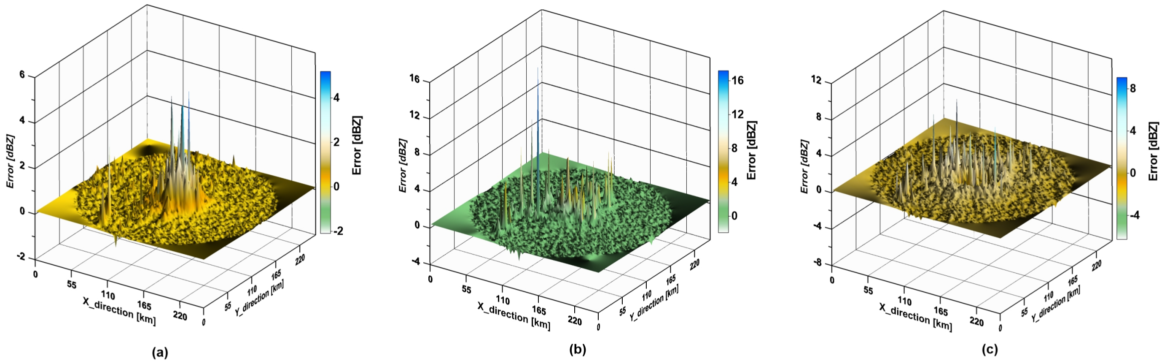


Figure 4.3 Averaged error distributions of filtered radar images, sorted by three rainy day. (a) for radar images from 26-May-2007, (b) for ones from 19-Jul-2008, (c) for ones from 26-Jul-2008.

4.3.2 Extracting Characteristics of Rain Cluster

Connected rainy pixels were determined from the filtered radar images, this was done by image segmenting method by which all the rainy pixels with same reflectivity are assembled to a cluster. In detail, rainy pixels sharing the same reflectivity value were labeled and clustered, then the contour of the cluster was extracted by Edge detection method (Figure 4.2a). Selection of the rainy pixels with same reflectivity obeyed the following rules:

- If the reflectivity of rainy pixel is lower than a given threshold then it is set to null.
- For each rainy pixel and its eight neighbors, if there are more than five ones are null, then it is set to null.
- If the pixel is 'spur', then the pixel is set to null. Here, the 'spur' pixels are those small, extraneous lines left after the skeletonization process for a binary image (Chanda and Majumder, 2004).
- If the area of identified segment is smaller than 9 km^2, then it is ignored.

Reflectivity threshold's selection for segmentation process was based on the radar reflectivity threshold level from Kronenberg et al. (2012) as in Figure 4.4. Through the figure, different radar reflectivity threshold level was marked with color, and each level was corresponded to different rainfall types. Here, 19 and 37 dBZ were chosen as the light and convective rain clusters, respectively. By this selection, rain clusters in rainfall events could be categorized as light and convective types which contained more rainfall information.

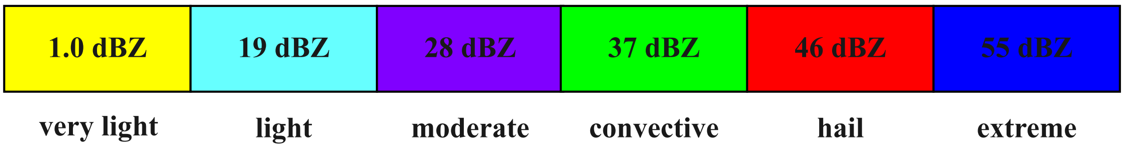


Figure 4.4 Radar reflectivity threshold level and corresponding rainfall types.

Then, an ellipse shape was fitted to the labeled segments (Figure 4.2b) and the physical&geometry characteristics were extracted (Figure 4.2c). The selected characteristics of identified rain cluster were defined as follows:

- **Area** [km^2]: number of pixels for identified rain cluster;
- **Cumulative rainfall** [mm]: areal total rainfall of the identified rain cluster per 5 minutes. For each rainy pixel contained in the identified rain cluster, the Z-R relationship-$Z = 256R^{1.42}$ is applied for converting reflectivity value into rain rate (in mm/5min), then the total rainfall of identified rain cluster is achieved by summing up rain rates of these pixels.
- **Max rain rate per 5 minutes** [mm]: peak rain rate of rainy pixel contained in identified rain cluster;
- **Mean rain rate per 5 minutes** [mm/km^2]: areal averaged precipitation of identified rain cluster, it is the ratio of rain cluster's cumulative rainfall and area as in equation (4.1):

  $$\mathbf{R_{mean}} = \frac{\mathbf{R_{total}}}{\mathbf{A}} \tag{4.1}$$

  where R_{mean} is the mean rain rate of the identified rain cluster, R_{total} and A represents the cumulative rainfall and area of rain cluster, respectively;
- **Number of convective pixels**: it is defined for reflecting internal growth or decay of rain cluster at different threshold levels. Here, according to the reflectivity threshold

definition presented in Figure 4.4, 37 dBZ is used as the threshold for calculating number of convective pixels of light rain clusters while 46 dBZ is used to confirm the number of convective pixels for convective rain cluster.

- **Eccentricity**: the ratio of minor axis length (l_{minor}) and major axis length (l_{major}) of identified rain cluster as in equation (4.2), it is used to describe how much the rain cluster shape approximates to a circular with a value range between 0 and 1. The greater of the value, the more ellipse shape that rain cluster looks like. Here, l_{minor} and l_{major} is calculated from the fitted ellipse shape of identified rain cluster.

$$\mathbf{Eccentricity} = \sqrt{\mathbf{1} - \frac{\mathbf{l_{minor}}^2}{\mathbf{l_{major}}^2}} \tag{4.2}$$

- **Center of mass** [km]: the center position of identified rain cluster which is weighted by rainy pixel's reflectivity contained in the identifed rain cluster as in equation (4.3).

$$\mathbf{C(X,Y)} = \frac{\sum\limits_{\mathbf{i=1}}^{\mathbf{n}}\sum\limits_{\mathbf{j=1}}^{\mathbf{m}} \mathbf{P(x_i,y_i)R(x_i,y_i)}}{\sum\limits_{\mathbf{i=1}}^{\mathbf{n}}\sum\limits_{\mathbf{j=1}}^{\mathbf{m}} \mathbf{R(x_i,y_i)}} \tag{4.3}$$

where C(X,Y) is the center of mass for the identified rain cluster, P(x_i, y_i) is the position of pixel contained in the identified rain cluster and R(x_i, y_i) is the pixel's reflectivity.

For all the extracted characteristics, they were finally stored in a relational database. In this study, an open source database processing system -PostgreSQL (Momjian, 2001) was applied for storing these characteristics (Figure 4.2d).

4.4 Rain Cluster Tracking Module

The aim of rain cluster tracking module is to connect rain clusters at different moments to ensure that they are into a series of consecutive time chains, and furthermore, to evaluate the spatial - temporal variability of their characteristics. As mentioned at the beginning of this chapter, most cell based tracking algorithms are developed for compatible with convective scenarios, or combined with other resources such as cloud-top, wind or temperature from numeric model output or satellite observations (e.g. Meyer et al., 2013; Nisi et al., 2014; Tüchler and Meyer, 2013; Wood-Bradley et al., 2012). Research domains where those algorithms applied are usually in a large spatial scale.

Another aspect which is merely considered for those algorithms is that how to cope with splitting and merging scenarios and how to evaluate growth and decay of rain clusters over their life cycles. Merging is one of the important phenomena for convective precipitating system development (Westcott, 1994). A merged rain cluster grows larger in shape and produces more precipitation than single ones. Some studies addressed the merging effect on super cells, squall lines and meso-scale convective systems (e.g. Lee et al., 2006a,b; Meixner, 2002). Compared to the merging phenomenon, fewer studies have investigated splitting of rain cluster. Bluestein et al. (1990) documented the initiation and behavior of splitting convective clouds, and concluded that this split is the result dynamic forcing effects rather than to rain water loading. Jewett and Wilhelmson (2006) investigated the role of splitting rain cells within squall lines using numerical simulations.

The tracking procedure of rain cluster can be categorized as pixel and object based types. For pixel based tracking approach, atmospheric conditions are based on Eulerian, pixel-based perspective. Here, the pixel size is set to be the radar's spatial resolution. An object based approach usually includes three general steps: storm identification; storm tracking and storm projection. For storm tracking step, rain clusters at former and next moment are connected with some matching rules (e.g. overlaps between rain clusters, area difference and distance between their center of mass).

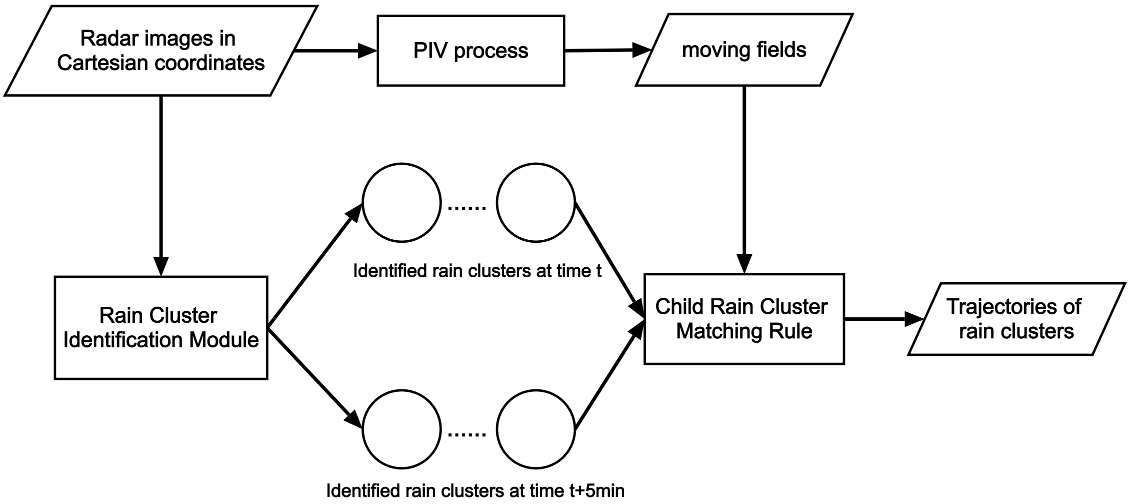


Figure 4.5 An processing illustration of rain cluster tracking module.

In this study, an object-based rain cluster tracking algorithm was developed (Figure 4.5). Firstly, the Particle Image Velocimetry was applied for deriving global motion vectors of two consecutive raw radar images (raw radar images are the ones without Median filtering process and no reflectivity threshold is set); secondly, at each time interval, rain clusters were identified and their characteristics were extracted; then, a sequence of rain cluster matching rules were implemented for identifying the most matched child rain clusters at the next time interval; finally, the rain clusters identified at two successive moments were connected according to their parent-children relationship as a time chain and stored into the relational database.

4.4.1 Particle Image Velocimetry

The Particle Image Velocimetry (PIV) is an optical method of flow visualization which is used to obtain instantaneous velocity measurements and related properties in fluids. The fluid is seeded with tracer particles which, for small particles which are enough sufficiently are assumed to faithfully follow the flow dynamics (the degree to which the particles faithfully follow the flow is represented by the Stokes number). The fluid with entrained particles is illuminated so that particles can be visible. The motion of the seeding particles is used to calculate speed and direction (the velocity field) of the flow being studied. The PIV method comprises a class of flow measuring techniques that are characterized by the recording of the displacement of small particles embedded in a region of fluid (Jahanmiri, 2011).

The PIV method can also be said as a consequence of the development in flow visualization techniques, which have contributed so importantly to the understanding of fluid flow phenomena. Similar with what have been introduced in Chapter 2 (see section 2.1 in Chapter 2), the PIV method can be selected as a kind of advection field tracking method for advection field's estimation at various spatial scale. In this study, the PIV method is implemented for simulating global motion vectors of consequent radar images, Figure 4.6 presented an simple illustration of the process for generating the global motion vectors by the PIV method.

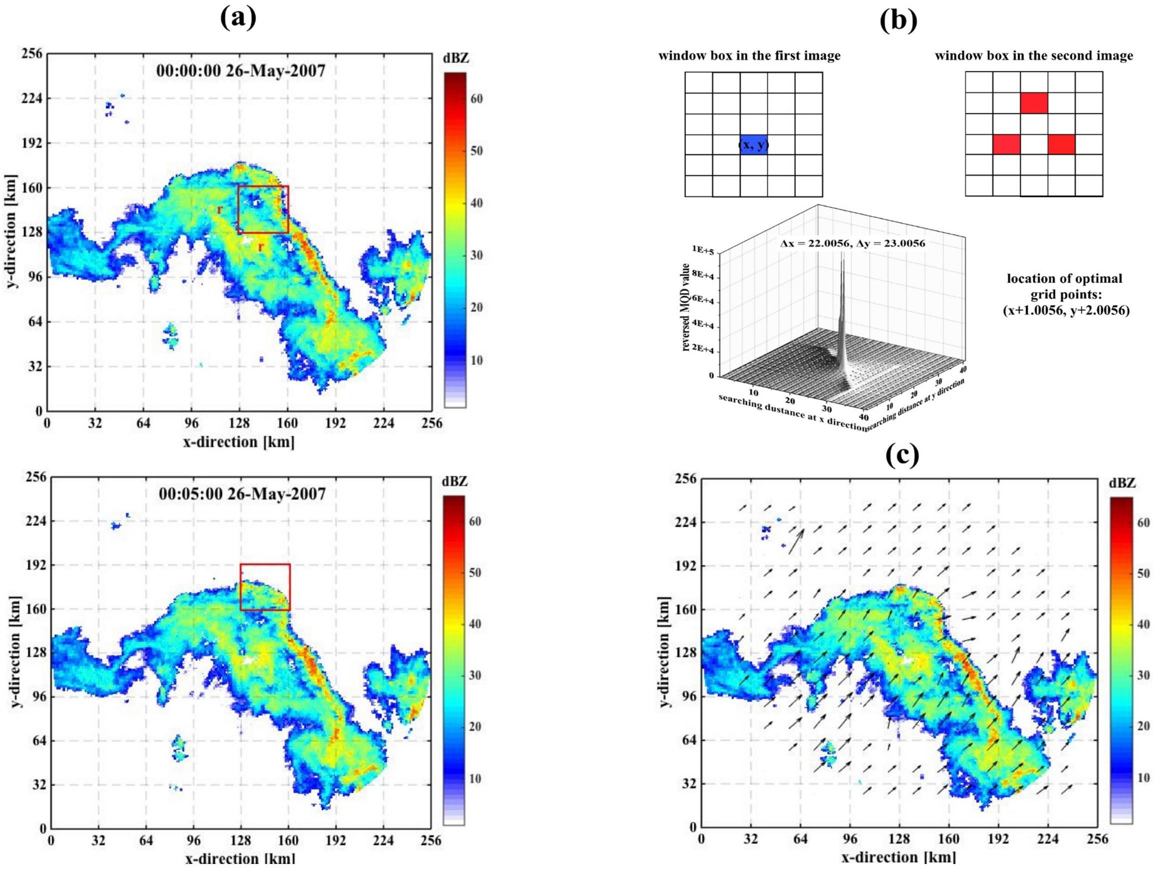


Figure 4.6 An simple illustration of the application of PIV method in global motion vector identification for radar images at 00:00:00 and 00:05:00 on 26-May-2007. (a) window boxes with the area of $r \times r$ are defined (rectangular with red color); (b) for any grid point in the window box at previous time (blue block), the MQD algorithm is applied for deducing its minimum reflectivity differences with grid points within any window box at next time (red blocks), the results are reversed and the solitary peak locations - Δx and Δy are corrected, the value of corrected Δx and Δy are 22.0056 and 23.0056 respectively, locations of the optimal grid point at time 00:05:00 on 26-May-2007 are calculated; (c) the global motion vectors are extracted and smoothed by median filter method.

It was similar to the TREC tracking algorithm and its improved form - COTREC, an interrogation box was defined firstly when applying the PIV. The interrogation box can be selected at various size, but default setting was 32×32 km^2 in this study. At each time interval, two consequent radar images f_1 and f_2 were selected, then the interrogation boxes with a size of R×R km^2 were defined which divide these radar images into a series of sub-regions. Here, the interrogation box with a size of 32×32 km^2 was confirmed which means that the whole radar image can be divided into 64 sub-regions (figure 4.6a).

In Figure 4.6b-c, a Minimum Quadric Differences (MQD) method was employed for extracting the optimal grid points at next time as in equation 4.4:

$$\mathbf{MQD(\Delta X, \Delta Y)} = \sum_{\mathbf{i=1}}^{\mathbf{N}} \sum_{\mathbf{j=1}}^{\mathbf{N}} \left| \mathbf{R_1(X_i, Y_j) - R_2(X_i + \Delta X, Y_j + \Delta Y)} \right| \tag{4.4}$$

where $R_1(X_i, Y_j)$ and $R_2(X_i, Y_j)$ are the reflectivity of grid points contained within the interrogation boxes of radar images at time t and t + Δt, respectively; Δx and Δy are the locations of minimum reflectivity difference at horizontal and vertical directions, respectively.

When applying the MQD algorithm, firstly, a searching distance - $d = 2 \times v_{\max} + 1$ was defined where $v_{\max}$ is the pre-set maximum velocity; then, the equation 4.4 was used for deducing the minimum reflectivity values of grid points, and Δx and Δy were calculated ($\Delta x, \Delta y \in d$). For simplified calculation, the reflectivity differences of grid points within window box were reversed. For guaranteeing the solitary peak locations could be calculated, Δx and Δy were corrected separately for 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. By this way, the optimal grid points at time t + Δt were identified whose locations could be presented as - ($x + \Delta x - \frac{d+1}{2}, y + \Delta y - \frac{d+1}{2}$). Finally, calculated global motion vectors were smoothed by the median filter algorithm.

Comparing to advection field tracking algorithms which are based on the maximum cross-correlation approach. The extracted global motion vectors based on the PIV method was more continuous and consistent. Through the practical applications, the cross-correlation based tracking algorithms do not work well for radar images which contain no rainy pixels (speckle images). Under such conditions, no global motion vectors will be generated. However, the calculating costs for tracking method proposed in this study was higher than those classical algorithms.

4.4.2 Rain Cluster Matching Rule

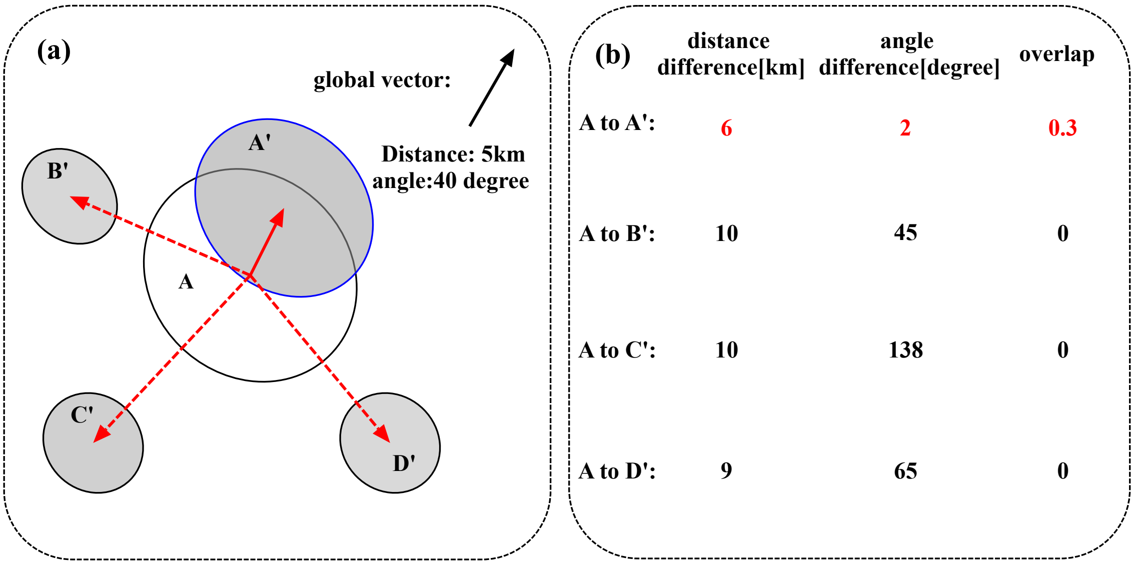
Before the introduction of rain cluster matching rule, some definitions were given firstly: For two radar images at two successive moments, rain clusters identified from previous moment's radar image are called parent clusters; rain clusters from next moment's radar image are called child clusters. For each identified parent cluster, rain cluster matching rule is employed to finding its most matched child clusters. Figure 4.7 illustrated the child rain cluster matching rule applied in this study. Here, the time lag of two moments is five minutes which is also the time interval of radar scanning. Main steps are listed as follows (more detailed information can be found in Appendix A):

(1) A boundary box is defined for each parent cluster; the range of the box should be larger than the size of the parent cluster. Here, the horizontal length of the box is the maximum horizontal length of the parent cluster plus 20 kilometers as well as the vertical length of the box.

(2) Number of child clusters falling into the box is counted. For each child rain cluster falling into the box, four characteristics are selected: area, cumulative rainfall, max and mean rain rate per 5 minutes. Another important characteristic in this step should be considered is the center of mass. By deriving center of mass for each child rain cluster, motion difference between child rain clusters and parent cluster could be obtained easily. Some additional characteristics, such as number of convective pixels and eccentricity are not considered because they are not obvious effects on the tracking results.

(3) For each child cluster falling into the box, there are two scenarios for finding the most matched child cluster: if there is only one child cluster falling into the box, and if it overlaps with the parent cluster, then it is the most matched child cluster; otherwise, following conditions are considered: firstly, distance and angle difference of center of mass between it and parent cluster is less than a given threshold when comparing to the mean value of derived global motion vector, then it is also the most matched child cluster; secondly, if there are two or more child clusters falling into the box which are not overlapped with the parent cluster, the matched rules are same with the first scenario, but one extra condition is included: calculating area differences between these child clusters and their parent, child clusters with the minimum area difference are the most matched ones.

Taking an example, for any two consequent radar images, in Figure 4.7a, A was the parent cluster which is identified at previous time step, dashed box was the defined boundary box of A, and $A^{'}$, $B^{'}$,$C^{'}$, $D^{'}$ were child clusters which falling into the boundary box of A. Overlaps, speed and direction differences between A and $A^{'}$, $B^{'}$, $C^{'}$, $D^{'}$ were calculated and the results were presented as in Figure 4.7b, according to the defined matching rule, $A^{'}$ should be the most likely child rain cluster for A (red row in Figure 4.7b).



	distance difference[km]	angle difference[degree]	overlap
A to A':	6	2	0.3
A to B':	10	45	0
A to C':	10	138	0
A to D':	9	65	0

Figure 4.7 A simply illustration for most matched rain cluster matching rule, the red solid arrow in (a) is the motion vector between the parent cluster A and its most likely child cluster A'

4.4.3 Life Stages Identification

During the life cycle of one rainfall event, physical and geometry features of rain cluster always varies, as mentioned in Chapter 1 (see section 1.3 in Chapter 1), three common stages can reflect these variations. As a matter of fact, rain cluster's variation is not only associated with its internal growth and decay, but also connected with outer clusters (e.g. merging or splitting). In this study, stages of rain cluster were extended, and five stages of rain clusters were identified during their life cycles, and their identification rules were listed as follows (Figure 4.8):

(a) **Initial**: If there is at least one child cluster but no parent cluster existed, it is a new generated rain cluster.

(b) **Tracking**: If there is only one child cluster and only one parent cluster existed, and this rain cluster passed over its whole life cycle without merging or splitting.

(c) **Merge**: If there are at least two parent clusters but only one child cluster existed, it could be said that it is a merged rain cluster.

(d) **Split**: Contrary to the definition for merge, for one rain cluster, if it has only one parent cluster but at least two child clusters, it could be said that this rain cluster is split.

(e) **Dissipate**: For one rain cluster, if it contained at least one parent cluster, but no child cluster existed, it could be said that this rain cluster is dissipated.

(f) **5 minutes life cycle**: Apart from five stages mentioned above, there are also some rain clusters with life cycle of only 5 minutes which are only observed in one radar image.

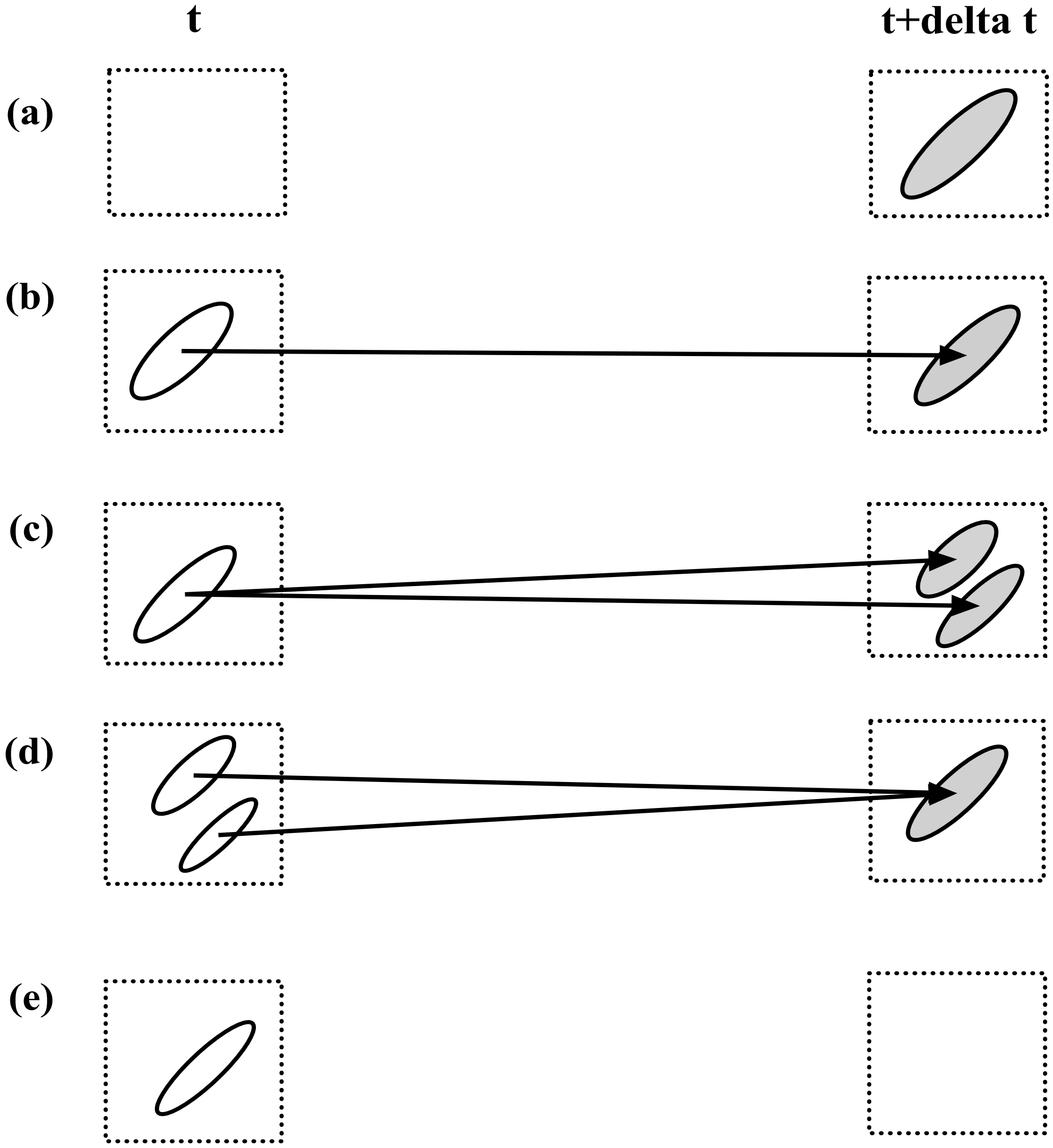


Figure 4.8 Life stages illustration of rain clusters during their life cycle. The transparent and grey parts in the figure presented parent and child rain clusters, respectively. (a) 'initial' stage; (b) 'tracking' stage; (c) 'Merge' stage; (d) 'split' stage; (e) 'dissipate' stage.

4.5 Semi-Lagrange Based Advection Scheme

For simply introduction of the Semi-Lagrange based advection scheme, one dimensional linear advection equation was presented firstly as in equation (4.5):

$$\frac{\partial \mathbf{z}}{\partial \mathbf{t}} + \mathbf{u}\frac{\partial \mathbf{z}}{\partial \mathbf{x}} = \mathbf{0} \tag{4.5}$$

where **u** is the constant coefficient parameter and z is the state of the system with its initial form presented as $z_0(x)$, x∈R. So the equation (4.5) can also be presented as:

$$\mathbf{z}(\mathbf{x},\mathbf{t}) = \mathbf{z_0}(\mathbf{x} - \mathbf{ut}) \tag{4.6}$$

Considering the discretization of equation (4.5) on a uniform one dimensional mesh grid with a spatial scale of Δx and a temporal scale of Δt. When the nodes in the mesh grid is presented by i and time intervals are presented by n, then the space-time locations of the node can be presented by $x_i = i\Delta x$, $t^n = n\Delta t$ with their approximate values computed by a numerical solution are presented as $z_i^n \approx z(x_i, t^n)$. The selected numerical solution is a standard finite difference methods based on approximation of differential operators by finite difference quotients.

Lagrange and Semi-Lagrange based advection scheme exploit instead the special feature of the advection equation, that is the representation of the exact solution in terms of the initial state. Particularly, it can be observed that the following equations hold considering without loss of generality the case of $u > 0$:

$$\mathbf{z}(\mathbf{x_i},\mathbf{t^n}) = \mathbf{z_0}(\mathbf{x_i} - \mathbf{un\Delta t}) = \mathbf{z_0}(\mathbf{x_i} + \mathbf{u\Delta t} - \mathbf{u(n+1)\Delta t}) = \mathbf{z}(\mathbf{x_i} + \mathbf{u\Delta t}, \mathbf{t^{n+1}}) \tag{4.7}$$

with a constant coefficient parameter u and with a initial state $z_0(x)$, x ∈ R. The system state at next time interval can be presented as:

$$\mathbf{z}(\mathbf{x_i},\mathbf{t^{n+1}}) = \mathbf{z_0}(\mathbf{x_i} - \mathbf{u(n+1)\Delta t}) = \mathbf{z_0}(\mathbf{x_i} - \mathbf{u\Delta t} - \mathbf{un\Delta t}) = \mathbf{z}(\mathbf{x_i} - \mathbf{u\Delta t}, \mathbf{t^n}) \tag{4.8}$$

Equation (4.8) provided the basis of Lagrange based advection scheme which is similar with what has been introduced in Chapter 1 (see equation (1.9) in Chapter 1). The special nature of the solution for equation (4.5) allows to use the knowledge of the solution in a mesh grid point at time n to derive the value of the solution at time n+1 at points of a mesh grid that has moved with the flow. Due to the fact that the mesh grid must be changed at each time interval, practical application of Lagrange based advection scheme is not very straightforward.

Equation (4.9) provided the basis of Semi-Lagrange based advection scheme. Correspondingly, the special nature of the solution for equation (4.5) is used to express the value of the solution at time n+1 at the mesh grid points in terms of the solution values at time n at points of a mesh grid that within a single time interval will be transported by the flow onto the computational mesh grid. This has the practical advantage that the mesh grid does not change in time, which is the one of the fundamental reasons for the much more wide-spread use of Semi-Lagrange based advection scheme rather than Lagrange based. Discrete definition of the Semi-Lagrange based advection scheme also can be obtained from equation (4.9).

$$\begin{aligned}
&\mathbf{z_i^{n+1}} = \mathbf{z^n_{i-u\frac{\Delta t}{\Delta x}}} = \mathbf{z^n_{i-k-\alpha}} \\
&\mathbf{u\frac{\Delta t}{\Delta x}} = \mathbf{k} + \alpha \qquad (4.9) \\
&\mathbf{k} = \left[\mathbf{u\frac{\Delta t}{\Delta x}}\right]
\end{aligned}$$

In equation (4.9), k and α are called the integral and fractional courant numbers, respectively. The expression $z^n_{i-k-\alpha}$ is to be interpreted as the value obtained from the approximate z_n values at the point $i\Delta x - u\Delta t$ by some interpolation procedure.

The earliest works which are concerned with methods for the advection equation by applying the propagation along characteristics may be presented by Courant et al. (1952), whose aim was on the numerical solution of the hyperbolic systems. In the meteorological application via Semi-Lagrange based advection scheme, a graphic integration technique was presented by FjØrtoft (1952). Wiin-Nielsen (1959) gave a more detailed analysis. Various scientific contributions have been established which presented the role of methods based on Semi-Lagrange based advection scheme from past decades. Applying the Semi-Lagrange based advection scheme for numerical precipitation forecasting has been developed since 2000's (e.g. Berenguer et al., 2011; Germann and Zawadzki, 2002; Mandapaka et al., 2012; Wang et al., 2013).

In scientific practices of radar based rainfall estimation and forecast, movements of advection fields are not always linearly. A rotating velocity field exists In some cases (e.g. storm events occurred in the tropical region), and the traditional velocity fields extrapolating method like Eulerian and Lagrange based advection scheme will not work so well. Semi-Lagrange based scheme can be solution. Considering equation (4.9), mathematical function of Semi-Lagrangian for radar based rainfall forecasting can be written:

$$\mathbf{z(X, t_0 {+} \tau) = z(X - \alpha, t_0)} \tag{4.10}$$

where z is the pixel's reflectivity or rain rate in position X at time t_0. The extrapolated pixel's reflectivity of rain rate in position X after τ leading time is equal to the pixel's reflectivity or rain rate in position $X - \alpha$ at time t_0. The extrapolated motion vector α is calculated according to the equation (4.11):

$$\alpha = \mathbf{\Delta t u(t_0, x - \frac{\alpha}{2})} \tag{4.11}$$

where $u(t_0, x - \frac{\alpha}{2})$ is the pixel's velocity in position $x - \frac{\alpha}{2}$. When the extrapolation starts, α = 0, the extrapolated motion vector is accumulated by N time intervals where the number N is divided by the extrapolating length τ.

4.6 Center of Mass's Extrapolating Method

According to the introduction in Chapter 2 (see section 2.2 in Chapter 2), center of mess's extrapolation is a kind of cell based QPF approach by which the rain cluster's center of mass at future time is extrapolated according to its locations at previous time intervals (e.g. TITAN and SCOUT system). The premise of most center of mess's extrapolating methods is the assumption that the extrapolated rain cluster at future time moves with a constant speed. Armed with this assumption, the linear extrapolation method is commonly applied. Nevertheless, measured moving speed for rain cluster is not constant in some cases and the extrapolated center of mass at future time exists high bias.

Figure 4.9 illustrated the linear or non-linear trajectories and the linear extrapolating results for their center of masses. Through the Figure 4.9a, for a measured relative linear trajectory, the linear extrapolated results presented high bias (the root mean squared error was [6.5, 13.3] km). It was the same with the non-linear trajectory (Figure 4.9b), and the root mean squared error for the extrapolated rain cluster's center of mass was [10.2, 4.3] km.

In this study, a position tracking and predicting algorithm- Kalman filter was applied and attempted to predict the rain cluster's center of mass taking into account the non-constant moving speeds of identified rain clusters.

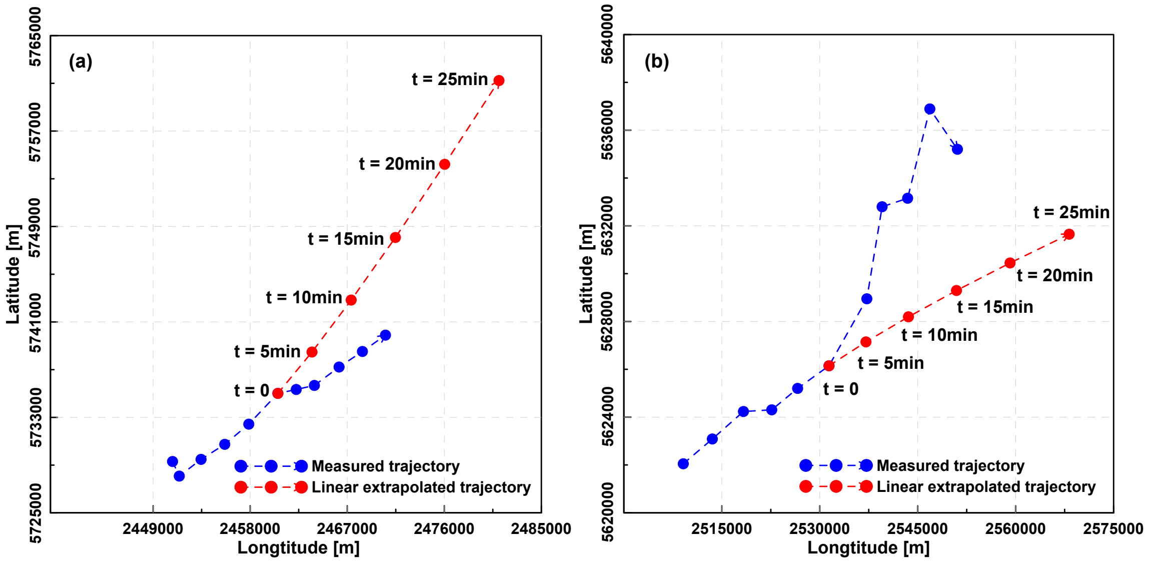


Figure 4.9 An illustration of measured linear (non-linear) rain cluster's trajectories and the linear extrapolating results for their center of masses, the trajectories were extracted from the developed precipitation objects identification and tracking algorithm. (a) for a measured relatively linear trajectory, (b) for a measured non-linear trajectory.

4.6.1 Basic Concepts for Kalman Filter

The concepts of probabilistic analysis in radar based rain cluster tracking is rarely applied. Conversely, it has been widely employed in other scientific domains for a long time (Bar-Shalom et al., 2004). The Kalman filter algorithm can be regarded as a statistical inversion method for solving linear filter of discrete data, in which an unknown state x of the system is estimated combining with some noisy measurements (Figure 4.10).

When applying the Kalman filter algorithm to predict rain cluster's center of mass, each rain cluster's position is treated as a discrete system. The unknown states x can be regarded as an assembly of characteristics (e.g. center of mass, moving speed, area, cumulative rainfall). The goal of kalman filter algorithm is to provide an optimal estimation for x at the next time interval by the noise measurements at the previous time step and a dynamic process for the temporal development of state parameters. In the practical application, the measurements can be acquired from the outputs of precipitation objects identification and tracking algorithm.

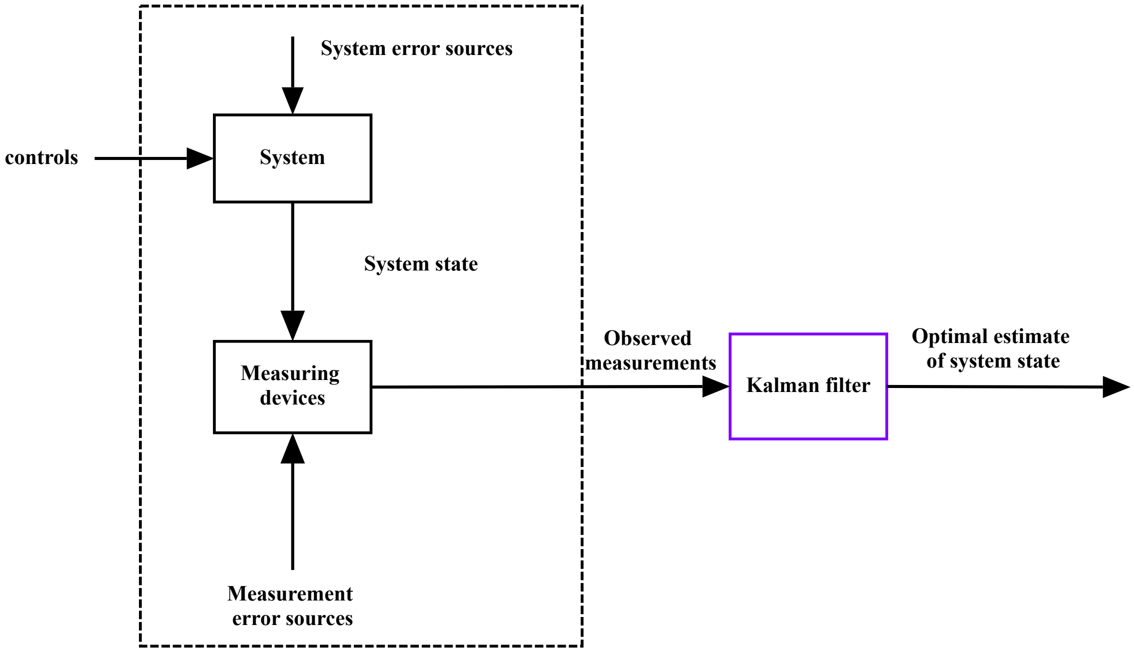


Figure 4.10 An illustration of kalman filter algorithm.

The principle of Kalman filter is that the system states can be presented as a linear stochastic state-space model with Gaussian noises which can be described as:

$$\begin{aligned} \mathbf{X_k} &= \mathbf{F_k X_{k-1} + W_{k-1}} \\ \mathbf{Y_k} &= \mathbf{H_k X_k + v_k} \end{aligned} \tag{4.12}$$

where F_k is the state transition matrix, H_k is the measurement matrix. w_{k-1} and v_k are the noisy state and measurement variables, respectively. F_k defines how the state evolves between successive time steps which can describe a system perfectly comparing to the deterministic model. The noisy variable w_k is added to describe the uncertainties of the system. Thus, w_{k-1} is also called as the processing noise and it assumed to follow the Gaussian distribution with the form of $w_{k-1} \sim N(0, Q_{k-1})$. The measurement matrix H_k defines how the model state is mapped from the defined state variables to the observations. The observation model H_k, in turn, defines how the model state is mapped from the state variables to the measurements. The noisy vector v_{k-1} is the measurement noise, which relates to the uncertainty or bias of the measurements y_k. Same with the process noise, it is also assumed to follow the Gaussian distribution with the form of $v_{k-1} \sim N(0, R_{k-1})$.

In most cases, both v_k and w_{k-1} are both assumed to follow the Gaussian distribution and independent on the initial state x_0.

The kalman filter estimates m_k and its covariance P_k for the unknown state x_k at k_{th} time step are estimated recursively with the following two phases algorithm:

Prediction step:

$$\begin{aligned} \mathbf{m_k^-} &= \mathbf{F_{k-1}\,m_{k-1}} \\ \mathbf{P_k^-} &= \mathbf{F_{k-1}\,P_{k-1}\,F_{k-1}^T + Q_{k-1}} \end{aligned} \tag{4.13}$$

Update step:

$$\begin{aligned} \mathbf{K_k} &= \mathbf{P_k^-\,H_k^T\,(H_k\,P_k^-\,H_k^T + R_k)^{-1}} \\ \mathbf{m_k} &= \mathbf{m_k^- + K_k(y_k - H_k\,m_k^-)} \\ \mathbf{P_k} &= \mathbf{(I - K_k\,H_k)\,P_k^-} \end{aligned} \tag{4.14}$$

The variables m_k^- and P_k^- in equation (4.13) denote the predicted state estimate and the corresponding covariance, which are updated with a given measurement y_k. In equation (4.14), K_k is the Kalman gain, which defines the system state with the initial state of m_0 and covariance of P_0. From Bayesian point of view, the estimates m_k and P_k can be regarded as the mean and covariance of the posterior distribution of the state x_k by a given noisy measurement y_k, the prior mean value - m_k^- and covariance - P_k^-.

4.6.2 Application of Kalman Filter Algorithm

In this study, the Kalman filter algorithm was applied for doing one hour ahead center of mass extrapolation for the identified rain cluster. The state variables were the rain cluster's center of mass with Cartesian coordinates which were marked as (x, y) including the corresponding velocity and accelerate elements which were marked as (v_x, v_y) and (a_x, a_y) separately, then the state matrix of kalman filter model could be written as $[x, v_x, a_x, y, v_y, a_y]^T$. The measurement variable was the rain cluster's center of mass, therefore the elements corresponding to the velocity and accelerate variables in the measurement matrix H_k were set to zeros. The velocity and accelerate was calculated based on the center of masses at previous time steps. Measurement matrix H_k could be presented as:

$$\mathbf{H_k} = \left\{ \begin{array}{cccccc} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{array} \right\} \tag{4.15}$$

and the measurement noise v_k was assumed to be a zero mean Gaussian distribution in this study with the covariance of:

$$\mathbf{R_k} = \left\{ \begin{array}{cc} r^2 & 0 \\ 0 & r^2 \end{array} \right\} \tag{4.16}$$

In equation (4.16), the parameter r describes the noise intensity of the measured elements (center of mass). The state transition model F_k and the process noise covariance Q_k of the discretized acceleration model whose noise was also assumed to follow a zero mean Gaussian distribution are listed in equation (4.17) and equation (4.18)

$$\mathbf{F_k} = \left\{ \begin{array}{cccccc} 1 & dt & dt^2 & 0 & 0 & 0 \\ 0 & 1 & dt & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & dt & dt^2 \\ 0 & 0 & 0 & 0 & 1 & dt \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right\} \tag{4.17}$$

$$\mathbf{Q_k} = \mathrm{q} \left\{ \begin{array}{cccc} \frac{1}{3}dt^3 & \frac{1}{2}dt^2 & 0 & 0 \\ \frac{1}{2}dt^2 & dt & 0 & 0 \\ 0 & 0 & \frac{1}{3}dt^3 & \frac{1}{2}dt^2 \\ 0 & 0 & \frac{1}{2}dt^2 & dt \end{array} \right\} \tag{4.18}$$

where dt is the time interval and q defined the processing noise intensity which could help to choose a reasonable parameter for the processing noise. Similar with the measurement noise, the value of q can be chosen subjectively.

4.7 Rain Cluster's Characteristics Predicting Method

At the beginning of Chapter 1, three stages of rain cluster over its life cycle were illustrated and they can be displayed as cumulus, mature and dissipation (see section 1.3 in Chapter 1). During rainfall events, one rain cluster is generated and then grows with its size enlarged and rain rate increased. When the inner structures become saturated, the growth of rain cluster slows and keeps stable, then it is in the mature stage. At the end of the rainfall events, the size and rain rate of rain cluster decrease till it disappears and this is known as the dissipation stage.

Figure 4.11 presented a conceptual illustration of three stages of rain cluster over its life cycle where the vertical axis was the mean value of rain cluster's characteristics and the horizontal axis was the duration of rain cluster.

Through the Figure 4.11, if we normalize all the rain cluster's trajectories, then a summary of stage variability of rain cluster's characteristics following the normalized durations. For example, if the rain cluster's trajectories with different durations are extracted, then the 'Normalized Duration Lines' of their characteristics can be calculated where the range of normalized duration is set from 0 to 1 and we can give a heuristic summary that the growth stage of a rain cluster is between 0 and 0.4, the mature stage is between 0.4 to 0.6 and the dissipation stage occurs between 0.6 to 1.

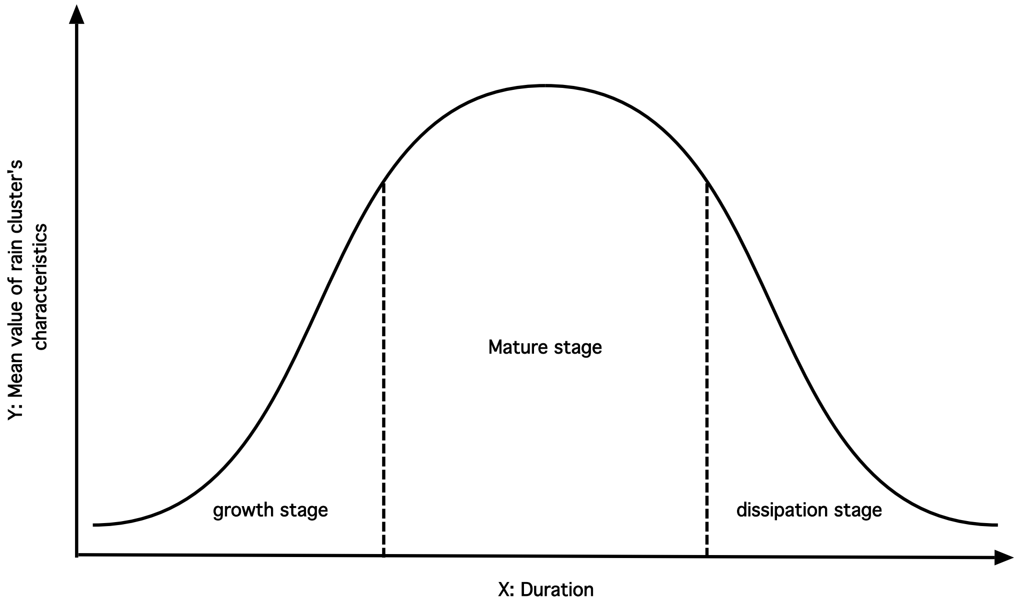


Figure 4.11 An conceptual illustration of three stages of rain cluster over its life cycle.

Trajectories with multi durations can be normalized into one standard time axis by the 'Normalized duration lines' method and the temporal developments for rain cluster's characteristics can be presented. By applying regressive fitting method, the fitted curve can be calculated and presented as a parabola function, the coefficient parameters a and b of the parabola function can also be confirmed as in equation (4.19):

$$\mathbf{V}_{\mathbf{characters}} = \mathbf{a}\,\mathbf{t}_{\mathbf{normal}}^{\mathbf{2}} + \mathbf{b}\,\mathbf{t}_{\mathbf{normal}} \tag{4.19}$$

where $V_{characters}$ is the normalized mean value of rain cluster's characteristics, a and b are the coefficient parameters of the fitted parabola function, respectively. t_{normal} is the normalized duration of rain cluster. According to the equation (4.19), predicting characteristics of rain cluster can be operated by equation (4.20):

$$\mathbf{V}_{\mathbf{forecast}} = \mathbf{a}\,\mathbf{V}_{\mathbf{now}}\,(\mathbf{t}_{\mathbf{normal}} + \frac{\mathbf{t}_{\mathbf{forecast}}}{\mathbf{T}})((\overline{\mathbf{t}_{\mathbf{normal}}} + \frac{\mathbf{t}_{\mathbf{forecast}}}{\mathbf{T}}) - \frac{\mathbf{b}}{\mathbf{a}}) \tag{4.20}$$

where $V_{forecast}$ is the predicted value at leading time $t_{forecast}$, $\overline{t_{normal}}$ is the median value of the normalized duration for three stages which are presented in Figure 4.11 (growth stage: 0.2, mature stage: 0.5 and dissipation stage: 0.6). V_{now} is the mean value of rain cluster's characteristics at previous time intervals. $T = \frac{t_{now}}{t_{normal}}$ is the ratio of current time t_{now} and median value of the normalized duration. As for the practical application, the V_{now} is averaged by the previous time interval's characteristic's values before the forecast starts. At each forecasting step, a new $V_{forecast}$ is generated and is combined with the former time step's value for updating V_{now}, then the updated V_{now} is inputted to the equation (4.20) and the new $V_{forecast}$ at next forecasting time is calculated. The rain cluster's characteristics which the proposed predicting method applied are: 'area', 'cumulative rainfall', 'max rain rate per 5 minute' and 'mean rain rate per 5 minute'.

4.8 Spatial Rainfall Verification Methods

The grid-point related error measuring is a problematic for the precipitation estimation and forecast. A classic example illustrating the flaw of grid-point based verification scores is the well-known 'Double Penalty' problem, which can be described as: prediction of a precipitation object with a correct size and structure might yield very poor verification scores. For example, a precipitation object is displaced slightly in space but the categorical verification scores penalize such a situation heavily. From a point of view of traditional verification methods, a displacement simply leads to a false alarm and it is also very poorly rated by a large root mean squared error (RMSE). Through these points of view, a forecast with a misplaced precipitation object is just as bad as a forecast that totally misses an event. Davis et al. (2006) illustrated some of the difficulties associated with diagnosing forecast errors using traditional verification approaches as in Figure 4.12.

Figure 4.12a illustrated a good forecast which the forecasted precipitation object (F) and observed one (O) were well matched; Figure 4.12b was a poor forecast which the forecasted and observed precipitation objects had large location deviations; in Figure 4.12c, the forecasted precipitation object had larger area and was offset to the observed one; in Figure 4.12d, the forecasted precipitation object was deviated to the observed one both in location and shape. Among the four examples illustrated in Figure 4.12, it appeared that case a was the best forecast.

Given the perceived differences in the performance, it was dismaying to note that all of the first four examples have identical basic verification statistics: Probability of Detection (POD), False Alarm Rate (FAR), Critical Success Index (CSI), and POD = 0, FAR = 1, CSI = 0. Thus, the verification technique is insensitive to the difference in location and shape errors. Similar insensitivity could be shown to be associated with timing errors. Figure 4.12e illustrated a very poor forecast from a variety of points of view with some skill (POD, CSI > 0; FAR < 1), suggesting it is a better forecast than the one depicted in Figure 4.12a.

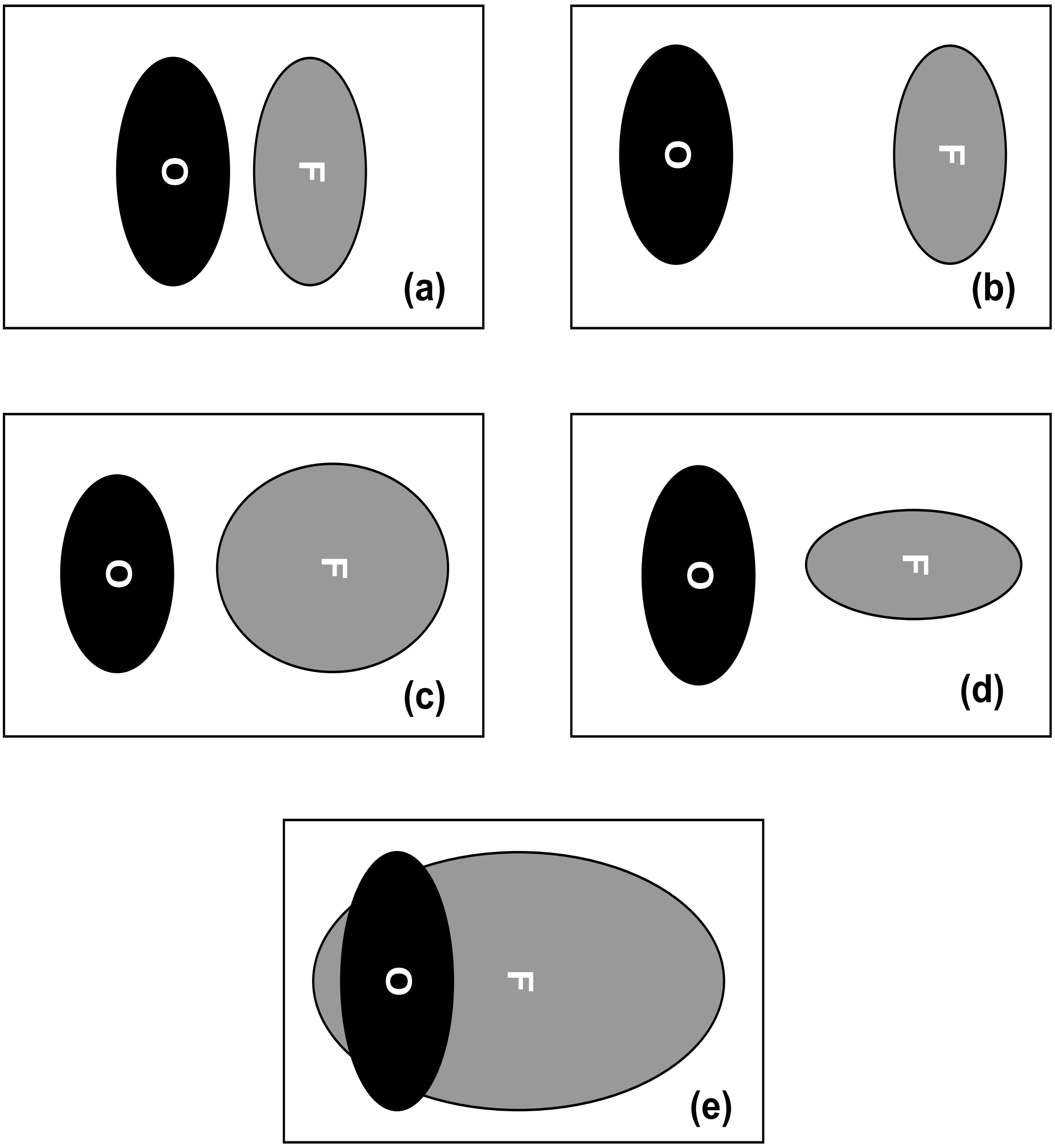


Figure 4.12 A schematic example of various forecasted and observed precipitation objects combinations.

However, a very fact that an event is somehow predicted should be positively evaluated because from a physical point of view, it is clearly a better forecast may totally miss an event. This issue becomes increasingly clear as the forecasting model's resolution increases. These high-resolution forecasting models can produce precipitation objects that are comparable to the radar information at various resolutions. Thus, complexity and a variety of structures generated by such models must be objectively scrutinized. High-resolution precipitation forecasts from numerical models may look quite realistic visually, and it may provide useful guidance for modelers and forecasters. However, usefulness of such high-resolute forecasts must be objectively quantified. Unfortunately, traditional verification scores are not much adapted for this purpose as already emphasized.

Despite substantial progress in observing precipitation patterns, characterization of their spatial patterns is always a challenge. While a great deal of effort has been put to statistical validation of grid based precipitation estimating results, uncertainties associated with geometry patterns of forecasted precipitation objects are not well researched.

Basic principles of spatial rainfall verification methods consist of relaxing an exact match to the observation at the fine scales. The spatially-based techniques stress the usefulness of the forecasts in analogy with the visual verification ('verification by eye') and focus on the verification of gridded forecasted data with the observations that is on the same grid as for radar based QPE. Such gridded observations have greater uncertainty at the higher resolutions than at the lower resolutions. Point-wise precipitation measurements, which are another source of verification, may suffer from an issue of representatives, particularly for precipitation objects existing highly variability. As a result, an estimated forecasting error is further biased by an observational uncertainty. In other words, a formally-obtained forecast error should be considered as practically meaningful. Spatially-based methods are built upon an idea of identifying weather events as 'objects' or 'features'. Under this perspective, the forecasted and observed rainfall values are not compared directly at the same locations (identical or near-by grids) but instead, the objects of interests are extracted from forecasted-observed data and then compared together so that verification statistics are obtained. A number of spatial verification methods have been proposed. Comparative review studies can be found in Ebert (2008) and Gilleland et al. (2009).

4.8.1 SAL Spatial Rainfall Verification Method

The SAL (Structure-Amplitude-Location) method is an object-based spatial rainfall verification approach for judging whether a QPF is good or not (Wernli et al., 2008). It aims at to address the following issues :

- It measures quantitatively three distinct aspects of the quality of a precipitation forecast in a previously specified area, integrated over time periods ranging from 1 to 24 hours.
- It takes into account the structure of the precipitation event (e.g., scattered convective cells, convective complex, frontal rain system), which is regarded as a direct fingerprint

of the physical nature of the event.

- It does not require a one-to-one matching between the identified objects and simulated ones in the forecast.
- It is close to a subjective visual judgment of the accuracy of a regional QPF.

The SAL methods is a three-component's feature based forecasting quality measurements, in which a forecast is treated in terms of Structure (S), Amplitude (A) and Location (L).The calculation of the three components of the SAL is outlined briefly in the following:

- **A** component corresponds to the normalized difference of the domain-averaged precipitation values of the modeled/forecasted objects - R_{mod} and the observed ones - R_{obs}:

  $$\mathbf{A} = \frac{\mathbf{D(Rmod) - D(Robs)}}{\mathbf{0.5\,[D(Rmod) + D(Robs)]}} \tag{4.21}$$

  where:

  $$\mathbf{D(R)} = \frac{\mathbf{1}}{\mathbf{N}} \sum\nolimits_{\mathbf{i,j \in D}} \mathbf{Ri,j} \tag{4.22}$$

  Here, D(**R**) denotes the domain average of the rain cluster R. The values of A are ranged from -2 to 2 and 0 denotes a perfect forecast in terms of amplitudes.

- **L** component of the SAL is consisted by two parts, L = L_1 + L_2 whereas the first term measures the normalized distance between the centers of mass of the modeled/forecasted and observed objects:

  $$\mathbf{L1} = \frac{\mathbf{|x(Rmod) - x(Robs)|}}{\mathbf{d}} \tag{4.23}$$

  where d denotes the largest distance between two boundary points of the considered domain and x(R) denotes the location for object R in this domain. Values of L_1 = 0 indicates that the centers of mass of the predicted and observed rain clusters are identical. The second part L_2 measures the averaged distance between the center of mass of the total objects and the individual. The L_2 can only differ from zero if either the observations or the forecast (or both) contain more than one object in the considered domain. The scaling of the L_2 is such that it has the same range as L_1, and hence the total location component L values vary between 0 and 2.

- **S** component is to compare the normalized volume for modeled/forecasted and observed objects. Such a measure captures information about the size and shape of objects. For each object R_n a scaled volume V_n is calculated as the following sum over all grid point values R(i, j):

$$\mathbf{Vn} = \sum\nolimits_{\mathbf{(i,j)\in Vn}} \frac{\mathbf{R(i,j)}}{\mathbf{Rn^{max}}} \tag{4.24}$$

where R_n^{max} denotes the maximum rain rate within the object. The volume V_n is calculated separately for all objects in the observational and forecast data sets. Then, the weighted mean of all objects' scaled rainfall volume, referred to as V, is determined for both data sets. Then, the S is defined as the normalized difference in V, analogously to the A component:

$$\mathbf{S} = \frac{\mathbf{V(Rmod) - V(Robs)}}{\mathbf{0.5\,[V(Rmod) + V(Robs)]}} \tag{4.25}$$

The scaling is again such that the possible range of values extends from -2 to 2. Positive values of S indicate that the predicted precipitation objects are too large and/or too flat; in contrast, negative values occur for too small and (or) too peaked objects.

4.8.2 Geometric Index Method

Geometric Index is a geometric framework for assessing spatial patterns of rain cluster, which is proposed by Aghakouchak et al. (2011). It constructed by some geometric concepts (e.g. minimum theoretical perimeter, convex hull and actual perimeter), three criteria indexes are defined for this method:

- **Connectivity index**: it is calculated based on the number of rain clusters (NC) and the total number of non zeros pixels or pixels above a given threshold (NP).

  $$\mathbf{C_{index}} = \mathbf{1} - \frac{\mathbf{NC - 1}}{\mathbf{\sqrt{NP} + NC}} \tag{4.26}$$

  where C_{index} is the connectivity index, NP is the number of (nonzero) pixels, and NC is the number of rain clusters. The Connectivity index is defined to compare multiple fields with respect to a reference image (e.g. simulated versus observed) and is not designed to compare patterns with significantly different spatial scales and/or number of pixels.

- **Shape index**:

  $$\mathbf{S_{index}} = \frac{\mathbf{P_{min}}}{\mathbf{P}} \tag{4.27}$$

  Shape index is used to distinguish rain clusters with different shapes. In this equation,

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

- **Area index**:

$$\mathbf{A_{index}} = \frac{\mathbf{A}}{\mathbf{A_{Convex}}} \tag{4.28}$$

Area index is defined based on the concept of convex hull (the boundary of the minimal convex set containing a finite set of points in space). It is the ratio of the area for rain cluster and area of its convex hull. In this equation, A is area of the pattern, and A_{convex} is the area of the convex hull.

CHAPTER 5

RESULTS

In this chapter, three day's radar images (26-May-2007, 19-Jul-2008 and 26-Jul-2008) were acquired for developing RCIT algorithm which was proposed in Chapter 4. Totally 9991 light rain clusters (reflectivity > 19 dBZ) and 3988 convective ones (reflectivity > 37 dBZ) were generated. Totally 1153 trajectories for light rain clusters and 508 ones for convective rain types were also extracted. Statistical analysis was applied for investigating spatial-temporal characteristics of identified rain clusters, and forecasting methods mentioned in Chapter 4 were also applied. Results were given in the following sections: In section 5.1, statistical descriptions for rain cluster's characteristics were presented. In section 5.2, mathematical functions for describing relations of rain cluster's characteristics were given. Section 5.3 presented the spatio-temporal analyzing results for identified rain clusters. Section 5.4 presented temporal developments of identified rain cluster's characteristics and 'Normalized Duration Lines' for rain cluster's characteristics were measured by regressive fitting method. In Section 5.5, verifying results of RCIT algorithm were presented by SAL and Geometry Index methods mentioned in Chapter 4. In Section 5.6, radar based precipitation forecasts for a leading time up to 2 hours were operated. Four forecasting methods based on the advection field tracking approach were implemented: the first one was the 'PIV_Semi-Lagrangian' wherein the motion vectors of advection field at previous time steps were generated by Particle Image Velocimetry method (see section 4.4.1 in Chapter 4), the future advection field was extrapolated by the Semi-Lagrangian scheme. The second method was similar to the first one but the advection field at the next time steps was extrapolated by the Lagrangian-Persistence scheme which was named as 'PIV_Lagrangian-Persistence'. The third and forth methods were based on TREC approach by which the motion vectors of advection field at previous time steps were generated by maximum cross-correlation way and the future advection field was extrapolated by Semi-Lagrangian and Lagrangian-Persistence schemes separately which were named as 'TREC_Semi-Lagrangian' and 'TREC_Lagrangian-Persistence', respectively. Additionally, predictions for the characteristics of rain clusters with a leading time up to one hour were also done by the methods proposed in Chapter 4. Verification results of forecasting and predicting methods were also presented in this section.

5.1 Descriptive Statistics of Rain Cluster's Characteristics

Descriptive statistics for rain cluster's characteristics included their minimum and maximum values, median value and the standard deviation as presented in Table 5.1. Empirical distributions of identified rain cluster's characteristics were presented in Figure 5.1 - 5.5 and some theoretical distributions were selected to fit their empirical distributions. Here, a multi-goodness of fit testing approach was proposed for the best theoretical distribution's searching, detailed description about this approach and its application was presented in Appendix B.

Table 5.1 Descriptive statistics of identified rain cluster's characteristics.

Characteristics	**min**		**max**		**Standard deviation**		**Median value**	
	>19dBZ	>37dBZ	>19dBZ	>37dBZ	>19dBZ	>37dBZ	>19dBZ	>37dBZ
Area [km^2]	10	10	18376	2855	1421.6	196.03	40	28
Cumulative rainfall [mm]	0.4	7.34	8851	4266.8	571.1	388.29	4.91	43.35
Max rain rate per 5 minutes [mm]	0.04	0.73	33.16	33.16	2.87	3.73	0.24	2.91
Mean rain rate per 5 minutes [mm/km^2]	0.04	0.73	6.19	8.39	0.31	0.88	0.1	1.5
Eccentricity	0	0	0.99	0.995	0.14	0.14	0.84	0.85

A high standard deviation existed for characteristic area, not only for identified light rain clusters but also for those convective ones. The value range of light rain cluster's area was between 10 and 18376 km^2, but most values were less than 40 km^2 (last column in Table 5.1). It was the same with area of convective rain clusters with a value range between 10 and 2855 km^2 and most of them were less than 28 km^2. A Generalized Pareto distribution (GPD) was found to fit the empirical distribution of log_{10} transformed area for both light and convective rain clusters (red lines in Figure 5.1). Equation (5.1) showed the density function of Generalized Pareto distribution:

$$\mathbf{f}(\mathbf{x}|\mathbf{k},\alpha,\theta) = (\frac{\mathbf{1}}{\alpha})(\mathbf{1}+\mathbf{k}\frac{(\mathbf{x}-\theta)}{\alpha})^{-\mathbf{1}-\frac{\mathbf{1}}{\mathbf{k}}} \tag{5.1}$$

where k is the shape parameter, θ and α are the location and scale parameters, respectively. For $\theta < 0$, $k > 0$ and for $\theta < x < \frac{\alpha}{k}$, $k < 0$. In the limit for $k = 0$, the GPD is an Exponential distribution. The GPD function assumes that all of the samples exceed a certain threshold. In this study, the threshold value (θ) for rain cluster's area was set to 10 km^2 due to the fact that the minimum value of 9 km^2 had been imposed in rain cluster identification procedure.

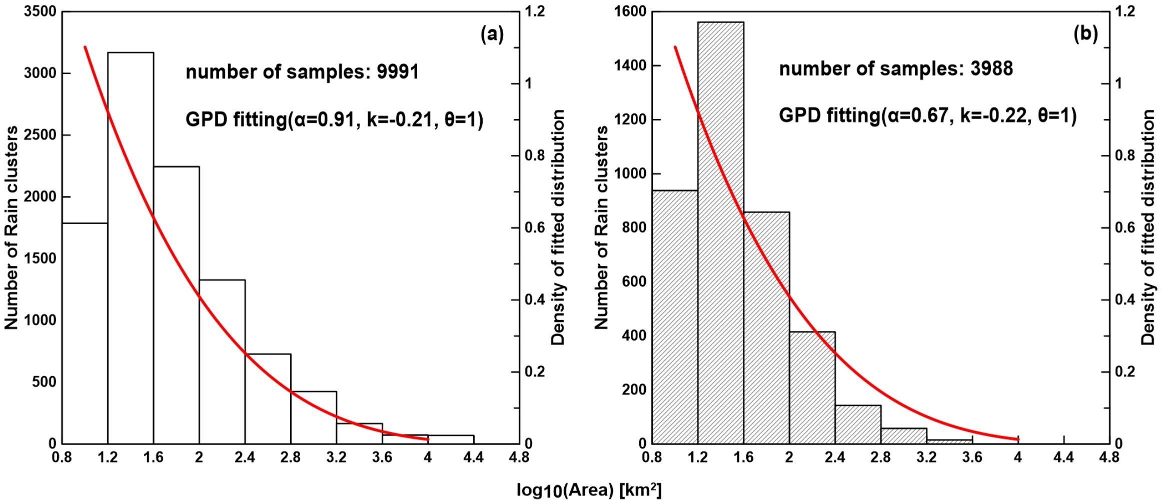


Figure 5.1 Empirical distribution of log_{10} transformed area fitted with its best matched theoretical distributions which were marked by red line. (a) for identified light rain clusters, (b) identified convective rain clusters.

For characteristic cumulative rainfall, high standard deviation also existed through the statistical descriptions (the forth column in Table 5.1). Its value range for identified light/convective rain clusters was obviously (0.4 to 8851 mm for light rain clusters and 7.34 to 4266.8 mm for convective rain clusters). Two theoretical distributions were found to fit the empirical distributions of their log_{10} transformed cumulative rainfall, respectively (Figure 5.2). For light rain clusters, distribution of their log_{10} transformed cumulative rainfall could be fitted with the Generalized Pareto distribution and for the convective ones, the Inverse Gaussian distribution (INV_Gauss) was the best theoretical distribution for fitting their log_{10} transformed cumulative rainfall.

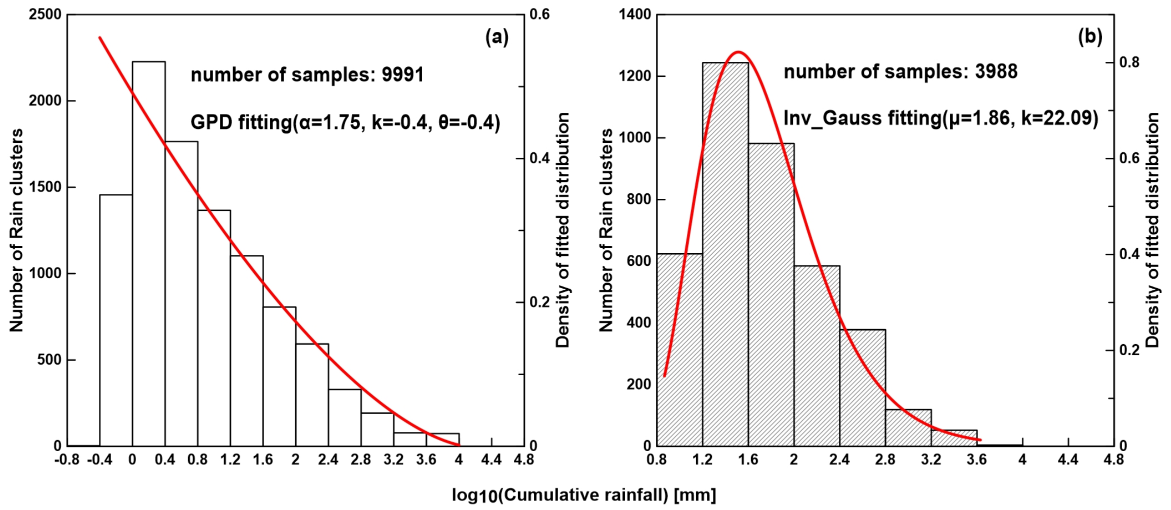


Figure 5.2 Same with Figure 5.1, but for log_{10} transformed cumulative rainfall.

Equation (5.2) presented the density function of Inverse Gaussian distribution. The higher median value of convective rain cluster's cumulative rainfall (43.35 mm) reflected that the identified convective rain clusters contained more rainy pixels with higher intensity compared to the light ones.

$$\mathbf{f}(\mathbf{x}|\mu,\lambda) = \left[\frac{\lambda}{2\pi\mathbf{x}^3}\right]^{\frac{1}{2}} \mathbf{e}^{\frac{-\lambda(\mathbf{x}-\mu)^2}{2\mu^2\mathbf{x}}} \tag{5.2}$$

for $x > 0$, $\mu > 0$ is the mean parameter and $\lambda > 0$ is the shape parameter. As λ tends to infinity, the Inverse Gaussian distribution is more like a normal distribution. In such case, the distribution for value of log_{10} transformed cumulative rainfall of convective rain clusters will present a symmetry trend.

For characteristic max rain rate per 5 minutes, its standard deviation was properly (2.87 mm for light rain clusters and 3.73 mm for convective ones) but their value range were also obviously (0.04 to 33.16 mm for light rain clusters and 0.73 to 33.16 mm for convective ones). Most rain cluster's max rain rate were small through their median value (0.24 mm for light rain clusters and 2.91 mm for convective ones). The Generalized Pareto distribution was selected as the best theoretical distribution for fitting empirical distribution of log_{10} transformed max rain rate of light rain clusters and the Generalized Extreme Value distribution (GEV) was the best for fitting the empirical distribution of log_{10} transformed convective rain cluster's max rain rate (Figure 5.3).

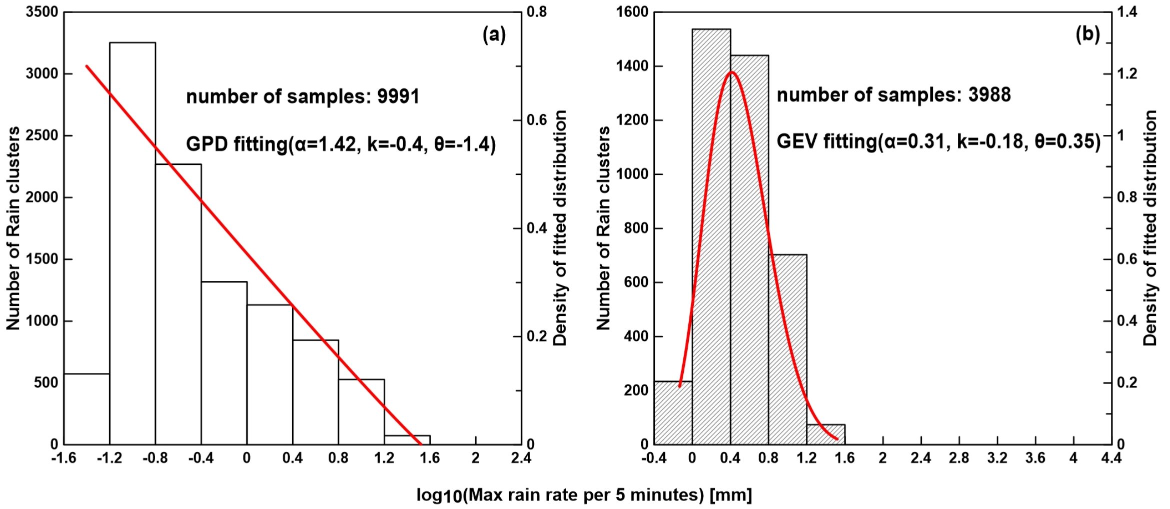


Figure 5.3 Same with Figure 5.1, but for log_{10} transformed max rain rate per 5 minutes.

Equation (5.3) presented the density function of GEV distribution:

$$\mathbf{f}(\mathbf{x}|\mathbf{k},\mu,\alpha) = \left(\frac{\mathbf{1}}{\alpha}\right)\mathbf{e}^{\left(-\left(\mathbf{1}+\mathbf{k}\frac{(\mathbf{x}-\mu)}{\alpha}\right)^{-\frac{\mathbf{1}}{\mathbf{k}}}\right)}\left(\mathbf{1}+\mathbf{k}\frac{(\mathbf{x}-\mu)}{\alpha}\right)^{-\mathbf{1}-\frac{\mathbf{1}}{\mathbf{k}}} \tag{5.3}$$

for $1+k\frac{(x-\mu)}{\alpha}$, $k>0$, the GEV distribution is the Fréchet distribution, while $k<0$, it corresponds to Weibull distribution. In the limit for $k=0$, the GEV distribution is Gumbel distribution, in that case, the density function is:

$$\mathbf{f}(\mathbf{x}|\mathbf{0},\mu,\alpha) = \left(\frac{\mathbf{1}}{\alpha}\right)\mathbf{e}^{\left(-\mathbf{e}^{\left(-\frac{(\mathbf{x}-\mu)}{\alpha}\right)}-\frac{(\mathbf{x}-\mu)}{\alpha}\right)} \tag{5.4}$$

The value range of characteristic mean rain rate per 5 minutes was more similar to characteristic max rain rate with 0.04 to 6.14 mm/km^2 and 0.73 to 8.3 mm/km^2 for light and convective rain clusters, respectively. Its High median value for identified convective rain clusters (1.5 mm/km^2) reflected that the areal averaged precipitation of convective precipitation objects obviously exceeded the light ones though their area were small than the light ones. Same theoretical distributions were found for fitting the empirical distributions of log_{10} transformed light/convective rain cluster's mean rain rate compared to the characteristic max rain rate (Figure 5.4).

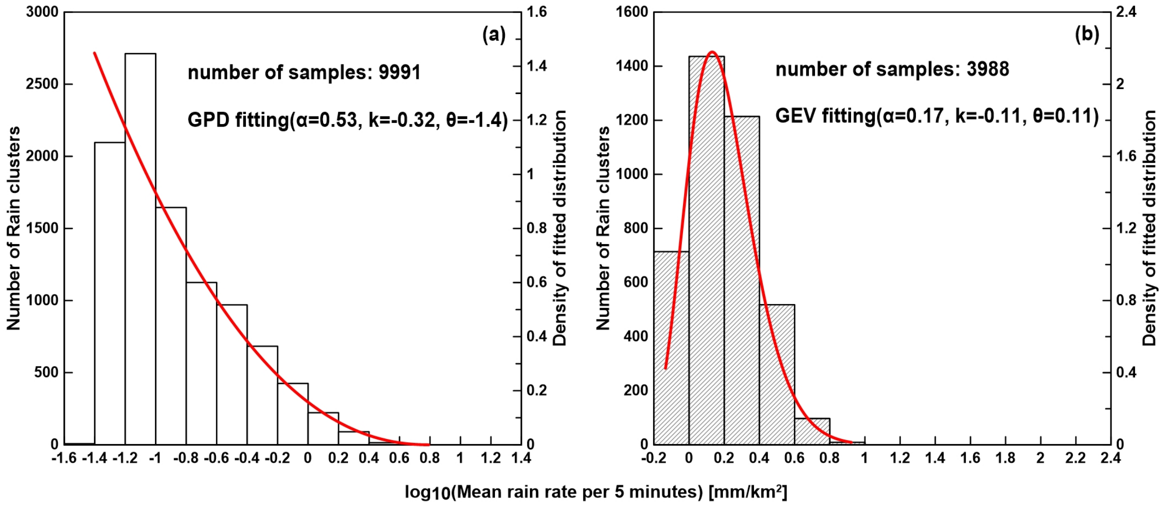


Figure 5.4 Same with Figure 5.1, but for log_{10} transformed mean rain rate per 5 minutes.

Value of characteristic eccentricity for identified light/convective rain clusters ranged from 0 to 1, and most of them were over 0.5 (median value was 0.84 for light rain clusters and 0.85 for convective ones) which reflected that the identified rain clusters, not only light ones but also the convective ones were more fitted to an ellipse shape. The Generalized Extreme Value distribution was found for fitting empirical distributions of their eccentricity (Figure 5.5).

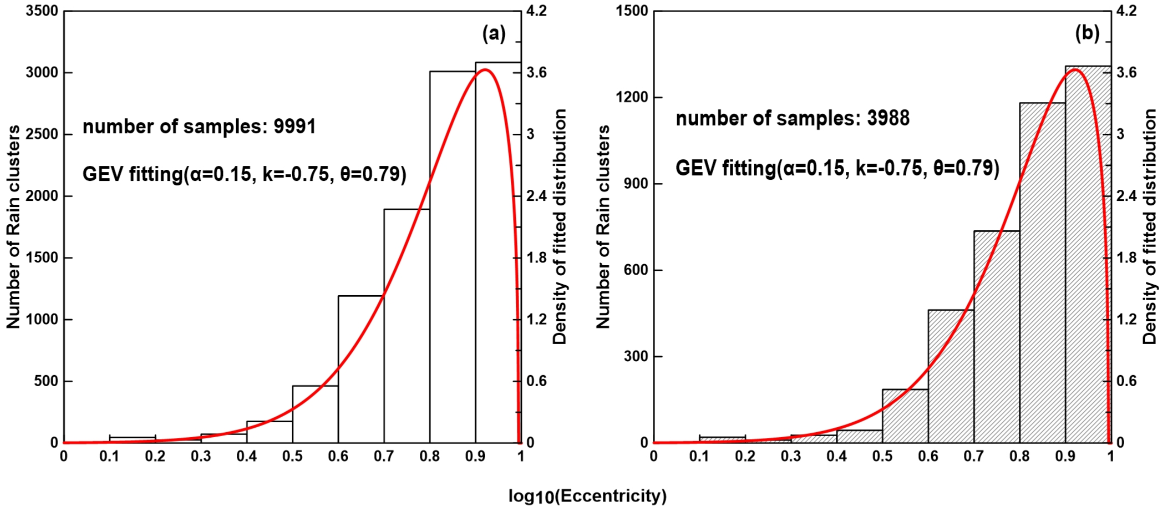


Figure 5.5 Same with Figure 5.1, but for characteristic eccentricity.

5.2 Mathematical Relations of Rain Cluster's Characteristics

The aim of the mathematical description for relations of rain cluster's characteristics is to quantify the inner structures of rainfall events by some simple mathematic functions. It is not an easy task. According to Waymire and Gupta (1981a,b), there is still no uniform mathematic way for describing their inner structure's relations because of different precipitating scenes and modeling tools based on these scenes. In this study, mathematic relations for describing the relationships of identified rain cluster's characteristics were investigated. Considering the aim of this section, these relationships included:

(a) 'Area - Cumulative rainfall', reflects the variability of integral rainfall volume following the variability of size for the identified rain clusters.

(b) 'Area - Number of convective pixels', reflects the variability of size following the internal rain rate's growth/decay for the identified rain clusters.

(c) 'Cumulative rainfall - Number of convective pixels', reflects the variability of integral rainfall volume following the internal rain rate's growth/decay for the identified rain clusters.

(d) 'Max rain rate per 5 minutes - Number of convective pixels', reflects the variability of peak rain rate following the internal rain rate's growth/decay for the identified rain clusters.

(e) 'Mean rain rate per 5 minutes - Number of convective pixels', reflects the variability of areal averaged precipitation following the internal rain rate's growth/decay for the identified rain clusters.

Figure 5.6-5.10 presented the scatter plots of log_{10} transformed characteristics of rain clusters and their fitting results (red curves in the figures).

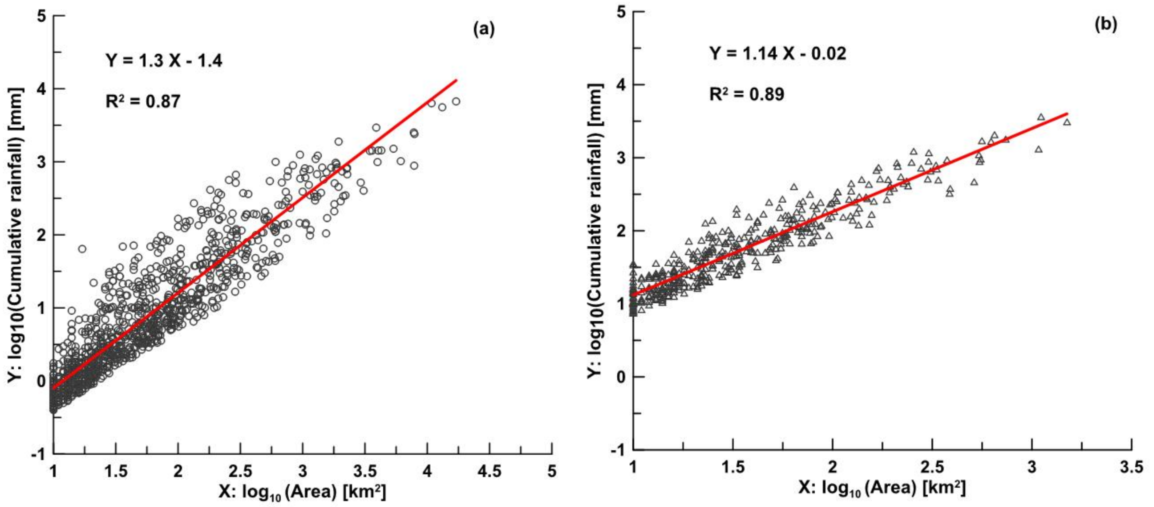


Figure 5.6 Scatter plots for log_{10} transformed area and cumulative rainfall and their linear fitting curve. (a) for identified light rain clusters, (b) for identified convective rain clusters

Through the Figure 5.6, the relation for log_{10} transformed area and cumulative rainfall presented a strongly linear trend, not only for the identified light rain clusters (the goodness fitting test results was 0.87), but also for the identified convective ones (the goodness of fit testing result was 0.89) which indicated that the cumulative rainfall of identified rain clusters presented a exponential growth, when the size of rain clusters touched to some threshold, the growth of their cumulative rainfall slowed down.

Number of convective pixels, one of the rain cluster's characteristic which is defined for reflecting the internal rain rate's growth/decay for identified rain clusters, was applied for investigating its relations to other physical and geometry characteristics of identified rain clusters. By doing this investigation, such question will be answered: Are the variability of value for rain cluster's physical and geometry characteristics from its internal intensity's growth/decay directly?

Figure 5.7-5.10 presented the relations between the number of convective pixels for light/convective rain clusters and log_{10} transformed characteristics: area, cumulative rainfall, max and mean rain rate per 5 minutes. Through their fitted curves, these relations followed an power-law function as in equation (5.5):

$$\log_{\mathbf{10}}(\mathbf{Characteristics}) = \mathbf{a}\,(\mathbf{N_p})^{\mathbf{b}} \tag{5.5}$$

where N_p represented the number of convective pixels for identified rain clusters, a and b are the coefficient parameters of power-law function.

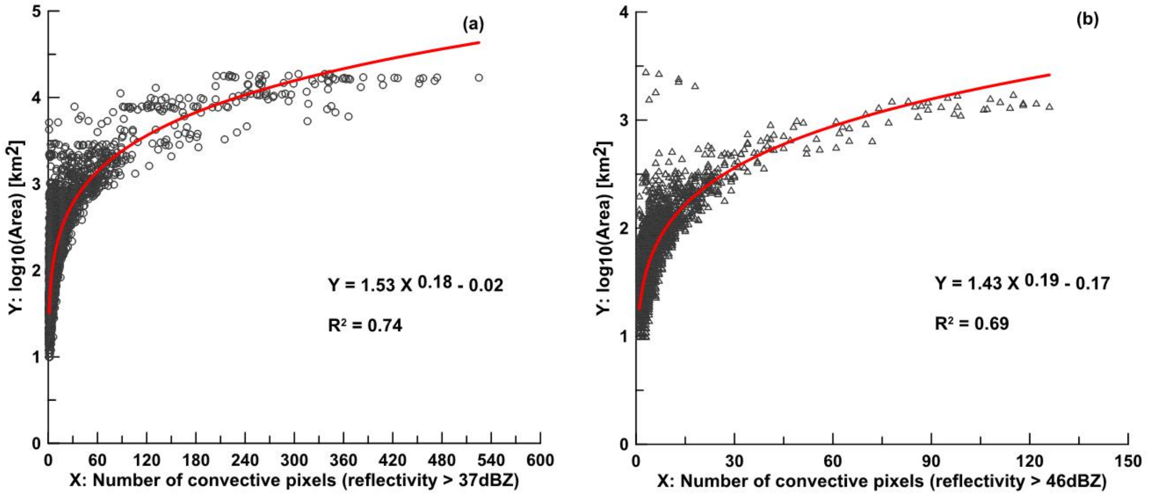


Figure 5.7 Scatter plots for log_{10} transformed area and number of convective pixels with their power-law fitting results. (a) for identified light rain clusters, (b) for identified convective rain clusters

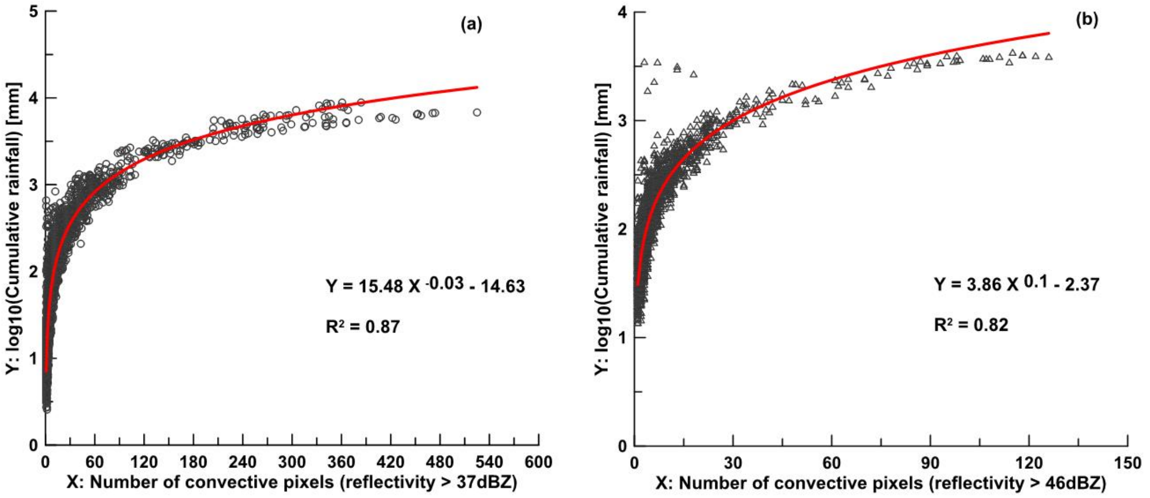


Figure 5.8 Same with Figure 5.7, but for log_{10} transformed cumulative rainfall.

Coefficient of determination for their regressive results reflected that: variability of identified rain cluster’s area and integral rainfall volume had robust connections with its internal rain rate’s growth/decay (goodness of fit testing was 0.74/0.69 for log_{10} transformed area and 0.87/0.82 for log_{10} cumulative rainfall). The goodness of fit testing results for relations between number of convective pixels and log_{10} transformed max and mean rain rate per 5 minutes were weak (0.52/0.42 for log_{10} transformed max rain rate per 5 minutes, 0.18/0.2 log_{10} transformed mean rain rate per 5 minutes). The question was answered: for identified light rain clusters, their size and integral rainfall volume's variability were strongly affected by their internal intensity's growth/decay, but whether the strong connections can be found for max/mean rain rates was blur.

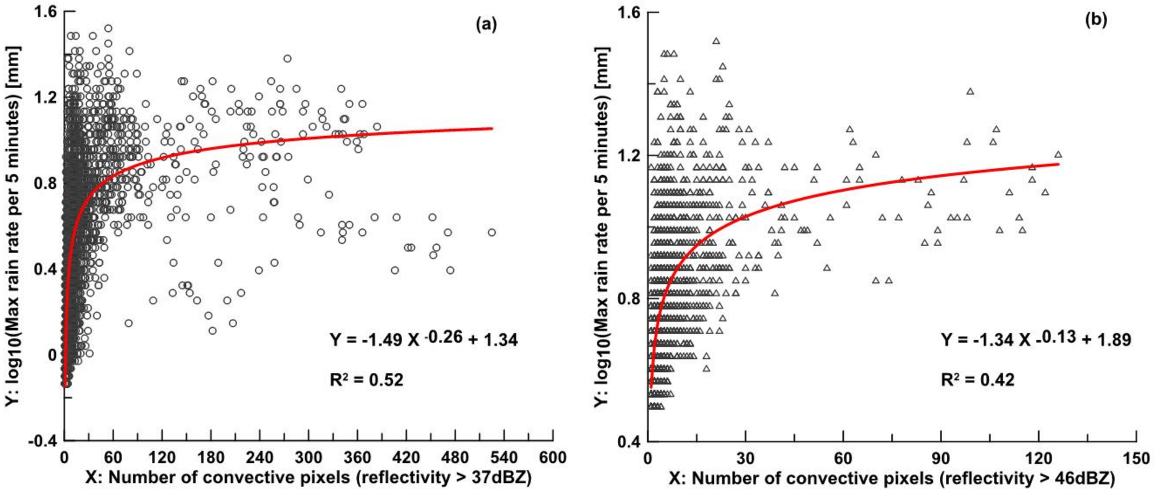


Figure 5.9 Same with Figure 5.7, but for log_{10} transformed max rain rate per 5 minutes.

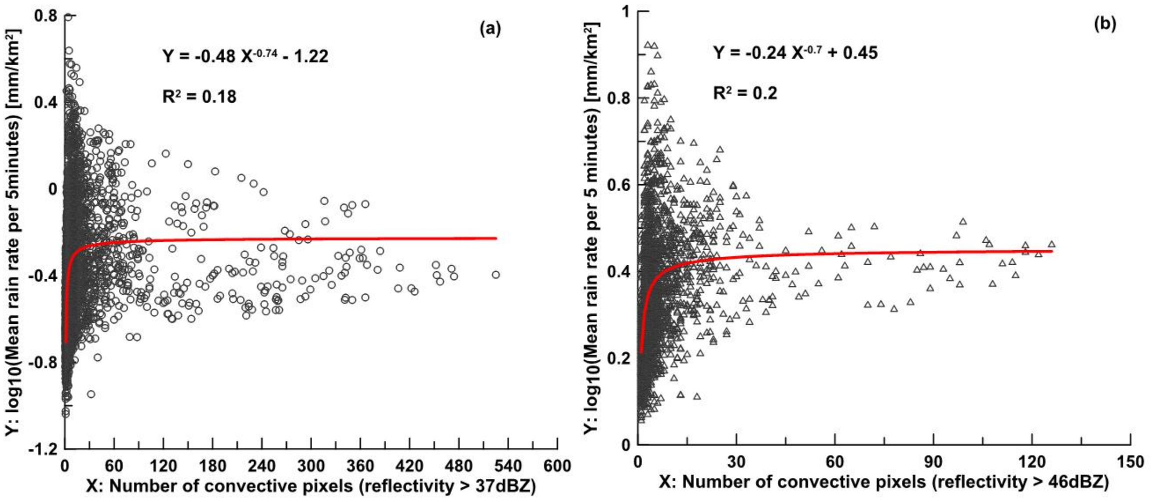


Figure 5.10 Same with Figure 5.7, but for log_{10} transformed mean rain rate per 5 minutes.

5.3 Spatial and Temporal Correlation Analyzing Results

The spatial and temporal correlation function for identified rain clusters was borrowed from Krajewski et al. (2000) as in equation (5.6) , and it was used for defining the precipitating process at the horizontal plane. In detail, spatial correlation is the cross-correlation of rain rate at two grids separated by a given spatial distance within the size of rain clusters (e.g., 1 km), which is aimed at studying the variability of rain rate as a function of space and its curve describes the decay of correlation index of rain rates following the different spatial scales and can be used for characterizing the spatial variability of precipitation objects. Correspondingly, the temporal correlation is the cross-correlation of rain rates at the same grid which is separated by a given time interval within the life cycle of identified rain clusters and curves of its correlation index describe the decay of rain rate's correlation index follows different temporal scales which is used to characterize the temporal variability of precipitation objects.

$$\mathbf{Corr}(\mathbf{\Delta s}, \mathbf{\Delta t}) = \frac{\mathbf{Cov}((\mathbf{I_{s,t}} - \mu), (\mathbf{I_{s+\Delta s, t+\Delta t}} - \mu))}{\alpha^{\mathbf{2}}} \tag{5.6}$$

where s and t represent the location and moment of rainy pixel in the identified rain clusters. Δs and Δt are the spatial and temporal scales. Cov is the covariance value for rain cluster's pixels at specified spatial and temporal scale. α is the sample variance of rain cluster's pixels.

The spatial scale along the prevailing directions of trajectories calculated from RCIT algorithm was selected for spatial correlation's calculation. For example, for rainy day 26-Jul-2008, the prevailing direction of calculated trajectories was from south-east to north-east. It was identified at the fourth quadrant in the Cartesian coordinate system with a negative value at the horizontal direction and a positive value at the vertical direction. Value of the spatial scale should be like $[-x, +y]$.

More details about the selection of spatial scale was presented in Appendix D. In this study, the correlation index for identified light/convective rain clusters in three selected rainy days were calculated as in Figure 5.11 with spatial scales from 1 to 10 km and temporal scales from 5 to 100 min.

Through the Figure 5.11, the correlation coefficients of rain rates for identified rain clusters got high values at small spatial or temporal scales, and decreased as the scale extending. The correlating index of identified rain clusters which followed the spatial and temporal scales could be fitted with a exponential function by applying the regressive method in which the spatial correlations of identified rain clusters were more fitted through the goodness of fit testing results, not only for light rain clusters (Figure 5.11a and c), but also for those convective ones (Figure 5.11b and d).

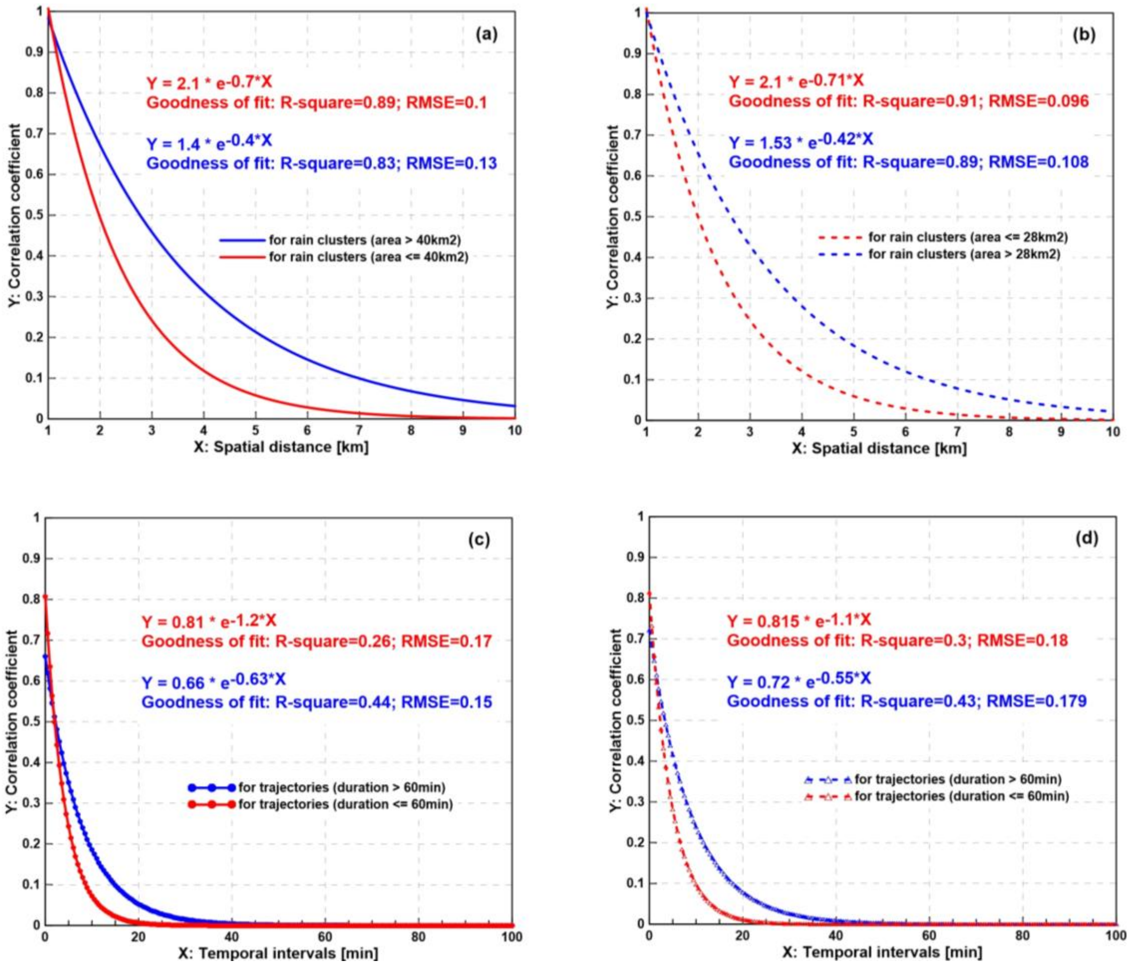


Figure 5.11 Spatial and temporal correlating curves for identified rain clusters, which is sorted by median values of their areas and duration of 60 min. (a) spatial correlation of identified light rain clusters, (b) same with (a), but for identified convective rain clusters, (c) temporal correlation of identified light rain clusters, (d) same with (c), but for identified convective rain clusters.

For identified light rain clusters with area above 40 km^2 (Figure 5.11a), their calculated spatial correlations got a higher value compared to those with their area less than 40 km^2, which indicated that the most identified light rain clusters with their area under 50^{th} percentile level were more smoothly than those with area above the same percentile level and this difference was more obviously when the spatial scale was less than 5 km. For the identified convective rain clusters with area above 28 km^2 (Figure 5.11b), their calculated spatial correlations also got higher values compared to those with area less than 28 km^2, but the correlation coefficients for the same spatial scale were smaller than those calculated for light rain clusters, which indicated that the most identified convective rain clusters with area under 50 $\%$ percentile level were smoother than those with area above at the same percentile level.

Identified convective rain clusters were more smoothly when they were compared to the identified light ones at the same spatial scale.

For identified light rain clusters with duration over 60 min (Figure 5.11c), their temporal correlation coefficients were higher than those with duration less than 60 min and the difference of their correlation coefficients was obviously within a temporal scale of 20 min, which indicated that the variability of identified light rain clusters with a shorter duration were more obviously than those with longer duration. Structures of these rain clusters might be more stable within a short temporal scale of 20 min. It was almost the same with what were calculated for the identified convective rain clusters (Figure 5.11d) excepted that the correlation coefficients were slightly higher than the light ones at the same temporal scale, which means a more stable inner structures of identified convective rain clusters than the identified light ones.

5.4 Temporal Developments for Identified Rain Clusters

5.4.1 Trajectory Analyzing Results

Trajectories for identified rain clusters were calculated and extracted by RCIT algorithm, which included spatial and temporal information about characteristics of identified rain clusters. According to the life stage's definition for rain clusters (see section 4.4.3 in Chapter 4), the calculated trajectories can be categorized in two types:

(a) **Single-tracks**. There is no rain cluster which merges or splits, and the calculated motion vectors for rain cluster's trajectory are properly without much deviation from the mean value of calculated global motion vectors.

(b) **Multi-tracks**. Rain clusters merge or split frequently with others, and the calculated motion vectors of trajectory show the obvious value deviation compared to the mean value of calculated global motion vectors. This scene is more complicated compared to 'single-tracks' type.

Totally 1153 trajectories for identified light rain clusters including 1043 ones of 'Single-tracks' type, and totally 508 trajectories including 478 ones of 'Single-tracks' type for identified convective rain clusters were calculated and extracted. Figure 5.12 presented the histograms of two type's trajectories for identified light and convective rain clusters, respectively.

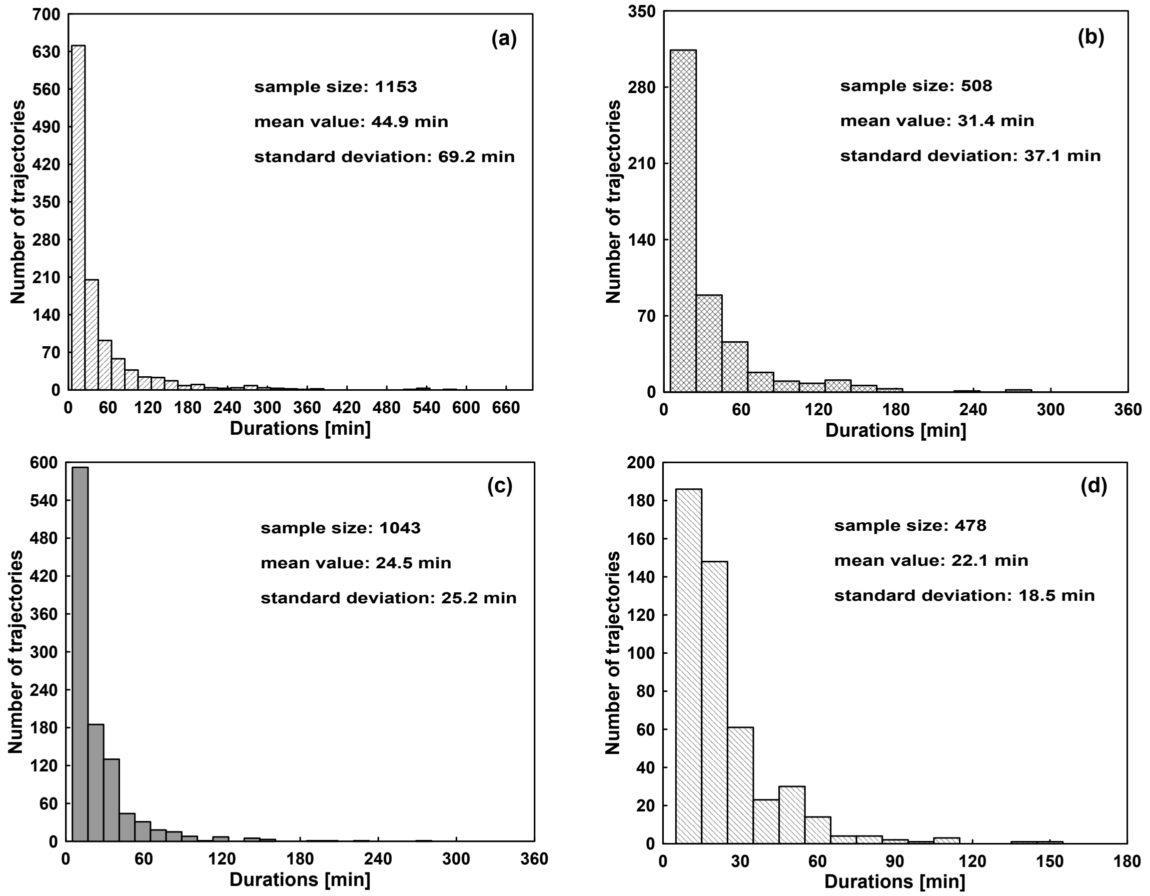


Figure 5.12 Histograms of calculated trajectory's durations. (a) all trajectories for identified light rain clusters, (b) same with (a), but for identified convective rain clusters, (c) 'Single-tracks' trajectories for identified light rain clusters, (d) same with (c), but for identified convective rain clusters.

Considering the durations of all trajectories for identified light rain clusters. Their mean value was 44.9 min with the maximum value of 655 min. Durations of 'Single-tracks' trajectories for identified light rain clusters were lower whose mean value was 24.5 min with the maximum value of 280 min. For identified convective rain clusters, mean value for their 'Single-tracks' trajectories was 22.1 min with a maximum value of 150 min.

Through the Figure 5.12, histograms of durations for calculated trajectories seemed to follow an exponential distribution which held a sharply decreasing trend, not only for the identified light rain clusters, but also for those convective ones. Most calculated trajectories hold a short duration, and the 'Single-tracks' ones were more. Trajectories with duration above 2 hours sparingly occurred.

Number of rain clusters with different stages over their life cycles was also summarized.

Table 5.2 presented these results, through the table, the identified rain clusters with 'track' stage were in dominant, 'merging' and 'splitting' ones occurred solely which indicated a stable inner structures for the identified rain clusters.

Table 5.2 Number of identified rain clusters with different stages, sorted by rain day.

stage	Light rain clusters (ref > 19dBZ)			Convective rain clusters (ref > 37dBZ)		
	26-May-2007	19-Jul-2008	26-Jul-2008	26-May-2007	19-Jul-2008	26-Jul-2008
Initial	166	462	471	81	133	277
Tracking	662	1889	1931	169	274	1019
Merging	14	59	52	1	2	30
Splitting	0	20	8	0	4	10
Dissipate	157	423	424	79	134	255
5 minute life cycle	481	2075	695	268	780	472

On rainy day 19-Jul-2008 and 26-Jul-2008, the number of identified rain clusters with '5 minute life cycle' was more than those from 26-May-2007, which means that the identified rain clusters from 26-May-2007 were more presented with the form of 'Single-tracks' type and persisted with long durations. Number of identified convective rain clusters for all stages was smaller than the light ones which means that the convective situation did not occur frequently in three rainy days. Particularly, for identified convective rain clusters, few splitting or merging scenes occurred compared to the identified light ones. The reasons might be that some convective rain clusters were only observed in one radar image, and most of them were embedded into the light ones which could split or merge easily.

Motion vectors of identified rain clusters over their life cycles were calculated based on the characteristics center of mass which was deduced by weighting the summarized coordinates of pixels in Cartesian coordinates at horizontal and vertical direction respectively and were presented as speed (unit: m/s) and direction (unit: degree) which were calculated according to the equation (5.7).

$$\begin{aligned}\mathbf{speed} &= \sqrt{(\mathbf{x_t}-\mathbf{x_{t+\Delta t}})^2+(\mathbf{y_t}-\mathbf{y_{t+\Delta t}})^2}\\ \mathbf{direction} &= \arctan(\frac{\mathbf{y_t}-\mathbf{y_{t+\Delta t}}}{\mathbf{x_t}-\mathbf{x_{t+\Delta t}}})\end{aligned} \tag{5.7}$$

where x_t and $x_{t+\Delta t}$ are the Cartesian coordinates of center of mass for identified rain clusters at previous and next time, respectively. y_t and $y_{t+\Delta t}$ are the Cartesian coordinates of center of mass for identified parent/child rain clusters at vertical direction. Figure 5.13 presented the empirical distributions of mean values of calculated motion vectors for identified rain clusters over their life cycles.

(a) For rainy day 27-May-2007 (the daily averaged motion vector was 2.11 m/s which was calculated from RCIT algorithm), the identified light rain clusters moved ahead the North-East direction with a mean speed of 11.1 m/s and the standard deviation of 4.79 m/s. The 'Single - tracks' ones hold a mean speed of 10.4 m/s with a standard deviation of 4.56 m/s. Identified convective rain clusters moved along the same direction but with a higher mean speed (12.8 m/s and standard deviation (7.46 m/s). The calculated 'Single - tracks' trajectories for identified convective rain clusters hold a mean speed of 12.4 m/s with the standard deviation of 6.88 m/s.

(b) For rainy day 19-Jul-2008 (the calculated daily averaged motion vector was 3.91 m/s), the identified light rain clusters moved ahead the East direction with a mean speed of 15.5 m/s and a standard deviation of 4.71 m/s. The calculated 'Single - tracks' trajectories hold a mean speed of 14.6 m/s with the standard deviation of 5.22 m/s. For identified convective rain clusters from this day, they had the same moving direction but with a lower mean speed (13.8 m/s) and a lower standard deviation (4.67 m/s), there was a mean speed was 13.7 m/s with the standard deviation of 4.88 m/s for their calculated 'Single - tracks' trajectories.

(c) For rainy day 26-Jul-2008 (the calculated daily averaged motion vector was 1.8 m/s), the identified light rain clusters moved ahead the North-West direction with a mean speed of 9.1 m/s and a standard deviation of 4.76 m/s. The calculated 'Single - tracks' trajectories hold a mean speed of 8.1 m/s with the standard deviation of 4.98 m/s. Identified convective rain clusters moved along South-East to North-West direction with a lower mean speed of 8.74 m/s and a lower standard deviation of 4.19 m/s. A mean speed of 7.74 m/s was deduced for their 'Single - tracks' trajectories with the standard

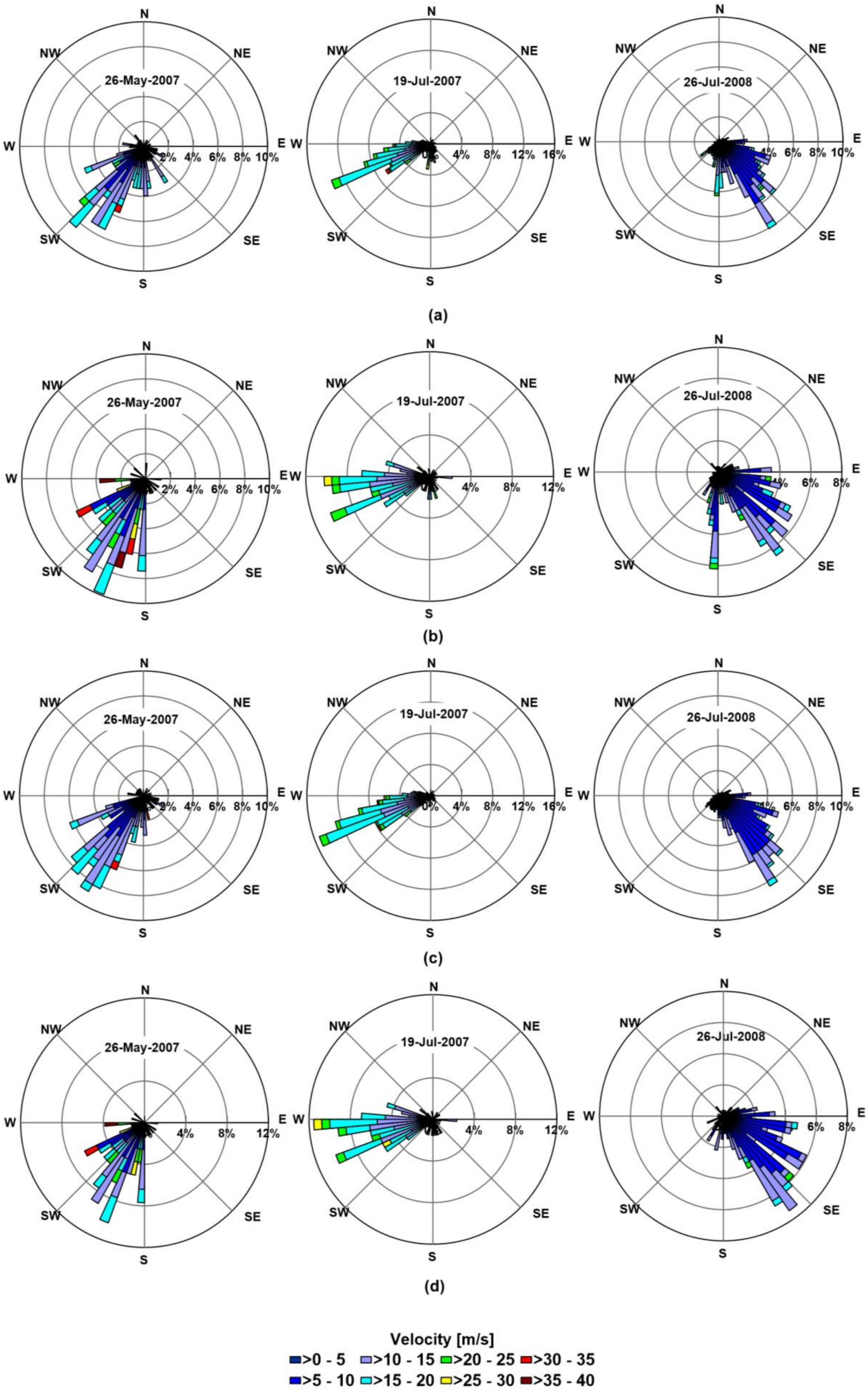


Figure 5.13 Empirical distributions of mean values of calculated motion vectors for identified rain clusters over their life cycles, sorted by three rainy day. (a) for all trajectories of identified light rain clusters, (b) same with (a), but for identified convective rain clusters, (c) for 'Single-tracks' trajectories of identified light rain clusters, (d) same with (c), but for identified convective rain clusters.

deviation of 3.47 m/s.

Results from the identified rain cluster's life cycle analysis indicated that the precipitation objects with low intensity hold longer durations compared to the identified convective ones, and their calculated trajectories without merging and splitting had shorter durations. The movements of precipitation objects in three rainy days also presented diversity: for rainy day 27-May- 2007 and 19-Jul-2008, mean value of motion vectors for identified light rain clusters was higher than the convective ones. Moving directions of identified rain clusters were along east which were conversely to those identified from rainy day of 26-Jul-2008. Moving directions of the identified rain clusters were same with what were observed from radar images.

5.4.2 Temporal Variability of Rain Cluster's Characteristics

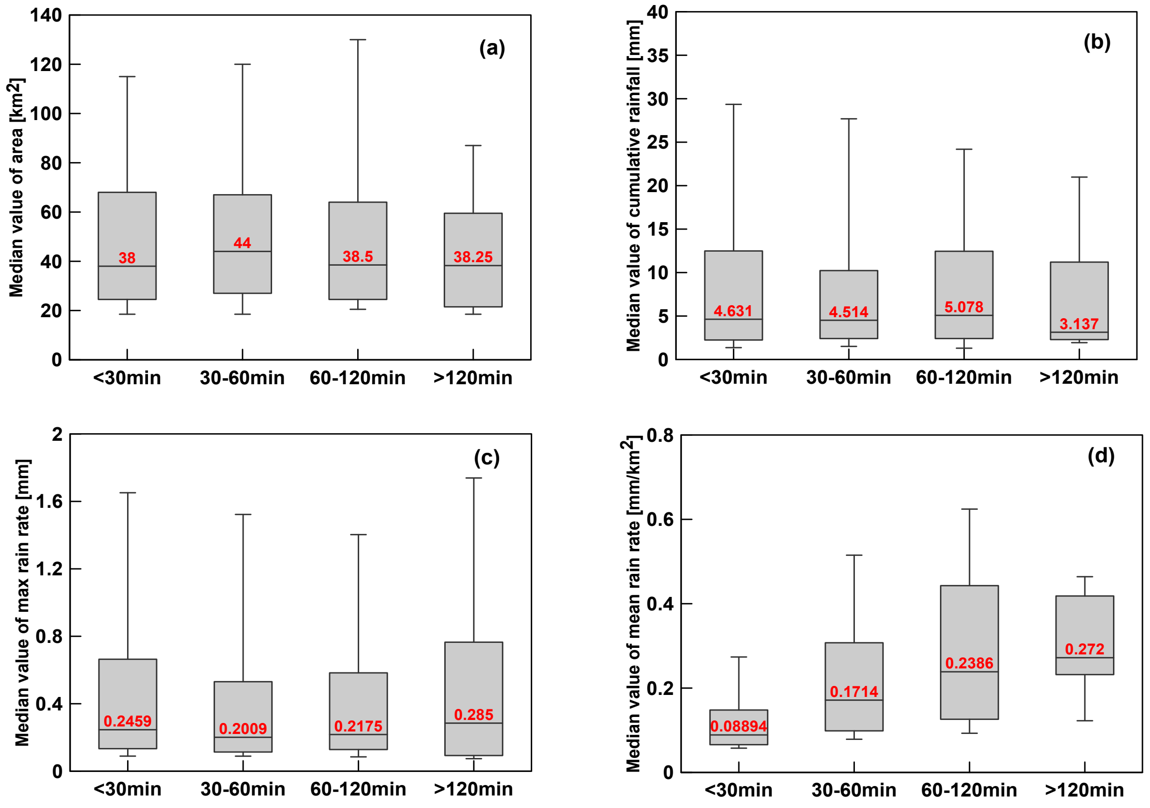


Figure 5.14 Box plots for median values of identified light rain cluster's characteristics, sorted by their durations. (a) for characteristic area, (b) for characteristic cumulative rainfall, (c) for characteristic max rain rate per 5 minutes, (d) for characteristic mean rain rate per 5 minutes.

Characteristics of identified rain clusters with same durations were sampled and grouped, then their median values were deduced. Figure 5.14 and Figure 5.15 presented the box plots of the median value for identified light/convective rain cluster's characteristics sorted by durations, respectively. Outliers (above 95% or under 5% quantile levels) in each class of box plot were removed.

Median value of characteristic area for identified light rain clusters with durations around 30 to 60 min was higher than others (Figure 5.14a), and decreased following the extended durations. Median value of cumulative rainfall presented a increasing trend for identified light rain clusters within a duration of 60 min and a maximum value was found around 60 tp 120 min (Figure 5.14b), then it decreased when identified rain cluster's duration over 120 min. Median value of characteristic max rain rate per five minute was higher for identified light rain clusters with durations less than 30 min or over 120 min and was lower for those with duration around 30 to 120 min (Figure 5.14c). An increasing trend was found for the median value of characteristic mean rain rate per 5 minutes (Figure 5.14d).

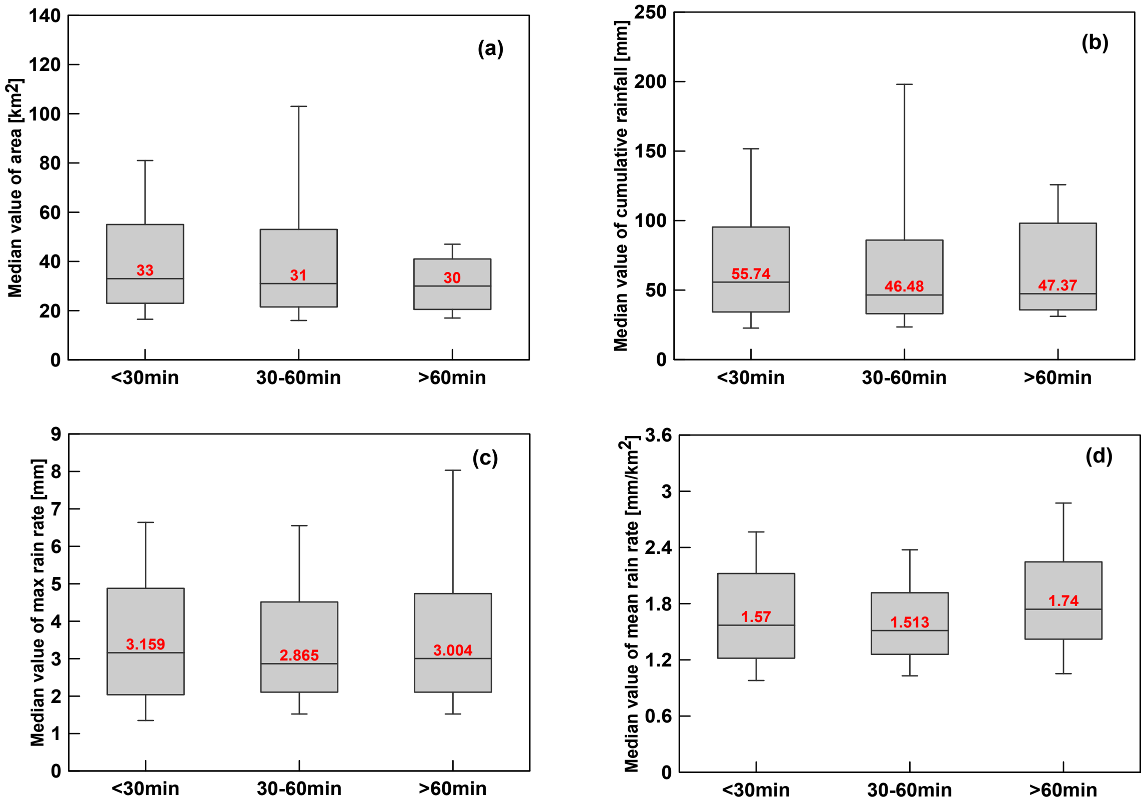


Figure 5.15 Same with Figure 5.14, but for characteristics of identified convective rain clusters.

For identified convective rain clusters, decreasing trend was found for median values of their characteristics area (Figure 5.15a), cumulative rainfall (Figure 5.15b) and max rain rate per 5 minutes (Figure 5.15c) when rain cluster's durations extended except that the median value of characteristic mean rate presented a slightly increasing trend.

5.4.3 Analyzing Results of Normalized Duration Lines

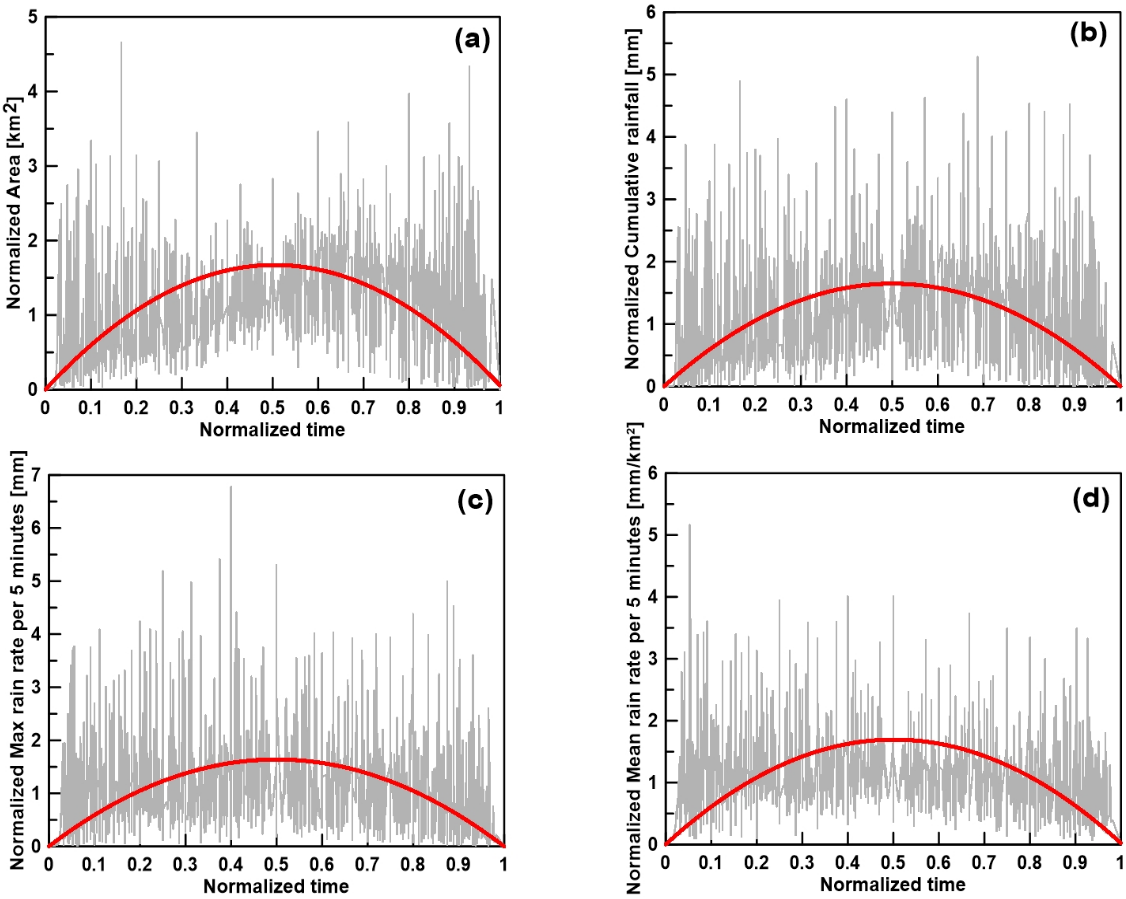


Figure 5.16 'Normalized Duration Lines' for characteristics of identified light rain clusters. (a) for characteristic area, (b) for characteristics cumulative rainfall, (c) for characteristic max rain rate per 5 minutes, (d) for characteristic mean rain rate per 5 minutes.

In this study, the 'Normalized Duration Lines' was calculated for reflecting the temporal variability of identified rain cluster's characteristics with different durations and was further applied for doing short term characteristic's prediction (see section 4.7 in Chapter 4). Based on the fitted parabola functions, prediction for the identified rain cluster's characteristics was possible. The calculating procedure for 'Normalized-Duration-Lines' were listed:

(a) Characteristics were derived from RCIT algorithm for rain clusters with 'Single - tracks' type, they were grouped and classified by the length of identified rain cluster's durations.

(b) For each trajectory in each group with the same duration, identified rain cluster's characteristics were calculated and normalized by their mean values. Meanwhile, the trajectory's duration belonging to this group was also normalized by the max duration length of this group.

(c) A robust regressive fitting method was applied and the parabola function was found for fitting characteristic's 'Normalized Duration Lines', and the coefficient parameters a and b for parabola function were determined.

More detail information about the method for calculating 'Normalized Duration Lines' of identified rain cluster's characteristics was presented in Appendix C.

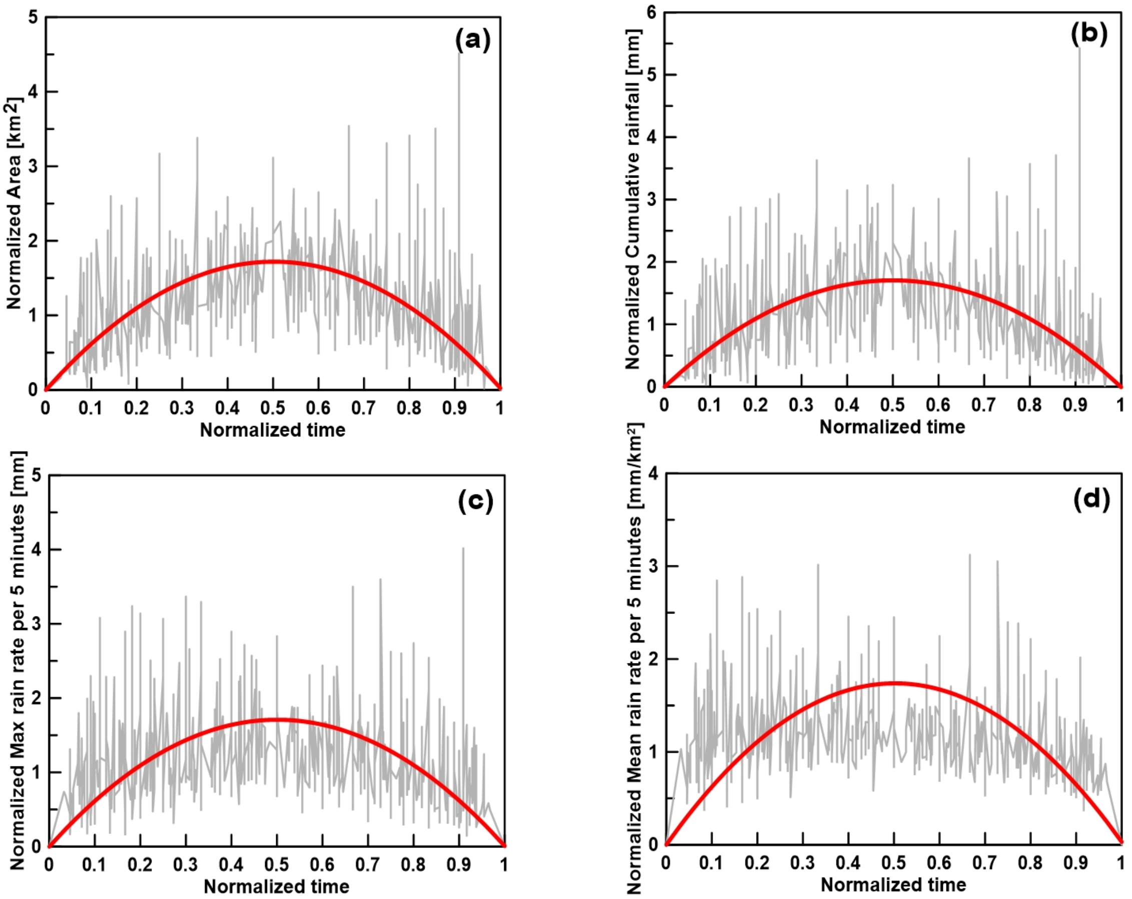


Figure 5.17 Same with Figure 5.16, but for characteristics of identified convective rain clusters.

Figure 5.16 and 5.17 gave the 'Normalized Duration Lines' for characteristics of identified light and convective rain clusters, gray lines in the figure were normalized value of identified rain cluster's characteristics which followed the normalized durations, red lines in the figure were the fitted curves with parabola functions. Table 5.3 presented the goodness of fit testing results for the 'Normalized Duration Lines' of rain cluster's characteristics by robust regressive method.

Through the goodness of fit testing results presented in Table 5.3, it was found that the parabola function was more robustly fitted to the characteristic's 'Normalized duration lines' for identified convective rain clusters compared to the light ones, and the best fits were for the characteristics area and cumulative rainfall. The goodness of fit testing results for parabola fitting of the identified rain cluster's characteristics indicated that the life stages of identified convective rain clusters more followed the 'Initial - Growing - Dissipating' order during their life cycles compared to the light ones.

Table 5.3 Goodness of fit testing results for the 'Normalized Duration Lines' of the rain cluster's characteristics.

Characteristics	**Reflectivity threshold**	**Parameters of parabola function**		**Goodness of fit testing results**	
		Parameter a	Parameter b	R-square	RMSE
Area [km^2]	> 19dBZ	-6.568	6.625	0.78	0.38
	> 37dBZ	-6.831	6.855	0.84	0.34
Cumulative rainfall [mm]	> 19dBZ	-6.582	6.584	0.72	0.49
	> 37dBZ	-6.837	6.827	0.83	0.36
Max rain rate per 5 minutes [mm]	> 19dBZ	-6.551	6.548	0.7	0.49
	> 37dBZ	-6.819	6.827	0.79	0.38
Mean rain rate per 5 minutes [mm/km^2]	> 19dBZ	-6.732	6.752	0.74	0.41
	> 37dBZ	-6.894	6.924	0.79	0.35

5.5 Verification Results for RCIT Algorithm

In this section, two object based spatial rainfall verification methods: 'SAL' (Structure-Amplitude-Location) and 'Geometric index' (Area-Connectivity-Shape) were implemented for evaluating the performance of developed RCIT algorithm. The modeled rain clusters were compared to those which were manually labeled from radar images.

5.5.1 SAL Verification Results

Considering the question: 'How similar are the modeled/forecasted objects to the observed ones according to a variety of descriptive criteria?'. The answer for this question is to extract several attributes (volume, intensity, shape) for modeled/forecast objects and observed objects separately and compare the similarity of these attributes,then the results are combined to give a synthesized summary for describing how well the modeled/forecast objects match the observed ones. This answer is also the principle of SAL method.

Taking precipitation objects as an example: Several characteristics of identified rain clusters are extracted. They are cumulative rainfall, mean rain rate per 5 minutes and center of mass. Correspondingly, same characteristics of rain cluster manually labeled from radar images are also extracted. For each identified/observed rain cluster pair, The normalized value of cumulative rainfall, mean rain rate per 5 minutes and center of mass is separately compared, the comparing results represent as the terms of Structure (S), Amplitude (A) and Location (L) respectively wherein the 'Structure' denotes how similar are the shape of the modeled precipitation objects to the observed ones, the 'Amplitude' denotes how similar is the averaged precipitation of the modeled precipitation objects to the observed ones and the 'Location' denotes how similar are the center of mass of the modeled precipitation objects to the observed ones. Then the performance of RCIT algorithm can be evaluated from the results of S, A and L components.

Figure 5.18 presented the SAL verification results for RCIT algorithm. Results were sorted by three rainy day. For each SAL plot in the Figure 5.18, the vertical axis denoted the A component, the horizontal axis denoted the S component, while the dots represented the value distribution for S and A components, the color scale for the dots denoted the L component. Each SAL plot in the figure was divided into four quadrants, the top-right part was the first quadrant, and following a counter clock- wise order, the top-left part was the second quadrant and the bottom-left, the bottom-right parts were the third and the forth quadrants, respectively. Each quadrant introduced the different descriptive criteria. Dots falling into the first quadrant indicated an over-estimation for both S and A components; dots falling into the second quadrant indicated an over-estimation for A component and an under-estimation for the S component; for dots falling into the third quadrant means an under-estimation for both S and A components; if the dots were falling into the fourth quadrant, this meant an over-estimation for the S component and an under-estimation for the A component. When the dots were falling on the zero point with a zero value of the L component, this meant a perfect match between modeled rain clusters and observed ones.

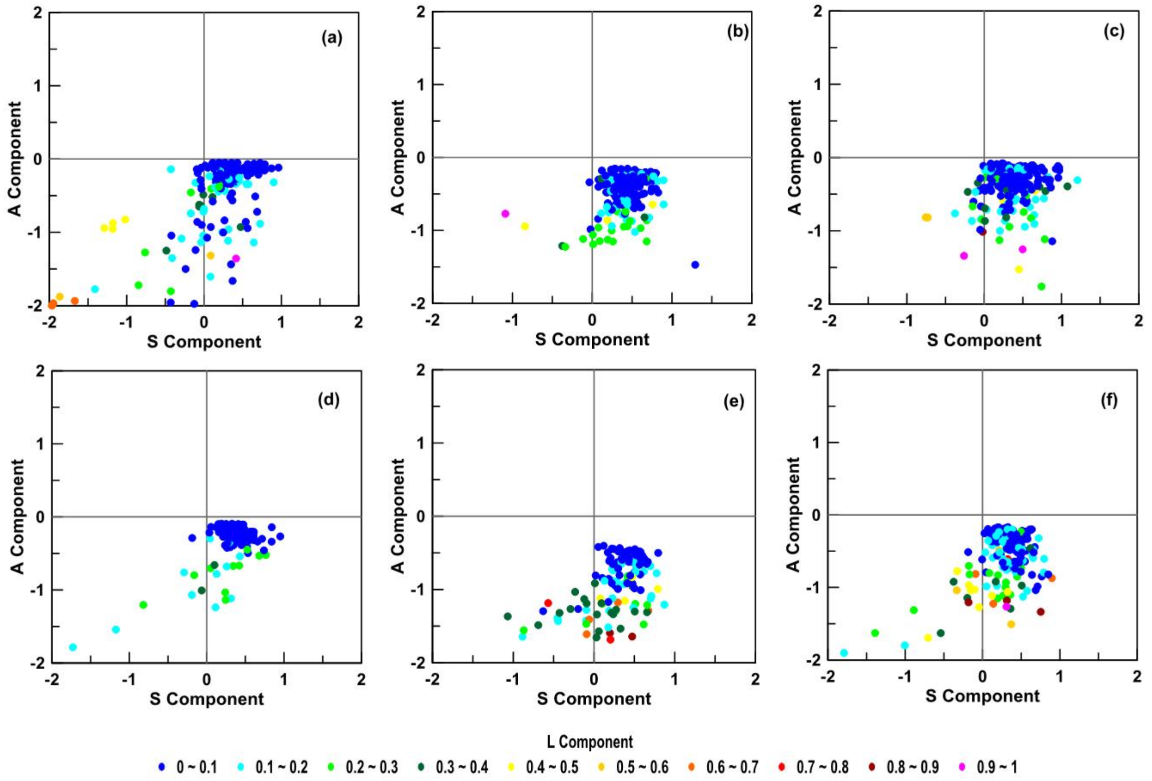


Figure 5.18 SAL verification results for RCIT algorithm, sorted by three rainy day. (a) for identified light rain clusters from 26-May-2007, (b) for identified light rain clusters from 19-Jul-2008, (c) for identified light rain clusters from 26-Jul-2008, (d) same with (a), but for identified convective rain clusters, (e) same with (b), but for identified convective rain clusters, (f) same with (c), but for identified convective rain clusters.

Through the Figure 5.18, most dots falling into the fourth quadrant, which meant an over-estimation for the S component and an under-estimation for the A component. Little dots falling into the third quadrant meant an over-estimation for the A component and an under-estimation for the S component. Through the value distribution of the L component, most modeled rain clusters were slightly miss matched to the observed ones.

For more quantitatively presenting the performance of the RCIT algorithm, the median value of three SAL components was calculated and the results were presented in Table 5.4. Through the Table 5.4, the RCIT algorithm produced more ’flat’ rain clusters with fewer rain rates compared to the observed ones according to the positive median value of S component and the negative median value of A component, not only for identified light rain clusters but also for the convective ones. Through comparisons among three day’s identified/observed rain clusters, identified rain clusters from 19-Jul-2008 were obviously miss matched to the observed ones. The median value of the L component indicated that a more or less deviation existed for weighted center's of all three day's identified rain clusters compared to the observed ones.

Table 5.4 Median value of three SAL components from verification results for RCIT algorithm, sorted by selected rainy days.

Rain days	**S component**		**A component**		**L component**	
	> 19dBZ	> 37dBZ	> 19dBZ	> 37dBZ	> 19dBZ	> 37dBZ
26-May-2007	0.272	0.327	-0.226	-0.288	0.0376	0.0423
19-Jul-2008	0.424	0.348	-0.352	-0.832	0.0544	0.104
26-Jul-2008	0.328	0.324	-0.301	-0.471	0.0504	0.0849

5.5.2 Geometric Index Verification Results

Geometric index method was used for the verification of RCIT algorithm by providing three geometry descriptive criteria: Area, Connectivity and Shape (see section 4.8.2 in Chapter 4). In this study, three geometry components were calculated and the results were shown as in Table 5.5. In the table, the label 'obs' represented the rain clusters which were manually labeled from radar images and the label 'sim' stranded for the identified rain clusters. Values of three components were divided into three quantile levels (25%, 50% and 75%).

Three geometry components calculated from the identified rain clusters were higher than those which were labeled from radar observations. For Connectivity component, the identified rain clusters were more closer to those from the manually labeled ones compared to the Area and Shape components at three percentile levels. Values of three components were more agreement between the identified and manually labeled light rain clusters than those calculated from identified convective ones. Agreements for identified and labeled rain clusters from rainy day 26-May-2007 was the best, and then were those from 26-Jul-2008 and the worst agreements for identified and labeled rain clusters were form 19-Jul-2008 at three percentile levels, this might be that most identified rain clusters from 19-Jul-2008 were smaller but contained more convective pixels. Through the summaries of the Geometric Index verification results which were presented in Table 5.5, the RCIT algorithm could well reflect the true geometry patterns (shape, inner connectivity and area) for light precipitation objects, but for the convective precipitation objects, other elements should be taken into the consideration.

Table 5.5 Three Geometric Index components at three quantile levels from the Geometric Index verification results for the RCIT algorithm, sorted by three rainy day.

		25% **quantile**				50% **quantile**				75% **quantile**			
	Rainy day	> 19dBZ		> 37dBZ		> 19dBZ		> 37dBZ		> 19dBZ		> 37dBZ	
		obs	sim	obs	sim	obs	sim	obs	sim	obs	sim	obs	sim
area	26-May-2007	0.06	0.16	0.07	0.2	0.11	0.28	0.14	0.3	0.34	0.55	0.2	0.4
	19-Jul-2008	0.03	0.05	0.01	0.02	0.06	0.09	0.02	0.03	0.11	0.14	0.03	0.25
	26-Jul-2008	0.03	0.09	0.02	0.04	0.08	0.17	0.04	0.06	0.19	0.32	0.06	0.13
connectivity	26-May-2007	0.85	0.97	0.84	0.97	0.89	0.98	0.9	0.98	0.93	0.99	0.94	1
	19-Jul-2008	0.72	0.91	0.81	0.94	0.75	0.95	0.87	0.98	0.83	0.97	0.96	1
	26-Jul-2008	0.8	0.94	0.84	0.95	0.83	0.97	0.88	0.97	0.86	0.98	0.94	0.99
shape	26-May-2007	0.19	0.29	0.19	0.31	0.27	0.39	0.27	0.36	0.46	0.58	0.32	0.47
	19-Jul-2008	0.15	0.16	0.09	0.12	0.21	0.24	0.11	0.14	0.29	0.31	0.13	0.37
	26-Jul-2008	0.14	0.21	0.1	0.14	0.23	0.32	0.15	0.18	0.37	0.43	0.19	0.26

5.6 Verification Results for Forecasting Methods

In this section, two newly radar-based short term precipitation forecasting methods: 'PIV_Semi-Lagrangian' and 'PIV_Lagrangian-Persistence' were developed and used to QPF with a leading time up to two hours. Their principles were described as follows:

(a) motion vectors at previous time interval was generated by the PIV algorithm, then it was interpolated into the grid with a size of 1×1 km^2.

(b) when forecasts started, the interpolated velocity fields were extrapolated in time and the advection fields were extrapolated by Semi-Lagrangian or Lagrangian persistence way.

Another two QPF methods: 'TREC_Semi-Lagrangian' and 'TREC_Lagrangian-Persistence' based on the TREC approach were also constructed as the comparisons. Proposed rain cluster's characteristics predicting method was also developed and applied for doing one hour ahead prediction. Additionally, the Kalman filter algorithm was applied for extrapolating the rain cluster's center of mass.

In order to evaluate the quality of developed forecasting methods in this study, three traditional verification parameters: Probability of Detection (POD), False Alarm Rate (FAR) and Critical Success Index (CSI) which were derived from the contingency table (Tessendorf and Einfalt, 2012) and the object-based spatial rainfall verification method - SAL were implemented, and the verification parameter - Root Mean Squared Error (RMSE) was also implemented for testing the performance of proposed rain cluster's predicting method.

Three traditional verifying parameters were introduced briefly. Being dependent on the comparison with an observed rain event, a forecast event can be categorized as happened (Yes) or non-happened (No). This categorization is non-probabilistic and discrete which can be presented by a 2×2 contingency table (Table 5.6)

Table 5.6 Contingency table for representing categorical precipitation forecasting results.

		Observations	
		Yes	**No**
Forecasted	**Yes**	a	c
	No	b	d

In the table, **a** presents that an rain event is forecasted and this event occurs, **b** presents that an rain event is forecasted and this event does not occur, **c** presents that an rain event is not forecasted and this event occurs, **d** presents that an rain event is not forecasted and this event does not occur. Then, the three verification parameters can be presented in equation(4.8).

$$
\begin{aligned}
&\mathbf{POD} = \frac{\mathbf{a}}{(\mathbf{a+b})} \\
&\mathbf{FAR} = \frac{\mathbf{c}}{(\mathbf{a+c})} \\
&\mathbf{CSI} = \frac{\mathbf{a}}{(\mathbf{a+b+c})}
\end{aligned}
\tag{5.8}
$$

5.6.1 Forecasts Setting Up

Eight periods with rainfall events occurred which were observed from Essen radar were selected for doing precipitation forecasting. Eight 'Single - tracks' trajectories calculated from RCIT algorithm were selected and eight ones for convective rain clusters were selected for doing characteristics prediction with a lead time up to one hour.

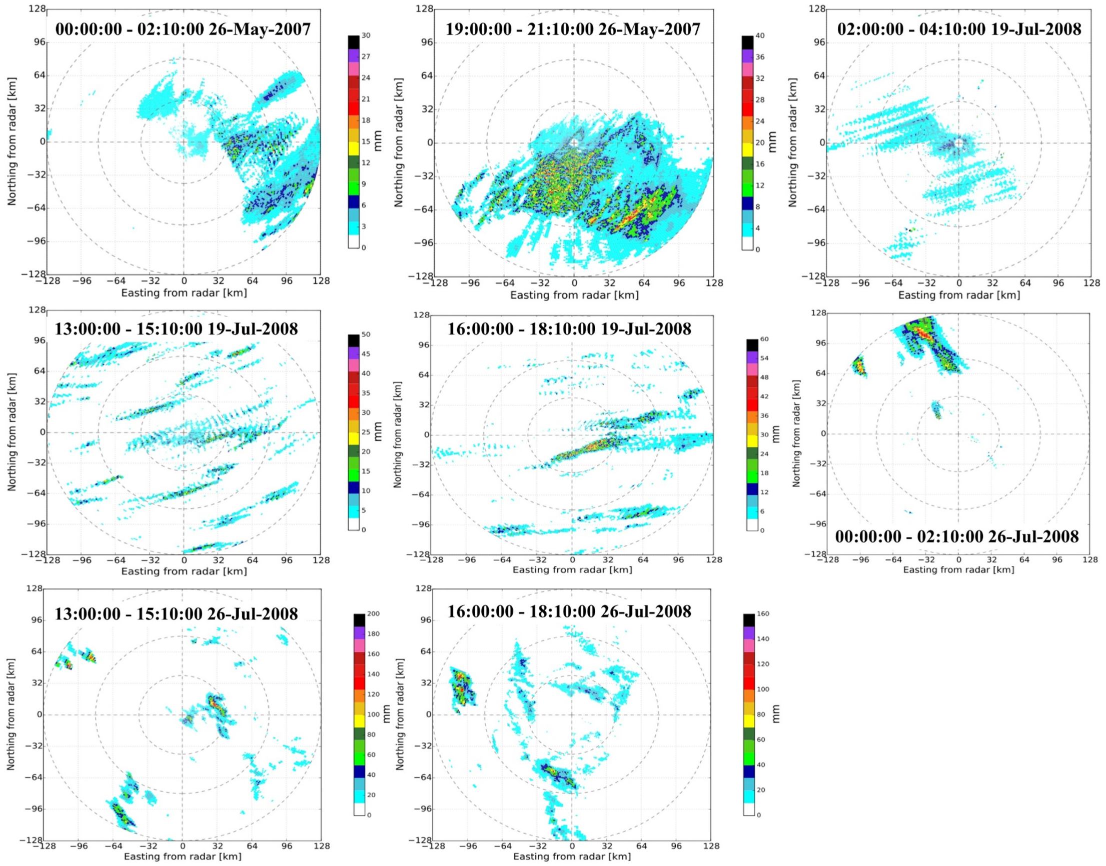


Figure 5.19 Cumulative rainfall for eight selected rainy periods, which were calculated from radar images.

Figure 5.19 and Figure 5.20 showed cumulative rainfall of eight rainy periods which were calculated from radar images and positions of selected ’Single - tracks’ trajectories. Table 5.7 presented the time periods of selected rainy periods for QPF and ’Single-tracks’ trajectories for characteristic predicting.

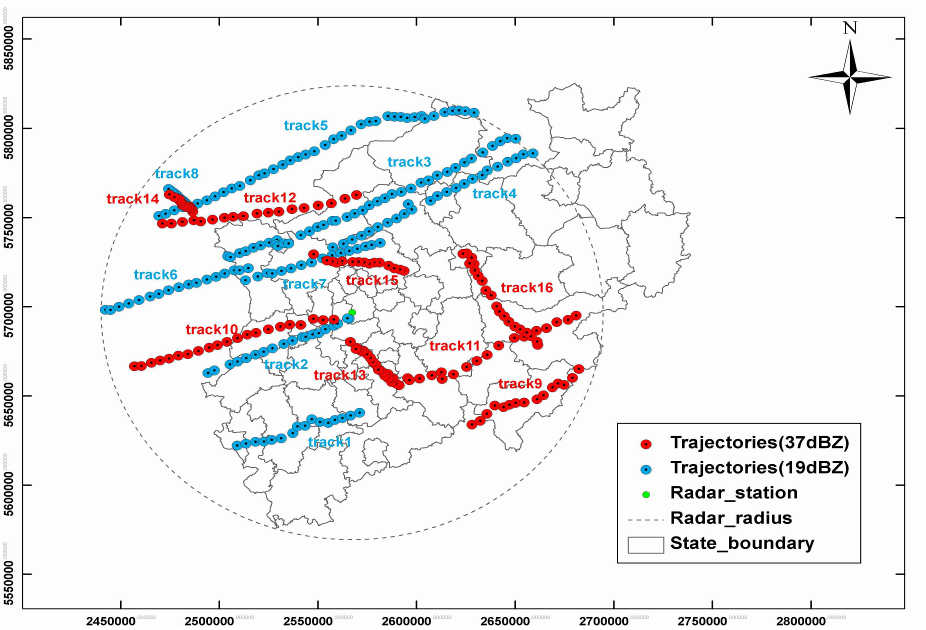


Figure 5.20 Locations of selected 'Single-tracks' trajectories.

Table 5.7 Selected rainy periods and trajectories for operating radar based QPF and rain cluster's characteristic prediction.

For QPF	**For rain cluster's characteristic prediction**			
	For light rain clusters		For convective rain cluster	
26-May-2007 00:00-02:10	track1	19-Jul-2008 01:45-02:55	track9	26-May-2007 00:00-01:10
26-May-2007 19:00-21:10	track2	19-Jul-2008 06:30-07:40	track10	19-Jul-2008 11:55-13:05
19-Jul-2008 02:00-04:10	track3	19-Jul-2008 10:30-11:40	track11	19-Jul-2008 14:50-16:00
19-Jul-2008 13:00-15:10	track4	19-Jul-2008 10:55-12:05	track12	19-Jul-2008 14:55-16:05
19-Jul-2008 16:00-18:10	track5	19-Jul-2008 11:40-12:50	track13	26-Jul-2008 06:20-07:30
26-Jul-2008 00:00-02:10	track6	19-Jul-2008 12:10-13:20	track14	26-Jul-2008 13:15-14:25
26-Jul-2008 13:00-15:10	track7	19-Jul-2008 13:50-15:00	track15	26-Jul-2008 16:20-17:30
26-Jul-2008 16:00-18:10	track8	26-Jul-2008 13:00-14:10	track16	26-Jul-2008 15:00-21:25

5.6.2 Results From Traditional Verification Method

Figure 5.21 showed the averaged value of three traditional verification parameters for four forecasting methods. Forecasted results by ’PIV_Semi-Lagrangian’ and ’PIV_Lagrangian-Persistence’ were better than those from ’TREC_Semi-Lagrangian’ and ’TREC_Lagrangian-Persistence’ within a forecasting length of 100 min. The verification differences for four QPF methods were not obviously when the forecasting length exceeded 100 min. Forecasting results within a lead time of 60 min were credibly but more for the results from ’PIV_Semi-Lagrangian’ method. Forecasting results from ’PIV_Lagrangian-Persistence’ method were more credible when the lead time was over 60 min. Convective QPF results (reflectivity > 37 dBZ, figure 5.21d-f) were more credible which indicated that newly developed forecasting methods applied in this study were more adopted to the convective QPF.

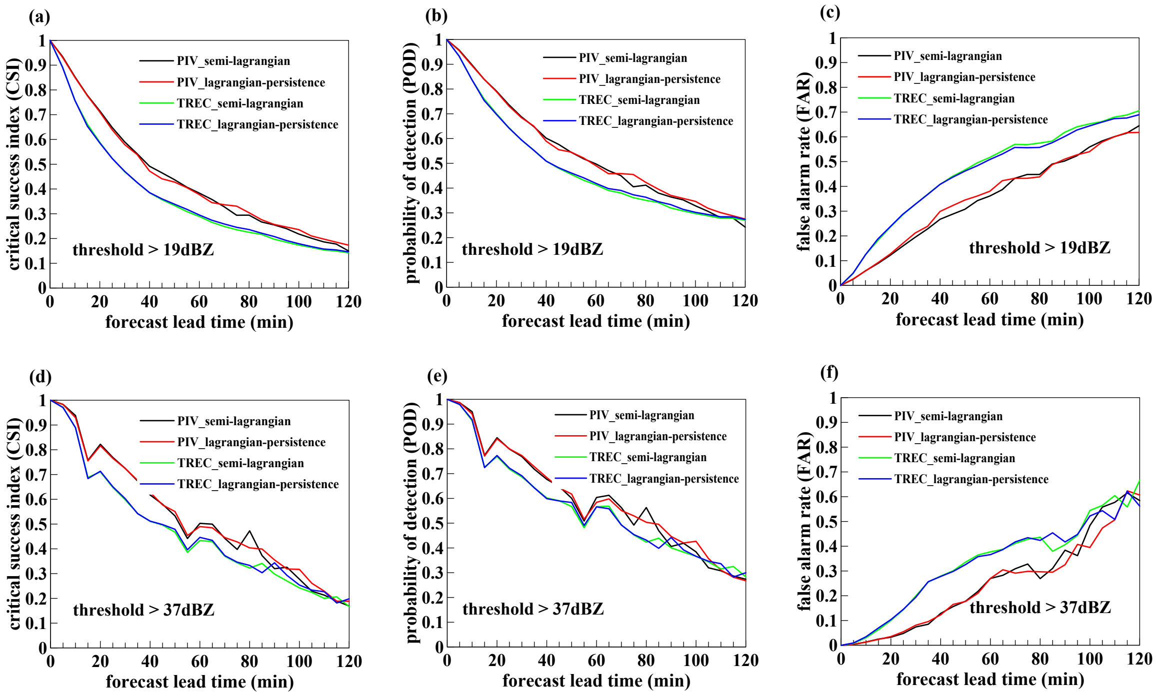


Figure 5.21 Averaged value of three traditional verification parameters for verfication results of four QPF methods. (a) for CSI score (ref > 19 dBZ), (b) for POD score (ref > 19 dBZ) (c) for FAR score (ref > 19 dBZ), (d) same with (a), but for convective forecasts (ref > 37 dBZ), (e) same with (b), but for convective forecasts (ref > 37 dBZ), (f) same with (c), but for convective forecasts (ref > 37 dBZ).

5.6.3 Results From SAL Verification Method

Figure 5.22 showed SAL verification results of four QPF methods respectively. In Figure 5.22a-d, most dots were scattered around the first, second and third quadrants with little ones falling into the fourth quadrant. Value distributions of three verification components indicated that the forecasted light rain clusters were more peaked with higher mean rain rates compared to the observed ones and their center of masses were miss matched to the observed rain cluster.

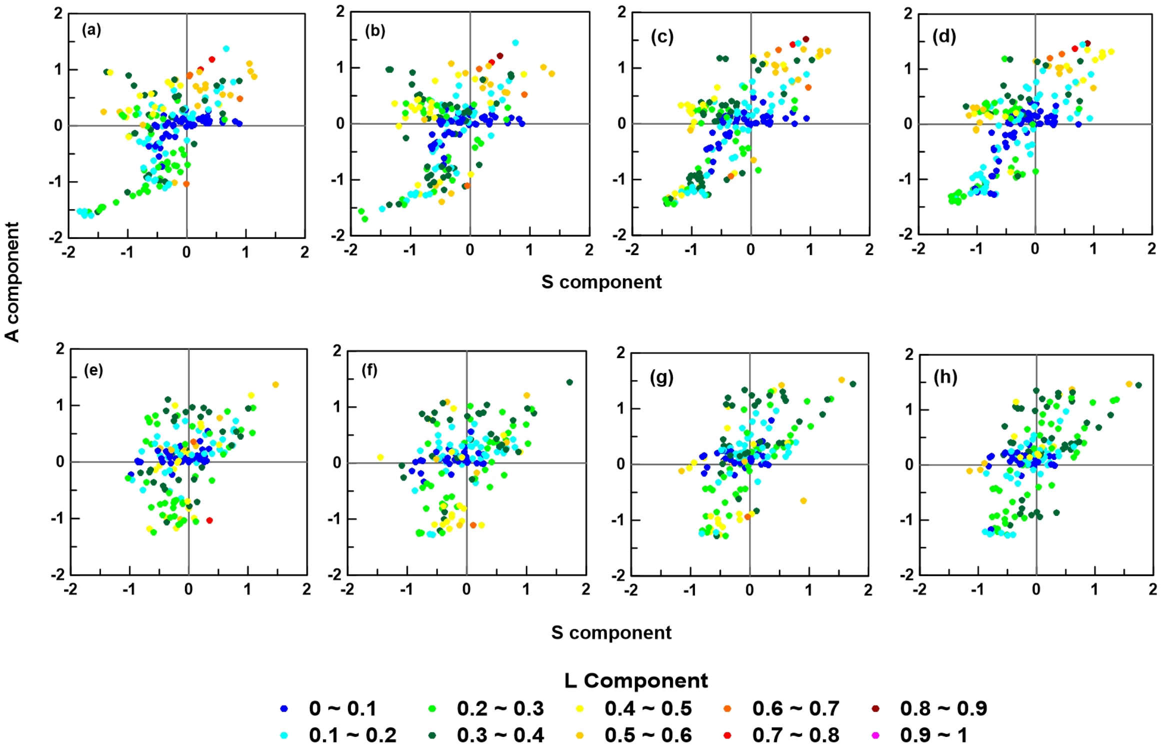


Figure 5.22 SAL verification results of four QPF methods. (a) for ’PIV_Semi-Lagrangian’ method (ref > 19 dBZ), (b) for ’PIV_Lagrangian-Persistence’ method (ref > 19 dBZ), (c) for ’TREC_Semi-Lagrangian’ method (ref > 19 dBZ), (d) for ’TREC_Lagrangian-Persistence’ method (ref > 19 dBZ), (e) same with (a), but for QPF with reflectivity > 37 dBZ, (f) same with (b), but for QPF with reflectivity > 37 dBZ, (g) same with (c), but for QPF with reflectivity > 37 dBZ, (h) same with (d), but for QPF with reflectivity > 37 dBZ.

In Figure 5.22e-h, value distributions of three SAL verification components were more around the zero points for forecasted convective rain clusters compared to the forecasted light ones which indicated that the forecasted convective rain clusters by four methods were more credible compared to the forecasted light ones by the same methods.

Table 5.8 presented the median value of three SAL components for four QPF methods. Forecasted light rain clusters (reflectivity > 19 dBZ) by ’PIV_Semi-Lagrangian’ method were more consist with the observed ones which the median value of S, A and L components were -0.057, -0.352 and 0.234, respectively and worst was from ’TREC_Lagrangian- Persistence' method with median values of three SAL components were -0.353, 0.135 and 0.186, receptively. For forecasted convective rain clusters (reflectivity > 37 dBZ), performances of four forecasting methods were almost same, but the best one was for the 'PIV_Lagrangian-Persistence' method for which median values of three SAL components were -0.085, 0.121 and 0.24, respectively. The worst forecasted results were from the 'PIV_Semi-Lagrangian' with median values of S, A and L components being -0.145, 0.125 and 0.251 respectively.

Table 5.8 Median value of three SAL components from SAL verifying results for four QPF methods.

	PIV_Semi-Lagrangian		PIV_Lagrangian-Persistence		TREC_Semi-Lagrangian		TREC_Lagrangian-Persistence	
	> 19dBZ	> 37dBZ	> 19dBZ	> 37dBZ	> 19dBZ	> 37dBZ	> 19dBZ	> 37dBZ
S component	-0.352	-0.145	-0.361	-0.085	-0.366	-0.09	-0.353	-0.074
A component	0.057	0.125	0.075	0.121	0.103	0.154	0.135	0.17
L component	0.234	0.251	0.259	0.24	0.195	0.23	0.186	0.231

Figure 5.23 presented the variability of median values of SAL components for four QPF methods at the different lead times. For light rainfall forecasts (threshold $>$ 19 dBZ), all of the forecasting methods produced small and peak rainfall fields (Fig 5.23a), for which the PIV_Semi-Lagrangian was worse when the lead time was between 35 and 60 min. At long lead time, the shape of the forecasted rainfall fields was still small and peak but for which the PIV_Lagrangian-persistence method was more worse. The precipitation of the forecasted rain clusters using the four methods was somewhat more overestimated with the lead time less than 60 min (Fig 5.23b), for which the amplitude value of developed methods (PIV_Semi-Lagrangian and PIV_Lagrangian-persistence) were more to the zero. However the amplitude value for method PIV_Lagrangian-persistence were more deviated to the zero level when the lead time was over 65 minutes. The center of mass of the forecasted rain clusters using developed methods were more displaced in comparison with those produced by the reference methods (Fig 5.23c).

For the convective forecasts (threshold $>$ 37 dBZ), the shape of forecasted rain clusters by the method - PIV_Semi-Lagrangian was more small and peak compared with the results of other three methods (Fig 5.23d). The precipitation of the forecasted rain clusters using the developed methods was also overestimated and the developed methods got the amplitude value which was more close to zero except for the forecasted rain clusters at lead time between 65 and 90 minutes (Fig 5.23e). It was, similarly, for forecasts of a threshold above 19 dBZ that the center of mass of the forecasted rain clusters using the four methods were more deviated when the lead times extended and there were no obvious value difference for location component between the developed methods and reference ones (Fig 5.23f).

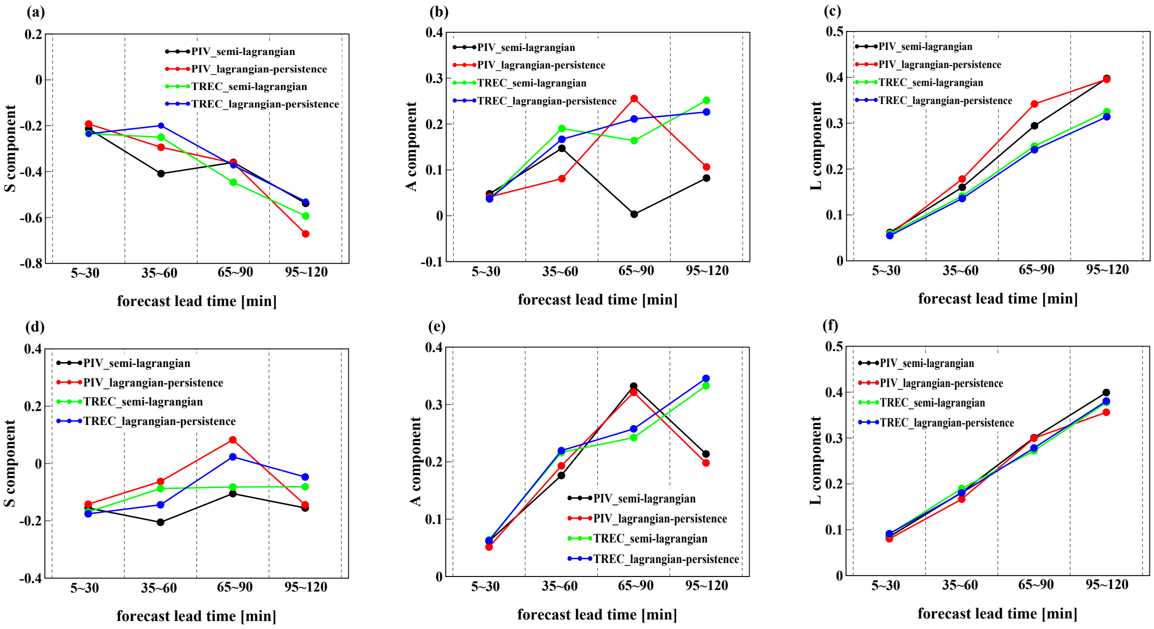


Figure 5.23 Variability of the median values of three SAL components for four QPF methods at different lead times. (a) for S component (reflectivity > 19 dBZ), (b) for A component (reflectivity > 19 dBZ), (c) for L component (reflectivity > 19 dBZ), (d) same with (a), but for QPF results with reflectivity > 37 dBZ, (e) same with (b), but for QPF results with reflectivity > 37 dBZ, (f) same with (c), but for QPF results with reflectivity > 37 dBZ

5.6.4 Verification Results for Characteristics Predicting Method

The proposed characteristic predicting method was applied for predicting rain cluster's characteristics with one hour ahead. The predicted characteristics were: Area, Cumulative rainfall, Max and Mean rain rate per 5 minutes. For this purpose, their mean values at three previous time intervals were calculated as the inputs of equation (4.20) and the forecast length (60 minutes by 5 minute intervals) was normalized and inputted as the time parameters of the prediction equation. At each forecast length, the forecasted characteristic was calculated and its mean value was updated by combing with the current forecasted value, then the new forecasted value at the next time interval was generated.

Figure 5.24 - 5.27 presented the predicted results for light rain cluster's characteristics which were compared to the selected trajectories. Figure 5.28 presented the extrapolated results for weighted center of light rain clusters by Kalman filter method (the vertical dashed line was the starting point of prediction). Table 5.9 presented the RMSE verification results of the predicting results for light rain clusters.

- Figure 5.24, for predicting results of characteristic area, high RMSE existed for track 1 track 4 and track 8 with values of 62.67, 99.422 and 150.633 km^2, respectively. From the comparison between forecasted/observed values, reason for resulting in such high errors was that characteristic values for these trajectories presented growing trends in their dissipation stages.

- Figure 5.25, for predicting results of characteristic cumulative rainfall, track 7 and track 8 presented the high bias, and their RMSE values were 63.75 and 241.884 mm, respectively. The high biases also occurred at the dissipation stage for their trajectories.

- Figure 5.26, Max rain rate predicting results for selected tracks showed the same biasing situation compared to the predicting results of characteristic of cumulative rainfall, higher RMSE values also existed for track 7 (5.258 mm) and track 8 (5.947 mm).

- Figure 5.27, for characteristic mean rain rate per 5 minutes, its predicting results for track 4, track 5, track 7 and track 8 existed high RMSE value (0.243, 0.221, 1.075 and 1.196 mm/km^2, respectively), and the predicting results for these trajectories at dissipation stage were lower than the observed ones.

- Figure 5.28, the center of mass extrapolating results for the eight selected trajectories showed no distinct bias (six and seven column in table 5.9), and their RMSE values floated at 1 km level.

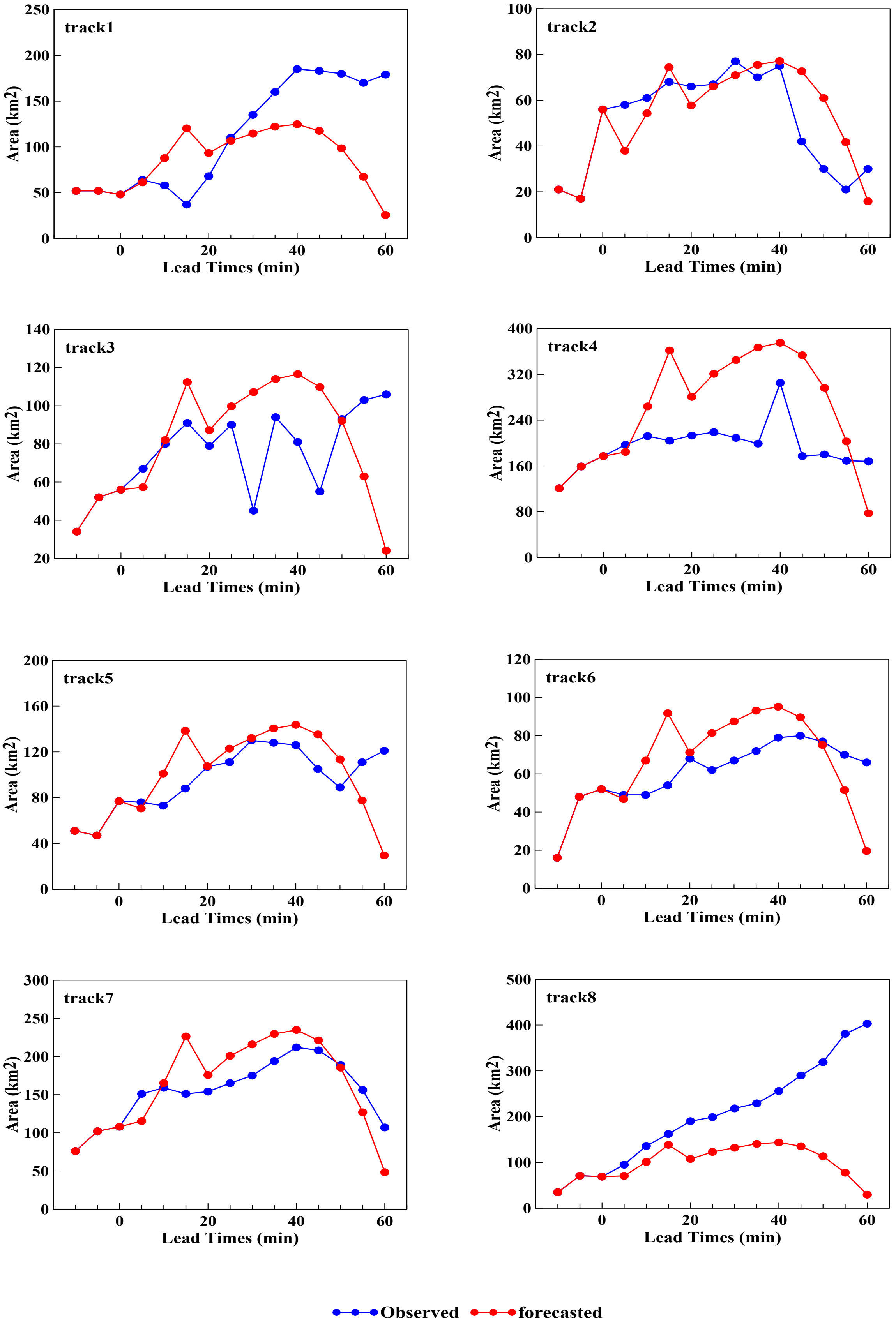


Figure 5.24 Predicting results for rain cluster's area (reflectivity > 19 dBZ).

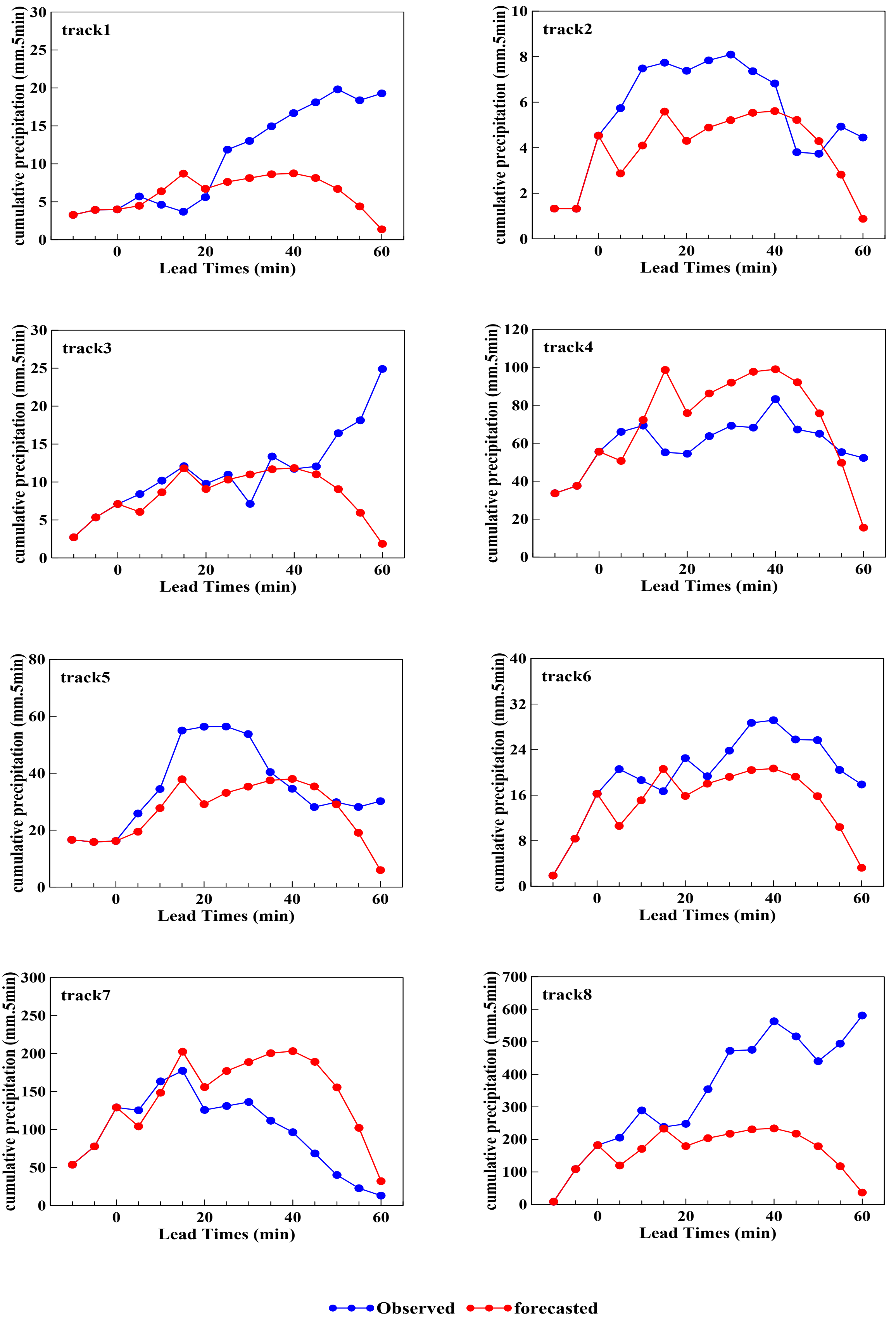


Figure 5.25 Same with Figure 5.24, but for characteristic cumulative rainfall.

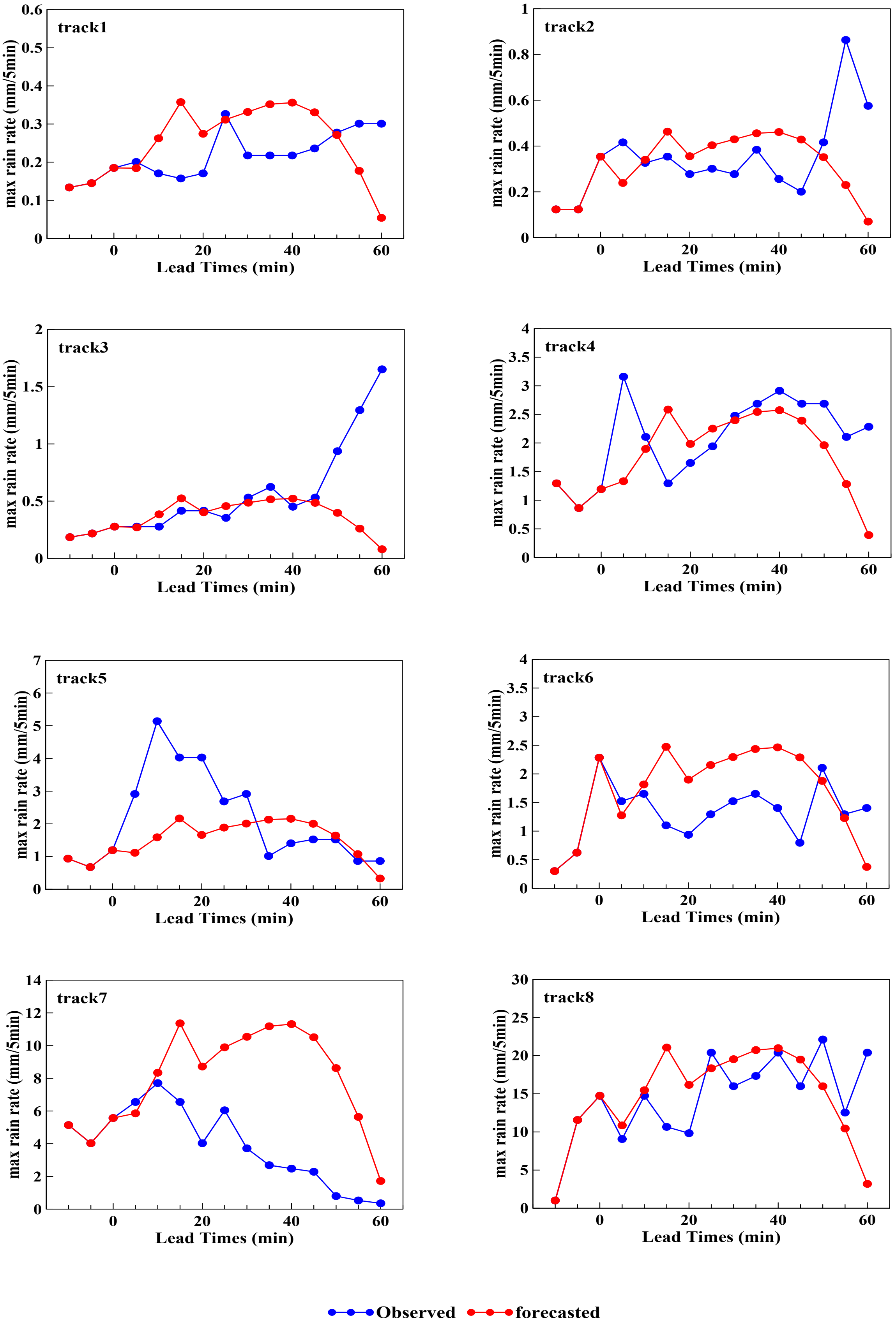


Figure 5.26 Same with Figure 5.24, but for characteristic max rain rate per 5 minutes.

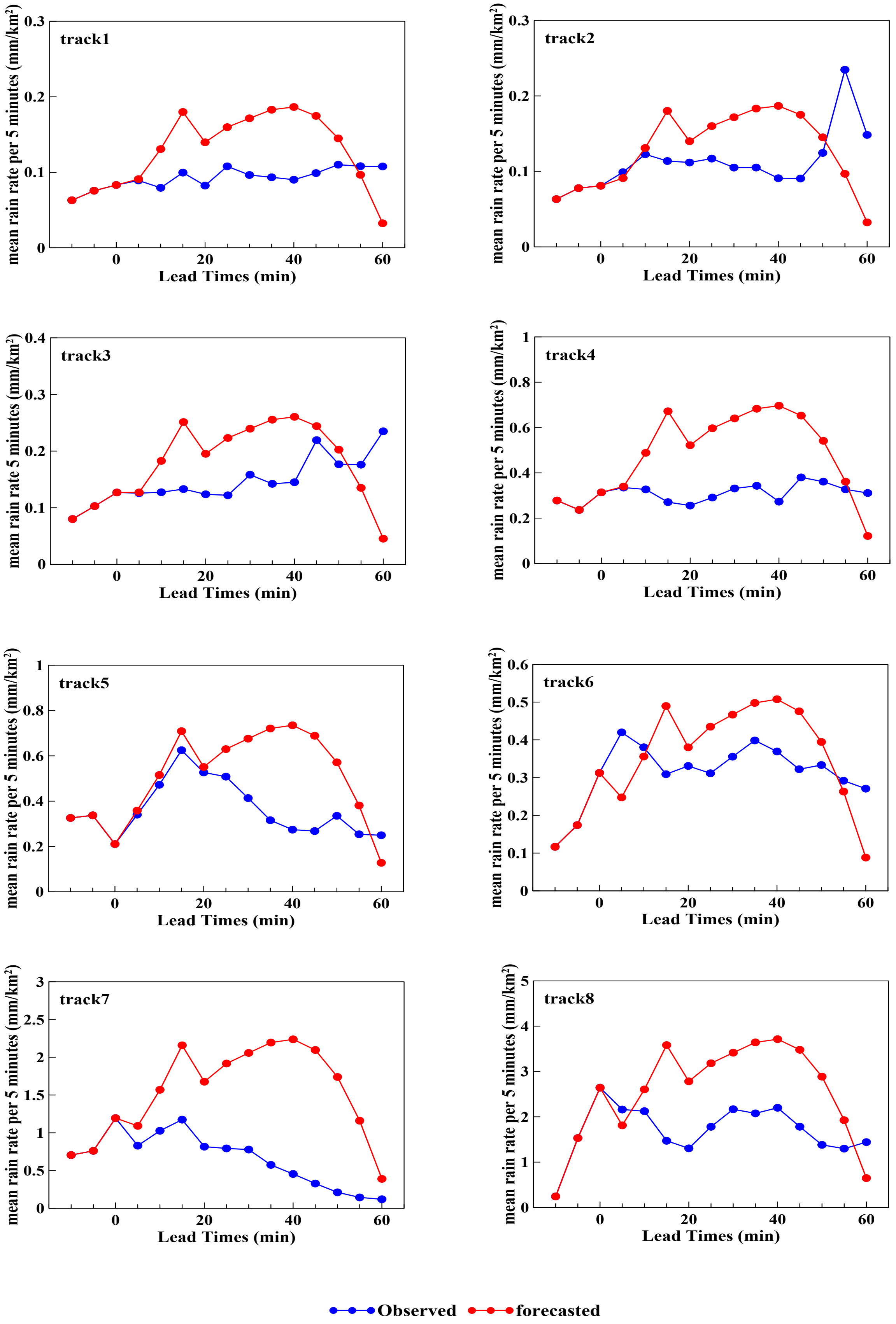


Figure 5.27 Same with Figure 5.24, but for characteristic mean rain rate per 5 minutes.

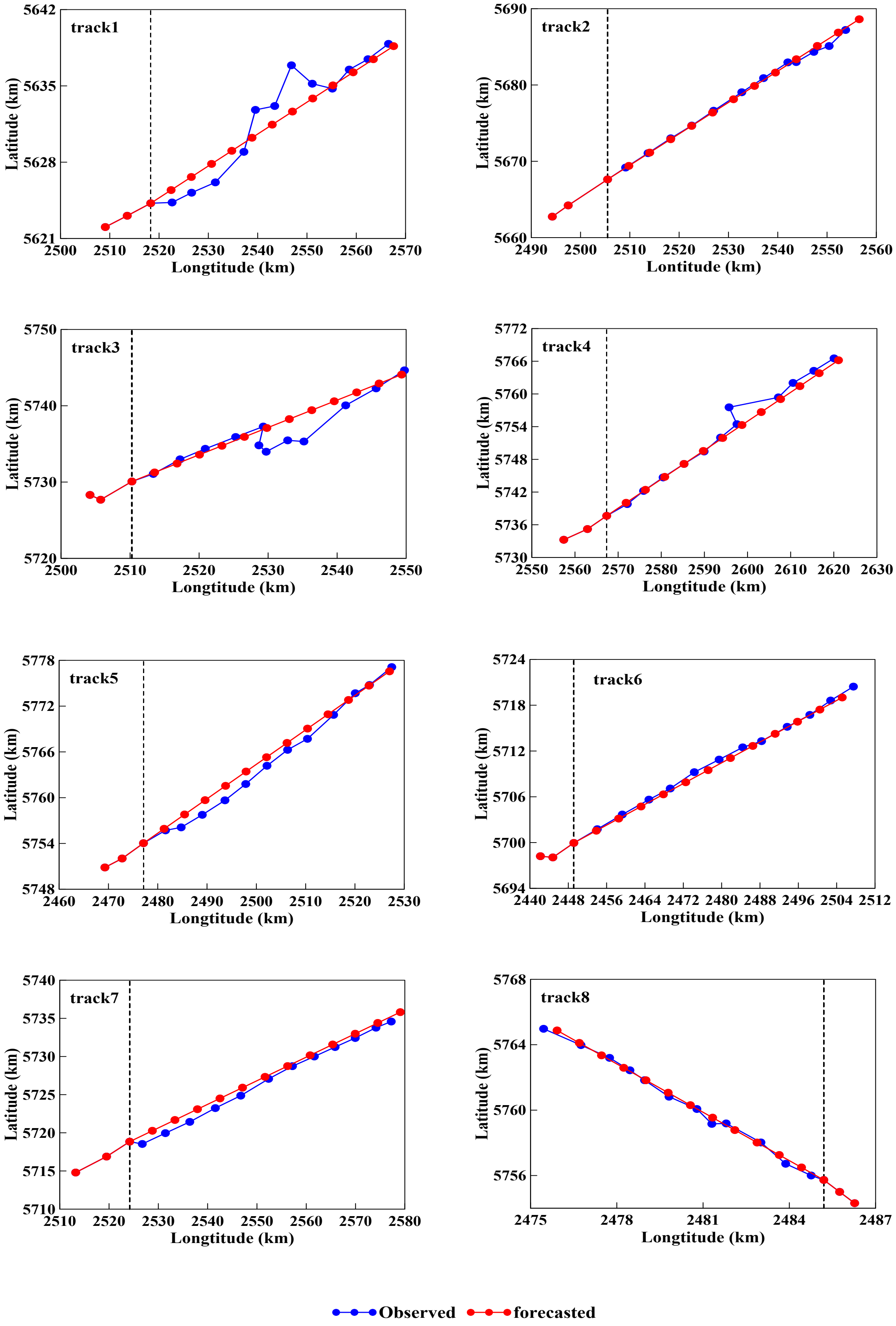


Figure 5.28 Predicting results for center of masses of rain clusters (reflectivity > 19 dBZ).

Table 5.9 RMSE verification results of proposed predicting method for light rain clusters (reflectivity > 19dBZ).

track id	Area	Cumulative rainfall	Max rain rate	Mean rain rate	Center of mass	
					x_coords	y_coords
track1	62.67	8.016	0.114	0.058	1.096	3.36
track2	14.503	2.237	0.237	0.067	0.525	0.58
track3	34.234	7.124	0.509	0.083	2.247	3.37
track4	99.422	21.346	0.827	0.243	1.527	0.66
track5	31.593	13.527	1.385	0.221	1.905	1.616
track6	19.79	7.257	0.788	0.111	1.202	1.469
track7	33.343	63.785	5.258	1.075	1.043	1.479
track8	150.663	241.884	5.947	1.196	1.617	2.349

Characteristics predicting results for convective rain clusters were presented in Figure 5.29 - 5.32 and the prediction results for their center of masses were presented in Figure 5.33. Table 5.10 summarized the RMSE verification results of the predicting results for convective rain clusters.

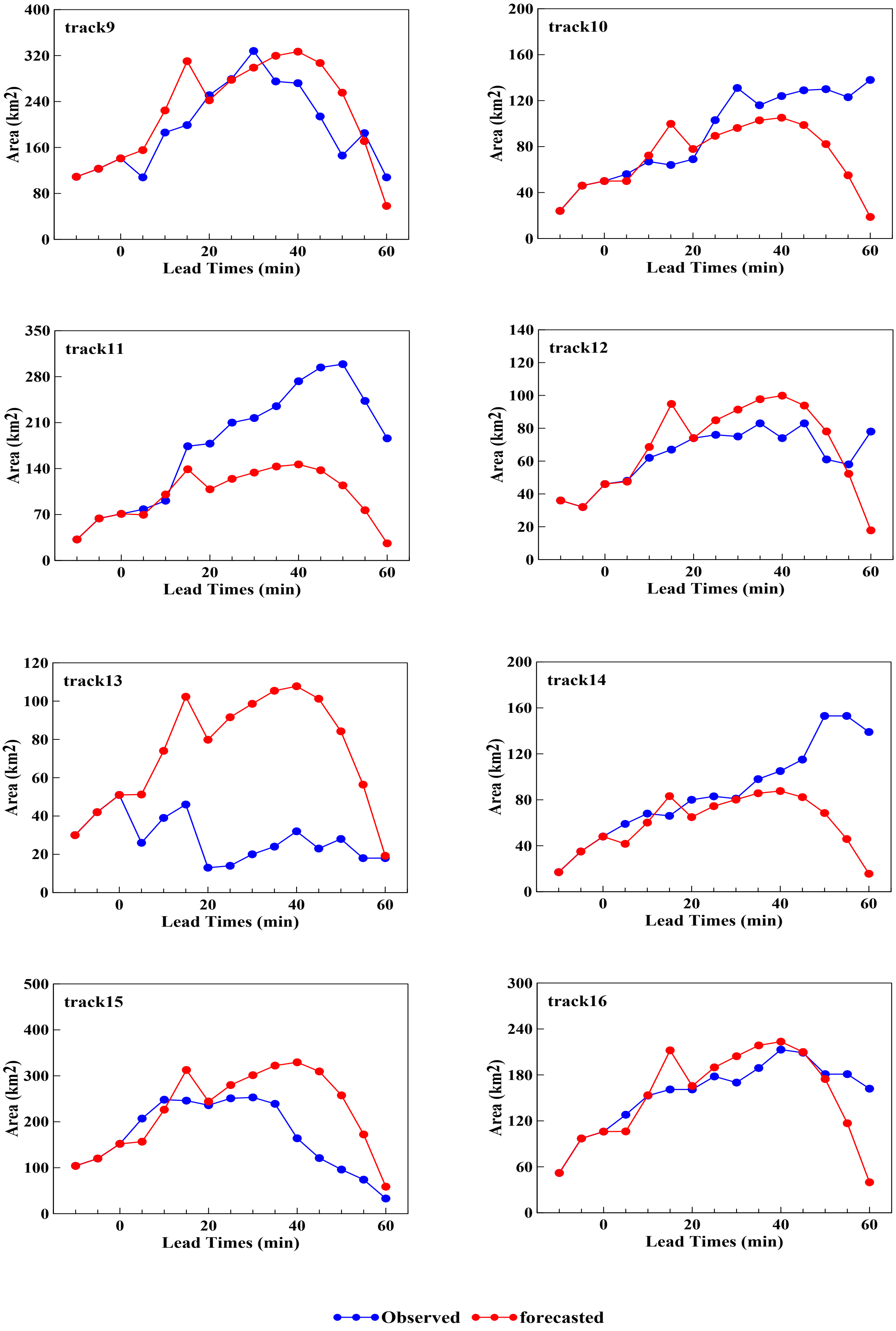


Figure 5.29 Predicting results for rain cluster's area (reflectivity > 37 dBZ).

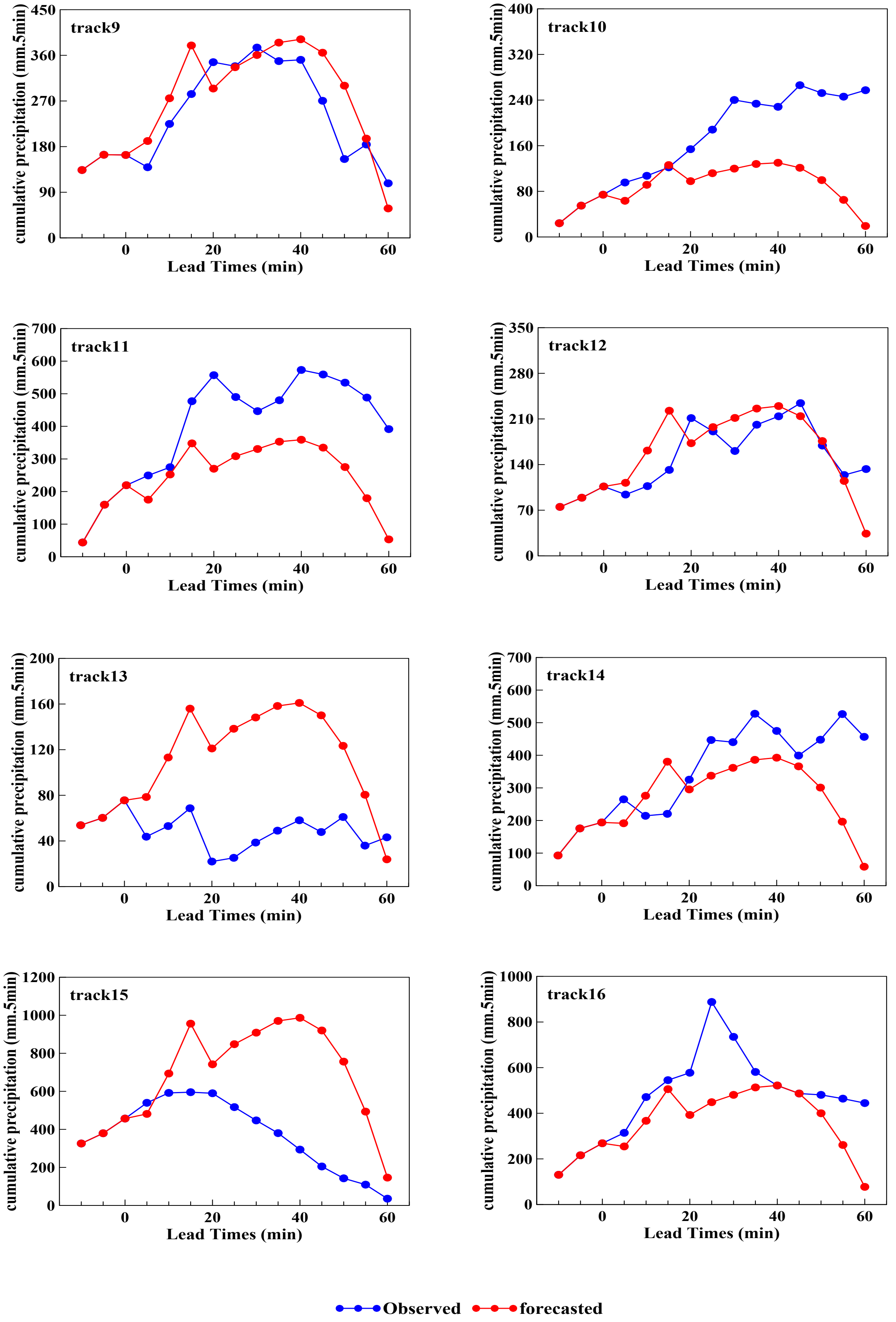


Figure 5.30 Same with Figure 5.29, but for characteristic cumulative rainfall.

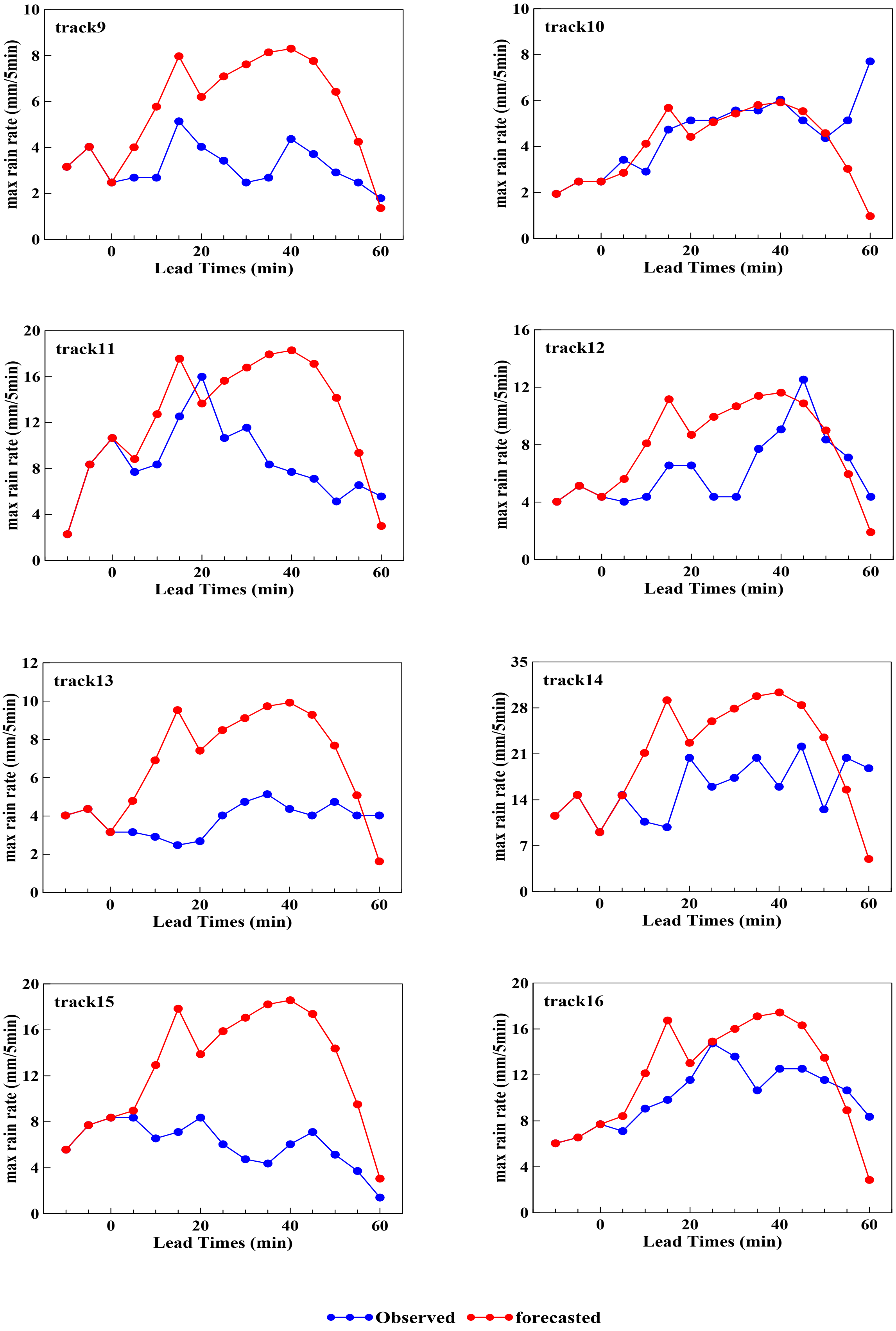


Figure 5.31 Same with Figure 5.29, but for characteristic max rain rate per 5 minutes.

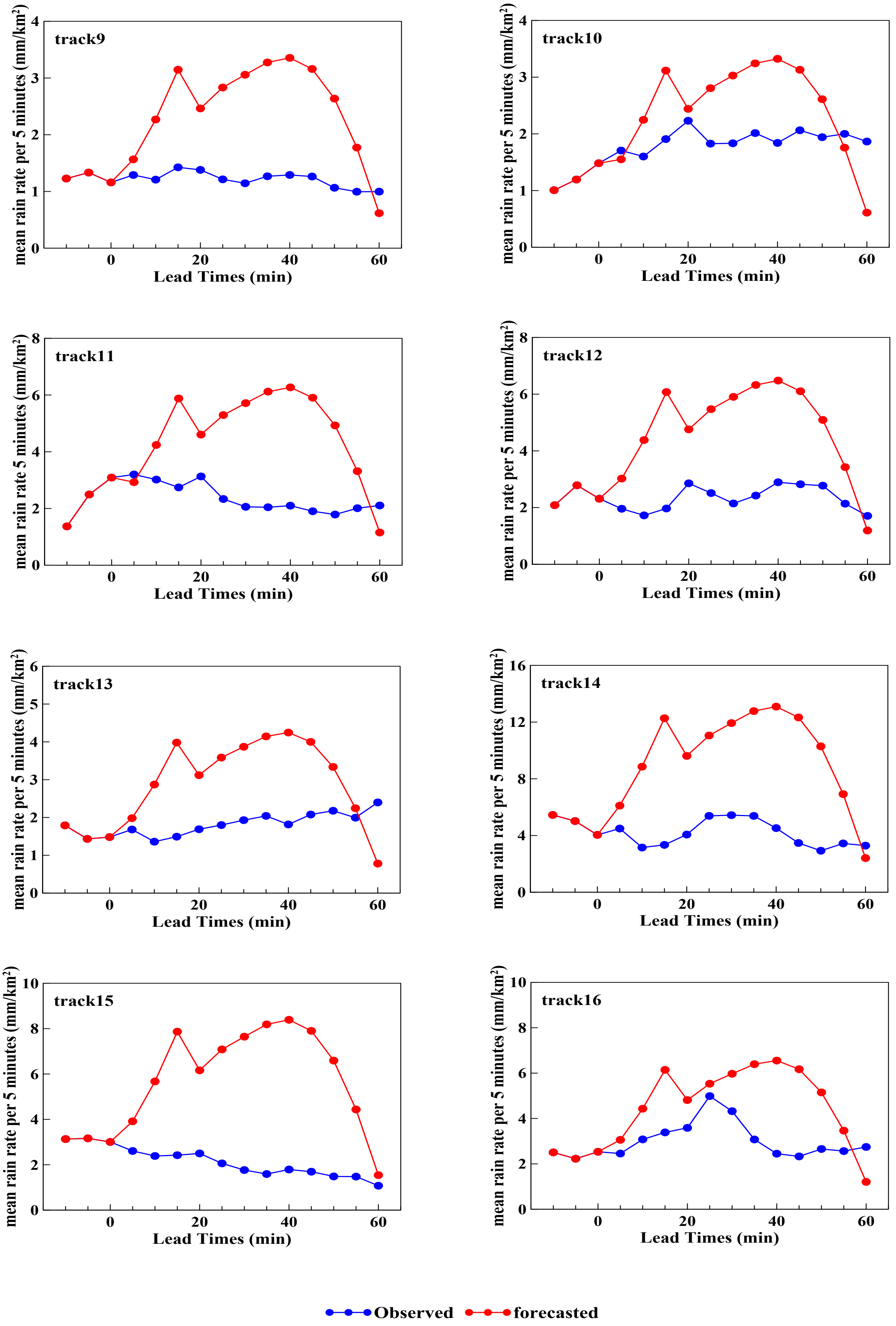


Figure 5.32 Same with Figure 5.29, but for characteristic mean rain rate per 5 minutes.

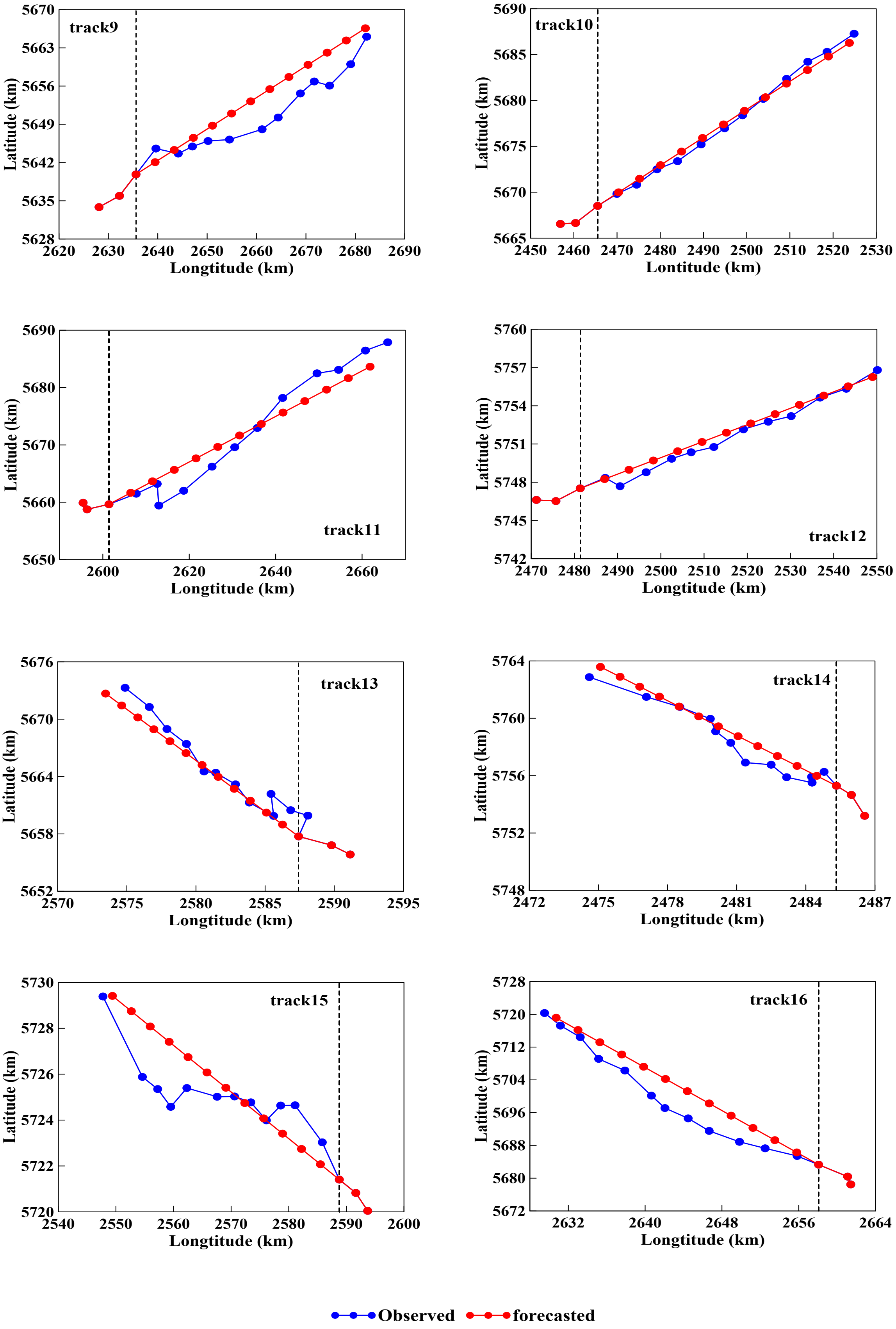


Figure 5.33 Predicting results for center of mass of rain clusters (reflectivity > 37 dBZ).

Table 5.10 RMSE verification results of proposed predicting method for convective rain clusters (reflectivity > 37dBZ).

track id	Area	Cumulative rainfall	Max rain rate	Mean rain rate	Center of mass	
					x_coords	y_coords
track9	54.96	59.333	3.068	1.335	0.832	1.544
track10	41.15	109.288	1.885	0.866	1.283	0.802
track11	102.261	189.869	5.791	2.559	1.989	2.23
track12	20.192	42.298	3.094	2.552	2.05	0.334
track13	54.663	75.861	3.872	1.541	0.536	1.091
track14	49.215	157.259	9.554	5.736	1.774	0.924
track15	88.292	397.172	8.214	4.289	1.074	0.96
track16	40.405	181.629	3.499	2.103	0.212	0.263

- In Figure 5.29, for the characteristic area, high RMSE value existed for track 11, the value was 102.61 km^2) and track 15 with the value of 88.292 km^2, through their predicting/observed comparison(the third column in figure 5.28), bias occurred at mutate and dissipation stage of their trajectories with a low predicting value.
- In Figure 5.30, for predicting results of the characteristic cumulative rainfall, their RMSE values for track 10, track 11, track 14, track 15 and track16 were higher than others obviously. For track 15, the RMSE value reached 397.172 mm. From predicted/observed comparisons, there were low predicting results existed for these trajectories except the track 15.
- In Figure 5.31, for max rain rate predicting results, high RMSE value existed for track 11, 14 and 15, and the RMSE values were 5.791, 9.554 and 8.214 mm, respectively. Through the comparison between the predicting and observed values, there were high predicted values existed for these trajectories.
- In Figure 5.32, over-high predicted results for characteristic mean rain rate per 5

minutes existed for all the selected trajectories, and through the comparison for their RMSE values, track 14 and track 15 were the highest ones with RMSE value of 5.736 and 4.289 mm/km^2, respectively.

- In Figure 5.33, for center of mass extrapolating results, their RMSE values were less than those from light rain clusters and floated around 1 km.

CHAPTER 6

DISCUSSION

Rainfall is one of the key factors in catchment's hydrological process whose variability on the modeling of a catchment response to rainfall forcing has long been a concern for hydrologists (Schoorens et al., 2014). As one type of rainfall measuring equipment, weather radar provides more useful information at high sptial-temporal resolution. Main objects of this study are the analysis of spatial and temporal characteristics of rain events observed from weather radar images and short-term quantitative precipitation forecasting as well as their verification.

A purely radar based precipitation object identification and tracking algorithm: RCIT was developed and applied for extracting characteristics of precipitation objects (rain clusters). The main advantage of RCIT algorithm is that it can produce more 'realistic' precipitation objects and reflect life stages of them (e.g. merging, splitting and internal growth&decay) by combing advection field and cell tracking approaches. Unlike the TREC (Rinehart and Garvey, 1978) and CO-TREC (Li et al., 1995) algorithm, but more similar with MTREC algorithm (Wang et al., 2007), estimating motion vectors by RCIT algorithm was based on Partial Image velocimetry (PIV) way, by which the extracted motion vectors were more robust and consisted, especially for radar images with speckle echo.

For assessing the efficiency of RCIT algorithm, radar images from Essen radar deployed in North Rhine Westphalia (NRW) were used as input, the object-oriented spatial rainfall verification methods: SAL and Geometric Index were implemented for evaluating the produced precipitation objects. As a weather variable, rainfall is often predicted as fields defined over a spatial domains (Casati et al., 2008). Spatial fields are characterized by a coherent structures which contain some physical/geometry features. Traditional verification methods (e.g. Critical Successor Index, Root Mean Squared Error) which are based on point by point comparison don't account for the intrinsic spatial correlation existing in them, while object based verification methods provides more diagnostic components which give more feedbacks on the physical nature of forecast error generated by forecasting methods (e.g. Gilleland, 2013; Schneider et al., 2014; Zacharov et al., 2013; Zimmer and Wernli, 2011). To the author's knowledge, It is the first time that applying objective verification methods for evaluating the performance of rain cell tracking algorithm.

Two novel advection field based forecasting methods: 'PIV_Semi-Lagrangian' and 'PIV_Lagrangian-Persistence' were developed for the aim of improving the capability of short-term quantitative precipitation forecasting. The developed QPF methods were also based on the PIV scheme and advection fields were extrapolated by Semi-lagrangian and Lagrangian-Persistence approach, respectively.

Moreover, precipitation objects (rain cluster) were identified by applying the RCIT algorithm and their properties were extracted. Questions proposed in this study (see section 1.6 in Chapter 1) were answered and discussed.

6.1 Inner Structures of Rain Events

Two types rain clusters were produced from the RCIT algorithm. For simulated light rain clusters (reflectivity $>$ 19 dBZ), their area and cumulative rainfall showed a high value range, but most were under 40 km^2 with the cumulative rainfall under 4.91 mm. Their peak and averaged rain rates were under 0.24 mm and 0.1 mm, respectively, and their shapes were more elliptical. Considering simulated convective rain clusters (reflectivity $>$ 37 dBZ), high value range also existed for their area and cumulative rainfall, and most were smaller than the light ones with the area under 28 km^2 but contained more precipitations (the median value was 43.35 mm). Their peak and average rain rates were higher than the light ones with the peak rain rate under 2.91 mm and the average rain rate under 1.5 mm. It was found that the simulated convective rain clusters were more elliptical (the median value was 0.85).

Distributions of the log_{10}transformed characteristics for simulated light rain clusters showed obvious diversity. For their area, cumulative rainfall, max and mean rain rate per 5 minutes, a Generalized Pareto distribution (GPD) could fit their empirical distributions, and for characteristic eccentricity, a Generalized Extreme Value distribution (GEV) was the best fit. Similar to the simulated light rain clusters, a Generalized Pareto distribution could also fit to the log_{10} transformed area of simulated convective rain clusters. Different from the simulated light rain clusters, the Inverse Gaussian distribution (Inv_Gauss) was found to fit log_{10}transformed cumulative rainfall of convective ones. For the log_{10} transformed max and mean rain rate of simulated convective rain clusters, a Generalized Extreme Value distribution (GEV) could fit. Similar to the light rain clusters, characteristics eccentricity of convective rain cluster also could be fitted with the GEV distribution. Distribution fitting results for characteristic area and eccentricity of simulated light&convective rain clusters were more agreement with Barnolas et al. (2010) which their study area were located in Catalonia region, but were different to the Log-Normal distribution from Karklinsky and Morin (2006) in southern Israel and Capsoni et al. (2008) in Padana Valley, Italy. Apart from the characteristic area, distribution diversity for other characteristics of simulated rain clusters were also investigated and could be described by their fitted theoretical distributions, which might be useful for further research (e.g. probabilistic precipitation estimation and forecast). Value distributions for characteristics of simulated rain clusters also suggested that the inner structures of rain events occurred in the NRW showed extremely unbalance: most light events were small in size and contained few precipitations, and it was more for convective events.

6.2 Relations of Precipitation Object's Characteristics

Relations for simulated rain cluster's characteristics were also investigated and quantified by some simple mathematic functions. It was found that the log_{10} transformed cumulative rainfall could be a linear function of log_{10} transformed area, not only for the simulated light rain clusters, but also for the convective ones, which indicated that the integral rainfall volumes of simulated precipitation objects presented a power-low increasing trend following the variability of their area.

Relations for log_{10} transformed characteristics (area, cumulative rainfall, max and mean rain rate) and characteristic number of convective pixels were also investigated. By this novel approach, connections between physical and geometric properties of precipitation objects and their internal growing&decaying trend could be established. It was found that the log_{10} transformed area, cumulative rainfall for the simulated light&convective rain clusters were found to be the power-law function of the character number of convective pixels, which indicated that the increase of size and integral rainfall volume for simulated rain clusters were directly affected by their internal growth, this increasing trend would retard when rain cluster's inner structure got stable. Through the goodness of fit testing results for relations between log_{10} transformed max&mean rain rate and the number of convective pixels, whether the variability of peak and average rain rate of simulated rain clusters were also directly affected by their internal growth&decay was still blur.

6.3 Spatio-Temporal Variability of Precipitation Objects

Spatio-temporal correlations of simulated rain clusters presented an exponential decreasing trend following the spatial and temporal scales, which was more similar to the results from Luini and Capsoni (2012) and Peleg et al. (2013). Generally speaking, the higher spatial correlation coefficient existed for those simulated rain clusters with big size, which indicated that those rain clusters were more smoothly than ones with small size at minor spatial scales, and all the rain clusters got smoothly when the spatial scale extended. Highly temporal correlation coefficients of simulated rain clusters with long durations at minor temporal scales indicated the stable structures for them compared to those with short durations. Through the goodness of fit testing results, spatial or temporal correlations of identified rain clusters could be fitted with an Exponential function as in equation (6.1).

$$\mathbf{Corr(\Delta s, \Delta t) = ae^{b(\Delta s, \Delta t)}} \tag{6.1}$$

where Δs, Δt represent the spatial and temporal scale, respectively. a and b are the coefficient parameters of Exponential function. These results were similar with (Ciach and Krajewski, 2006; Haile et al., 2009) in which the spatial and temporal variations of rainfall obtained by rain gauges or radar were also described by the Exponential function.

6.3.1 Life Cycles of Precipitation Objects

Durations of simulated rain clusters presented some extreme value distributions which indicated that most rain events had short durations less than one hour. Events with very long durations (e.g. above 2 hours) rarely existed. It should be noted that the simulated rain clusters in this study were from rain events which occurred in summer season (May to September), affecting by the warm air mass from Atlantic Ocean and terrain, convective rain events occurs frequently in North Rhine Westphalia in such season, rain fields identified from such events usually live in a short life cycle (e.g. Berg et al., 2013; Feng et al., 2012). Averaged moving speed of the simulated light rain clusters were higher than the calculated daily averaged moving speed, which was similar to the simulated convective ones.

6.3.2 Variability for Characteristics of Precipitation Objects

Results from temporal analysis of simulated rain cluster's characteristics indicated that the size and integral rain volumes of light rain clusters presented an increasing trend as their durations extended, which was conversely for the convective cases. Areal averaged precipitation of the simulated light&convective rain clusters also presented increasing trend as durations extended, but their peak rain rates varied not so obviously. In this study, stages of precipitation object were simulated by applying the ’Normalized-Duration-Line’ method which was based on the assumption of classical life stages of thunder storms, unlike the results from Weusthoff and Hauf (2008a), it was found that the normalized duration lines of simulated rain cluster’s characteristics were found to be fitted by the parabola function, and stages of convective rainfall events followed more with ’Growth - Decay - Dissipation’ compared to the light cases through the goodness of fit testing results.

6.4 Discussion of Verification Results

6.4.1 Performance of RCIT Algorithm

As mentioned at the beginning of this chapter, two object-oriented spatial rainfall verification methods: SAL and Geometric Index were applied for testing the performance of developed precipitation identification and tracking algorithm. Through the verification results, the RCIT algorithm produced more ’flat’ light rain clusters (reflectivity $>$ 19 dBZ) with fewer precipitations compared to those which were manually labeled from radar images. Center of masses for simulated light rain clusters were more or less miss matched to those manually labeled ones. It was the same to the simulated convective rain clusters (reflectivity $>$ 37 dBZ) which had much fewer precipitations and more deviated center of masses.

It could be demonstrated that the developed algorithm performed well for representing the geometry properties of precipitation objects by comparing to those manually labeled from radar images, not only for the modeled light precipitation objects but also for the convective ones.

Through the median value of three SAL components (see table 5.4 in Chapter 5), the A component got the lowest value, which indicated that the precipitation for the simulated light&convective rain clusters were obviously under-estimated. The reason might be due to the applied median filtering algorithm that not only reduced the noisy pixels contained in radar images but also filtered some pixels with normal reflectivity.

6.4.2 Performance of QPF Methods

According to the traditional verification results (Critical Success Index, CSI), developed QPF methods - ’PIV_Semi-Lagrangian’ and ’PIV_Lagrangian-Persistence’ got high score rates in comparison to the reference methods, not only for the forecasts with low intensity, but also for the convective cases. It seemed that the developed QPF methods improved the forecasting quality of the study area.

Was it true as we expected? Through the object-oriented verification results, forecasted light rain clusters were more ’peaked’ by four QPF methods in this study, and precipitations of forecasted light rain clusters were more close to those manually labeled from radar images. Center of masses of the forecasted light rain clusters were more or less miss matched to the labeled ones. It was the same to forecasted convective rain clusters but they were more ’peaked’ with higher precipitations and obviously miss matched center of masses. When taking rainfall fields as the verification descriptors, performance of developed forecasting methods were not so outstanding compared to the referenced methods.

Median value of three SAL components for the forecasted precipitation objects (see table 5.8 in Chapter 5) indicated that the structure of the forecasted rain clusters from four forecasting methods were mostly deviated to the manually labeled ones in light rainfall forecasting cases, then was the center of mass, the best was the precipitation. In connection with such cases, center of masses of rain clusters forecasted by the developed forecasting methods were more miss matched to the labeled ones in comparison to the reference methods. For convective rainfall forecasting cases, center of masses of the forecasted rain clusters from four QPF methods were mostly miss matched to the manually labeled ones, and the forecasted rain clusters produced from developed forecasting methods were more ’peaked’ and their center of masses were more miss matched compared to the reference methods.

Variability of SAL components following different lead times (see figure 5.23 in Chapter 5) indicated that:

(a) For the short-term rainfall forecasts with low intensity, the shape and center of mass of the precipitation objects might be the key factors affecting the predictability of developed methods (PIV_Semi-Lagrangian and PIV_Lagrangian-persistence) according to the scores of three SAL components.

(b) While considering the convective rainfall forecasts, precipitation and center of mass of the precipitation objects were the main constraints for the performance of developed methods.

Additionally, through the practical application in this study, it was found that the objective verification method provided a more detail approach for analyzing the uncertainties of radar based QPF, which might provide another view for the purpose of improving the quality in short-term QPF.

6.4.3 Performance of Characteristic Predicting Methods

The proposed predicting method in this study gave a possibility for forecasting the diversity of rain event's features as Area, Cumulative rainfall, Max and Mean rain rate. The rain cluster's center of mass predicted results indicated that the Kalman filter algorithm was suitable which provided another possibility of extrapolating the center of mass of rain field (Rossi et al., 2015). However, it is emphasized that the kalman filter algorithm used in this study was only suitable for those rain cluster's trajectories which were more linearly, for complete non-linear trajectories, more measuring elements should be added into the algorithm (e.g. Abrecht et al., 2015; Jatoth et al., 2015).

6.5 Limitations and Outlooks

It was undeniable that limitations still existed in this study which can be concluded as follows:

- The RCIT algorithm developed in this study is purely depended on radar images. Sources of uncertainty based on radar rainfall estimation include spurious echoes, attenuation of the radar signal, beam blockage, anomalous propagation, variations in the vertical profile of reflectivity (VPR), conversion of radar reflectivity to precipitation rate through empirically or statistically derived Z–R relationships (where Z is the reflectivity and R is the rainfall intensity) which may not be representative for different rainfall intensities and precipitation processes (Liguori and Rico-Ramirez, 2014). Therefore, a rainfall adjustment procedure based on radar and rain gauge should be operated in the future work for improving the accuracy of rainfall estimation (Einfalt and Lobbrecht, 2012).

- For the RCIT algorithm and rainfall forecasting methods applied in this study, motion vectors of advection fields are estimated based on Particle Image velocimetry method, which is belong to a statistical approach. How reality of the estimated motion vectors are? In future work, other methods should be applied for reproducing more realistic motion vectors (e.g. Bowman et al., 2013; Korsholm et al., 2015).

- As already introduced in Chapter 1, radar based precipitation estimation and forecasting is the main purpose of this study, moreover, how to applied the results of this study into catchment rainfall-runoff simulation and flood protection is another important researching aspect which is put forward at the end of this thesis. In detail, precipitation objects have been modeled and analyzed which gave us a possibility to acquire more knowledges of inner behavior of rain events, and are the urban flood's generation directly affected by the variability of rainfall event's properties in state of North Rhine Westphalia? Factors which influenced precipitation forecast's quality have been investigated and identified, and can the forecast's quality be improved when these factors are taken into account? These are what should be careful thought in future work. Here, this thesis gave a general outlook:(a) Examinations for runoff's sensitivity to the characteristics of precipitation objects which are separately modeled from weather radar and rain gauges need to be done. (b) Uncertainties of radar based precipitation forecast needs to be modeled (e.g. Berenguer and Zawadzki, 2008; Vincendon et al., 2011), and implemented to the existing forecasting methods for improving the forecast quality.

CHAPTER 7

SUMMARY

Rainfall is not only one of the most natural processes on the earth, but also an important factor of flood generation. Quantitative precipitation estimation (QPE) and forecast (QPF) can improve our understandings about this process and its interaction with local hydrological elements, and more, give an effective warning before hazards occur. The aim of this study is the investigation of spatial-temporal variability of characteristics for rainfall events occurred in state of North Rhine Westphalia (NRW), Germany and the evaluation of radar based short term precipitation forecasting methods. For such purpose, high-resolution rainfall data sets are acquired from C-band Essen radar station belonging German Weather Service (DWD).

A precipitation objects identification and tracking algorithm- RCIT (Rain Cluster Identification and Tracking) is proposed and developed for fulfilling requirements of QPE. Spatial and temporal characteristics of rainfall events are extracted and analyzed by the proposed method. In another aspect, for the requirements of short term QPF, two advection field tracking based QPF methods: 'PIV_Semi-Lagrangian' and 'PIV_Lagrangian-Persistence' are developed wherein the past velocity fields are estimated by 'Particle Image Velocimetry' (PIV) method and the advection fields are extrapolated by Semi-Lagrange and Lagrange-Persistence schemes separately. Additionally, a predicting method for event's characteristics is proposed and a Kalman filter algorithm is also implemented for event's location forecasting. Two object-oriented spatial rainfall verification methods: SAL and Geometric Index are employed for providing detail information about the qualities of proposed algorithm and forecasting methods.

To this end, results show that the characteristics of rainfall events existed obviously spatial and temporal variability. Most rainfall events were small in size with durations of a few minutes to one hour and the internal growth and decay for their characteristics were obvious during the event's life cycles. The RCIT algorithm produced 'large' and 'flat' precipitation objects, which contained less precipitations in comparison with what were measured from radar images, and mass centers of the modeled objects were slightly deviated to the measured ones. Forecasted precipitation objects were 'small' and 'peaked' in comparison with the measured ones bay the developed forecasting methods, which contained more precipitation than the measured objects and their mass centers were more or less deviated to the measured ones. Performance of newly developed QPF methods: 'PIV_Semi-Lagrangian' and 'PIV_Lagrangian-Persistence' performed better than reference methods through the traditional verification results, nevertheless, shape and mass center of precipitation objects obviously effected the performances of developed forecasting methods for light rainfall forecasting, while precipitation object's rainfall volume and mass center were the key factors which affect the predictability of convective rainfall forecasting.

Zusammenfassung

Niederschlag ist nicht nur einer der wichtigsten natürlichsten Prozesse auf der Erde, sondern stellt darüber hinaus einen wichtigen Faktor bei der Entstehung von Hochwassern dar. Eine quantitative Niederschlagsabschätzung (QPE) und –Vorhersage (QPF) trägt nicht nur dazu bei, unser Verständnis dieses Prozesses und die Wechselwirkungen mit hydrologischen Prozessen zu verbessern, sondern liefert auch effektive Möglichkeiten zur Vorwarnung, noch bevor Katastrophen entstehen.

Ziel dieser Arbeit ist es, die räumliche und zeitliche Variabilität von Niederschlagsereignissen im Bundesland Nordrhein Westphalen (Deutschland) zu untersuchen und die radarbasierten Methoden zur Kurzzeitniederschlagsvorhersage zu bewerten. Dazu wurden hochaufgelöste Niederschlagsdaten der C-Band Radarstation Essen des Deutschen Wetterdienstes (DWD) verwendet.

Ziel ist es weiterhin, einen Algorithmus zur Identifikation und für das Tracking von Niederschlagsobjekten (RCIT, Rain Cluster Identification and Tracking) zu entwickeln um die Anforderungen der quantitativen Niederschlagsschätzung zu erfüllen. Mit dieser Methode wurden räumliche und zeitliche Charakteristika der Niederschlagsereignisse ermittelt und analysiert. Um den Anforderungen der quantitativen Niederschlagsvorhersage gerecht zu werden wurden darüber hinaus zwei Vorhersagemethoden entwickelt, das „PIV_Semi-Langrangian“ und das „PIV_Langrangian Persistence“ Verfahren. Die Geschwindigkeitsfelder werden dabei über „Particle Image Velocimetry“ (PIV) bestimmt, während die Advektion getrennt über Semi-Lagrange und Lagrange Persistence Verfahren extrapoliert werden.

Weiterhin wurde eine Vorhersagemethode für die Charakteristik der Niederschlagsereignisse entwickelt sowie ein Kalman-Filteralgorithmus zur Vorhersage der Lokalisierung dieser

Ereignisse angewandt. Um detaillierte Informationen zur Qualität des Algorithmus und der Vorhersagemethode zu gewinnen, wurden zwei objekt-orientierte räumliche Niederschlagsverifizierungsmethoden genutzt: SAL und Geometric Index.

Die Ergebnisse zeigen, dass die Eigenschaften der Niederschlagsereignisse offensichtlich räumlich und zeitlich variabel sind. Die meisten Niederschlagsereignisse sind von geringer Größe und dauern nur wenige Minuten bis eine Stunde an. Das interne Wachstum wie auch dessen Zerfall sind während des Event-Lebenszyklus deutlich erkennbar. Der RCIT Algorithmus produziert große und vertikal flache Niederschlagsobjekte, welche im Vergleich zu den Radarmessungen weniger Niederschlag enthalten und auch die Massenschwerpunkte der modellierten Objekte weichen leicht von den Messungen ab. Die vorhergesagten Niederschlagsobjekte sind im Vergleich mit den Messungen kleiner und vertikal spitz zulaufend, enthalten mehr Niederschlag und die Schwerpunkte weichen mehr oder weniger stark von den Messungen ab.

Die neu entwickelten QPF Methoden „PIV_Semi-Lagrangian“ und „PIV_Lagrangian-Persistence“ zeigen eine im Vergleich mit den Referenzmethoden und deren Verifizierungsverfahren eine verbesserte Leistung. Für leichte Niederschlagsereignisse sind der gemittelte Niederschlag und der Massenschwerpunkt des Niederschlagsobjektes die beiden Faktoren, die die Qualität der entwickelten Vorhersagemethode beeinflussen, während für konvektive Ereignisse das Niederschlagsvolumen und der Massenschwerpunkt die Schlüsselfaktor darstellen.

REFERENCE AND ON-LINE SOURCES

Abrecht, D. G., Schwantes, J. M., Kukkadapu, R. K., McDonald, B. S., Eiden, G. C., and Sweet, L. E. (2015). Real-time noise reduction for mössbauer spectroscopy through online implementation of a modified kalman filter. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 773:66–71.

Aghakouchak, A., Nasrollahi, N., Li, J., Imam, B., and Sorooshian, S. (2011). Geometrical characterization of precipitation patterns. *Journal of Hydrometeorology*, 12(2):274–285.

Akaike, H. (1998). Information theory and an extension of the maximum likelihood principle. In *Selected Papers of Hirotugu Akaike*, pages 199–213. Springer.

Austin, G. and Bellon, A. (1974). The use of digital weather radar records for short-term precipitation forecasting. *Quarterly Journal of the Royal Meteorological Society*, 100(426):658–664.

Austin, G. and Bellon, A. (1982). Very-short-range forecasting of precipitation by the objective extrapolation of radar and satellite data. *Nowcasting, ed. KA Browning, Academic Press, New York.*

Bar-Shalom, Y., Li, X. R., and Kirubarajan, T. (2004). *Estimation with applications to tracking and navigation: theory algorithms and software.* John Wiley & Sons.

Bárdossy, A. and Das, T. (2008). Influence of rainfall observation network on model calibration and application. *Hydrology and Earth System Sciences Discussions*, 12(1):77–89.

Barnolas, M., Rigo, T., and Llasat, M. C. (2010). Characteristics of 2-d convective structures in catalonia (ne spain): an analysis using radar data and gis. *Hydrology and Earth System Sciences*, 14(1):129–139.

Bartels, H., Dietzer, B., Malitz, G., Albrecht, F., and Guttenberger, J. (2005). Kostra-dwd 2000 - extreme heavy precipitation for germany (1951-2000) - update report. *Offenbach: German Weather Service, Hydrometeorology*.

Bellon, A. and Austin, G. (1984). The accuracy of short-term radar rainfall forecasts. *Journal of hydrology*, 70(1):35–49.

Berenguer, M., Sempere-Torres, D., and Pegram, G. G. (2011). Sbmcast–an ensemble nowcasting technique to assess the uncertainty in rainfall forecasts by lagrangian extrapolation. *Journal of Hydrology*, 404(3):226–240.

Berenguer, M. and Zawadzki, I. (2008). A study of the error covariance matrix of radar rainfall estimates in stratiform rain. *Weather and Forecasting*, 23(6):1085–1101.

Berg, P., Moseley, C., and Haerter, J. O. (2013). Strong increase in convective precipitation in response to higher temperatures. *Nature Geoscience*, 6(3):181–185.

Bluestein, H. B. (1992). *Synoptic-dynamic Meteorology in Midlatitudes: Observations and theory of weather systems*, volume 2. Taylor & Francis.

Bluestein, H. B., McCaul Jr, E. W., Byrd, G. P., Walko, R. L., and Davies-Jones, R. (1990). An observational study of splitting convective clouds. *Monthly weather review*, 118(6):1359–1370.

Bowler, N. E., Pierce, C. E., and Seed, A. (2004). Development of a precipitation nowcasting algorithm based upon optical flow techniques. *Journal of Hydrology*, 288(1):74–91.

Bowler, N. E., Pierce, C. E., and Seed, A. W. (2006). Steps: A probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled nwp. *Quarterly Journal of the Royal Meteorological Society*, 132(620):2127–2155.

Bowman, K. P., Lin, J. C., Stohl, A., Draxler, R., Konopka, P., Andrews, A., and Brunner, D. (2013). Input data requirements for lagrangian trajectory models. *Bulletin of the American Meteorological Society*, 94(7):1051–1058.

Bruni, G., Reinoso, R., van de Giesen, N. C., Clemens, F. H. L. R., and ten Veldhuis, J. A. E. (2015). On the sensitivity of urban hydrodynamic modelling to rainfall spatial and temporal resolution. *Hydrology and Earth System Sciences*, 19(2):691–709.

Byers, H. R. and Braham Jr, R. R. (1948). Thunderstorm structure and circulation. *Journal of Meteorology*, 5(3):71–86.

Capsoni, C., D'Amico, M., and Locatelli, P. (2008). Statistical properties of rain cells in the padana valley. *Journal of Atmospheric and Oceanic Technology*, 25(12):2230–2244.

Casati, B., Wilson, L., Stephenson, D., Nurmi, P., Ghelli, A., Pocernich, M., Damrath, U., Ebert, E., Brown, B., and Mason, S. (2008). Forecast verification: current status and future directions. *Meteorological applications*, 15(1):3–18.

Chanda, B. and Majumder, D. D. (2004). *Digital image processing and analysis*. PHI Learning Pvt. Ltd.

Chiang, Y.-M., Chang, F.-J., Jou, B. J.-D., and Lin, P.-F. (2007). Dynamic ann for precipitation estimation and forecasting from radar observations. *Journal of Hydrology*, 334(1):250–261.

Ciach, G. J. and Krajewski, W. F. (2006). Analysis and modeling of spatial correlation structure in small-scale rainfall in central oklahoma. *Advances in Water Resources*, 29(10):1450 – 1463.

Collier, C. and Hardaker, P. (1996). Estimating probable maximum precipitation using a storm model approach. *Journal of hydrology*, 183(3):277–306.

Cooper, M. and Fernando, A. (2009). The effect of the raingauge distribution on stormwater models. In *18th World IMACS Congress and MODSIM09 International Congress on Modelling and Simulation*, pages 3500–3506.

Courant, R., Isaacson, E., and Rees, M. (1952). On the solution of nonlinear hyperbolic differential equations by finite differences. *Communications on Pure and Applied Mathematics*, 5(3):243–255.

Davis, C., Brown, B., and Bullock, R. (2006). Object-based verification of precipitation forecasts. part i: Methodology and application to mesoscale rain areas. *Monthly Weather Review*, 134(7):1772–1784.

Denoeux, T. and Rizand, P. (1995). Analysis of radar images for rainfall forecasting using neural networks. *Neural Computing & Applications*, 3(1):50–61.

Dixon, M. and Wiener, G. (1993). Titan: Thunderstorm identification, tracking, analysis, and nowcasting-a radar-based methodology. *Journal of Atmospheric and Oceanic Technology*, 10(6):785–797.

Ebert, E. E. (2008). Fuzzy verification of high-resolution gridded forecasts: a review and proposed framework. *Meteorological applications*, 15(1):51–64.

Einfalt, T., Arnbjerg-Nielsen, K., Golz, C., Jensen, N.-E., Quirmbach, M., Vaes, G., and Vieux, B. (2004). Towards a roadmap for use of radar rainfall data in urban drainage. *Journal of Hydrology*, 299(3):186–202.

Einfalt, T., Denoeux, T., and Jacquet, G. (1990). A radar rainfall forecasting method designed for hydrological purposes. *Journal of Hydrology*, 114(3):229–244.

Einfalt, T., Hatzfeld, F., Wagner, A., Seltmann, J., Castro, D., and Frerichs, S. (2009). Urbas: forecasting and management of flash floods in urban areas. *Urban Water Journal*, 6(5):369–374.

Einfalt, T. and Lobbrecht, A. (2012). Compositing international radar data using a weight-based scheme. *IAHS-AISH publication*, pages 20–25.

Einfalt, T., Scheibel, M., and Mittelstädt, R. (2008). Analysis of a damage producing flash flood in the wupper area. *11th ICUD*.

Ekstrom, M. P. (1984). *Digital Image Processing Techniques*. Academic Press, Inc., Orlando, FL, USA.

European Environment Agency (EFA) (2015). Nationally designated areas (CDDA). http://www.eea.europa.eu/data-and-maps/data/nationally-designated-areas-national-cdda-10/. Online; accessed 23 November 2015.

Feng, Z., Dong, X., Xi, B., McFarlane, S. A., Kennedy, A., Lin, B., and Minnis, P. (2012). Life cycle of midlatitude deep convective systems in a lagrangian framework. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 117(D23).

FjØrtoft, R. (1952). On a numerical method of integrating the barotropic vorticity equation. *Tellus*, 4(3):179–194.

Fuyou, T., minghu, C., yaping, Z., and yanfei, Y. (2010). An investigation into the effect of rain gauge density on estimating the areal rainfall using a radar gauge calibration algorithm. *Acta Meteorologica Sinica*, 68(5):717–730.

Georgakakos, K. (1986). On the design of national, real-time warning systems with capability for site-specific, flash-flood forecasts. *Bulletin of the American Meteorological Society*, 67(10):1233–1239.

German weather service (DWD) (2006). KOSTRA-DWD 2000 – Coordinated heavy precipitation regionalization evaluations. –Offenbach: self-publishing DWD.

German weather service (DWD) (2014). Climate Data Center (CDC) of the DWD. ftp://ftp-cdc.dwd.de/pub/CDC/. Online; accessed 01 July 2014.

Germann, U., Berenguer, M., Sempere-Torres, D., and Zappa, M. (2009). Real—ensemble radar precipitation estimation for hydrology in a mountainous region. *Quarterly Journal of the Royal Meteorological Society*, 135(639):445–456.

Germann, U. and Zawadzki, I. (2002). Scale-dependence of the predictability of precipitation from continental radar images. part i: Description of the methodology. *Monthly Weather Review*, 130(12):2859–2873.

Germann, U., Zawadzki, I., and Turner, B. (2006). Predictability of precipitation from continental radar images. part iv: Limits to prediction. *Journal of the atmospheric sciences*, 63(8):2092–2108.

Gilleland, E. (2013). Testing competing precipitation forecasts accurately and efficiently: the spatial prediction comparison test. *Monthly Weather Review*, 141(1):340–355.

Gilleland, E., Ahijevych, D., Brown, B. G., Casati, B., and Ebert, E. E. (2009). Intercomparison of spatial forecast verification methods. *Weather and Forecasting*, 24(5):1416–1430.

Gires, A., Onof, C., Maksimovic, C., Schertzer, D., Tchiguirinskaia, I., and Simoes, N. (2012). Quantifying the impact of small scale unmeasured rainfall variability on urban runoff through multifractal downscaling: A case study. *Journal of Hydrology*, 442–443:117 – 128.

Golding, B. (1998). Nimrod: A system for generating automated very short range forecasts. *Meteorological Applications*, 5(01):1–16.

Grecu, M. and Krajewski, W. (2000). A large-sample investigation of statistical procedures for radar-based short-term quantitative precipitation forecasting. *Journal of hydrology*, 239(1):69–84.

Haile, A. T., Rientjes, T., Gieske, A., and Gebremichael, M. (2009). Rainfall variability over mountainous and adjacent lake areas: The case of lake tana basin at the source of the blue nile river. *Journal of Applied Meteorology and Climatology*, 48(8):1696–1717.

Han, L., Fu, S., Zhao, L., Zheng, Y., Wang, H., and Lin, Y. (2009). 3d convective storm identification, tracking, and forecasting-an enhanced titan algorithm. *Journal of Atmospheric and Oceanic Technology*, 26(4):719–732.

Handwerker, J. (2002). Cell tracking with trace3d—a new algorithm. *Atmospheric Research*, 61(1):15–34.

Heistermann, M., Jacobi, S., and Pfaff, T. (2013). Technical note: An open source library for processing weather radar data (wradlib). *Hydrology and Earth System Sciences*, 17(2):863–871.

Jahanmiri, M. (2011). Particle image velocimetry: Fundamentals and its applications. Technical report, Department of Applied Mechanics. 58.

Jatoth, R. K., Gopisetty, S., and Hussain, M. (2015). Performance analysis of alpha beta filter, kalman filter and meanshift for object tracking in video sequences. *International Journal of Image, Graphics and Signal Processing (IJIGSP)*, 7(3):24.

Jewett, B. F. and Wilhelmson, R. B. (2006). The role of forcing in cell morphology and evolution within midlatitude squall lines. *Monthly weather review*, 134(12):3714–3734.

Johnson, J., MacKeen, P. L., Witt, A., Mitchell, E. D. W., Stumpf, G. J., Eilts, M. D., and Thomas, K. W. (1998). The storm cell identification and tracking algorithm: An enhanced wsr-88d algorithm. *Weather and Forecasting*, 13(2):263–276.

Karklinsky, M. and Morin, E. (2006). Spatial characteristics of radar-derived convective rain cells over southern israel. *Meteorologische Zeitschrift*, 15(5):513–520.

Korsholm, U. S., Petersen, C., Sass, B. H., Nielsen, N. W., Jensen, D. G., Olsen, B. T., Gill, R., and Vedel, H. (2015). A new approach for assimilation of 2d radar precipitation in a high-resolution nwp model. *Meteorological Applications*, 22(1):48–59.

Krajewski, W. F., Ciach, G. J., McCollum, J. R., and Bacotiu, C. (2000). Initial validation of the global precipitation climatology project monthly rainfall over the united states. *Journal of Applied Meteorology*, 39(7):1071–1086.

Kronenberg, R., Franke, J., and Bernhofer, C. (2012). Classification of daily precipitation patterns on the basis of radar-derived precipitation rates for saxony, germany. *Meteorologische Zeitschrift*, 21(5):475–486.

Kropp, J., Block, A., Reusswig, F., Zickfeld, K., and Schellnhuber, H. (2006). Semiquantitative assessment of regional climate vulnerability: The north-rhine westphalia study. *Climatic Change*, 76(3-4):265–290.

Lang, P. (2001). Cell tracking and warning indicators derived from operational radar products. In *Proc., 30th Int. Conf. on Radar Meteorology, Munich, Germany, Amer. Meteor. Soc*, pages 245–247.

Lee, B. D., Jewett, B. F., and Wilhelmson, R. B. (2006a). The 19 april 1996 illinois tornado outbreak. part i: Cell evolution and supercell isolation. *Weather and forecasting*, 21(4):433–448.

Lee, B. D., Jewett, B. F., and Wilhelmson, R. B. (2006b). The 19 april 1996 illinois tornado outbreak. part ii: Cell mergers and associated tornado incidence. *Weather and forecasting*, 21(4):449–464.

Lempio, G., Einfalt, T., Lobbrecht, A., and Lempio, G. (2012). Considerations for compositing radar data from three countries. In *ERAD 2012—The Seventh European Conference on Radar in Meteorology and Hydrology. Toulouse, France, 24–29 June*.

Li, H., Corzo Perez, G. A., Martinez, C. A., and Mynett, A. E. (2013). Self-learning cellular automata for forecasting precipitation from radar images. *Journal of Hydrologic Engineering*, 18(2):206–211.

Li, L., Schmid, W., and Joss, J. (1995). Nowcasting of motion and growth of precipitation with radar over a complex orography. *Journal of applied meteorology*, 34(6):1286–1300.

Li, P.-W., Wong, W.-K., Cheung, P., and Yeung, H.-Y. (2014). An overview of nowcasting development, applications, and services in the hong kong observatory. *Journal of Meteorological Research*, 28:859–876.

Liang, Q., Feng, Y., Deng, W., Hu, S., Huang, Y., Zeng, Q., and Chen, Z. (2010). A composite approach of radar echo extrapolation based on trec vectors in combination with model-predicted winds. *Advances in Atmospheric Sciences*, 27:1119–1130.

Liang, X., Jianzhong, G., and Leung, L. R. (2004). Assessment of the effects of spatial resolutions on daily water flux simulations. *Journal of Hydrology*, 298(1):287–310.

Liguori, S. and Rico-Ramirez, M. A. (2014). A review of current approaches to radar-based quantitative precipitation forecasts. *International Journal of River Basin Management*, 12(4):391–402.

Liu, J., Bray, M., and Han, D. (2013). A study on wrf radar data assimilation for hydrological rainfall prediction. *Hydrology and Earth System Sciences*, 17(8):3095–3110.

Lopez, V., Napolitano, F., and Russo, F. (2005). Calibration of a rainfall-runoff model using radar and raingauge data. *Advances in Geosciences*, 2(2):41–46.

Luini, L. and Capsoni, C. (2012). The impact of space and time averaging on the spatial correlation of rainfall. *Radio Science*, 47(3).

Mácaa, P. and Torfs, P. (2009). The influence of temporal rainfall distribution in the flood runoff modelling. *Soil and Water Research*, 4(Special Issue 2):S102–S110.

Mandapaka, P. V., Germann, U., Panziera, L., and Hering, A. (2012). Can lagrangian extrapolation of radar fields be used for precipitation nowcasting over complex alpine orography? *Weather and Forecasting*, 27(1):28–49.

Marchi, L., Borga, M., Preciso, E., and Gaume, E. (2010). Characterisation of selected extreme flash floods in europe and implications for flood risk management. *Journal of Hydrology*, 394(1–2):118 – 133.

Massey Jr, F. J. (1951). The kolmogorov-smirnov test for goodness of fit. *Journal of the American statistical Association*, 46(253):68–78.

Mecklenburg, S., Joss, J., and Schmid, W. (2000). Improving the nowcasting of precipitation in an alpine region with an enhanced radar echo tracking algorithm. *Journal of Hydrology*, 239(1):46–68.

Meixner, T. (2002). Spatial patterns in catchment hydrology observations and modelling. *Vadose Zone Journal*, 1(1):202–203.

Méndez-Antonio, B., Caetano, E., Rivera-Trejo, F. G., Rodríguez, R. A. C., Watts, C., and Soto-Cortés, G. (2013). Weather radar data and distributed hydrological modelling: an application for mexico valley. *Open Journal of Modern Hydrology*, 3(02):79.

Meyer, V., Höller, H., and Betz, H. (2013). Automated thunderstorm tracking: utilization of three-dimensional lightning and radar data. *Atmospheric Chemistry and Physics*, 13(10):5137–5150.

Milan, M., Schüttemeyer, D., Venema, V., and Simmer, C. (2009). Assimilation of radar precipitation and satellite data into a nwp model using a physical initialisation scheme. In *EGU General Assembly Conference Abstracts*, volume 11, page 4033.

Mishra, A. (2013). Effect of rain gauge density over the accuracy of rainfall: a case study over bangalore, india. *SpringerPlus*, 2(1).

Momjian, B. (2001). *PostgreSQL: introduction and concepts*, volume 192. Addison-Wesley New York.

Morin, E., Goodrich, D. C., Maddox, R. A., Gao, X., Gupta, H. V., and Sorooshian, S. (2006). Spatial patterns in thunderstorm rainfall events and their coupling with watershed hydrological response. *Advances in Water Resources*, 29(6):843–860.

Moseley, C., Berg, P., and Haerter, J. O. (2013). Probing the precipitation life cycle by iterative rain cell tracking. *Journal of Geophysical Research: Atmospheres*, 118(24):13–361.

Nakakita, E., Ikebuchi, S., Nakamura, T., Kanmuri, M., Okuda, M., Yamaji, A., and Takasao, T. (1996). Short-term rainfall prediction method using a volume scanning radar and grid point value data from numerical weather prediction. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 101(D21):26181–26197.

Nisi, L., Ambrosetti, P., and Clementi, L. (2014). Nowcasting severe convection in the alpine region: the coalition approach. *Quarterly Journal of the Royal Meteorological Society*, 140(682):1684–1699.

Olsen, R., Rogers, D. V., and Hodge, D. B. (1978). The ar b relation in the calculation of rain attenuation. *Antennas and Propagation, IEEE Transactions on*, 26(2):318–329.

Peleg, N., Ben-Asher, M., and Morin, E. (2013). Radar subpixel-scale rainfall variability and uncertainty: lessons learned from observations of a dense rain-gauge network. *Hydrol. Earth Syst. Sci*, 17(6):2195–2208.

Pierce, C., Ebert, E., Seed, A., Sleigh, M., Collier, C., Fox, N., Donaldson, N., Wilson, J., Roberts, R., and Mueller, C. (2004). The nowcasting of precipitation during sydney 2000: an appraisal of the qpf algorithms. *Weather and Forecasting*, 19(1):7–21.

Quirmbach, M., Einfalt, T., and Langstädtler, G. (2012). Climate change analysis of precipitation data for north rhine-westphalia. *Atmospheric Research*, 109–110(0):1 – 13.

Radhakrishna, B., Zawadzki, I., and Fabry, F. (2012). Predictability of precipitation from continental radar images. part v: Growth and decay. *Journal of the Atmospheric Sciences*, 69(11):3336–3349.

Rinehart, R. E. (1990). *Radar for Meteorologists: You, Too Can be a Radar Meteorologist*. Department of Atmospheric Sciences, University of North Dakota.

Rinehart, R. E. and Garvey, E. T. (1978). Three-dimensional storm motion detection by conventional weather radar. *Nature*, 273(5660):287–289.

Rossa, A., Liechti, K., Zappa, M., Bruen, M., Germann, U., Haase, G., Keil, C., and Krahe, P. (2011). The cost 731 action: A review on uncertainty propagation in advanced hydro-meteorological forecast systems. *Atmospheric Research*, 100(2):150–167.

Rossi, P. J., Chandrasekar, V., Hasu, V., and Moisseev, D. (2015). Kalman filtering–based probabilistic nowcasting of object-oriented tracked convective storms. *Journal of Atmospheric and Oceanic Technology*, 32(3):461–477.

Rousseeuw, P. J. and Leroy, A. M. (2005). *Robust regression and outlier detection*, volume 589. John Wiley & Sons.

Ruzanski, E. and Chandrasekar, V. (2012). An investigation of the short-term predictability of precipitation using high-resolution composite radar observations. *Journal of Applied Meteorology and Climatology*, 51(5):912–925.

Schlatter, T. W. (2000). Variational assimilation of meteorological observations in the lower atmosphere: a tutorial on how it works. *Journal of atmospheric and solar-terrestrial physics*, 62(12):1057–1070.

Schneider, S., Wang, Y., Wagner, W., and Mahfouf, J.-F. (2014). Impact of ascat soil moisture assimilation on regional precipitation forecasts: A case study for austria. *Monthly Weather Review*, 142(4):1525–1541.

Schoorens, J., Emmanuel, I., LE Balier, V., Guillon, A., and Elahaut, B. (2014). Influence of rainfall spatial variability on hydrological modelling: case study on two urban catchments. In *ERAD2014-8th European Conference on Radar in Meteorology and Hydrology*.

Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics*, 6(2):461–464.

Seed, A. (2003). A dynamic and spatial scaling approach to advection forecasting. *Journal of Applied Meteorology*, 42(3):381–388.

Shimizu, S. and Uyeda, H. (2012). Algorithm for the identification and tracking of convective cells based on constant and adaptive threshold methods using a new cell-merging and-splitting scheme. *Journal of the Meteorological Society of Japan. Ser. II*, 90(6):869–889.

Singh, V. P. (1997). Effect of spatial and temporal variability in rainfall and watershed characteristics on stream flow hydrograph. *Hydrological Processes*, 11(12):1649–1669.

Singh, V. P. and Woolhiser, D. A. (2002). Mathematical modeling of watershed hydrology. *Journal of hydrologic engineering*, 7(4):270–292.

Smith, M. B., Koren, V. I., Zhang, Z., Reed, S. M., Pan, J.-J., and Moreda, F. (2004). Runoff response to spatial variability in precipitation: an analysis of observed data. *Journal of hydrology*, 298(1):267–286.

Sokol, Z. (2006). Nowcasting of 1-h precipitation using radar and nwp data. *Journal of Hydrology*, 328(1):200–211.

State Office for Nature, Environment and Consumer Protection of North Rhine Westphalia (2010). Climate and Atmosphere. http://www.lanuv.nrw.de/kfm-indikatoren/index.php?mode=gruppe&aufzu=1. Online; accessed 16 March 2010.

State Office for Nature, Environment and Consumer Protection of North Rhine Westphalia (2015). Environmental data. http://www.naturschutz-fachinformationssysteme-nrw.de/natura2000-meldedok/de/downloads.

Tessendorf, A. and Einfalt, T. (2012). Ensemble radar nowcasts-a multi-method approach. *IAHS-AISH publication*, pages 311–316.

Tüchler, L. and Meyer, V. (2013). Systematic investigations of intense convective precipitation events on european scale based on radar-and lightning-cell tracking. In *EGU General Assembly Conference Abstracts*, volume 15, page 9155.

Tuttle, J. D. and Foote, G. B. (1990). Determination of the boundary layer airflow from a single doppler radar. *Journal of Atmospheric and oceanic Technology*, 7(2):218–232.

Vincendon, B., Ducrocq, V., Nuissier, O., and Vié, B. (2011). Perturbation of convection-permitting nwp forecasts for flash-flood ensemble forecasting. *Natural Hazards and Earth System Science*, 11(5):1529–1544.

Vischel, T. and Lebel, T. (2007). Assessing the water balance in the sahel: Impact of small scale rainfall variability on runoff. part 2: Idealized modeling of runoff sensitivity. *Journal of hydrology*, 333(2):340–355.

Wang, G., Liu, L., and Ruan, Z. (2007). Application of doppler radar data to nowcasting of heavy rainfall. *Journal of Applied Meteorological Science*, 18(2):388–395.

Wang, G., Wong, W., Liu, L., and Wang, H. (2013). Application of multi-scale tracking radar echoes scheme in quantitative precipitation nowcasting. *Advances in Atmospheric Sciences*, 30(2):448–460.

Waymire, E. and Gupta, V. K. (1981a). The mathematical structure of rainfall representations: 1. a review of the stochastic rainfall models. *Water Resources Research*, 17(5):1261–1272.

Waymire, E. and Gupta, V. K. (1981b). The mathematical structure of rainfall representations: 2. a review of the theory of point processes. *Water Resources Research*, 17(5):1273–1285.

Weather in National Weather Service (NWS) (2014). JetStream - Online School for Weather in National Weather Service (NWS). http://www.srh.weather.gov/jetstream/index.htm. Online; accessed 20 June 2014.

Wernli, H., Paulat, M., Hagen, M., and Frei, C. (2008). Sal-a novel quality measure for the verification of quantitative precipitation forecasts. *Monthly Weather Review*, 136(11):4470–4487.

Westcott, N. E. (1994). Merging of convective clouds: Cloud initiation, bridging, and subsequent growth. *Monthly weather review*, 122(5):780–790.

Weusthoff, T. and Hauf, T. (2008a). Basic characteristics of post-frontal shower precipitation rates. *Meteorologische Zeitschrift*, 17(6):793–805.

Weusthoff, T. and Hauf, T. (2008b). The life cycle of convective-shower cells under post-frontal conditions. *Quarterly Journal of the Royal Meteorological Society*, 134(633):841–857.

Wiin-Nielsen, A. (1959). On the application of trajectory methods in numerical forecasting. *Tellus*, 11(2):180–196.

Wood-Bradley, P., Zapata, J., and Pye, J. (2012). „cloud tracking with optical flow for short-term solar forecasting “. *Solar Thermal Group, Australian National University, Canberra, Australija.*

Wright, D., Smith, J., Villarini, G., and Baeck, M. (2013). Estimating the frequency of extreme rainfall using weather radar and stochastic storm transposition. *Journal of Hydrology*, 488:150 – 165.

Wuensch, K. (2011). Chi-square tests. In Lovric, M., editor, *International Encyclopedia of Statistical Science*, pages 252–253. Springer Berlin Heidelberg.

Yakir, H. and Morin, E. (2011). Hydrologic response of a semi-arid watershed to spatial and temporal characteristics of convective rain cells. *Hydrology and Earth System Sciences*, 15(1):393–404.

Zacharov, P., Rezacova, D., and Brozkova, R. (2013). Evaluation of the qpf of convective flash flood rainfalls over the czech territory in 2009. *Atmospheric Research*, 131:95–107.

Zahraei, A., Hsu, K.-l., Sorooshian, S., Gourley, J., Lakshmanan, V., Hong, Y., and Bellerby, T. (2012). Quantitative precipitation nowcasting: a lagrangian pixel-based approach. *Atmospheric Research*, 118:418–434.

Zahraei, A., Hsu, K.-l., Sorooshian, S., Gourley, J. J., Hong, Y., and Behrangi, A. (2013). Short-term quantitative precipitation forecasting using an object-based approach. *Journal of Hydrology*, 483:1–15.

Zappa, M., Beven, K. J., Bruen, M., Cofino, A. S., Kok, K., Martin, E., Nurmi, P., Orfila, B., Roulin, E., Schröter, K., et al. (2010). Propagation of uncertainty from observing systems and nwp into hydrological models: Cost-731 working group 2. *Atmospheric Science Letters*, 11(2):83–91.

Zimmer, M. and Wernli, H. (2011). Verification of quantitative precipitation forecasts on short time-scales: A fuzzy approach to handle timing errors with sal. *Meteorologische Zeitschrift*, 20(2):95–105.

ACKNOWLEDGEMENTS

I would like to express my special appreciation and thanks to my two PhD advisors: Professor Dr. Achim Schulte and Dr. Thomas Einfalt, you have been a tremendous mentors for me. I would like to thank you for encouraging my research and allowing me to grow as a research scientist. Your advice on my research have been invaluable. Another thank is to Professor Dr. Karl Tilman Rost: the second reviewer of this dissertation, thank you very much for taking your valuable time in appraising this thesis. Your comments are great helpful for my further research. Here, a Chinese phrase is quoted to thank all of you again for your guidence - '*The Higher You Look Up To It, The Harder You Studied It*'.

I would also like to thank Dr. Christian. Reinhardt-Imjela for your disinterested help about my study, I still remember your words when I started my PhD study: 'Scientific research is like of building a house, the more you put in it, the more it is rugged and beautiful'. I will still hold it in my future study and work.

Thanks to my colleges in this work group: Dr. Kai Hartman, Dr. QingBo Cheng, Msc. Venna Purv Purohit, Dr. Jens Bölsher and my Chinese colleges: Dr. Jialiang Cai, Dr. Zhili Zhao, Dr. Yubin Zhao, Dr. JiZhou Zhao, Msc. Dongwei Wang, you gave me lots of help and courages in my research and life. Thanks to Msc. Alrun Jasper-Tönnies from Hydro & meto Gmbh & Co, KG, thank you for your help in radar data analysis and suggestions in my research. Additionally tanks to Freie Universität Berlin and Chinese Scholarship Council for providing me the opportunity and foundations of studying in Germany. The five year's study and life in this beautiful country will be the most unforgettable experience in my life.

A special thanks to my family. Words can not express how grateful I am to my dear father and mother for all of the sacrifices that you've made on me. Your prayer and confidence for me was what sustained me thus far. Another thanks to Dr. Barbara Demandt for your touching care and help for my life in Berlin.

APPENDIX A

ILLUSTRATION OF RAIN CLUSTER MATCHING RULE

Rain cluster matching rule is defined in this study for searching most matched child rain clusters (see section 4.3 in Chapter 4). The detail procedure is illustrated as following steps:

a. Global motion vectors V_t of each radar image are derived at every five minutes by the PIV method, and their mean value V_{mean} is calculated as in equation (A.1).

$$\mathbf{V_{mean}} = \frac{\sum\limits_{\mathbf{i=1}}^{\mathbf{n}} \mathbf{V_t(i)}}{\mathbf{n}} \tag{A.1}$$

where n is number of grids divided from PIV procedure.

b. Rain clusters are identified (section 4.3 in Chapter 4) and their characteristics are derived: Area (A), Cumulative rainfall (I_{total}), Max rain rate per 5 minutes (I_{peak}) and Mean rain rate per 5 minutes (I_{mean}).

c. For each rain cluster at time t (R_t), a boundary box is defined with a horizontal range of $[x-d, y+d]$ and a vertical range of $[y-d, y+d]$, where d is a predefined range, in this study d is 8 km.

d. For rain clusters at time $t+\Delta t$ ($R_{t+\Delta t}$), searching the ones which are intersected with boundary box defined in step c.

e. For each $R_{t+\Delta t}$ fallen into the boundary box, vector (v_x, v_y) between its weighted center and center of mass of R_t are calculated. Meanwhile, the overlap (C)between each $R_{t+\Delta t}$ and $R_{t+\Delta t}$.

f. Three cases are considered in this step:

(1) If there is only one $R_{t+\Delta t}$ in the boundary box, then if $C > 0$ then $R_{t+\Delta t}$ is the most matched child cluster. If $C = 0$ but length of (v_x, v_y) $<$ length of 3 $\times$ length of V_{mean} and angle of (v_x, v_y) minus the angle of $V_{mean} < 45^o$, then

$R_{t+\Delta t}$ is the most matched child rain cluster.

(2) If there are at least two $R_{t+\Delta t}$, for each one, if $C > 0.2$, then it is the most matched child rain cluster, if $C < 0$, the conditions is same with case one but the angle difference is less than 20^o; if above conditions are not fulfilled, then the ones which have less difference in A are the most matched child rain clusters.

The child rain cluster matching rule developed here is similar with the rule that proposed by Weusthoff and Hauf (2008b), but is more adapt to local conditions, particularly, in this study, the global motion vector calculating results by the PIV method keeps persisting feature of wind fields, which is more robust than results calculated by maximum correlating way.

APPENDIX B

MULTI GOODNESS OF FIT TEST

The idea of Goodness of Fit test (GOF) is to test whether a data set is well fitted with a predefined distribution which gives the highest probability of producing the observed data. Based on such idea, series GOF test method were developed with commonly applied ones are:

$a.$ **Chi-Square test** (χ^2 test) (Wuensch, 2011), which is used to test if the samples of data comes from a population with a specific distribution. It can be applied to any univariate distribution whose cumulative distribution function can be calculated. The Chi-Square test is defined for the hypothesis - H_0: The data follow a specified distribution; H_1: The data do not follow the specified distribution. Test static is operated on data which is pre-divided into k bins and can be described by equation (B.1):

$$\mathbf{x^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}} \tag{B.1}$$

where O_i is the observed frequency for bin i and E_i is the expected frequency for bin i. The expected frequency is calculated by:

$$\mathbf{E_i = N(F(Y_u) - F(Y_l))} \tag{B.2}$$

In equation (B.2), F is the cumulative distribution function for the distribution being tested, Y_u is the upper limit for class i, Y_l is the lower limit for class i, and N is the sample size.

The test statistic follows, approximately, a chi-square distribution with k-c degrees of freedom where k is the number of non-empty cells and c is the number of estimated parameters for the distribution by added one. For example, for a three parameter Weibull distribution, c value is 4. Therefore, the hypothesis that the samples of data are from a population with the specified distribution is rejected if $x^2 > x^2_{1-\alpha,k-c}$, whereas $x^2_{1-\alpha,k-c}$ is the Chi-Square critical value with k - c degrees of freedom and significance level a.

$b.$ **Kolmogorov-Smirnov test** (**K-S** test) (Massey Jr, 1951), this type test method is based on the empirical distribution function (ECDF). Given N ordered data points $Y_1, Y_2, ..., Y_n$, the ECDF is defined as in equation (B.3):

$$\mathbf{E_N} = \frac{\mathbf{n(i)}}{\mathbf{N}} \tag{B.3}$$

where n(i) is the number of points less than Y_i and Y_i are ordered from smallest to largest value. This is a step function that increases by $1/N$ at the value of each ordered data point. Same with Chi-Square test, K-S test is based on the hypothesis: H_0: the data follow a specified distribution; H_1: the data do not follow the specified distribution. The statistic test of it can be described as in equation (B.4):

$$\mathbf{D} = \max_{\mathbf{1 \leq i \leq N}} (\mathbf{F(Y_i)} - \frac{\mathbf{i-1}}{\mathbf{N}}, \frac{\mathbf{i}}{\mathbf{N}} - \mathbf{F(Y_i)}) \tag{B.4}$$

where F is the theoretical cumulative distribution which must be a continuous distribution, and it must be fully specified. The hypothesis of K-S test which regards the distributional form is rejected if its test statistic (D) is greater than the critical value obtained from a table. There are several variations of these tables in the literature that use somewhat different scalings for the K-S test statistic and critical regions. These alternative formulations should be equivalent, but it is necessary to ensure that the test statistics is calculated in a way that is consistent with how the critical values were tabulated.

Though widely applied in distribution fits test, flaws exists for these two comm goodness-of-fit test methods. The Chi-Squared test method depends on specifying the number of histogram classes into which the data will be grouped, and there is no good rule that gives the correct number to use. It also makes some assumptions that only come close to being valid when data set is provided. The Kolmogorov-smirnoff test method is deigned to test the goodness of fit for distributions with defined parameter values, not those where the parameters are estimated from the observed data. Corrections are possible for only a very few types of distribution.

None of these goodness of fit statistics penalize distributions for the number of parameters they use. Thus, a distribution with more parameters may well fit the observed data better because it has a lot more flexibility in shape than a distribution with fewer parameters, but the apparently improves over-fitting problem. Another problem is that none of these methods can correctly handle truncated, censored or binned data and they can not give a proper statistical weighting to the plausibility of each candidate distribution.

For finding the good and reliable fitted distributions of rain cluster's characteristics, in this thesis, several goodness of fit test methods are employed, which are Kolmogorov-Smirnov test (**K-S** test) method, Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC).

a. **Akaike Information Criterion** (AIC) (Akaike, 1998) is based on the use of Kullback-Leible's information as the discrepancy measure between the true distribution and the approximating distributions: $M_i = g_i(x, p_1, p_2, ..., p_n)$. The AIC for ith candidate distribution can be computed as in equation (B.5):

$$\mathbf{AIC} = -\mathbf{2}\prod(\theta) + \mathbf{2p} \tag{B.5}$$

where $\prod(\theta)$ stands for maximum log-likelihood of data set's sample, p is the parameter's number of candidate distribution. When the sample size n is samll, with respect to the number of estimated parameter P_i. The smaller for value of AIC, the best fitting result for candidate distribution.

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

$$\mathbf{BIC} = -\mathbf{2}\prod(\theta) + \ln(\mathbf{n})\mathbf{p} \tag{B.6}$$

where the symbols are equal to equation (B.6). And smaller value of BIC means the candidate distribution can well fit empirical distribution.

Totally fifteen candidate distributions are used for fitting empirical distribution of characteristics for rain clusters.The goodness of fit test results for candidate distributions of rain cluster's characteristics are presented in Table B.1 and Table B.2.

Table B.1 Goodness-of-fit testing results of candidate distributions, for light rain cluster's characteristics

Distributions	Area			Cum rainfall			Max intensity			Mean precip			Eccentricity		
	BIC	AIC	K-S	BIC	AIC	K-S	BIC	AIC	K-S	BIC	AIC	K-S	BIC	AIC	K-S
GPD	13907	13885	0.05	23522	23501	0.07	17383	17361	0.09	6339	6317	0.07	-6323	-6345	0.31
GEV	15574	15553	0.04	23686	23665	0.04	17811	17790	0.06	6676	6655	0.05	-14936	-14958	0.02
Inv_Gauss	16063	16048	0.06	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
Brinbaum Saunders	16087	16073	0.06	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
Log-Normal	16124	16109	0.06	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
Log-Logistic	16568	16553	0.07	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
Gamma	16750	16736	0.08	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
Nakagami	17613	17599	0.09	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
t Location-Scale	18390	18368	0.12	25781	25760	0.08	19913	19892	0.12	9469	9447	0.11	-11540	-11561	0.10
Logistic	18435	18420	0.12	25875	25861	0.09	20263	20248	0.10	9561	9547	0.10	-11317	-11331	0.10
Rician	18652	18638	0.11	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
Weibull	18653	18639	0.13	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
Normal	18894	18879	0.11	25800	25786	0.09	19904	19890	0.12	9469	9455	0.12	-10275	-10290	0.10
Extreme Value	23878	23864	0.22	29591	29576	0.17	22968	22954	0.16	13246	13231	0.18	-13842	-13857	0.06
Exponential	31197	31189	0.43	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	15731	15723	0.44

Table B.2 Same with table B.1, but for convective rain cluster's characteristics.

Distributions	Area			Cum rainfall			Max intensity			Mean precip			Eccentricity		
	BIC	AIC	K-S	BIC	AIC	K-S	BIC	AIC	K-S	BIC	AIC	K-S	BIC	AIC	K-S
GPD	3069	3050	0.06	6156	6137	0.10	2794	2775	0.13	-1850	-1869	0.10	-2663	-2681	0.32
GEV	3797	3779	0.05	5890	5871	0.03	2517	2498	0.05	-1989	-2008	0.03	-6079	-6098	0.03
Inv_Gauss	4075	4063	0.07	5871	5858	0.03	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
Brinbaum Saunders	4082	4069	0.07	5873	5860	0.03	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
Log-Normal	4090	4077	0.07	5892	5879	0.03	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
Log-Logistic	4250	4238	0.08	6104	6092	0.03	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
Gamma	4313	4300	0.08	5999	5986	0.04	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
Nakagami	4605	4593	0.09	6180	6168	0.06	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
t Location-Scale	4857	4838	0.12	6501	6482	0.07	2656	2637	0.06	-1762	-1781	0.05	-4685	-4704	0.11
Logistic	4862	4849	0.12	6537	6525	0.07	2872	2859	0.06	-1606	-1618	0.06	-4581	-4594	0.11
Rician	4959	4946	0.11	6456	6444	0.08	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
Weibull	5179	5166	0.14	6532	6519	0.07	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
Normal	5002	4990	0.11	6503	6490	0.08	2647	2635	0.06	-1770	-1783	0.05	-4112	-4125	0.11
Extreme Value	6907	6895	0.21	7968	7955	0.15	3427	3414	0.10	-702	-714	0.12	-5614	-5626	0.06
Exponential	11487	11481	0.48	12426	12420	0.40	NULL	NULL	NULL	NULL	NULL	NULL	6348	6341	0.45

For the character - area, ’Generalized Pareto’ distribution was the best fitted candidate with minimum BIC value of 13907/3069, the minimum AIC value of 13885/3050 and the K-S value of 0.05/0.06. For character cumulative precipitation derived from rain cluster above 19 dBZ, ’Generalized Pareto’ distribution is the best candidate with minimum BIC vaue of 23522, the minimum AIC value of 23501 and maximum K-S value of 0.07, for the ones derived from rain cluster above 37 dBZ, ’Inverse Gaussian’ distribution is the best one with minimum BIC value of 5873, the minimum AIC value of 5860 and the K-S value of 0.03. For character maximum intensity of rain cluster above 19 dBZ, ’Generalized Pareto’ distribution is the best one with minimum BIC value of 17383, the minimum AIC value of 17361 and K-S value of 0.09, for the same character derived from rain cluster above 37 dBZ, ’Generalized Extreme Value’ distribution was fitted to sample of data with the minimum BIC value of 2517, the minimum AIC value of 2498 and K-S value of 0.05. ’Generalized Pareto’ distribution is the best one for fitting sample of character areal mean precipitation from rain cluster above 19 dBZ with the min BIC value of 6339, the minimum AIC value of 6317 and K-S vaue of 0.07, but for those from rain cluster above 37 dBZ, ’Generalized Extreme Value’ distribution was the best one with the minimum BIC value of -1989, the minimum AIC value of -2008 and the K-S value of 0.03. For character eccentricity, ’Generalized Extreme Value’ distribution was chosen as the best fitted distribution with the minimum BIC value of -14936/-6079, the minimum AIC value of -6345/-6098 and K-S value of 0.31/0.03.

APPENDIX C

NORMALIZED DURATION LINE CALCULATING PROCEDURE

’Normalized Duration Lines’ for rain cluster’s characteristic prediction is calculated for ’Single-tracks’ which are generated by the RCIT algorithm, detail steps are listed as follows:

$a.$ Searching trajectories with same duration, here, each trajectory is defined as $track_i = [P_{t1}, P_{t2}, ... P_{tm}]$, where P_{tj} is the characteristic value at each time step tj; i is the id of trajectory and $i \in [1, n]$; duration of trajectory is $5 * m$.

$b.$ For time vector $T = [t_1, t_2, ... t_m]$ of trajectories with same duration, its normalized time vector T_{normal} is calcuated as in equation (C.1):

$$\mathbf{T_{normal}} = \frac{\mathbf{T} - \mathbf{T}_{\min}}{\mathbf{5m} - \mathbf{T}_{\min}} \tag{C.1}$$

where T_{min} is the minimum value of T.

$c.$ For each trajectory with same duration, its normalized vector is calculated as in equation (C.2):

$$\mathbf{track_{normal}} = \frac{\mathbf{track_i} - \mathbf{P}_{\min}}{\mathbf{mean}(\mathbf{P})} \tag{C.2}$$

$d.$ After step b and c, each normalized characteristics of rain cluster is grouped according to the normalized time vector within a range $[0, 1]$.

$e.$ Fitting method is applied and the coefficient parameters a and b of fitted parabola function are calculated for each characteristic’s normalized duration line, here, the Bi-Square robust way is used for fitting method (Rousseeuw and Leroy, 2005) .

Appendix D

Spatial Correlation Calculating Procedure

Spatial correlation calculation is based on the spatial distance of the rain cluster along its prevail moving direction, following steps illustrated procedure for calculating spatial correlation:

a. Each rain cluster in radar image is made of sets of Cartesian grids and each grid is surrounded by eight neighboring grids. According to their positions, each neighbouring grid is labeled and presented as in Figure D.1.

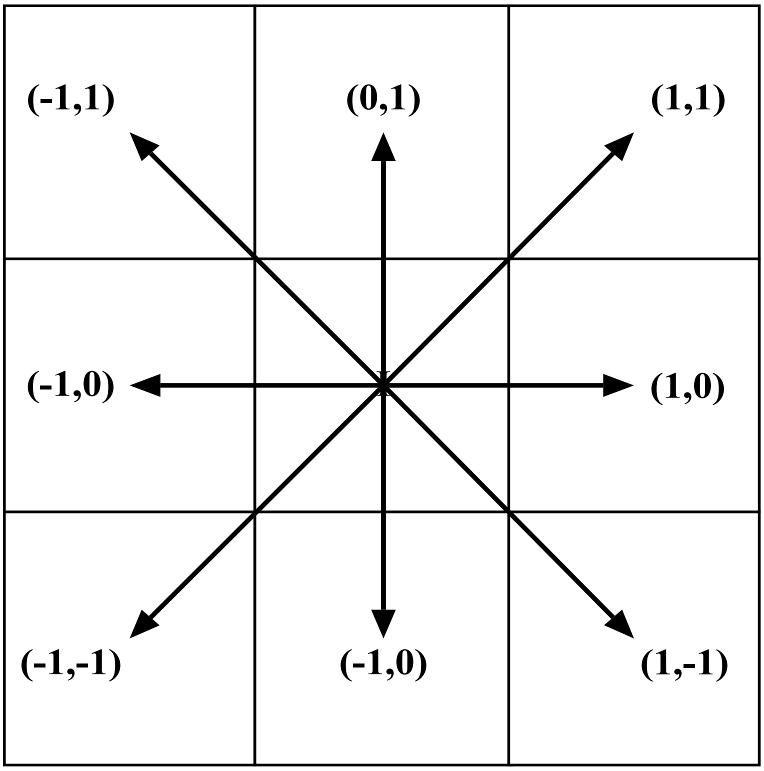


Figure D.1 Illustration of spatial distance in different direction.

In the figure A.1 each grid is labeled with unit vector, and the arrow means direction of the vector.

b. According to the prevail wind direction which is calculated by **RCIT** algorithm (figure 5.13 in Chapter 5), the spatial distance is chosen along that direction, e.g. for rainy day 27-May-2007 its prevail moving direction is from south-west to north-east for rain cluster above 19 dBZ, then the unit spatial distance is the vector of grid **(1,1)**; for rainy day 28-Jul-2008 its prevail moving direction is from south-east to north-west for rain clusters above 19 dBZ, then the unit spatial distance is the vector of grid **(-1,1)**. Finally, the spatial correlation is calculated based on equation 5.6 (in section 5.3 of Chapter 5).