UTRECHT UNIVERSITY

Graduate School of Natural Sciences

Climate Physics master thesis

Identifying atmospheric conditions that accompany "sparkles", i.e. small scale, intermittent lightning discharges near the top of thunderstorms.

First examiner:
Aarnout van Delden

Second examiner:
Michiel Baatsen

Supervisors KNMI:
Jelle Assink
Hidde Leijnse

Candidate:
Reinaart van Loon
1802038

In cooperation with:
KNMI
ASTRON

Brian Hare
Olaf Scholten

August 16, 2024

Abstract

Many aspects of lightning, including the initiation and propagation, remain poorly understood. The Low-Frequency Array (LOFAR) radio telescope, based in the North-East part of the Netherlands, allows to study lightning with unprecedented detail. Utilizing an array of antennas and an interferometry-based method, LOFAR can capture high resolution lightning images every $\tilde{2}0$ min.

On June 18 2021, LOFAR image small-scale, intermittent lightning discharges near thundercloud tops, which are named "sparkles". Similar observations in the U.S. indicate that sparkling clouds are associated with strong updrafts and overshooting cloud tops. Although hypothesized to be related to charge pockets, the exact mechanisms for sparkle behaviour remain unclear.

This research aims to improve the understanding of sparkles by comparing the sparkle events, on June 18 2021 in the Netherlands as imaged by LOFAR, with ambient atmospheric conditions, using data from the polarimetric C-band radar on Borkum island (Germany). A fuzzy logic hydrometeor classification algorithm provides insights into the particles aloft. Beside visual analysis, a spatio-temporal clustering algorithm distinguishes sparkles from other lightning structures and allows for a statistical comparison between the atmospheric conditions around sparkles and other lightning structures.

Radar images reveal that major sparkle clouds are located above mesoscale vortices. Furthermore, sparkles occur in areas of high radar reflectivity, horizontal divergence, turbulence, and graupel. Visual inspection shows that sparkles are often located near the edge of high-reflectivity volumes and near volumes with graupel. These findings suggest that turbulent conditions in the vicinity of graupel are responsible for sparkles. Beside the possibility of mixing between oppositely charged graupel and ice crystal volumes, it is also hypothesised that either local inductive charge separation or local charge sedimentation could be responsible for sparkle behaviour. In order to test this hypotheses, one might employ a high resolution atmospheric simulation that incorporates the inductive charging mechanism.

The accuracy of these results is limited by radar resolution, particularly in vertical height, and attenuation. Additionally, the statistical significance and confidence in the persistence of sparkle clouds is limited. Unexplained cloud top signatures in

radar differential reflectivity values warrant further investigation.

Contents

1. Introduction

Lightning is a mysterious phenomenon that has intrigued people for a long time. Historical artifacts dating as far as ancient Mesopotamia reveal that these powerful and mysterious flashes have long been associated with godly interventions. With the current scientific knowledge of electromagnetism, plasma-physics, chemistry and meteorology, we can explain many aspects of the initiation and propagation processes. Nevertheless, more mysteries in the field of lightning physics behold. Additional to the fundamental research to soothen our curiosity, understanding of the physical mechanisms of lightning also has a very practical application for meteorology. If causal relations between meteorological conditions and lightning can be established, observations of lightning may be used as a proxy of the atmosphere. For example to quantify the atmospheric composition[1], to study the dynamics inside thunderstorms[2], or to indicate hazardous thunderstorms[2]. Unfortunately, the road towards a thorough understanding of lightning is not straightforward. Research is challenging because of the transient nature of lightning, the large range of spatial and timescales involved, and the harsh conditions in thunderstorms.

Lightning measurements

Lightning itself can be measured both acoustically and via electromagnetic waves. The latter constitutes a very broad spectrum within the electromagnetic spectrum and therefore, there is a large variety of measurement techniques. Beside pictures with cameras from the ground or satellites, radio waves are the primary source for lightning detection networks. A number of organisations operate networks of radio antennas in a relatively low frequency band, e.g. Meteorage, Vaisala and EartNetworks. They can not only cover a large area, but also localise lightning flashes, distinguish different types of lightning, and estimate the peak electrical current of a flash. Although this type of data is a valuable asset for research into lightning statistics, there are some major limitations to such lightning detection networks. Lightning with small electrical currents are poorly resolved, and the system performs best for the less frequent cloud-to-ground type of lightning[3].

There are also systems that image lightning in a higher frequency band. For

example LMA networks [4], and in the Netherlands, the LOw Frequency ARray (LOFAR)[5]. The latter has managed to image lightning with an unprecedented resolution and this has contributed to the research into the fine structures of lightning, e.g. Hare *et al.* [6] in the discovery of so-called "needles", or Liu *et al.* [7] in the initial stages of lightning.

Sparkles

On June 18 2021, LOFAR captured 9 images of thunderstorms that crossed the Northern part of the Netherlands. These images are very rich in terms of lightning structures. The images also show much lightning activity at high altitudes, exceeding 10 km, which is relatively uncommon for the Netherlands. Between 9 and 14 km altitude, volumes are filled with small lightning structures. The small lightning structures seem disconnected from larger structures, occur intermittently throughout individual images, and the volumes seem to persist across different images. The observations of these small-scale, high-altitude, intermittent discharges are presented in Scholten *et al.* [8] and they name them "sparkles".

Observations similar to sparkles are documented by Emersic *et al.* [9], Calhoun *et al.* [10] and MacGorman *et al.* [2]. They established that small-extent, intermittent discharges may occur continuously above strong vertical updrafts where the cloud tops shoot beyond the tropopause. In this context, it is hypothesised that the small discharges are limited in size by pockets of charge. Although several mechanism are proposed to explain such charge pockets, no conclusive evidence is presented.

In Scholten *et al.* [8], evidence is presented that some of the sparkles are related to positive leader lightning structures that are otherwise invisible in LOFAR images. Many sparkles, however, cannot directly be attributed to positive leaders. Therefore, the question remains what processes inhibit the growth into larger structures but, at the same time, continuously initiate sparkles.

Research objectives

The aim of this research project is to improve the understanding of the mechanisms that cause sparkles. This is attempted by comparing data from a polarimetric Doppler radar on Borkum island (Germany) with sparkle events imaged by LOFAR during thunderstorms on June 18 2021. The study into the radar data is summarized by the following three research questions:

(a) What are the dynamical features of the thunderstorms that produce sparkles?

(b) What is the location of the sparkles with respect to the thunderstorm's clouds?

(c) What hydrometeors are near sparkles and in the clouds that produce sparkles?

Combining this information with previous studies, this thesis ultimately attempts to contribute to the main research question:
What charging mechanisms and charge structures lead to sparkles?

It would not only be of scientific value to improve the understanding of sparkles. The potential link between sparkles and very strong updrafts, particle composition, turbulence, and cross-tropopause mixing; would mean that sparkle observations might be used respectively; for the identification of hazards that accompany very strong convection, proxies of cloud tops, and quantification of stratospheric intrusion.

Outline

In the next chapter, relevant theory is discussed in the context of lightning. Previous research on sparkles and related observations are also extensively discussed in this chapter. Chapter 3 covers the methods that are used to process the LOFAR images and radar data, and how these data sources may be compared. Subsequently the results are presented and discussed in chapter 4 from both a visual and a computational approach. Conclusions are drawn in the final chapter 6. The appendix of this thesis provides supplementary figures of LOFAR images, an elaborate description of polarimetric radar variables, and comprehensive figures of the polarimetric radar variables.

2. Theory

This theory chapter serves three purposes. Firstly, the theory is introduced to understand the phenomenon of sparkles. This involves general background theory of thunderstorms in Section 2.1 and theory of lighting in Section 2.2. Secondly, Section 2.3 discusses some measurement techniques for lightning, and elaborates on the details of LOFAR and its lightning imaging procedure. Thus, the context is provided to appreciate the LOFAR images that are used in this thesis. Finally, Section 2.4 focuses on sparkles, including insights from previous studies on observations that seem similar to sparkles.

For the sake of conciseness, the theory on polarimetric radar variables is omitted from this section and is instead provided in appendix B.

2.1 Thunderstorms

Thunderstorms refer to very strong convective situations that produce lightning. Beside the onset of convection, there are multiple dynamical effects that can enhance convection to thunderstorm levels. Some microphysical aspect of thunderclouds, which play a role in lightning processes, are also described below. Theory about electrification and discharge processes can be found in Section 2.2.

Convection

Convection is a buoyancy effect that is caused by the instability of the atmosphere. Such an instability means that when an air parcel is displaced vertically, the displacement is amplified by buoyancy effects. The two variables that determine the stability of air, are the potential temperature and water content. When the potential temperature decreases with increasing altitude, an air parcel that is displaced upward will be warm with respect to its surrounding and thus have a relatively low density. Therefore it will start to rise upward. Vice versa, a parcel that is displaced downward will be cold with respect to its surrounding and negative buoyancy will make it sink.

The pressure in the atmosphere decreases with height and this influences the

temperature of air. According to the first law of thermodynamics, the temperature decreases when work is done by the expansion of an air parcel that is displaced to a lower pressure at higher altitude. Vice versa, the temperature increases when negative work is done by the compression of an air parcel that is displaced to a higher pressure at lower altitude. In order to accommodate for the pressure changes in the atmosphere, the potential temperature θ, i.e. the temperature of air parcels when they were to be at the same pressure, is used to determine the atmospheric stability of a vertical column.

The water vapour content influences the atmospheric stability by releasing latent heat upon condensation. When an air parcels cools upon expansion, the heat release would warm the air and thus increase the upward buoyancy. To accommodate for this effect, the equivalent potential temperature θ_e, i.e. θ of an air parcel if all the latent heat would be released, is used to determine the atmospheric stability over the whole troposphere. High moisture levels are a key ingredient for the strong convection in thunderstorms.

The troposphere is the bottom $\tilde{1}0$ km layer of the atmosphere with the stratosphere above. These two layers are separated by the tropopause, i.e. an abrupt change in the vertical θ gradient[11]. A very positive vertical θ gradient in the stratosphere constrains the convection into the stratosphere. It is the lack of convection, a significant agent of vertical mixing, that is conducive for the different chemical composition of the stratosphere. The latter contains significantly less water vapor and short lived gases that originate from the earths surface. Very positive vertical θ gradients may also be found below the tropopause. Such "capping inversions" also constrain convection. In a thunderstorm, the convection typically reaches beyond any capping inversions. In such case, air parcels near the ground reach the "level of free convection" (LFC) where they have positive buoyancy up to the statosphere. Such convection is called "deep convection". The Convective Available Potential Energy (CAPE) is a measure that indicates the instability of the atmosphere and indicates the kinetic energy that air parcels may acquire during deep convection.

One may define a "level of neutral buoyancy" (LNB) as the altitude where an air parcel that is lifted adiabatically from the ground, has neutral buoyancy with respect to the surrounding air. For thunderstorm situations, the LNB is sometimes near or above the tropopause. Because of the acquired momentum in deep convec-

tion, clouds may reach above the LNB. Such clouds are called "overshooting cloud tops", and they are limited in height by the negative buoyancy beyond the LNB, i.e. in the stratosphere.

Two other terms that are related to the convective vertical motion are 'updrafts' and 'downdrafts', referring to volumes of upward and downward air motion.

Hydrometeors

Hydrometeors are the water particles in the atmosphere. By definition clouds are composed of hydrometeors. In thunderstorms, hydrometeors of various kinds may be present. The interaction between water vapor and different hydrometeors is complex and an extensive description is beyond the scope of this report. Some terminology is introduced below in order to understand the electrification processes described in Section 2.2. Unless otherwise stated, this text is based on Saunders *et al.* [12] and one is advised to consult this source for further details.

There are three major processes that determine the size and number of hydrometeors: (1) phase changes, (2) collisions between hydrometeors, and (3) mechanical interactions with the wind. Phase changes depended on thermodynamics and the water vapor content. Water vapor can condense or freeze upon particles into the liquid or ice phase, respectively. In the case of deposition, ice crystals may grow into complex structures. The rate of growth, as described in further detail in Section 2.2, is important for ionization within hydrometeors and thus electrification processes. Inversely, hydrometeors may decrease in size by evaporation and sublimation. Evaporation and sublimation are favored by higher temperatures. Given that the (absolute) temperature of the atmosphere generally decreases with increasing altitude, the water vapour content also generally decreases with increasing altitude.

Although low temperatures are favorable for condensation, deposition or freezing of liquid water, these processes do not occur spontaneously. Hydrometeors or aerosols act as a catalyst for these phase changes. Water droplets for example, can sustain a liquid phase in sub-freezing temperatures. Droplets of "supercooled liquid water" may be found down to temperatures of -40 °C. Typically, supercooled liquid water freezes instantaneously when it collides with ice. This is called accretion.

The collision of hydrometeors mainly depends on number densities and the

relative velocities between hydrometeors. It should be expected that two kinds of hydrometeors that differ in terminal fall velocity, i.e. the vertical velocity when gravitational force and air friction balance, collide more often than hydrometeors that fall coherently at the same velocity. Turbulent motion also disturb coherent velocities and generally enhance hydrometeor collisions. Upon collision, hydrometeors may merge, rebound, scatter or potentially a combination of those processes. Liquid particles are more likely to merge than to rebound.

A wide variety of hydrometeors can be found in thunderstorms. Strong circulations and updrafts within thunderstorms allow particles to be suspended longer before falling out. This gives the hydrometeors more time to grow in size. Additionally, the turbulent conditions caused by convection promote the number of collisions[13]. Graupel, a common type of hydrometeor in thunderstorms, is associated with electrification processeses. Therefore, it is worthwhile to discuss this type of hydrometeor in further detail.

Graupel is a type of hydrometeor resulting from accretion of supercooled liquid water droplets unto falling ice crystals. This process results in nearly spherical ice particles with a diameter up to 8 mm. The frozen droplets on the surface of graupel are called "rime" and growth process of graupel is called "riming". Riming is driven by the higher terminal fall velocity of graupel compared to small liquid water droplets. This leads to collisions as graupel falls with respect to the droplets. In contrast,"unrimed" ice crystals, which are often found at the same places as graupel, have a terminal fall velocity similar to supercooled liquid water droplets. Consequently, there are fewer collisions between graupel and ice crystals and Thus, the growth of ice crystals is primarly limited to the deposition of water vapor. When ice crystals collide with graupel, both being ice particles, they are relatively likely to rebound. However some liquid water may be transferred upon collision. The reason for this is the melting of ice upon impact, or the presence of a liquid water layer on the hydrometeor surface during moist conditions.

2.2 Lightning

Lightning refers to the electric discharge through air. In nature it occurs when clouds of hydrometeors or dust produce an electric field large enough for dielectric breakdown of air. Once ionized, air has a much higher conductivity which allows for charge to flow down the gradient of the electric field. This phenomenon may be observed acoustically or by the emitted electromagnetic waves, albeit in the visible spectrum. The length and timescales associated with lightning have a large range, not only among different types of lightning, but also within the development of a single event. Moreover, the timescales of lightning are very small with respect to the human perception and sometimes even at the limit of the instrumental capacities. Given these complications, and the difficulties in replicating and modelling of the atmosphere where lightning occurs, there remain many questions in field of lightning science.

In this thesis, the lightning phenomenon of interest occurs in a convective thunderstorm. This subsection elaborates on the charging and discharge processes that occur in such convective thunderstorms. Furthermore, given the specific case study in this study, lightning is discussed in the context of the Netherlands. Unless otherwise stated, the source for this Section is [14], and the reader is advised to read this for a more comprehensive description.

2.2.1 Charge separation

Charge separation is the driver of lightning. It is the clouds of charge that induce the electric field that cause dielectric breakdown and charge flow. Charge separation happens at different scales. At the micro-physical scale, hydrometeors exchange charge. At the mesoscale, charge is distributed and accumulated by sedimentation: under the influence of gravity and the ambient wind velocity, particles of different shapes, sizes and densities may be advected differently and accumulate at different places. If there is a non-zero average charge transfer among different types of hydrometeors and sedimentation separates these different hydrometeors, this results in clouds of net charge.

Early observations of thunderstorms by Mason [15] showed that charge separation is associated with precipitation in the form of graupel, at temperatures between $0^{\circ}C$ and -40$^{\circ}C$, and with a tripole charge structure. Figure 2.1 illustrates a simple model. Graupel pallets have a higher terminal fall velocity than the much

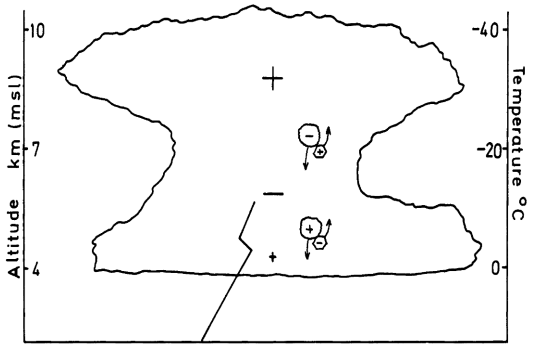

Figure 2.1: Figure from Saunders [16]. Charge transfer between graupel pellets (large particles moving downward) and ice crystals (small particle moving upward) leads to a tripole charge distribution. In this case graupel charges positively and negatively at lower and higher temperatures respectively.

smaller ice crystals and therefore the ice crystals are lofted upward in the updraft and to the anvil cloud. In combination with charge exchange at the micro-physical scale, i.e. positive upon graupel at low altitudes and negative upon graupel at higher altitudes, the top of the cloud is positively charged. The positively charged region at low altitudes results from a change of sign of the micro-physical charging mechanism such that graupel at the lower altitudes carries positive charge to the lower part of the cloud.

Laboratory experiments have proven two micro-physical charging processes that are viable for atmospheric conditions, on the appropriate timescales and providing the charging rate associated with lightning: non-inductive and inductive charge separation [16]. These processes are described below.

Non-inductive charge separation

The micro-physical process that is generally considered to be dominant in charge separation is non-inductive charge transfer upon rebounding collisions between riming graupel and ice crystals[16]. As outlined by Dash *et al.* [17], this process relies on the difference of vapor deposition rates between graupel and ice crystal. During fast vapor deposition, an outer layer of OH^- ions is formed because a disordered crystal structure allows the H^+ ions to diffuse to the hydrometeor's core of ice. It is thus the water vapor flux that determines the surface charge density. Upon collision, the surface of the particles melts. This allows for two types of charge transfer:(1) liberated ions will equilibrate over both liquid surfaces, and (2) upon separation the two particles take an approximate equal share of the liquid interface. Consequently, the particle which has the highest vapor deposition rate and thus the

highest negative charge density on its surface will lose more negative charge, and a net positive charge remains. It is important to note that the estimated timescale of collision of 0.1 ms allows for the exchange for the outer, negative charge, and is too short for the diffusion of the positive ions of the hydrometeor's core of ice.

This leaves the question what causes the difference in vapor deposition rate between riming graupel and ice crystals. According to Saunders *et al.* [12], the vapor deposition rate of ice crystals is solely determined by vapor diffusion from the environment. For riming graupel, super-cooled liquid water freezes at the surface, and therefore not only the vapor diffusion from the environment, but also vapor diffusion from the freezing droplets contributes to the vapor growth rate. The accretion of super-cooled liquid water droplets also releases latent heat, therefore warming the graupel surface and decreasing the diffusion rate. Various laboratory experiments have shown that the sign of charge separation, and thus the ratio of the vapor growth is determined by the temperature and liquid water content[16]. Within the vertical column of a thunderstorm's cloud, the temperature and liquid water content generally decreases. It is generally believed that, within the vertical, a sign reversal of the charge separation of the non-inductive charge mechanism explains the tripole charge structure as depicted in figure 2.1.

Inductive charge separation

The inductive charge mechanism relies on an ambient electric field that induces a charge polarization withing particles. This would not change the overall charge of individual particles, but merely redistribute it such that ions of opposite sign are accumulated at opposite ends. When two particles touch, the charge is redistributed over the whole. If the particles subsequently disconnect, both are left with a net charge. An example is illustrated in figure 2.2. Similar to this example, it is intuitive to comprehend why inductive charging enhances the electric field in updrafts. The smaller particles carried upwards rebound at the bottom of falling graupel and bring positive (negative) charge upwards in an downward (upward) pointing electric field. Vice versa, the graupel are charged negatively (positively) for an downward (upward) pointing electric field and carry this charge downward upon falling.

Laboratory experiments have shown that the collision of liquid drops would not suffice for substantial charge separation because of partial coalescence, i.e. merging of hydrometeors[12]. The low conductivity of ice also makes that charge transfer is

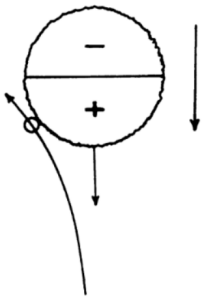

Figure 2.2: Example of an ice crystal (small particle), rebounding from a graupel pellet(large particle) polarized by an ambient electric field. The ice crystal gains a positive charge because it is in contact with the positive side of the large particle, leaving the graupel pellet negatively charged. The subsequent movement of the two charges particles enhances the ambient electric field. Figure from Saunders *et al.* [12].

too slow for ice-ice rebounding[18]. According to Mansell *et al.* [18], considerable charge transfer may be achieved by rebounding of liquid water droplets on graupel in dry growth mode. In a wet growth conditions liquid drops would coalesce with the liquid surface around graupel. Although the inductive charging mechanism is generally considered to be too weak to electrify a storm from fair weather conditions, it may play a major role when there is a substantial ambient electric field.[12]

Turbulence and charge separation

Generally, turbulence is associated with charge diffusion and therefore considered to weaken the electric field strength. However, Mareev *et al.* [19] demonstrate that in turbulent conditions, the inductive charging mechanism is enhanced significantly such that the threshold for dielectric breakdown is reached much faster. They use a model, employing two sizes of ice-particle, to quantify the effect of turbulence on both the inductive and non-inductive charging mechanism. Without turbulence, the linear increase of the non-inductive charging mechanism is dominant. For the inductive charging mechanism, the electric field increases exponentially over time.

The findings of Mareev *et al.* [19] may be understood when the direction of rebounds are considered. For the non-inductive mechanism, the charge that is carried upon particles after rebound is independent of the direction. After all, the charge transfer is determined by particle characteristics only. Turbulence might increase the number of rebounds, but not the rebounds that align with the charge sedimentation. Turbulence also increases the rebounds that oppose the charge sed-

imentation and thus oppose the overall cloud charging. For the inductive charging mechanism, the direction of rebound does affect the charge upon the particles. In principle, most types of rebounds would carry charge to increase the ambient electric field. Therefore, the increased number or rebounds in turbulent conditions allows for a net current independent of sedimentation effects.

2.2.2 Charge structure

The tripole model as illustrated in figure 2.1 is a relatively common observation[20], but there are also ample observations of more complex charge structures. This is not surprising, as convective systems do not only constitute of single updraft. As Bruning *et al.* [21] demonstrate, the evolution of the updraft and the advection from the active charge region may genera complicated charge structures. Some thunderstorms also exhibit a polarity that is inverted from figure 2.1. Bruning *et al.* [22] establish that this is primarily the result of changes of thermodynamics state within the updraft, influencing the non-inductive charging mechanism, and the vertical displacement of the updraft.

Charged screening layer

Not only the hydrometeors can carry charge, clear air contains charged ions. The electric field that is induced by charged clouds attracts or repels ions and thus a layer of ionized air is formed around charged clouds[23]. This so called "charged screening layer" can (partially) neutralize the electric field beyond the cloud. Some of the ions in the screening layer attach to the charged hydrometeors and can thus also neutralize the charge within the cloud.

2.2.3 Lightning discharge

The topic of lightning discharge is complex and involves physical processes at many temporal and spatial scales. A full description is beyond the scope of this report. This subsection discusses some aspects of lighting discharge that are relevant for the study. This includes the initiation of lightning, and some lightning phenomena that are key for the interpretation of the data that is used. Unless explicitly stated, the information is conform to the extensive works of Mazur [24] and Dwyer *et al.* [14].

Lighting is the result of the electric field that becomes strong enough such that an electrical current can propagate through air. In such case, the electrical break-

down ionizes the air, separating electrons from atoms. The mixture of free electrons and positive ions is much more conductive than non-ionized "virgin" air. An electric current also produces heat and thermal ionization may follow, increasing the amount of plasma and thus the conductivity. This process consist of multiple stages and realizations.

Corona discharges are the initial stage of electrical breakdown. In this type of discharge the air is ionized, but the plasma is relatively cold, not very conductive and barely luminous. Lab experiments have demonstrated two modes of corona discharges, when the electric field is increased gradually a corona glow may be observed while corona streamers are the result of fast changes in the electric field.[24] The corona glow often constitutes a cone, thus the name *corona*, that fans out from pointy ends of a charged surface. Similar to a charged screening layer, ions in the air might neutralize the electric field and thus inhibit further ionization. Corona glows are a common observation during thunderstorms, for example at the ends of vegetation. Corona streamers constitute filaments of current that last for a short time. The filaments may branch out to form a fan shape.

When the electric field is strong enough, corona streamers will be superseded by a channel of hot plasma called a leader. It is generally acknowledged that the ignition of leaders is a consequence of the heating by the combined current of many corona streamers. However, the initiation process is still an active field of research, e.g. Sterpka *et al.* [25]. The electric fields that are measured in clouds are weaker than the thresholds measured in laboratory measurements for leader formation.[26] Multiple theories have been developed that may explain a sufficient local enhancement of the electric field. To our current knowledge, it is most credible that the electric field in clouds is locally enhanced by a cascade of positive corona streamers[27], potentially in combination with local electric field enhancement by turbulent mixing[10] and charge polarization within single hydrometeors[26]. After initiation, both positive leaders and negative leaders may develop. Not only do these carry opposite charge, they propagate in different manners.

Positive leaders propagate relatively straight forward. Positive corona streamers fan out from the tip of the previously established positive leader. At the tip, where the streamers focus, the air is ionized such that the positive leader is extended. This allows for a relatively continuous growth of the positive leader channel over time. [24]

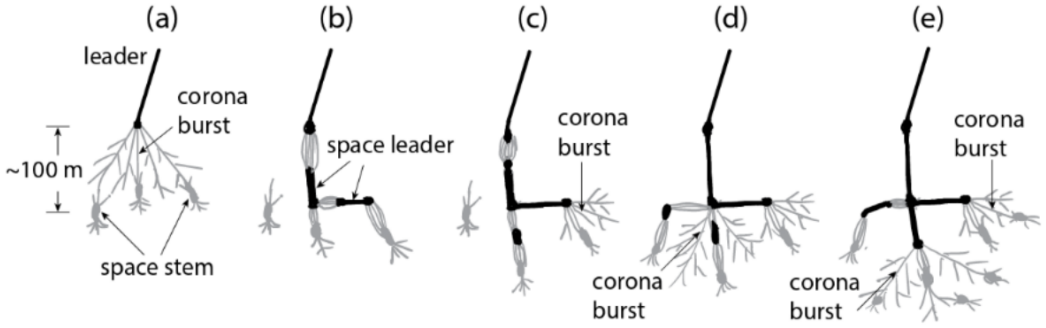

Figure 2.3: Figure from Liu *et al.* [27]. A schematic diagram of negative leader propagation. (a)A corona burst occurs at the tip of the negative leader may form space stems ahead of the negative leader. Space leaders are composed of relatively cool plasma and are the intercept of bidirectional corona streamers. (b)Some space leaders heat up sufficient to become very conductive space leaders.* (c)Space leaders grow in length and may connect.* (d)The connection between the negative leader and space leader leads to a current pulse and subsequent corona burst at the space leader end.* (e)Space leaders further ahead attach resulting in another current pulse and corona burst.* (*) previous steps happen simultaneously. Black colours represent hot plasma channels, thus leaders, and grey colours represent cold plasma, thus streamers and space stems.

The growth of negative leader channels is not as continuous, but involves bursts of corona streamers and plasma channels that are disconnected from the main leader. These so called "space leaders" are formed ahead of the negative leader, grow bidirectionally under influence of the positive and negative streamers at its opposite ends, and subsequently may attach to the negative leader. At attachment, a sudden charge pulse follows form the mismatch in the charge between the space leader and negative leader.[24] This may briefly illuminate the whole lightning channel and produce corona burst which may be involved in the development of new space leaders. A schematic diagram of negative leader propagation is depicted in figure 2.3.

Positive and negative leaders are charged positively and negatively respectively to their surroundings. Generally, they are bounded to clouds that are charged in an opposite sign[2]. In some cases, leader channels from a cloud connect to the earth, albeit through an object. For such cloud-to-ground (CG) lightning, much of the charge that was built up along leaders can now quickly drain away to the ground. The quick drainage of charge is called the "return stroke" and is often accompanied by explosive illumination and noise due to the heating by the large electrical current.

More frequently, leader channels do not connect to the ground, but live only within or between clouds. Such intra-cloud or inter-cloud (IC) lightning leaves a

"track" of charged plasma that cools down when the electrical currents decreases. Sometimes, the track is re-ignited. The propagation of the new leader is much faster because the cool plasma track is still much more conductive from the virgin air that the initial leader worked through. Some literature refers to this phenomenon as "dart leaders", e.g. Scholten *et al.* [8], because of their fast propagation. Multiple dart leaders are often observed consecutively through the same channel and there are also ample observations of dart leaders on CG leader tracks. In such case, return strokes can follow dart leaders. The fast consecutive nature of dart leaders is commonly observed as the flickering of lightning.

The cooling down of leader channels is also associated with another distinct type of lightning phenomenon called "needles". Hare *et al.* [6] show that dart leaders sometimes connect small flickering negative leaders. This may be explained by charge accumulation along a positive leader that is disconnected from the negative leader. As the channel between the positive and negative leader cools down, a positive leader may continue to carry away negative charge of the surrounding cloud. Instead of draining the negative charge into the negative leader, charge accumulates along the positive leader. Some of the negative charge is drained as needles, .i.e. small flickering negative leaders surrounding the positive leader.

Transient luminous events are a category of lightning that extends beyond the troposphere into the stratosphere and sometimes the ionosphere. This category includes phenomena like "sprites", i.e. cold plasma discharging that illuminates the air between the tropo- and ionosphere; "jets", i.e. blue lightning extending from the top of charged clouds; "elves", i.e. luminous circular-expanding cold plasma discharge in the ionosphere; and "sprite halos", i.e. horizontal ionization capping sprites.

Another category of lightning discharges that are worthwhile discussing are Compact Intra-cloud Discharges (CIDs), which produce some of the strongest radio signals among the different lightning phenomena. They may be observed in two different manners, namely bipolar electric field changes or short duration radio bursts. The former observation is also referred to in literature as Narrow Bipolar Events (NBEs), Trans-Ionospheric Pule Pairs (TIPPs) when observed by satellites, and sferics in the relatively low frequency radio regime. The relation between bipolar electric field change and radio burst is not fully understood. However, from the short duration of CIDs it is inferred that charge flows quickly through a hot leader

channel of a length in the order of a kilometer. This theory lacks visual observations that should accompany such a mechanism. Sometimes, leader channels are preceded by CIDs and therefore they are considered to be a type of initiation process. CIDs are most frequently observed in the upper part of clouds near or in strong convective updrafts.

2.3 Lightning detection and imaging

This section touches upon the topic of lightning observations with electromagnetic antennas. First providing the reader with relevant background information, this section subsequently describes the technical details of LOFAR and the lightning imaging technique that was used for the lightning images used in this thesis.

2.3.1 Electromagnetic (EM) lightning observations

The signal of lightning is different from the charge buildup within clouds. The latter may induce a slow change in the electric field, which can be measured and quantified by absolute measurements of the electric field. Lightning, on the other hand, induces transient changes of electrical currents and for this reason may emit EM waves. As described in section 2.2, lightning discharges occur at a range of different time and spatial scales, ranging from small corona streamers with small currents, to very large currents through long CG leader channels. In consequence, different phenomena may radiate EM waves at different wavelengths. A comprehensive description of lightning detection in the EM spectrum may be found in Rakov [28]. This source is also the primary source of the subsequent subsections about lightning observations.

Lightning produces EM waves primarily in the radio spectrum from below 1 Hz to 300 MHz, but also in the microwave spectrum, the visible spectrum, and even as very high frequency gamma rays. Within the EM radio spectrum, three main regimes are distinguished in lightning detection: (1) very low frequency (VLF), frequency $f = 3 - 30$ kHz and wavelengths of $\lambda = 100 - 10$ km; (2) low frequency (LF), $f = 30 - 300$ kHz and $\lambda = 10 - 1$ km; and (3) very high frequency (VHF), $f = 30 - 300$ MHz and $\lambda = 10 - 1$ m. As a direct consequence of the wavelengths, the different measurement regimes have different properties. The larger the wavelength, the further the EM waves travel and the better they refract around objects large objects, including the earths surface. On the other hand, smaller wavelength allow for higher resolution imaging of lightning. Furthermore, longer distance detection networks rely on relatively horizontal propagation of high amplitude EM waves. Consequently, such systems are most sensitive for strong electrical currents in the vertical direction. Thus, VLF and LF are most suitable for CG lightning.

2.3.2 VLF and LF lightning measurements

Modern lightning detection systems at (inter)national scale often employ a network of multiple EM antennas over large areas that cover the VLf and LF range. Such system are capable of localization, classification, and estimation of the peak current of a lightning event. Although this type of data is a valuable asset for research into lightning statistics, there are some major limitations to such lightning detection networks. Lightning with small electrical currents are poorly resolved, and the system performs best for the less frequent cloud-to-ground type of lightning. [3]. Pédeboy *et al.* [29] found, for example, that the Meteorage network, providing data to various European meteorological organisation including the KNMI, can detect CG lightning with an accuracy in the order of a hundred meters. IC lightning detection is much more uncertain in terms of the detection, classification and a localisation accuracy of a few kilometers.

2.3.3 VHF lightning observations

Networks of VHF antennas have very different capabilities from the lower frequency networks. The small wavelengths allow to image lightning with many data points in space and time. A lightning "image" refers to a collection of 4D points, three spatial and a temporal dimension, that correspond to VHF pulses emitted by lightning. In accordance to the wavelengths, the VHF EM pulses are believed to originate from sudden corona bursts. Such burst occur, for example, along the propagation of negative leaders, but not in the propagating positive leaders. For this reason, positive leaders are only visible in the VHF spectrum by dart leaders re-igniting a positive leader channel, or by needles.

In order to get a 4D lightning image, arrays of antennas are employed. Lightning Mapping Arrays (LMAs) are type of lightning imaging networks that have been the object of renowned studies into the field of lightning meteorology (e.g. as presented in section 2.4. LMAs are typically composed of 10-15 stations, each located at 15-20 km distance from each other. Using a time-of-arrival technique, LMAs can image lightning with an accuracy in the order of 30 meter.

2.3.4 LOFAR

The Low Frequency Array (LOFAR) is an array of radio antennas that was originally build for astronomical purposes. High resolution images are the product of

the high number of antennas, a phased-array design and an interferometry technique. Additionally, LOFAR has some unique digital processing techniques, making it suitable for many different kinds of observations[5].

There are two kinds of antennas involved in LOFAR, low band antennas (LBAs) focussing on the 10-90 MHz frequencies and high band antennas (HBAs) focussing on the 110-250 MHz frequencies. These antennas are grouped in stations that are located mainly in the northern part of of the Netherlands, but also across Europe. All Dutch LOFAR station are shown in figure 2.4. Each station constitutes 96 dipole LBAs and 48 HBA units called "tiles". The core of LOFAR, a circular area called the "superterp", is located in Exloo in the Netherlands and is composed of 7 stations. More details on the antenna configurations and characteristics may be found in Haarlem *et al.* [5].

The frequency range of LOFAR overlaps with the VHF spectrum of lighting. For astronomical observations, lightning may be a source of noise. However, LOFAR may also be employed for lighting imaging as an objective. A brief description follows of the characteristics and algorithms of the LOFAR lightning imaging technique based on Hare *et al.* [30] and Scholten *et al.* [31]. For lightning imaging, only the Dutch LOFAR LBAs are used in the 30-80 MHz frequency band. This corresponds to the VHF range in the field of lightning detection. Each LBA is dual-polarized with an inverted V-shape as depicted in figure 2.5, and the antennas are sampled at 200 MHz, thus every 5 ns. About 12 LBAs are used per LOFAR station and the data strorage is not continuous. A so called 'transient buffer board' (TBB) stores about 5 seconds of data per antenna. An external trigger, namely discharges as registered by lightninmaps.org, is used to write out the TBB data to a permanent storage. This trigger has a time delay of about a second. Complications during storage processes are compensated for by the large number of antennas. Galactic background radiation is used for amplitude calibration of the antennas.

With this LOFAR setup, lightning images can be captured about every 20 minutes, each containing useful data up to about 2 seconds. Further in the report, the property of capturing relatively short snapshots is referred to as the "triggered nature" of LOFAR.

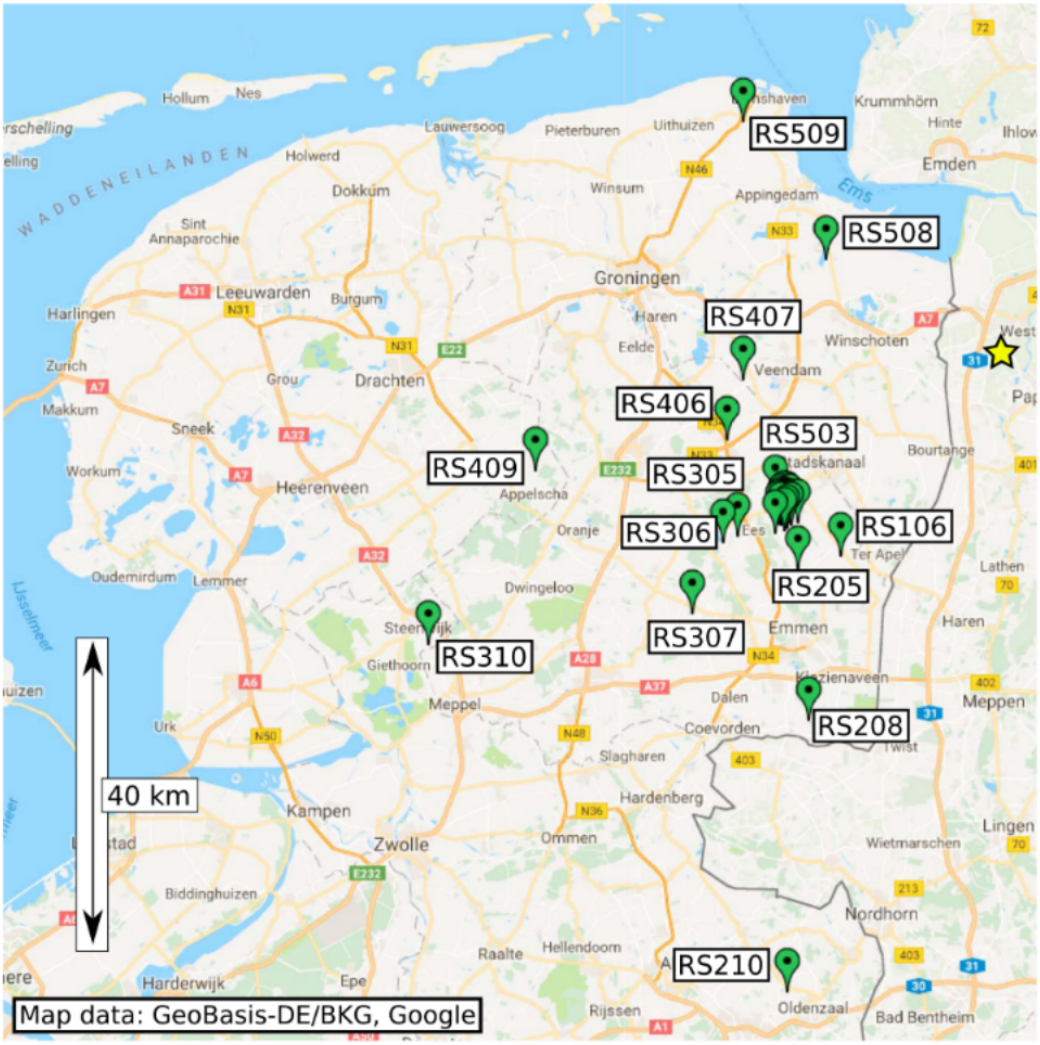

Figure 2.4: Figure by Hare *et al.* [30]. Map of LOFAR stations in the Netherlands. The core can be distinguished as the location with the densely packed stations. The yellow star may be ignored.

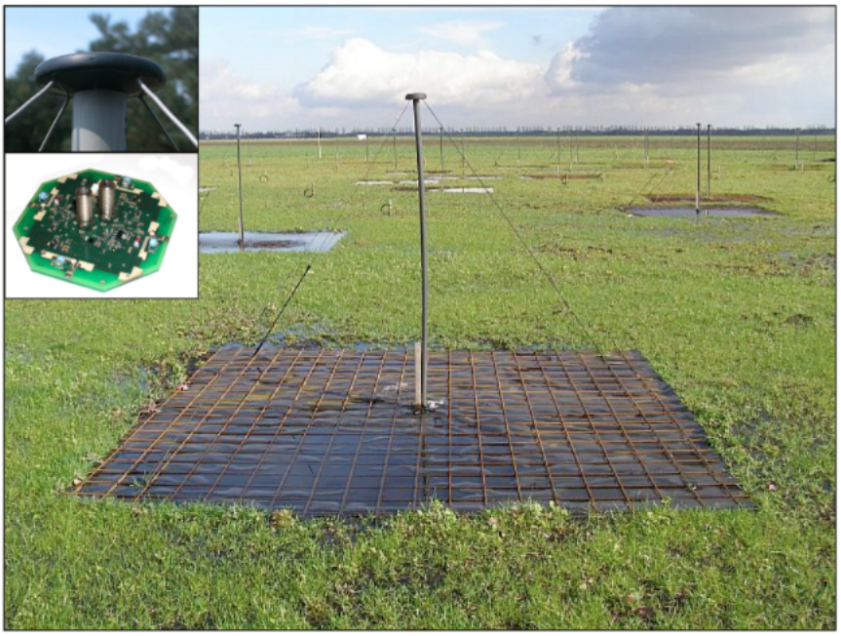

Figure 2.5: Figure by Haarlem *et al.* [5]. A single LOFAR low band antenna (LBA). Approximate size of 1 cubic meter.

Calibration

The algorithm to produce lighting images from the raw LBA LOFAR data starts with a digital filtering process that reduces the interference of unwanted radio sources. Subsequently, pulses are determined computationally. A pulse may contain multiple peaks that correspond to individual VHF emitting events.

In order to reach a high image resolution, an accurate time calibration is needed among different stations as well as individual LBAs. This is achieved with an algorithm that, for each lightning image, attempts to find suitable calibration pulses, corresponding to VHF sources, between a reference and nearby antennas; calculates a first guess of the 4D point sources fitting those pulses; iteratively adds more antennas, and updates the suitable pulses and corresponding point sources; and finally calibrates all antennas by simultaneously fitting the optimal pulses in time and space. For source localisation, the difference between the calculated and measured pulse arrival times are minimized. Visual inspection is needed in order to correlate signals between different stations. Scholten *et al.* [31] finds for about 15 sources in a single image, a root mean square error of 1-2 ns in terms of pulse timing.

Imaging

When the antennas are accurately calibrated in time, an algorithm is employed to image point sources. This involves a procedure, inspired by the Kalman filter, to gradually include stations of increasing distance to the LOFAR core. The algorithm first selects, within short time windows in the signal of the reference antenna, distinctive pulses. In order of decreasing intensity of the selected pulses, a first initial guess of the source location is calculated with nearby antennas; a chi-square fitting procedure based on the initial guess yields a more accurate source location; iteratively, antennas at larger distances are included in the fitting procedure if they contain the same distinctive pulse; and the procedure is repeated for the next largest pulse. The time window is shifted such that the whole length of the LOFAR data is imaged. This procedure can image about 200 VHF sources per millisecond with an accuracy that is estimated to be in the order of a meter. It is a subjective matter what threshold criteria are employed on the data quality, balancing between accuracy and the number of points to include.

Besides the interferometry technique described above, another imaging technique employs a beamforming technique for 3D VHF intensity imaging[32]. This technique is not further discussed because it is not relevant for the data in this thesis.

2.4 Sparkles

On June 18 2021, LOFAR captured 9 images of thunderstorms that crossed the North of the Netherlands. These images are not only very rich in terms of VHF sources and lightning structures, they also show much lightning activity at high altitudes, exceeding 10 km. Such high altitude VHF activity is rarely observed in the Netherlands as so called High Altitude Negative Leaders (HANLs)[33]. Some of the LOFAR images show very different structures, volumes between 9 and 14 km altitude filled with small structures that seem disconnected from larger structures. They occur intermittently throughout individual images, and these volumes also seem to persist across different images.

Scholten *et al.* [8] present their observation of the small-scale, high-altitude, intermittent discharges and they name them "sparkles". They define them as "spatially small (not exceeding a few 100 m) intermittent discharges near the top of the clouds where this activity is seen over extended periods lasting the full duration of a LOFAR recording". In the remainder of this report, "sparkle clouds" refer to volumes filled with sparkles. It is yet to be determined if sparkles actually only reside inside the cloud in a meteorological sense that is filled with hydrometeors. It is also yet to be determined if sparkles would be a distinct entity as a lightning phenomenon. As for the LOFAR images, an appreciable continuum of sparkles sizes, structures, propagation speeds and intensities are observed within sparkle clouds. At least, the observation of multiple VHF sources that constitute individual sparkles yields confidence that they are not artefacts of the LOFAR imaging technique.

Scholten *et al.* [8] observe that some sparkles are later connected by dart leaders and therefore it may be assumed that those sparkles were connected to positive leader tracks. This would indicate a very similar mechanism to the needles that were discovered by Hare *et al.* [6]. Unlike needles, there is no evidence that sparkles exhibit flickering. Needles have only been observed before at lower altitudes and thus it is hypothesized that the difference with sparkle characteristics is caused by the atmospheric conditions at high altitudes. Although a needle-like mechanism is potent, the majority of the sparkles cannot be ascribed to positive leader tracks.

2.4.1 Similar observations

There are multiple other observations that match the characteristics of sparkle. A description follows from the most compelling cases, namely observations by the LMA in Oklahoma USA and camera pictures showing blue luminous events (BLUEs).

BLUEs

BLUEs are lightning events that are perceived as blue by the human eye. Corona streamers are believed to be responsible for the strong 337 nm, blue light component[34]. Within the definition of BLUEs are "blue jets" and "blue starters", transient lightning phenomena that extend upwards from the clouds up to 40 and 25 km respectively; but also small blue discharges that are associated with NBEs (Section 2.2). The latter shows resemblance to sparkles. Some examples follow:

Lyons *et al.* [35] identified "gnomes" and "pixies" in a lightning campaign involving photography. These terms correspond respectively to blue lightning extending 1 km beyond the cloud top, and isolated small blue discharges in the dome of the cloud.

Chanrion *et al.* [36] present pictured observations of relatively small blue discharges at the top of overshooting cloud tops. These occur over an area of 4-9 km^2, and are an order magnitude smaller than pixies, and have a duration up to 125 ms. From the observations they infer that the discharges were generated at the screening layer and that the discharges are the results of corona streamers. It is hypothesized that the sparkle-like behaviour follows as a consequence to the lower threshold for electric breakdown at high altitudes; turbulence to increase the electric field locally; and a weak screening layer to restrict the size.

Soler *et al.* [34] compare pictures of BLUEs with electric field measurements at the ground. They find the events to be composed of single or occasionally multiple pulses with a duration of 1.6-3.6 ms. By modelling the diffusion of light in the clouds, they infer that the discharges occur 1.7- 6.5 km below the cloud top interface. This is consistent with EM field measurements. It is also suggested that the BLUEs correspond to NBEs.

All of the above studies found that BLUEs were isolated from leader structures and their intermittent nature is also in line with sparkle observation. Nevertheless, it is not entirely clear to what extent the observations of BLUEs and sparkles

are related. Although the observations by Chanrion *et al.* [36] are much larger and longer than the LOFAR sparkles, this may be related to differences between the observation methods. LOFAR's antennas are only expected to measure corona stream bursts, while gradual streamer formation may be observed better in the visual spectrum. It is also noteworthy that LOFAR images show a much higher sparkle frequency and density than observed for BLUEs. Absorption in the visual spectrum might diminish the number count of BLUEs. Moreover, a high sparkle density may not be distinguished in 2D camera pictures. The absorption may also bias the inferred location of BLUEs such that only the light at the edge or interface of clouds may be perceived by a camera.

Oklahoma LMA observations

Emersic *et al.* [9] reports of a thunderstorm in Oklahoma in 2006 and the corresponding lightning observations. The storm dynamics and micro-physical characteristics are derived from the high resolution National Weather Radar Testbed Phased-Array Radar (NWRT-PAR) with a 1 minute time resolution. LMA data reveals, besides very high lightning flash rates, two episodes of transient, isolated VHF sources at high altitudes above the main updraft. These episodes lasted for about 4 minutes each and the timing matched well with increasing altitude of high radar reflectivity values, thus updraft intensification. Whereas the first episode was located 1 to 4 km above the 30 dB reflectivity contour, the second episode was surrounding the interface at the top of the 30 dB reflectivity contour.

Calhoun *et al.* [10] presents a case study of a thunderstorm in Oklahoma 2004, comparing lightning data with images from two mobile C-band radars. They observe high altitude, isolated VHF sources above and somewhat downstream a strong updraft. These VHF sources persist continuously at a relatively low rate for more than an hour. Increasing rates of the high altitude VHF sources were observed to match increasing updraft strength.

Calhoun *et al.* [37] builds upon the observations by Calhoun *et al.* [10]. A high resolution atmospheric model of the same storm, not only assimilating the radar data, but also incorporating electrification and discharge processes, is studied to understand the lightning observations. Although the model does reproduce charge pockets and the corresponding small extent lightning flashes near the updraft, it did not show the small-scale high-altitude flashes that resemble sparkles.

MacGorman *et al.* [2] analyse five cases of thunder storms that the exhibit small-extent, isolated, high-altitude VHF sources. In either the KUON polarimetric radar or NWRT-PAR images, they find a consistent relation between the cloud top height and even a significant correlation of 0.86 between the top height of the 18 dB radar reflectivity contour and the maximum altitude of the sparkle-like discharges. Similar to Calhoun *et al.* [10], the cloud of sparkle-like discharges had a domed appearance capping the updraft.

A number of hypotheses are proposed in these three LMA studies that relate to sparkle characteristics:

- All three studies acknowledge that the threshold for electric breakdown is lower at the atmospheric conditions at higher altitudes. This could potentially increase the flash frequency at high altitudes.
- All three studies mention an interaction of the charges cloud with the charged screening layer aloft. This could locally enhance the electric field by decreasing the distance between volumes of opposite charge. Calhoun *et al.* [10] hypothesize that turbulent eddies at the top of overshooting clouds could fold the screening layer into the cloud. More specifically, MacGorman *et al.* [2] suggests that, due to interaction of the updraft with the horizontal wind speed, the charged screening layer could fold quasi periodically into the cloud. Neither of the studies present direct observations or physical models to show such folding.
- Although not directly related to the sparkle-like VHF sources, Calhoun *et al.* [10] hypothesize that the small extent and high frequency of flashes near the updraft are caused by charge pockets. If charge is distributed homogeneously only within small pockets, a flash feeding on such a charge pocket would be inhibited to grow into large structures. According to Calhoun *et al.* [10], charge pockets near the updraft core may be a result of turbulence, due to large horizontal shear in the updraft velocities, mixing the heterogeneous charge structure within an updraft. The complexity in the updraft's charge structure could be caused by the entrainment of cold, dry air into the updraft and variability of latent heating. As such, the temperature and liquid water content would vary within the updraft column and thus the sign of the non-inductive charging mechanism.
- Calhoun *et al.* [10] hypothesize that the high charge density at the top of the

updrafts may be related to the sparkle-like characteristics. In the overshooting cloud top, particles accumulate due to the decreasing vertical velocities. This theory does, however, not explain the limited spatial extent of sparkles. Therefore, MacGorman *et al.* [2] adjacently hypothesize that within one updraft, variability in the liquid water content, temperature, and time for precipitation growth, would lead to a variability in the charge carried upon graupel. At the cloud top, accumulation of the heterogeneously charged graupel pellets could result in charge pockets of sufficient charge density for sparkles

There are similarities between the characteristics of sparkles as observed by LOFAR and the sparkle-like VHF sources observed by LMA, namely the small spatial length-scale and high-altitude. Although the triggered nature of LOFAR inhibits to draw firm conclusions, the persistence of sparkles within and between LOFAR images make it at least likely that sparkles endure for similar timescales as observed by the Oklahoma LMA network, i.e. the duration of updraft surges. The higher sparkle density and frequency in some LOFAR images, might either be explained by the higher resolution or meteorological conditions. If sparkles are indeed the same phenomenon as observed with the LMA network, one might expect a similar relation to updraft surges as shown by Emersic *et al.* [9], Calhoun *et al.* [10] and MacGorman *et al.* [2].

3. Methods

This chapter elaborates on the methods that were specific for this thesis. Section 3.1 describes an algorithm used to computationally classify LOFAR sources as sparkles. Section 3.2 describes the radar data and the algorithms to compare the radar data with LOFAR. Section 3.3 describes some important aspects to consider the hydrometeor classification (HMC) algorithm applied to the radar data. A more thorough description of the HMC algorithm can be found in Appendix B. Since the LOFAR images were processes outside the scope of this study, the relevant information of LOFAR has previously been discussed in Section 2.3.4.

3.1 Sparkle classification algorithm

The LOFAR data consists of VHF point sources in space and time. Visually, it is often easy to determine what points belong to the same lightning structures. Since the LOFAR images contain thousands of VHF sources, it would be very cumbersome to classify sparkles from other structures manually. Inspired by the clustering algorithm as described by Fuchs *et al.* [38], this thesis also makes use of the scikit-learn package DBSCAN function[39]. This function acts much like a basic distance clustering algorithm, connecting points that are within a defined distance of each other, but with a more stringent condition on remote points. Remote points only belong to a cluster when they are within reach of another point that, by itself, is within distance of a minimum number of other points. One can choose a distance threshold to accommodate for the clustering needs.

If tuned right, the DBSCAN algorithm can find large lightning structures as clusters of VHF sources. If so, the VHF sources that are not clustered, and are located above 8 km altitude are classified as sparkles. The altitude threshold of 8 km was chosen to eliminate clustering errors below the altitudes associated with sparkles. Moreover, it could be that small, sparkle-like lightning structures at lower altitudes are physically different from the actual sparkles at cloud top altitudes. Therefore, low altitude sparkle classifications might contaminate the analysis.

Clustering parameters

Lighting flashes propagate through time and space. Therefore, a simple euclidean distance of the three spatial dimensions would not be suitable to cluster lightning structures. Instead, a four dimensional euclidean distance if used for clustering. The three spatial dimensions (x, y and z) are normalized with a factor d_0 and the temporal dimension (t) is normalized with a factor t_0. The threshold for clustering D is thus defined as:

$$D^2 = (x/d_0)^2 + (y/d_0)^2 + (z/d_0)^2 + (t/t_0)^2 \tag{3.1}$$

Although Fuchs *et al.* [38] use normalization factors $d_0 = 3$ km and $t_0 = 0.15$ s, these parameters are not necessarily universal. This would depend on the VHF point and the flash density. As for the LOFAR data, having a much higher resolution than the LMA data, it could be expected that these parameters are too coarse.

In this thesis, clustering parameters were chosen on behalf of a trial-and-error process. The aim of this process was that the classification algorithm selects, per individual LOFAR images, as many sparkles as possible, but with as little contamination of other lightning structures. A visual inspection, identifying small-scale discharges above 8 km as sparkles, and identifying large lightning structures such as negative leaders and dart leaders, was used to judge the sparkle classifications. When tolerable, the parameters were chosen to be similar between different LOFAR images. This has led to two sets of parameter options: $d_0 = 2000$ m , $t_0 = 0.3$; and $d_0 = 1000$ m , $t_0 = 0.15$. The common factor between the parameter of $d_0/t_0 = 2/3 \cdot 10^4$ m/s is accidental. The optimal parameters per LOFAR image, are listed in table 3.1. It should be emphasized that these parameters are not per se optimal and might be tuned in case of a more stringent definition of sparkles. There is no objective metric available in this study to quantify the performance of the sparkle classification algorithm and the chosen set of parameters.

3.2 Radar data processing

The radar data that is used in this report comes from a dual-polarimetric c-band radar, located on the island of Borkum, Germany. This specific radar was chosen

Table 3.1: Table with appropriate clustering parameters for each LOFAR image.

Time LOFAR	d_0 (m)	t_0 (s)
17:46:58	2000	0.3
18:06:44	2000	0.3
18:26:40	1000	0.15
18:39:23	2000	0.3
18:58:48	2000	0.3
19:17:36	2000	0.3
19:37:29	1000	0.15
19:54:24	1000	0.15
20:11:51	2000	0.3

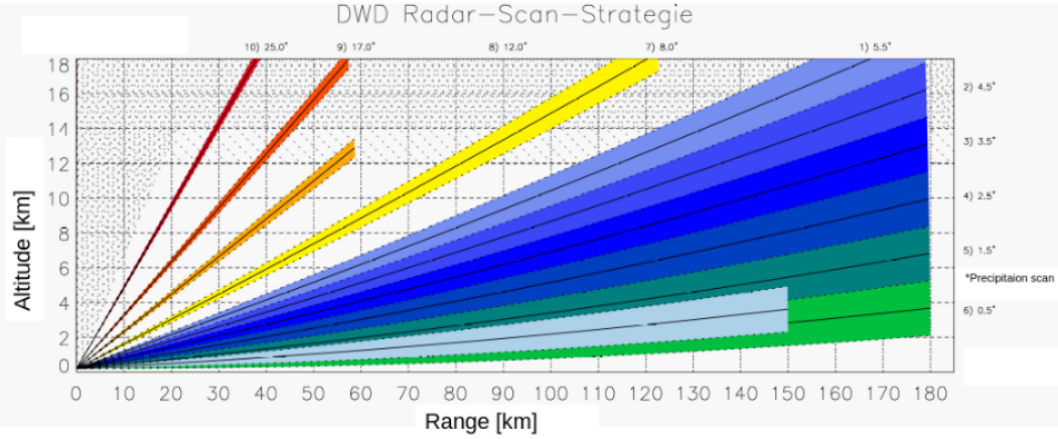

Figure 3.1: Adapted version of figure from Frech *et al.* [40]. Scanning strategy of the DWD radar on Borkum. Different radar beams and the 1° beam width are represented by shadings. The corresponding angles are shown at the side of the axes. *The precipitation scan is not used in this study and the remainder of the report.

given its favourable location and range with respect to the location of interest, i.e. the North of the Netherlands. The technical details of this radar can be found in Frech *et al.* [40]. Some key aspects of the radar are listed:

1. C-band frequency range of 5600-5650 MHz, with 0.4 and 0.8 µs pulses.
2. Volume scan with 10 elevation angles between 0.5° and 25° with a beam width of 1°, and a range bin separation of 250 meter. See figure 3.1 for the scanning strategy.
3. The complete volume scanning procedure is repeated every 5 minutes.
4. Raw data is processed by the German Meteorological Service (DWD) into standard radar variables with MURAN software.

The polarimetric radar variables that are used for this report are elaborately discussed in appendix B. In brief the different polarimetric radar variables can be

interpreted as follows:

- The horizontal reflectivity Z_h represents the amount of backscatter from the horizontally polarized radar signal. Higher values indicate either more hydrometeors, larger hydrometeors or more liquid (with respect to ice) in hydrometeors.
- The differential reflectivity Z_dr is the ratio of the horizontal and vertical polarized reflectivity signal, Z_h and Z_v respectively. Generally positive values indicate oblate hydrometeors. Vice versa, negative values indicate elongated hydrometeors.
- The zero time-lag correlation coefficient ρ_{hv} is the correlation between the Z_h and Z_v signal within one radar resolution volume. Decreased values from unity indicate heterogeneity in the alignment or shapes of the hydrometeors.
- The specific differential phase K_{dp} is the phase shift between the horizontal and vertically polarized radar signals upon propagation through a resolution volume. Positive values indicates oblate hydrometeors and negative values indicate elongated hydrometeors.

K_{dp} is computed from from the raw, noisy total differential phase ψ_{dp} as provided by DWD. An interpretation of the differential phase and details on the method to compute K_{dp} can be found in appendix B.

Spatial data processing

Throughout the analysis, the LOFAR data is compared spatially with radar data. There are two important factors to consider for a valid spatial analysis, the georeferencing of the radar data and the time difference between radar and LOFAR data.

For georeferencing the radar data, the wradlib.georef.georeference algorithm[41] has been applied. This algorithm makes uses of the equations provided Doviak *et al.* [42] that take into account the curvature of the earth and the refractive index of the atmosphere. The altitude of a radar resolution volume is given as:

$$h = \sqrt{r^2 + (k_e a)^2 + 2 r k_e a \sin\theta} - k_e a \tag{3.2}$$

With r the radial range from the radar, θ the radar scanning azimuth, a the radius of the earth, and k_e a factor that compensates for the refractive index of the atmosphere.

The radial distance of a radar beam projected on the earth surface if given as:

$$s = k_e a \arcsin \frac{r \cos \theta}{k_e a + h} \tag{3.3}$$

A standard value of $k_e = 4/3$ is used as the refractive index compensation factor. Some literature refers to $k_e a = a_e$ as the equivalent earth radius, since mathematically a_e corresponds to a situation without refraction.

For comparison, LOFAR data is converted to the azimuthal equidistant projection that is native to the radar georeferencing described above. All figures that include radar data are in this azimuthal equidistant projection.

In order to deal with the time difference Δt between LOFAR data and radar data, an advection scheme is applied to the radar data to match the time of LOFAR images. In this study, no attempt is done to perform a temporal interpolation of the radar data. Instead, using the time of the LOFAR image of interest, the radar sweeps are collected that are closest in time, one of each elevation angle. Subsequently, the horizontal wind velocities of the ERA5 data of June 18 2021, are linearly interpolated in time and space unto the location of each radar data point. This yields the horizontal wind vector $\vec{v}_i$, corresponding to each resolution volume i. These values are used to move the georeference $\vec{x}_i$ of the each radar data point over a distance of $\Delta\vec{x}_i = \Delta t_i \cdot \vec{v}_i$. Here, Δt_i is the time difference between the LOFAR data and each radar data point.

In practice, the advection scheme has moved radar data up to 3 km. This is not surprising given the maximum time difference of 150 seconds for the 5 minute radar scanning procedure and the high altitude wind velocities of more than 30 m/s. Within one sweep there may be significant differences in advection shifts. Especially between the first and last scanning azimuth which are adjacent but have a time difference of 20 seconds. The interception of the tropopause might also be visible in the advection shift, because of the significant wind veering and shear within this region.

To reduce computational cost, the ERA5 altitudes are replaced by the average altitude of each model level. As a results, the ERA5 data is structured on a Cartesian grid and the interpolation is much less costly.

Quantitative comparison with LOFAR

In order to quantitatively compare the radar data near sparkles and other lightning structures, the algorithm in for this study collects (category A) radar data points in proximity of sparkles, and (category B) radar data points in proximity of other LOFAR VHF sources. Any radar data points that were in both near sparkles and near other VHF sources, thus both in category A and B, are removed from category B. As such, the category A dataset can be interpreted as an atmospheric volume conducive to sparkles. Dataset B, on the other hand, can be interpreted as an atmospheric volume that is only conducive to other types of lightning.

The *scipy.spatial.KDTree.querry_ball_point* function was used for collecting the radar points within a radial distance of the LOFAR data. In all results, a radial distance of 2 km is used as a threshold for "proximity". This is a subjective choice, that seems reasonable keeping in mind the poor vertical radar resolution up to 4 km, the temporal radar resolution of 5 minutes, and the uncertainties in advection over such timescales.

3.3 Hydrometeor classification

In order to characterize the atmosphere in terms of hydrometeors, the Wradlib fuzzy-logic hydrometeor classification (HMC) algorithm is employed with 2-dimensional trapezoidal membership functions. As explained elaborately in appendix B, a fuzzy-logic HMC-algorithm calculates, for each resolution volume, the probability that the volume is filled with a specific type of hydrometeors. The resolution volume is subsequently classified as the type of hydrometeor that has the highest probability. For example, when the HMC-algorithm finds a probability of 0.8 for graupel, 0.5 for dry snow and 0.2 for wet snow, it is classified as graupel. The classification skill can be quantified with two metrics: (1) the value of the highest probability and (2) the difference between the highest and second highest probability. The former is indicative of how good the radar data fits the membership functions, thus a higher probability is better. The latter is indicative of how good the algorithm can distinguish between the best and second best choice. Given the fact that the membership

functions are adjacent and partially overlap, and taking into account potential mixing of different hydrometeor types, values larger than 0.3 are considered as a good performance in terms of distinction.

In order to judge on the skill of the hydrometeor classification, two metrics are deployed in this research. Visually, the maximum probability of classification is plotted. Examples may be found in section 4.2 and appendix C. Secondly, a probability aggregate over all included resolution volumes is computed as:

$$P_{HM,aggregate} = \frac{1}{N} \sum_{n=0}^{N} P_{HM,n} \tag{3.4}$$

Here N is the number of radar resolution volumes included, and $P_{HM,n}$ is the probability of hydrometeor type HM for resolution volume n. $P_{HM,aggregate}$ combines both metrics to quantify the skill, but is only useful when the HMC statistics show dominant mode, i.e. if the majority of the data points is classified as a particular HM type. The significance of the dominant mode is both determined by the absolute value, and the difference between the $P_{HM,aggregate}$ values of the dominant and alternative HM types.

The polarimetric variables that are included in the HMC are Z_h, Z_{dr}, ρ_{hv}, K_{dp} and the temperature T. The latter is obtained from the ERA5 reanalysis for June 18 2021. All polarimetric radar variables are weighted equally. The membership functions have been retrieved from Wradlib [43] and were constructed based on radar signal simulations by Marzano *et al.* [44].

4. Results

4.1 Thunderstorm description

Just like many meteorological phenomena, thunderstorms can be characterized on different spatial and temporal scales. This section's focus is on the description of the thunderstorm at the mesoscale, i.e. the time and length scales associated with convective systems. The mesoscale features are relevant with regard to the aim of this study, because charge buildup is a consequence of convection. Moreover, as illustrated below, mesoscale features of convective systems can be retrieved from polarimetric radar data.

In figure 4.1, the radar reflectivity and VHF sources classified as sparkles are projected on the horizontal plane, for all nine LOFAR images. As first glance, it is evident that the sparkles are always close to areas of high reflectivity. From the combined radar and LOFAR images, we can also distinguish three convective systems of interest, as marked with A, B and C in figure 4.1.

It can be observed that first, at 17:46 and 18:26 UTC, system A produces most sparkles. Starting at the 18:39 UTC image, system B produces sparkles as well. A qualitative view of the LOFAR images (Appendix A) shows that the sparkles classified in cell A at 18:39 and 18:58 UTC, and the sparkles classified in cell B at 19:17 are probably erroneous and are likely part of larger lightning structures. Beside some classification errors at 17:46, 18:39 and 19:37 UTC, system C does produces a few VHF sources that are small and isolated. However, it lacks persistence and significance of sparkle-like VHF sources. As can been seen in appendix A, system C does produces significant amounts of lightning that is not classified as sparkles. Especially at lower altitudes.

In order to evaluate the evolution of the convective cells with respect to lightning, the maximum altitude of the $Z_h = 40$ dB contour was estimated. The choice for the $Z_h = 40$ contour is somewhat arbitrary. However, it is expected that only graupel, hail or very large rain can produce such high reflectivity values. At the top of clouds, these types of hydrometeors are associated with updrafts[45]. The

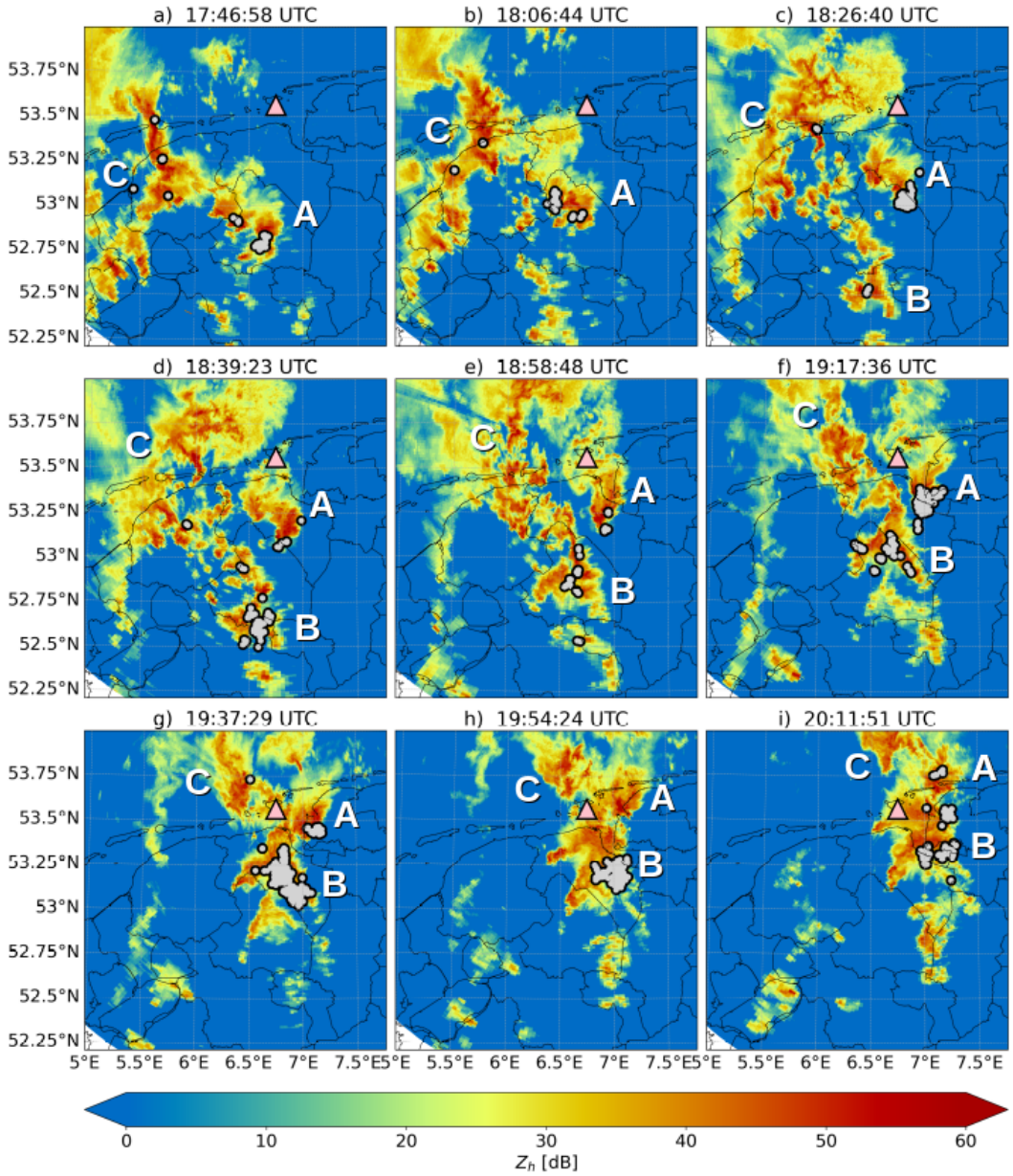

Figure 4.1: Radar reflectivity Z_h (elevation angle 1.5°) and sparkle LOFAR sources projected on the horizontal plane, for all nine LOFAR images.

results are presented in table 4.1 in the H_{40dB} columns. This table also serves as overview of the presence of lightning in sparkles throughout the different LOFAR images.

4.1.1 Convective system A

The precursor of system A emerges around 15:00 UTC in the province of Limburg. At first, it is a solitary unstructured system until it splits when it reaches the city Zwolle around 17:00 UTC. The left-moving convective cell merges with another convective system and seems to trigger more convection. This is the precursor of system C. The right-moving forms system A. About the time of the first LOFAR image, another system showing relatively weak Z_h values, overtakes cell A and they merge. Until 18:30 UTC, system A exhibits some ambiguous splitting behaviour (figure 4.1 panels a and b) after which a left moving-cell diminishes quickly in intensity (figure 4.1 panel e) and the right moving-cell intensifies. Around 19:30, system B approaches system A and system A loses some of its structure at the west side. At the South-East side, the convective systems seems to stay organized until about 20:11 UTC.

4.1.2 Convective system B

On a line extending south of cell A, convection is constantly triggered from about 16:30 UTC onwards. Some of the smaller cells merge with cell A and some die out. Around 18:00 UTC, some 30 km south of convective system A, three convective cells increase in intensity. In the course of half an hour:

- The northern cell remains structured and a single lightning flash is imaged by LOFAR at 18:26 UTC.
- The central cell has split, but the left- and right-mover (now western and eastern cells) stay close.
- South of the central cell, more convection is triggered in its wake.
- The southern cell get unstructured and loses intensity.

Now the central cell increases in intensity and sparkles are imaged at 18:39 UTC above the eastern updraft. Not much later, the northern and central cell connect via a narrow line of high reflectivity. Instead of single updrafts, the system forms a line of convection. After 19:00 when the northern part approaches system A, the convective intensity shifts to the south. It is two updrafts in the centre and in the east side of system B that remain intense and produces many sparkles at 19:37 UTC, 19:54 and 20:11 UTC. At the same time the southern cell is approaching and connecting to the central part of system A. In the west and south part of system

LOFAR time	Convective system A			Convective system B		
	Lightning	Sparkles	H_{40dB}	Lightning	Sparkles	H_{40dB}
17:46:58	Yes	Yes	Right-mover: ∼9.5 km Left-mover: ∼8 km	No	No	
18:06:44	Yes	Yes	left mover: ∼10 km right-mover: ∼7 km	No	No	
18:26:40	Yes	Yes	<13 km	Yes	No	7 km < H_{40dB} <8 km
18:39:23	Yes	No	∼12.5 km	Yes	Yes	∼10.5 km
18:58:48	Yes	3 or 4	11 km < H_{40dB} <14 km	Yes	Yes	∼10.5 km
19:17:36	Yes	Yes	<12 km	Yes	No	∼10.5 km
19:37:29	Yes	Yes	<12.5 km	Yes	Yes	∼12 km
19:54:24	No	No	8 km< H_{40dB} <12 km	Yes	Yes	∼12.5 km
20:11:51	Yes	Yes*	∼10 km	Yes	Yes	10 km < H_{40dB} <12.5 km

Table 4.1: Table listing if the different convective systems exhibit lighting and sparkles at the time of LOFAR images. The H_{40dB} altitude is an estimation of the maximum altitude of $Z_h = 40$ dB values. *These 10 sparkle classifications are dubious.

A, prominent vortices are observed. The updrafts that are associated with such rotation did not produce sparkles at the time of LOFAR images.

4.2 Visual sparkle analysis

In order to understand the relation between the atmospheric conditions and sparkles, radar variables are compared in time and place with the LOFAR data. This has resulted in a number of observations:

(a) Z_h: Sparkles are generally located in or near high 40 dB values at high altitudes, often near a Z_h edge of high reflectivity gradients.

(b) HMC: According to the HMC-algorithm, sparkles are generally found close to the interface between graupel/hail and ice crystals.

(c) Z_{dr}: Where there are sparkles, the volumes of high Z_h values are often outlined with negative Z_{dr} values. Although not unique for sparkles, cloud tops exhibiting sparkles sometimes show high Z_{dr} values where the radar beams cuts through the tropopause.

(d) v_{rad}: Sparkles are generally located near cloud top divergence.

(e) w_{rad}: Sparkles are generally located near $>W_{rad} = 6$ m/s values. These values mainly outline the cloud tops.

For a meaningful interpretation, some key aspects of the radar data were considered:

- The Z_{dr} values are very prone to attenuation. This is very visible at places where the radar beams propagate through high reflectivity values at altitudes below 4 km. This artefact can be expected since at these altitudes, the liquid composition or liquid outer layer can make hydrometeors both very oblete and reflective. Since Z_{dr} is used in the HMC-algorithm, care should be taken in its interpretation.
- The v_{rad} values are masked for Z_h below 0 dB, for no meaningful velocity is measured when there are no particles and thus no reflectivity. In the raw radar data, data points with $Z_h < 0$ dB give misleading -32 m/s v_{rad} values and therefore, these are omitted from all figures.
- The v_{rad} values are prone to aliasing. Therefore, strong contrast in the radial direction between very high (red) and very low (green) values are sometimes aliasing errors. This is not always evident, especially for v_{rad} jumps of approximately half the v_{rad} range.

- w_{rad} represents the variance of the radial wind velocities within a radar resolution volume. Such variance might be large for turbulent conditions, but also for non-turbulent conditions when the velocity gradient is high. Although a large velocity gradient generally leads to turbulence, turbulence is suppressed by stable stratification. As a results, the tropopause may accommodates for large vertical gradients in horizontal wind velocities between the tropo- and stratosphere. In such case, one would find high w_{rad} values in or near the tropopause.

  However, it should also be considered that w_{rad} is computed as a weighted second moment , taking Z_h values for weighting[42]. This means that, within one resolution volume, a sub-volumes of relatively high Z_h values gives a higher contribution to the w_{rad} value. Since there are barely any hydrometeors in the stratosphere, it generally is transparent for C-band radars. As such, when part of a resolution volume is filled by a cloud and the other part by the stratosphere, the cloud contributions should dominate the w_{rad} value.

  Turbulence could potentially mix across the tropopause. In such case, the momentum flux would decrease any previous wind shear.

  In view of these arguments, it can be assumed that, when radar resolution volumes with enhanced w_{rad} values are non-transparent, this indicates the presence of turbulence. Perhaps, when turbulence is decreasing after some cross-tropopause mixing, the w_{rad} radar signal might be enhanced by a non-turbulent wind shear component.

  In practice, many radar images do show increased W_{rad} where the radar beams are expected to intersect the cloud-stratosphere interface. It is not clear to what extent these increased values represent turbulence or cross-tropopause wind shear.

The following exemplary radar images, system A at 19:17:36 and system B at 19:37:29, and corresponding descriptions serve to appreciate the observations listed above.

Convective system A: 19:17:36 UTC

Figure 4.2 shows sparkles projected on the top view of a radar sweep and projected upon a vertical cross-section. From the left panel, it is observed that the sparkles are in or close to the main cloud with high reflectivity values of the 17° radar sweep. A closer look at the cross-section in the right panel shows that many sparkles are ac-

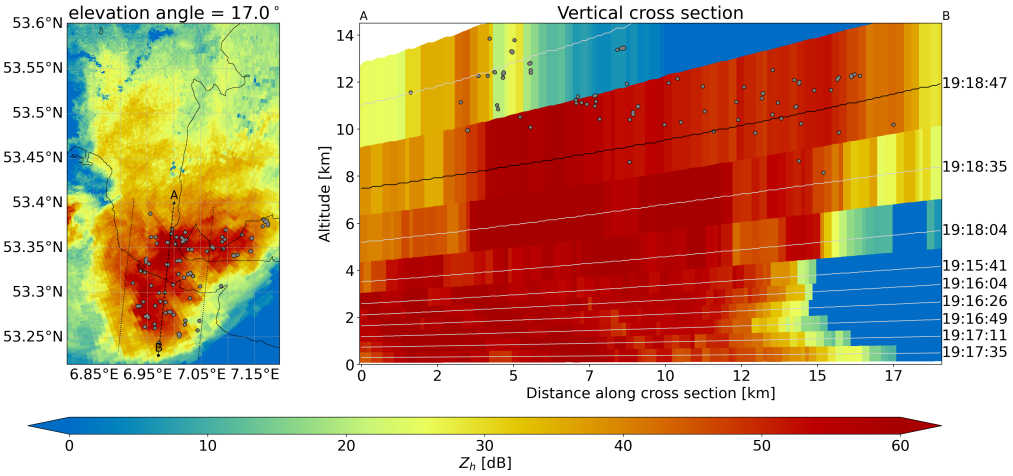

Figure 4.2: Z_h at 19:17 UTC and LOFAR sources classified as sparkles. The left panel shows a top view of an elevation angle of 17° with sparkles projected unto the horizontal plane. The right panel shows a vertical cross-section with sparkles within 5 km projected upon this plane. The vertical cross-section runs along the solid black line in the left panel from A to B. In the right panel, the radar beams from separate radar sweeps are indicated with grey lines, and a black line corresponding to the sweep depicted in the left panel. The actual time of the radar sweeps (mean time for sweep data in cross-section) are indicated at the right-side of the axes.

tually located at significantly greater altitude than the 17° radar sweep, especially when one considers the true radar resolution volume which is much less tall than in the cross-section display. A different cross-section and elevation angle of 25° in figure 4.3 shows that most of the high altitude sparkles are inside the high reflectivity values. Few incidental sparkles, are found at relatively low reflectivity values below 20 dB.

Figure 4.4 shows radar variables Z_h, Z_{dr}, v_{rad}, w_{rad}, as well as the hydrometeor classification (HMC) for system A at 19:17 UTC. The observation of the negative Z_{dr} outline (observation (c)) is present in this figure, but barely visible. The remarkably positive Z_{dr} is illustrated somewhat above the 12 km altitude contour. In the v_{rad} image (panel c) it can be observed that around sparkles, the radial velocity values diverge. This either indicates physical divergence or rotation, when the values diverge in the radial or azimuthal direction respectively. Furthermore, where Z_h values are relatively low, high w_{rad} values in panel d clearly indicate the location where the radar sweep cuts the cloud edge. Around sparkles, the w_{rad} are also enhanced. A final observation from figure 4.4, panel e, is that the HMC-algorithm identifies the radar resolution volume near sparkles either as hail (HL), graupel/hail (GH), or ice crystals (vertically aligned: VC, horizontally aligned: HC). Besides the sparkles that are already present at the GH-ice interface, it is very likely that others are at this

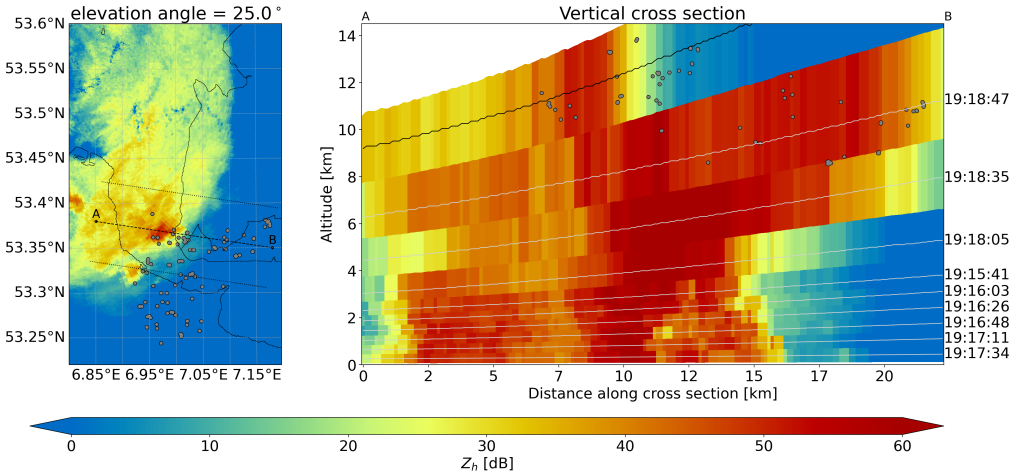

Figure 4.3: Identical to figure 4.2, but for an elevation angle of 25° and another vertical cross-section.

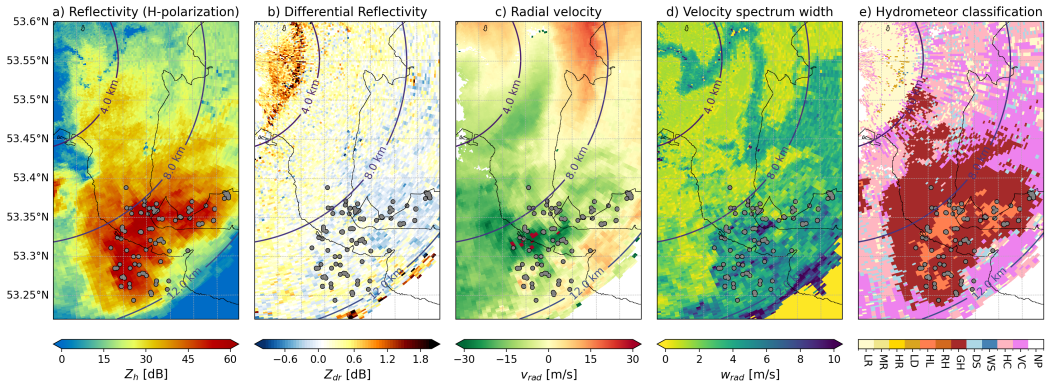

Figure 4.4: Topview of radar Z_h, Z_{dr}, v_{rad}, w_{rad} and HMC, at 19:17 UTC for an elevation angle of 17°. Sparkles are projected upon the horizontal plane. The horizontal extent is equivalent to figure 4.2 and 4.3. Contour lines correspond to the altitude of radar data points in km.

interface as well if the altitude difference between sparkles and resolution volumes are considered.

When one considers the radar images at an elevation 25° as depicted figure 4.5, one finds that the sparkles also exist in areas that the HMC-algorithm has classified as ice crystals or snow. Comparing the W_{rad} image between the 17° and 25° elevation angles, there is a much broader band of relatively high $W_{rad} > 6$ m/s values in figure C.2. Although the low Z_h values could be interpreted as the transparent contribution of the stratosphere, the substantial Z_{dr} values reveal that it is probably not solely the tropopause contribution to the high w_{rad} values. The positive and negative Z_{dr} values could be indicative of ice crystal orientation due to the ambient electric or wind field.

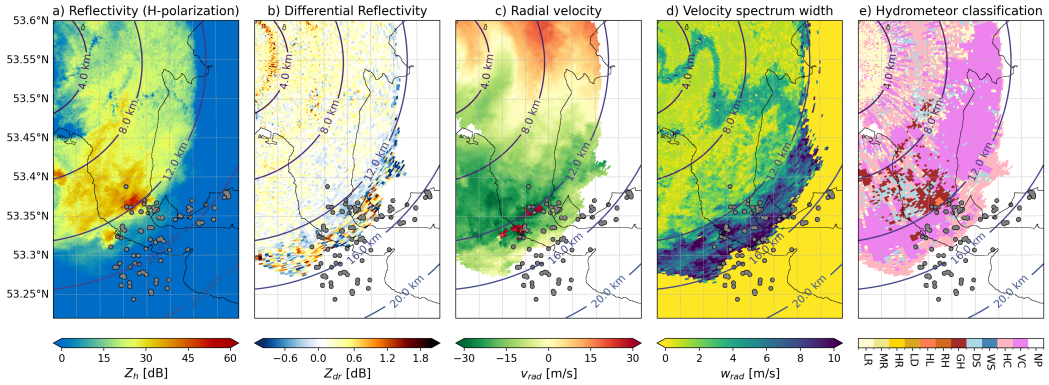

Figure 4.5: Similar to figure 4.4, but for an elevation angle of 25°.

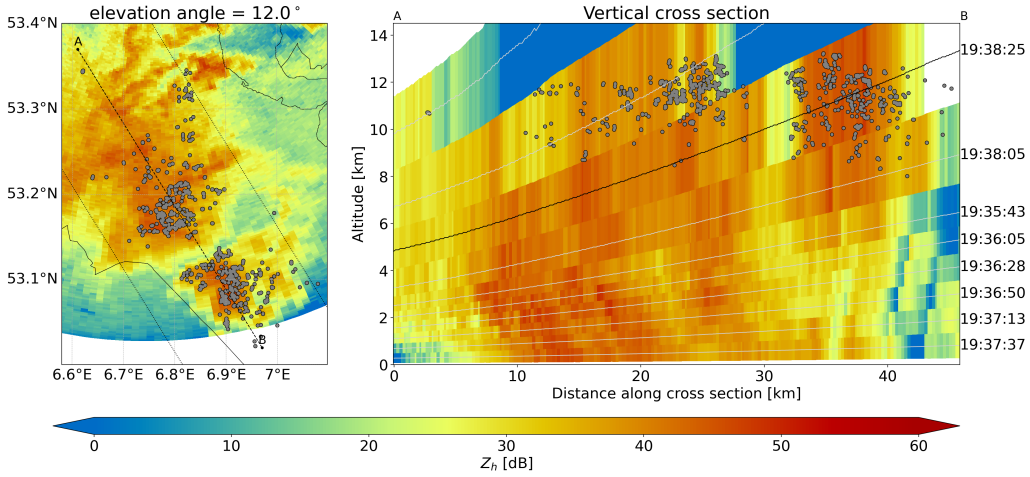

Figure 4.6: Similar to figure 4.2, but at 19:37:29 UTC, with a different horizontal extent that shows convective system B, and with a 12° elevation angle.

Convective system B: 19:37:29 UTC

Figure 4.6 shows sparkles projected on the top view of a radar sweep and projected upon a vertical cross-section. Also in this image, the sparkles are mainly located on top top and surrounding $Z_h > 40$ dB values. Other radar variables Z_{dr}, v_{rad}, w_{rad} are presented in figure 4.7 and 4.8.

Figure 4.7 and 4.8 show radar variables Z_h, Z_{dr}, v_{rad}, w_{rad}, as well as the hydrometeor classification (HMC) for system B at 19:37 UTC. The former with an elevation angle of 12° and the latter with an elevation angle of 17°. In both figures, observation (c) can be observed as the negative Z_{dr} values near the high Z_h values surrounding sparkles. Furthermore, the enhanced w_{rad} values and the graupel-ice (GH-VC/HC) interface are clearly visible near sparkles.

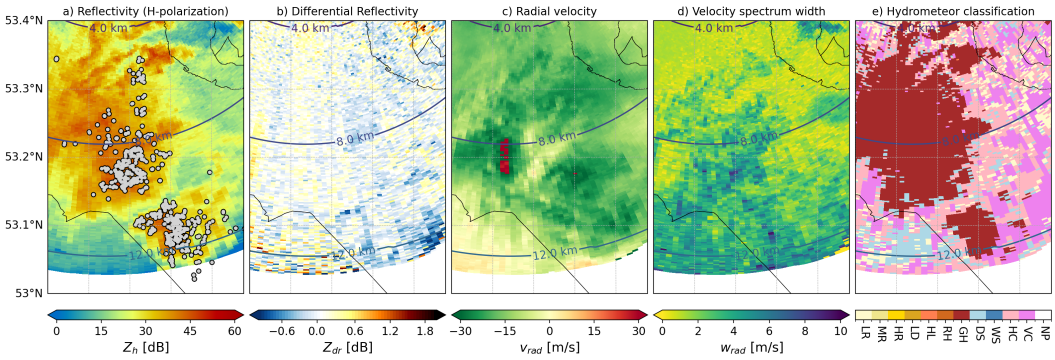

Figure 4.7: Topview of radar Z_h, Z_{dr}, v_{rad}, w_{rad} and HMC, at 19:37 UTC for an elevation angle of 12°. The horizontal extent matches figure 4.6. Sparkles have not been imaged to maintain visibility. It is advised to use figure 4.6 as a reference for sparkle regions.

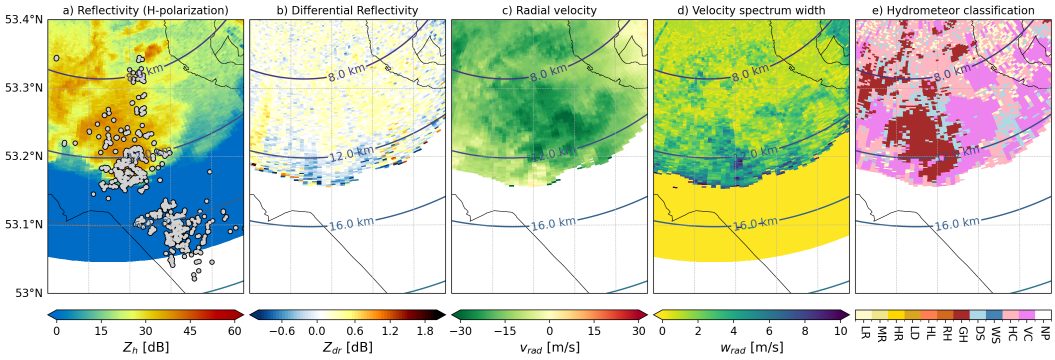

Figure 4.8: Similar to figure 4.7, but for an elevation angle of 17°.

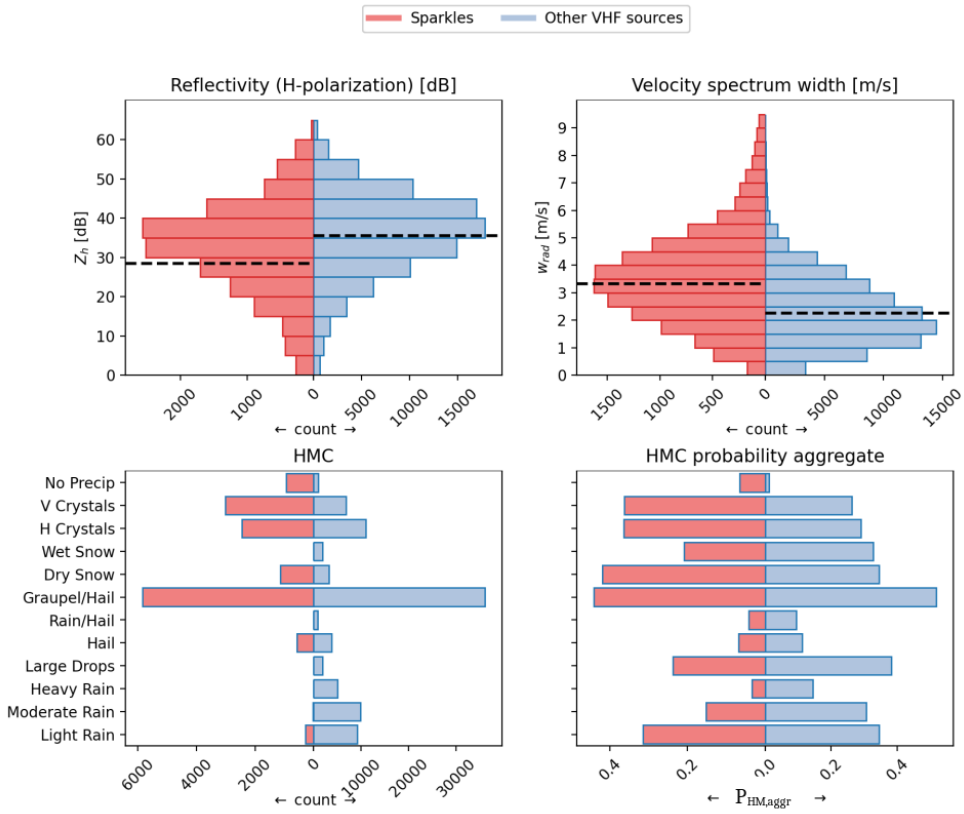

Figure 4.9: Histograms showing the radar data within 2 km from sparkles, and radar data above 8 km within 2 km distance in the horizontal direction. The dashed lines show the means for Z_h and w_{rad}. The HMC probability aggregate is according to section 3.3.

4.3 Quantitative sparkle analysis

According to the method as described in section 3.2, radar variables have been collected within a distance of 2 km with respect to VHF sources classified as sparkles and other VHF sources. For Z_h, w_{rad}, and the HMC-algorithm, the results are presented as histograms in figure 4.9.

From figure 4.9 it can be observed that the Z_h values are generally lower around to sparkles and w_{rad} values are generally higher. Around sparkles, there are generally more ice crystals instead of different rain types. The HMC probability aggregate (section 3.3) shows that the HMC-algorithm has relatively low skill, with high probabilities of dry snow, ice crystals, and light rain.

One problem of the analysis presented in figure 4.9 is that the sparkle data has an altitude bias. After all, the sparkles are defined as VHF sources above 8 km. Other VHF sources are primarily found at lower altitudes. This is even more en-

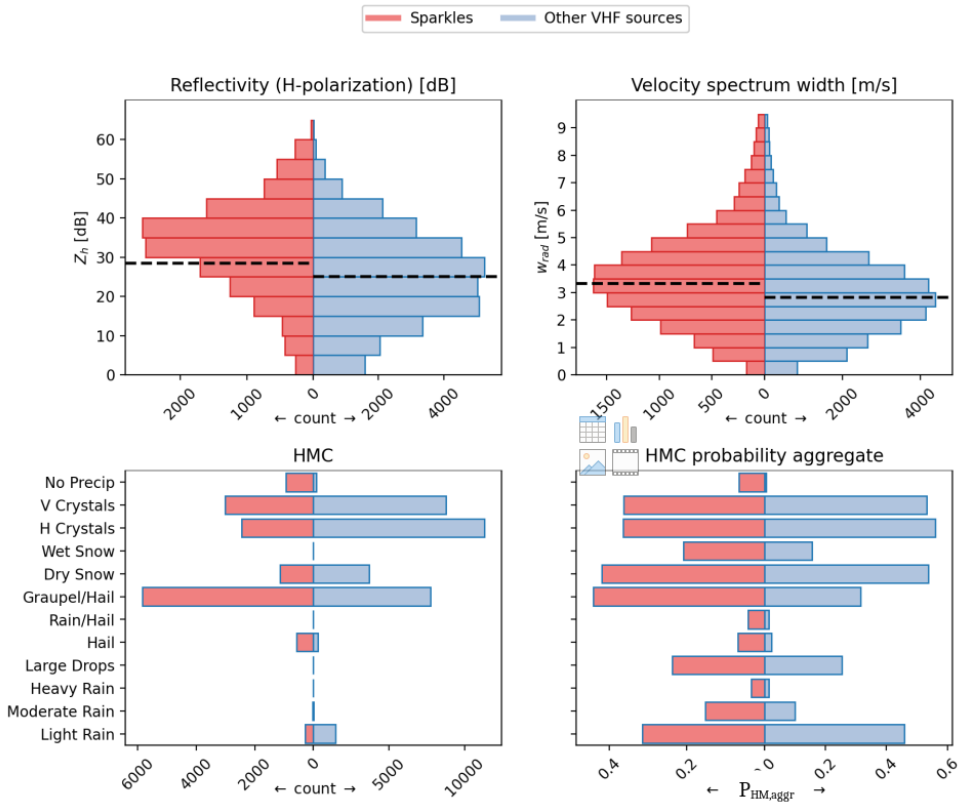

Figure 4.10: Similar to figure 4.9, but the radar data of other VHF sources is selected within 2 km in the horizontal, above 8 km, and Z_h>0.

hanced by the choice to classify radar data near both sparkle and other VHF sources as sparkle data. This is problematic since atmospheric conditions are highly correlated with altitude. In order to eliminate the altitude bias, data is also collected close to other VHF sources in the horizontal plane, but above 8 km. Data with Z_h<0 dB are omitted such that data outside the thunder clouds is ignored. The results are presented in figure 4.10.

Figure 4.10 illustrates that in thunderclouds above 8 km and above lightning, far away from sparkles or if no sparkles, the Z_h values are generally lower and there is less graupel and hail. The difference for w_{rad} is less pronounced than in figure 4.9. One might question if the difference near and far away from sparkles is still significant. The difference in the mean of 0.5 m/s is not large taking into account the variance. However, it should be noted that the tail of extreme values is much higher for radar data near sparkles.

5. Discussion

In the previous chapter, three different types of data were presented:(1) a description of the thunderstorm,(2) a qualitative analysis of the radar data with respect to sparkles, and (3) a quantitative comparison between radar data near sparkles and radar data near other lighting structures. These sources of data have provided information about the sparkles with respect with the storm dynamics, the relative location of sparkles with respect to the clouds, and the characteristics of the clouds near sparkles in terms of the hydrometeors and turbulence. In order to give a more thorough understanding of the thunderstorm dynamics, the first section of this chapter provides an interpretation of the convective systems seen in the radar images. Subsequently all results are combined in order to address the main research question: What charging mechanisms and charge structures lead to sparkles? Successive to this "synthesis" section, Section 5.3 reviews the performance of the sparkles classification algorithm and Section 5.4 discusses the interpretation and limitations of the hydrometeor classification algorithm. Other limitations of this study and an outlook for future research are provided in Section 5.5 and Section 5.6 respectively.

5.1 Convective system characterization

5.1.1 Ambient thunderstorm conditions

Thunderstorms are favoured by higher CAPE values. As can be seen in figure 5.1, the ERA5 at 18:00 UTC shows high CAPE larger than 1000 J/kg throughout most of the Netherlands. At some places, very high values larger than 3200 J/kg are found. The CAPE values increase between 15:00 and 18:00 UTC.

The arrows in figure 5.1 represent the wind at different altitudes according to the ERA5 data. They show much wind shear and veering of the wind direction with height. At 15:00 UTC, the wind veers clockwise with increasing altitude throughout the whole of the Netherlands. At 18:00 and 21:00 UTC (latter not shown) there is barely either no wind veering in the Netherlands, or at least not clockwise. Only at the border with Germany (East) in the provinces of Drenthe and Groningen there is

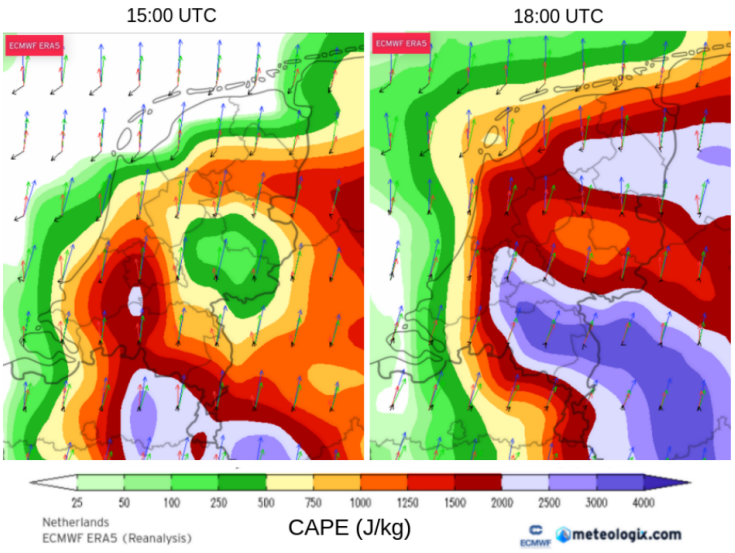

Figure 5.1: Overview of CAPE (shading) and wind (arrows) in the Netherlands on June 18 2021. Wind vectors correspond to 10 m (black), 850 mbar (red), 500 mbar (green) and 300 mbar (blue). Adaption from figures retrieved from meteorologix.com.

some minor veering between the lower (red) and higher altitude (green and blue) wind velocities.

5.1.2 System A: Supercell

Supercells are convective systems that are characterized by a strong, isolated updraft that lives for a long period. By definition supercells exhibit a mesocyclone, i.e. "a region of vertical vorticity with a characteristic width of 3-8 km and magnitude of $O(10^{-2})$ s^{-1}"[46], throughout the lower part of the updraft.

Beside high CAPE values to drive the convection, supercell organisation is the result of low altitude wind, wind shear and clockwise (at the Northern hemisphere) veering of the wind with altitude. More specifically, this combination of factors ensure that the precipitation is not interfering adversely to the low level updrafts, and that cold outflow from downdrafts do not interfere with the inflow of moist warm air. In such a case, a frontal system may develop much like a synoptic scale cyclone. The inflow of warm air would flow in through the warm sector to be carried aloft in the updraft. A schematic diagram is given in figure 5.2.

There are multiple distinct features that are indicative of a supercell. Most of the precipitation is often seen in the "forward flank downdraft" as a elongated area downstream of the updraft (with respect to the ambient wind), and as the "rear flank downdraft" in an area upstream of the updraft. Both features are indicated in

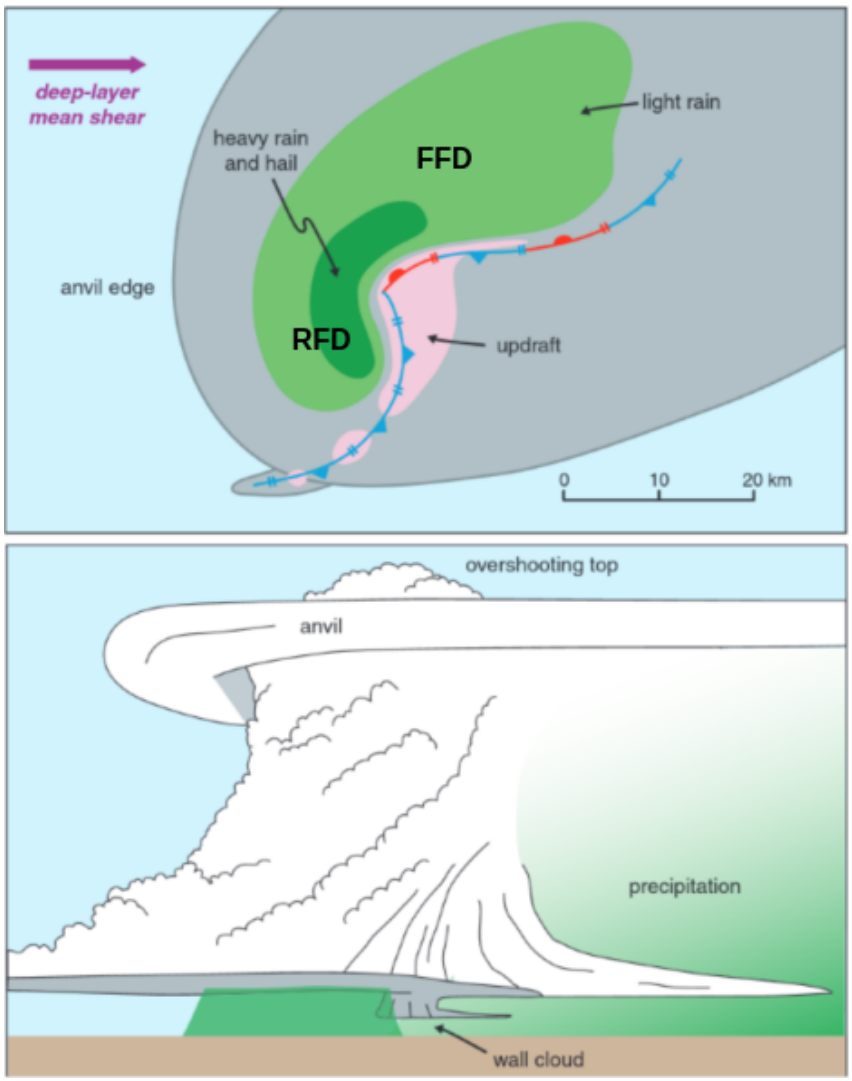

Figure 5.2: Figure from Markowski *et al.* [46] with minor adaption. Schematic diagram of supercell precipitation and cloud structure. Top: horizontal cross section at low altitude. Bottom: visual appearance from the viewpoint of an observer south of the storm. Green colors indicate (heavy) precipitation, grey en white indicates the cloud, pink indicates the updraft, and fronts are indicated by the blue and red barbed lines. The FFD and FFD are indicated by annotations.

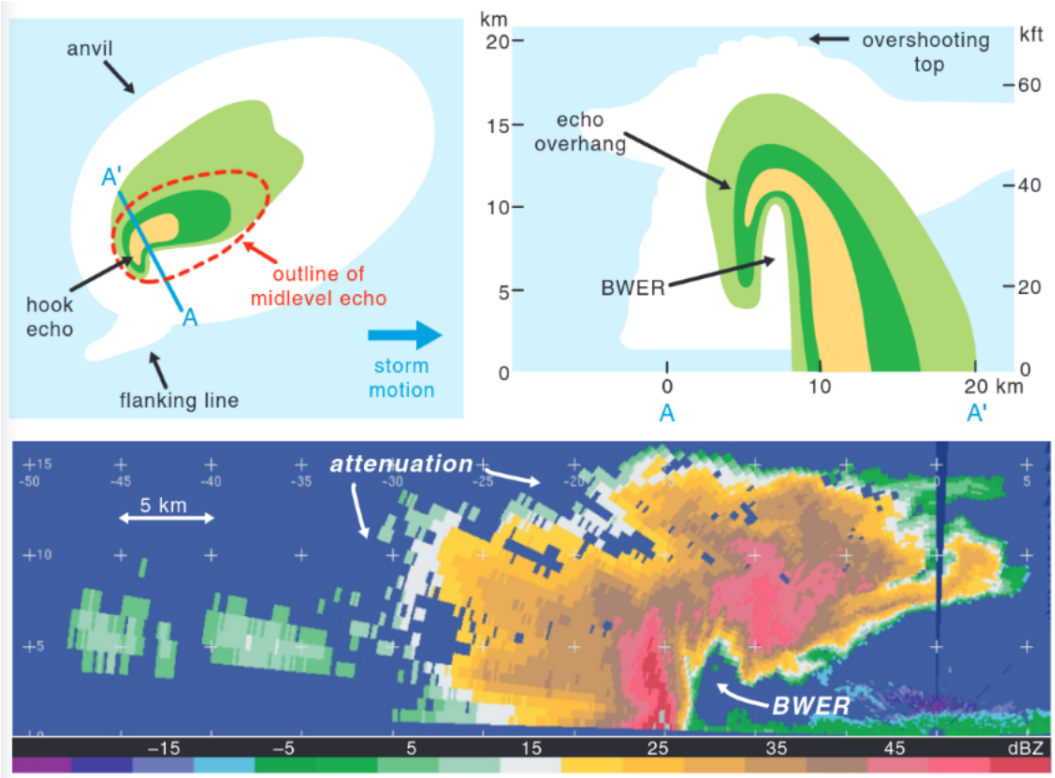

Figure 5.3: Figure from Markowski *et al.* [46] with minor adaption. Top: schematic diagram of a supercell. Green and yellow shading shows moderate and high radar reflectivity values at low altitude (left) and in a vertical cross section fom A-A' (right). White indicates the cloud. Bottom: vertical cross section of radar reflectivity, approximately along the line A-A' of an actual supercell. The Hook Echo and BWER are indicated.

figure 5.2. A "hook echo" and sometimes a "Bounded Weak Echo Region" (BWER) are generally considered as indications of the main updraft location[47]. These features are indicated in figure 5.3.

From about 18:30 until about 19:45, convective system A displays features that are reminiscent of a FFD, RFD, a BWER and a hook echo. This can, for example be seen in figure 5.4, corresponding to the time of the LOFAR image in panel f from figure 4.1. The BWER is even better visible in the vertical cross section in the right panel of figure 5.5. Given these structural features, the sustained high Z_h values, and the sustained high altitude, convective system A could be considered a supercell. The timing and location of system A, matches well with the ambient thunderstorm conditions according to the ERA5 data. The cell is sustained as in the North-East German border where the CAPE values are relatively high, and where there is some clockwise wind veering to promote the organization of convection into a supercell structure.

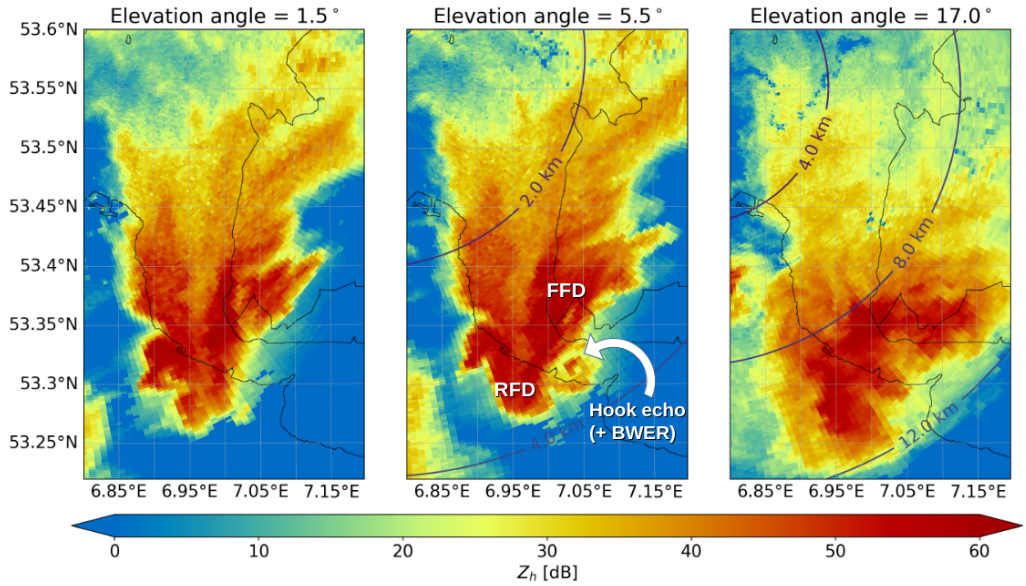

Figure 5.4: Z_h top view at 19:17 UTC for elevation angles of 1.5°, 5.5° and 17°, with annotations of supercell features: FFD, RFD, Hook Echo and the BWER. Contour lines indicate the altitude of radar data.

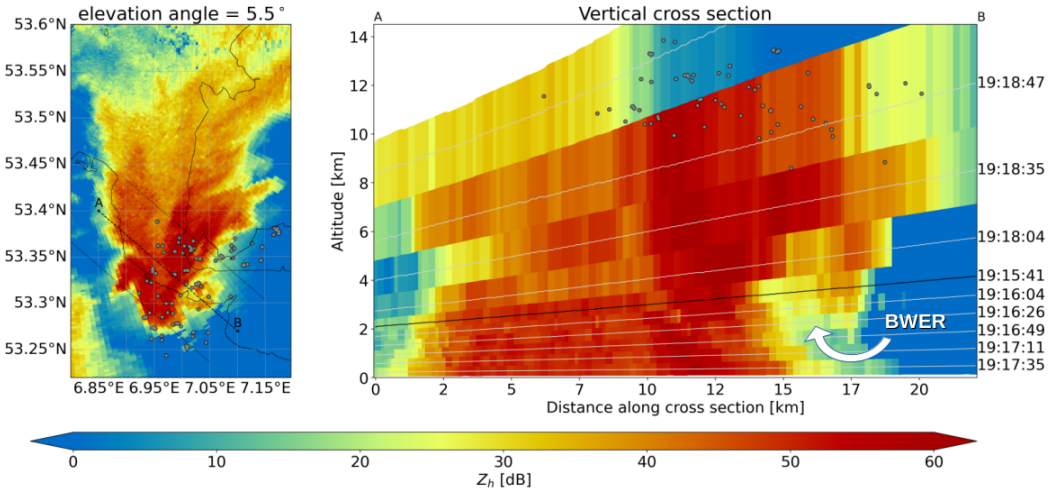

Figure 5.5: Similar to figure 4.2, but for an elevation angle of 5.5° and a different vertical cross-section. The BWER is indicated by an annotation.

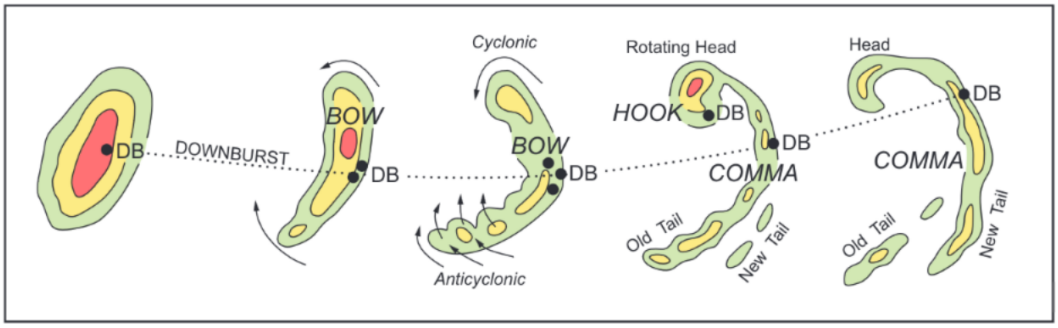

Figure 5.6: Figure from Markowski *et al.* [46]. Schematic diagram of a Bow Echo. Green, yellow and green shading indicates little, moderate and intense precipitation.

5.1.3 System B: Bow Echo

Sometimes convection organizes itself into a large systems with large scale convective updrafts extending over horizontal length scales more than 100 km. An example of such a "mesoscale convective system" (MCS) are so-called "Bow Echos". Bow Echos are associated with MCSs that produce most severe weather and they are characterised by strong rear inflow at mid-altitude, "bookend vortices", i.e. vortices at the ends of the line of convection[46]. The bookend vortices enhance the rear inflow and give the characteristic bow shape that are seen in radar images. A schematic diagram of a Bow Echo, showing the evolution over time, is given in figure 5.6.

From about 18:45 until 20:15, system B sustains a large area of high reflectivity values. Therefore, one could consider this consider system to be a MCS. This observation would also be in line with the ambient thunderstorm conditions as depicted in figure 5.1: there is a large area of relatively high CAPE values matching the approximate location of convective system B and the strong wind shear lacking wind veering would favor MCS organisationMarkowski *et al.* [46].

The large arrow shape and the vortices in the left 'arm pit' (e.g. in figure 4.1 panel h) of convective system B are reminiscent of a Bow Echo. The smaller system that overtakes the main system from south could also be an indication of the strong rear inflow associated with Bow Echos. For the time being, this would be the only known case of a Bow Echo producing sparkles or sparkle-like discharges.

When high elevation radar scans are considered, e.g. figure 4.7, it can be seen that at the South-East end of convective system B has a relatively solitary overshooting cloud top. This raises the question: To what extent is the solitary overshooting top related to the large scale convection associated with Bow Echos? One

possible explanation for the solitary position of the outer updraft could be the vertical wind profile with respect to the position of the MCS. The ambient wind (figure 5.1) at high altitudes is almost parallel to the Western part of the Bow Echo, while the high altitude wind is almost perpendicular with respect to the Eastern part of the Bow Echo. Therefore, the Western part of the Bow Echo can be expected to get a more "parallel stratiform" character, whereas the Eastern part of the system should get a more "leading stratiform" character[46]. This would also explain, for the Western part of Bow Echo, the relatively narrow shape of the upwind and the somewhat messy organisation more downwind (e.g. figure 4.1 panel h). According to Markowski *et al.* [46], leading stratiform MCS allow for more discrete convective cells.

5.2 Synthesis

A remarkable feature that is seen in the radar data, are the nonzero Z_{dr} values at sparkle altitudes. Especially surrounding the high reflectivity values, the Z_{dr} values are negative. Such negative values might be indicative of vertically aligned ice crystals, possibly due to the ambient wind or an electric field. The latter case is thought to be of transient nature, only preceding lighting flashes on short notice[47]. In the case of a cloud in sparkling mode, producing very high flash densities and frequencies, one may expect continuous realignment of the ice crystals. The negative Z_{dr} values may therefore very well be a signature of sparkles. Unfortunately, the Z_{dr} values are very prone to attenuation. Generally the values are biased negatively beyond high Z_{dr} values. This makes one wonder if the negative Z_{dr} values around sparkles are physical or just neutral Z_{dr} values that suffer from attenuation effects. A more thorough analysis of the attenuation and possibly an algorithm to correct for it, could help to interpret the radar Z_{dr} images.

Placing the results in the context of previous work, many similarities are found. The correlation between high reflectivity values and sparkle-like lightning is also found by Emersic *et al.* [9], Calhoun *et al.* [10] and MacGorman *et al.* [2]. They find these discharges to form a cap on top of a very strong, high reaching vertical updraft. With the radar data presented in this work, lacking in vertical resolution and not containing 3D wind velocities, it is not possible to unambiguously determine the relative location of sparkles with respect to the updraft. However, from the radar images showing at high altitude both high Z_h values, likely caused by grau-

pel, and divergence, it is at least likely that the sparkles are above updrafts. In cell A, the relative location of the BWER and hook echo indicating the main updraft, is also in line with the sparkle location aloft.

All of the aforementioned studies also specifically associate the sparkle-like discharges with intensifying updrafts. The triggered nature of LOFAR lightning imaging complicates a temporal analysis of sparkles. However, if the convective systems that produce sparkles are indeed a supercell and a squall line, it may be expected that updrafts were intense. The story line of the mesoscale events in convective system A, such as the development of a BWER, hook excho, FFD and RFD, would suggest intensifying updrafts around the time of sparkles. For convective system B, there are little mesoscale features that indicate the updraft intensification near sparkles. If any, the sparkles are clearly not above the left bookend vortex and perhaps above a barely visible right bookend vortex.

A reason to believe that the sparkles are indeed located at the top of updrafts is the radar signature of graupel near sparkles. These are only to be found at these altitudes if lofted upward by strong updrafts. If the updraft strength were to be subsided before the sparkle observations, the graupel could be expected to have fallen to lower altitudes. Following this line of reasoning, one would expect that for the strong updrafts associated with sparkles, graupel is brought close or into the turbulent, diverging overshooting top.

With the presented data, it is not possible to prove a causal relation between graupel and sparkles. It may that graupel at sparkle altitudes is only a by-product of the true mechanism. In this case the sparkles might originate, as suggested in previous studies, from the interaction between a cloud of (charged) ice crystals and a charged screening layer. The data in this thesis is inadequate to evaluate this hypothesis.

If there were to be a causal relation between graupel in the overshooting cloud top and sparkles, four mechanisms may be considered:

5.2.1 Graupel and charged screening layer

As suggested by Calhoun *et al.* [10], the proximity of the charged screening layer and the lofted charge in the updraft might enhance the electric field. The necessity for graupel might be rooted in the charge density or the sign of the charge carried by graupel. In the latter case, it would not be surprising if the charged screening

layer above the graupel volume, would be oppositely charged to the graupel in proximity.

This hypothesis leaves two options to explain the small spatial extent of sparkles. Perhaps the micro-physical conditions of the charged screening layer do not allow for full leader breakdown. Alternatively, charge pockets might be formed by the turbulent conditions in overshooting tops and the periodic folding of the screening layer, as suggested by MacGorman *et al.* [2].

5.2.2 Variability in charge sedimentation

Turbulent conditions in an overshooting cloud top can have multiple effects on the charge structure. One possible explanation for charge pockets is related to charge sedimentation. This is a necessary condition for cloud charging. If strong updrafts loft graupel particles into the more turbulent regions of the cloud top, the velocities are much more variable at the altitude of sedimentation. This would, in turn, lead to variability in charge sedimentation and thus charge pockets. Such a variable charge sedimentation would be most pronounced at the top of the graupel tower, for at the sides the replenishment by overhead graupel sedimentation reduces net charge separation.

5.2.3 Graupel-ice mixing

It is also conceivable that charge pockets are the result of mixing volumes of graupel and ice in the turbulent conditions in an overshooting cloud top. This mechanism is somewhat contradictory in the sense that turbulent mixing would also blur the graupel-ice interface and thus neutralizes the charge separation. Therefore, this mechanism would be most fruitful for eddies with large length-scales.

5.2.4 Inductive charge separation

Although the non-inductive charge separation is generally considered to dominate the charging mechanism in thunderclouds, inductive charging might be considerable in certain conditions. Mareev *et al.* [19] shows that for turbulent conditions, an ambient electric field and hydrometeors of adequate size, the inductive mechanism may lead to exponential growth of the electric field. Taking into account the variability of the velocities in the turbulent overshooting cloud top, the inductive charging mechanism has the potential to quickly charge local pockets of much turbulence. Since the inductive mechanism relies on increased particle rebounds, this

mechanism is not so much increased by large eddies, as it is by small eddies.

5.3 Sparkle Classification

Given the clustering parameters as presented in Section 3.1, the results of the sparkle classification algorithm provide satisfactory results. An example in figure 5.7 shows that all sources from the "confetti cloud" are classified as sparkles. A thorough inspection shows that many of the classified sparkles, although separated in time, seem to be clustered in space. Moreover, classified sparkles are sometimes near, on top of, or extending from large lightning structures. Therefore, one could question if the sparkle classifications are always justified. Such a judgement depends, however, on a definition of sparkles in terms of spatial and temporal separation that is, at the present time, not formally defined.

There are other situations where the classification algorithm clearly does not perform well. Such an example is illustrated in figure 5.8 showing a zoom-in of the LOFAR image at 19:17 UTC. The algorithm has classified quite some VHF sources as sparkles that, upon visual inspection, are part of larger lightning structures. The reason for such misclassification lies in the low spatial density of VHF points. This may occur in LOFAR data for dart leaders or positive leaders.

It is expected that in clouds with high sparkle density, such as in figure 5.7, false positives will have marginal effect on a radar analysis. If one does not employ an unrealistic precision for radar data selection, the high density of true sparkles would mark most of the relevant volume anyway. For volumes with no or few sparkles, false positives might obscure or bias the radar analysis. Whereas a quality assessment of sparkle classification is taken into account in the consecutive qualitative radar analysis, the bare output of the sparkle classification is used in the quantitative analysis in section 4.3.

5.4 HMC algorithm

The results that are obtained from the HMC algorithm are presented in previous subsections. This subsection serves to review on the interpretation and the performance of the algorithm.

Per radar resolution volume, probabilities are calculated for all hydrometeor types. As explained in section 3.2, the performance of the fuzzy logic HMC-algorithm

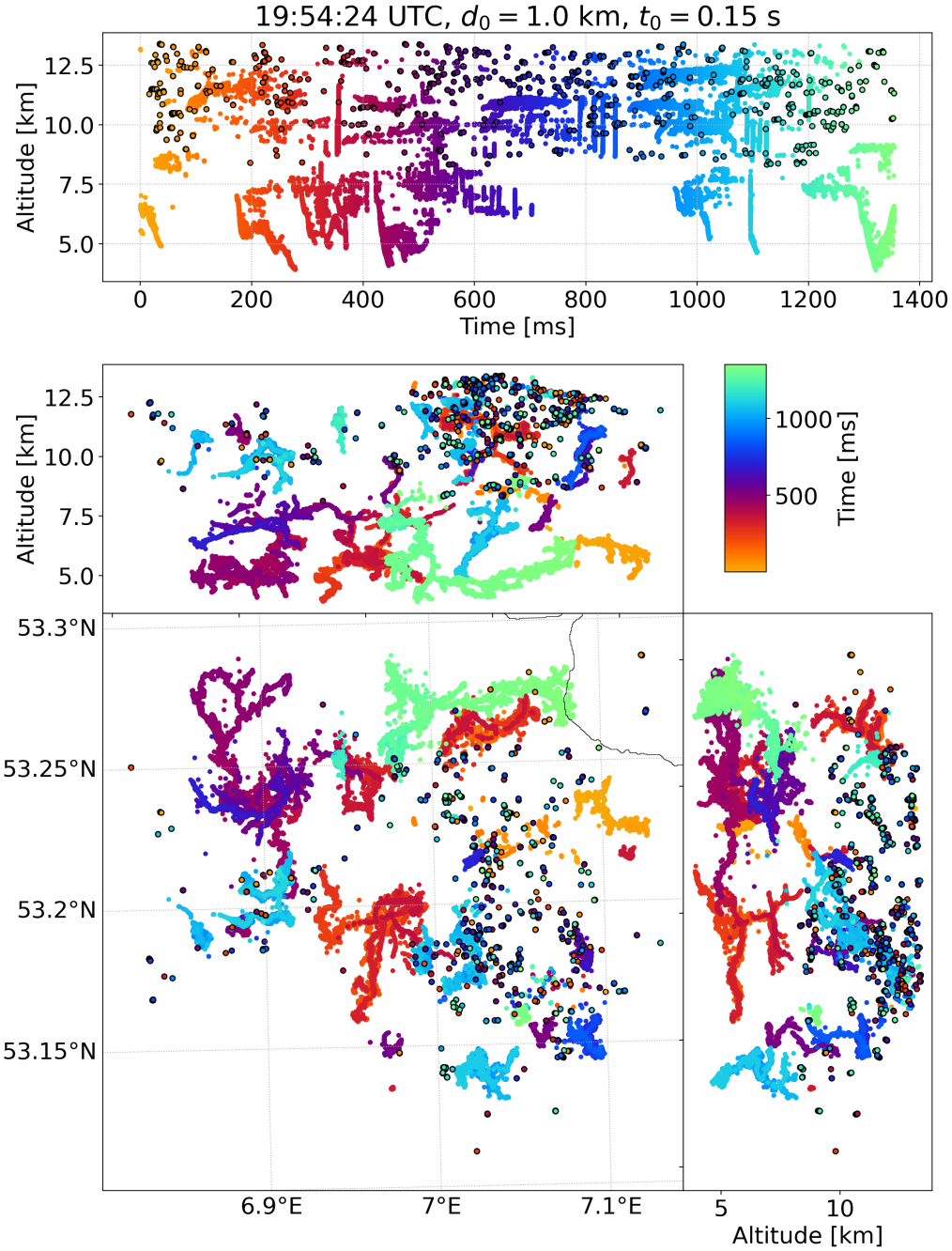

Figure 5.7: LOFAR VHF sources at 19:54 UTC. Colours indicate the time with respect to the image reference time. Sparkles as classified with the algorithm described in section 3.1 are emphasized. In the upper panel, the altitude is plotted versus the time. The other axes display projections of sparkles on the horizontal and two orthogonal planes. d_0 and t_0 correspond to the parameters that are used in the clustering algorithm.

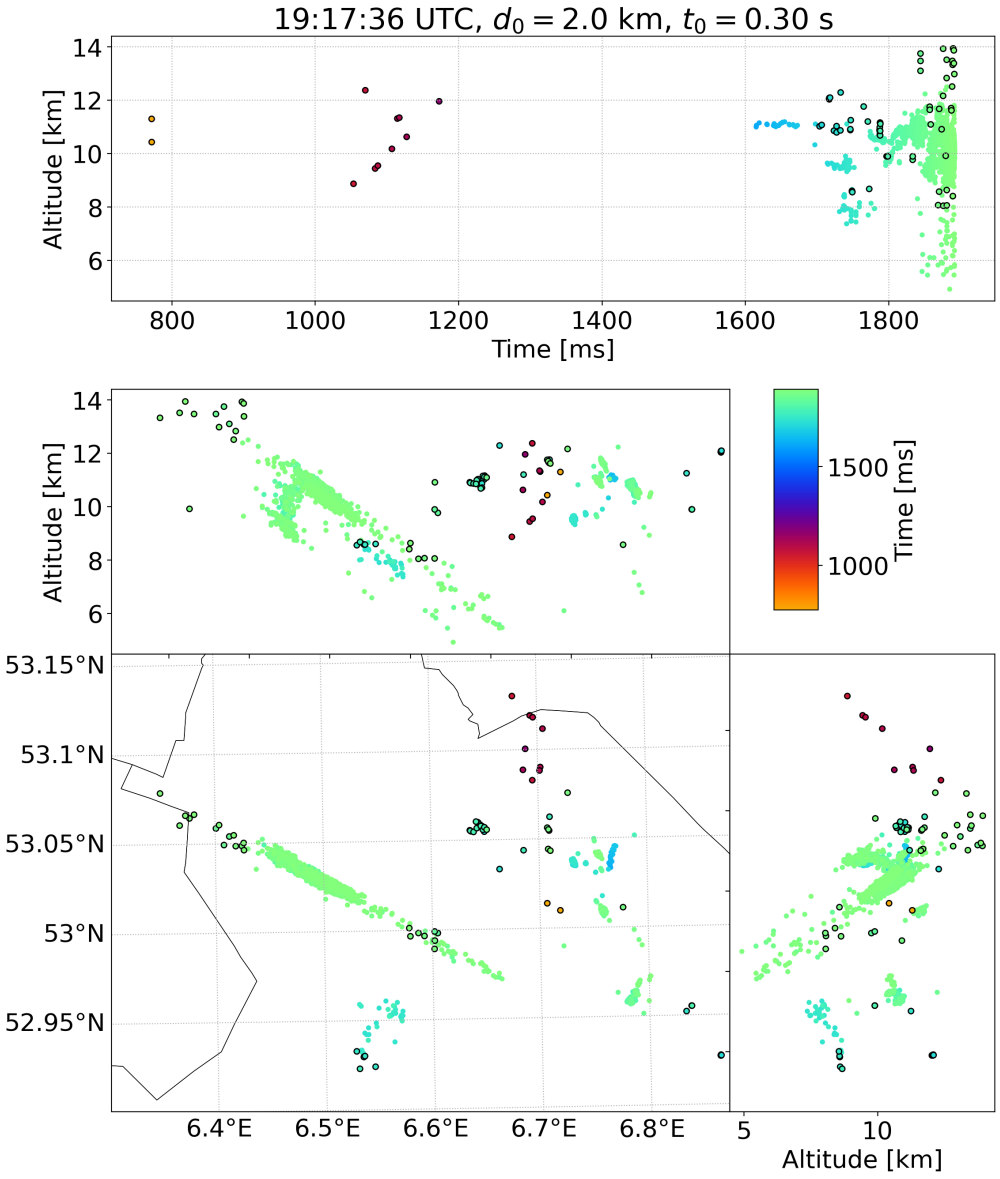

Figure 5.8: Similar to figure 5.7, but at 19:17 UTC, and for a different horizontal extent. Most VHF sources in light green, i.e. after 1800 ms, seem to be part of a large lightning structure and thus the sparkle classifications after 1800 ms are erroneous.

are quantified with two metrics. Namely the maximum HM probability and the probability aggregate as defined in equation 3.4. The former can be found in the figures in appendix C. The latter are given figures 4.9 and 4.10.

Although the maximum HM probability shows to be relatively high (>0.6) in most areas of interest, the results could use some nuance. Figures 4.9 and 4.10 show that the graupel and dry snow probabilities are very close. This artefact probably arises because of the overlapping areas of the Z_h and T membership functions.

Another aspect to consider for the fuzzy-logic algorithm is that is was originally designed for S-band radars. These radars are less prone to Mie scattering effects because of their larger wavelength [48]. Although previous work[44], [49], [50] has shown the potential of fuzzy-logic algorithms for C-band radars, attenuation remains a source of error. In this thesis, no attempt is made to compensate for attenuation. Qualitative review of the data shows that the attenuation is non-existent or at least minimal for the altitudes of sparkles. For better results, one may attempt a more complicated scheme to weight the different polarimetric variables, along the lines of Dolan *et al.* [51].

Another means to improve the HMC-algorithm is to incorporate the linear depolarization ratio. This could not only help in discriminating between different types of ice, it may also be used for identification and quantification of the riming process.[52] If a radar does not provide this variable, one may attempt to retrieve a similar metric using the method as suggested by Ryzhkov *et al.* [52].

5.5 Limitations

There are multiple limitations to this study and these are discussed in order to appreciate the results accordingly. Limitations of the sparkle classification algorithm and HMC algorithm are discussed in previous subsections 5.3 and 5.4.

A major shortcoming of the radar data from the Borkum radar is the vertical resolution. At sparkle altitudes, this may exceed 4 km. Therefore, the estimations of the cloud tops are very coarse and it remains a question to what extent the sparkles live in or on top of the clouds.The reason for the poor vertical resolution is mainly the scanning strategy, taking steps from 5.5°, to 8°, to 12°, to 17, to 25°. The beam width of 1° also increases the resolution when further away from the radar, but this is only limited to 1.7 km at 100 km from the radar.

Another artefact of the Borkum radar are the significant attenuation effects of the C-band radar frequency. Especially at low levels, were large oblate water may be expected, attenuation is clearly visible. For example in the Z_{dr} values. In this study, no attempt is made to correct for the attenuation effects. The algorithm to compute the K_{dp} values, found significant δ values. Especially at low altitudes. This is a clear indication that Mie scattering also played a role. This should be taken into account when attempting attenuation correction.

The v_{rad} values of the Borkum radar proved to be prone to aliasing. Although not impossible with careful examination and some imagination, the aliasing has complicated the analysis of the mesoscale dynamics.

The quantitative results presented in section 4.3 were collected within a 2 km range of the sparkles as given by the sparkle classification algorithm. This seems like a reasonable range given the vertical radar resolution, uncertainties in the advection of the radar data, and the maximum time difference of 2.5 minutes between radar and LOFAR data.

Although 2.5 minutes is short on the timescale of a mesoscale convective system, it is not sure if this also holds for the cloud top. According to high resolution radar data presented in MacGorman *et al.* [2], the cloud top height may vary in the order of a kilometer over a time span of 2 minutes. The same data also shows that this variability has a zero-mean over more than 3 hours. Comparison among the different case studies by MacGorman *et al.* [2], shows mixed results on the persistence of cloud tops. What is evident from these cases, is that maximum altitudes of high Z_h contours are much more variable throughout time than lower Z_h contours. This may be expected as the large hydrometeors that are causing the high Z_h values are quicker to fall when updrafts subside.

It is to be concluded that there is some unknown uncertainty due to the time difference between LOFAR and radar data. For this reason, no attempt is made to quantify the uncertainty that is bound to the 2km radial distance.

5.6 Outlook

Given the data presented in this study, attenuation correction could seriously improve the results. This may not only improve the results of the HMC algorithm, it could also disclose to what extent the observed Z_{dr} values near sparkles are at-

tenuation effects. Moreover, Z_{dr} and K_{dp} columns that are now barely visible by attenuation, could be used to get more insights into the evolution of updrafts[53].

Another way to improve the understanding of sparkles with the current data, would be to delve into the literature about the cloud tops of supercells and bow echos. This thesis has given some compelling evidence that convective system A and convective system B could be classified as a supercell and bow echo respectively. However, a thorough study of the consequences of these observations is lacking.

Although the radar data of this study does indicate enhanced turbulence around sparkles, it is not possible to determine to which extent any of the hypothesized mechanisms (Section 5.2) are actually important for sparkles. There are a number of possibilities for further research to test the credibility of the hypotheses.

Firstly, one could consider to use high resolution atmospheric simulations of dynamics in the overshooting cloud tops and the spatial distribution of hydrometeors. The variability in the hydrometeor distribution and the responsible mechanisms, mixing at the interface or variability in the sedimentation, might elucidate the credibility of hypothesis 5.2.3 and 5.2.2. In such case one should at least incorporate hydrometeors with the dynamical properties of graupel and ice crystals. If possible, one should assimilate radar data into this atmospheric model along the lines of Calhoun *et al.* [37]. This would allow to compare the position of sparkles with the hydrometeor distribution.

Secondly, one could use a similar atmospheric simulation as described above, but including electrification and discharge effect. In order to test the credibility of hypothesis 5.2.4 one should at least incorporate the inductive charging mechanism. Perhaps along the line of Mareev *et al.* [19]. In order to get a correct ambient electric field for the inductive charging mechanism, it may be important to model the whole convective system including the non-inductive charging mechanism along the lines of Calhoun *et al.* [37].

Thirdly, high resolution radars data might be of use to test the variability of hydrometeors, turbulence and the position of the screening. Mobile radars such as used for the analysis in MacGorman *et al.* [2] may be used to collect the appropriate data. Especially with regard to attenuation effects of the intense precipitation in the RFD and FFD, one could position the radars appropriately. Since the sparkles are associated with cloud tops, one should also consider to focus the radar on this area.

Perhaps an advection scheme in combination with cell tracking software, based on operational radar and satellite images, could predict the location of overshooting clouds. This might serve as a target for a cloud top radar scanning strategy to get the best possible resolution.

6. Conclusions

This study attempts to increase the understanding of the mechanisms that lead sparkles, i.e. intermittent, small scale lightning discharges near the top of thunderstorms. Radar data was analysed with respect to these discharges.

On June 18 2021, thunderstorms crossed the Netherlands, and the Low Frequency Array of radio antenna's imaged lightning from these storms with a very high resolution (1 m and 0.1 ms). The lightning images reveal small-scale discharges, up to a few hundred meters in size, occurring intermittently throughout time and at relatively high altitudes near the tropopause. Scholten *et al.* [8] name this type of discharges sparkles.

Previous observations of sparkle-like discharges include blue optical events (BLUEs) and discharges observed by a Lightning Mapping Array in Oklahoma USA. Studies have associated these discharges with intensifying updrafts and hypothesize that they result from interaction with a charged screening layer above the cloud tops, pockets of charge limiting growth into larger structures, lower threshold for electric breakdown at high altitudes, accumulation of charged graupel at the top of updrafts, and formation of charge pockets through due to the variability of graupel charging in the updraft.

In this study, radar data from a meteorological radar on Borkum Island (Germany), was analysed with respect to sparkles, both visually and quantitatively. Radar data was advected to match the timing of LOFAR images. An algorithm was developed to create vertical cross sectional images of the radar data, and a fuzzy logic hydrometeor classification (HMC) algorithm based on 2D membership functions determined the type of hydrometeors in the atmosphere. A clustering algorithm distinguished sparkles from other lightning structures, allowing for a computational comparison between polarimetric radar data in proximity to sparkles and radar data in proximity of other lightning structures.

This study appears to be the first to computationally differentiate sparkles from other lightning structures, making the quantitative radar analysis of different lightning types based on polarimetric data a novel contribution to lightning science.

Consistent with previous studies, sparkles were found near high reflectivity values at thigh altitudes. Such values are associated with strong updrafts, lofting of graupel, and overshooting cloud tops. The HMC confirms that sparkles are often found near volumes of graupel. Near sparkles, high values of the radar velocity spectrum width indicate turbulent conditions. These finding imply that the combination of turbulent conditions and graupel lofted by strong updrafts into the overshooting cloud top, may be important factors for sparkles. Negative radar differential reflectivity values often outlined clouds with sparkles, potentially indicating vertical ice crystal alignment due to ambient electric fields. This may be a signature of the cloud charge, but uncertainty about attenuation effects prevent an unambiguous interpretation.

The two convective systems that produced sparkles resembled a supercell and a bow echo. The latter is the first known case of a bow echo to produce sparkles or sparkle-like discharges. It remains unclear what this finding could indicate about the nature of sparkles.

Multiple mechanisms are suggested that could explain charge pockets in cloud tops. Beside the existing hypotheses of the interaction with a charged screening layer, turbulent mixing of the updraft, and graupel accumulation at the top of updrafts; it is hypothesized that (1) turbulent conditions at the cloud tops could lead to variability in the graupel sedimentation and thus the charge distribution, and (2) that turbulent conditions might lead to charge pockets through the inductive charging mechanism.

In order test the credibility of the proposed mechanisms,it is recommended to use high-resolution atmospheric simulations of thunderstorms that incorporate turbulence, and various hydrometeors such as graupel and ice crystals. Building on the simulations by Calhoun *et al.* [37], one should (1) explicitly model the inductive charging mechanism in the context of turbulence, perhaps following the approach of Mareev *et al.* [19], and (2) increase the model resolution near cloud tops. Additionally, high resolution radars, perhaps focusing on the overshooting cloud tops, would also be valuable in combination with future LOFAR observations.

Acknowledgements

I would like to conclude my master's thesis and, with it, my master's in Climate Physics by expressing my gratitude to all those who have supported me along the way.

This research project would not have been possible without the guidance of my supervisors, Jelle Assink, Aarnout van Delden, Hidde Leijnse and Michiel Baatsen. I am especially grateful for to my daily supervisor at the KNMI, Jelle Assink, for getting me on board and his insightful feedback. Aarnout van Delden has provided the optimism and encouragement to work through my project, and I appreciate his flexibility and input throughout the project. Hidde Leijnse has been a critical pivot throughout the project to nudge me in the right direction. Finally, Michiel Baatsen's storm chasing photographs and expertise about thunderstorm dynamics have been a great source of inspiration.

A special thanks goes to Brian Hare and Olaf Scholten, for providing the beautiful LOFAR images that were central to my thesis. Their patience and enthusiasm regarding my work has been invaluable, and I look forward to future collaborations.

I would also like to acknowledge Jan Fokke Meirink and Nathalie Theeuwes, two colleagues at the KNMI, for their assistance with satellite and HARMONIE data, respectively. Although this data was not included in the final report, their help was very valuable and greatly appreciated.

My colleagues at the KNMI, especially those in the seismology department, have made the past year enjoyable and welcoming. Therefore, I am thankful for their hospitality.

OOn an organizational level, I am grateful to the KNMI and ASTRON for facilitating this research project. The KNMI by providing a workplace, and ASTRON generously provided the LOFAR data.

Finally, I would like to express my gratefulness for the support of my beloved parents, sisters, friends, and of course, my dear lover Danielle. Without you, I

would not be the happy person that I am.

Appendices

A. LOFAR images

This part of the appendix serves so give an overview of the sparkles with repects to other lightning structures.

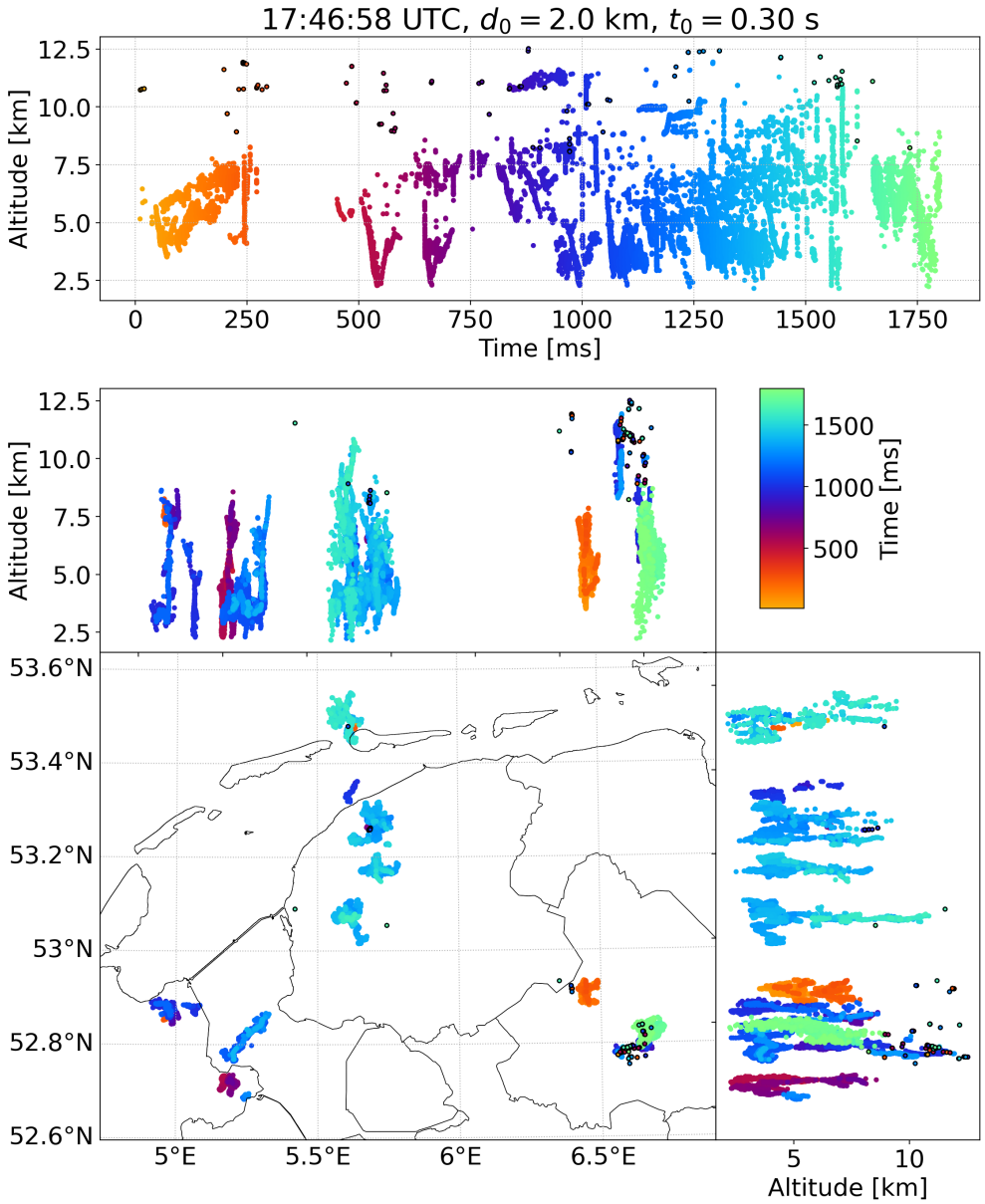

Figure A.1: LOFAR VHF sources at 17:46:58 UTC. Colours indicate the time with respect to the image reference time. Sparkles as classified with the algorithm described in section 3.1 are emphasized. In the upper panel, the altitude is plotted versus the time. The other axes display projections of sparkles on the horizontal and two orthogonal planes. d_0 and t_0 correspond to the parameters that are used in the sparkles classification algorithm.

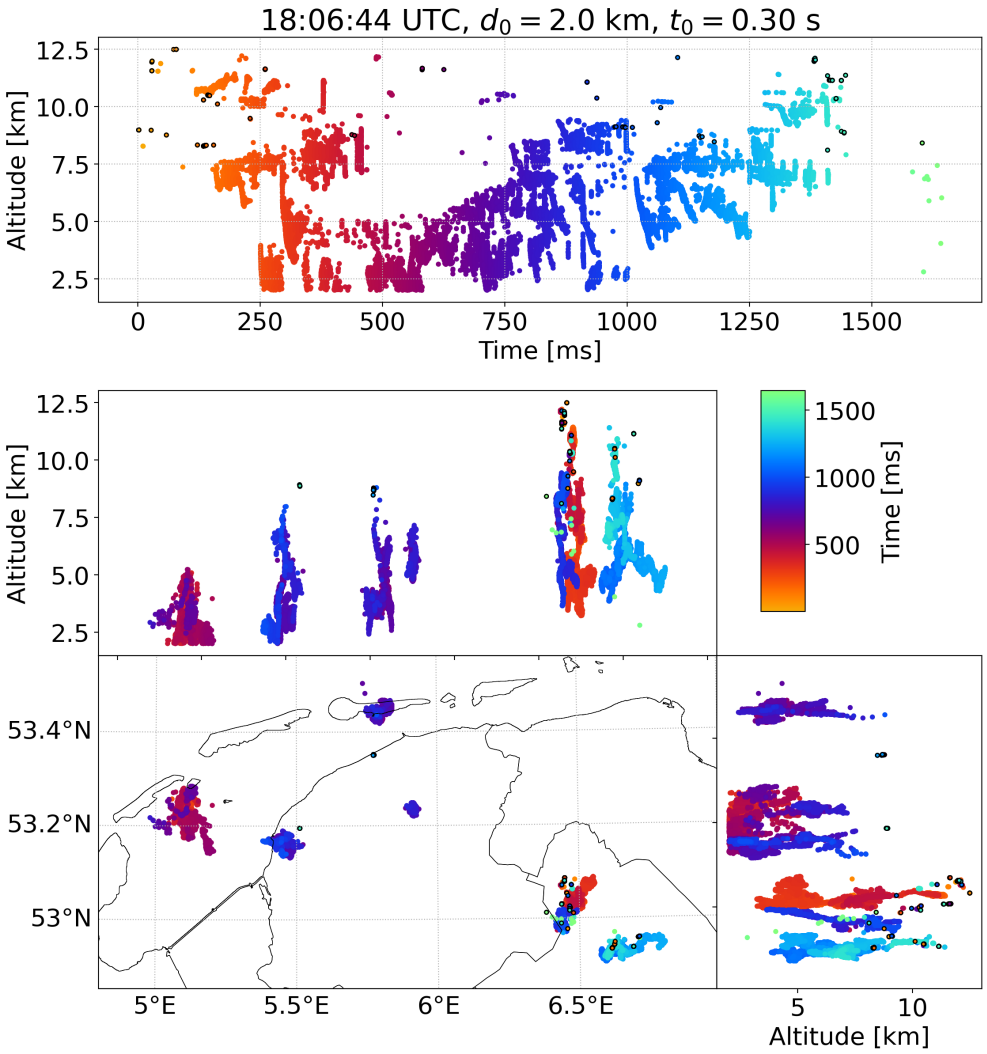

Figure A.2: Similar to figure A.1, but for the image at 18:06:44 UTC.

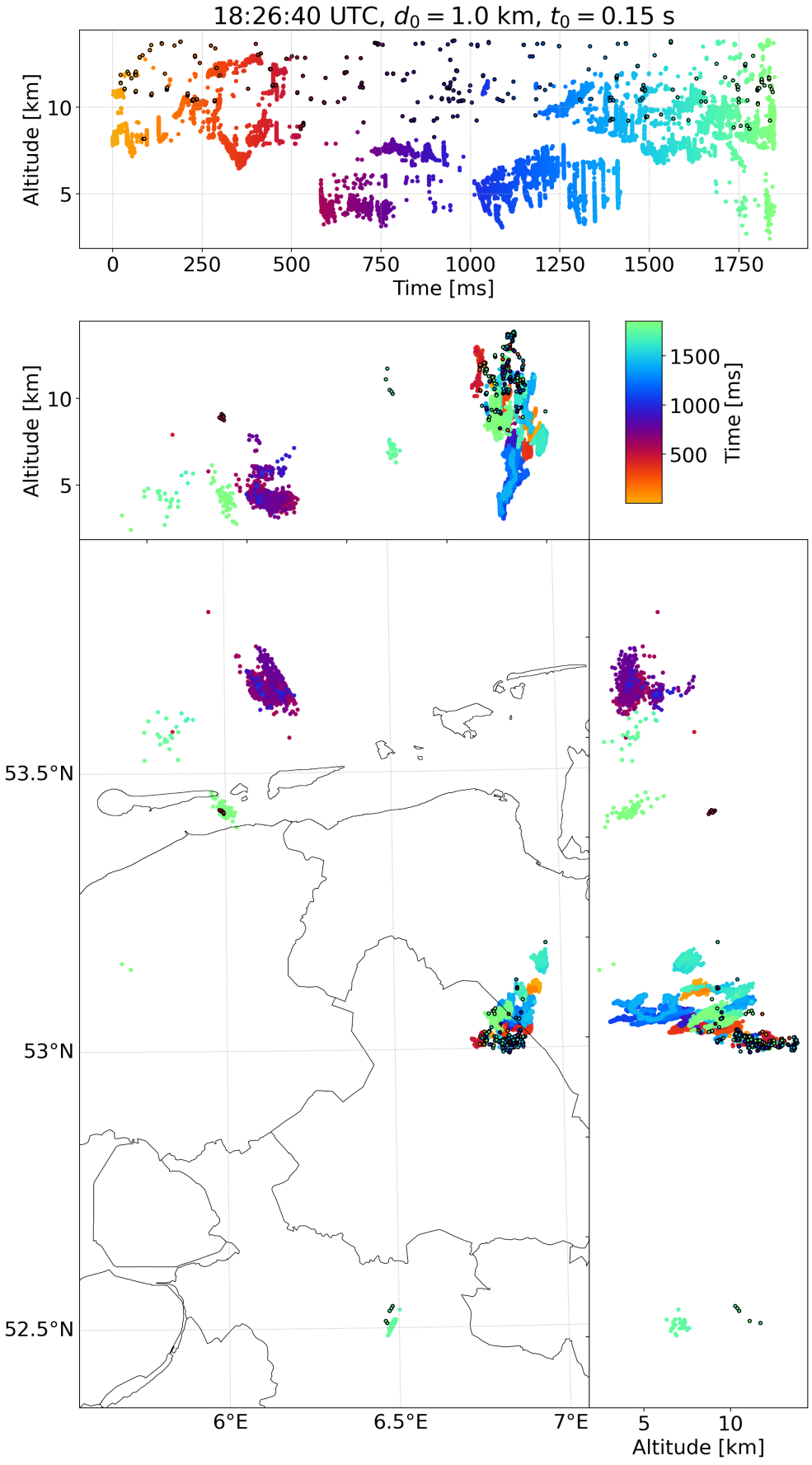

Figure A.3: Similar to figure A.1, but for the image at 18:26:40 UTC.

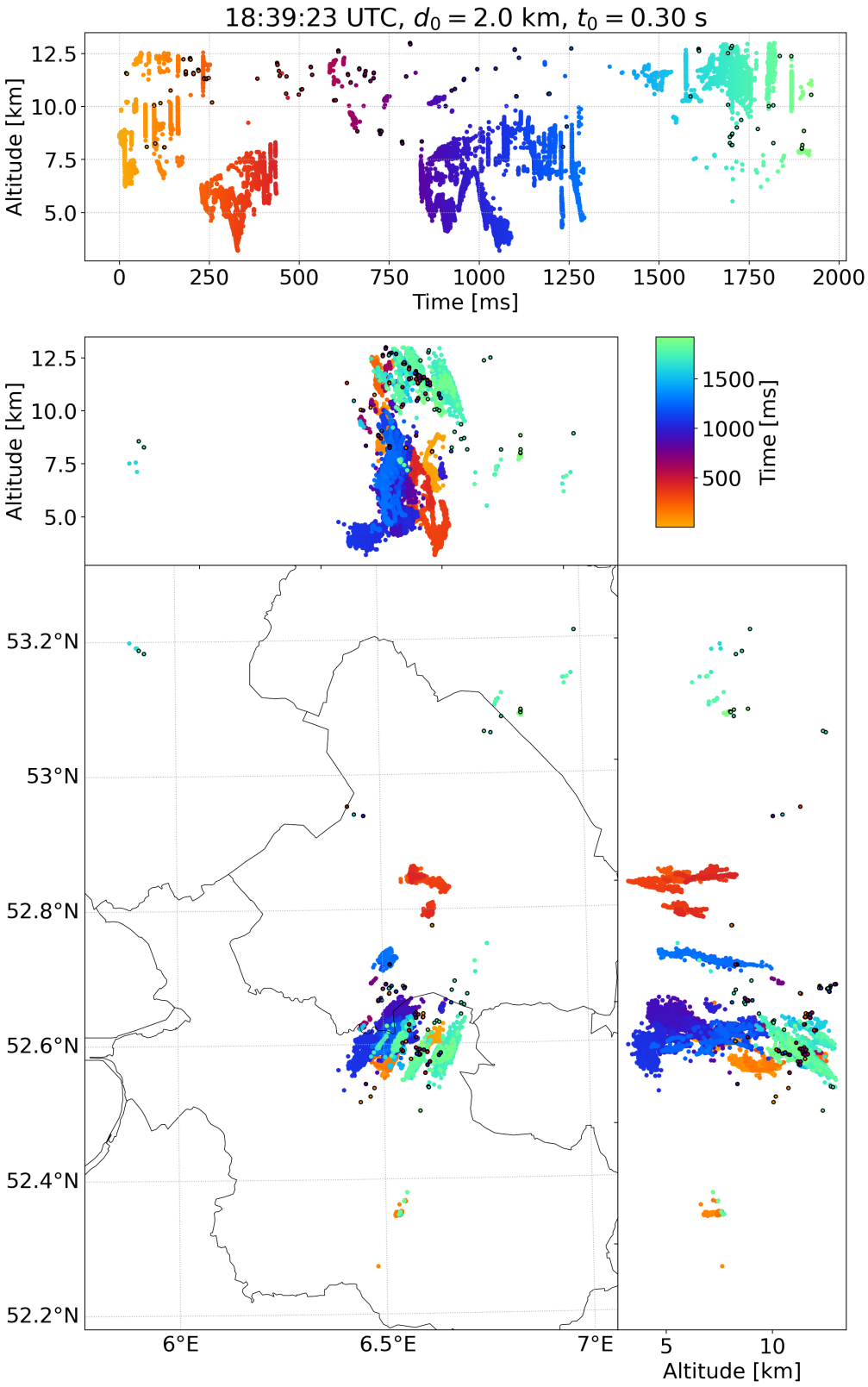

Figure A.4: Similar to figure A.1, but for the image at 18:39:23 UTC.

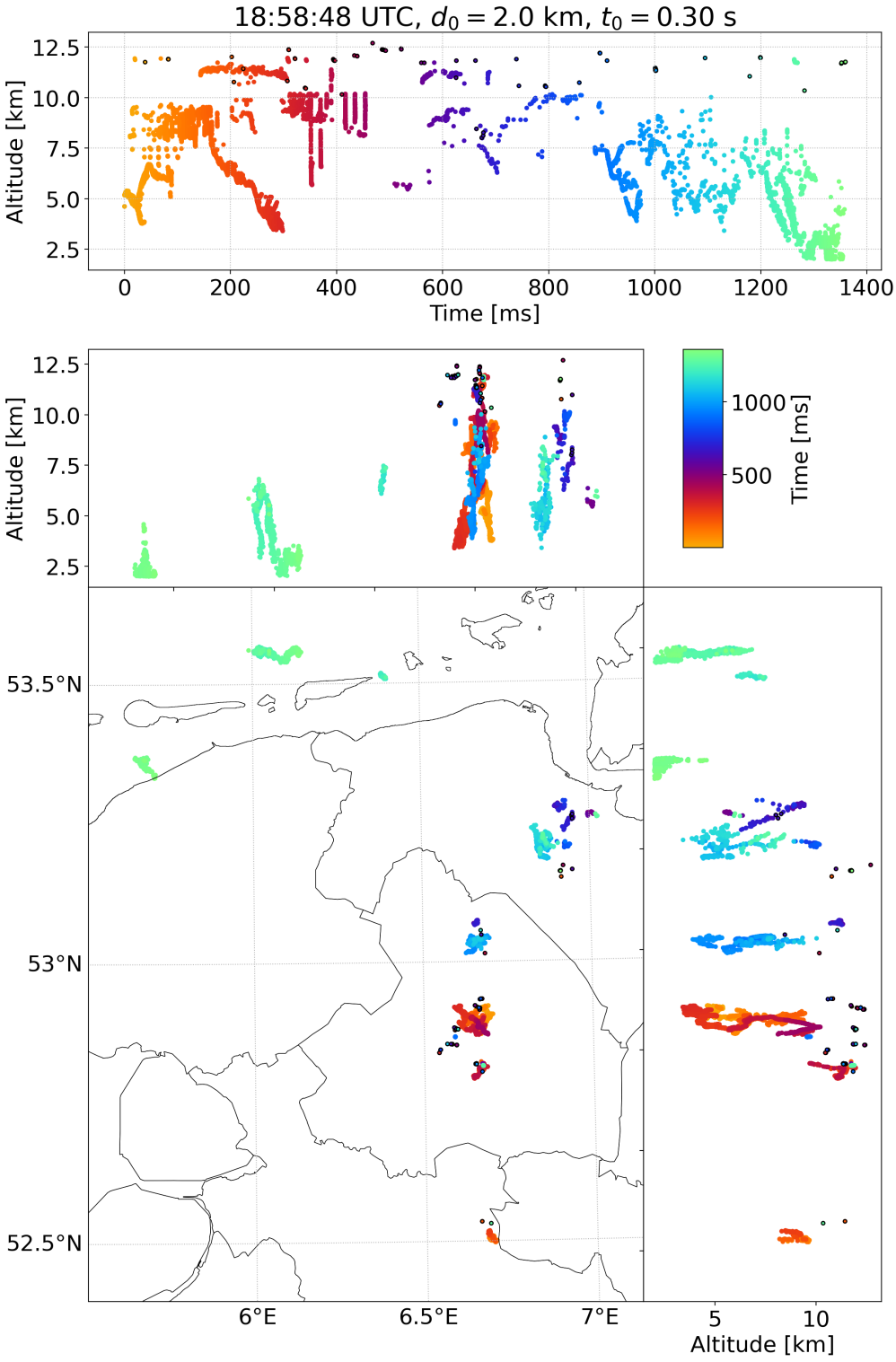

Figure A.5: Similar to figure A.1, but for the image at 18:58:48 UTC.

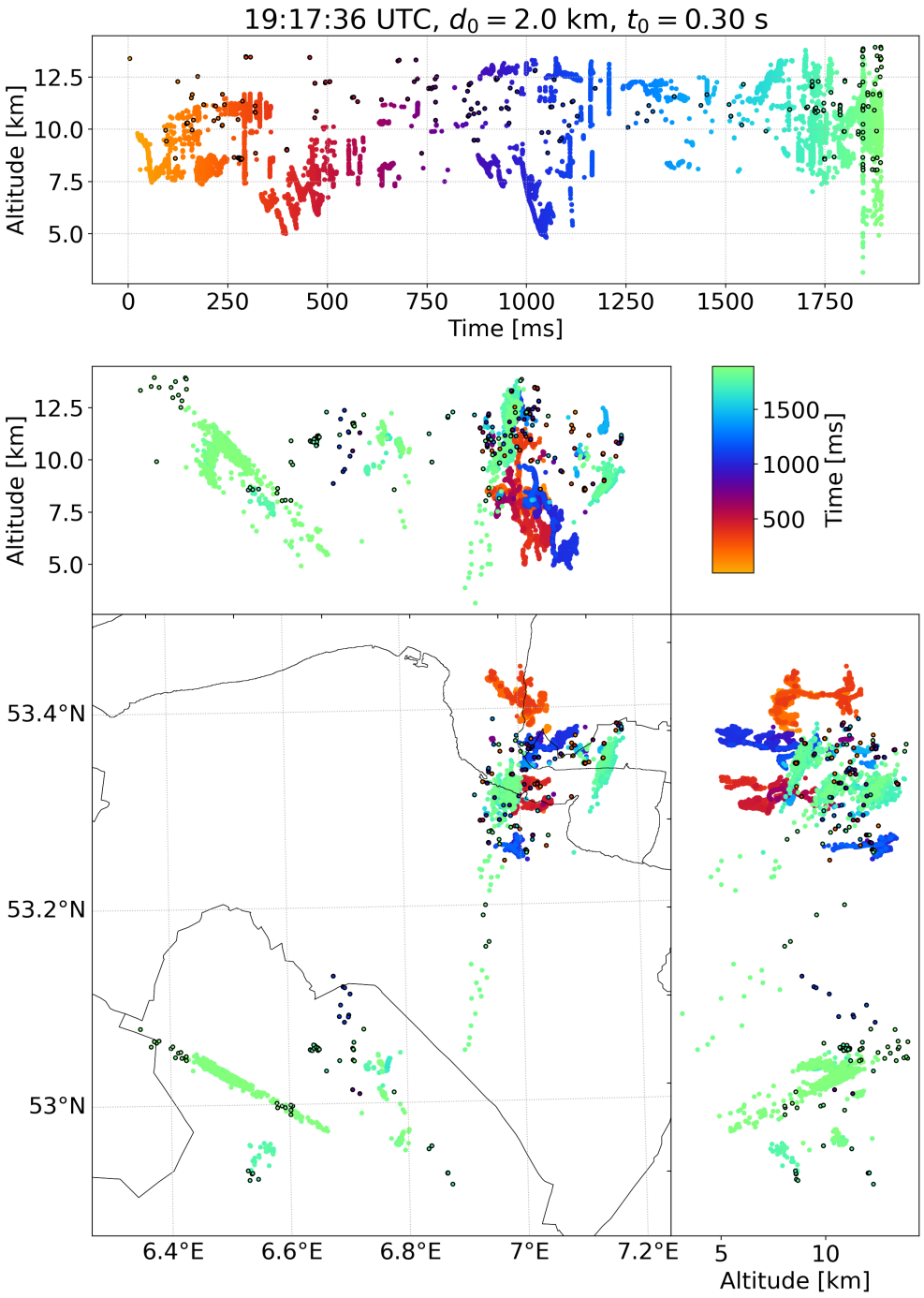

Figure A.6: Similar to figure A.1, but for the image at 19:17:36 UTC.

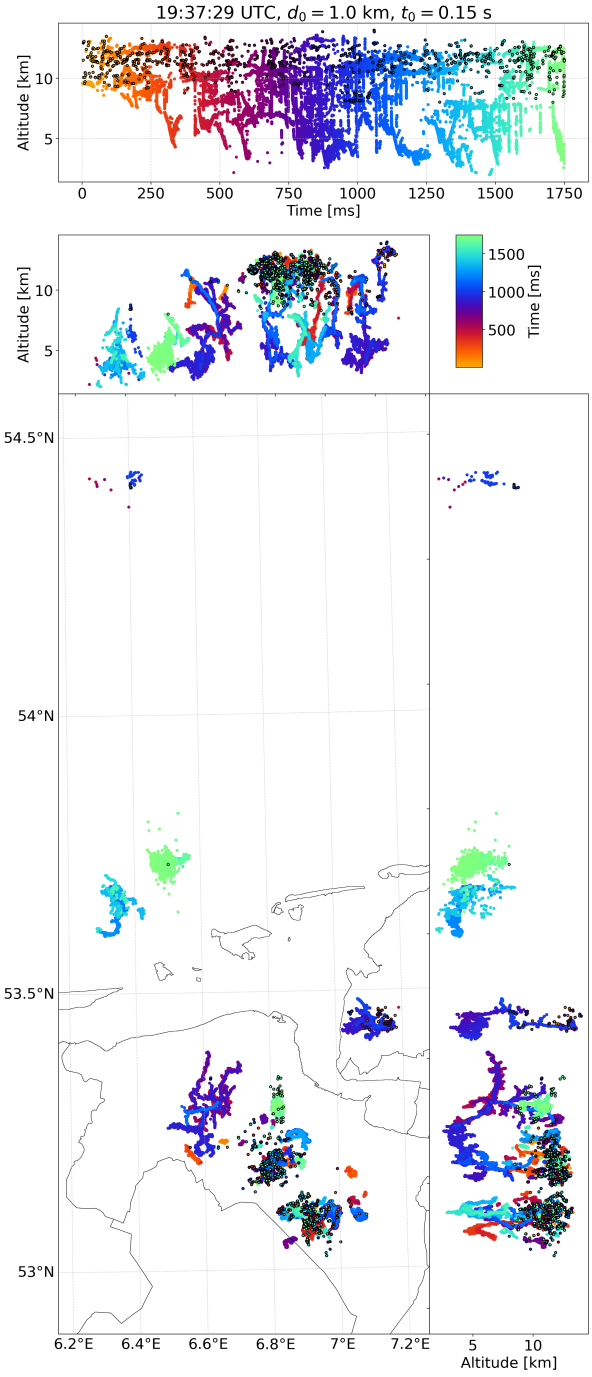

Figure A.7: Similar to figure A.1, but for the image at 19:37:29 UTC.

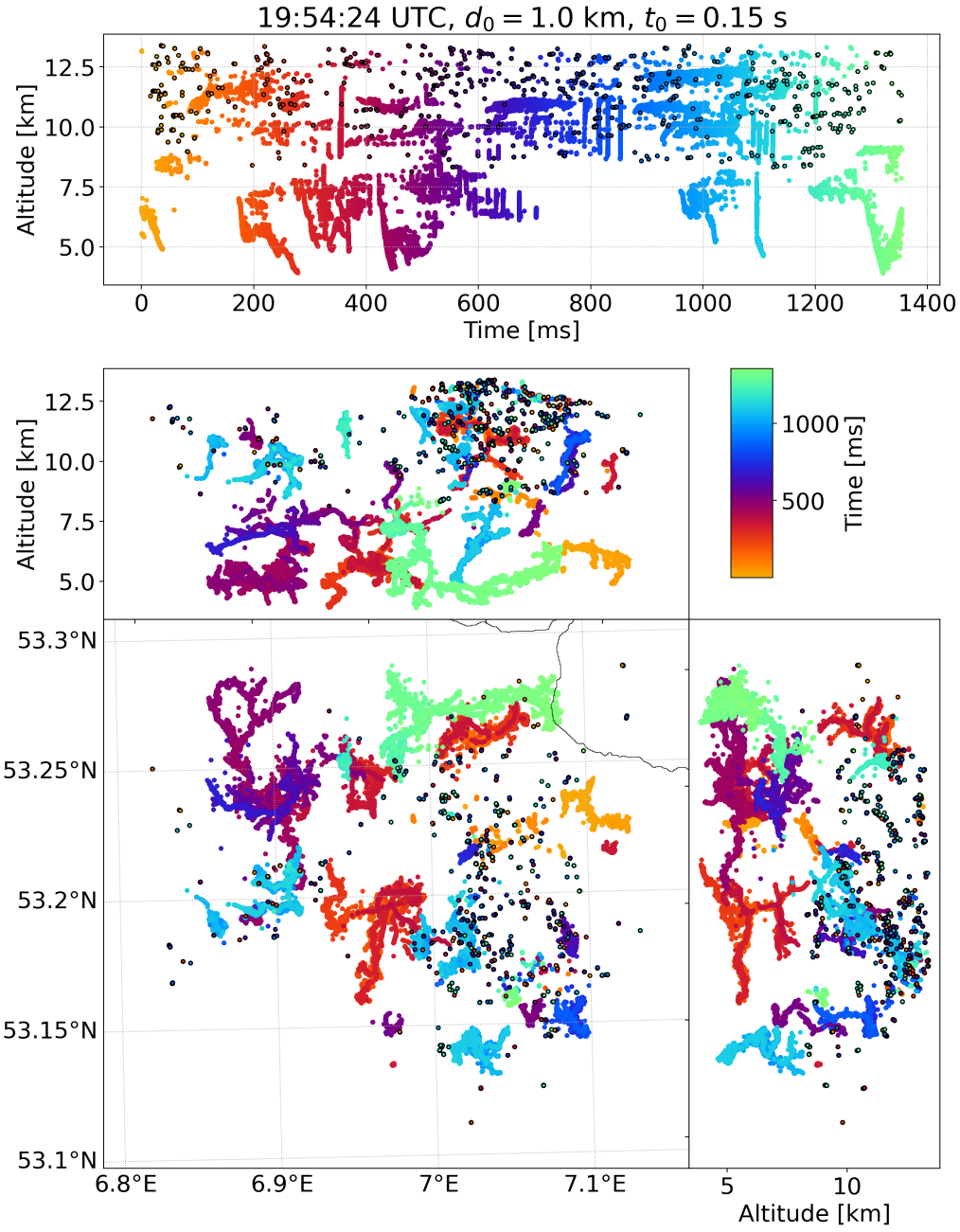

Figure A.8: Similar to figure A.1, but for the image at 19:54:24 UTC.

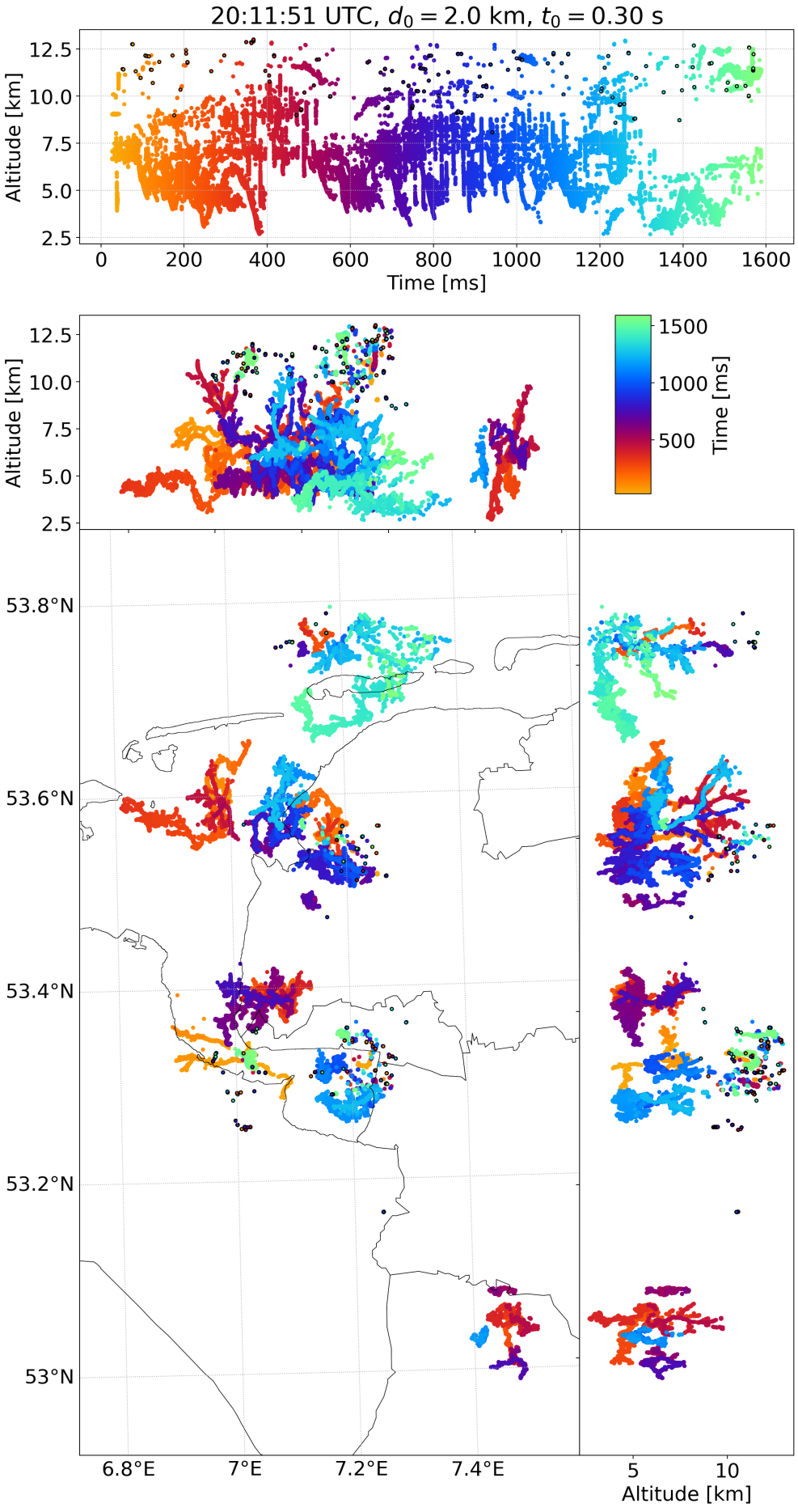

Figure A.9: Similar to figure A.1, but for the image at 20:11:51 UTC.

B. Radar variables

For this thesis, data is used from the German dual-polarimetric C-band frequency radar located on the island of Borkum. This appendix describes the polarimetric radar principles, and elaborates on the polarimetric variables used in this thesis and the interpretation thereof. The primary sources of this section are Doviak *et al.* [42], Straka *et al.* [48], and Kumjian [45] and KUMJIAN [54]. The former for general aspects of the Doppler radars, and the others for the characterisation and interpretation of polarimetric variables. The reader is advised to study these sources for a more comprehensive overview. Complementary to this part of the appendix, comprehensive figures of the radar variables can be found in appendix C.

Scanning procedure

Dual-polarimetric radars emit electromagnetic pulses with two orthogonal polarizations and measure the backscattered signal. A horizontally (H) polarized wave with an electric field component in the horizontal direction, and a vertically (V) polarized wave with with an electric field component orthogonal to both the propagation and horizontal direction. It is important to note that the V-polarization is not necessarily vertical with respect to the horizontal. When the radar beam is directed at an angle with respect to the horizontal, the V-polarization has an equal angle with the geometrical vertical direction.

Processing of the signals along one beam allows to differentiate volumes along the radial direction. By changing the beam angle, it is possible to get a more comprehensive volume scan of the three dimensional atmosphere. For the remainder of this report, the range refers to the distance along the radial direction, the resolution volume refers to the volume that is probed by a single beam within a range bin, the elevation angle refers to the angle with respect to the horizontal, and the azimuth refers to an angle with respect to the geographic North. While some radars operate in Range Height Indicator (RHI) mode, in which case the elevation angles are scanned before moving to another azimuth, the radar data in this thesis operates in Plan Position Indicator (PPI) mode. In this case, all azimuths are scanned before the radar moves to another elevation angle. For the remainder of this report, a sweep

refers to a scan of different azimuths for a fixed elevation angle.

Inherent to remote sensing, the signal may be altered along the way to and back from the resolution volume. The process called attenuation refers specifically to the power loss due to propagation and scattering along the beam path. Although partially compensated during computation, attenuation should be accounted for in the interpretation of certain variables.

Scattering properties

Backscatter upon particles is fundamental to radar observation. The properties of backscatter from a single particle is ultimately dependent on the size, geometry, orientation, and dielectric properties of particles.

First it is important to recognize that there are two major scattering regimes. The Rayleigh regime refers to particles that are much smaller than the radar wavelength (diameter $D \lesssim 1/10$). Such small particles act as a dipole antenna. The amount of backscatter is determined by the backscatter cross section. This cross section is the apparent area with respect to the wave polarization. For particles that are oblate, elongated along the H-dimension, the backscatter of a H-polarized wave is dominant over the backscatter of a V-polarized wave. The opposite holds for a particle that is elongated along the vertical dimension. A cross-polar signal, perpendicular to the incident polarization, may even be scattered by a particle when it is not symmetrical in the co-polar or cross-polar plane.

For particles that are not small with respect to the wavelength, "Mie scattering" applies. For Mie scattering, resonance effects and internal reflections are highly dependent on geometry. In the so called "Mie regime", the Mie scattering of the large particles may strongly enhance, and therefore complicate the scattering properties.

The dielectric properties of particles influence the scattering properties in multiple ways. The higher refractive index of (lower density) ice with respect to water, gives a smaller backscatter cross section. For Rayleigh scatterers, this results in lower backscatter intensity and accompanying polarization effects for (lower density) ice particles. The (complex) refractive index also affects Mie scattering, but with more complexity. For the refractive index determines the effective wavelength inside the hydrometeor and therefore influences the hydrometeors apparent geometry and thus the resonance scattering. In some situations ice particles can be coated with liquid water. Generally, this enhances the backscatter intensity and, especially

when the liquid water coating is not distributed evenly, the polarization effects.

Doppler velocity

The Doppler velocity is calculated from the phase shift that is acquired by the incident wave in a resolution volume upon scattering on hydrometeors. Such a phase shift is a direct consequence of the hydrometeors velocity along the propagation direction of the incident waves. For this reason, the a radar Doppler velocity is also referred to as the radial velocity V_{rad}. Besides V_{rad}, an average over the resolution volume, the variability of V_{rad} may be considered. The velocity spectrum width W_{rad} is defined as the square root of the second central moment, thus the variance, of the the radial Doppler velocity within a resolution volume. Enhanced W_{rad} values can be the result of turbulence, wind shear, and falling hydrometeors.[55]

Reflectivity

Generally, the backscattered power is proportional to the volume integral of the backscatter cross section perpendicular to the pulse propagation direction. Consequently a volume with more and larger particles will give a higher reflectivity signal. Because of dielectric effects, (lower density) ice particles yield a lower backscatter signal than liquid water particles or particles with liquid coating. For a dual-polarimetric radar, the shape or orientation of particles may determine the ratio of the H- and V-polarized reflectivity , Z_h and Z_v respectively. In the Rayleigh scattering regime, oblate particles give a relatively high Z_h value whereas particles that are elongated with respect to the vertical, would give a relatively high Z_v signal. For this reason, the differential reflectivity Z_{dr}, i.e. the logarithmic ratio of Z_h and Z_v, may be indicative of the statistical shape or orientation of hydrometeors. Falling droplets, for example, will be oblate, yielding positive Zdr values. On the contrary, V-aligned ice crystals would yield negative Z_{dr} values. Z_{dr} values suffer both from the calibration errors of the Z_h and Z_v values and attenuation, especially when Z_h or Z_h are relatively low.

Differential phase

Besides attenuation reducing the signal power, the propagation through particles may also change the phase of the radar signals. The differential phase ϕ_{dp} is therefore considered a propagation variables and is defined as the phase difference between the H- and V-polarized signals. During propagation through a volume of

statistically oblate particles in the Rayleigh regime, the H-polarized wave will propagate slower and subsequently acquire a larger phase shift yielding positive ϕ_{dp} values. Vice versa for a volume with statistically V elongated particles, ϕ_{dp} will be negative. Since ϕ_{dp} works cumulative along the radar beam, a specific differential phase K_{dp} can be constructed as half of the (discrete) radial derivative of ϕ_{dp}. K_{dp} may therefore be interpreted as a variable that is intrinsic to the radial space between the resolution grid points representing the statistical anisotropy. Since K_{dp} depends on the phase and not the reflectivity, it is not sensitive for attenuation, radar calibration, particles that are statistically isotropic, and size distribution variations.

Besides during propagation, a phase shift may also be acquired upon resonance backscattering by hydrometeors in the Mie regime. This phenomenon, also known as the backscatter differential phase δ, can be observed as significantly enhanced or ϕ_{dp} values. Therefore, for large hydrometeors, ϕ_{dp} also contain intrinsic information about the resolution volume. For K_{dp}, δ effects might add a lining around the volumes of large hydrometeors. Since the sign of δ can potentially change over a narrow band of hydrometeor sizes and shapes, it can potentially be used to for hydrometeor characterisation. For the same reason, δ can obscure the interpretation of ϕ_{dp} and K_{dp}.

B.0.0.1 K_{dp} processing

Since DWD provided only the total differential phase ψ_{dp}, the Wradlib python package[41] has been used for filtering and computing ϕ_{dp} and K_{dp} values. The result is highly depend on filtering and correction choices, this subsection describes the specifics on the ψ_{dp} processing details.

In general ψ_{dp} is noisy. The general approach used to retrieve a more meaningful ϕ_{dp} and computing the K_{dp} signal follows the method as described by Vulpiani *et al.* [56]. The described method is an iterative process based upon 4 four steps:

1. K_{dp} is calculated using a finite difference method on the ψ_{dp} signal: K'_{dp}
2. K_{dp} values are filtered for realistic values such that $\kappa_1 < K'_{dp} < \kappa_2$. When this condition is not met, K'_{dp} is replaced with a zero value. This method attempts to remove noise and backscatter phase δ_{hv}. Potential phase folding should be taken into account to prevent unnecessary data loss.
3. $\hat{\phi}_{dp}$ is reconstructed by integration of the filtered K'_{dp}.

4. K_{dp} is computed from $\hat{\phi}_{dp}$: $\hat{K}_{dp}$

It is important to realize that the computation of K_{dp} makes use of a smoothing window. This has a twofold purpose. Not only does it smooth the noisy ψ_{dp} signal, it also deals with values outside the $\kappa_1 - \kappa_2$ range. Once the $\kappa_1 < K'_{dp} < \kappa_2$ is not met, K'_{dp} values get replaces by zero's. These nonphysical zero values are compensated by smoothing of the neighbouring values. For this reason, Vulpiani *et al.* [56] suggests to iterate over steps 3 and 4.

Another important artefact to consider is that phase measurements are prone to phase folding. The phase measurements are limited to 180°, such that when a phase shift is increased beyond 180°, it jumps to −180°. The opposite holds for the phase decreasing from −180°, that results in a jump to 180°. Phase jumps always lead to K'_{dp} values outside the $\kappa_1 - \kappa_2$ range and subsequently data loss. Therefore, it is useful to try to unfold the phase folds into a useful signal.

Although powerful because of the smoothing, and because of the independence of ψ_{dp} calibration offsets, the method by Vulpiani *et al.* [56] as manifested in the *wradlib.dp.phidp_kdp_vulpiani* has a major shortcoming during phase unfolding. The unfolding is based on K'_{dp} computed with a central finite difference method. This method duplicates unwanted artefacts around the artefacts itself. This may lead to unforeseen phase unfolding. Moreover, the suggested unfolding algorithm only permits a one time phase addition of 360°. Therefore, it cannot cope with multiple folds or a fold in the decreasing −180° direction. In such cases, data is lost because of the $\kappa_1 < K'_{dp} < \kappa_2$ restriction on K_{dp} by step 2. Finally, the Vulpiani *et al.* [56] method does not include possibilities for noise masks. The iteration of smoothed k_{dp} does compensate partially for random noise. The phase unfolding is, however, before appropriate smoothing.

To increase the performance on phase folds, the following algorithm was applied for ψ_{dp} processing:

1. Changes in ψ_{dp} are neglected for low reflectivity values. E.g. where $Z_h < z'$ or $Z_v < z'$.

2. Phase unfolding by a modified *wradlib.dp._unfold_phidp_naive* python function based on Wang *et al.* [57]. This algorithm takes ρ_{hv} and the standard preceding standard deviation as arguments, to differentiate between noise and actual phase folds. The algorithm has bee modified in order to tune the thresholds on the realistic propagation specific phase κ_1 and κ_2, the threshold

for fold recognition K_{fold}, the threshold for the window standard deviation σ_w.

3. K_{dp} values are filtered for realistic values such that $\kappa_1 < K'_{dp} < \kappa_2$. When this condition is not met, K'_{dp} is replaced with a zero value.
4. $\hat{\phi}_{dp}$ is reconstructed by integration of the filtered K'_{dp}.
5. K_{dp} is computed from $\hat{\phi}_{dp}$: $\hat{K}_{dp}$

For this study, the values for the filtering parameters in the above method are: $\kappa_1 = -5^\circ$, $\kappa_2 = 20^\circ$, $z' = 0$ dB, $\sigma_w = 0.8$. The size of the rolling window computation along the radar beam is 11 bins.

Correlation coefficient

Once there is variability in hydrometeor shape and orientation within a resolution volume, this may be captured by the correlation coefficient ρ_{hv} at zero time lag. It is defined as the normalized covariance between the the H- and V-polarized signals and equals unity for perfectly correlated signals. Once the H and V signals do not vary simultaneously within a resolution volume, ρ_{hv} is reduced towards zero. Although independent of the radar sensitivity calibration, particle concentration and propagation effects, the interpretation of ρ_{hv} is not straight forward for Mie scatterers. Backscatter differential phase δ may significantly decorrelate the H- and V-polarized signals, leading to low ρ_{hv} values.

Hydrometeor classification

Using the variables described above, it may be possible to characterise the hydrometeors inside a resolution volume. Not only in terms of type, but sometimes also in terms of quantity, size, geometry or orientation. Since in situ measurements are sparse, most of the logic behind hydrometeor classification is based on physics, modelling studies, on-ground disdrometers or upward pointing radars. Besides the radar variables, the temperature is also employed as a predictor for the hydrometeor type, for it is a key factor determining the thermodynamic phase of water. Although not explicitly described in this section, the theory behind qualitative hydrometeor characterisation described by Straka *et al.* [48] and Kumjian [45] are elaborated on for the interpretation of the results in section 4.

Besides the manual interpretation of radar images, algorithms may be employed to compute the presence of different hydrometeor types. One method for such a hy-

drometeor algorithm would be to separate the multi-dimensional polarimetric variable space into distinct regions for different bulk hydrometeors classes k, e.g. hail, rain, or snow. Ideally, these regions would occupy a volume in an M-dimensional space of polarimetric variables Y. Such thresholds might be constructed by (subjective) optimization of a probability density function mapping the probability of correct classification in the M-dimensional Y-space. As Zrnić *et al.* [58] notes, such a probability function is not available. For this reason, Zrnić *et al.* [58] describes a realization of a workable alternative algorithm, namely the fuzzy logic hydrometeor classification.

A simple approach is to construct probability mappings $f_k(Y_i)$, for class k given a value Y_i. A weighted sum S_k of the separate probabilities $f_k(Y_i)$ would subsequently provide information for the classification:

$$S_k = \frac{\sum_{i=1}^{M} A_i f_k(Y_i)}{\sum_{i=1}^{M} A_i} \tag{B.1}$$

Here A_i is the weight for the different polarimetric radar variables Y_i. The hydrometeor class j is decided as j with the maximum value of S_k:

$$j = \arg\max_k S_k \tag{B.2}$$

The confidence of the classification can be estimated from both the value of S_k, representing a probability, and the difference between the maximum S_k and the second highest S_j. Note that f_i is not a probability density function, for the probability is not given by the integral of f_i over Y_i, but by the value of f_i. In theory, f_i could be any function, but it is a natural choice for f_i to be a trapezoid with values between 0 and 1. Such a trapezoid represents a "likely" region for j where f_i equals unity, and linear transition areas separating "impossible" from the "likely" Y_i values. Besides f_i, the choice of the weights A_i is of importance for the outcome. One might for example decide on a lower weight for a variable with lower confidence or less predictability.

A key difference between the M-dimensional threshold volume and the weighted sum described above, is that the latter does not account for any correlation between Y_i. The former would allow for much more tight differentiation between classes k. One way to incorporate some correlation between Y_i is two construct 2-dimensional probability functions $f_k(Y_i, Y_j)$ as the product of a one dimensional probability function and a constrained 2-dimensional probability function (equation B.3).

$$f_k(Y_i, Y_j) = f_k(Y_i) \cdot f_k(Y_j|Y_i) \tag{B.3}$$

Here, both $f_k(Y_i)$ and $f_k(Y_j|Y_i)$ are both trapezoidal shaped and may be determined by comparing scatter diagrams in a $Y_i - Y_j$-plane, (manual) skill optimization, and expert opinion.

C. Radar images

Figure C.1 shows a comprehensive set of radar variables. In order to determine the quality of the HMC-algorithm, the variables Z_{dr}, ρ_{hv}, K_{dp} and the HMC maximum probability HMC$_probability$ are provided. The purpose of $\phi^*_{dp,unfolded}$ and δ_{hv} is to provide insights in the quality of the K_{dp} algorithm.

Figure C.2 shows a comprehensive set of radar variables with the same reasoning as for figure C.1. However, the higher elevation angle of 25° is relevant for sparkles at higher altitudes as depicted in fig 4.3.

Similar to figure C.1 and figure C.2, the polarimetric variables are shown in figure C.3 and C.4.

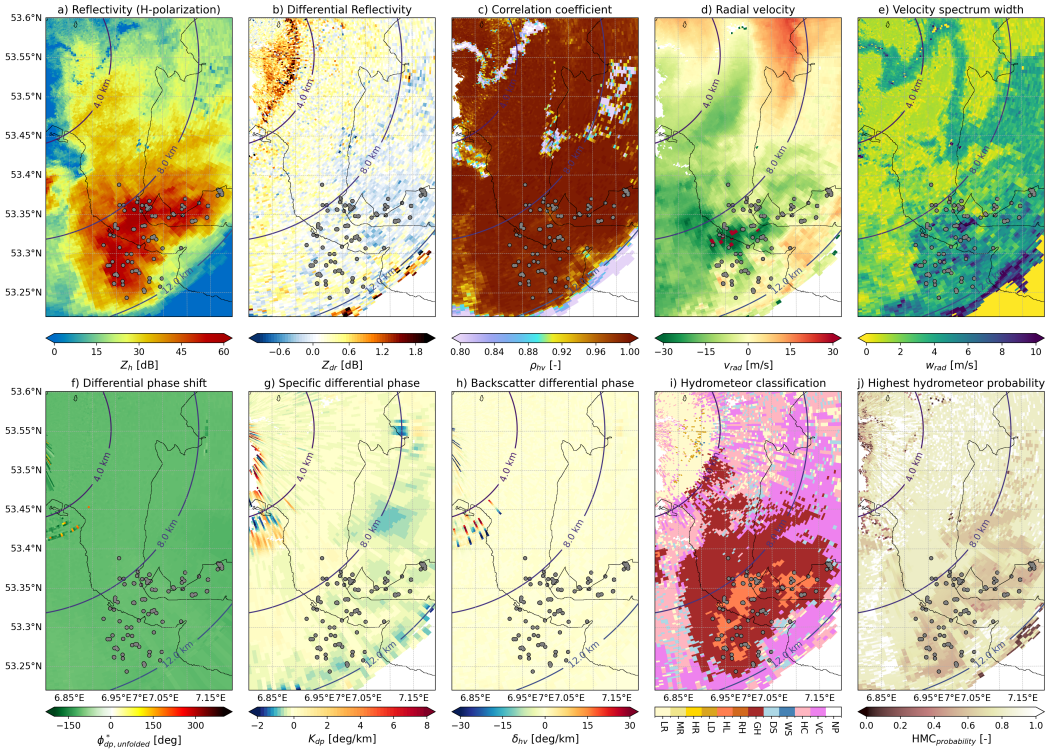

Figure C.1: Radar Z_h, Z_{dr}, ρ_{hv}, v_{rad}, w_{rad}, $\phi^*_{dp,unfolded}$, K_{dp}, δ_{hv}, HMC and HMC$_{probability}$ projected upon the horizontal plane for an elevation angle of 17°. The horizontal extent is identical to figure 4.4.

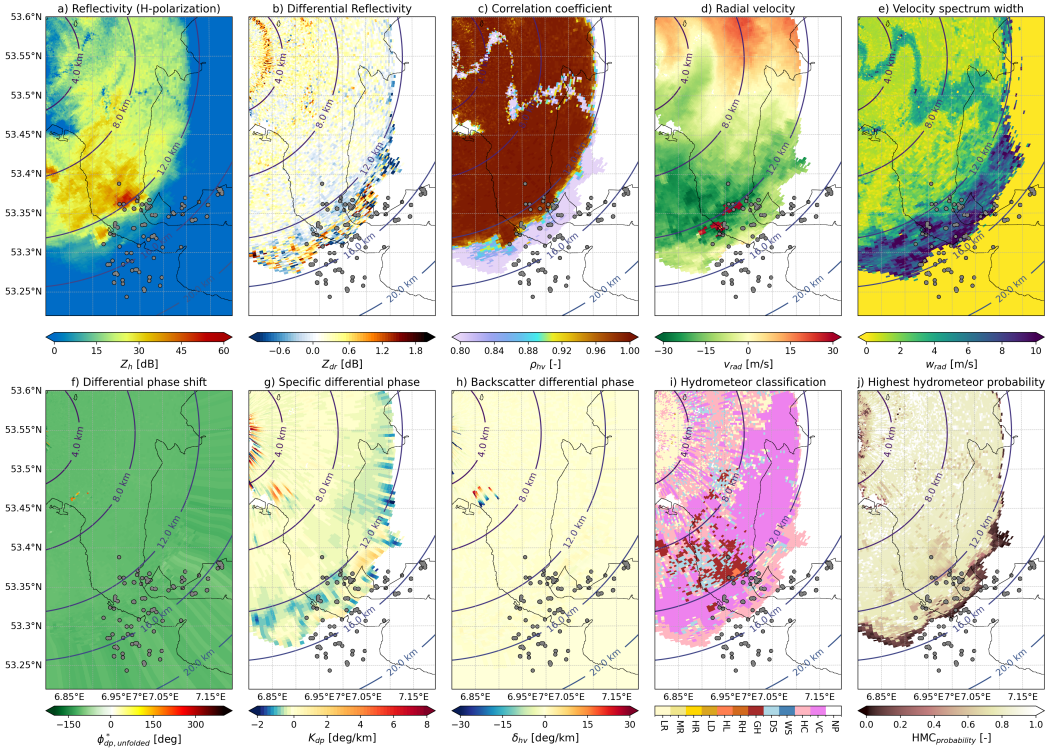

Figure C.2: Similar to figure C.1, but with an elevation angle of 25° and sparkles projected upon the horizontal plane of panel a.

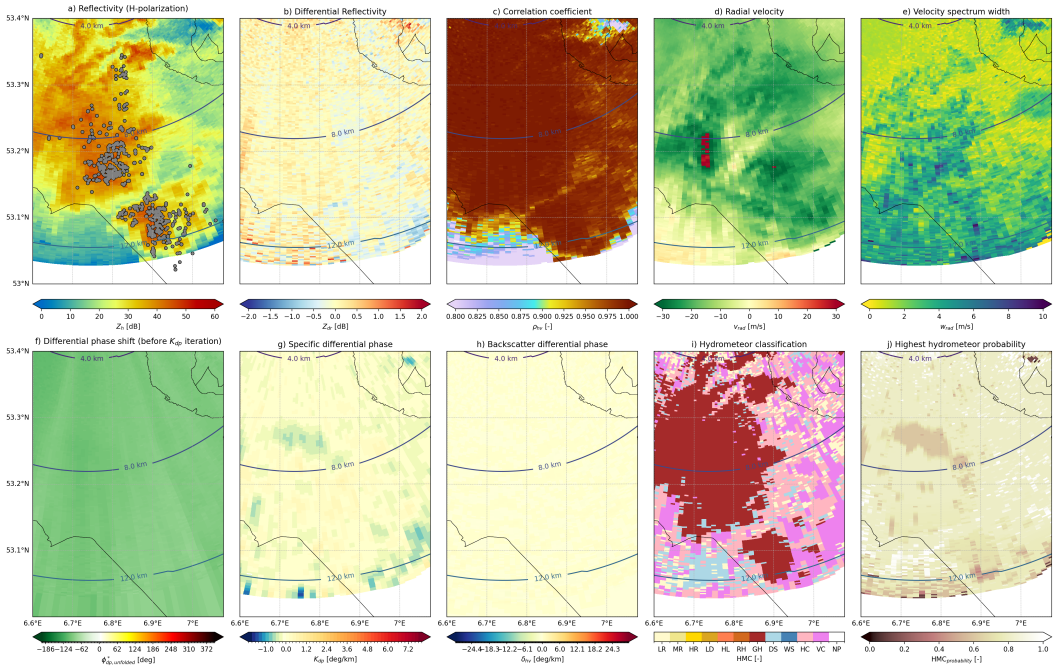

Figure C.3: Radar Z_h, Z_{dr}, ρ_{hv}, v_{rad}, w_{rad}, $\phi^*_{dp,unfolded}$, K_{dp}, δ_{hv}, HMC and $\text{HMC}_{probability}$ projected upon the horizontal plane for an elevation angle of 17°. The horizontal extent is identical to figure 4.7.

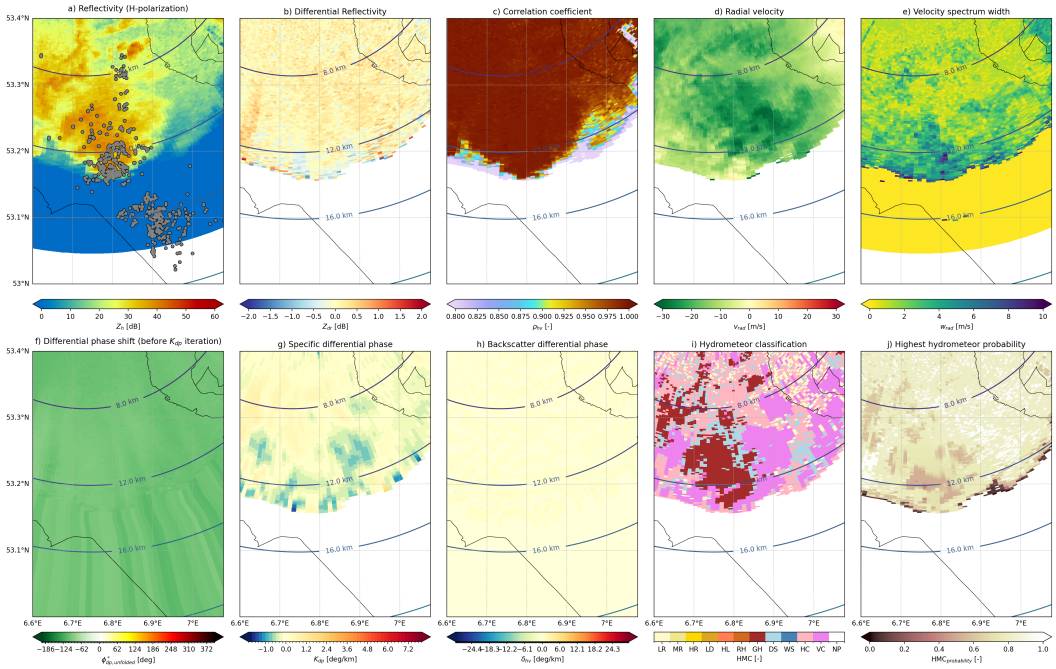

Figure C.4: Similar to figure C.3, but with an elevation angle of 25° and sparkles projected upon the horizontal plane of panel a.

Bibliography

[1] W. Deierling, W. A. Petersen, J. Latham, S. Ellis, and H. J. Christian, "The relationship between lightning activity and ice fluxes in thunderstorms," *Journal of Geophysical Research: Atmospheres*, vol. 113, no. D15, 2008.

[2] D. R. MacGorman, M. S. Elliott, and E. DiGangi, "Electrical discharges in the overshooting tops of thunderstorms," *Journal of Geophysical Research: Atmospheres*, vol. 122, no. 5, pp. 2929–2957, 2017.

[3] M. Bonnet, "Status of lightning detection: Performances and limitations of existing systems," *International Colloquium of Ligtning and POwer Systems*, 2014.

[4] R. J. Thomas, P. R. Krehbiel, W. Rison, S. J. Hunyady, W. P. Winn, T. Hamlin, and J. Harlin, "Accuracy of the lightning mapping array," *Journal of Geophysical Research: Atmospheres*, vol. 109, no. D14, 2004.

[5] M. P. van Haarlem, M. W. Wise, A. Gunst, G. Heald, J. P. McKean, J. W. Hessels, A. G. de Bruyn, R. Nijboer, J. Swinbank, R. Fallows, *et al.*, "Lofar: The low-frequency array," *Astronomy & astrophysics*, vol. 556, A2, 2013.

[6] B. Hare, O. Scholten, J. Dwyer, T. Trinh, S. Buitink, S. Ter Veen, A. Bonardi, A. Corstanje, H. Falcke, J. Hörandel, *et al.*, "Needle-like structures discovered on positively charged lightning branches," *Nature*, vol. 568, no. 7752, pp. 360–363, 2019.

[7] N. Y. Liu, O. Scholten, B. M. Hare, J. R. Dwyer, C. F. Sterpka, I. Kolmašová, and O. Santolík, "Lofar observations of lightning initial breakdown pulses," *Geophysical Research Letters*, vol. 49, no. 6, e2022GL098073, 2022.

[8] O. Scholten, B. M. Hare, J. Dwyer, N. Liu, C. Sterpka, J. Assink, H. Leijnse, and S. t. Veen, "Small-scale discharges observed near the top of a thunderstorm," *Geophysical Research Letters*, vol. 50, no. 8, e2022GL101304, 2023, e2022GL101304 2022GL101304. DOI: https://doi.org/10.1029/2022GL101304. eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2022GL101304. [Online]. Available: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2022GL101304.

[9] C. Emersic, P. Heinselman, D. R. MacGorman, and E. Bruning, "Lightning activity in a hail-producing storm observed with phased-array radar," *Monthly Weather Review*, vol. 139, no. 6, pp. 1809–1825, 2011.

[10] K. M. Calhoun, D. R. MacGorman, C. L. Ziegler, and M. I. Biggerstaff, "Evolution of lightning activity and storm charge relative to dual-doppler analysis of a high-precipitation supercell storm," *Monthly Weather Review*, vol. 141, no. 7, pp. 2199–2223, 2013.

[11] American Meteorological Society. "Glossary of meteorology." Accessed: 2024-07-05. (2024), [Online]. Available: https://glossary.ametsoc.org.

[12] C. Saunders and S. Peck, "Laboratory studies of the influence of the rime accretion rate on charge transfer during crystal/graupel collisions," *Journal of Geophysical Research: Atmospheres*, vol. 103, no. D12, pp. 13 949–13 956, 1998.

[13] J. C. Souza and E. C. Bruning, "Assessment of turbulence intensity in different spots of lightning flash propagation," *Geophysical Research Letters*, vol. 48, no. 21, e2021GL095923, 2021.
[14] J. R. Dwyer and M. A. Uman, "The physics of lightning," *Physics Reports*, vol. 534, no. 4, pp. 147–241, 2014.
[15] B. Mason, "A critical examination of theories of charge generation in thunderstorms," *Tellus*, vol. 5, no. 4, pp. 446–460, 1953.
[16] C. Saunders, "Charge separation mechanisms in clouds," *Planetary Atmospheric Electricity*, pp. 335–353, 2008.
[17] J. Dash, B. Mason, and J. Wettlaufer, "Theory of charge and mass transfer in ice-ice collisions," *Journal of Geophysical Research: Atmospheres*, vol. 106, no. D17, pp. 20 395–20 402, 2001.
[18] E. R. Mansell, D. R. MacGorman, C. L. Ziegler, and J. M. Straka, "Charge structure and lightning sensitivity in a simulated multicell thunderstorm," *Journal of Geophysical Research: Atmospheres*, vol. 110, no. D12, 2005.
[19] E. A. Mareev and S. O. Dementyeva, "The role of turbulence in thunderstorm, snowstorm, and dust storm electrification," *Journal of Geophysical Research: Atmospheres*, vol. 122, no. 13, pp. 6976–6988, 2017.
[20] E. R. Williams, "The tripole structure of thunderstorms," *Journal of Geophysical Research: Atmospheres*, vol. 94, no. D11, pp. 13 151–13 167, 1989.
[21] E. C. Bruning, W. D. Rust, D. R. MacGorman, M. I. Biggerstaff, and T. J. Schuur, "Formation of charge structures in a supercell," *Monthly Weather Review*, vol. 138, no. 10, pp. 3740–3761, 2010.
[22] E. C. Bruning, S. A. Weiss, and K. M. Calhoun, "Continuous variability in thunderstorm primary electrification and an evaluation of inverted-polarity terminology," *Atmospheric Research*, vol. 135, pp. 274–284, 2014.
[23] W. Hoppel and B. Phillips, "The electrical shielding layer around charged clouds and its role in thunderstorm electricity," *Journal of Atmospheric Sciences*, vol. 28, no. 7, pp. 1258–1271, 1971.
[24] V. Mazur, *Principles of lightning physics*. IoP Publishing, 2016.
[25] C. Sterpka, J. Dwyer, N. Liu, B. Hare, O. Scholten, S. Buitink, S. ter Veen, and A. Nelles, "The spontaneous nature of lightning initiation," *Geophysical Research Letters*, vol. 48, no. 23, 2021.
[26] D. Petersen, M. Bailey, W. H. Beasley, and J. Hallett, "A brief review of the problem of lightning initiation and a hypothesis of initial lightning leader formation," *Journal of Geophysical Research: Atmospheres*, vol. 113, no. D17, 2008.
[27] N. Y. Liu, O. Scholten, B. M. Hare, J. R. Dwyer, C. F. Sterpka, I. Kolmašová, and O. Santolík, "Lofar observations of lightning initial breakdown pulses," *Geophysical Research Letters*, vol. 49, no. 6, e2022GL098073, 2022.
[28] V. Rakov, "Electromagnetic methods of lightning detection," *Surveys in Geophysics*, vol. 34, pp. 731–753, 2013.
[29] S. Pédeboy, P. Barnéoud, E. Defer, and S. Coquillat, "Analysis of the intracloud lightning activity detected with low frequency lightning locating systems," in *25th International Lightning Detection Conference, 7th International Lightning Meteorology Conference*, 2018, pp. 12–15.
[30] B. Hare, O. Scholten, A. Bonardi, S. Buitink, A. Corstanje, U. Ebert, H. Falcke, J. Hörandel, H. Leijnse, P. Mitra, *et al.*, "Lofar lightning imaging: Map-

ping lightning with nanosecond precision," *Journal of Geophysical Research: Atmospheres*, vol. 123, no. 5, pp. 2861–2876, 2018.

[31] O. Scholten, B. Hare, J. Dwyer, C. Sterpka, I. Kolmašová, O. Santolík, R. Lan, L. Uhlíř, S. Buitink, A. Corstanje, *et al.*, "The initial stage of cloud lightning imaged in high-resolution," *Journal of Geophysical Research: Atmospheres*, vol. 126, no. 4, e2020JD033126, 2021.

[32] O. Scholten, B. Hare, J. Dwyer, N. Liu, C. Sterpka, S. Buitink, T. Huege, A. Nelles, and S. ter Veen, "Time resolved 3d interferometric imaging of a section of a negative leader with lofar," *Physical Review D*, vol. 104, no. 6, p. 063 022, 2021.

[33] O. Scholten, B. Hare, J. Dwyer, N. Liu, C. Sterpka, S. Buitink, A. Corstanje, H. Falcke, T. Huege, J. Hörandel, *et al.*, "Distinguishing features of high altitude negative leaders as observed with lofar," *Atmospheric Research*, vol. 260, p. 105 688, 2021.

[34] S. Soler, F. J. Pérez-Invernón, F. Gordillo-Vázquez, A. Luque, D. Li, A. Malagón-Romero, T. Neubert, O. Chanrion, V. Reglero, J. Navarro-Gonzalez, *et al.*, "Blue optical observations of narrow bipolar events by asim suggest corona streamer activity in thunderstorms," *Journal of Geophysical Research: Atmospheres*, vol. 125, no. 16, e2020JD032708, 2020.

[35] W. A. Lyons, T. E. Nelson, R. A. Armstrong, V. P. Pasko, and M. A. Stanley, "Upward electrical discharges from thunderstorm tops," *Bulletin of the American Meteorological Society*, vol. 84, no. 4, pp. 445–454, 2003.

[36] O. Chanrion, T. Neubert, A. Mogensen, Y. Yair, M. Stendel, R. Singh, and D. Siingh, "Profuse activity of blue electrical discharges at the tops of thunderstorms," *Geophysical Research Letters*, vol. 44, no. 1, pp. 496–503, 2017.

[37] K. M. Calhoun, E. R. Mansell, D. R. MacGorman, and D. C. Dowell, "Numerical simulations of lightning and storm charge of the 29–30 may 2004 geary, oklahoma, supercell thunderstorm using enkf mobile radar data assimilation," *Monthly Weather Review*, vol. 142, no. 11, pp. 3977–3997, 2014.

[38] B. R. Fuchs, E. C. Bruning, S. A. Rutledge, L. D. Carey, P. R. Krehbiel, and W. Rison, "Climatological analyses of lma data with an open-source lightning flash-clustering algorithm," *Journal of Geophysical Research: Atmospheres*, vol. 121, no. 14, pp. 8625–8648, 2016.

[39] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, *et al.*, "Scikit-learn: Machine learning in python," *the Journal of machine Learning research*, vol. 12, pp. 2825–2830, 2011.

[40] M. Frech, M. Hagen, and T. Mammen, "Monitoring the absolute calibration of a polarimetric weather radar," *Journal of Atmospheric and Oceanic Technology*, vol. 34, no. 3, pp. 599–615, 2017.

[41] M. Heistermann, S. Jacobi, and T. Pfaff, "Technical note: An open source library for processing weather radar data (wradlib)," *Hydrology and Earth System Sciences*, vol. 17, pp. 863–871, 2013. DOI: 10.5194/hess-17-863-2013.

[42] R. J. Doviak *et al.*, *Doppler radar and weather observations*. Elsevier Inc, Academic Press, 1993.

[43] Wradlib, *2-dimensional membership functions for c-band hydrometeor classification*, https://github.com/wradlib/wradlib-data/blob/main/misc/msf_cband_v2.nc, NetCDF data file, 2024. [Online]. Available: https:

//github.com/wradlib/wradlib-data/blob/main/misc/msf%5C_cband%5C_v2.nc.

[44] F. S. Marzano, D. Scaranari, M. Celano, P. Alberoni, G. Vulpiani, and M. Montopoli, "Hydrometeor classification from dual-polarized weather radar: Extending fuzzy logic from s-band to c-band data," *Advances in Geosciences*, vol. 7, pp. 109–114, 2006.

[45] M. R. Kumjian, "Principles and applications of dual-polarization weather radar. part i: Description of the polarimetric radar variables.," *Journal of Operational Meteorology*, vol. 1, 2013.

[46] P. Markowski and Y. Richardson, *Mesoscale meteorology in midlatitudes*. John Wiley & Sons, 2011.

[47] C. R. Homeyer and M. R. Kumjian, "Microphysical characteristics of overshooting convection from polarimetric radar observations," *Journal of the Atmospheric Sciences*, vol. 72, no. 2, pp. 870–891, 2015.

[48] J. M. Straka, D. S. Zrnić, and A. V. Ryzhkov, "Bulk hydrometeor classification and quantification using polarimetric radar data: Synthesis of relations," *Journal of Applied Meteorology and Climatology*, vol. 39, no. 8, pp. 1341–1372, 2000.

[49] A. Overeem, R. Uijlenhoet, and H. Leijnse, "Full-year evaluation of non-meteorological echo removal with dual-polarization fuzzy logic for two c-band radars in a temperate climate," *Journal of Atmospheric and Oceanic Technology*, vol. 37, no. 9, pp. 1643–1660, 2020.

[50] G. Vulpiani, M. Montopoli, L. D. Passeri, A. G. Gioia, P. Giordano, and F. S. Marzano, "On the use of dual-polarized c-band radar for operational rainfall retrieval in mountainous areas," *Journal of Applied Meteorology and Climatology*, vol. 51, no. 2, pp. 405–425, 2012.

[51] B. Dolan, S. A. Rutledge, S. Lim, V. Chandrasekar, and M. Thurai, "A robust c-band hydrometeor identification algorithm and application to a long-term polarimetric radar dataset," *Journal of Applied Meteorology and Climatology*, vol. 52, no. 9, pp. 2162–2186, 2013.

[52] A. Ryzhkov, S. Y. Matrosov, V. Melnikov, D. Zrnic, P. Zhang, Q. Cao, M. Knight, C. Simmer, and S. Troemel, "Estimation of depolarization ratio using weather radars with simultaneous transmission/reception," *Journal of Applied Meteorology and Climatology*, vol. 56, no. 7, pp. 1797–1816, 2017.

[53] C. H. B. Lo, T. H. Stein, R. W. Scovell, C. D. Westbrook, T. Darlington, and H. W. Lean, "Use of z dr columns for early detection of severe convection within the operational radar network of the united kingdom," *Meteorological Applications*, vol. 31, no. 1, e2159, 2024.

[54] M. R. KUMJIAN, "Principles and applications of dual-polarization weather radar. part ii: Warm-and cold-season applications.," *Journal of Operational Meteorology*, vol. 1, 2013.

[55] R. J. Doviak, D. S. Zrnic, and D. S. Sirmans, "Doppler weather radar," *Proceedings of the IEEE*, vol. 67, no. 11, pp. 1522–1553, 1979.

[56] G. Vulpiani, M. Montopoli, L. D. Passeri, A. G. Gioia, P. Giordano, and F. S. Marzano, "On the use of dual-polarized c-band radar for operational rainfall retrieval in mountainous areas," *Journal of Applied Meteorology and Climatology*, vol. 51, no. 2, pp. 405–425, 2012.

[57] Y. Wang and V. Chandrasekar, "Algorithm for estimation of the specific differential phase," *Journal of Atmospheric and Oceanic Technology*, vol. 26, no. 12, pp. 2565–2578, 2009.

[58] D. S. Zrnić, A. Ryzhkov, J. Straka, Y. Liu, and J. Vivekanandan, "Testing a procedure for automatic classification of hydrometeor types," *Journal of Atmospheric and Oceanic Technology*, vol. 18, no. 6, pp. 892–913, 2001.