universität wien

MASTERARBEIT | MASTER'S THESIS

Titel | Title

Initiation areas of orographic convection in the Alpine region with a focus on the relation to synoptic weather patterns

verfasst von | submitted by

Daniel Schlager BSc

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien | Vienna, 2024

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 066 614
Studienrichtung lt. Studienblatt | Degree programme as it appears on the student record sheet:	Masterstudium Meteorologie
Betreut von | Supervisor:	Dott. Dott. ric. Stefano Serafin

Kurzfassung

Diese Masterarbeit liefert neue Klimatologien in Zusammenhang mit hochreichender Konvektion, insbesondere der Auslöse konvektiver Zellen, in einem vordefinierten Gebiet (die Alpenregion, 42-49°N und 4-19°E) und für einen Zeitraum von 12 Jahren (Radarreflektivitätsdaten) bzw. 16 Jahren (Blitzdaten). Besonderes Interesse besteht dabei an der Beziehung zwischen der Konvektionsauslöse (CI) und dem vorherrschenden synoptischen Strömungsregime. Sowohl Radar-, als auch Blitzdaten bieten Vor- und Nachteile bei der Analyse von Gewittern (z.B. von Radarstrahlen abgeschattete Gebiete, möglicherweise kurze Dauer der elektrischen Aktivität einer Zelle). Daher werden in dieser Arbeit die beiden Datensätze nicht nur einzeln behandelt, sondern auch gemeinsam, in der Hoffnung, robustere Klimatologien zu erhalten. Hierbei wird ein Ansatz gewählt, der auf der Identifikation und Verfolgung konvektiver Zellen während ihrer Lebensdauer mit Tracking-Algorithmen basiert - ein Verfahren, das typischerweise im Nowcasting verwendet wird. Der bereits bestehende, rein radarbasierte Tracking-Algorithmus T-DaTing (Thunderstorm Detection and Tracking) wird erweitert und an die Nutzung von Blitzdaten angepasst. Außerdem wird ein neues Modul hinzugefügt, das sich mit der Kombination von Radar- und Blitzdaten zu sogenannten Mischzellen beschäftigt. Dies beinhaltet eine Transformation beider Originaldatensätze auf dasselbe raumzeitliche Gitter und ein Glätten der Blitzdaten durch einen Gaußschen Filter, um zellähnliche Objekte zu erzeugen. Darüber hinaus wird eine physikalisch sinnvolle Wahl der T-DaTing-Eingabeparameter für die Verwendung von Blitzdaten vorgestellt.

Die Ergebnisse zeigen, dass Klimatologien von Mischzellen stark von Radarsignalen beeinflusst werden und somit sichtbare Signale der Probleme aufweisen, denen Radar-Messungen in Gebirgsregionen unterliegen. Klimatologien, die ausschließlich auf Blitzdaten basieren, tendieren dazu, robuster zu sein und zeigen keine unphysikalischen Muster, da Blitzsensoren einen nahezu homogenen Blitzerkennungsgrad im gesamten untersuchten Gebiet haben. Häufigkeiten und Klimatologien von 11 verschiedenen synoptischen Strömungstypen werden ebenso diskutiert wie Klimatologien, die nach Monat und Tageszeit stratifiziert sind und einen ausgeprägten saisonalen Zyklus mit einem Maximum im Juli und einen täglichen Zyklus mit einem Maximum in den Nachmittagsstunden zeigen. Darüber hinaus werden Zellen, die einen hohen Einfluss auf die Menschheit und die Umwelt haben (sehr intensive, langanhaltende oder weit ziehende Zellen) explizit untersucht, wobei Gebiete identifiziert werden, die prädestiniert sind für Zellauslöse oder Zellauflösung. Im abschließenden Teil werden einige Statistiken zu Zelleigenschaften und deren Abhängigkeiten sowie deren Beziehung zu den synoptischen Strömungstypen präsentiert, z.B. zeigt sich, dass die Stichprobenverteilungen der Zelllebensdauer und der maximalen Zellgröße gut durch Exponentialfunktionen beschrieben werden. Es wird gezeigt, dass die maximale Blitzdichte und die maximale Zellgröße stark korrelieren: eine Verdopplung der Zellgröße führt zu einer ungefähren Verdopplung der Blitzdichte und damit zu viermal mehr Blitzen im gesamten Gewitter.

Abstract

This thesis provides new climatologies on deep moist convection, especially on the initiation of convective cells, in a predefined domain (the Alpine region, 42-49°N and 4-19°E) and for a period of 12 (radar reflectivity data) and 16 years (CG lightning data) respectively. One particular aspect of interest is the relation between CI and the prevailing synoptic flow regime. Radar and lightning data both offer advantages and shortcomings in the analysis of storms (e.g. radar beam blockage, possibly short period of a cell's electrical activity). Therefore, in this work the 2 datasets are not only processed individually, but also combined in hope of creating more robust climatologies. Hereby, an approach based on identifying and tracking convective cells during their lifetime with tracking algorithms is chosen, a procedure typically used in nowcasting. The already existing purely radar-based tracking algorithm T-DaTing (Thunderstorm Detection and Tracking) is being extended and adapted towards the use of lightning data. Furthermore, a new module dealing with the combination of radar and lightning data to so-called mixed cells is added. This also requires a transformation of both original datasets onto the same spatiotemporal grid and a smoothing of lightning data by a Gaussian filter to create cell-like objects. Moreover, a physically reasonable choice of T-DaTing input parameters for the use of lightning data is presented.

The results show that climatologies of mixed cells are strongly affected by radar signals, and thus show visible signals of the problems radar measurements face in mountainous areas. Climatologies based on lightning data alone tend to be more robust and do not show unphysical patterns, because lightning sensors provide a nearly homogeneous detection efficiency across the studied domain. Frequencies and climatologies for 11 different synoptic flow types are discussed, as well as climatologies stratified by month and time of day, revealing a pronounced seasonal cycle with a maximum in July, and a daily cycle with a maximum in the afternoon hours. In addition, cells bearing high impact on people and the environment (very intense, long-lasting or far-moving cells) are explicitly investigated, thereby detecting areas prone to cell initiation and others prone to cell decays. In a final part, some statistics on cell properties and their interdependence and relation to synoptic flow types are presented, e.g. revealing that the sample distributions of cell lifetime and maximum cell size are well described by exponential functions. Maximum lightning flash density and maximum cell size are shown to be strongly correlated: a doubling in cell size leads to an approximate doubling of flash density and hence 4 times more flashes in the whole system.

Acknowledgements

First of all, I would like to express my deepest gratitude to my supervisor Dr. Stefano Serafin for his constant and valuable support, his ongoing guidance, feedback and encouragement. He not only greatly helped in casting my principal ideas into a proper thesis topic, but also showed so much patience with me being a student and a part-time employee at the same time.

I am also extremely grateful to Rudolf Kaltenböck from Austro Control (ACG) for providing me with data of their weather radar network and the rights to use them. Similarly, Wolfgang Schulz from the Austrian Lightning Detection and Information System (ALDIS) is hereby greatly thanked for supplying cloud-to-ground lightning flash data. Moreover, I would like to acknowledge the straightforward download possibility of ERA5 reanalysis data in the Climate Data Store of the Copernicus Climate Change Service.

Many thanks go to Dr. Monika Feldmann for taking time to answer my questions on the T-DaTing tracking algorithm introduced in one of her studies. I am also thankful to all contributors for making publicly available such great Python-packages as pySTEPS (PySteps developers, 2018-2024) and Cartopy (Met Office, 2010 - 2015) including Natural Earth - free vector and raster map data.

I would like to extend my heartfelt thanks to my partner Sanjiv Shrestha, who supported me in all phases of this work, whether they were hard or easy. He continuously motivated me and gave valuable feedback. And special thanks also got to my parents and all family members for their constant moral support, motivation and time.

Lastly, thanks also go to all other friends, study mates and people that were directly and indirectly involved in bringing this work to an end.

Contents

List of Figures

List of Tables

1. Introduction and Scientific Status Quo

Deep moist convection (hereafter DMC) and its related meteorological phenomena (such as heavy precipitation, hail, lightning strikes, strong winds or even tornadoes) are among weather events that have the highest impact on people, infrastructure and the environment. Therefore, it is crucial to know about and ideally even predict the origin, moving behaviour and internal evolution of convective cells. One particular aspect is the relation between the location of origin and the current synoptic conditions.

The two most common meteorological parameters that immediately come into mind when thinking about thunderstorms are precipitation (remote sensed via radars) and electrical activity (remote sensed via lightning detection sensors). Both of these data types offer advantages and shortcomings in the analysis of storms. On the one hand radar signals can be measured at many stages of a cell's lifetime, but radar beams may be blocked by mountains or other obstacles. While lightning sensors have relatively uniform detection capabilities, some storms may not be electrically active. This work aims at combining the advantages of radar and lightning data and should add a new view on DMC in the Alpine region by providing hopefully more robust climatologies (such as the spatiotemporal distribution of convective initiation, hereafter CI). An approach based on identifying and tracking convective cells during their lifetime with tracking algorithms is chosen, as this offers the possibility to not only analyse the initiation of storms but also their subsequent evolution. Although such algorithms are predominantly applied for nowcasting, it is justified to use them in a climatological sense, if the same method on homogeneous datasets is implemented.

Having this objective in mind, Chap. 1 gives an overview of DMC basics, followed by a selection of previous scientific studies on DMC climatologies. Furthermore, the basic structure of tracking algorithms is explained, examples of already existing algorithms are given and their characteristic features contrasted. In Chap. 2 the aim of this thesis is explained and research questions are formulated. In Chap. 3 comes a thorough description of the datasets and methods used to answer these questions, including a discussion on the principles of a radar-based tracking algorithm proposed by Feldmann et al. (2021). Moreover, the steps needed to adapt this algorithm for the use of other data sources are presented. This rather rare approach of using a tracking algorithm to come up with climatologies on DMC leads to a number of interesting results shown and interpreted in Chap. 4. Apart from the focus on DMC initiation, another one is laid on the behaviour of high impact cells (very intense, long-lasting or far-moving ones) and some basics of cell properties and their dependence on each other are given. Finally, conclusions drawn from the method and results are summarized in Chap. 5, including a list of the most relevant findings.

1.1. Deep moist convection in mountainous areas

To describe the conditions under which thunderstorms may develop an ingredient-based approach can often be found in the literature. For example, Johns and Doswell (1992) mention 3 major ingredients: (i) a sufficiently deep moist layer in the lower or mid parts of the troposphere, (ii) a steep enough lapse rate (decrease of temperature with height) and (iii) lifting of air in the moist layer far enough in order to release the instability created by the other two.

Conditional instability

The process of DMC initiation is often described by assuming and analysing the behaviour of a lifted parcel of air. To better understand this concept, some definitions need to be made (following the nomenclature in Kirshbaum et al. (2018)):

- $\Gamma = -\partial T_p/\partial z$ ... parcel lapse rate, the decrease in temperature T of a parcel of air when being lifted a vertical distance of z
- $\gamma = -\partial T_e/\partial z$ ... environmental lapse rate, the decrease in temperature T of a given atmospheric state along the vertical coordinate z
- $\Gamma_d \approx g/c_p \approx 9.8\ Kkm^{-1}$ ... dry adiabatic lapse rate (where g is Earth's gravitational acceleration and c_p is the heat capacity of air at constant pressure), cooling rate of unsaturated parcels of air when lifted adiabatically
- Γ_s ... saturated adiabatic lapse rate, cooling rate of a vertically lifted parcel of air in a saturated adiabatic process (involving condensation and latent heat release)
- LCL ... lifting condensation level, level at which a surface-based parcel of air reaches saturation when being lifted
- LFC ... level of free convection, level at which a lifted parcel of air becomes warmer than its environment and therefore positively buoyant
- LNB ... level of neutral buoyancy, level at which a freely ascending parcel of air reaches the same temperature as the environment again
- CAPE ... convective available potential energy, integrated positive buoyancy between LFC and LNB (can be seen as a combined measure for the ingredients instability and moisture (Holley et al., 2014))
- CIN ... convective inhibition, integrated negative buoyancy below LFC

Now, one can assume a parcel of air near the ground that is subject to some sort of lift. At first the unsaturated parcel will cool according to Γ_d, until it reaches its LCL. There condensation of water vapor occurs releasing latent heat, which causes the cooling rate along the further ascent to be lowered to Γ_s. When there is a point at which the temperature of the parcel T_p becomes higher than the temperature of the environment T_e (the LFC), conditional instability ($\Gamma_s < \gamma < \Gamma_d$) is released. The parcel becomes positively buoyant and due to its CAPE accelerates away from its original position without further need of an external source of lift, forming the prototype of a convective updraft. This process can also be studied with the use of atmospheric sounding data printed on thermodynamic diagrams.

It should be kept in mind, that the description above involves simplifications. For example, entrainment and detrainment processes between the lifted parcels and the surrounding air are ignored that could possibly reduce buoyancy.

Sources of lift

Apart from the latent instability, some source of lift is needed for the onset of DMC. Markowski and Richardson (2010) suggest several types of forced ascent, such as air mass boundaries (synoptic fronts, drylines, outflow boundaries of preexisting storms, sea breezes) or orographic circulations driven by differential heating of elevated or sloped terrain. Here mountainous areas come into play. Kirshbaum et al. (2018) divide mechanically and thermally forced trigger mechanisms of moist orographic convection. Examples of mechanically driven triggers are shown in the top 3 schematics of Fig. 1.1. In an unblocked flow, air can directly ascend the slope (a), whereas in a blocked flow convergence and possible uplift occur upstream of the mountain (b) or in form of lee-side convergence (c). Thermally driven triggers are the onset of anabatic winds during the day and convection over the crest (d) versus katabatic winds and convection near the mountain base during the night (e). Finally, combined trigger mechanisms include an optimal overlay of thermally driven flows at the slopes with gravity-wave ascent aloft (f) or cold pools of already existing storms interacting with the impingent flow (g).

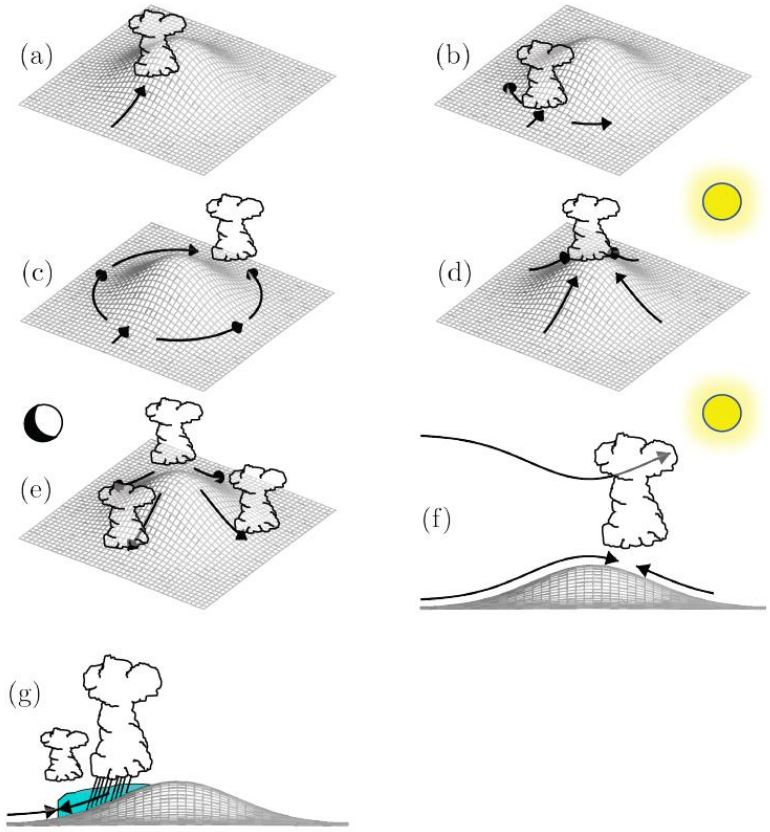

Figure 1.1.: Orographic convection can be triggered in a variety of ways: mechanical (a-c), thermal (d-e) or combined (f-g). Figure from Kirshbaum et al. (2018), for details see text.

1.2. Climatologies of DMC in the Alps and beyond

The variety of effects described in Chap. 1.1 - relating orography, lifting of air and the onset of DMC - leads to a clear impact of mountainous areas such as the Alps onto thunderstorm activity, which makes them preferential regions of research. Climatologies of convective activity can be computed using various observations and parameters, e.g. including flash density. Hereby, high flash density rates tend to occur near the edges of mountain ranges, whereas right above the main crest lower flash numbers are counted (e.g. see Fig.18 from Anderson and Klugmann (2014), that shows accumulated monthly flash numbers for 2008-2012 measured by the Arrival Time Differing NETwork (ATDnet) operated by Met Office). Also in Anderson and Klugmann (2014), a discussion of lightning detection efficiency is included, which is said to be not absolutely uniform across all of Europe, but still relatively continuous.

Apart from Anderson and Klugmann (2014) there are many other studies on DMC climatologies based on different storm features and different regions. An extensively researched area is e.g. the Po Valley and the adjacent Friuli Venezia Giulia region in Northern Italy, as the latter accommodates one of Europe's cloud-to-ground (CG) lighting maxima (Feudale and Manzato, 2014). In this study a climatology for 18 years of CG lightning data is computed and its variations in time and space (divided into coastal areas, plains and mountains) are studied. Cacciamani et al. (1995) and Costa et al. (2001) examine the synoptic, dynamic and thermodynamic conditions that favor the development of thunderstorms in the Po valley area, finding that the interaction between synoptic scale forcings and orographic effects plays a major role. An analysis of various instability indices and their applicability to the region is given in Manzato (2003), derived from atmospheric soundings near the city of Udine. Also climatologies on thunderstorm occurrence and rainfall (Manzato, 2007) as well as on hail (Manzato, 2012) exist. The latter is of special interest, as a combination of several factors makes the area prone to great agricultural losses in connection to severe hail events (Morgan, 1973).

Zooming out of Northern Italy to the whole Alpine region, climatologies of CG flashes and convective initiation (CI) events are provided by Manzato et al. (2022). This is done for the years 2005 to 2019 and the results are stratified by 3 orography classes, time of the day and time of the year. CI events are identified through applying a spatiotemporal mask on the CG flash data, which in turn involves the choice of a set of empirical thresholds. Moreover, a difference measure between the spatial CG and CI distributions is computed to distinguish between areas of thunderstorm activity predominantly due to in situ initiation or due to advection of already existing cells. The methods and results by Manzato et al. (2022) serve as a basis for the work that is presented in this thesis.

Going to even larger scales, a study by Taszarek et al. (2019) needs to be mentioned. They try to develop a pan-European thunderstorm climatology using data of 39 years and multiple data sources, such as soundings, surface observations, lightning data, reanalysis data and severe weather reports. The results include a temporal and spatial analysis of thunderstorm occurrence and peak activity times, as well as a climatological trend analysis. Moreover, it is stated that lightning is the most objective data source, but it has shortcomings due to its rather short period of availability and a lack of sensors in parts of Europe.

Also country-wise climatologies can be found. For example, Antonescu and Burcea (2010) provide a Romanian CG lightning climatology based on 4 years of data and a spatial resolution of 20km. In the next 2 paragraphs studies focusing on Austria will be presented, using lightning and radar data respectively.

Determining thunderstorm tracks in Austria based on lightning data

Bertram and Mayr (2004) analyse cloud-to-ground lightning data of Austria's lightning detection system ALDIS for the summers of 1993-1999, where only a part of the domain is chosen to reach relatively uniform detection efficiency. On days with at least 3000 flashes in the domain a tracking approach is carried out, which consists of (i) filtering weak storms, (ii) identifying storm centers and (iii) connecting storm centers between subsequent times. The resulting tracks are categorized into six classes of different 700 hPa-winds, which is supposed to be the steering level for thunderstorms. Finally, the dominating track paths for each flow type are depicted in their Fig.14. The resulting patterns show remarkable differences for each flow type, but rather similar areas of origin, namely mountain ranges of moderate altitude around 2500 m above sea level.

Radar-based severe storm climatology for Austria

Kaltenboeck and Steinheimer (2015) analyse thunderstorm activity based on images of accumulated radar derived precipitation fields of the Austrian weather radar network during the convective seasons of 2008-2012. The results are stratified based on atmospheric instability (i.e. CAPE), vertical wind shear and the choice of four flow classes on the 500 hPa level. Major findings include that westerly and southerly flows are related to more widespread storm development over Austria and strong deep-layer shear tends to favor organized line patterns over wide areas.

In Fig. 1.2, an image of the accumulated precipitation during the months April to September of 2008-2012 is provided. Artificial structures in the precipitation field are visible that are related to various difficulties radar measurements in complex terrain face. For example, Germann et al. (2022) mention harsh weather conditions at mountain tops, challenging maintenance of remote stations and, most importantly, the reflection and shielding of the radar beam by other mountains.

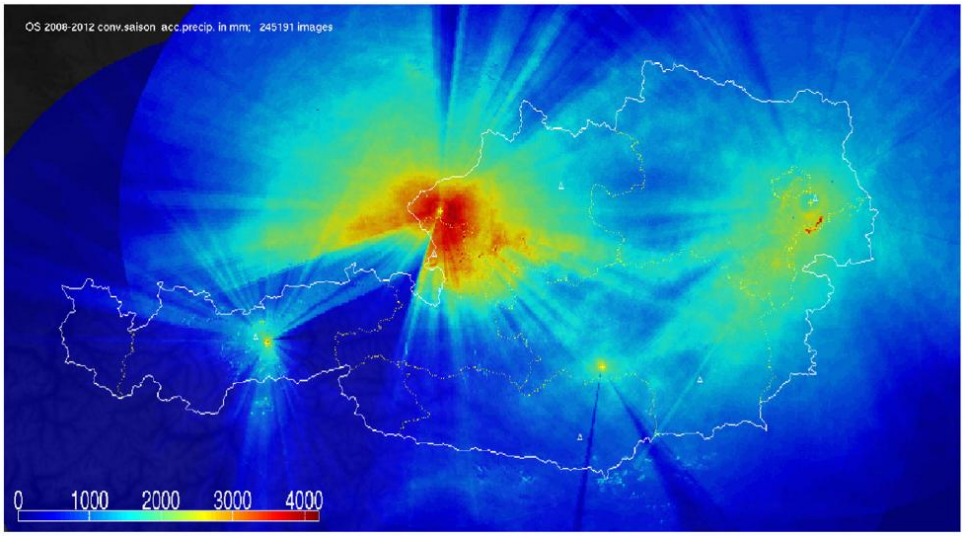

Figure 1.2.: Radar derived precipitation sum for five convective seasons, that shows artificial structures due to challenges of radar measurements in complex terrain (Kaltenboeck and Steinheimer, 2015). Used with permission[1].

[1]Reprinted from Atmospheric Research, 158-159, Rudolf Kaltenboeck and Martin Steinheimer, Radar-based severe storm climatology for Austrian complex orography related to vertical wind shear and atmospheric instability, 216-230, Copyright (2015), with permission from Elsevier.

Climatologies making use of satellite imagery

A source that has not been mentioned so far is data retrieved by satellites, especially their infrared channels as these allow the same measurements during day and night. When using stallite data for thunderstorm detection, the challenge is separating cumulonimbus anvils from other non-convective clouds.

Morel and Senesi (2002a) try to build a European climatology of mesoscale convective systems (MCS) based on infrared satellite images. They do so by running 2 algorithms, one for cloud-tracking and one for the discrimination between convective and non-convective clouds. The cloud-tracking requires 3 thresholds, a temperature and an area threshold for defining MCS cloud shields and an overlapping threshold to attribute each MCS to its equivalent in successive images. This is one of the rare examples of the usage of tracking algorithms to come up with DMC climatologies (this thesis should add another, but instead of data retrieved from satellites radar and lightning data are used). The discrimination algorithm then uses a parameter based on brightness temperature gradients at the MCS edges to seperate convective from non-convective cases, as these tend to be sharp for developing systems. The results in (Morel and Senesi, 2002b) suggest a mainly continental occurrence of MCS in Europe, often triggered by orographic features and being in phase with the diurnal cycle of radiative heating. Typical are a propagation towards the east-north-east and a life time of around 5.5 hours (from approximately 3 pm to 9 pm LST).

Other applications of satellite-based data include Funatsu et al. (2018) (analysing convective overshooting), Levizzani et al. (2010) (computing a warm-season climatology of precipitating clouds) or Cancelada et al. (2020) (back-tracking severe convective systems to their origin). Though possible, using satellite data is beyond the scope of this thesis.

Benefits and shortcomings of using different datasets

Already with a limited number of studies presented, different datasets seem to provide different challenges when it comes to analysing DMC. To conclude this chapter, in Tab. 1.1 advantages and disadvantages of the use of radar (R), lightning (L) and satellite data (S) are contrasted.

	Advantages	Disadvantages
R	Signal appears early in the life cycle of a convective cell and shows a rather smooth evolution with time.	Measurements face various difficulties (such as beam blockage) and are therefore not uniform across the domain.
L	Detection efficiency is homogeneous across the domain to a high degree.	Electrical activity within cell lifetime can be rather short. Even severe cells do not necessarily produce lightning.
S	As clouds represent the first stage of cell development, initial detection lies close to actual initiation location. Detection is uniform across the domain.	Cumuli triggering can be shaded by clouds at higher altitudes (e.g. by anvils of pre-existing storms). Resolution of all-day available (non-visible) satellite channels may be too coarse.

Table 1.1.: Remote sensing techniques and their benefits and shortcomings for analysing DMC initiation and evolution.

1.3. Algorithms for detection and tracking of convective cells

Many different tracking algorithms can be found in the literature, all having their individual set ups and specialities. Still, the majority of them, if not all, follow a few basic steps:

1. Detection/identification of cells
2. Advection/propagation of identified cells to the next timestep
3. Matching the advected cells with the newly identified cells of the next timestep (thereby in some way accounting for cell splits and mergings)

Four examples of already existing tracking algorithms are given in the following paragraphs. Even though all of the chosen algorithms are purely radar-based, the implementation of the 3 basic steps varies greatly not only when it comes to their basic principals but also to their level of complexity. The tracking algorithm used in this thesis (T-DaTing) will be introduced and explained in more detail in Chap. 3.2. T-DaTing has been more recently developed, is publicly available as a Python-package and has already been applied to parts of the Alps.

For this first short insight though, TITAN (Thunderstorm Identification, Tracking, Analysis and Nowcasting by Dixon and Wiener (1993)), SCIT (Storm Cell Identification and Tracking by Johnson et al. (1998)), TRACE3D (by Handwerker (2002)) and CELLTRACK (by Kyznarová and Novák (2009)) have been chosen.

First differences already appear when looking at the form of the input data. While TITAN uses 3D reflectivity data that is transformed onto a cartesian grid, SCIT and TRACE3D use a polar coordinate system. CELLTRACK relies on a cartesian grid transformation again, but only uses a maximum reflectivity field, so 2D input data.

When it comes to detecting cells, TITAN and CELLTRACK mainly rely on two thresholds, one for a minimum reflectivity (35 dBZ and 44 dBZ, respectively) and the other for a measure of extent (a volume of at least 50 km^3 and an area of 4 km^2, respectively). Although, in CELLTRACK a detection scheme with a variable reflectivity threshold is also available. TRACE3D also uses a variable reflectivity threshold in relation to close lying maxima, whereas SCIT not only takes different reflectivity thresholds into account, but also thresholds for many other parameters such as segment length or overlap. Furthermore, there is also a difference in how cell centroids are calculated, which can be simple geometrical centres (e.g. CELLTRACK) or reflectivity-weighted ones (e.g. TITAN).

Advection of identified cells can for instance rely on motion fields of other algorithms (CELLTRACK) or motion vectors of previous time steps (SCIT and TRACE3D). TITAN does not even need an advection part, as the matching is based on a combinatorial optimisation scheme including e.g. size and travelled distance of a cell. On the other hand SCIT's matching approach is rather simple, using the smallest distance between a newly identified cell and all advected cells in the surroundings. Distance is also considered in CELLTRACK, but a shape similarity measure is added. And TRACE3D implements a stepwise process of dropping the least credible assignment until only one final connection remains.

Finally, splits and mergings can be considered through searching for significant changes in cell area (TRACE3D) or though finding origins/endpoints of tracks lying within areas assigned to another cell (TITAN and CELLTRACK).

Tracking algorithms using more than one dataset

So far, only tracking algorithms for analysing radar data have been presented. An algorithm that aims to work for radar and lightning data is introduced in Steinacker et al. (2000). After transforming both datasets to the same spatiotemporal grid, 3 major steps are carried out: at first, a smoothing of the respective input field with a simple Gaussian filter, where in Eq. 1.1 n is the distance from the centered grid point in multiples of the grid width and k is a factor that defines the filter width (and therefore the scale of the system to be analysed, i.e. individual convective cells or cell clusters):

$$g = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{n}{k}\right)^2\right] \tag{1.1}$$

g serves as a weighting factor for the field variables. Secondly, maxima in the input field are identified via setting certain thresholds of lightning density and precipitation rate. Finally, a multi-step selection procedure is carried out to find the most likely displacement vectors between maxima of two consecutive time steps.

Radar and lightning data can not only be processed independently by tracking algorithms, but also in a combined form as in Meyer et al. (2013). There, by adapting a method for tracking and monitoring satellite-data based clouds (Cb-TRAM, Zinner et al. (2008)) and radar data (rad-TRAM, Kober and Tafferner (2009)) a new algorithm for lightning (li-TRAM) is introduced. Then the output of both rad-TRAM and li-TRAM is combined towards a tracking and monitoring algorithm for electrically charged cells (ec-TRAM). In detail, rad-TRAM and li-TRAM identify individual cell objects for each dataset and ec-TRAM combines these radar and lightning cells by searching for spatial overlaps. Hereby, the outline of the new so-called ec-TRAM cells is defined by following the outermost contours of the underlying radar and lightning cells. Finally, ec-TRAM cells are tracked following the schematic in Fig. 1.3, where the track of an ec-TRAM cell at each time step (in dashed contours) is following the track of the underlying cell that appears to be biggest after the connection (for example, at connection 4, the lightning cell track is used as cell L2 afterwards is bigger than cell R1). After analysing the results obtained with this new combined algorithm, it is even stated that "the hybrid method ec-TRAM outperforms rad-TRAM and li-TRAM regarding reliability and continuous assessment of strom tracks especially in more complexly developing storms". Using more than one data set therefore seems to improve automatic cell tracking.

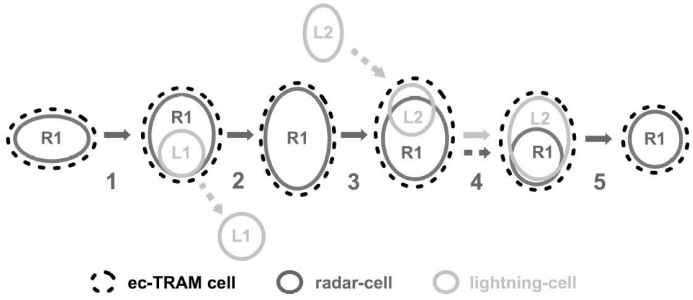

Figure 1.3.: Schematic of tracking ec-TRAM cells. Figure from Meyer et al. (2013).

2. Aims of this Thesis

The overall objective of this work is to gain new insight into the occurrence, especially the initiation, and evolution of deep moist convection (DMC) in the Alpine region during recent years. Also a stratification by type of synoptic flow regime is sought.

Hereby, data should be derived from tracks of individual convective cells, calculated by a cell detection and tracking algorithm. As previously stated, this approach of calculating climatologies is not used very often, but could offer the potential for new and interesting insights. It needs to be mentioned, that an approach involving tracking algorithms is more demanding than e.g. just adding up flash counts to flash density distributions. But it also provides the advantage of keeping temporal information that can be used to assess cell evolution or distinguish preferential regions of convection initiation from regions of frequent cell passages and areas of predominant cell decay.

Due to its public availability as part of the Python-package pySTEPS (PySteps developers, 2018-2024), the algorithm T-DaTing (Thunderstorm Detection and Tracking) proposed by Feldmann et al. (2021) and based on Hering et al. (2004) is adapted to the needs of this work. As using different data sets seems beneficial for tracking the evolution of DMC (compare Meyer et al. (2013) described in Chap. 1.3), the purely radar-based T-DaTing is also tuned towards the use of lightning data and a combined tracking of radar and lightning cells (introducing so-called mixed cells). A description of T-DaTing, the thresholds that need to be set and the major changes being made is given in Chap. 3.

It is the hope that a combined run can benefit from the uniform efficiency of lighting detection in areas where radar beams are partly or fully blocked. On the other hand radar data offers the possibility of early cell detection and of keeping severe convective events in the analysis even if no electrical activity is at play, e.g. in heavy showers of rain.

Finally, Fig.8 from Manzato et al. (2022) showing a lightning CI climatology should be compared to the climatologies generated in this thesis, as the underlying lightning dataset is the same (except for a one-year time difference). The evaluation method differs though.

Research questions

1. How can a tracking algorithm originally designed for one data source be objectively adapted for the use different data sources (e.g. radar and lightning)?

2. How much do DMC climatologies benefit from combining radar and lightning data?

3. How do DMC initiation and evolution in the Alps vary with synoptic flow regimes, time of year, time of day and other stratification variables?

4. How much do cell numbers vary in relation to cell lifetime, size and intensity?

3. Data and Methods

3.1. Radar, lightning and synoptic flow data

Thunderstorm occurrence and characteristics can vary over a wide range of spatial and temporal scales. Even whole convective seasons can extensively differ from one another when being compared interannually. Thus, for deriving climatological statements on DMC in the Alps it is reasonable to process datasets that are as long as possible, ideally homogeneous and for the greater part free of system outages. In the following paragraphs the three acquired data sources for this work are presented.

Radar data

Radar data is provided by Austro Control (ACG), Austria's main air service provider. The Austrian weather radar network consists of 4 main stations - 2 at lower altitudes and 2 at mountain sites above 2000 m above sea level (Kaltenboeck, 2012). These have been upgraded to C-band dual polarised radars with a peak power of 500 kW in the years 2011 to 2013. The radars are installed at Rauchenwarth in the east, Feldkirchen in the north, Zirbitzkogel in the south and Patscherkofel in the west.

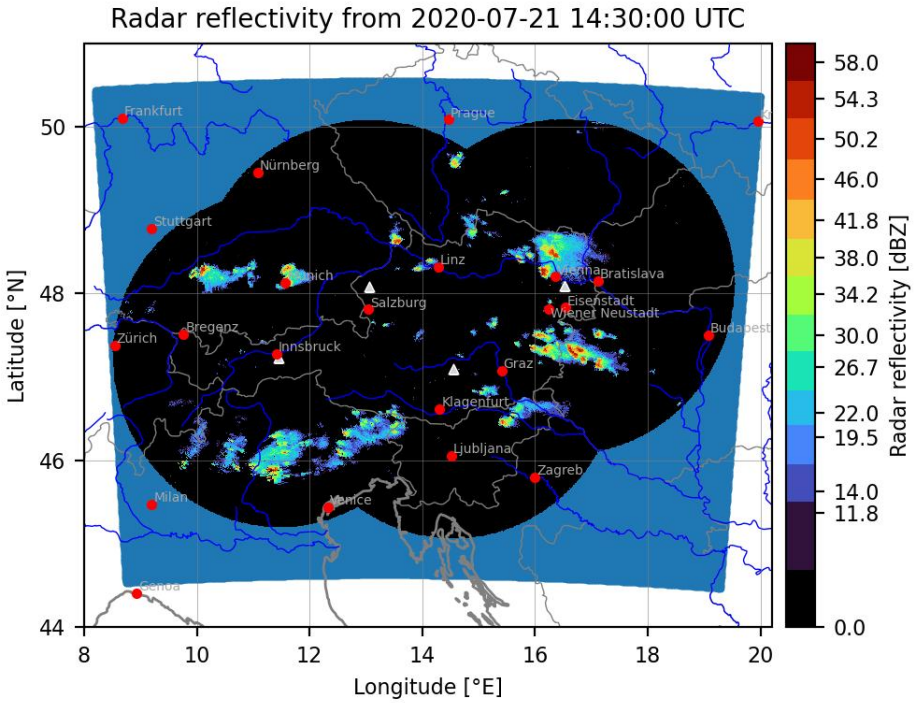

Figure 3.1.: Example image of original radar data for a mid-afternoon case in July 2020.

Before the mentioned change of radar hardware, measurements were taken with a bottom to top scan strategy of 16 elevations lasting around 5 minutes. Since then an interleave scan strategy has been followed, consisting of 2 half scans with 8 elevations each that overlap to a full volume scan approximately every 5 minutes. As these scans are carried out from top to bottom, the lower elevations with the highest range (of more than 200 km) are the most recent ones.

For this study, 2D maximum reflectivity measurements from the 4 main radar stations are used. The data has been received as a series of monthly netCDF-files for 12 consecutive years (01 January 2009 - 31 December 2020) with a temporal resolution of 5 minutes, a spatial resolution of around 1 km (on a cartesian grid) and 14 discrete reflectivity levels. An example image is shown in Fig. 3.1.

Radar data is usually not always free of system outages, which is why statistics on radar availability have been computed. The radar with the least missing images is Rauchenwarth, its availability reaches about 98.4%. For the other three the percentages are similar: 94.5% for the radar at Patscherkofel, 93.7% for the one at Zirbitzkogel and 93.6% for the one in Feldkirchen. The distribution of outage times for individual months can be seen in Fig. 3.2a. For the period of January to May and in December all radar sites show rather short outage times of hardly more than a total of 15 days in the 12 year period. From June to November however, all radars except for the one at Rauchenwarth experience much longer periods of missing data. To analyse that further, outage times have been plotted on a daily basis, an example for Patscherkofel being shown in Fig. 3.2b. From there it can be seen that most of the missing availability comes from one summer/autumn in 2013, and similar effects can be found for Zirbitzkogel in 2012 and for Feldkirchen in 2011. This is an effect of the system upgrade in those years, which unfortunately coincides with the convective season. Aside from the maintenance periods, the datasets are continuous enough to theoretically provide robust DMC climatologies, if it were not for other radar deficiencies such as beam blocking.

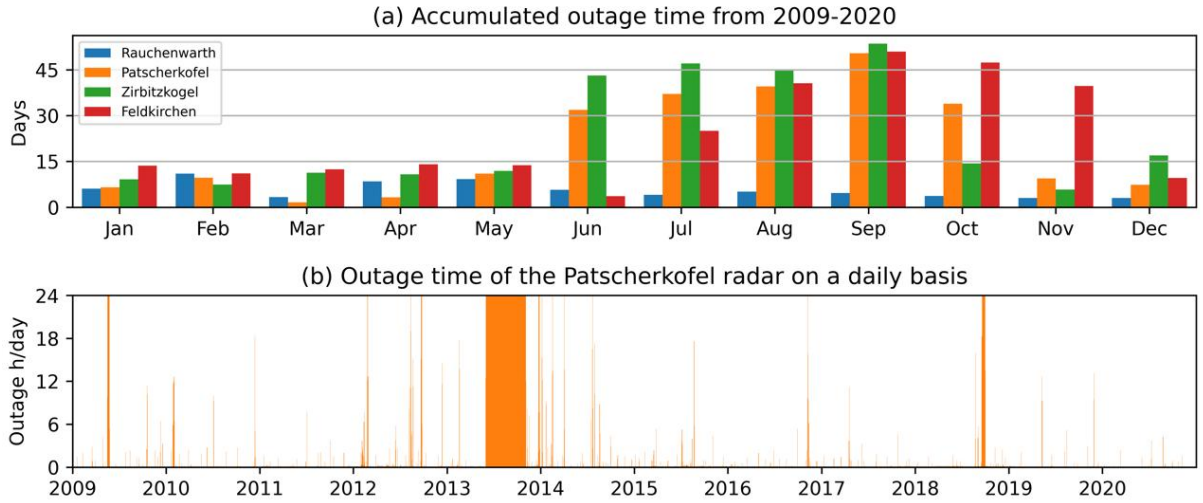

Figure 3.2.: Statistics on the availability of Austrian radars. (a) on a monthly basis for all 4 radars and (b) on a daily basis for the radar at Patscherkofel.

Lightning data

Lightning data is gathered by EUCLID, the EUropean Cooperation for LIghtning Detection, and provided by ALDIS, the Austrian Lightning Detection and Information System. In the way of combining data from several networks a complete dataset for the Alpine region and its surroundings (42-49°N and 4-19°E) is formed. In 2020 around 180 sensors were part of the EUCLID network (Manzato et al., 2022), where the ones located in the mentioned domain are indicated as red crosses in Fig. 3.3. For this domain, the spacing of sensors and their type are homogeneous and the orography is said to have no impact on the network's performance (Manzato et al., 2022). Furthermore, studies suggest that cloud-to-ground (CG) flash detections in Austria provide a median location accuracy of around 100 m and a detection efficiency of around 97% (Schulz et al., 2016). While CG data is homogeneously available for at least 15 years, the collecting period for intra-cloud (IC) lightning data in the EUCLID network is much shorter, dating back until 2015 (Manzato et al., 2022). For a study on climatological scales as this thesis, it is therefore reasonable to limit the analysis to CG data.

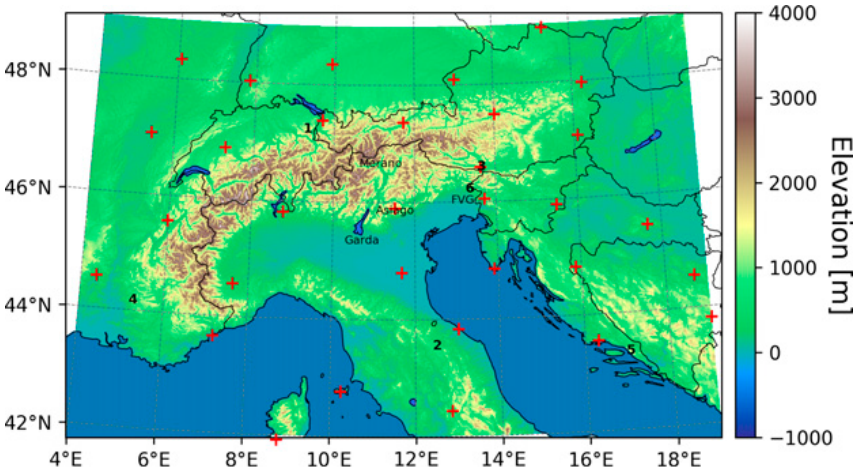

Figure 3.3.: Domain of lightning data, its orography shading and EUCLID network sensor locations (indicated as red crosses). Figure from Manzato et al. (2022).

The original dataset for this work contained information on single CG lightning flashes in a period of 16 consecutive years (01 January 2005 - 31 December 2020). The data was then mapped onto a lat-lon-grid by counting flash numbers for each grid box. This transformation yielded monthly netCDF-files with a temporal resolution of 10 minutes (timestamp marks the beginning of each period) and a spatial resolution of 0.02° in the North-South direction and 0.025° in the West-East direction. The latter corresponds to a pixel size of approximately 2x2 km².

Transformation of radar and lightning data

To achieve general comparability between both datasets and to be able to combine them in a mixed cell approach, a transformation onto the same spatial grid is necessary. Temporal consistence can on the other hand easily be reached by processing every second slice of the radar data only. An example for the changes in both datasets is given in Fig. 3.4.

The spatial transformation for radar data is a simple nearest neighbour interpolation onto the more regular lightning grid. This comes at a cost, because the resolution is going down by about a factor of 2.

For lightning data, the transformation involves more work, as in a relatively fine grid flash counts of one and the same cell can be separated by empty pixels (that is usually not the case for reflectivity fields). To create a cell-like pattern it is therefore necessary to introduce a smoothing procedure based on a Gaussian filter, similar to the one described in Steinacker et al. (2000). As a first step flash counts are converted into a flash density (with units $km^{-2}h^{-1}$)). Following the suggestions in Fig.2 and Fig.3 of Steinacker et al. (2000), a filter with a standard deviation of $\sigma = 1.0$ is chosen and this filter is being truncated at 10% of the central value of the Gaussian curve (i.e. at $\sqrt{-2\ln 0.1} \approx 2.146$ standard deviations). This corresponds to a filter width of 2 grid boxes in each direction, the value of one grid point is therefore spread out to 24 surrounding ones around it (in a rectangular configuration). It is important to note that during this process the value of one grid box is just distributed to several boxes and the sum of the new values remains the same as the original value.

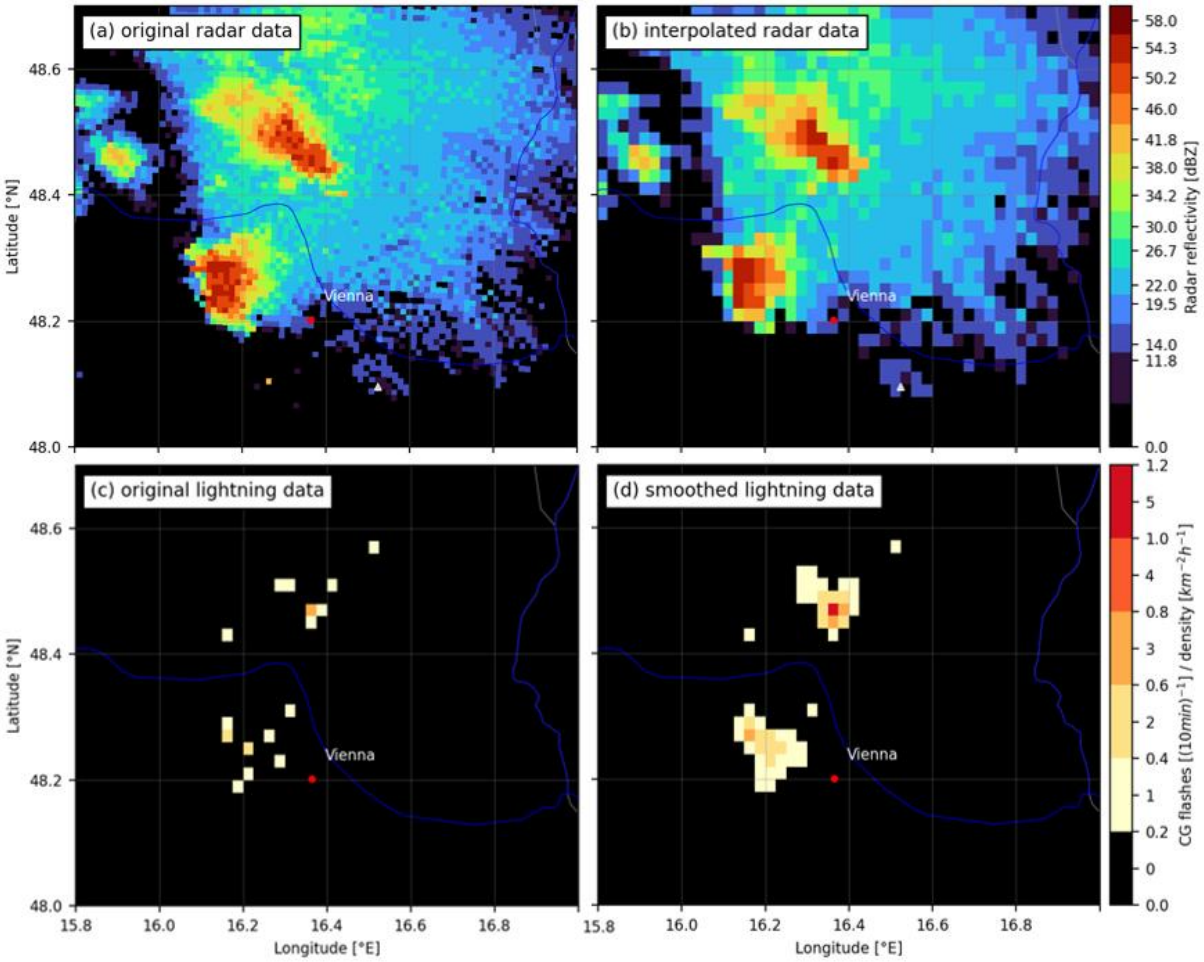

Figure 3.4.: Effects of interpolating radar data and smoothing lightning data with a Gaussian filter for a case of 2 pronounced thunderstorms in the Vienna region on 21 July 2020, 14:30 UTC. Resolution of the original radar data (a) goes down by a about a factor of 2 after the interpolation (b). Scattered flash counts (c, colorbar integers) become contiguous patches of smoothed flash density (d, colorbar decimals).

Synoptic flow data

Winds at mid-tropospheric levels are found to be the main drivers for the direction of thunderstorm propagation, where different studies often take the 700 hPa pressure levels as steering level (e.g. see remarks in Bertram and Mayr (2004)). Hence, it makes sense to stratify the resulting climatologies based on the flow direction on these levels. For this purpose an automated weather type classification scheme is applied, based on the so-called Jenkinson-Collison classification in the version of Otero et al. (2018) and available as a Python package named jcclass (Herrera-Lormendez, 2022). Mean sea level pressure data is suggested as input field, from which geostrophic winds and vorticity are being calculated. The dominant circulation type is then attributed to each grid point by looking at wind and vorticity values at 16 other points, that are uniformly distributed around it and extending out as far as 10° in each direction (compare second figure in Herrera-Lormendez (2022)). Moreover, it is possible to choose a broad (27) or narrow (11, Tab. 3.1) range of individual weather types.

Class	-1	0	1	2	3	4	5	6	7	8	9
Flow type	LF	A	NE	E	SE	S	SW	W	NW	N	C

Table 3.1.: Weather classification scheme output and the respective synoptic flow types (Herrera-Lormendez, 2022). LF stands for low flow, A for anticyclonic, C for cyclonic and all other letters for the commonly used wind directions.

Although the original version of the algorithm uses ground level data, it can be argued to use geopotential data at higher altitudes as a starting point (as friction tends to become less and less important with height and the geostrophic assumption better fulfilled). For stratifying the resulting climatologies of this work, 700 and 850 hPa are chosen as steering levels and the respective geopotential data for 12 UTC of each day in the period 2005-2020 is obtained from the ERA5 reanalysis. Then the flow type for each ERA5 grid point is computed, first for the whole lightning domain and second for the part of it where radar data is available. Finally, the most frequent classes for both domains and both levels as well as their respective pixel fractions of the whole domain are saved for further use.

3.2. The radar-based tracking algorithm T-DaTing

In Chap. 2 the tracking algorithm T-DaTing has been introduced briefly. Feldmann et al. (2021) used it as one part of a study on mesocyclone occurrence and frequency in the Alpine region, where it was tuned on the detection of supercells in the Swiss operational radar network.

As most tracking algorithms, T-DaTing also follows the 3 steps presented in Chap. 1.3. Detection of cells is based on a multi-threshold approach with different reflectivity values, minimum cell size and minimum distance from each other. Advection of the identified cells is then performed via an optical flow estimation (Pulkkinen et al., 2019) implemented in the Python-package pySTEPS (PySteps developers, 2018-2024). Finally, the matching is primarily based on a spatial overlap threshold of advected cells and newly identified ones in the next time step. Important details of T-DaTing are explained in the following sections, whereas an in-depth-listing of all changes being made to the original code is provided in the appendix B.

Input and thresholds

T-DaTing requires 2 main input variables, (i) a 3D-array of shape $t \times m \times n$ referring to time, latitude and longitude, and (ii) a list of corresponding timestamps. Furthermore, a number of thresholds need to be set, for this study mostly following the suggestions in Feldmann et al. (2021):

1. t_{min} - minimal number of cell identifications in consecutive time steps in order for a track to be saved. The default is 3 time steps corresponding to a cell life time of at least 10 minutes (Feldmann et al. (2021) use the same number for a dataset with the same temporal resolution of 5 minutes).

2. R_{min} - lower threshold for object detection. Higher values will be treated as potential cells, lower values are set to 0.0 (as for all background pixels). The default is 34 dBZ, which - having the 14 discreet reflectivity levels in Fig. 3.1 in mind - comes closest to the 35 dBZ proposed in Feldmann et al. (2021).

3. R_{max} - upper threshold for object detection. Higher values will be set to this value, in order to avoid unrealistic splitting of intense cells with more than one reflectivity maximum into several adjoining smaller cells. The default is 50 dBZ, which is closest to the 51 dBZ in Feldmann et al. (2021).

4. ΔR - minimal reflectivity difference on the path of least change between two identified maxima, in order to split a cell into smaller ones (comparable to the concept of topographic prominence for mountains). The default is 6 dB, which is the same value as in Feldmann et al. (2021).

5. A_{min} - minimal area for object detection. Smaller cells will be discarded. The default is 5 pixels, corresponding to an area of approximately 20 km². This value is smaller than the 50 km² applied in Feldmann et al. (2021), because the focus there was not on cell initiation but on tracking larger, supercellular convection.

6. R_p - minimal value of peak reflectivity in detected objects. Cells reaching this intensity in neither of their pixels will be discarded. The default is 41 dBZ, which is the same value as in Feldmann et al. (2021).

7. d_{min} - minimal distance between two identified maxima. For maxima closer to each other (either between different identified objects or within the same object) the weaker ones are being discarded. The default is 5 pixels, corresponding to a distance of approximately 10 km, which is the same value as in Feldmann et al. (2021).

Other than that, a few technical variables dealing with information on previous runs as well as the output and saving method can be specified.

Cell detection - the D in DaTing

Cell detection is carried out for each time step independently. First, weak reflectivity signals lower than R_{min} (35 dBZ) are excluded and very strong ones are saturated at R_{max} (50 dBZ) to avoid excessive cell splitting. Then, a unique ID is given to each remaining connected area (or cell), followed by discarding areas smaller than the parameter A_{min} ($\approx$ 20 km^2) and/or weaker than R_p (41 dBZ), as these could indicate short-lived cells or measurement errors.

As a next step, all local maxima in the reflectivity field are identified based on a method also used for image manipulation (van der Walt et al., 2014). For application of the ΔR-parameter, during this process the paths of least change between all possible maxima are computed, allowing only direct connections between pixels and excluding diagonal ones. Maxima that fulfill the ΔR-criterion can still be discarded if their distance from each other is smaller than d_{min} ($\approx$ 10 km). Step by step the weakest of these maxima is discarded (and if it is the only one in a cell also the whole cell) until no close lying maxima remain. The intensity of a maximum is thereby defined as the average reflectivity value of the maximum and its surrounding 8 pixels.

At this stage of the algorithm, every maximum corresponds to exactly one final cell, but still more than one maximum can fall into the same connected area. To actually carry out the splitting of such areas and find the bordering lines between its different cells, an inverse watershed algorithm is applied (Beucher and Lantuejoul (1979) and van der Walt et al. (2014)). Maxima are used as the watershed seed points. After the splitting, the check for size (A_{min}) and intensity (R_p) of the cell needs to be repeated.

Finally, the output of this first algorithm part is created. This includes a domain-sized 2D-array showing the position of all detected cell areas labelled with unique cell IDs, and a table of the most important parameters for each cell. This table consists of: (i) the cell ID, (ii) the current timestamp, (iii) the x- and y-coordinates of all pixels in the cell, (iv) the x- and y-coordinates of the cell centroid, (v) the highest reflectivity within the cell, (vi) the contour lines of the cell (van der Walt et al., 2014) and (vii) the cell area in terms of pixels.

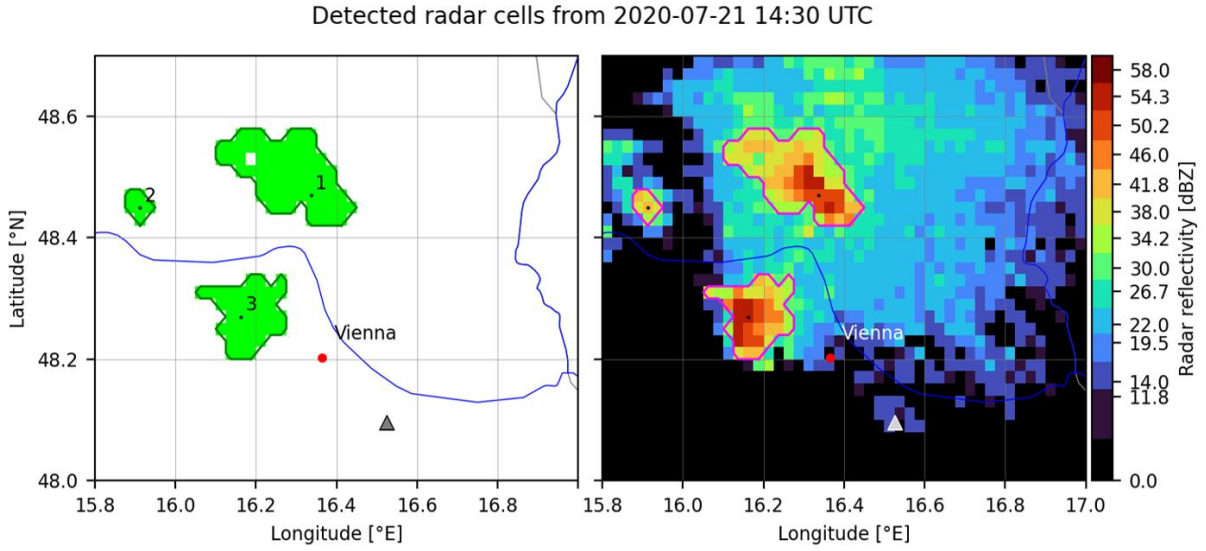

Figure 3.5.: Outcome of 3 cell detections of variable size and intensity in the Vienna region. Cell outlines in pink follow the 34 dBZ minimum reflectivity threshold (small weaker parts inside a cell are allowed, but not counted as cell area).

A carefully considered calculation of cell centroids is essential to avoid zigzagging of resulting tracks. Contrary to the simple averaging of pixel coordinates in the original algorithm, in this work a more advanced approach is followed. From the original unsaturated reflectivity field cell cores are computed, defined as all parts of a cell reaching a value of its peak reflectivity minus ΔR (6 dB). From these parts the largest one that includes a pixel of peak reflectivity is chosen as the main cell core, and from this main cell core the reflectivity weighted average is set to be the centroid. In rare cases (curved cell cores) this point can lie outside of the whole cell, which is corrected by choosing the nearest pixel inside the cell as centroid.

An example for the outcome of the detection part of T-DaTing is given in Fig. 3.5. Especially for cell 1 the benefit of the more sophisticated centroid calculation shows clearly (compared to a purely geometric approach as in the original algorithm).

Cell tracking - the T in DaTing

Cell tracking requires a series of at least 3 consecutive input frames, in order for cell motion to be calculated through an optical flow method included in the PySteps package (PySteps developers, 2018-2024). From the several available options the Lucas Kanade optical flow algorithm is chosen (Lucas and Kanade (1981), also used in Feldmann et al. (2021)). As the flow estimation needs 2 frames prior to the current one, the detected cells in the first 2 time steps cannot be included in the track analysis.

From the third time step onwards, the actual tracking is started by propagating identified cell objects. This is done by averaging the optical flow motion vectors of each pixel in the cell. Coordinates of pixels leaving the domain in the course of advection are set to the coordinates of the domain edges.

Next, advected cells are compared with the newly identified ones. The new cell with the biggest spatial overlap is chosen to be the successor, and its ID is changed to match the previous cell's ID. However, this only happens if the fraction of the overlap is greater than 40%. A little lower value is chosen as in Feldmann et al. (2021) (50%), as cells in their developing phase are smaller and can have more variability compared to mature supercells. If no match for an advected cell is found, the cell is supposed to have decayed and the associated track is ended. On the other hand, a new track is initiated, if no predecessor for a newly identified cell can be found.

It can be, that more than one possible match is found, which corresponds to a splitting or merging of cells. In case of a merging the ID of the largest involved cell is passed on and its track continued. For a splitting event, the larger cell after the splitting is chosen to be the successor, and for the other new tracks with new IDs are initiated.

Sometimes it also happens, that visibly obvious connections on the radar images fail the overlap criterion from above. To help reduce such missed connections, an extra combination procedure has been added to T-DaTing. The essential point is, if there is a track ending at one time step and a new track is starting in the next one, the overlap fraction is reduced to $> 0\%$, which means only one pixel must be common between the advected cell and the newly identified one. Again, if more than one possible match is found, the one with the largest overlap is processed further.

Until here, the results of the tracking are still organised as tables of cells sorted by time (one table for one time step and different track IDs). In a final step, the data rows are re-arranged into tables of tracks sorted by ID (one table for one track and different time steps), and tracks shorter than t_{min} (3 identifications) are discarded. An exception from this rule applies only at the beginning or the end of a monthly data file, where a track could continue in the previous or the next month. To cope with this issue, a new transition-function has been added to the algorithm.

Output and storing

The final output of T-DaTing involves 3 parts: (i) a list of label images and (ii) a list of cell properties as described in the subsection on detection, and (iii) a list of these cell properties reorganised in tables for individual tracks. Results are stored in 4 data files:

- A file of concatenated cell tables sorted by time (using the Python-module pickle).
- A file of concatenated track tables sorted by ID (using the Python-module pickle).
- A netCDF file containing all label images (dims: time, latitude and longitude).
- A netCDF file containing the most relevant track information (dims: number of track and track length). That includes x- and y- coordinates of the cell centroid, timestamp of occurrence, maximum reflectivity and area in terms of pixels.

3.3. Parameter adaptation for the use of lightning data

In the original form, T-DaTing was intended for the use of radar data, particularly for the MeteoSwiss radar network (PySteps developers, 2018-2024). But with the right choice of input data (smoothed flash density) and the right tuning of thresholds, the algorithm also works for detecting and tracking lightning cells. Choosing appropriate thresholds is a challenge though, particularly due to the fact that radar reflectivity is measured on a logarithmic scale, whereas flash density is a linear quantity. Below the final values are stated next to the aspects considered to make the choice objective and reproducible:

- $A_{min} = 5\ px$ and $d_{min} = 5\ px$: The minimal size and distance for a cell to be detected do not change depending on the dataset, as long as the same grid is being used.
- $t_{min} = 2$: As 3 identifications for the 5 min resolution of radar data correspond to a cell life time of at least 10 minutes, the same life time for the 10 min resolution of lightning data is reached after a minimum of 2 identifications.
- $FL_p = 0.240\ km^{-1}h^{-1}$: In part (d) of Fig. 3.4 it can be seen, that single flashes remain in the smoothed dataset as outliers. These outliers should not be identified as separate cells by T-DaTing, i.e. each cell must at least comprise 2 flashes. This can be ensured by setting FL_p a little higher than the peak density of a single flash. When computing this value its change across the domain needs to be considered, as not all boxes in a lat-lon-grid are of equal size. The smallest boxes are located at the northern edge of the domain, leading to the highest peak density value of $0.2398\ km^{-1}h^{-1}$, which is in a final step rounded up to the value given above.
- $FL_{max} = 3.1\ km^{-1}h^{-1}$: Here comparisons are made for a test month (July 2020). First, the number of radar pixels exceeding 50 dBZ (and therefore being subject to saturation) is calculated. This corresponds to approximately 0.27% of all radar pixels different from 0. Now, FL_{max} for the lightning dataset is chosen in a way that leads to the same percentage of saturated pixels.

- $FL_{min} = 0.136\ km^{-1}h^{-1}$: The choice of a suitable lower threshold for cell detection is more difficult. It is based on a contingency table approach for the mentioned test period and carried out in the part of the domain that encompasses both radar and lightning data (essentially Austria and its surroundings).

	$R > 40\ dBZ$	$R < 40\ dBZ$
$FL > FL_{min}$	A	B
$FL < FL_{min}$	C	D

Table 3.2.: Contingency table used to find the minimum flash density FL_{min}.

It is assumed that a connection between radar reflectivity and the onset of CG lightning exists, where a good predictor seems to be a 40 dBZ threshold (e.g. Yang and King (2010)). In the common domain all pixels are divided into classes, representing the 4 entries of the contingency table. Pixels with radar reflectivity exceeding 40 dBZ and lightning flash density exceeding a specific FL_{min} raise the number in entry A of the contingency table (Tab. 3.2). Entry B is raised when the 40 dBZ criterion is met, but flash density remains below FL_{min}. Likewise, entry C includes pixels with strong enough lightning signals and too weak radar signals. Finally, entry D contains pixels with no strong enough signals whatsoever. Although D is usually the highest number, it is not important. From the other 3 entries a bias measure can be computed with Eq. 3.1, that is greater than 1 for radar exceeding lightning pixels and between 0 and 1 for lightning exceeding radar pixels. FL_{min} is now varied in magnitude until a configuration is found that shows a bias as close to 1.0 as possible. This value is found to be $FL_{min} = 0.136\ km^{-1}h^{-1}$.

$$BIAS = \frac{A + B}{A + C} \tag{3.1}$$

One further annotation: the same result also shows up when strong enough radar and lightning pixels are just counted for each time frame and not individually compared, i.e. they must not be co-located. This is physically reasonable as some convective cells are not electrically active and others appear in areas of the domain that are prone to radar beam blockage.

- $\Delta FL = 0.079\ km^{-1}h^{-1}$: The physical implications of varying this parameter are not obvious at first sight. Therefore another numerical solution based on contingency tables is needed, hereby implemented for whole cell objects instead of individual pixels.

  Radar and lightning cell numbers are compared for each time frame. The analysis is reduced to radar cells with a reflectivity $R > 50\ dBZ$, as not all cells are electrically active (also compare the definitions of intense cells in Chap. 4.3). The smaller number is then raising entry A of the contingency table, the gap to the larger number raises entry B when there are more radar than lightning cells or entry C when there are more lightning than radar cells, and D is kept empty as it is not needed for the bias calculation following Eq. 3.1.

Having fixed all of these 7 parameters already allows T-DaTing runs with flash density as input variable. The algorithm code itself does not need to be changed.

3.4. Combination of radar and lightning data

Climatologies on deep moist convection (DMC) can vary considerably depending on the data source (see figures in Taszarek et al. (2019)). Though when being combined, climatologies tend to benefit from the use of multiple data sources (as for the ec-TRAM detection and tracking algorithm in Meyer et al. (2013)). After making T-DaTing available for the use of lightning data in Chap. 3.3, it should also be possible to process a combination of radar and lightning data. For this, the main part of the algorithm does not even need to be changed, but it has to be extended.

The basic concept is to group previously detected, individual radar and lightning cells together to form so-called mixed cells. These mixed cells can then be tracked in the same manner as for the separate datasets. The grouping follows the work described in Meyer et al. (2013) and illustrated in Fig. 1.3. It is based on choosing the outermost boundary of overlapping cells as the new contour for a mixed cell.

The implementation into T-DaTing involves the following steps: at first, lightning and radar cells have to be detected separately. This can be done in-place or by loading the output files of previous runs. Then, for each cell the corresponding cells of the other data type are searched, having an overlap of at least 1 pixel. Hereby, even more than 1 match is allowed, as a lightning cell could for example be the link between two close lying radar cells. In an iterative procedure, all cells connected via a continuous train of such links are then combined, resulting in one mixed cell having one unique mixed ID. At this point, 4 types of mixed cells can be found, consisting of: (i) exactly 1 radar and 1 lightning cell (a single thunderstorm cell), (ii) a mixture of radar and lightning cells (multicell clusters), (iii) 1 single radar cell (rain showers with no electrical activity) and (iv) 1 single lightning cell (thunderstorm in an area with blocked radar view). Especially the latter could offer potential for improvement of DMC climatologies.

To be able to leave the tracking part of T-DaTing unchanged, similar input needs to be created for mixed cells, i.e. a labels image and a table of detected cell properties. The labels image is filled with the newly assigned mixed cell IDs in their corresponding areas, taking care that this is done only in parts of the domain common for radar and lightning data. Also, a final check of the size and intensity criteria for cells is carried out, where each mixed cell needs to comprise at least $A_{min} = 5\ px$, of which at least 1 pixel must exceed $R_p = 41\ dBZ$ or $FL_p = 0.240\ km^{-1}h^{-1}$. The table of cell properties is similar to the one described in Chap. 3.2 and consists of: (i) the cell's ID, (ii) the timestamp, the coordinates of (iii) all pixels and (iv) the centroid, the maximum (v) reflectivity and (vi) flash density, (vii) the contour lines and (vii) the area. The centroid is a weighted average of all cell pixels. For mixed cells made up of a combination of radar and lightning cells, the weights of both datasets are multiplied by one half. For mixed cells of only one data type, the weights are calculated based on this data type only. Again, for curved cells with a centroid outside of the cell area, the nearest pixel in the cell is assigned as centroid.

The remaining part of the tracking procedure for mixed cells follows exactly the same steps as described for radar data in Chap. 3.2. This concludes the remarks on the methods of this thesis. Their implementation in practice can be summarized as follows:

1. Make radar and lightning data comparable (interpolation vs. smoothing)
2. Customise the tracking algorithm T-DaTing according to the needs of this thesis
3. Conduct a T-DaTing run for a 12 year period of radar data (2009-2020)
4. Adapt parameters of T-DaTing towards the use of lightning data

5. Conduct a T-DaTing run for a 16 year period of lightning data (2005-2020)
6. Add a T-DaTing module for mixed cells comprising either or both of the datasets
7. Conduct a T-DaTing run for a 12 year period of mixed data (2009-2020)
8. Stratify and interpret the results in Chap. 4

Carrying out these steps results in a large number of data files, one for each of the 3 data sources, each month of the 12 or 16 years of data, and each of the 4 output methods (cell tables, track tables, label images and most relevant track information). From this big dataset all of the figures and statistics presented in Chap. 4 have been generated.

4. Results, Analyses and Interpretation

On the following pages some of the most interesting aspects hidden in the track data of radar, lightning and mixed cells will be presented. This involves applying different stratification methods (e.g. based on cell type, synoptic flow, time of year and day, cell lifetime, maximum cell size, maximum cell intensity or cell moving distance).

At first, in Chap. 4.1 examples of the detection and tracking output are shown to prove the overall operability of T-DaTing in its adapted form. Then a detailed chapter on cell initiation follows (Chap. 4.2), including a focus on the relation to synoptic flow and the temporal evolution of DMC during a day and year. In Chap. 4.3 analyses are carried out for cells having a high impact on people and the environment (very intense, long-lasting or far-moving cells). And finally, Chap. 4.4 provides images and statistics on how different cell properties vary, and if and how they are related to each other and to the synoptic flow.

4.1. A first look into the tracking data

To get familiar with the data, it is a good start to look into the calculation results of one single day. After that the output of one month can be better understood, before the whole data set of 12 or 16 years respectively is analysed.

Cell tracks of one day in July 2020

From Fig. 3.5 it can already be seen that T-DaTing is able to *detect* features in a reflectivity field, similar to what human inspection would define as precipitating convective cells. Still, no example of the algorithms tracking capability has been shown so far. In Fig. 4.1 now, all tracks of a day in July 2020 are plotted, colored according to the time of their initiation, with the initiation location marked as a black dot. Short tracks stemming from short-living single cells are present as well as long tracks of apparently organised storms, especially for cells originating in Southern Bavaria. The DMC evolution on this day starts near the Sarntal Alps tough, a region that is known for being a center of thunderstorm activity within the Alps. In the course of the day some cells to the northeast appear, followed by cells over the Alpine forelands. During the late evening activity concentrates on the eastern edge of the Alps with cells moving towards the adjoining plains.

At first sight it seems that there are no big differences between the tracks of the radar cells (a) and mixed cells (b). And this dominating effect of radar data on mixed cells will be seen in many other figures. When looking more closely, though, some details seem interesting. At first, the number of mixed cells is slightly going down in comparison to radar cells, apparent from the figure titles and the histogram panels based on four different periods of a day - late night, before noon, after noon and early night. This can be attributed to the combination of overlapping cells during the mixed cell calculation. The biggest difference is cell tracks originating near the southwestern boundary of the domain, over parts of Switzerland and Italy. These cells stem from lightning activity alone as radar visibility is blocked by mountains in this area. Lastly, the tracks of mixed cells tend to be a little smoother than the ones of radar cells, which could be due to the different centroid calculations of the two cell types.

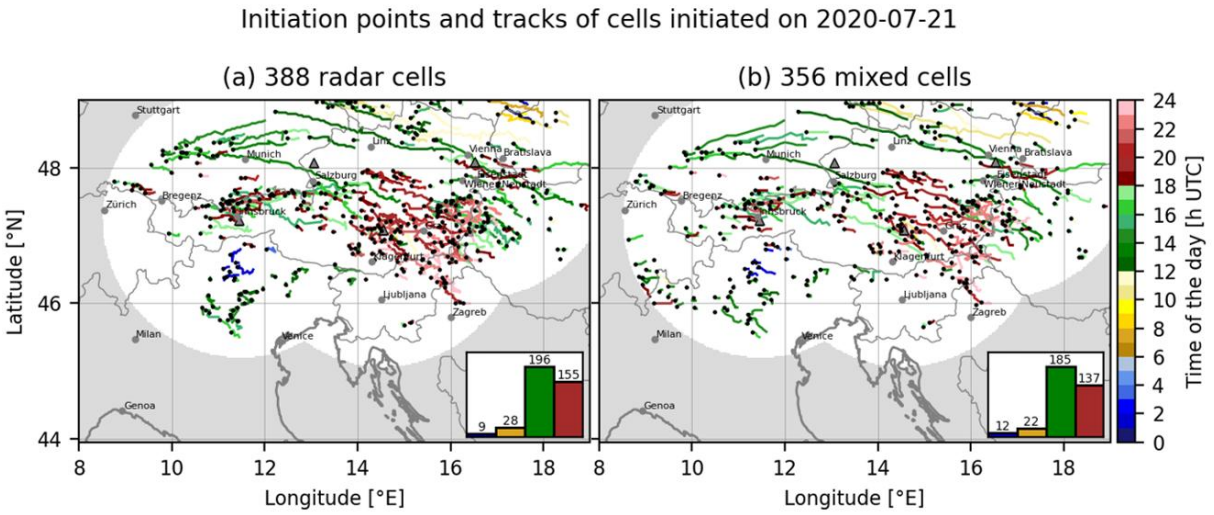

Figure 4.1.: Example of radar and mixed cell tracks for one specific day (21 July 2020). Colors according to time of day of the cell's initiation. Tracks numbers for 6h-periods are given as a histogram in the bottom right corner of each image.

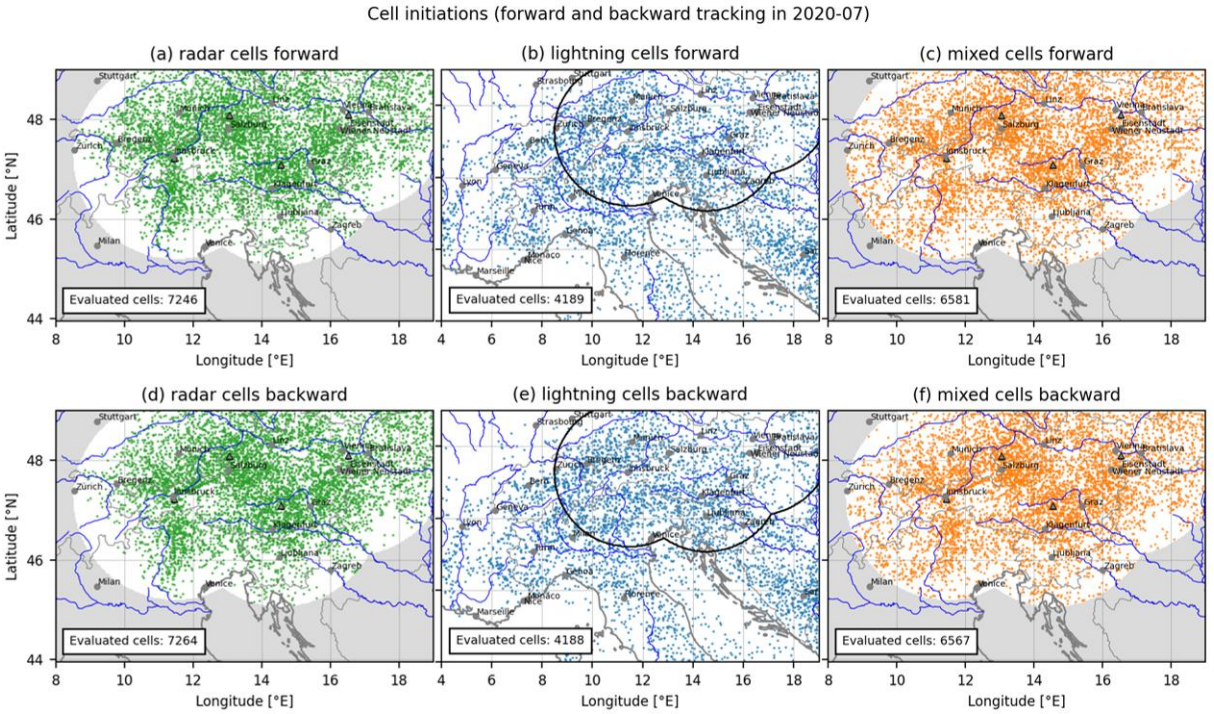

Figure 4.2.: Scattered cell initiation points showing no significant change of tracking radar, lightning and mixed cells forward and backward in time. Data for July 2020.

Handling of cell splits and mergings

The previously mentioned dominant radar influence on mixed cells is not just characteristic of a specific day. It is also visible in the data of the whole month (Fig. 4.2) and it can partly be explained by the lower number of lightning cells ($\approx$ 4200) in comparison to radar cells ($\approx$ 7200), despite being distributed over a much larger domain.

In this figure the initiation locations for all 3 data types are shown, based on 2 different calculation strategies: (i) providing T-DaTing with a sequence of images forward in time (CI is equivalent to the first track point, panels a, b and c), and (ii) a sequence backward in time (CI at last track point, panels d, e and f). The reason behind doing this is to get a feeling of whether it matters how T-DaTing is dealing with cell splittings and mergings, one of the basic steps of tracking algorithms as described in Chap. 1.3. Reversing time is equivalent to turning splittings into mergings and vice versa. Therefore, if the results of both calculation strategies do not differ substantially, cell splittings and mergings play only a minor role in creating the dominant patterns of CI. This is the case for the data in Fig. 4.2. The numbers of detected tracks vary only marginally on the order of 0.1%.

4.2. Preferred areas of CI in the Alps

After getting familiar with the output of T-DaTing, in this chapter different CI climatologies will be presented and stratified by aspects influencing the occurrence of DMC, such as synoptic flow type or time of year and day.

4.2.1. Overall climatologies for the 3 data types

Before stratifying the data, *all* detected CI events for the 3 data types are plotted as density maps in Fig. 4.3. Here it needs to be noted, that figures of this type come with a different resolution depending on the size of the input sample in order to bring out the main features of the data. For the maps in Fig. 4.3 a resolution of 2x2 grid boxes was chosen (which results in an area of approximately 4.2 x 4.2 km^2). The respective number of grid boxes and the size of a pixel in this and all upcoming figures of this type are given in the figure caption.

Panel (a) for radar data suggests maxima of precipitating convective cell initiation north of the Alps (especially near the city of Salzburg), in a streak reaching from the Carinthian Nock Mountains across the edge of the Alps in Styria to the Vienna Basin near Wiener Neustadt, at the mountain slopes of Friuli-Venezia Giulia in Italy and next to it in Slovenia, and finally above the previously mentioned Sarntal Alps in South Tyrol.

Some of the visible patterns cannot be explained by natural causes, but can be attributed to the measurement setup. For example, areas of reduced radar beam coverage are seen to the southeast of the radar at Zirbitzkogel (denoted by the grey triangle next to Klagenfurt). However, the largest regions of limited radar sight are located in the western parts of Austria, particularly in Vorarlberg, the neighbouring provinces of North Tyrol and the triangular area between Innsbruck and Salzburg. From this figure it also arises, that the Patscherkofel radar has best viewing capabilities in the direction of the main valleys, namely the Inn valley to the west and east and the Wipp valley to the south, providing direct view over the Brenner pass onto the Sarntal Alps. Altogether, the pattern of radar cell CI resembles quite precisely the one in Fig. 1.2 showing radar-derived precipitation sums of convective seasons.

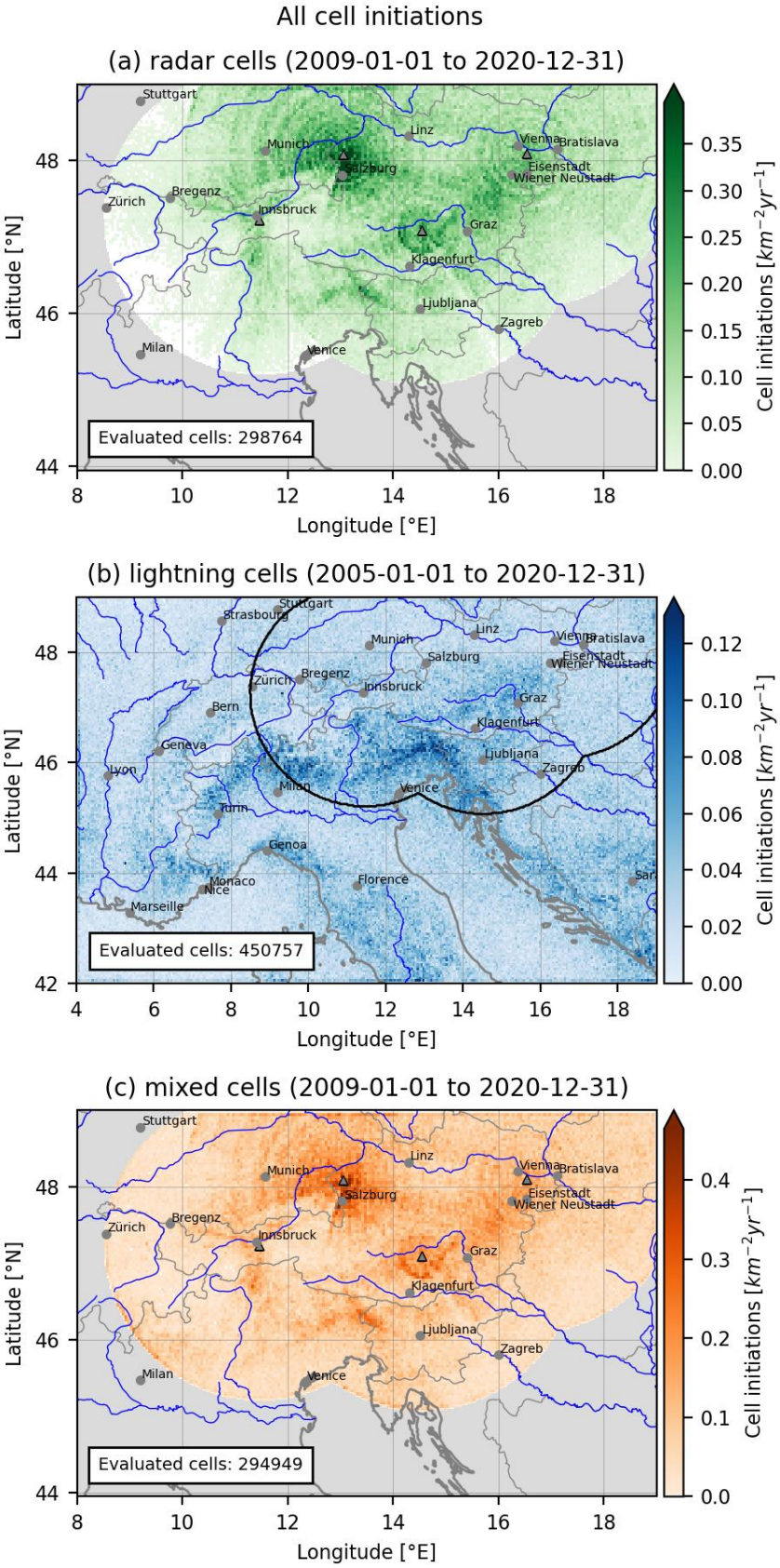

Figure 4.3.: Initiation climatologies for radar, lightning and mixed cells (no stratification, 2x2 grid boxes combined, 17.3 km² average pixel size).

But it is not only the blocked areas that are apparent. Also the circular shapes around the individual radar locations are not natural variations, but stem from the different elevations of the radar scan strategy. Their existence also indicates that cells tend to artificially gain and lose strength on radar images when passing through certain areas. For cells in a specific intensity range (around $R_p = 41\ dBZ$, the minimal value of peak reflectivity) this leads to them being detected multiple times in a broken manner. In turn, this could lead to an overestimation of CI based on radar data in areas that are prone to such circular patterns. This overestimation near radar sites together with the underestimation in areas subject to beam blockage further distorts the real CI climatologies.

Panel (b) of Fig. 4.3 on lightning-based CI further supports this hypothesis, as the maxima of radar-based CI around Salzburg and Vienna are not visible there. Contrary to that the maximum near the transition zones between the Alps and plains in Northern Italy is even much more pronounced. Generally jumps in terrain height seem to be preferred areas of CI (to a minor degree also visible at the northern and western edges of the Alps, as well as for other mountain ranges such as the Apennines in Italy and the coastal areas of Croatia and Bosnia-Herzegovina). On the other hand, there is a pronounced CI minimum in the mid part of the Western Alps, where mountains typically are higher and synoptic flow tends to be blocked more often (compare Fig. 1.1b). Also it can be mentioned, that this lightning CI climatology created with the use of a tracking algorithm resembles closely Fig. 8 and Fig. 9 from Manzato et al. (2022), where a clustering of CG flashes was carried out to define CI events. Even small-scale CI hotspots over certain mountains such as Säntis in Switzerland and Dobratsch in Austria are visible in both studies. But the figures in Manzato et al. (2022) not only show a similar pattern, also the CI density values are similar when being converted to the same unit (peak values in Fig. 8 and Fig. 9 correspond to $0.04 - 0.17\ km^{-2}yr^{-1}$).

Another advantage of lightning data is visible in flat parts of the domain where lightning CI events are distributed uniformly, contrary to the radar CI events where the distance from radar locations plays an important role. This is the immediate effect of the relatively homogeneous detection efficiency of lightning sensors. These advantages of lightning data come at a cost, that is the lower magnitude of cell initiations per grid box (the colorbar ranges indicate that there are approximately 3 times more radar and mixed cell initiations than lightning cell initiations). To sum up, the potential of lightning data would be great to fill the gaps in the radar image, if it was not for the lower number of cells.

Panel (c) showing mixed cell CI events provides further evidence, that the number of lightning CI events is just too small to significantly alter the pattern created by the radar data. Only in areas out of sight for the radar beams the climatology visually improves, especially for cells entering the domain north of Milan. These entering cells also lead to a thin stripe of increased CI density at the domain borders, which clearly is an artefact, but it already hints at the dominant direction of propagation of cells - coming from the west-southwest and moving towards the east-northeast. More on this is going to follow in Chap. 4.2.2 on synoptic flow, but first some aspects of the mixed CI pattern should be elaborated further.

A deeper look into mixed cells

Due to the design of mixed cells, either comprising both data sources or just one, it is possible to split the mixed CI climatology into 3 distinct parts (Fig. 4.4): (a) cells that only exceed the peak reflectivity criterion of radar data $R_p = 41\ dBZ$, but not the one on flash density $FL_p = 0.240\ km^{-1}h^{-1}$, (b) vice versa and (c) cells that fulfill both criteria. These 3 parts make up approximately 56%, 11% and 33% of all mixed cells respectively, again showing the big

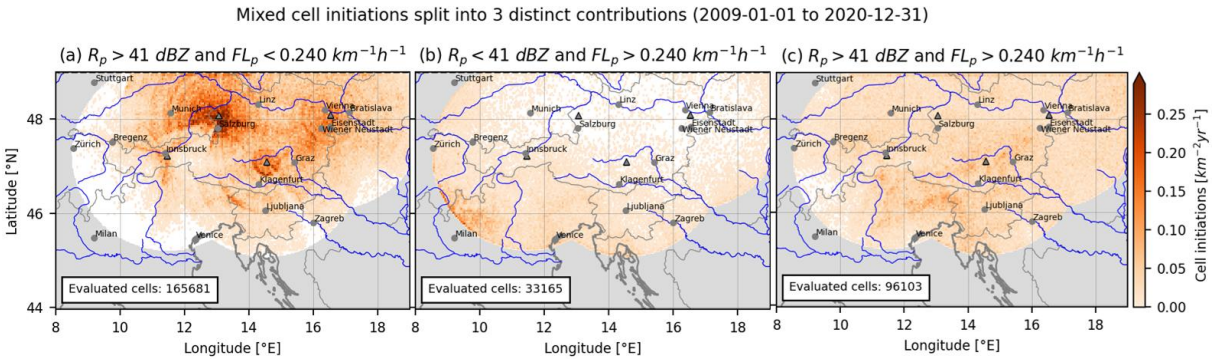

Figure 4.4.: Initiation climatologies for mixed cells (split into cells with strong radar signals only (a), strong lightning signals only (b) and both signals (c), 2x2 grid boxes combined, 17.3 km² average pixel size).

influence of radar data on mixed cells. It should be kept in mind though, that cells sorted into category (a) can still include lightning signals weaker than the FL_p threshold. Likewise, cells in category (b) can produce precipitation with a reflectivity below 41 dBZ.

From panel (a) it is immediately clear that the strong maximum around Salzburg entirely stems from cells that were stored due to their sufficiently strong radar signal, i.e. showers of rain without much electrical activity or even errors of the evaluation method (multiple detections of the same cell) based on inadequate measurements and different radar beam elevations.

Panel (b) however, can serve as a good first proxy on areas that are subject to beam blockage by mountains, with pronounced maxima near the domain borders, but also some intra-Alpine areas. This makes it somewhat of a complementary image to panel (a).

Finally, panel (c) shows a much more homogeneous CI climatology coming close to the one derived from lightning data in Fig. 4.3. In fact, the sum of (b) and (c) would be closest to it, as it would contain all mixed cells having strong enough lightning signals to be stored in the T-DaTing procedure.

4.2.2. CI and its relation to synoptic flow

Until here all analyses were carried out for the entirety of cells of the respective data type. Now, a first stratification of the data is sought, based on weather types represented by the synoptic flow in 850 and 700 hPa. As mixed cells are substantially influenced by the radar contribution, the following results are given for radar and lightning cells only.

Each of the cells is assigned to a flow type based on the day of its initiation and the dominant flow calculated at noon (compare Chap. 3.1). The frequency of flow types for convective cells (not for days per year!) is shown in Fig. 4.5. Whereas hardly any CI events in relation to easterly flow occur, they are much more characteristic for westerly and southwesterly flows. Apart from the purely directional classes, many CI events are also assigned to cyclonic and anticyclonic flow (presence of high or low pressure systems near the Alps) as well as low flow cases. At the 850 hPa level low flow cases are more than twice as frequent compared to the 700 hPa level. This could hint at mesoscale mountain winds (e.g. Alpine pumping) playing a greater role at lower levels than the synoptic flow. To prevent the sample sizes of each class from varying too much, the 700 hPa level is chosen as steering level in the following figures.

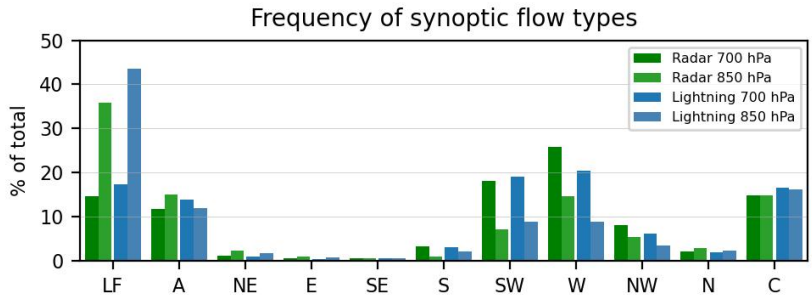

Figure 4.5.: Radar and lightning cells are very unevenly distributed across synoptic flow types on the 700 and 850 hPa steering levels (including a low flow (LF), anticyclonic (A) and cyclonic case (C)).

CI climatologies for lightning cells and 9 synoptic flow classes are depicted in Fig. 4.6. The classes NE, E and SE are combined in order to increase the sample size to 7752 cells, which is still a little less than the 8284 cells for the next highest class N. S and NW are also rather sparsely populated, whereas cell numbers for A, C, LF, SW and W all range between 60000 and more than 90000 CI events. Fig. 4.7 shows another aspect of the same data, i.e. a dimensionless ratio between the CI frequency in a synoptic class and the global CI frequency. Doing so, a relative enhancement of CI in a specific class appears to have values greater than 1 (bright colors), whereas a suppression leads to values between 0 and 1 (dark colors). In comparison to showing absolute numbers, this relative approach reveals dominant patterns also for the sparsely populated classes.

For the low flow case in panels (a) CI is predominantly attached to mountain areas, mostly the edges of high mountain ranges as the Alps. But maxima are also visible for the Balkans, the Apennines and smaller areas of elevated terrain such as Corsica. Especially in the relative image maxima along the Jura mountains at the French-Suisse border, the Vosges near Strasbourg, the Black Forest and the Swabian mountains in Germany appear and to a minor extent also the elevated parts of Mühl- and Waldviertel in Northern Austria. One reason behind this pattern could be the favorable overlap of slope winds and plain-to-mountain winds.

In the anticyclonic setting (b) CI maxima are also located in mountainous regions but shifted towards their centers. The underlying physical process is again mountain-induced wind systems that inject moister boundary layer air to higher levels and therefore above its LFC. Far from the mountains a source of lift is often lacking or even counteracted by large-scale subsidence. That is why there are only a few lightning CI events over the plains in Hungary or Northern Italy. However, the relative image reveals an otherwise unnoticed and unexpected enhancement of CI to the northwest of the domain.

The almost complementary pattern is visible in panels (c) for cyclonic flow. In this case the main maxima occur near the coastlines to the northeast of the Adriatic and Tyrrhenian sea, which has several reasons. Firstly, this flow type is closely related to cyclogenesis in Northern Italy after cold front passages from the west. At their southern edges these cyclones often lead to winds that can take up moisture over the sea before moving over land. In parts of Italy and Croatia these winds additionally encounter lift due to terrain jumps, favoring the onset of DMC.

The most common case of westerly flow (g) shows one strong maximum stretching from parts of Italy north of the river Po across Slovenia to Istria in Croatia. To interpret this (and any other) feature in Figs. 4.6 and 4.7, frictional veering needs to be accounted for, i.e. the change of wind direction with height. For instance, in case of westerly winds at 700 hPa, near-surface

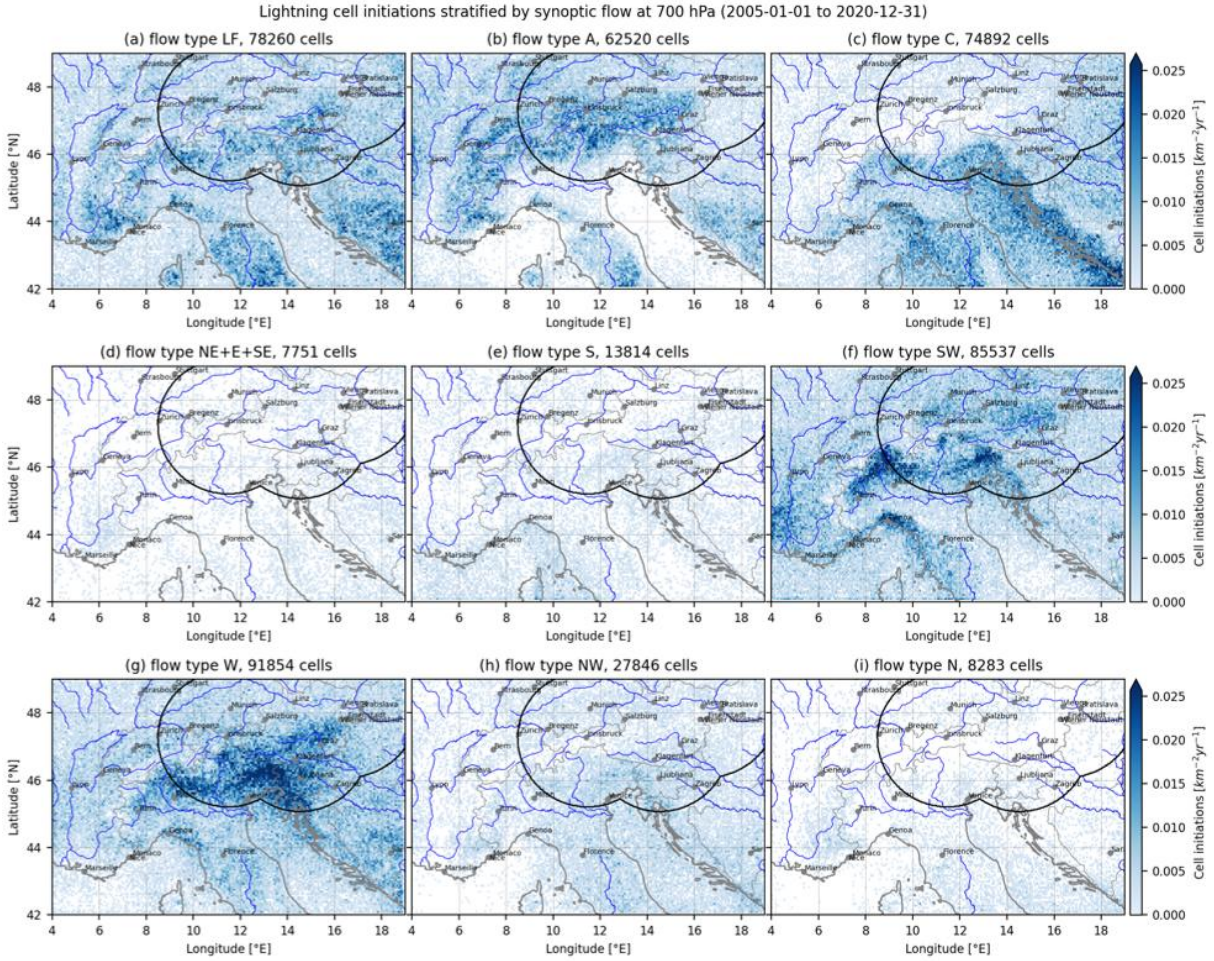

Figure 4.6.: Initiation climatologies for lightning cells (stratified by synoptic flow type, 3x3 grid boxes combined, 39.0 km^2 average pixel size).

winds tend to have a southwesterly component. For the main maximum in (g) this translates to low-level advection towards the mountains. A minor second maximum is visible in the southeast of Austria. This region is known for thunderstorms originating when cold air moves across the Alps and over the moister, warmer air on the lee side, therefore creating instability. Surprisingly, on the upwind side of the Alps the number of CI events for this flow type is rather low, only a minor enhancement signal is visible there (even tough the mountains provide a source of lift to impinging winds).

The SW pattern in (e) includes both types of maxima. Those on the windward side of mountains appear e.g. over Northern Italy, Tessin or bending around the Alps from the Rhone valley over the French Riviera to the Gulf of Genoa (the latter best visible in the relative image). On the other hand, maxima in the lee of the Alps can be detected over Southern Germany and parts of Austria in both figures. Especially the maximum between Bregenz, Innsbruck and Munich is known as an initiation region for strong thunderstorms moving through the northern forelands of the Alps, not rarely showing supercell features. A reason for the differences of westerly and southwesterly flow could be the predominance of warm and moist air coming from the Mediterranean (in the SW case) versus colder air from the Atlantic (in the W case).

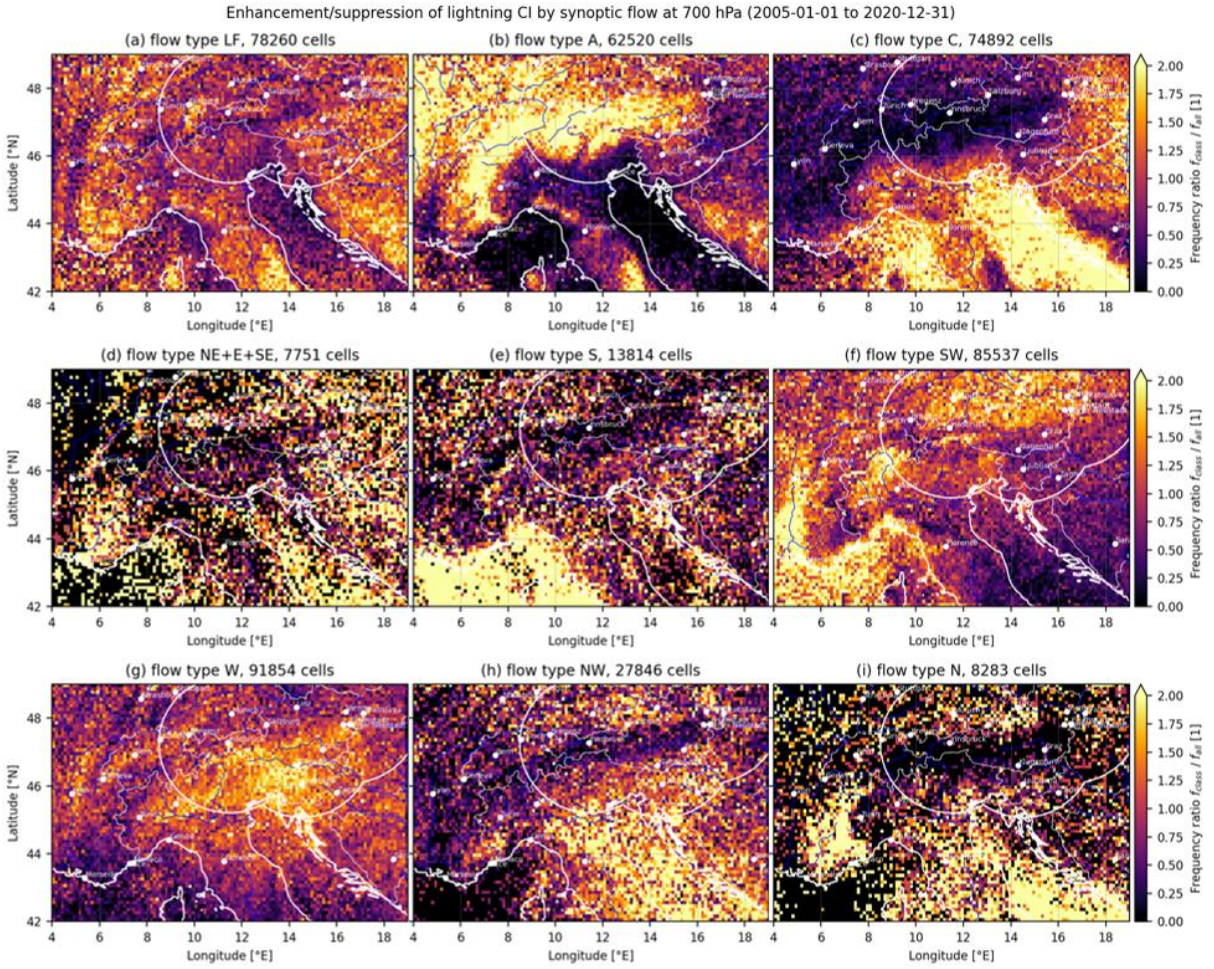

Figure 4.7.: Same data as in Fig. 4.6, but as a frequency ratio to highlight the relative enhancement or suppression of CI for each synoptic flow type (5x5 grid boxes combined, 108 km^2 average pixel size).

Finally, northwesterly flow (h) tends to favor DMC above and around the Adriatic Sea, possibly due to colder air moving across warmer surfaces again. All other cases (N, S, NE+E+SE) are too sparsely populated to get robust information out of the absolute numbers of CI events. In the relative image, however, a clear enhancement of CI is visible over the Tyrrhenian Sea for southerly flow, but over the Adriatic Sea for northerly flow. A second pronounced enhancement region for the N case appears over the Provence in France. For the combined easterly case there is still no clear signal in the relative image, but a lot of noise.

A last feature to mention is that the previously introduced Sarntal Alps serve as a confined, but powerful CI hotspot for westerly, southwesterly and anticyclonic winds.

The CI patterns of radar cells in Fig. 4.8 are somewhat similar to the ones of lightning cells. The LF case shows maxima at the Alpine edges, A towards the interior of and C around the Alps. Also these 3 plus the W and SW case are the most densely populated ones, whereas N, S and NE+E+SE show only a small number of CI events. Nevertheless, some differences exist, the most striking one being the maximum around Salzburg that occurs predominantly in the cyclonic, southwesterly and westerly flow types. As stated previously this is most likely a

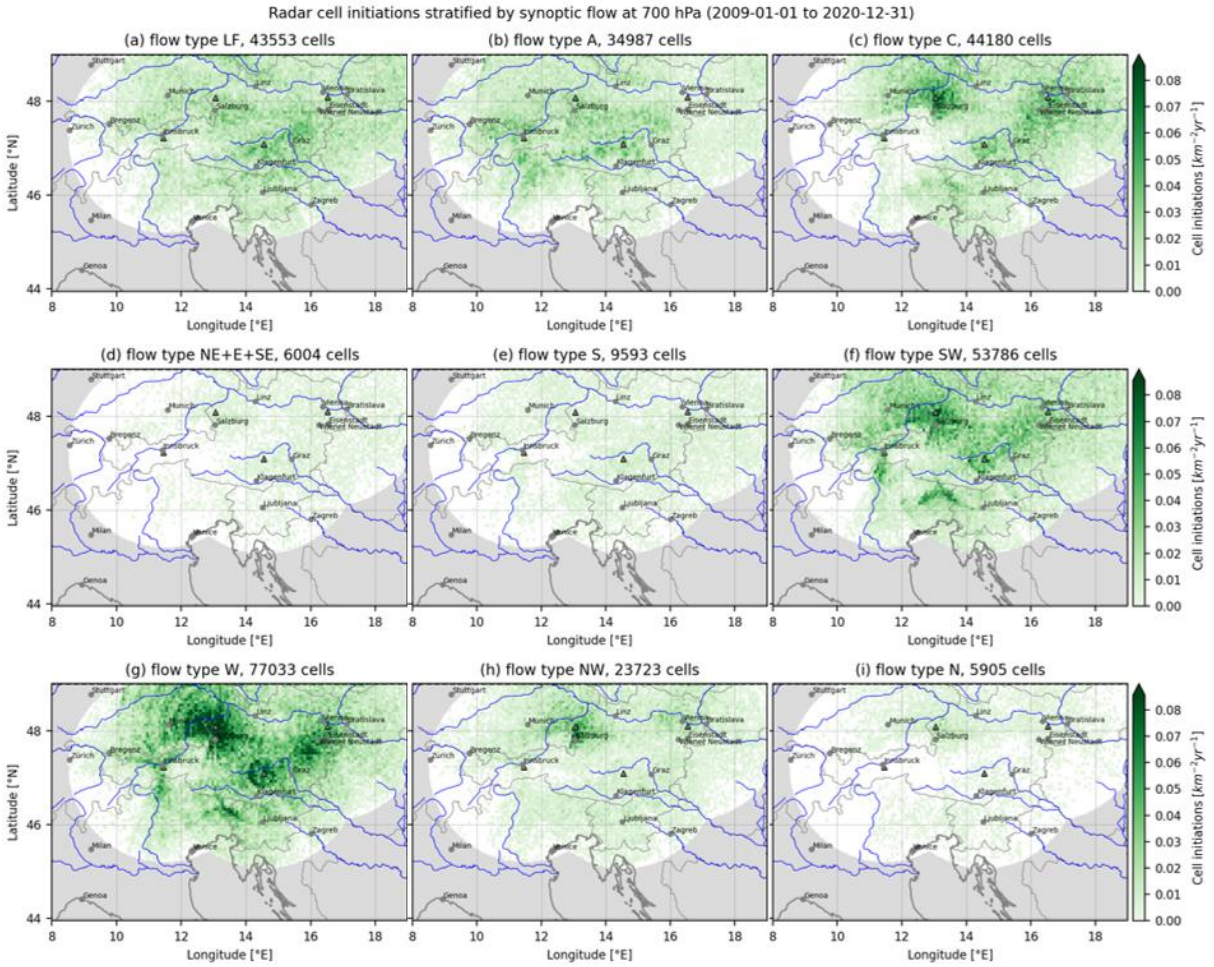

Figure 4.8.: Initiation climatologies for radar cells (stratified by synoptic flow type, 3x3 grid boxes combined, 39.0 km^2 average pixel size).

combination of really existing showers of rain without electrical activity and shortcomings of the measurement and data evaluation strategies (these systematic errors also being the main reason for not providing an enhancement/suppression analysis of radar cells). A similar, but not so strong signal is also visible for the NW case, possibly representing showers forming in the cooler air behind cold fronts. One last thing to mention for the W case is the maximum over the so-called Semmering-Wechsel-region between Wiener Neustadt and Graz. This region often experiences the first cell initiations in the course of a day, with cells moving across parts of Burgenland and into Hungary later on.

4.2.3. CI varying with time of year and day

Another way of stratifying the climatologies is using a time measure, as DMC experiences great yearly and daily variations. The frequencies of CI events in the course of a year and a day are presented in Fig. 4.9. Panel (a) shows weak convective activity in the beginning of a year until April, when DMC starts to develop more frequently. Cell numbers then amplify until they reach their peak in July. The following decline is more rapid for radar than for lightning cells, which is due to fact that DMC shifts towards the Mediterranean in the course of autumn and this region

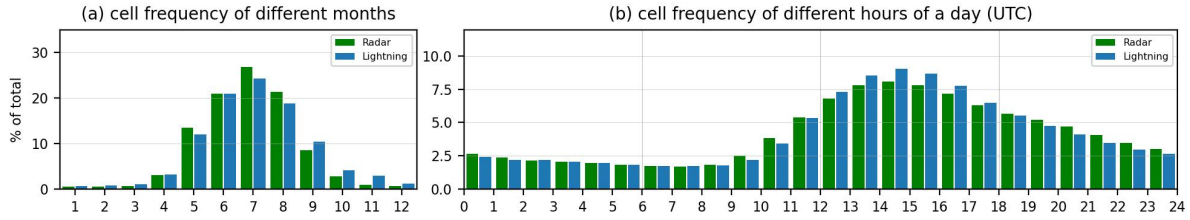

Figure 4.9.: Radar and lightning activity follows pronounced seasonal and daily cycles.

is only covered by the lightning domain of this work. During a typical day (in a climatological sense) the CI minimum occurs in the morning hours, followed by a steep incline over noon until the maximum for both datasets is reached between 14 and 15 UTC (corresponding to 16 and 17 CEST). Then activity is starting to go down again more and more slowly until it reaches the next minimum in the following morning.

Similar to the analyses on synoptic flow types, in Figs. 4.10 and 4.11 absolute and relative lightning CI climatologies for 6 months (April to September) in the warm season are presented. Panels (a) show the spring onset of DMC predominantly over land, as land is warming quicker than large water bodies (due to their different heat capacities). Interestingly, a prominent minimum can be seen right over the Alpine main ridge, which hints at anticyclonic flow not playing a major role during that time of the year - at least the air is not warm or moist enough for mountain-induced flows to lift the air to its LFC.

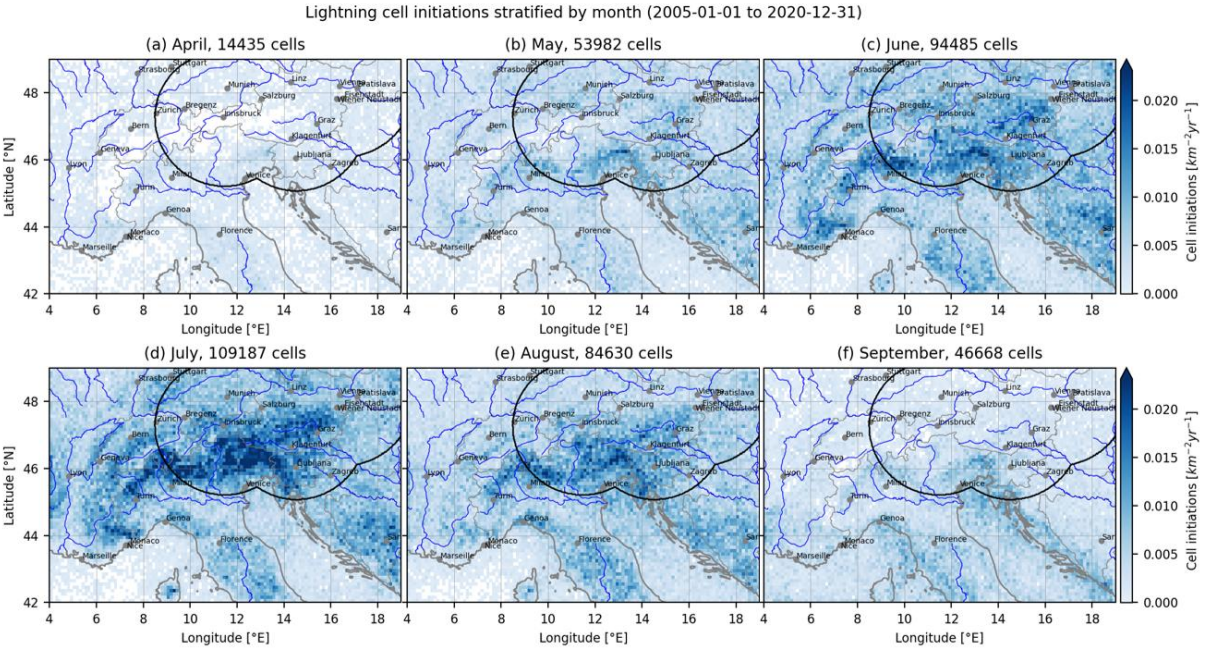

Figure 4.10.: Initiation climatologies for lightning cells during the warm season (stratified by month, 5x5 grid boxes combined, 108 km² average pixel size).

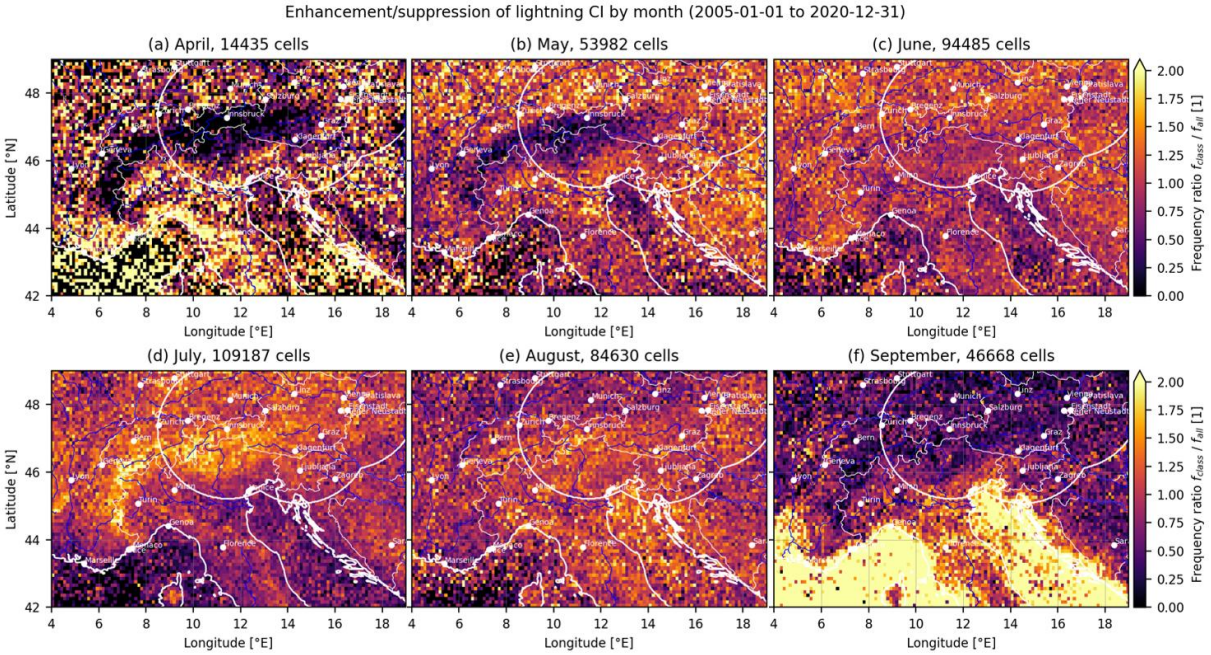

Figure 4.11.: Same data as in Fig. 4.10, but as a frequency ratio to highlight the relative enhancement or suppression of CI for each month (5x5 grid boxes combined, 108 km² average pixel size).

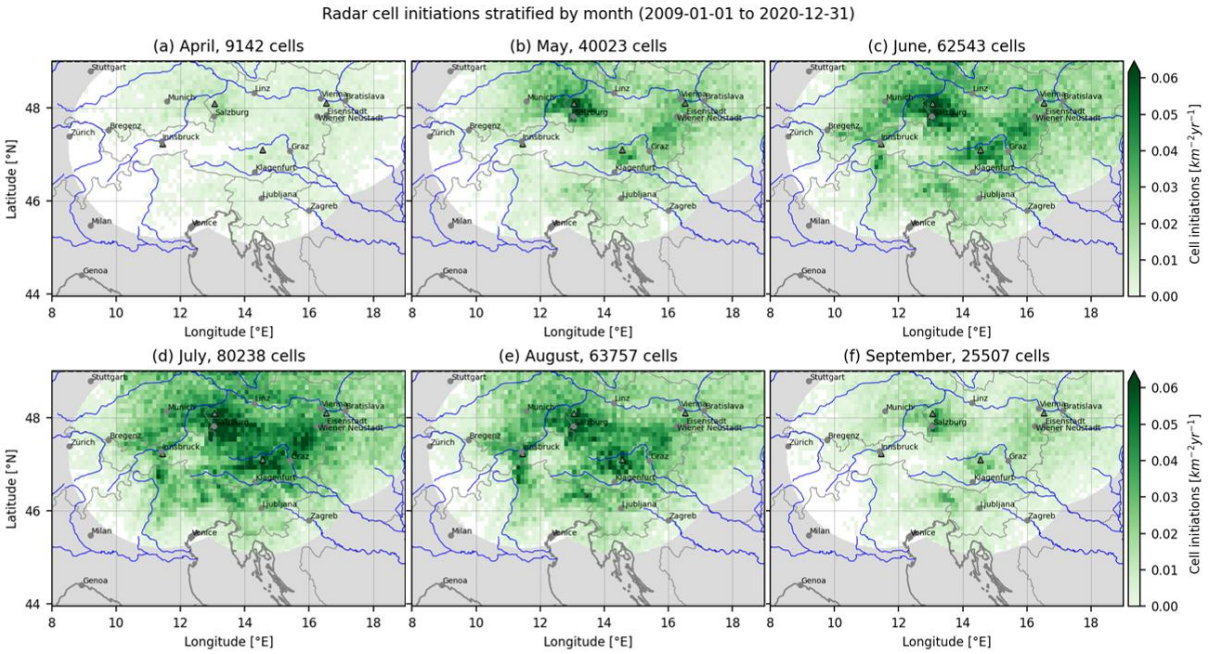

Figure 4.12.: Initiation climatologies for radar cells during the warm season (stratified by month, 5x5 grid boxes combined, 108 km² average pixel size).

For absolute cell numbers this minimum also occurs in the other 5 months - sometimes more, sometimes less pronounced. Apart from that the patterns for May (b), June (c), July (d) and August (e) do not vary extensively, only the magnitude does following the yearly cycle in Fig. 4.9a, culminating in a peak in July. The main maximum in all months extends along the southern edge of the Alps, mainly covering areas in Northern Italy but also parts of Switzerland, Austria, Slovenia and Istria in Croatia. Within Austria the maximum lies directly to the south of the main Alpine ridge, again hinting at the onset of DMC by cold air from the northwest moving over warmer moist air in the southeast. Minor maxima are seen not only for the northern edge of the Alps, but also for all other mountain ranges in the domain, which suggests that orography generally is the main driver for deep convection in the area. However, the relative importance of the Alps varies, as can be seen in Fig. 4.11. Until July the Alps continuously gain importance, before the Mediterranean region becomes the main driver of CI in September.

This shift towards the Mediterranean is also visible for absolute cell numbers and can be explained by warm water still providing enough energy and moisture for near-surface air to become conditionally unstable. The dominant maxima at the northwest-southeast oriented coastlines can be explained be the prevailing winds and resemble the cyclonic case in Fig. 4.6. This hints at the first significant cyclogenesis events over Northern Italy in the commencing autumn. The remaining months during the cold season are not depicted here, but they all show a similar pattern as September with a gradual decrease of intensity until a minimum in January and February is reached. By then, the sea has also cooled down significantly and is not serving as a major DMC driver anymore.

For radar data, areas covered by the sea only make up a negligible fraction of the entire domain. This leads to an even more consistent pattern of CI between different warm season months as can be seen in Fig. 4.12. Maxima on the one hand occur in physically undisputed CI hotspots (such as parts of Styria, the Semmering-Wechsel-region, the transition zone towards the Friuli-Venezia Giulian plains or the border area between Tyrol and Bavaria). On the other hand some areas favoring CI may be missed due to beam blockage (towards Switzerland) or overestimated due to the applied measurement and data evaluation strategies (as explained in Chap. 4.2.1 for the maximum near Salzburg). A last interesting thing to notice is, that the Sarntal Alps maximum appears to be more prominent in June, July and August than in the other months. This statement holds even generally for all features related to the Patscherkofel radar. The physical reason behind this so far remains unclear.

Coming from analysing DMC variations throughout a year to variations during a day, the results for lightning CI events are shown in Figs. 4.13 and 4.14. The primary time for CI is the afternoon and early evening depicted in panel (c), which is in accordance with the statements in the beginning of this chapter. Moreover, the dominating already well-known maxima stick to the orography. Still, a view features are worth being noted.

Firstly, the late night case in panels (a) resembles the September and winter cases in the previous analysis for a whole year. This is due to the same underlying process, warmer water heating cooler air aloft until conditional instability is released. It does not matter whether the temperature differences in that process stem from the course of the seasons or the circadian rhythm. In panels (b) for the late morning and noon cases regions with an unusually early onset of DMC during the day can be identified. Maxima of absolute cell numbers stretch across the Istrian peninsula in Croatia and Southwestern Slovenia or the parts of the Balkans. Moreover, the already mentioned early initiation of cells in the Semmering-Wechsel-region in Eastern Austria is to a minor degree visible. In the relative image also the maximum in Mühl- and Waldviertel can be seen, and the overall importance of the Mediterranean is highlighted.

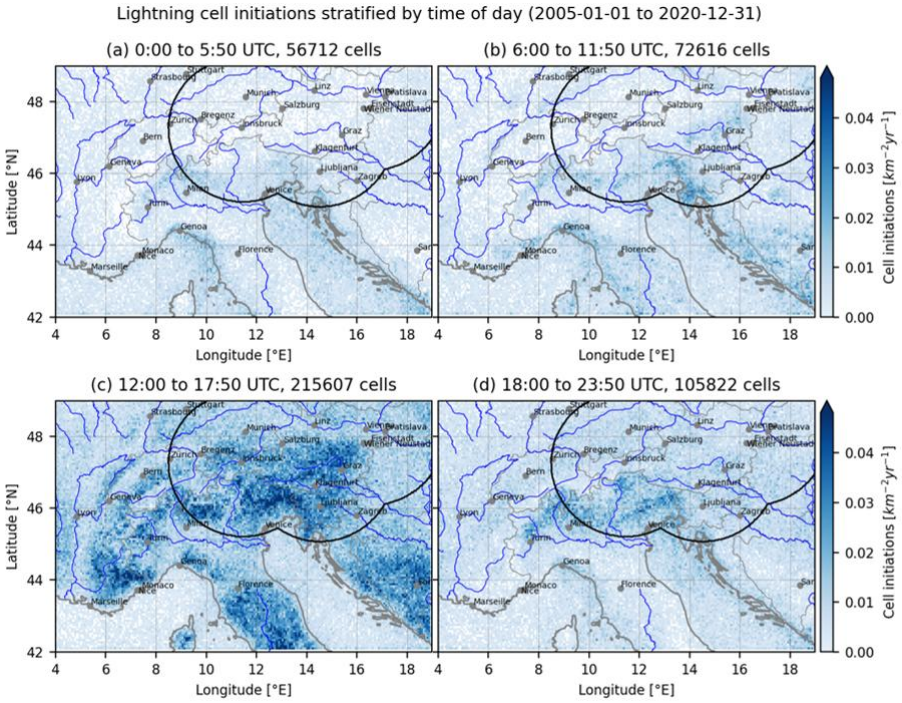

Figure 4.13.: Initiation climatologies for lightning cells (stratified by time of day, 3x3 grid boxes combined, 39.0 km^2 average pixel size).

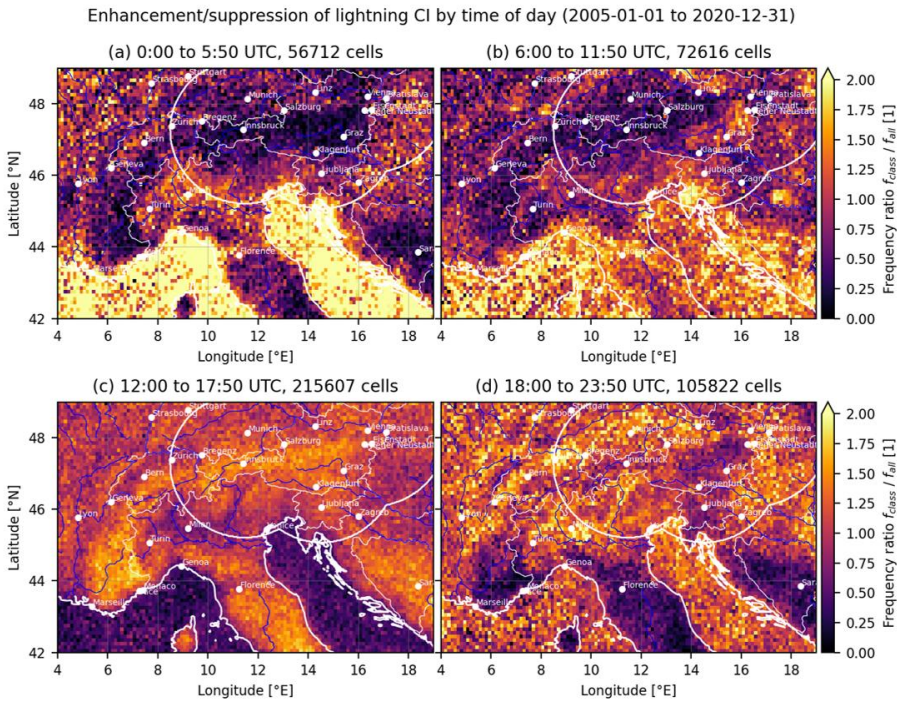

Figure 4.14.: Same data as in Fig. 4.13, but as a frequency ratio to highlight the relative enhancement or suppression of CI for each time of day class (5x5 grid boxes combined, 108 km^2 average pixel size).

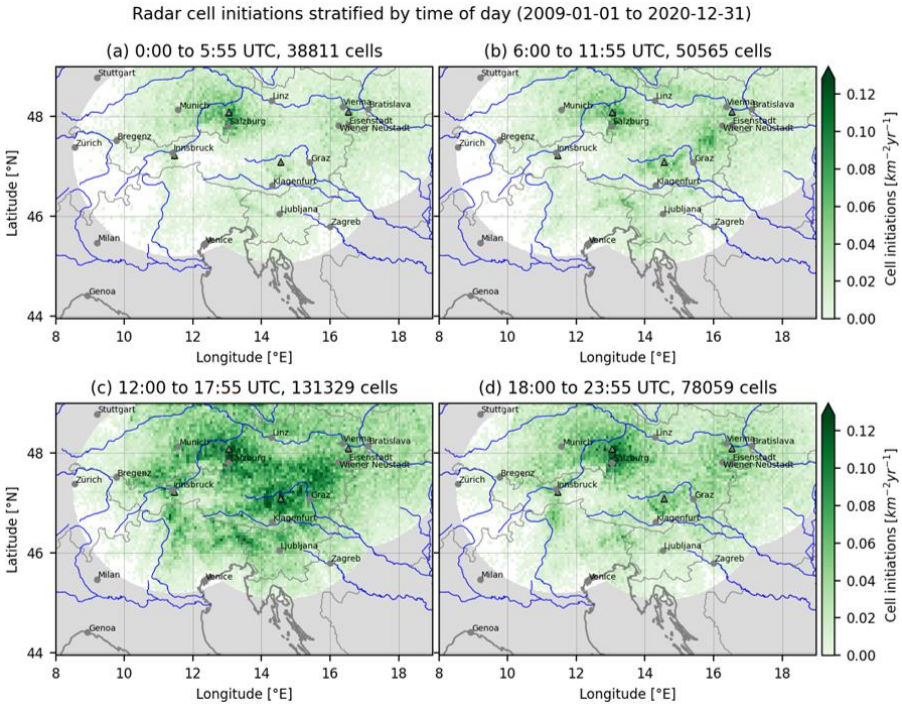

Figure 4.15.: Initiation climatologies for radar cells (stratified by time of day, 3x3 grid boxes combined, 39.0 km^2 average pixel size).

For the peak cases in afternoon and early evening (c) the orography exerts the biggest influence on CI again. What is new in comparison to other presented lightning climatologies is a strong, but temporally short maximum in the French Provence, that is not there in the other 3 panels. Finally, in the course of the evening (d) CI is enhanced over the plains, possibly due to the effects of the onset of katabatic mountain winds. Absolute cell numbers remain highest in Northern Italy including the Sarntal Alps and in the Swiss Tessin. Later at night thunderstorm activity shifts back over the sea.

The figure on radar CI during different periods of the day (Fig. 4.15) mostly supports the statements given for lightning CI, e.g. the early onset of cells in the Semmering-Wechsel-region as visible in panel (b). In the night, however, the area showing the highest activity does not lie to the south of the main Alpine ridge, but to the north. This could again be an artefact due to reduced radar capabilities over Northern Italy, but based on personal experience it happens quite often that showers of rain traverse the northern forelands from west to east also during the night.

4.3. Definition and behaviour of high impact cells

In the previous chapters the focus was entirely laid on the initiation of convective cells. Now, also the temporal behaviour of cells after their initiation should come into play, together with an analysis of cells that influence people and the environment in a severe or long-lasting manner. These so-called “high impact” cells are defined based on 3 properties:

1. **Very intense** cells resulting in potential danger no matter whether they move or stay close to a specific place. To analyse differences in the behaviour of weak and intense cells different reflectivity and flash density thresholds are set and the outcoming climatologies compared.

2. **Long-lasting** cells bearing high impact even when they are not extremely intense (either through affecting large areas and many people if they move, or by ongoing heavy precipitation at a specific place when they do not move). The definition of "long-lasting" hereby follows suggestions in Markowski and Richardson (2010), more precisely their equation 8.2, that sets a limit to the lifetime τ of single cells:

$$\tau \approx \frac{H}{w_0} + \frac{H}{v_t} \tag{4.1}$$

It is the sum of the time it takes air to rise to its equilibrium level (where the scale height of the atmosphere $H \approx 10\ km$ serves as an estimate for the height of the LNB and $w_0 \approx 5 - 10\ ms^{-1}$ is the updraft velocity) plus the produced precipitation to fall to the ground (where $v_t \approx 5 - 10\ ms^{-1}$ is the terminal fall speed of rain droplets). This results in a lifetime τ between 30 and 60 minutes. Hence, in order to avoid most single cells in the analysis a threshold of 60 minutes is applied.

3. **Far-moving** cells that influence large regions and many people independently from its other parameters. The chosen threshold of 40 km here is a compromise between an ideally even longer distance and a sufficiently large remaining sample size for plotting the climatologies. It has to be interpreted as the distance between the first and the last point of a cell track.

Intense mixed cells

Mixed cells offer the possibility of using more than one thunderstorm property in the definition of intensity. On the one hand cells producing heavy precipitation ($R > R_{max} = 50\ dBZ$) at some point in their lifetime can cause damage, on the other hand cells showing substantial electrical activity ($FL > FL_p = 0.240\ km^{-2}h^{-1}$) can be seen as impactful. Consequently cells exceeding either or both of these thresholds are included in the analysis, leading to the results shown in Fig. 4.16.

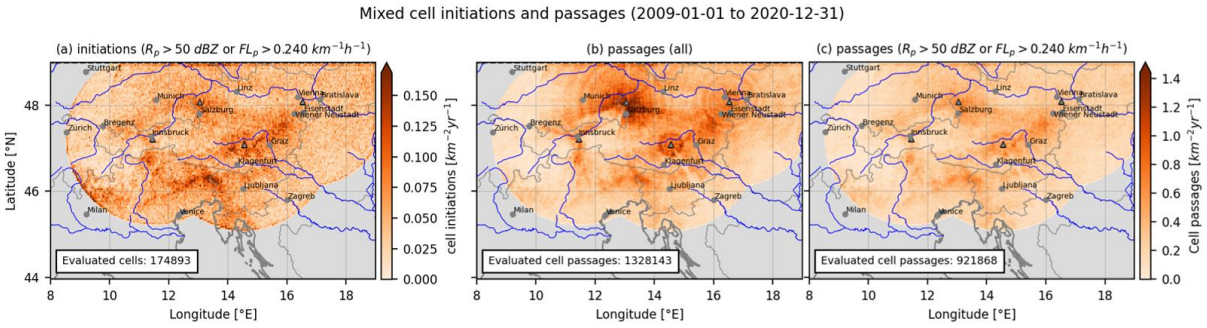

Figure 4.16.: Initiation (a) and passages (b and c) climatologies for mixed cells (stratified by maximum cell intensity, 2x2 grid boxes combined, 17.3 km² average pixel size).

In panel (a) the mixed cell initiation pattern (compare Fig. 4.3) is filtered for intense cases. The influence of the originally dominant Salzburg maximum is thereby greatly reduced, which means a large fraction of it consists of rather weak radar-only cells. This also supports the hypothesis that individual cells with a peak reflectivity close to $R_p = 41\ dBZ$ tend to be tracked more than once due to the variations of radar sensitivity in the area. Besides this delicate aspect, the mixed CI pattern inevitably resembles the lightning pattern. It shows major maxima along

the southern edge of the Alps (stretching as far as Croatia), in Southeastern Austria and over the Sarntal Alps, a minor maximum over Southern Bavaria and a pronounced minimum over the intra-Alpine areas in Switzerland. Furthermore, it needs to be recalled that the narrow maxima at the domain borders are artificial due to cells entering the domain.

Features similar to the ones in the initiation climatology can also be observed in the climatology of cell passages (panels (b) and (c)), i.e. the location of cell centroids not only at the beginning of a track, but also at all subsequent track points. The biggest difference is that the transition between features in the images of passages tends to be smoother, which stems from the 5 times higher number of sample points. It should be noted, that this number is also a first guess value for the average lifetime of a cell, corresponding to $5 \times 10\ min = 50\ min$ (more on cell lifetime follows in Chap. 4.4).

Intense radar and lightning cells

In this section the climatologies of passages are shown for each data type individually, filtered by 3 (subjectively chosen) intensity thresholds. Panel (a) in Fig. 4.17 provides passages of all radar cells that were tracked and stored by T-DaTing ($R_p > 41\ dBZ$). Again the pattern resembles the one of cell initiations, but has a more sharp and less grainy look. As the intensity threshold is raised to (b) $R_p > 50\ dBZ$ and (c) $R_p = 58\ dBZ$ (the greatest reflectivity value in the provided data files) the maximum shifts from Salzburg to Styria, similar to what was observed for mixed cells. However, the maxima in Italy are underestimated, as this is the area that benefits most from the inclusion of lightning data into mixed cells.

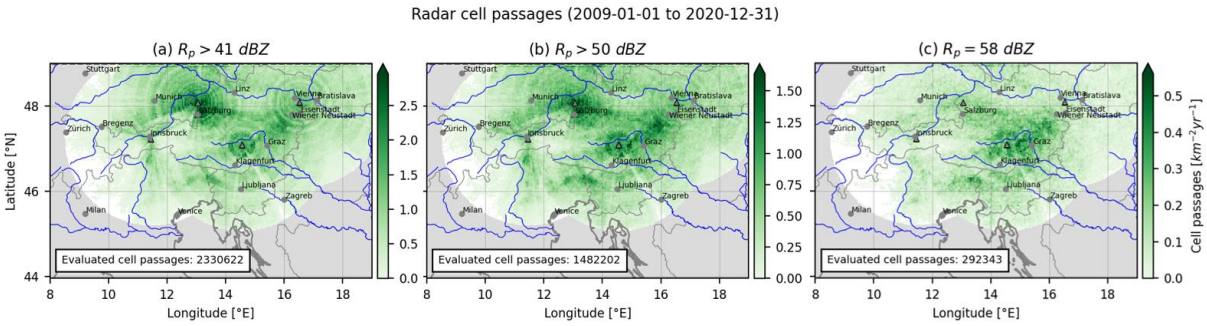

Figure 4.17.: Passages climatologies for radar cells (stratified by maximum cell intensity, 2x2 grid boxes combined, 17.3 km^2 average pixel size).

Cell passage climatologies can best tell us something about how often a region experiences strong convective activity (CI climatologies on the other hand give us a hint, but do not tell us exactly where cells move to after being initiated). Hence, Fig. 4.17 states that Styria and the bordering areas of Carinthia and Southern Lower Austria are prone to the heaviest precipitation events in Austria.

This last statement essentially also holds for the climatologies of lightning cell passages in Fig. 4.18. It should be remembered though, that the analyses of this work do not give an estimate on how many lightning strikes occur, just on how many events there are that are accompanied by lightning strikes. In this figure filtering values were chosen based on 2 thresholds already used by T-DaTing: (i) the minimum necessary flash density of lightning cells $FL_p > 0.240\ km^{-2}h^{-1}$ and (ii) the saturation value $FL_p > 3.1\ km^{-2}h^{-1}$. Apart from the maximum in Austria it is interesting to see how the maxima along the southern edge of the Alps evolve differently. While

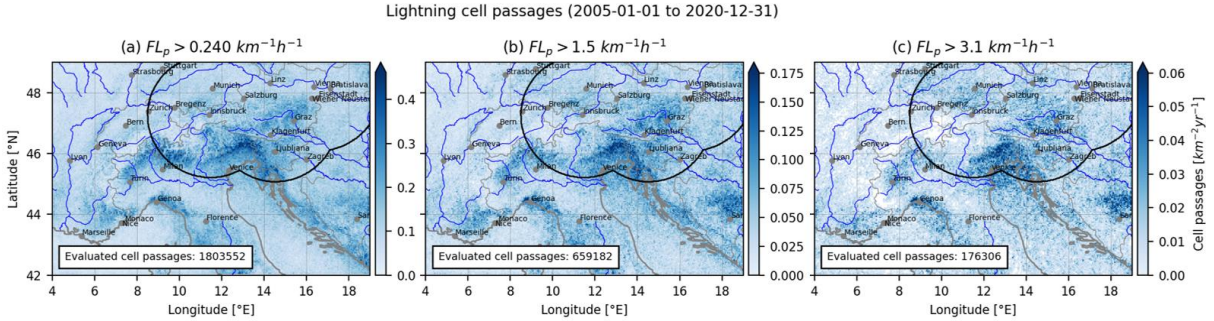

Figure 4.18.: Passages climatologies for lightning cells (stratified by maximum cell intensity, 2x2 grid boxes combined, 17.3 km^2 average pixel size).

the western maximum (north of Turin and Milan) loses influence for very lightning-intense cells, the maximum in Friuli-Venezia Giulia remains very prominent and even stretches out over the Adriatic Sea. This in addition to the climatology on lightning CI suggests that if cells initiated at the edge of the Alps move southeastwards over the plains and the sea, they are often the most electrically active ones. Other regions frequently experiencing lightning-intense storms are parts of Bosnia-Herzegovina and the area around Genoa.

Long-lasting radar and lightning cells

Now a different type of high impact is considered: cells having a continuous track of more than 7 points, i.e. having a long lifetime of more than 60 minutes. As a result single cells are excluded and better organised cells are kept in the analysis. The remaining radar cells make up a fraction of 15.8% of all cells, the respective fraction for lightning cell is 14.1%. For the purpose of this analysis, the track initiation locations are compared with the respective locations of cell decay. The difference between these 2 climatologies then hints at regions particularly prone to DMC onset (more initiations than decays) or to cell vanishing (more decays than initiations). The results are given in Fig. 4.19.

In panels (a) and (c) the initiations and decays of long-lasting radar cells are shown, which at a first glance seem to be rather similar and show the already well-known maxima. The difference plot in panel (e) however brings up some interesting, probably unexpected features. The most impressive one is the bipolar structure of the Salzburg maximum, showing cells mostly being initiated towards Bavaria while Upper Austria seems a dominant region of cell decay. This may in part be attributed to the already several times mentioned deficiencies (measurement and evaluation strategies), but it also resembles real cases of organised cells (sometimes supercells) that propagate eastwards through the Alpine forelands. Going further east, another hotspot of initiation is the Waldviertel region, which would as such not be immediately visible from panel (a). These cells usually tend to move towards the low lands of Eastern Austria in the Vienna region, where there is indeed a corresponding area of decay.

Coming to the southern parts of the Alps, regions favoring initiation are again visible in Fruli-Venezia Giulia as well as in a streak from East Tyrol over Styria to the Semmering-Wechsel-region, symbolising CI from cooler air moving across the main Alpine ridge. Finally, another bipolar feature is visible south of Innsbruck, which in part seems realisitic (Sarntal Alps), but is also a consequence of the narrow valley offering limited sight for the Patscherkofel radar (cells moving into this streak from the west causing artificial initiations and cells moving out to the east causing

decays). Similarly, one should be careful when interpreting signals close to the domain borders, as they are influenced by entering and leaving cells and the dominant direction of propagation.

For the lightning cell initiations and decays in (b) and (d) it also needs a closer look to spot relevant differences, but in (f) some features appear very clearly. First, there is a bipolarity for all major maxima in Italy, showing primary initiation regions towards the southwest and decay regions at the northeast. This supports once more the statement on the predominant direction of cell propagation. A similar structure can also be seen along the Croatian coastline, favoring CI over the Adriatic Sea and cell decay over land. Another predominant initiation area is located at the same region in Bosnia-Herzegovina as the image on very lightning-intense cell passages

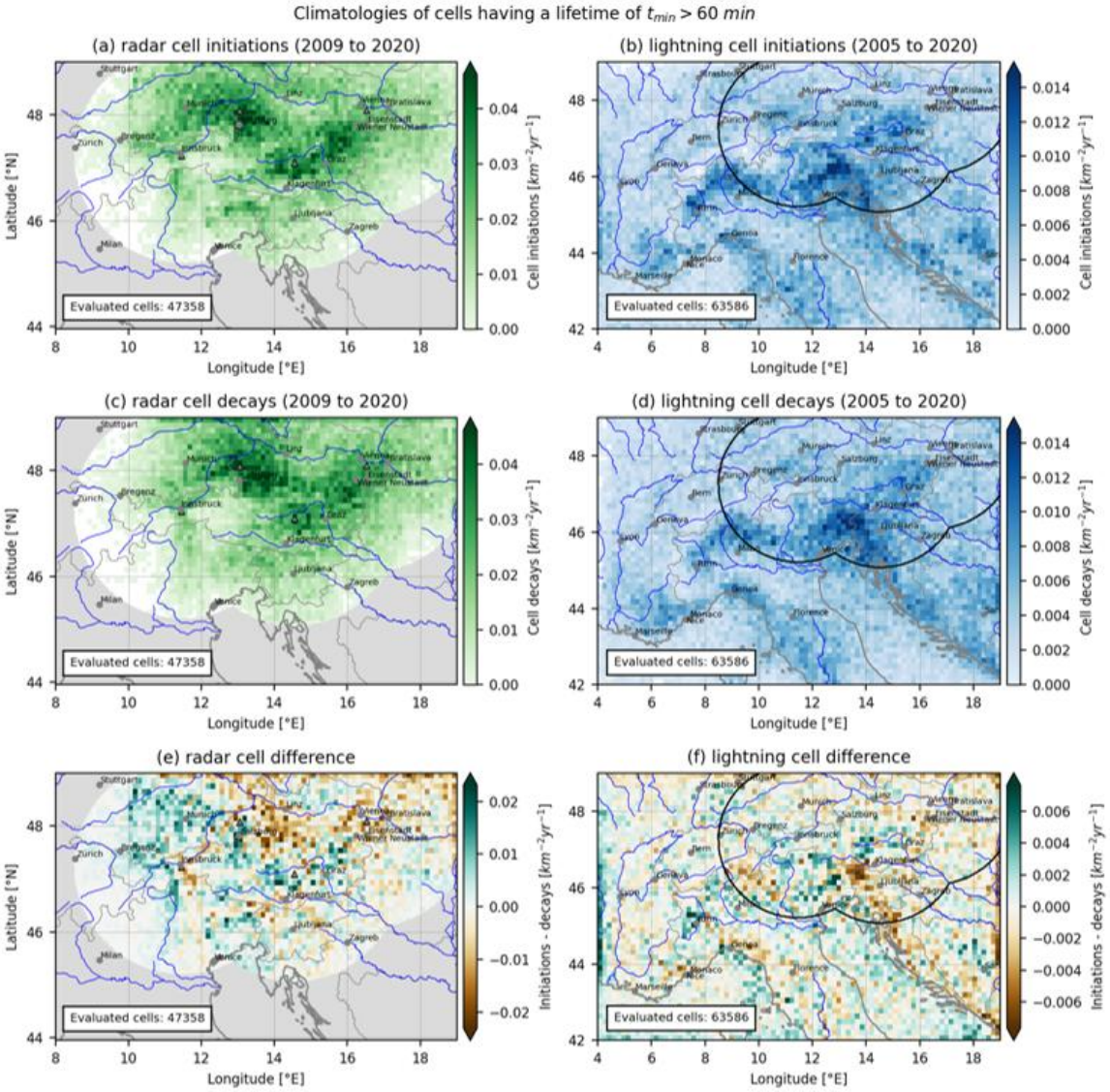

Figure 4.19.: Initiation and decay climatologies of radar (5x5 grid boxes combined, 108 km² average pixel size) and lightning (7x7 grid boxes combined, 212 km² average pixel size) cells lasting > 60 min. Positive values in the difference plot represent more initiations, negative ones more decays.

showed above. On the other hand, the maxima north of the Alps are not as clearly visible as for the radar data. Still the Black Forest in Germany, the region between Bregenz and Innsbruck and the Waldviertel in Lower Austria can be interpreted as minor initiation centers.

Far-moving radar and lightning cells

Reducing all cells to the ones propagating a distance of at least 40 km means keeping a fraction of 10.4% radar and 6.8% lightning cells only. This difference and also the one for long-lasting cells show, that strong precipitation in a convective cell is more common and lasts longer than electrical activity. The resulting climatologies of far-moving cells are depicted in Fig. 4.20.

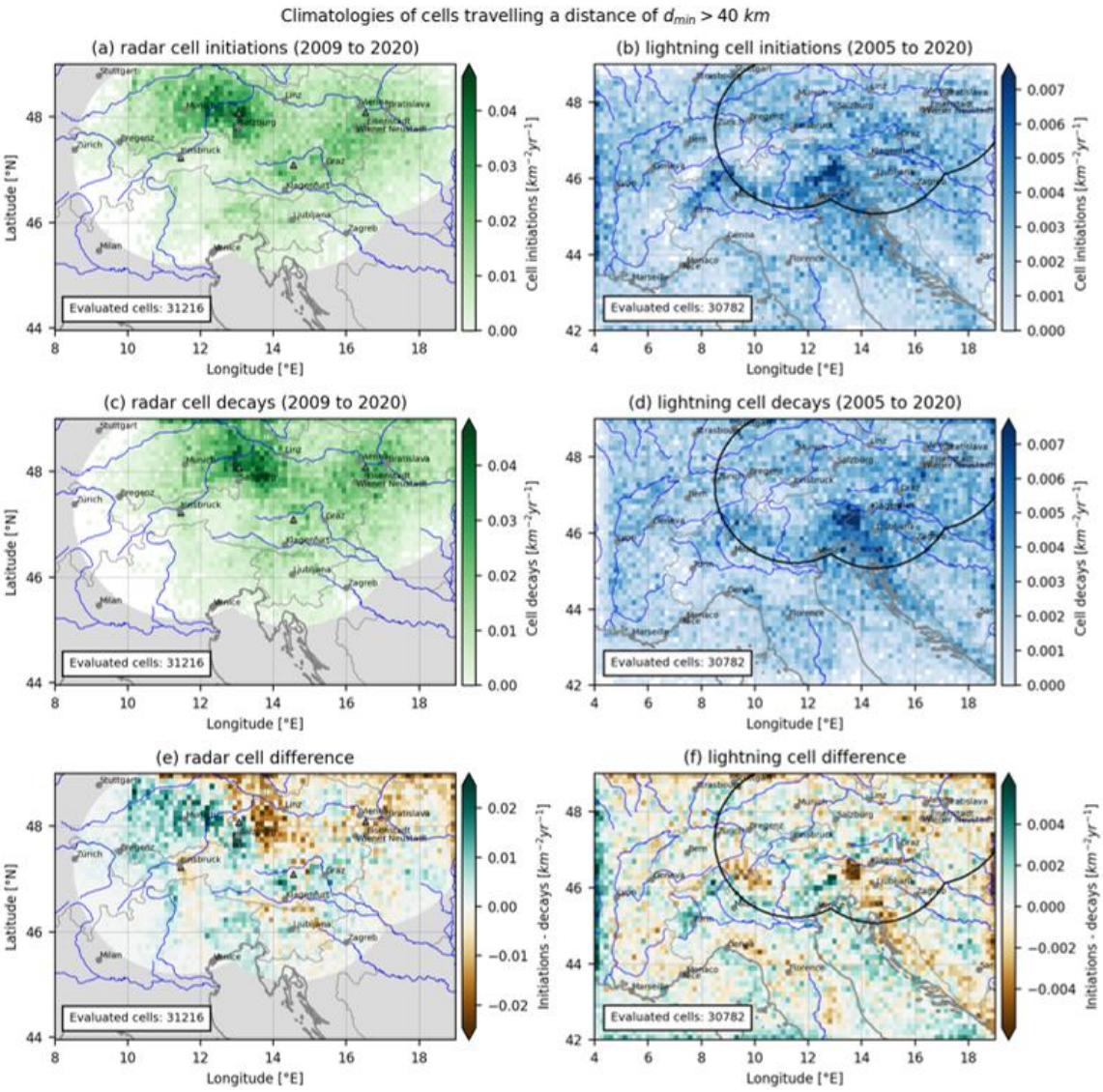

Figure 4.20.: Initiation and decay climatologies of radar (5x5 grid boxes combined, 108 km² average pixel size) and lightning (7x7 grid boxes combined, 212 km² average pixel size) cells moving > 40 min. Positive values in the difference plot represent more initiations, negative ones more decays.

The overall structure of the figures on long-lasting and far-moving convective cells does not vary much. However, the maxima around Salzburg in the radar case and the Italian-Slovenian border in the lightning case tend to become even stronger, not only in the initiation and decay plots but also in the difference plots. This is most likely caused by the unbundling of initiation and decay regions in comparison to Fig. 4.19, where some cells could last long, but also remain close to their initial position (e.g. back-building storms).

To sum up, the main statement here is that far-moving lightning cells predominantly propagate from Friuli-Venezia Giulia to the northwestern edge of Slovenia, whereas for far-moving radar-only cells a maximum is found in the norhtern Alpine forelands. To what extent this maximum is real or influenced by the discussed overestimation of radar cell initiations remains unclear.

4.4. Some statistics on cell properties and evolution

So far most of the analyses carried out were presented on spatial maps. In this chapter different cell properties will be investigated and interpreted, trying to find variations in space and time, as well as correlations between some of those properties.

Electrical activity in mixed cells

Fig. 4.4 already showed that mixed cells can be split into 3 distinct parts, namely cells being kept in the analysis due to their intense enough radar parts (56%), due to their lightning parts (11%) and cells having both (only 33%). These differences can now be analysed even further by comparing the percentages of radar and lightning activity in mixed cells in relation to the total track length. Fig. 4.21 compares the period of radar or lightning activity on one axis (2 classes each) with the total lifetime of a cell on the other. Also thin lines representing 20, 40, 60 and 80 percent of the total lifetime are shown.

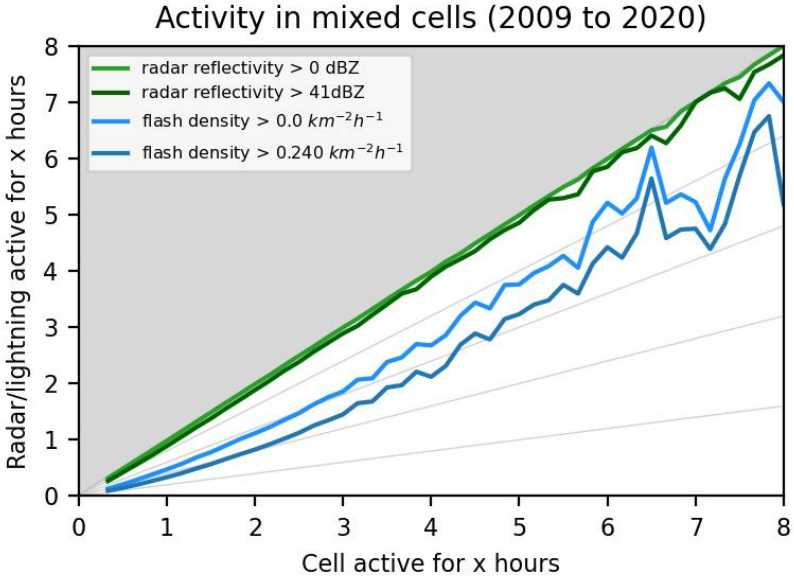

Figure 4.21.: Fractions of the lifetime of mixed cells exceeding different radar and lightning maximum intensity thresholds. Thin grey lines are drawn 20% apart.

First of all, it can be seen that most of the mixed cells contain radar signals, as the green lines are close to the diagonal (especially when allowing for weak echoes < 41 dBZ). Lightning activity though is not as common. For example, when looking at short-lived cells to the left of the plot only around 50% of all cells include some sort of lightning signal, hardly 40% have a peak flash density $FL_p > 0.240\ km^{-2}h^{-1}$. When moving towards longer cell lifetimes, these percentages go up to around 70% for cells lasting 5 hours. The lines become increasingly jumpy though, as the number of underlying cells gets smaller. The fact that most mixed cells are lasting < 3 h can also be seen when adding the respective percentages in the above paragraph: $11\% + 33\% = 44\%$, which matches the dark blue line in the left plot half.

To summarize the main message of this figure, it can be said that although cell numbers go down for long-lasting cells the percentage of electrical activity goes up. Both, the high electrical activity and the long lifetime, hint at a better organised structure of such cells.

Cell lifetime, size and intensity in relation to synoptic flow

Lifetime, maximum size and maximum intensity are 3 cell properties that have already been used in the definition of high impact cells in Chap. 4.3. In the next sections the distribution of cell numbers in relation to these 3 properties will be investigated, starting with a stratification by the 11 flow types introduced in Tab. 3.1.

At first, Fig. 4.22 shows cell percentages for individual classes of cell lifetime (for all 3 data and all 11 flow types). Thick green, blue and orange lines denote the average distribution of each flow type. This and the upcoming figures should be interpreted as histograms. Percentages instead of absolute cell numbers are used to better highlight possible differences between flow types. Each class is 10 minutes wide and characterised by the lower boundary (i.e. an x value of 30 min represents cell lifetimes between 30 and 39 min). Furthermore, as cell numbers go down nearly exponentially with cell lifetime, a logarithmic scale is used (symmetrical log) that also allows for zeros in the data. It involves a transition between a purely logarithmic part and a linear part, evident as a change in the slope of all lines below 3%.

Apart from the exponential decay of cell numbers, no significant differences between flow types can be observed in relation to cell lifetime, even independent from the respective data type. Only very few cells last longer than 3 hours, supporting the conclusions drawn from the numbers and Fig. 4.21 in the previous section. The steep drop in the left of each plot stems from the track length criterion, keeping only cells with a lifetime > 10 mins in the analysis.

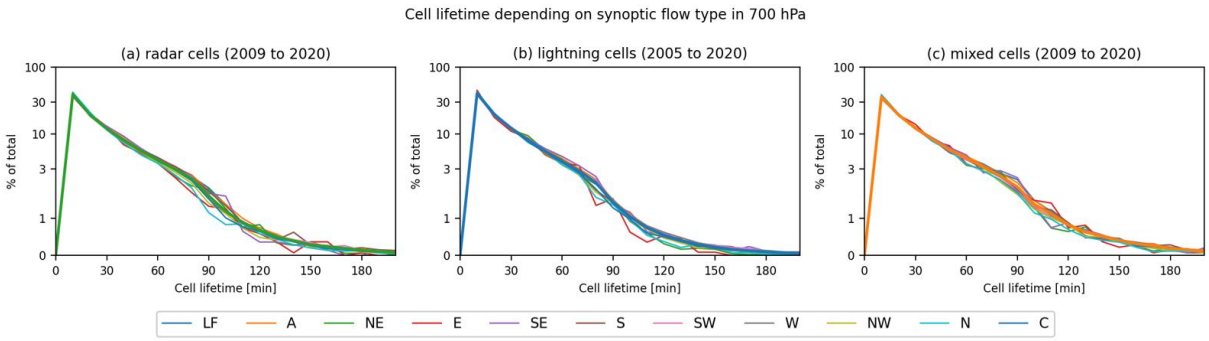

Figure 4.22.: Distributions of radar, lightning and mixed cell lifetimes in relation to synoptic flow. Thick green, blue and orange lines denote the respective sum of all flow types.

Coming to the relation between cell numbers and maximum cell size, the distributions are not as uniform anymore, although again an overall exponential decline is visible in Fig. 4.23 (each class is 40 km² wide). For radar cells in panel (a) and mixed cells in (c) the features are rather similar and show the largest cells originating under flows from west to south. In turn the cyclonic flow type seems to be a good proxy for the overall distribution. Finally, cells remain comparably small for easterly flows, flows from the north and the anticyclonic case. This could be an indication that in such environments cells do not organise that efficiently.

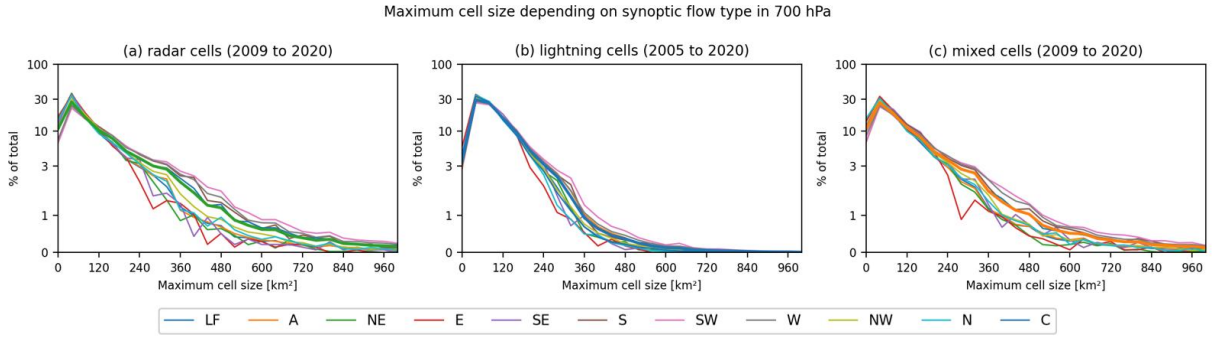

Figure 4.23.: Distributions of radar, lightning and mixed cell maximum sizes in relation to synoptic flow. Thick green, blue and orange lines denote the respective sum of all flow types.

Panel (b) on lightning cells shows a somewhat similar order of flow types (with the largest cells for the SW case and the smallest ones for the E case). But the range of lines is not as wide and cells generally tend to be smaller, as the lines are more rapidly declining. Contrary to the figures on cell lifetime, the lines here do not start at 0 as the minimum area criterion allows for cell sizes between 20 and 40 km².

Lastly, the distributions for maximum cell intensity are given in Fig. 4.24, splitting up mixed cells into a radar and a lightning part. The class widths are 4 dBZ and 0.25 $km^{-2}h^{-1}$ respectively.

For radar cells in panel (a) only 5 classes are available due to the discreteness of the original reflectivity data: 41.8, 46.0, 50.2, 54.3 and 58.0 dBZ. There is a clear trend for cells in the (anti)cyclonic flow types to be more intense, the presence of a low or high pressure system above the Alps therefore seems to favor the production of intense precipitation. For low pressure systems this could be explained by the added synoptic scale uplift of air, for high pressure systems it is

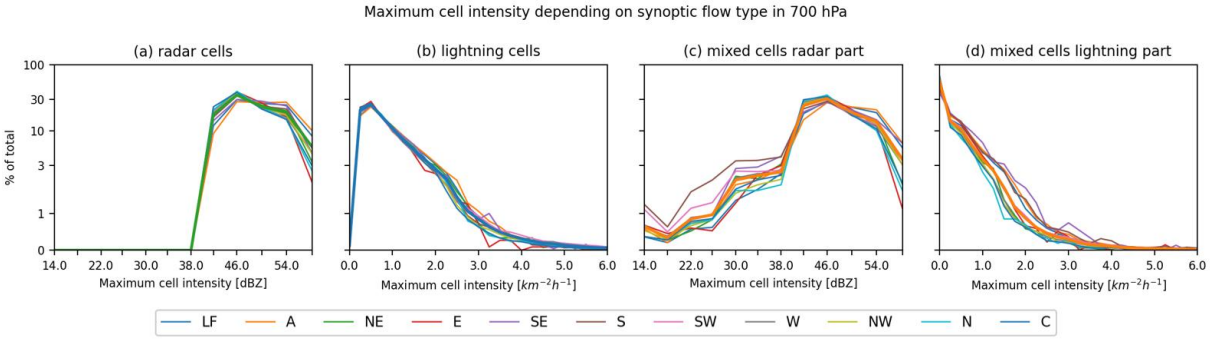

Figure 4.24.: Distributions of radar, lightning and mixed cell maximum intensities in relation to synoptic flow. Thick green, blue and orange lines denote the respective sum of all flow types.

more surprising. On the other side, comparably weak cells mostly occur in low flow situations (or the sparsely populated northerly flow case). The radar part of mixed cells in panel (c) resembles the data for pure radar cells in the high intensity range. However it also allows for reflectivities weaker than 41 dBZ, if the corresponding lightning part fulfills its minimum intensity criterion ($FL_p > 0.240\ km^{-2}h^{-1}$).

Apart from the again exponential decline, the maximum intensity of pure lightning cells in panel (b) does not show much dependence on the synoptic flow type (only a slight favoring of the cyclonic and anticyclonic cases). Interestingly, this in not the case for the lightning part of mixed cells in panel (d), where the dependence on the flow type is stronger. Moreover, it is not only the (anti)cyclonic case that stands out, but also the southeasterly and southerly flows. This difference can only be explained by the different underlying domains. Therefore stronger cells in the SE and S class must appear around Austria than they do for the whole Alpine region (though still their number is limited). Finally, in this panel the lowest intensity class is again populated with cells that do not meet the minimum intensity criterion for lightning data, but the one for radar data.

Cell lifetime, size and intensity in relation to month and time of day

It is also worth taking a look at the relation between cell lifetime, maximum size and maximum intensity on the one side and time of year and time of day on the other side. However, this analysis is carried out in less detail and the results are combined in Fig. 4.25, which can be interpreted in the same way as before.

Panels (a) and (b) show the dependence on cell lifetime for all 3 data types (radar, lightning and mixed), stratified by the time of cell initiation. Contrary to the images on flow classes the 12 or 24 thin lines representing each month or hour are colored equally, which at least allows statements on the variability. In panels (c) and (d) the dependence on the maximum cell size are similarly depicted, followed by an analysis of radar intensity in panels (e) and (f), and of lightning intensity in panels (g) and (h).

Hereafter, the most important messages of this figure are given:

- For all 3 cell properties the variability of different months is greater than the one of different hours. This means the average size, intensity and lifetime of a cell depend much more on the season than they do on the time of day. Hence, cells initiated during a subsequent day and night do statistically not vary as much as cells initiated in summer and spring or fall.
- The maximum cell size does depend much more on the respective time measure than the cell lifetime does, similar to what was observed for synoptic flow types.
- When looking at the thick overall lines for each data type, it again becomes clear that mixed cells are essentially radar dominated as the respective lines do not vary much. Especially for the maximum size lightning cells show a different behaviour, as there tends to be a higher percentage of rather small cells (again similar to the analysis on flow types). This has to do on the one hand with the design of lightning cells (Gaussian smoothing), but it could also mean not all parts of convective cells are electrically active.
- The splitting of mixed cell intensities into a radar and a lightning part requires careful interpretation of the figures in the bottom row. As 56% of mixed cells have a maximum lightning intensity below 0.240 $km^{-2}h^{-1}$, only the remaining 44% are distributed across large parts of panels (g) and (h), leading mixed cells to look less electrically intense. For better comparability the percentages of mixed cells can be approximately doubled, which causes the differences to be considerably less striking.

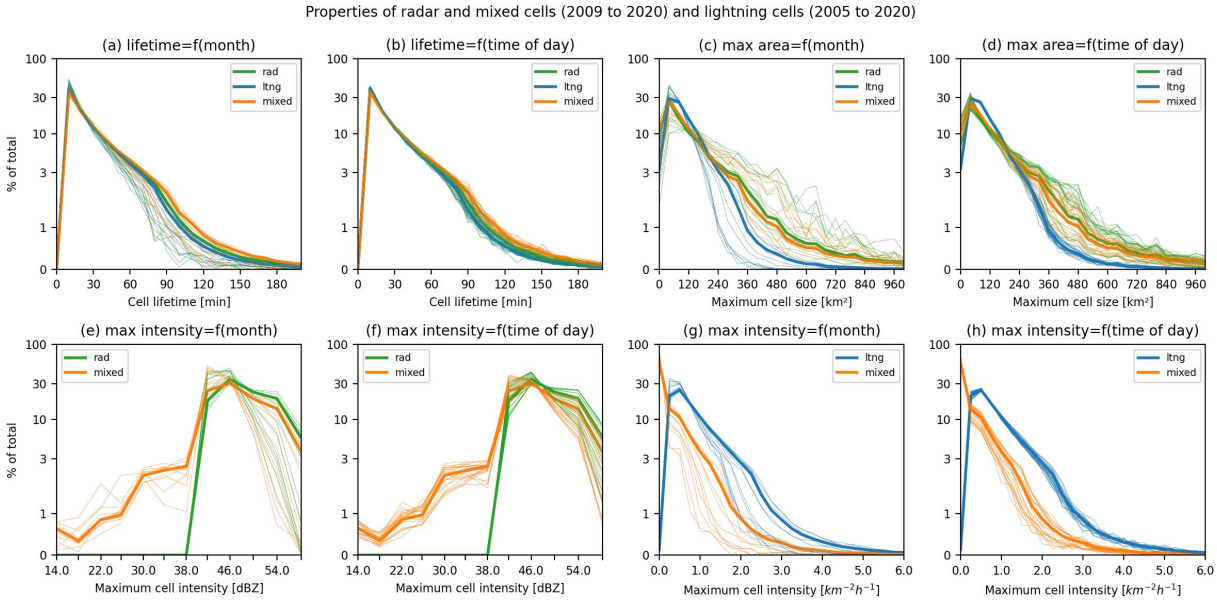

Figure 4.25.: Distributions of radar, lightning and mixed cell lifetimes, maximum sizes and maximum intensities in relation to month and time of day. Thick green, blue and orange lines denote the respective sum of all flow types.

Correlations between cell lifetime, size and intensity

The results chapter is concluded by studying possible correlations between the 3 main cell properties in form of 2D histograms and a linear regression analysis in Fig. 4.26. The maximum cell intensity is depicted as a function of the maximum size (panels (a) and (d)) and as a function of the cell lifetime (panels (c) and (f)). And cell lifetime is depicted as a function of the maximum size (panels (b) and (e)). The bin widths for radar data are 4 dBZ, 40 km^2 and 5 min, the ones for lightning data are 0.1 $km^{-2}h^{-1}$, 40 km^2 and 10 min. For each combination of property and data type the regression function and the corresponding coefficient of determination are added in a box.

The analysis on maximum radar cell intensity and maximum size in panel (a) does not result in any substantial correlation. Furthermore, it shows the limitations of using only a discrete number of reflectivity levels, and it implies the existence of few, but very large radar cells around the 50 dBZ range. These very large cells are not detectable for lightning data in panel (d). Generally the correlation between maximum lightning cell intensity and maximum size is the strongest of all 6, the R^2 coefficient implying around 60% of the variations in intensity can be explained by the cell size. Following the regression line, a doubling of the cell size approximately results in a doubling of the flash density and consequently 4 times more lightning flashes in the whole system.

In terms of cell lifetime being a function of the maximum cell size, again different behaviour of radar (b) and lightning cells (e) is found. For radar data there seems to be a weak correlation in 2 directions, as comparably small but long-lasting cells are just as frequent in the sample as large cells with a limited lifetime. For lightning cells in panel (e) the distribution is more narrow, hinting at larger cells generally lasting longer, but the correlation is not as high as for cell intensity.

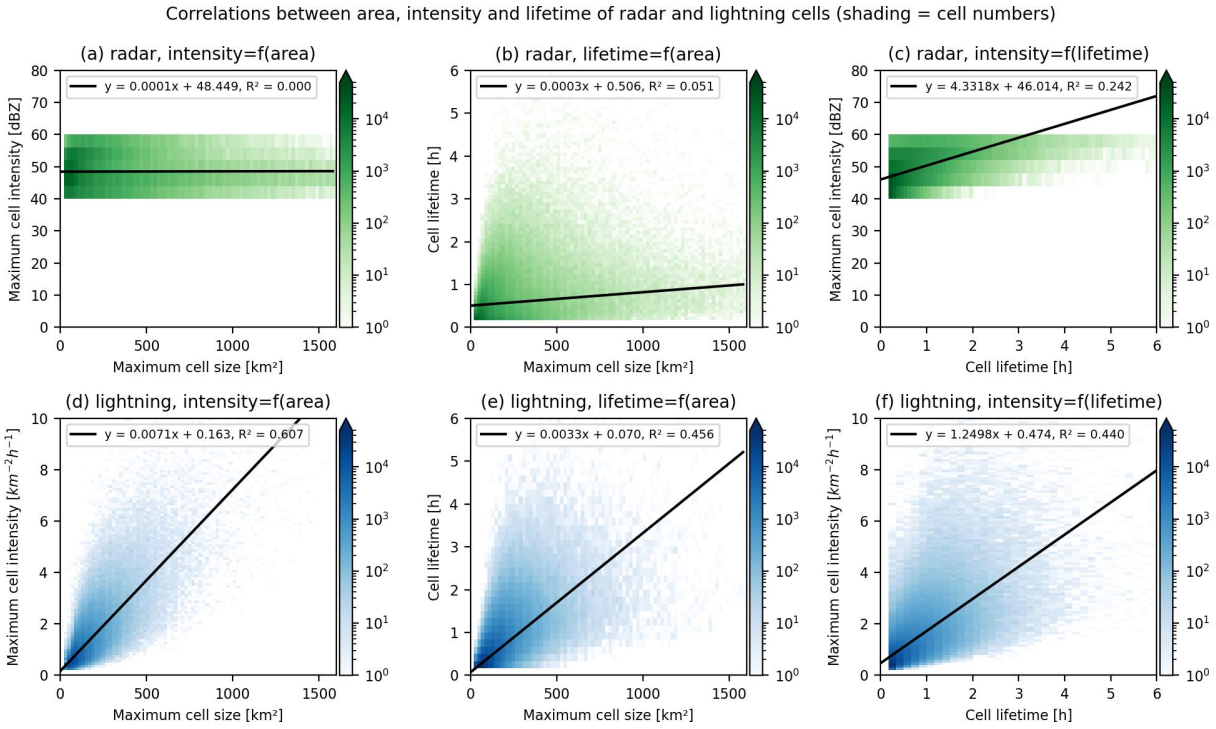

Figure 4.26.: Correlations between radar and lightning cell lifetimes, maximum sizes and maximum intensities. Linear fits are included as a simple attempt to quantify the degree of interdependence between respective pairs of cell properties.

The last possible combination of maximum cell intensity and lifetime finally suggests similar results for radar and lightning data, i.e. an increase of intensity with lifetime. Whereas radar cells with a peak reflectivity of 41 dBZ hardly last longer than 1.5 hours, beyond 5 hours are observed for cells reaching 58 dBZ. For these very intense radar cells the decrease of cell numbers with lifetime (or maximum size) is not exponential any more, but shows a rather uniform distribution over the first few hours (or square kilometers). The degree of intensity variability explainable by the lifetime of lighning cells in panel (f) is similar to the one between lifetime and size. Doubling the cell lifetime does statistically imply a near doubling of flash density, which again hints at the better organised structure of long-lasting cells.

5. Conclusions

The previous pages have shown that spatiotemporal climatologies of DMC as well as analyses of cell properties offer a manifold of stratification and interpretation possibilities and are hence ideal study objects for the occurrence and behaviour of convective cells, particularly thunderstorms.

In the course of this work a variety of interesting, sometimes even surprising features and relationships has been found. The most relevant statements, especially but not only drawn from climatologies on lightning data, are summarised and grouped in the following lists:

1. General statements on DMC:
   - DMC in the Alpine region is mainly driven by the orography (compare all CI climatologies, Fig. 4.3, 4.6, 4.8, 4.10, 4.12, 4.13 and 4.15) and reaches a peak during July and times between 14 and 15 UTC (Fig. 4.9).
   - The dominant direction of cell propagation in the Alpine region is from the west-southwest towards the east-northeast (compare signals of cells entering and leaving the domain, e.g. in Fig. 4.20f).
   - Strong precipitation in a convective cell is more common and lasts longer than electrical activity (e.g. Fig. 4.21 or the percentages of long-lasting and far-moving radar and lightning cells).
   - CI generally seems to be favored near prominent jumps in terrain height, not necessarily in the middle of a mountain range (e.g. Fig. 4.3 or Fig. 4.6acfg).
   - The Sarntal Alps in South Tyrol are a confined, but powerful CI hotspot in the middle of the Alps for some configurations (including W, SW and A flow types in Fig. 4.6 and 4.8, or summer months in Fig. 4.10 and 4.12).
2. Basic statements on radar, lightning and mixed cells:
   - The strong maximum around Salzburg in the radar climatologies mostly stems from rather weak, electrically inactive convective cells, combined with the discussed overestimation due to radar scan strategies (Fig. 4.4).
   - Only 33% of mixed cells are made up of both sufficiently strong radar and lightning parts. 56% have a strong radar part, 11% a strong lightning part only (compare cell numbers in Fig. 4.4).
   - Most mixed cells contain a radar signal. The percentage of electrical activity goes up for long-lasting mixed cells, while cell numbers simultaneously go down (Fig. 4.21).
3. Statements in relation to synoptic flow:
   - For both, radar and lightning data, CI events are most likely for flows at the 700 hPa level coming from the west or southwest together with low flow cases and situations in which a low or high pressure system is located near the Alps. Hardly any CI events occur in easterly flows (Fig. 4.5).

- For the low flow case CI is predominantly attached to mountain range edges, for anticyclonic cases it is centered over the mountain ranges (both depending on mountain-induced mesoscale circulations, Fig. 4.6ab and 4.8ab) and for cyclonic cases CI maxima are shifted towards the Mediterranean - closely related to cyclogenesis over Northern Italy (Fig. 4.6c).
- For westerly and southwesterly flow CI maxima are observed on the windward side and sometimes even in the lee of mountain ranges (e.g. Fig. 4.6fg).

4. Statements in relation to time of year and day:
   - In spring the first onset of DMC usually happens over land (quick warming of land masses, Fig. 4.10a). After its main season in summer, DMC shifts towards the sea/Mediterranean area in autumn/winter (water remains comparably warm for a longer time, Fig. 4.10f).
   - In the course of a day regions of predominantly early DMC onset are Istria in Croatia and, for Austria, the Semmering-Wechsel-region (Fig. 4.13b and 4.15b). After its afternoon peak, DMC last longest in and around Northern Italy (Fig. 4.13d). During the second half of the night maxima shift to the sea (again profiting from water being relatively warmer than air, Fig. 4.13a).
5. Statements in relation to intense, long-lasting and far-moving cells:
   - The region with the most frequent radar cell passages in the domain lies close to Salzburg, whereas the most intense radar cells move across Styria (Fig. 4.17). For lightning data it shows that cells originating at the Friuli-Venezia Giulian edge of the Alps and subsequently moving southeastwards over the plains and the Adriatic sea are the most electrically active ones (Fig. 4.18).
   - Often a bipolar structure of areas with predominant cell initiation and cell decay is observed for cells lasting $>$ 60 min (Fig. 4.19ef) and moving $>$ 40 km (Fig. 4.20ef).
6. Statements on cell property statistics:
   - Cell numbers go down nearly exponentially with increasing cell lifetime (Fig. 4.25ab), maximum size (Fig. 4.25cd) and maximum flash density (Fig. 4.25gh). For maximum radar reflectivity the behaviour is different due to the limit number of levels´(Fig. 4.25ef).
   - Cell lifetime does not depend much on different flow types (Fig. 4.22), maximum cell size (Fig. 4.23) and maximum cell intensity (Fig. 4.24) instead do so. For example, the largest cells originate under flows from west to south, and the presence of a low or high pressure system (flow types C or A) seems to favor the production of intense precipitation.
   - The average size, intensity and lifetime of a cell depend much more on the season than they do on the time of day. Moreover, the maximum cell size does depend much more on the respective time measure than the cell lifetime does (Fig. 4.25).
   - There is a pronounced correlation between maximum lightning cell intensity and maximum size. Around 60% of the variations in intensity can be explained by the cell size. A doubling of the cell size approximately results in a doubling of the flash density and consequently 4 times more lightning flashes in the whole system (Fig. 4.26d).

- Larger lightning cells generally tend to last longer (Fig. 4.26e).
- For both, radar and lightning cells, there is an increase of intensity with cell lifetime (Fig. 4.26cf). For lightning cells a doubling of the cell lifetime does statistically imply a near doubling of flash density.

The results show that (i) it is justified to use tracking algorithms in a climatological sense, provided that homogeneous datasets are used, and (ii) it is beneficial to incorporate more than one dataset into the analyses, provided that the same method is used. To serve these constraints, the originally purely radar-based tracking algorithm T-DaTing has been successfully adapted towards the use of lightning data. This included choosing physically reasonable input parameters/thresholds and smoothing the lightning data in a way that creates detectable and trackable cell objects. Now in principle, this strategy could also be applied for other data types related to DMC, such as data retrieved from satellites or hail signatures.

Similar to what Taszarek et al. (2019) found, climatologies based on lightning data have proven to be the most robust ones. On the other hand radar climatologies face a number of problems and shortcomings. For example, an underestimation of CI events in mountainous areas such as the Alps is observed due to radar beam blockage, while there is an overestimation near radar sites stemming from the scan strategy (using discrete elevation angles and causing different cell detection sensitivities in neighbouring areas). The resulting radar CI pattern still resembles the radar-derived precipitation sums in Fig. 1.2 (originally from Kaltenboeck and Steinheimer (2015)), proving that these radar shortcomings do not relate to the use of tracking algorithms, but are intrinsically anchored in the data.

Here it also should be recalled that Meyer et al. (2013) state climatologies tend to benefit from the use of multiple data sources. In principal this seems valid, as radar data provides long-lasting, rather smooth and therefore easily trackable features and lightning sensors offer a relatively homogeneous detection efficiency. The potential of lightning data would therefore indeed be great to fill the gaps in the radar image caused by radar beam blockage, which in this work leads to the introduction of mixed cells. But in practice the number of lightning CI events is just too small to significantly alter the pattern created by the radar data. Only when excluding mixed cells with radar signals < 50 dBZ and no electrically active parts, the resulting climatologies resemble the more robust ones of lightning data.

Bibliography

Anderson, G. and D. Klugmann (2014), A European lightning density analysis using 5 years of ATDnet data. *Natural Hazards and Earth System Sciences* **14**(4), 815–829. doi:10.5194/nhess-14-815-2014.

Antonescu, B. and S. Burcea (2010), A Cloud-to-Ground Lightning Climatology for Romania. *Monthly Weather Review* **138**(2), 579–591. doi:10.1175/2009MWR2975.1.

Bertram, I. and G. J. Mayr (2004), Lightning in the eastern Alps 1993-1999, part I: Thunderstorm tracks. *Natural Hazards and Earth System Sciences* **4**(4), 501–511. doi:10.5194/nhess-4-501-2004.

Beucher, S. and C. Lantuejoul (1979), Use of Watersheds in Contour Detection, International workshop on image processing: Real-time Edge and Motion Detection/Estimation. URL: `https://people.cmm.minesparis.psl.eu/users/beucher/publi/watershed.pdf`. Last access: 8 August 2024.

Cacciamani, C., F. Battaglia, P. Patruno, L. Pomi, A. Selvini, and S. Tibaldi (1995), A Climatological Study of Thunderstorm Activity in the Po Valley. *Theoretical and Applied Climatology* **50**(3-4), 185–203. doi:10.1007/BF00866116.

Cancelada, M., P. Salio, D. Vila, S. W. Nesbitt, and L. Vidal (2020), Backward Adaptive Brightness Temperature Threshold Technique (BAB3T): A Methodology to Determine Extreme Convective Initiation Regions Using Satellite Infrared Imagery. *Remote Sensing* **12**(2), 337. doi:10.3390/rs12020337.

Costa, S., P. Mezzasalma, V. Levizzani, P. P. Alberoni, and S. Nanni (2001), Deep convection over Northern Italy: synoptic and thermodynamic analysis. *Atmospheric Research* **56**(1-4), 73–88. doi:10.1016/S0169-8095(00)00091-0.

Dixon, M. and G. Wiener (1993), TITAN: Thunderstorm Identification, Tracking, Analysis, and Nowcasting - A Radar-based Methodology. *Journal of Atmospheric and Oceanic Technology* **10**(6), 785–797. doi:10.1175/1520-0426(1993)010<0785:TTITAA>2.0.CO;2.

Feldmann, M., U. Germann, M. Gabella, and A. Berne (2021), A characterisation of Alpine mesocyclone occurrence. *Weather and Climate Dynamics* **2**(4), 1225–1244. doi:10.5194/wcd-2-1225-2021.

Feudale, L. and A. Manzato (2014), Cloud-to-Ground Lightning Distribution and Its Relationship with Orography and Anthropogenic Emissions in the Po Valley. *Journal of Applied Meteorology and Climatology* **53**(12), 2651–2670. doi:10.1175/JAMC-D-14-0037.1.

Funatsu, B. M., J.-F. Rysman, C. Claud, and J.-P. Chaboureau (2018), Deep convective clouds distribution over the Mediterranean region from AMSU-B/MHS observations. *Atmospheric Research* **207**, 122–135. doi:10.1016/j.atmosres.2018.03.003.

Germann, U., M. Boscacci, L. Clementi, M. Gabella, A. Hering, M. Sartori, I. V. Sideris, and B. Calpini (2022), Weather Radar in Complex Orography. *Remote Sensing* **14**(3), 503. doi:10.3390/rs14030503.

Handwerker, J. (2002), Cell tracking with TRACE3D - a new algorithm. *Atmospheric Research* **61**(1), 15–34. doi:10.1016/S0169-8095(01)00100-4.

Hering, A. M., C. Morel, G. Galli, S. Sénési, P. Ambrosetti, and M. Boscacci (2004), Nowcasting thunderstorms in the Alpine region using a radar based adaptive thresholding scheme. In: *Proceedings of ERAD (2004)*, pp. 206–211, Copernicus Publications.

Herrera-Lormendez, P. (2022), PedroLormendez/jcclass: version 0.0.4. URL: `https://pedrolormendez-jcclass.readthedocs.io/en/latest/index.html`. Last access: 8 August 2024. doi:10.5281/zenodo.7025220.

Holley, D. M., S. R. Dorling, C. J. Steele, and N. Earl (2014), A climatology of convective available potential energy in Great Britain. *International Journal of Climatology* **34**(14), 3811–3824. doi:10.1002/joc.3976.

Johns, R. H. and C. A. Doswell (1992), Severe Local Storms Forecasting. *Weather and Forecasting* **7**(4), 588–612. doi:10.1175/1520-0434(1992)007<0588:SLSF>2.0.CO;2.

Johnson, J., P. L. MacKeen, A. Witt, D. E. Mitchell, G. J. Stumpf, M. D. Eilts, and K. W. Thomas (1998), The Storm Cell Identification and Tracking Algorithm: An Enhanced WSR-88D Algorithm. *Weather and Forecasting* **13**(2), 263–276. doi:10.1175/1520-0434(1998)013<0263:TSCIAT>2.0.CO;2.

Kaltenboeck, R. (2012), New generation of dual polarized weather radars in Austria. In: *ERAD 2012 - The Seventh European Conference on Radar in Meteorology and Hydrology*.

Kaltenboeck, R. and M. Steinheimer (2015), Radar-based severe storm climatology for Austrian complex orography related to vertical wind shear and atmospheric instability. *Atmospheric Research* **158-159**, 216–230. doi:10.1016/j.atmosres.2014.08.006.

Kirshbaum, D. J., B. Adler, N. Kalthoff, C. Barthlott, and S. Serafin (2018), Moist Orographic Convection: Physical Mechanisms and Links to Surface-Exchange Processes. *Atmosphere* **9**(3), 80. doi:10.3390/atmos9030080.

Kober, K. and A. Tafferner (2009), Tracking and nowcasting of convective cells using remote sensing data from radar and satellite. *Meteorologische Zeitschrift* **18**(1), 75–84. doi:10.1127/0941-2948/2009/359.

Kyznarová, H. and P. Novák (2009), CELLTRACK - Convective cell tracking algorithm and its use for deriving life cycle characteristics. *Atmospheric Research* **93**(1-3), 317–327. doi:10.1016/j.atmosres.2008.09.019.

Levizzani, V., F. Pinelli, M. Pasqui, S. Melani, A. G. Laing, and R. E. Carbone (2010), A 10-year climatology of warm-season cloud patterns over Europe and the Mediterranean from Meteosat IR observations. *Atmospheric Research* **97**(4), 555–576. doi:10.1016/j.atmosres.2010.05.014.

Lucas, B. D. and T. Kanade (1981), An Iterative Image Registration Technique with an Application to Stereo Vision. In: *Proceedings of Imaging Understanding Workshop*, pp. 121–130.

Manzato, A. (2003), A climatology of instability indices derived from Friuli Venezia Giulia soundings, using three different methods. *Atmospheric Research* **67-68**, 417–454. doi:10.1016/S0169-8095(03)00058-9.

Manzato, A. (2007), The 6 h climatology of thunderstorms and rainfalls in the Friuli Venezia Giulia Plain. *Atmospheric Research* **83**(2-4), 336–348. doi:10.1016/j.atmosres.2005.08.013.

Manzato, A. (2012), Hail in Northeast Italy: Climatology and Bivariate Analysis with the Sounding-Derived Indices. *Journal of Applied Meteorology and Climatology* **51**(3), 449–467. doi:10.1175/JAMC-D-10-05012.1.

Manzato, A., S. Serafin, M. M. Miglietta, D. Kirshbaum, and W. Schulz (2022), A Pan-Alpine Climatology of Lightning and Convective Initiation. *Monthly Weather Review* **150**(9), 2213–2230. doi:10.1175/MWR-D-21-0149.1.

Markowski, P. and Y. Richardson (2010), *Mesoscale Meteorology in Midlatitudes*. John Wiley & Sons, Chichester, West Sussex, UK, 407p.

Met Office (2010 - 2015), *Cartopy: a cartographic python library with a Matplotlib interface*. Exeter, Devon. URL: `https://scitools.org.uk/cartopy`.

Meyer, V. K., H. Höller, and H. D. Betz (2013), Automated thunderstorm tracking: utilization of three-dimensional lightning and radar data. *Atmospheric Chemistry and Physics* **13**(10), 5137–5150. doi:10.5194/acp-13-5137-2013.

Morel, C. and S. Senesi (2002a), A climatology of mesoscale convective systems over Europe using satellite infrared imagery. I: Methodology. *Quarterly Journal of the Royal Meteorological Society* **128**(584), 1953–1971. doi:10.1256/003590002320603485.

Morel, C. and S. Senesi (2002b), A climatology of mesoscale convective systems over Europe using satellite infrared imagery. II: Characteristics of European mesoscale convective systems. *Quarterly Journal of the Royal Meteorological Society* **128**(584), 1973–1995. doi:10.1256/003590002320603494.

Morgan, G. M. J. (1973), A General Desription of the Hail Problem in the Po Valley of Northern Italy. *Journal of Applied Meteorology and Climatology* **12**(2), 338–353. doi:10.1175/1520-0450(1973)012<0338:AGDOTH>2.0.CO;2.

Otero, N., J. Sillmann, and T. Butler (2018), Assessment of an extended version of the Jenkinson–Collison classification on CMIP5 models over Europe. *Climate Dynamics* **50**(5-6), 1559–1579. doi:10.1007/s00382-017-3705-y.

Pulkkinen, S., D. Nerini, A. A. Pérez Hortal, C. Velasco-Forero, A. Seed, U. Germann, and L. Foresti (2019), Pysteps: an open-source Python library for probabilistic precipitation nowcasting (v1.0). *Geoscientific Model Development* **12**(10), 4185–4219. doi:10.5194/gmd-12-4185-2019.

PySteps developers (2018-2024), pySTEPS - The nowcasting initiative. URL: `https://pysteps.readthedocs.io/en/stable/index.html`. Last access: 8 August 2024.

Schulz, W., G. Diendorfer, S. Pedeboy, and D. R. Poelman (2016), The European lightning location system EUCLID – Part 1: Performance analysis and validation. *Natural Hazards and Earth System Sciences* **16**(2), 595–605. doi:10.5194/nhess-16-595-2016.

Steinacker, R., M. Dorninger, F. Wölfelmaier, and T. Krennert (2000), Automatic Tracking of Convective Cells and Cell Complexes from Lightning and Radar Data. *Meteorology and Atmospheric Physics* **72**(2-4), 101–110. doi:10.1007/s007030050009.

Taszarek, M., J. Allen, T. Púcik, P. Groenemeijer, B. Czernecki, L. Kolendowicz, K. Lagouvardos, V. Kotroni, and W. Schulz (2019), A Climatology of Thunderstorms across Europe from a Synthesis of Multiple Data Sources. *Journal of Climate* **32**(6), 1813–1837. doi:10.1175/JCLI-D-18-0372.1.

van der Walt, S., J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Gouillart, T. Yu, and the scikit-image contributors (2014), scikit-image: image processing in Python. *PeerJ* **2**, e453. doi:10.7717/peerj.453.

Yang, Y. H. and P. King (2010), Investigating the Potential of Using Radar Echo Reflectivity to Nowcast Cloud-to-Ground Lightning Initiation over Southern Ontario. *Weather and Forecasting* **25**(4), 1235–1248. doi:10.1175/2010WAF2222387.1.

Zinner, T., H. Mannstein, and A. Tafferner (2008), Cb-TRAM: Tracking and monitoring severe convection from onset over rapid development to mature phase using multi-channel Meteosat-8 SEVIRI data. *Meteorology and Atmospheric Physics* **101**(3-4), 191–210. doi:10.1007/s00703-008-0290-y.

A. Abbreviations and Symbols

List of abbreviations:

A	Anticyclonic synoptic flow type
ACG	Austro Control GmbH
ALDIS	Austrian Lightning Detection and Information System
CAPE	Convective Available Potential Energy
C	Cyclonic synoptic flow type
CEST	Central European Standard Time
CG	Cloud-to-Ground lightning
CI	Convective Initiation
CIN	Convective Inhibition
DMC	Deep Moist Convection
ECMWF	European Centre for Medium-Range Weather Forecasts
ERA5	Fifth generation ECMWF reanalysis for the global climate and weather
EUCLID	European Cooperation for Lightning Detection
IC	Intra-Cloud lightning
LCL	Lifting Condensation Level
LF	Low Flow synoptic flow type
LFC	Level of Free Convection
LNB	Level of Neutral Buoyancy
LST	Local Standard Time
ltng	Token for lightning cells
MCS	Mesoscale Convective System
mixed	Token for mixed cells
netCDF	Network Common Data Form (file format)
rad	Token for radar cells
pySTEPS	Python framework for Short-Term Ensemble Prediction Systems
T-DaTing	Thunderstorm Detection and Tracking (implemented in pySTEPS)
UTC	Coordinated Universal Time

List of symbols:

A_{min}	T-DaTing input: minimal area of a detected object
c_p	Heat capacity of air at constant pressure ($\approx 1005\ Jkg^{-1}K^{1}$)
dBZ	Logarithmic unit for radar reflectivity Z (dB = decibel)
$dims$	Dimensions of a netCDF file
d_{min}	T-DaTing input: minimal distance between maxima
ΔFL	T-DaTing input: minimal flash density difference on path between maxima
FL_{max}	T-DaTing input: upper flash density threshold, saturation value
FL_{min}	T-DaTing input: lower flash density threshold
FL_p	T-DaTing input: minimal peak flash density

g	Earth's gravitational acceleration ($\approx 9.81\ ms^{-1}$)
g	Weighting factor for Gaussian smoothing
H	Scale height of the atmosphere ($\approx 10\ km$)
k	Factor defining the filter width for Gaussian smoothing
n	Distance from the centered grid point in multiples of the grid width
nan	Not a Number
px	Pixel
R^2	Coefficient of determination
ΔR	T-DaTing input: minimal reflectivity difference on path between maxima
R_{max}	T-DaTing input: upper reflectivity threshold, saturation value
R_{min}	T-DaTing input: lower reflectivity threshold
R_p	T-DaTing input: minimal peak reflectivity
T_e	Temperature of the environment
t_{min}	T-DaTing input: minimal length of a track (in time steps)
T_p	Temperature of a parcel of air
v_t	Terminal fall speed of rain droplets
w_0	Updraft velocity
z	Vertical distance
γ	Environmental lapse rate
Γ	Parcel lapse rate
Γd	Dry adiabatic lapse rate ($\approx 9.8\ Kkm^{-1}$)
Γ_s	Saturated adiabatic lapse rate
σ	Standard deviation
τ	Lifetime of single cells

B. Detailed T-DaTing Changes

On the following pages the changes made to the T-DaTing algorithm code are contrasted in form of tables, starting with the detection part of the algorithm tstorm.py.

Original code from pySTEPS:	**Code adapted to this work:**
Input parameters (for resolution of 1 km): minref=35, maxref=48, mindiff=6, minsize=50, minmax=41, mindis=10	Input parameters (for a resolution of 2 km): minref=34, maxref=50, mindiff=6, minsize=5, minmax=41, mindis=5
In the saturation process, connected areas of having more than 3 maxref pixels are raised to maxref+1.	Not included.
The size of the structure element for the detection of maxima varies with mindis and includes all surrounding pixels.	Only the 4 directly neighbouring pixels (not the diagonals) are used for the calculation of the path of least change.
In the longdistance-function deleting maxima lying too close to another is essentially carried out randomly.	Maxima with lower intensities are deleted, based on the average of the maximum itself and its 8 surrounding pixels (at the edges only the maximum itself is chosen).
The initiation of the watershed algorithm (to split cells) accounts for cases with no reflectivity-free areas in the domain. Watershed lines are computed, but not used.	As there are always enough empty pixels in the domains of this study, the initiation process just involves setting nan's to 0's. Watershed lines are not computed.
Not included.	The removing of too small ($<$ minsize) or too weak ($<$ minmax) cells is repeated after the cell splitting process.
In the getprofile-function no nan-setting outside of the domain boundary is applied and cell contour lines are calculated a little more narrowly (0.5).	Areas outside of the radar domain are set to nans instead of 0s and cell contour lines are a little wider (0.8).
The calculation of cell centroids is a simple averaging of cell coordinates with the result always being rounded down (which is not necessarily the nearest integer).	Centroids are based on a reflectivity weighted average being rounded to the nearest integer. The weights are taken from cell cores, the top parts of cells having a reflectivity greater than the respective maximum minus the mindiff parameter. If several cores exist for one cell, the one with the highest maxima and subsequently the largest area is chosen. If the centroids lie outside of e.g. curved cells, the nearest pixel inside the cell is chosen.

The next table includes the changes in the tracking part of the algorithm tdating.py:

Original code from pySTEPS:	**Code adapted to this work:**
The package skimage is loaded, but its only use remains unimportant as contour lines for advected cells are not stored.	The additional packages datetime and netCDF4 are loaded, while skimage is not loaded.
Not included.	New input parameters are added, for the inclusion of previous runs (max-id, track-list and last-period) as well as for the storing of the output (save, savestring, lon and lat).
The provided cell-list and label-list of previous runs are checked for same length.	The lists are checked for having length 2 (as this is the only portion needed for unbroken tracking).
Max-id is always initialised as 0, leading to ambiguities in different months. Moreover, after each time step a list of newly identified IDs is computed and max-id is treated unnecessarily complicated.	When including previous runs, max-id is provided as an input parameter, so no cells can have the same ID. Also in the initiation phase, max-id is reduced to the maximum of the labels image and does not need to be accounted for in more detail as the first 2 time steps are not stored.
In the advect-function table columns for maximum reflectivity and contour lines are added and partly filled, but they are not used anymore.	The calculation of max-ref and cont is omitted.
In the match-function a specific overlap of 40% is implemented. Matches with 0s or nans are not specially treated.	The overlap is allowed to take a variable value. 0s in the matching ID list are excluded. Nan-matches are counted and an informative print statement on their number is issued, together with taking the next largest matching ID in the further process.
The new labels image has 0s outside of the radar domain boundaries.	Regions outside of the domain boundaries are set to nan.
Not included.	A combination-function is added to reduce missed connections. Tracks ending in one time step and tracks originating in the next are searched. Advect terminating cells and match them to the new cells in the next time step with a reduced overlap criterion ($> 0\%$). If several possible matches exist, use the one with the biggest overlap. For the found matches not only the track IDs are continued, but also max-id is lowered.

In the couple-track-function always the whole cell list is searched for cells belonging to a specific track, which leads to very high computational costs for T-DaTing runs of long periods.	A procedure is implemented that makes sorting cells into tracks much quicker. First, the lowest and highest track ID in the dataset are extracted. Then the number of time steps of interest is narrowed down by remembering when the previous track started. The search is stopped as soon as the cell of interest is not found in the next time step anymore.
All tracks need to fulfill the mintrack criterion.	Tracks belonging to cells that appear in the very first and last instance of a monthly data file are kept even if they are shorter, as they could extend into the previous or next data file.
Not included.	A transition-function is added to account for cells "living" across 2 monthly data files. First, the IDs of such cells are extracted and the corresponding tracks are searched for in the old and new track-lists. The track parts in the new track-list are deleted and appended to the corresponding track parts in the old track-list (this ensures that the track can be found in the results file named after the month of its initiation). Lastly, another check of the mintrack criterion is carried out after connecting related cells.
Not included.	A save-results-function is added that saves the cell-list, the old and the new track-list into pickle files, the label images into a netCDF file (with the dims time, x and y), and the most important track data into another netCDF file (with the dims number of track and track length). The latter includes the centroid, the timestamp, the maximum reflectivity/flash density and the area of each cell.

Apart from these changes in the detection and tracking parts of T-DaTing, also a new code block on the design of mixed cells was added in tdating.py. In principal, it works as follows (a less technical and therefore more comprehensive description can also be found in Chap.3.4):

Code created for this work:
The detection of radar and lightning cells remains unchanged and is carried out by tstorm.py. Alternatively, this information can also be loaded from previous output files.
In the cell-lists of each data type a column is added and then filled with IDs of overlapping cells of the other data type. Contrary to above more than one overlap is allowed, only 0s and nans are excluded from the analysis.
Then an empty labels image for mixed cells is created, having nans outside of the common domain. Filling this labels image with mixed cell IDs first requires an iterative procedure, jumping back and forth the connection between overlapping radar and lightning cells until no new radar and lightning IDs can be added to the cell cluster. All pixels of this cluster are then marked in the labels image with unique mixed ID.
For all mixed cell clusters in a time step the respective minsize and minref criteria are checked to ensure tracking only strong enough cells.
Also an empty cell-list is created, similar to the ones introduced in the get-profile-function of tstorm.py. The only differences are another type of weighted centroid calculation, and 2 maximum intensity columns instead of 1, for reflectivity and flash density.
From here on the standard T-DaTing algorithm may be followed also for mixed cells.