

We thank the two reviewers very much for the time taken to review our manuscript and for the helpful and constructive comments. We are pleased that the reviewers approved the manuscript to be well written with a solid scientific argumentation. In preparing a revised manuscript, we have considered all major and minor suggestions listed in the reviews.

Reviewer #1

In their study, Kunz et al look at the ambient conditions during hail storm events in Central Europe. Different data sources are included in the analysis: Radar composites are used to track storm cells; the ESWD data base provides the hail diameter during storms; and the ERA-Interim reanalysis is used to characterize the environmental conditions at hail storm (HS) locations. The manuscript is well written, the scientific argumentation solid, and the balance of text and figures well motivated. Still, some aspects of the study should be addressed before it can be considered suitable for publication in NHESS. The major points are listed below (Major Concerns), some minor comments then follow in 'Minor Comments'.

Major Concerns

1. Resolution of ERA-Interim: Severe convective storms have a small spatial scale and can be triggered by small-scale topographic or other environmental features. How does this fit with the rather coarse spatial and temporal resolution of the ERA-Interim dataset? Of course, I notice that the authors are aware of this issue and they mention it at several places in the manuscript. Still, there remains the uncertainty whether the DLS (deep-level shear), SRH (storm-relative helicity) and SLI (surface lifted index) in ERA-Interim are really representative for the exact location where hail is recorded? I wonder whether - the authors should show some results of the high-resolution analysis data set that is mentioned in section 5.1, which however is not shown in the paper. - a limited time period (1 year?) could be chosen and the analysis of the manuscript be repeated based on, e.g., high-resolution ECMWF analysis data. I certainly don't expect the authors to repeat the whole analysis with higher-resolution data; but it would be nice to see how well the environmental conditions are really represented in ERA-Interim. I can imagine, for instance, that ERA-Interim does not capture the environmental conditions very well over the Massif Central, or over the Black Forest. A critical discussion, optimally complemented by a reduced analysis at higher resolution, should be included in the manuscript.

We followed this suggestion and included a new subsection (5.3 Effects of model resolution on convective parameters) where we discuss exemplarily two quantities – lapse rate and DLS – for CoastDat-3, a high-resolution reanalysis (~10 km, 1 hour resolutions) dynamically downscaled from ERA-Interim. The results, presented in the new Figure 11, show a larger variability of the fields, whereas the magnitude and the spatial distribution are almost unchanged.

Note that reanalysis is used only to estimate the convective environment, which has a much lower spatial variability compared to SCS probability. Storm's features are quantified from the tracking algorithm (also SRH). Triggering mechanisms are not considered as these are not reliably modelled even with high-resolution NWP models. (In the meantime we have performed 7-years hindcast runs using COSMO-EU and COSMO-DE for Europe and Germany with a resolution of 7 and 2.8 km, respectively. The results will be shown in a next paper.)

2. Identification of fronts: On P5,L118-120 some details about the front identification are given. The methodology to do so is well established and fine also for this study. Nevertheless, some aspects of the identification might be worthwhile to carefully address in the manuscript: (i) First, only cold fronts are considered, i.e. any convective storm associated with a warm front is not included? Is there a specific reason why the analysis is restricted to cold fronts? (ii) Stationary or slowly moving

fronts (< 3 m/s) are not included in the attribution, based on the (fair) assumption that synoptic fronts associated with extratropical cyclones typically move with reasonable speed? Still, the application of this criterion also might be a little artificial. (iii) Thermal fronts, or fronts associated with land-sea contrasts, are explicitly excluded. I guess that there are good reasons why to do that, but they should be more clearly discussed in the paper. (iv) Finally, pure geometric conditions are used (e.g., minimum length of 500 km) to separate 'significant' fronts from more 'fragmented' ones. I wonder, however, whether this also brings in some bias, because I can imagine that frontal fragmentation can be favored as the 'pure' and elongated synoptic fronts originating in the North Atlantic move in over Europe and move land inward. Frontal fragmentation might also be enhanced over orography, e.g. the Massif Central or the Black Forest. I am aware that objective identification of fronts is a difficult matter. But at least I would appreciate if the authors carefully address these limiting factors in the discussion.

We have expanded Section 2.3 Cold front detection and included some additional explanations in the results section. (i) Warm-fronts may be relevant for setting of the preconvective conditions. But due to the slow ascending warm air in combination with warm-air advection aloft, they are not relevant triggers for SCS. For this reason, we concentrate exclusively on cold fronts. We have added a comment at the beginning of Section 4. (ii-iii) Sea-breeze fronts and/or thermal boundaries, e.g. from Alpine pumping, are related and restricted to specific terrain characteristics. When considering a larger area as in our case, it does not make sense to mix the various types of fronts. We have added a comment (also about the thresholds) in the Section 2 Data and Method and included a short paragraph at the beginning of the Section 4 (SCS and fronts; L274-280). (iv) The reviewer is right that our approach, in particular the fragmentation, brings in some bias. However, the criteria are necessary to filter out thermal contrasts that are not related to fronts with their lifting and cross-circulation. This issue is now briefly discussed in the related paragraph of Section 6 Discussion. (We think about a follow-up paper where we will consider all the above suggestions and separate the fronts according to different criteria/characteristics).

3. SCS vs HS; frontal vs. non-frontal: The focus of this study clearly is on the HS, whereby these HS are classified according to their hail diameter based on the ESWD reports. That's fine, but brings in also some subjectivity. More specifically, how reliable/robust is the analysis if it is based on the maximum recorded hail size? Furthermore, Figure 1 also shows that the number of HS is small compared to the objectively identified SCS. Of course, I see that exactly these HS are at the heart of the manuscript's analysis. But, at some stage while reading the manuscript I 'defined' what I considered for myself as the most interesting part of the study. And, honestly, it is not the categorization according to the hail size and the track length. To me, the most interesting aspect is the distinction between frontal and nonfrontal storms! I then wondered whether this would not be a better storyline than the existing one. Basically, the story would start with a climatology of the SCS over Central Europe; then, the next question addresses the fact whether the frontal tracks are associated with front or not. If this attribution is done, a next section could deal with the different characterization of the frontal vs the non-frontal storms. Here, the track length, the track duration, and also the presence of hail or not (including its size) could be discussed. Of course, I know that this storyline kind of reverses what the authors do now. One advantage of restructuring would be that the weakest part -- at least in my opinion --, i.e. the hail diameter, does not stand at the beginning, but only appears at the end of the chain. I would certainly not enforce the authors to reorder everything in the manuscript, but they might give it a thought.

(1) We considered not the mean, but the maximum hail diameter as this is usually reported in the ESWD; added in Section 2.1. (2) Unfortunately not all hail events are reported in the ESWD (and not all SCS tracks are associated with large hail); reasons for this are a significant underreporting (particularly in France), but also small hail with diameters less than 2 cm that are not included in the

ESWD. We have added an explanation at the beginning of Section 3 (L233-239). (3) We focused on hailstorms and their properties because to the best of our knowledge no study so far has investigated ambient conditions for different classes of hail size and track length in Europe. Our knowledge about hailstorms including their frequency and characteristics is still very limited. Note that in Germany but also in several other European regions/countries hail causes large economic losses and the highest share of insurance losses of all natural hazards. Most relevant for the damage potential of hailstorms are hailstone sizes and track length. (4) The idea of rearranging the manuscript is appealing if one focuses on the fronts. We have discussed this extensively with some of the co-authors. But as the focus should be still on hail for the reasons listed above, we have decided to keep the previous order.

4. Application in forecasting: This final major concern is a more 'philosophical' one. The basic aim of the study is to identify the severe and damaging storms, whereby the severity is (obviously) a combination of hail diameter and storm track length (footprint). These storms are taken as the basis of the analysis, and then the environmental conditions of these storms are considered. From a forecasting perspective, the reverse questions is much more interesting. In short, if we have an environment with this and that environmental conditions, what is the chance that we get a storm with a severe footprint? Please note that this reversed perspective is not exactly the same as the one studied in the manuscript, and this should be carefully discussed. Stated in a more practical way: How can the results of this study be used in an operational weather forecast? How would be the performance of such a forecasting algorithm if it were based on the environmental parameters used in this study? I think that the authors do a first step in this direction by quantitatively looking in P13,L378-383 at the predictive skill of certain parameter combinations. This part of the text should, however, be more detailed. The remained rather unclear to me how the predictive skill is determined...

This is a very interesting point the reviewer raised and highly relevant for forecasting / nowcasting. We are currently working on that topic in another project together with the German Weather Service DWD, where we include a subset of environmental conditions in the nowcasting algorithm to get an estimate about the expected future life time of the convective cells (a publication will follow). In case of hail events, however, it's not possible to reliably quantify the probability of non-events because of a lack of comprehensive hail observations (see comments about the underreporting). We have added a paragraph in the Conclusion Section where we discuss this issue. After discussion with the co-authors, we have decided to delete L378-383 about predictive skill because it gives not more information (except of skill score values) compared to the paragraph above.

Minor Comments

P2,L24: large hail -> Does this refer to large hail stones, or does it refer to intense hail? It is not completely clear.

This refers to large hailstones; we have changed the sentence accordingly: "...large hail with a diameter of at least 2 cm..."

P2,L30: What exactly is meant by footprint?

The term footprint refers to the hail swath; we have added an explanation: "hail swath (envelope encompassing all hail streaks; footprint)"

P2,L56: What is the temporal resolution of the cell tracking algorithm, more precisely: what is the resolution of the input fields?

The input fields (radar data) have a temporal resolution of 5 min (see L138).

P2,L68: Does the propagation direction of hailstorms depend on hail size? → This question comes a little out of the blue, i.e. it was not motivated by the preceding paragraphs of the introduction. Is there a good reason to assume that the storm movement's direction depends on the hail size? If so, discuss this in the introduction.

We agree and deleted this question.

P4,L103: The SRH is calculated based on the storms movement (c_x, c_y) , and based on the ERAInterim

It may seem strange to mix up the input parameter. However, the tracking algorithm directly compute the motion vector \vec{c} for each of the detected cells, which are much more reliable compared to indirect methods such as that proposed by Bunkers et al. (2000). This is the reason why we combined the two input parameters. We have modify this paragraph (new L115-121) in order to clarify this.

P5,L141: Several studies have provided evidence that suggests this threshold to provide sufficient skill for hail detection → Rephrase in clearer way.

We have included a new sentence: "Several studies have provided evidence that this lower threshold is suitable to identify hail in radar data."

P5,L148: Here, the SCS tracks are first introduced. Possibly, it would be nice to learn already at this place about the typical characteristics of these tracks? For instance, about their length and duration? This is, in my view, reasonable because it allows already to set this characteristics to set into perspective of the limited spatial and temporal resolution of the ERA-Interim dataset.

To avoid any misinterpretation at the beginning of the manuscript, we have removed all results from the Data and Methods. Instead, we have included a new paragraph at the beginning of Section 3 that briefly summarizes the results of the data combination (L233-239).

P6,L156: A minor detail: Wouldn't it be easier to take the spherical distance between the frontal points and the storm points instead of this (approximate) formula? It is approximate because the latitude actually changes along the storm track and along the front.

That's almost the same. Changes in the latitude are considered by the cos term.

P7,L211: "The frequency of both SCS and HS events shows a large spatial variability with an overall increase from north-to-south, i.e., with the distance to the Atlantic" → I Am not sure whether I see this effect as clearly as it is stated in the text. There is a distance-to-Atlantic effect near the coast, but further land inward it is less clear to me... P7,L212: From Figure 1 it is not clear that the regional hot spots are downstream of the Massif Central or the Black Forest. It is not clear because we don't see in this figure the direction of the wind.

We have reformulated the whole paragraph (now L241-249).

P9,L250: not the individual fronts are done here → Sounds a little strange! It becomes clear with the succeeding sentence.

We have deleted this part because it's not necessary.

P10,L287: fronts create hail-conducive conditions mainly through two effects: along-front advection of moisture at lower levels leading to larger CAPE, and higher wind speed aloft enhancing vertical wind shear → What about the propagation of the front itself? This will also induce vertical lifting, as does also the ageostrophic circulation associated with fronts? I would also assume that these effects depend on the intensity of the fronts. In short, a more comprehensive discussion of the frontal effects would be nice -- as is actually done in the later discussion/conclusion sections of the paper.

We have expanded the discussion of the fronts in the discussion section and deleted the paragraph about the frontal modification of the environment in Sect. 4.2

P10,L291: fronts also largely determine the orientation of the tracks → Sounds a little strange to me! Rephrase, e.g., to: the tracks are strongly coupled to the (typically eastward) propagation of the fronts.

We changed this sentence accordingly. However, we have modified and moved the whole paragraph to the discussion section (now L526-542).

P12,L389: Based on the ERA-Interim winds, right-moving and left-moving supercells are distinguished. Can this reliably be done based on the coarse ERA-Interim winds? Note that the ERA-Interim winds are representative for approximately 80 x 80 km². The storm itself is, on the other hand, much smaller.

Because of the storm's small-scale dimensions, we quantified the shift vector of the convective cells from the radar tracking algorithm. The wind field in 500 hPa, on the other hand, is modified only marginally by local-scale mechanisms but rather the result of the synoptic-scale setting of the pressure systems, which is sufficiently resolved by ERA-Interim. We have add a comment (L455-458).

P11,L314: Considering the magnitude of DLS, it is found that the values are quite low with a mean of 12.5m/s around the HS events. Several authors have shown that organized convection capable of producing larger hail develops only in sheared environments above around 10m/s →Just for curiosity, myself not being an expert on storm dynamics. I would expect that there is an intermediate range of DLS that favors an intense and sustained convective activity. Too weak, but also too strong wind shear would hinder the convective storm development. It would be interesting to discuss this aspect somewhat in the manuscript, if my naive thought is correct.

Interesting. In the literature there is no upper limit reported. The findings of other project of my group do not support the hypothesis of such an intermediate range of DLS. As we haven't investigated this issue in our paper, we won't discuss this here. -- But I'll keep that in mind for future research.

Reviewer #2

The manuscript is well-written, correctly structured and methodologically sound. It lies within the scope of Natural Hazards and Earth System Sciences, as it considers the detection and modelling of a natural hazard (hail), with severe socioeconomic impacts. The authors present an empirical study with clear objectives, which involve the association of Severe Convective Storms (SCS) to fronts, the differences in the characteristics of frontal SCS differ to the non-frontal ones, the effect of environmental conditions to hail diameter and track lengths and whether the propagation direction of hailstorms is associated to hail size. Overall, the manuscript is of high quality and should be considered for publication in Natural Hazards and Earth System Sciences with few minor suggestions for revision presented below.

Figure 1: Consider splitting the Figure to two, as in Figure 4. In this manner, the spatial patterns in Germany will be easier to detect and present (color scale).

The reviewer is right that the spatial pattern in Germany is difficult to detect. We have modified the color scale, but used the same scale for the entire area to allow direct comparison.

L298: Seeing Figure 5, one might think 50-150 km class should be used instead. Is there another reason for choosing a 50-100 km class?

I fully agree, but the thresholds were set so that each sample contains at least 35 events. The classes as proposed would result in too small sample sizes for the long tracks (27 / 29 / 35 events for $D < 3$ / 3-4 / ≥ 5 cm instead of 64 / 72 / 50 events). We have included a comment at the beginning of Section 5: "When defining the threshold values, it was taken into account that each class contains at least 50 events – except of the class $L = 50\text{--}100$ km and $D \geq 5$ cm (Table 1). Using other thresholds, for example, 150 km instead of 100 km as suggested by the diameter–length relation shown in the boxplot (Figure 5) would result in too small sample sizes with less than 30 events."

L353-354: It would be less confusing to define the classes as $D < 3$ cm, $D < 75$ km and $D > 3$ cm, $D > 75$ km
Changed as suggested (L436-437)

L356, L370: Figures 10 and 11 are the other way around.

Thanks (this was a Latex compiler error)

L365-382: Perhaps consider removing this section. It does not contribute much to the findings and overall discussion and could be misleading. Statistical evaluation needs to be addressed also in terms of uncertainty analysis, which is not quantitatively presented in the manuscript.

The reviewer is right, we did not consider uncertainty analysis, mainly because of the unreported HS events that are unknown. We have therefore deleted the last paragraph of the statistical evaluation.

Ambient conditions prevailing during hail events in central Europe

Michael Kunz^{1,2}, Jan Wandel¹, Elody Fluck^{1,*}, Sven Baumstark^{1,+}, Susanna Mohr^{1,2}, and Sebastian Schemm³

¹Institute of Meteorology and Climate Research (IMK), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

²Center for Disaster Management and Risk Reduction Technology, KIT, Germany

³Institute for Atmospheric and Climate Science, ETH Zürich, Zürich, Switzerland

* now at: Department of Earth and Planetary Sciences, Weizmann Institute of Science, Rehovot, Israel

+ now at: Heine + Jud, Stuttgart, Germany

Correspondence: Michael Kunz (michael.kunz@kit.edu)

Abstract. Around 26,000 severe convective storm tracks between 2005 and 2014 have been estimated from 2D radar reflectivity for parts of Europe, including Germany, France, Belgium, and Luxembourg. This event-set was further combined with eyewitness reports, environmental conditions and synoptic-scale fronts based on ERA-Interim reanalysis. Our analyses reveal that on average about a quarter of all severe thunderstorms in the investigation area were associated with a front. Over complex
5 terrains, such as in southern Germany, the proportion of frontal convective storms is around 10–15%, while over flat terrain half of the events require a front to trigger convection.

Frontal storm tracks associated with hail on average produce larger hailstones and have a longer track. These events usually develop in a high-shear environment. Using composites of environmental conditions centered around the hailstorm tracks, we found that dynamical proxies such as deep-layer shear or storm-relative helicity become important when separating hail
10 diameters and, in particular, their lengths; 0-3 km helicity as a dynamical proxy performs better compared to wind shear for the separation. In contrast, thermodynamical proxies such as Lifted Index or lapse rate show only small differences between the different intensity classes.

1 Introduction

Severe convective storms (SCS) are responsible for almost one-third of total damage by natural hazards in Germany and central
15 Europe (MunichRe, 2020). Examples of recent major loss events include the two supercells on 27/28 July 2013 related to the depression *Andreas* with economic losses of EUR 3.6 billion mainly due to large hail (Kunz et al., 2018) or storm clusters during *Ela* on 8-10 July 2014 with economic losses of EUR 2.6 billion (SwissRe, 2015) caused by both large hail and severe wind gusts (Mathias et al., 2017). Given the major damage associated with SCS, particularly due to large hail, there is a considerable and increasing need to better understand the local probability of SCS, their intensity and relation to prevailing
20 atmospheric precursors.

Several authors have attempted to establish relations between SCS/hailstorms and favorable atmospheric environments (for Europe: Manzato, 2005; Groenemeijer and van Delden, 2007; Kunz, 2007; Sánchez et al., 2009; Mohr and Kunz, 2013; Púčik et al., 2015; Sánchez et al., 2017; Madonna et al., 2018, among others). Hail-conductive environments have been estimated

either from proximity soundings or from model data/reanalysis, both available over several decades and, depending on the spatial resolution, on a regional, continental, or global scale. According to Púčik et al. (2015), for example, **large hail with a diameter of at least 2 cm** most likely form in environments with high values of CAPE and bulk wind shear. While the former is directly related to the intensity of the updraft, the latter is decisive for the organization's form of the convective systems – single cells, multicells, supercells, mesoscale convective systems (MCS; Markowski and Richardson, 2010). In addition, several studies have suggested that SCS preferentially occur during specific weather regimes, such as European **or Scandinavian** blocking or teleconnection patterns (Aran et al., 2011; García-Ortega et al., 2011; Kapsch et al., 2012; Piper et al., 2019; Mohr et al., 2019). However, to date, no study has investigated environmental conditions according to hailstone size and **hail swath (envelope encompassing all hail streaks; footprint)**, despite their relevance to overall storm damage.

Forecast experience has shown that synoptic fronts, particularly cold fronts during the summer months, can significantly modify the convective environment, primarily due to increasing convective available energy (CAPE) and decreasing convective inhibition (CIN) in combination with cross-frontal circulations leading to lifting and enhanced vertical wind shear. By combining hailstorm tracks determined from radar data over Switzerland between 2002 and 2013 with front detections (Schemm et al., 2015) based on the Consortium for Small Scale Modelling (COSMO) analysis, Schemm et al. (2016) found that up to 45% of storms in northeast and southern Switzerland were associated with a cold front. They concluded that mainly wind-sheared environments created by the fronts provide favorable conditions for hailstorms in the absence of topographic forcing.

Difficulties in analyzing environmental conditions prior to or during hailstorms usually arise from insufficient direct hail observations that may serve as ground truth. The number of ground weather stations is too small to reliably detect all SCS. High-density hailpad networks exist in only a few regions across Europe (e.g., Merino et al., 2014; Hermida et al., 2015) and therefore cannot be used to reproduce entire hailstorm footprints. In order to compensate for this monitoring gap, remote sensing instruments such as satellite (Bedka, 2011; Punge et al., 2017; Ni et al., 2017; Mroz et al., 2017), lightning (Chronis et al., 2015; Wapler, 2017), or radar (Holleman et al., 2000; Puskeiler et al., 2016; Nisi et al., 2018) due to their area-wide observability are used to estimate the frequency and intensity of SCS. In particular, weather radars can give some indications of hail occurrence using either radar reflectivity above a certain threshold (e.g., Mason, 1971; Hohl et al., 2002) or at specific elevations in combination with different height specifications (melting level, -20° C environmental temperature, top of the storm cell; Waldvogel et al., 1979; Smart and Alberty, 1985; Witt et al., 1998). While observations by dual-polarization radars offer better predictions for hail (e.g., Heinselman and Ryzhkov, 2006; Ryzhkov et al., 2013; Ryzhkov and Zrnicek, 2019) these systems have been installed in Europe only recently and cannot be used for climatological studies.

Another important data source for hail is severe weather reports from trained storm spotters or eye-witnesses that are pooled into severe weather archives such as the European Severe Weather Database (ESWD; Dotzek et al., 2009). Although reporting is selective and biased towards population density and available spotters, these reports provide valuable information about the intensity of the various convective phenomena associated with SCS such as maximum hail diameter. The combination of these reports with storm tracks estimated from radar observations allows us to reconstruct entire footprints of SCS and/or hailstorms.

In our study, we have reconstructed SCS tracks from two-dimensional (2D) radar reflectivity using a cell tracking algorithm during a 10-year period (2005–2014) over central Europe including France, Germany, Belgium, and Luxembourg. **As our**

focus in on SCS, we considered only tracks above a reflectivity of $Z \geq 55$ dBZ, a threshold frequently used as hail
60 criterion (e.g., Holleman et al., 2000; Hohl et al., 2002; Kunz and Kugel, 2015; Puskeiler et al., 2016). In order to
include additional information on the maximum hail diameter of the SCS, a sub-sample of hailstorms (HS) was created
by combining the radar-derived SCS tracks with ESWD hail reports.

Afterward, we investigate characteristics and environmental conditions at the time and location of the events unfolding
for different classes of hail diameter, track lengths (lifetime) and the relationship with synoptic-scale fronts. Environmental
65 conditions are assessed by constructing composites of meteorological fields from ERA-Interim reanalysis centered around the
location of a single storm. **To estimate the effects of subgrid-scale spatial variations on environmental conditions, for
example by disturbances induced by orographic features or by temperature and moisture advection, we additionally
used coastDat-3 reanalysis with a resolution about six times higher compared to ERA-interim.**

The main scientific questions of our study are the following:

- 70 – How frequent are SCS associated with a front?
- Do the characteristics of SCS **associated with a synoptic cold front** differ from those without a front?
- How do the environmental conditions in terms of thermodynamical and dynamical parameters differ between hail diam-
eter classes, track lengths, and frontal/non-frontal events?
- **How does a higher model resolution affect the environmental conditions around the SCS?**
- 75 – ~~Does the propagation direction of hailstorms depend on hail size?~~

The paper is structured as follows: Section 2 introduces the datasets and methods used. Section 3 deals with the frequency of
SCS and HS and Section 4 examines the role of synoptic cold fronts and convective storms. Section 5 statistically investigates
environmental conditions prevailing around the storms for different classes of hail size and track length. Section 6 synthesizes
and summarizes the major findings, while the most important conclusions are drawn in Section 7.

80 2 Data and Methods

The investigation area is central Europe, including Germany, France, Belgium, and Luxembourg, from 2005 to 2014, where
data were available. Since SCS/HS in Europe occur mainly in the summer half-year (SHY; Berthet et al., 2011; Punge and
Kunz, 2016; Púčik et al., 2019), all analyses refer to the period from April to September.

2.1 ESWD hail reports

85 The ESWD, managed and maintained by the European Severe Storms Laboratory (ESSL), is the only multinational database
and by far the largest archive of hail reports in Europe. Quality-checked reports of SCS and related phenomena originate from
storm chasers and trained spotters, sometimes supplemented by newspaper reports. In our study, we consider the reported

maximum hail diameters of all quality levels (70.4% of all reports were confirmed, 29.0% were at least plausibility checked). This includes both large hail with a diameter of at least 2 cm usually given in increments of 1 cm (in rare cases of 0.5 cm) and hail layers with a depth of at least 10 cm, regardless of hail diameter. In those cases, and when a hail size is not specified (usually in case of small hail), the diameter is set to 1 cm.

During the 10-year investigation period, a total of 4,577 reports of severe hail in the study area is available. Most reports stem from Germany (76.5%), followed by France (21.1%), Belgium (1.7%), and Luxembourg (0.7%). This distribution does not reflect the occurrence probability of SCS but is primarily due to ESSL originally being a German initiative.

Because of the large spatial extent of the study area in a west-east direction, we converted the time stamps for the daily cycle analysis (only for that; cf. Fig. 2) from UTC into Local time (LT) by adding $\Delta t = 24\text{h}/360^\circ\text{lat} = 4 \text{ min per degree}$ starting from 0° lat .

2.2 Reanalyses

Atmospheric conditions prevailing over a larger area around the SCS tracks are studied using ERA-Interim (Dee et al., 2011) reanalysis from the European Center for Medium-Range Forecast (ECMWF). This data set, which was also used for the detection of synoptic cold fronts (see Sect. 2.3), is represented as spherical harmonics at a T255 spectral resolution (approx. 80 km) on 60 vertical levels from the surface up to 0.1 hPa with a temporal resolution of 6 hours. **In order to estimate the effects of the model resolution on the dynamic and thermodynamic environmental conditions, we additionally used high-resolution coastDat-3 reanalysis data for selected variables. This second reanalysis from the Helmholtz Zentrum Geestach (HZG) has a spatial and temporal resolution of 0.11° (approx. 10 km) and 1 hour, respectively. It was produced by dynamically downscaling ERA-Interim using COSMO in climate mode (CCLM; Rockel et al., 2008).**

Mesoscale environments of the hailstorm tracks are characterized by severe storm predictors representing both thermodynamical and dynamical conditions. We tested and applied several convection-related parameters, but focus here only on those proxies with the highest prediction skill: Surface Lifted Index (SLI) representing latent instability (Galway, 1956), Lapse-Rate (LR) as the temperature difference between 700 and 500 hPa representing potential instability (only for coastDat-3), Deep-Layer Shear (DLS) as the difference of the wind vectors between 500 hPa and the surface, and 0–3 km Storm-Relative Helicity (SRH) quantified by:

$$\text{SRH} = \int (\mathbf{v}_h - \mathbf{c}) \cdot (\nabla \times \mathbf{v}_h) dz \quad (1)$$

$$= \int \left[-(u - c_x) \left(\frac{\partial v}{\partial z} \right) + (v - c_y) \left(\frac{\partial u}{\partial z} \right) \right] dz \quad (2)$$

where $\mathbf{v}_h = (u, v)$ is the horizontal wind vector and $\mathbf{c} = (c_x, c_y)$ the (constant) cell motion vector, which is usually estimated from a semi-empirical relation such as that from Bunkers et al. (2000). As the convective cell tracking algorithm directly computes \mathbf{c} for each SCS/HS event (see next Sect. 2.4), we used these values to quantify SRH in addition to the vertical wind shear provided by ERA-Interim. Helicity is a measure of the degree to which the direction of motion is aligned with the (horizontal) vorticity of the environment $\omega_h = \nabla \times \mathbf{v}_h$ (Markowski and Richardson, 2010). Only stream-

120 **wise vorticity, which is a prerequisite for supercells bearing the largest hailstones, contributes to SRH (Thompson et al., 2007).**

2.3 Cold front detection

Synoptic-scale cold fronts are detected in ERA-Interim based on the method outlined in Schemm et al. (2015), which is briefly summarized here. To identify and locate fronts in the reanalysis, we used the thermal front parameter (TFP; Renard and Clarke, 1965; Hewson, 1998) defined as

$$\text{TFP} = -\nabla|\nabla\theta_e| \cdot \frac{\nabla\theta_e}{|\nabla\theta_e|}, \quad (3)$$

where θ_e denotes the equivalent potential temperature at 850 hPa, a widely-used choice in the forecasting community, **which also neglects sea-breeze fronts**. The first term in Eq. 3 represents the gradient of the frontal zone ($|\nabla\theta_e|$), which must be higher than $4 \text{ K (100 km)}^{-1}$. The second term is the unit vector of the θ_e gradient. The TFP hence captures changes of the gradient of the frontal zone along the gradient itself. The frontal zone is strongest where $\text{TFP} = 0$, and its leading-edge is where $\text{TFP} = \text{Max}$. **For the detection of propagating synoptic fronts, which are in the focus here because of their relevance for convection triggering, we require all fronts to have a length of at least 500 km and a minimum advection speed of 3 m s^{-1} . These two thresholds may seem somewhat artificial or arbitrary. But as shown by Schemm et al. (2015), their implementation sufficiently remove land-sea contrast and thermal boundaries from Alpine pumping from the dataset and limit the data to fronts typically associated with extratropical cyclones.**

2.4 Radar data and storm tracking

Tracks of SCS are identified from two-dimensional (2D) radar reflectivity based on the precipitation scan at low elevation angles. Radar data with a spatial and temporal resolution of 1 km and 5 min, respectively, were provided by Meteo France and by German Weather Service (DWD) as entire radar composites. Whereas all 17 German radars operate in C-band, 19 radars in France are C-band and five each are S-band and X-band. The area in France covered by the S-band radars is rather small ($< 5\%$ of the total area) compared to that captured by C-band, and mainly restricted to the southwest (S-band radars at Opoul, Nîmes, Bollene, and Collobrieres). Because of the dominance of C-band radars, we did not distinguish between the two radar types. X-band radars, exclusively operating in the Maritime Alps in southeast France, are not considered due to their strong attenuation of the radar signal.

145 Storm tracks were reconstructed by applying a modified version of the cell tracking algorithm TRACE3D originally designed for 3D reflectivity in spherical coordinates (Handwerker, 2002). Thus, TRACE3D has to be modified to rely on 2D radar reflectivity in Cartesian coordinates (Fluck, 2017). The tracking algorithm first identifies all convective cells (reflectivity core, RC) embedded into larger “regions of intense precipitation” (ROIP; Handwerker, 2002). Afterward, the weighted center (barycenter) of all RCs is tracked spatially over subsequent time intervals dt by establishing a temporal connection between 150 the detected RCs. For each RC, a 2D shift velocity vector \mathbf{v}_T is calculated in different ways, depending on whether and over what distance an RC has already been detected in previous scans. The new position of the RC is estimated from $\mathbf{s}_T = \mathbf{v}_T \cdot \delta t$

within a certain search radius r , which depends on the length of s_T and the distance to the closest neighboring RC. This process is repeated for all subsequent scans until the complete track of a convective cell is reconstructed. The algorithm considers different processes such as cell splitting or merging. Correction algorithms are implemented for undesired radar effects such as the bright band or anomalous propagation (so-called Anaprop). In addition, we eliminated all single grid points with high radar reflectivity, but without lightning within a radius of 10 km. This filter is based on the assumption that SCS are always accompanied by lightning. Note that the filter only eliminates single spurious signals but keeps the tracks that are composed of numerous radar grid points.

In our analyses, we considered only storm tracks above a threshold of $Z \geq 55$ dBZ, referred to as the Mason (1971) criterion for hail detection. Several studies have provided evidence that this lower threshold is suitable to identify hail in radar data (e.g., Holleman et al., 2000; Hohl et al., 2002; Kunz and Kugel, 2015; Puskeiler et al., 2016). However, high radar reflectivity does not guarantee to have hail on the ground, mainly because of potential melting hailstones and the relation $Z \sim D^6$, where D is the hail size diameter. For example, the evaluation of radar-derived cell tracks with damage data from two insurance companies by Puskeiler et al. (2016) has shown that the Mason (1971) criterion provides a satisfactory probability of detection (62 and 55%), but also a high false alarm rate (35 and 40%). This means that our SCS sample based on this criterion consists mainly of hailstorms, but also includes some heavy rain events (see Sect. 2.5.2 for the definition of the HS sample).

During the investigation period, 26,012 SCS tracks were identified.

Each SCS event, defined as an entire track reconstructed by the tracking algorithm, contains the following parameters: center (lat/lon) of the track including date and time, mean angle, width, total length and duration; the latter two quantities allowing us to compute the storm motion vectors c required for SRH (cf. Eq. 1). For further details on the tracking and the results, see the study by Fluck (2017).

2.5 Combination of SCS tracks with other parameters

2.5.1 Combination of SCS with fronts

To match the SCS tracks with synoptic front detections (cf. Sect. 2.3), we first compute the minimum horizontal distance d_i between the two events:

$$d_i = \sqrt{(a_i \cdot \cos(lat \cdot 2\pi/360) \cdot l)^2 + (b_i \cdot l)^2}, \quad (4)$$

where a_i is the longitudinal distance between a frontal grid point i and the grid points of an individual storm track, b_i is the same for the latitude, lat is the position (latitude) of the storm track, and l is the (constant) distance of 1° latitude (≈ 111.32 km). The \cos function in the equation takes into account the poleward convergence of the lines of longitude. For each front detection, we compute the distance d_i to all grid points defining the track of an SCS identified in the same 6-hour period. The minimum of all d_i , thus $d_{min} = \min(d_i)$ defines the minimum distance between the front and the related SCS.

Frontal SCS are defined as those events where a front is located within a search radius of $R = L/2 + 200$ km (L is the length of an SCS track) around the storm track, i.e. when $d_i < R$. Assuming a front to act as a potential trigger for convection, the

185 distance between the two events must be limited (Trapp, 2013). For this reason and because of the low temporal and spatial resolution of the front detections, we set the constant part of R to 200 km. Note that changing this part to a value of 300 or 400 km has no significant effect on the results. The constant part in R ($L/2$) considers only the time of the center of the SCS for the synchronization between the two events. The longer L , the larger the temporal and spatial difference between tracks and fronts can be and, thus, the larger R must be.

190 To account for temporal coincidence, we consider the timestamp of the SCS centers that must be within the period of the front detections (00, 06, 12, 18 UTC). When the SCS center is exactly between the ERA-Interim run times (03, 09, 15, 21 UTC), both time frames are used in the calculations of d_i . Since the front detections are available for six-hour intervals only, the time difference between the centers of the SCS and the fronts is at most three hours. Considering the start time of the SCS instead of that at the center has only a small marginal effect on the results because of both the low temporal resolution of the
195 reanalysis and the comparatively short duration of the SCS tracks (exponential distribution; 73% of all SCS have a duration of 2 hours and less).

2.5.2 Combination of SCS tracks with ESWD data

The SCS tracks derived from the radar composites are additionally combined with the ESWD reports to assign each track a maximum hail diameter. **This step not only ensures that the resulting subsample “hailstorms (HS)” consists of hail events**
200 **solely but also merges hailstorm tracks and maximum hail diameters.** This is done by considering both the date/time and the horizontal distance d_i between a certain track and the nearest ESWD report in the same way as described above for the fronts. Only ESWD reports with $d_{min} \leq 10$ km to the closest grid point are considered; these storms are hereafter referred to as hailstorm (HS) events or tracks. A tolerance of 10 km is necessary for two reasons: In some cases, the ESWD reports do not give an exact position; and hailstones falling to the ground may drift with the horizontal wind over distances of several
205 kilometers (Schuster et al., 2006). When an ESWD report coincides with several tracks, we further considered the time of the report if specified. Still unclear cases (around 100 events corresponding to 2% of all cases) were not considered in the event set. If more than one ESWD report is assigned to a single storm track, we considered only that with the maximum reported hail diameter.

~~The combination of SCS tracks with ESWD reports substantially reduced the sample size to 985 HS tracks. The main reasons
210 for the much low number of HS compared to SCS events are the underreporting of hail events, especially over France, and some SCS tracks accompanied by heavy rain and not by hail. Nevertheless, this sample size is still sufficient for the investigation of environmental conditions for different intensity classes.~~

For all investigations, we separated the maximum hailstone diameter into three different classes (samples): $D < 3$ cm (48,0% of all HS tracks), $3 \leq D \leq 4.5$ cm (37.0%), and $D \geq 5$ cm (15.0%).

215 2.5.3 Composite construction

The investigation of the environmental conditions around the HS tracks is based on composites of convection-related parameters from ERA-Interim. **The composites are obtained by averaging the environmental fields of moving spatial windows**

of 800 km in latitude and longitude around the center of individual HS tracks (i.e., ± 400 km to the north, south, east, and west from the center of the track). The center of the composites represents the location of all HS tracks. Similar composites have already been used by Graf et al. (2011) to investigate Central European tornado environments. The effect of latitudinal dependence on the horizontal difference between the grid points in the reanalysis is considered by transferring the latter to Cartesian coordinates with a grid resolution of approximately 50 km. As mentioned above, using the start location instead of the center does not affect the results because of the limited spatial extent of the tracks (mean lengths of frontal/non-frontal HS tracks are 56.8/96.2 km, respectively). In addition, due to the low resolution of the ERA-Interim data, it can be assumed that the convective environment is not modified by ongoing convective storms. Temporal coincidence is ensured by using the reanalysis fields with the smallest time difference to the HS events. Therefore, the largest time difference between the environmental conditions and the HS events is 3 hours.

The single ERA-Interim fields are averaged either for all events or for different categories of events related to hail diameter classes, HS track lengths, and frontal vs. non-frontal HS events. Since most of the HS events propagate from southwest to northeast (67.6% between 180° and 270°), we have not aligned the fields accordingly. Note, however, that according to a test where this was realized, the results remained essentially the same.

3 Frequency of SCS and HS

During the investigation period, 26,012 SCS tracks were identified. The combination of those tracks with ESWD reports substantially reduced the sample size to 985 HS tracks. The main reason for the much lower number of HS compared to SCS events is an underreporting of hail events, especially over France (Groenemeijer et al., 2017), where only 828 ESWD reports are available during the investigation period compared to 3,022 for Germany (note that most of the hailstorms are captured by various reports). Furthermore, an unknown part of the SCS events is accompanied only by small hail (less than 2 cm), which is not reported in the ESWD, or even just by heavy rainfall. Nevertheless, this sample size is still sufficient for the investigation of environmental conditions for different intensity classes.

3.1 Spatial distribution of SCS and HS events

The frequency of both SCS and HS events shows a rather high spatial variability, but also some larger contiguous spatial patterns. In general, their frequency is lowest near the coast and highest inland. Most pronounced is the large hotspot of SCS events southeast of the center of France near the Massif Central. Other hotspots of SCS and HS events can be found in southwest Germany between Black Forest and Swabian Jura or in the southeast near the Ore Mountains. Given a southwesterly flow direction usually predominant on hail-prone days in both France and Germany (Vinet, 2001; Kunz and Puskeiler, 2010; Piper et al., 2019), most of these hotspots are located over and downstream of the mountain ranges. Over France, SCS tracks are much more frequent compared to Germany (Fig. 1). By contrast, HS tracks are far more frequently detected over Germany due to more available reports. Nevertheless, Figure 1 suggest a relationship between the two records: regions with an increased SCS frequency also show an increased HS frequency and vice versa.

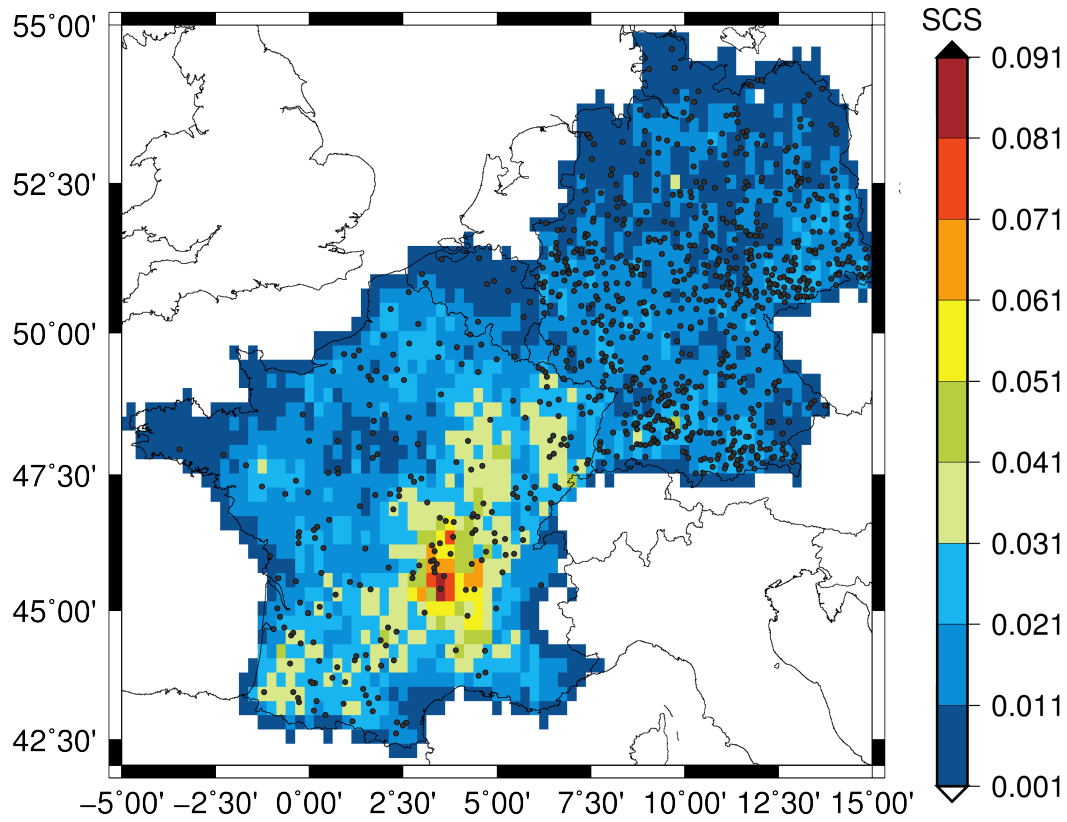


Figure 1. Number of SCS per year (center of each track) interpolated on a $0.25^\circ \times 0.25^\circ$ (color bar) and HS (black dots) between 2005 and 2014 over the investigation area (France, Germany, Belgium and Luxembourg).

250 3.2 Daily and seasonal cycle

Both HS and SCS events (the latter not shown) feature pronounced seasonal and diurnal cycles with a maximum in the afternoon in the warmest months of July and August. While the number of HS is lowest in April and September and dominated by smaller-sized hail, the months of May to July are similar with the highest number of HS events of the diameter class $D \geq 5$ cm in June (Fig. 2a). A comparison of the three summer months shows that events with large hail are rarest in July. Reasons for this counterintuitive result might be a decrease in frontal events, which have low hail sizes on average (cf. Sect. 2.5.1), or reduced reporting in this month due to summer holidays.

The diurnal cycle is much more pronounced than the seasonal cycle. The minimum number of HS events occurs in the early morning hours between 03 and 09 LT and the maximum in the afternoon between 15 and 18 LT (Fig. 2b). The largest increase occurs between 12 and 15 LT and the largest decrease after 21 LT. A total of 841 events, which correspond to 85.4% of the HS sample, is registered in the period from 12 to 21 LT.

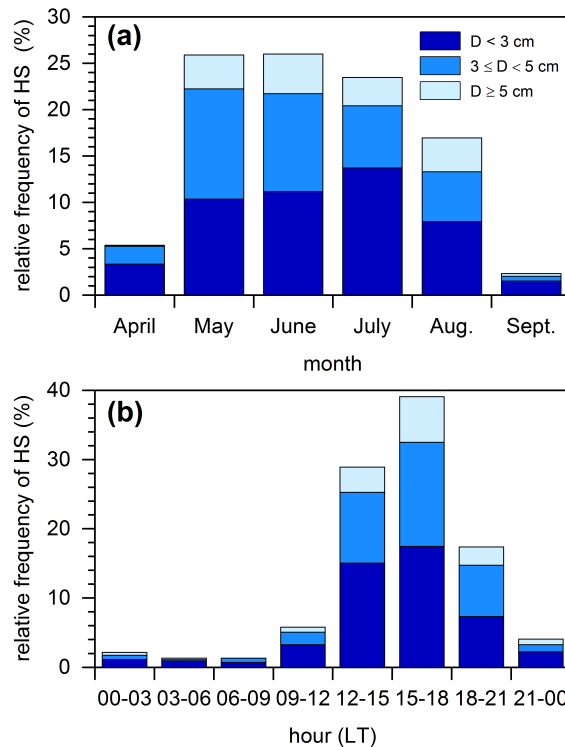


Figure 2. (a) Seasonal and (b) diurnal (local time, LT) cycle of HS tracks (SHY, 2004–2014) depending on the hail-size diameter according to ESWD reports.

A separation of the diurnal cycle according to the hail diameter shows that during the first half of the day (00–12 LT), most events are associated with hail smaller than 5 cm. Especially from 03 to 09 LT, hailstones are the smallest of the entire day. This result, however, must be treated with care because of the low number of events in that period (26 events) in combination with the potential underreporting by spotters in the night. During noon and afternoon, the proportion of hail with a diameter of at least 5 cm increases, and the highest probability of occurrence is between 15 and 18 LT. In the evening and night (18–00 LT), the relative proportion of large hail remains almost constant.

The pronounced diurnal cycle of the HS probability (Fig. 2b) is closely linked to the warming of near-surface layers of air and the associated increase in lapse rate / CAPE together with a decrease in CIN (Markowski and Richardson, 2010). In addition, triggering mechanisms such as low-level flow convergence in the wake of thermally-induced circulation over complex terrain or inhomogeneities in land cover are also connected to the diurnal temperature cycle. Studies using radar reflectivity or lightning detections found similar diurnal cycles for most of the area except for the Mediterranean (e.g., Wapler, 2013; Nisi et al., 2016; Piper and Kunz, 2017).

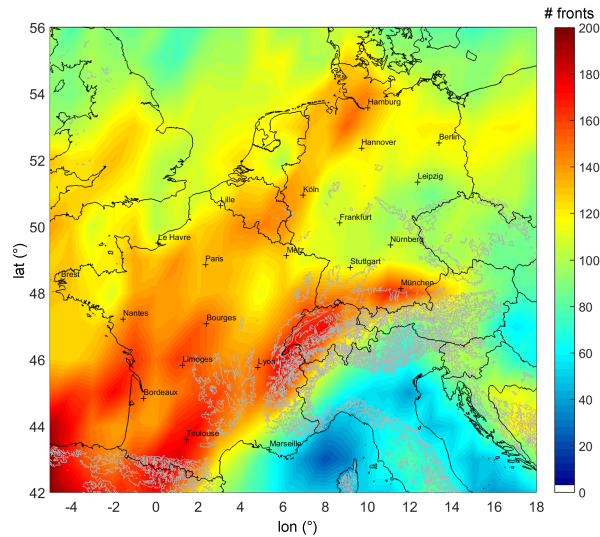


Figure 3. Number of synoptic-scale fronts per $1^\circ \times 1^\circ$ area between 2005 and 2014 (SHY) based on ERA-Interim reanalysis according to Schemm et al. (2015). Grey isolines represent the terrain (600, 1200, 1800, and 3600 m asl).

4 SCS associated with synoptic cold fronts

275 **Because of their relevance for SCS triggering, we investigate in the following the relation between synoptic cold fronts**
with a significant length typically associated with extratropical cyclones and SCS/HS events. Warm fronts are not
considered here because they are not important triggers for convection. This is mainly due to their reduced cross-
frontal circulation and the resulting slow ascend, deduced through the Sawyer-Eliassen equation (Emanuel, 1985), in
combination with warm air advection aloft, which has a stabilizing effect. Because of their limitation to a specific
territory, we also do not consider regional-scale land-sea contrasts, sea-breeze fronts and thermal boundaries from
280 **Alpine pumping in the analysis.**

4.1 Cold front climatology

The investigation area is frequently affected by synoptic-scale cold fronts. The number of fronts per grid point of the size $1^\circ \times 1^\circ$ during the 10-year investigation period ranges between 85 in eastern Germany and 175 near the Pyrenees (Fig. 3). Overall, front density in France is larger than in Germany.

285 During their propagation, cold fronts tend to weaken over land mainly because of friction in the lowest layers and horizontal mixing of air mass properties. Usually, they also dissolve when the air from the warm sector has entirely lifted (occlusion). As the largest fraction of fronts affecting central Europe propagates in east- to south-east directions, their detectable density gradually decreases in the same direction. In addition, an elevated front density can be found on the western and northern side (upstream) of large mountains such as the Pyrenees, Massif Central, and the Alps. These large mountain ranges tend to slow

290 down the propagation of fronts, leading to an elevated frequency upstream when counting the time steps where a front prevails
 and not the individual fronts as is done here (Schemm et al., 2016). Thus, slow propagating fronts may be repeatedly detected
 and counted during the time steps of ERA-Interim (6 hours). In contrast, fronts occur less frequently downstream of larger
 mountains as well as at a greater distance to the sea, where the increasing continentality acts to weaken or even dissolve the
 fronts.

295 4.2 Occurrence of frontal SCS and HS tracks

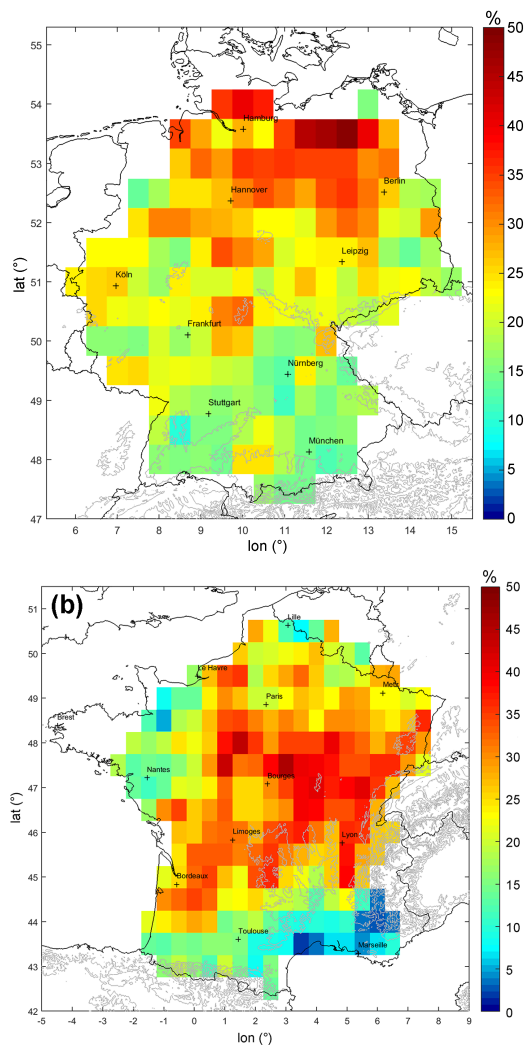


Figure 4. Share of frontal SCS (relative to all SCS; $r \leq 200$ km) over (a) Germany and (b) France for $0.5^\circ \times 0.5^\circ$ (SHY, 2005–2014). Grid points containing less than 50 SCS tracks (see Fig. 1) were left white.

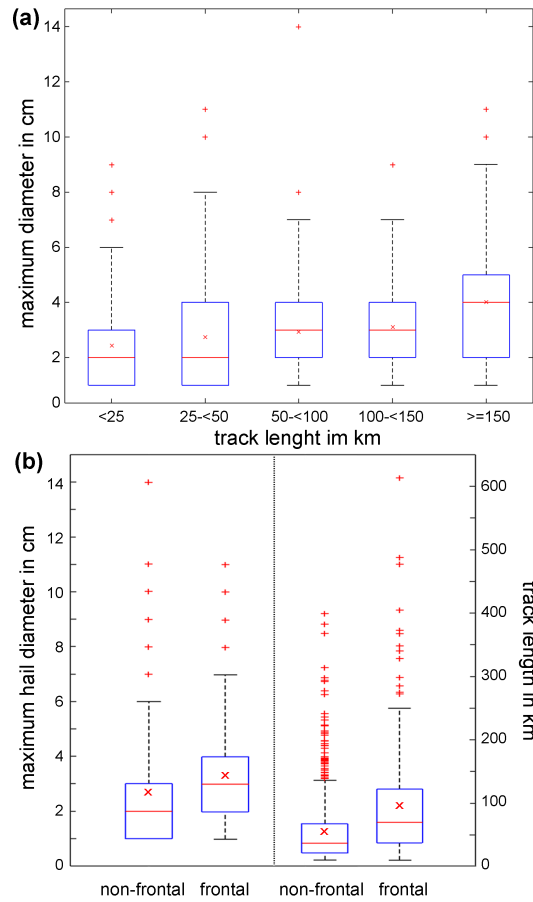


Figure 5. Boxplots showing (a) HS track lengths vs. maximum hail diameter according to ESWD reports and (b) maximum diameter (left part) and track length (right part) for HS events with/without a synoptic-scale cold front. Indicated in the boxplots are interquartile range (blue box); median and mean values (red line and red x); upper/lower 25% percentile \pm interquartile range \times 1.5 (black lines); data points outside of this range are marked as outliers (red crosses).

To assess the role of synoptic cold fronts in the probability and properties of SCS, we first discuss the spatial distribution of the ratio of frontal SCS relative to all SCS events. This ratio is computed independently for each single grid point with a size of $0.5^\circ \times 0.5^\circ$. Averaged over the entire area of Germany and over the 10-year study period, 18.9% of all SCS tracks are related to a cold front; for France, the ratio is slightly higher with 22.4% (Fig. 4). The most conspicuous feature in the spatial distribution of the frontal streaks is the strong gradient in the south-to-north direction, particularly over Germany. For example, while in the German northeast (Mecklenburg Lake Plateau) the share of frontal SCS reaches the highest value of 50%, it decreases to less than 10% in southern Germany over the Black Forest (SW Germany) and the region south of Nuremberg (SE Germany). In France, the spatial distribution of frontal SCS is smoother than in Germany. Most striking in France is the extended maximum

of the frontal share of around 45% northeast of the domain's center and several minima with only a few percent near the coasts
305 of both the North Atlantic and the Mediterranean.

If we compare the proportion of frontal SCS both with both the distribution of all SCS tracks (Fig. 1) and with the frontal
density (Fig. 3), the opposite behaviour is often observed. In several regions with an increased number of fronts and/or SCS
events, the number of frontal SCS is low and vice versa. This is especially true for Germany, but also for parts of France.
Over complex terrain such as in SW Germany (Black Forest) or Southern France (Massif Central), where frontal SCS are
310 comparatively rare, it can be assumed that orographically-induced vertical lifting is often sufficient to trigger convection so
that a front is not necessary.

Considering HS instead of SCS events, we found that an even higher number, namely 25% of all HS tracks across the
entire study domain are connected to a synoptic cold front. Because of the small number of HS track detections, especially in
France (cf. Fig. 1), we do not show this relation here. Note, however, that if only areas with a sufficient number of events are
315 considered, the spatial distributions of frontal HS and SCS tracks are quite similar.

For the HS events, a relation is found between the length of the tracks as detected by the radar algorithm and the maximum
observed hail diameter (Fig. 5a). While the mean diameter for a length of $L < 50$ km is around 2 cm, it increases to around
3 cm for $50 \leq L < 150$ km and to 4 cm for $L \geq 150$ km. Furthermore, the distributions of both quantities, maximum diameters
and track lengths, differ between frontal and non-frontal tracks. Mean diameters are 3.3 cm in for frontal events and 2.73 cm
320 for the others (Fig. 5b, left part). For hail size diameter classes of < 2 , 2–3.5, 4–5.5, and ≥ 6 cm, the ratio between frontal
and non-frontal events is 16.7, 23.1, 35.8, and 34.7%, respectively (not shown; note that the finer classes are used only in this
example). This means, the higher the probability of a nearby front, the larger is the hailstone diameter on average.

Differences between frontal and non-frontal HS events are also found for the length and mean propagation direction of
the tracks. While frontal HS tracks have a mean length of 96.2 km (interquartile range of 40–125 km), non-frontal tracks are
325 almost half shorter with 56.8 km (25–65 km; Fig. 5b, right part). Non-frontal HS events have a mean propagation angle of 215°
(interquartile range 185° – 255°), whereas those connected to a front propagate slightly more to the east with a direction of 232°
(interquartile range 217° – 258° ; not shown). Note that the interquartile range $\times 1.5$ is almost twice for non-frontal compared to
frontal HS). In that latter range of angles, also the largest hailstones can be observed.

~~These results suggest that fronts create hail-conducive conditions mainly through two effects: along-front advection of
330 moisture at lower levels leading to larger CAPE, and higher wind speed aloft enhancing vertical wind shear. The latter creates
an environment that favors the development of organized, more persistent thunderstorms such as multicells and supercells,
which plausibly explain the larger diameters and track lengths for HS events connected to a front compared to those without
a front. Furthermore, fronts also largely determine the orientation of the tracks, which results from the predominating general
flow direction from the west sector. Differences in prevailing environmental conditions are investigated in the next section.~~

Table 1. Number of HS events in the respective classes of maximum hail size diameter D and track length L .

	$L < 50$ km	$L = 50\text{--}100$ km	$L > 100$ km
$D < 3$ cm	311	98	64
$D = 3\text{--}4.5$ cm	190	102	72
$D \geq 5$ cm	63	35	50

335 5 Environmental conditions of HS tracks

Environmental conditions prevailing during HS events are investigated using SLI, DLS, and SRH from ERA-Interim reanalysis (see Sect. 2.2). The composites presented in the following show the mean fields of the respective parameter around the center of the HS tracks (see Sect. 2.5.3). To examine environmental conditions depending on the intensity of the HS events, we further divided the HS sample into 9 subsamples according to the observed hail diameter D (< 3 cm, $3\text{--}4$ cm, ≥ 5 cm) and track length L (< 50 km, $50\text{--}100$ km, > 100 km). **When defining the threshold values, it was taken into account that each class contains at least 50 events – except of the class $L = 50\text{--}100$ km and $D \geq 5$ cm (Table 1). Using other thresholds, for example, 150 km instead of 100 km as suggested by the diameter–length relation shown in the boxplot (Figure 5) would result in too small sample sizes with less than 30 events.** Further subdivision, for example according to the time of occurrence, was not carried out. Although scientifically interesting, this would further reduce the sample sizes, particularly the most interesting high-intensity classes.

5.1 Mean composites of environmental conditions

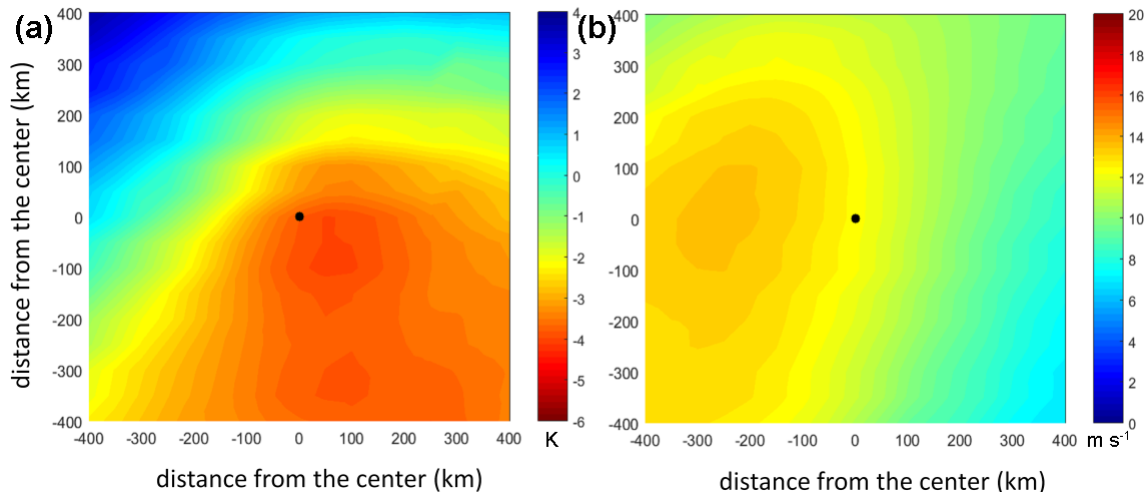


Figure 6. Composite analyses showing the average values of (a) SLI and (b) DLS from ERA-Interim in moving spatial windows centered at the track location (center) for all 985 HS events between 2005 and 2014 (SHY; see Fig. 1).

Averaged over all classes of HS events, the SLI around the center of the tracks has a value of -3.8 K in the mean (Fig. 6a), indicating a high potential for convective storms (e.g., Manzato, 2003; Kunz, 2007). SLI has its absolute minimum about 140 km southeast of the events, but the difference to the center, on average of 0.2 K , is almost negligible. Overall, a significant
350 increase in convection favoring conditions can be observed from the northwest of the HS center to the southeast. While these conditions prevail over 400 km to the south and east of the center, the area to the north and west sees higher and positive values of SLI, thus stable conditions, at approximately 100–200 km distance already. The SLI field occurs rather smooth, mainly because of the low resolution of ERA-Interim (cf. Sect. 5.3).

The vertical wind shear (DLS) has its maximum about 250 km to the west of the HS centers in an upstream direction (Fig. 6b).
355 This spatial difference is plausible because a trough frequently prevails to the west of the events. Since DLS is dominated by the wind speed aloft (500 hPa), a trough with an associated jet manifests itself by a maximum in DLS. Considering the magnitude of DLS, it is found that the values are quite low with a mean of 12.5 m s^{-1} around the HS events. Several authors have shown that organized convection capable of producing larger hail develops only in sheared environments above around 10 m s^{-1} (e.g., Weisman and Klemp, 1982; Markowski and Richardson, 2010; Dennis and Kumjian, 2017). This is one of the reasons to further
360 subdivide the whole sample as mentioned above and shown in the next paragraph.

5.2 Environmental conditions depending on hail size and track length

Separating the hail events according to their intensity allows a detailed view of the prevailing environmental conditions. The SLI composites show a slight decrease (higher instability) around the center of the HS events from small hail with shorter tracks (SLI $\approx -3.7\text{ K}$) to large hail with longer tracks (SLI $\approx -4.5\text{ K}$; Fig. 7). The strongest decrease in stability occurs for increasing
365 hail diameter, while the composites are less sensitive to variations in track lengths. In all cases, the lowest instability prevails to the southeast of the hail events as was already found for the mean composite (cf. Fig. 6). Despite favorable environments for SCS, which predominate all classes, the highest instability in the case of larger hail is an indicator for higher updraft speed within the thunderstorm cloud, which is a prerequisite for the growth to large hailstones.

The distance between the location of the events and the location of the highest instability is greater for longer tracks than for
370 shorter ones, but only in case of small to medium-sized hail. At this point one may speculate that the reason for this shift might be related to the role of cold fronts, considering that longer tracks and larger hailstones are more often connected to a cold front as discussed in the previous section (cf. Fig. 5). The role of cold fronts versus environmental conditions will be investigated in the next section.

In contrast to the thermodynamical proxy SLI, the dynamical quantity DLS shows significantly pronounced differences between the nine HS categories (Fig. 8). Even though DLS also distinguishes between the diameter classes, the largest differences
375 are found for the three-length classes. For example, DLS has a mean value of 17 m s^{-1} for long tracks in the smallest diameter class ($D < 3\text{ cm}$), which is almost twice as high compared to short tracks with the same diameter class (8.5 m s^{-1} ; Fig. 8, upper row). The same applies to the other diameter classes. For long tracks with large hail, DLS reaches values of about 20 m s^{-1} and is thus in the range of the values given in the literature (e.g., Weisman and Klemp, 1982; Thompson et al., 2007; Markowski
380 and Richardson, 2010). The area of the highest DLS values is located several hundred kilometers to the west of the HS events

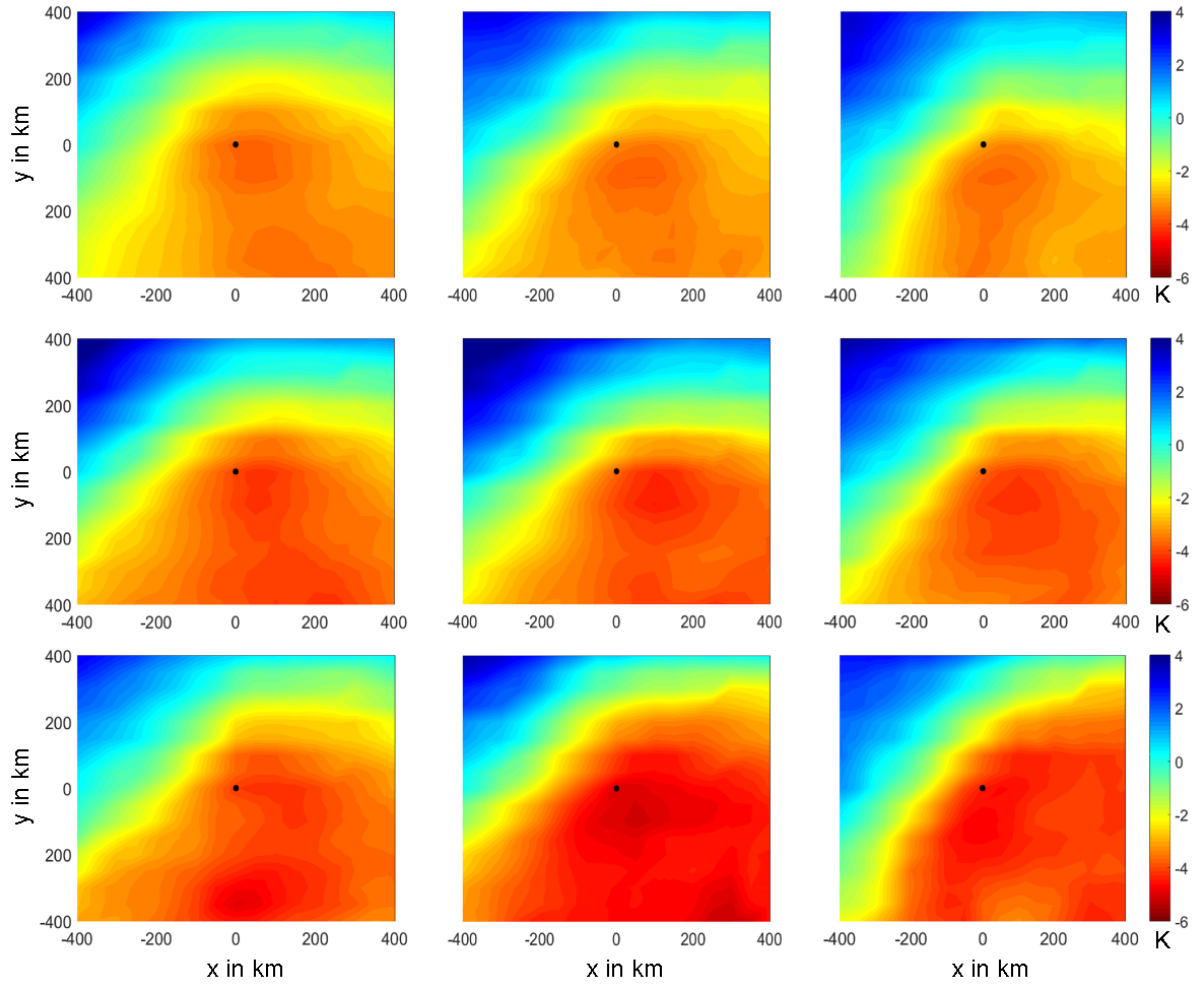


Figure 7. Composite analyses of SLI related to maximum observed hail diameters of $D < 3$ cm (top row), $3\text{--}4.5$ cm (middle row), and ≥ 5 cm (bottom row) and for track lengths of $L < 50$ km (left column), $50\text{--}100$ km (middle), and ≥ 100 km (right). The sizes of the subsamples are listed in Table 1.

on average. For large hail, the DLS maxima are even higher and further away from the HS events. These events are usually triggered by upper-level troughs to the west, associated with higher wind speed at mid-troposphere levels. One may argue that a relationship between DLS and track length prevail per se since both are dominated by the wind speed aloft. Note, however, that the separation of DLS applies not only to track length, but also to storm duration (not shown here, but see Wandel, 2017).

385 In addition to DLS, SRH has been suggested by several authors (e.g., Thompson et al., 2007; Kunz et al., 2018) to be an important proxy not only for the prediction of tornadoes but also for large hail. In our composite analyses, SRH (Fig. 9) shows even more pronounced differences between the nine HS categories compared to DLS. Hail events with shorter tracks on average are in a range between 0 and $50 \text{ m}^2 \text{ s}^{-2}$. By contrast, longer tracks have much higher mean values between 84

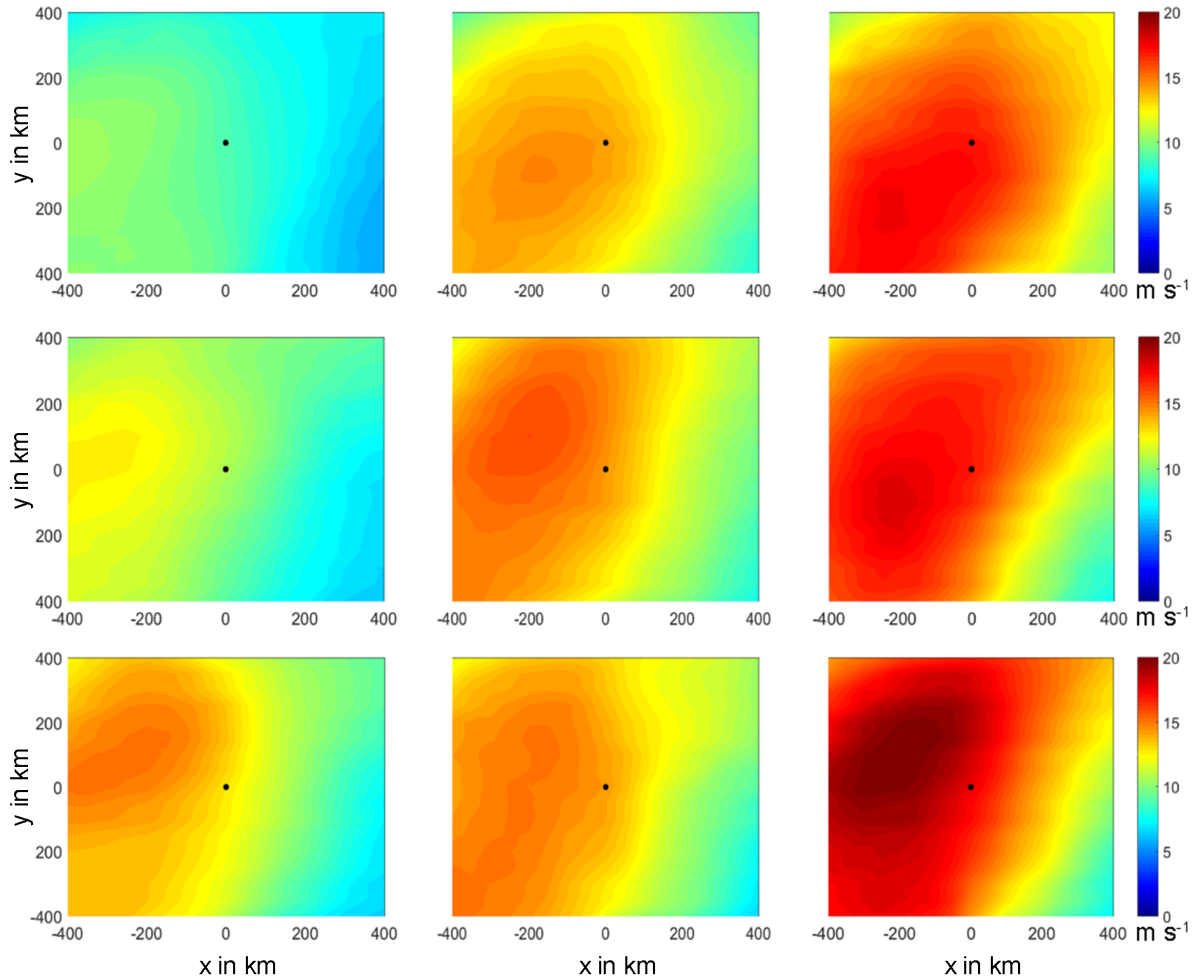


Figure 8. Same as Fig. 7, but for 0–500 hPa DLS.

and $116 \text{ m}^2 \text{ s}^{-2}$. **According to the investigations of proximity soundings by Thompson et al. (2007), such environments favor the development of weakly tornadic and nontornadic supercells – provided that sufficient CAPE is present.** In addition, there is also an increase in SRH from small to large hail, which is weaker compared to the trend in the length classes. Interestingly, the highest SRH values occur directly at or near the location of the hail event and not on the upstream side as was the case for DLS.

To further investigate which of the dynamical parameter, SRH or DLS, best distinguishes the HS intensity, we consider only the two categories that correspond to the highest and lowest damage potentials: smaller hail with $D < 3 \text{ cm}$ combined with short track length of $L < 50 \text{ km}$ and large hail with $D \geq 5 \text{ cm}$ combined with longer tracks of more than 100 km (high-intensity events). Environmental parameters are computed by the mean of the 3×3 ERA-Interim grid points centered around the HS locations.

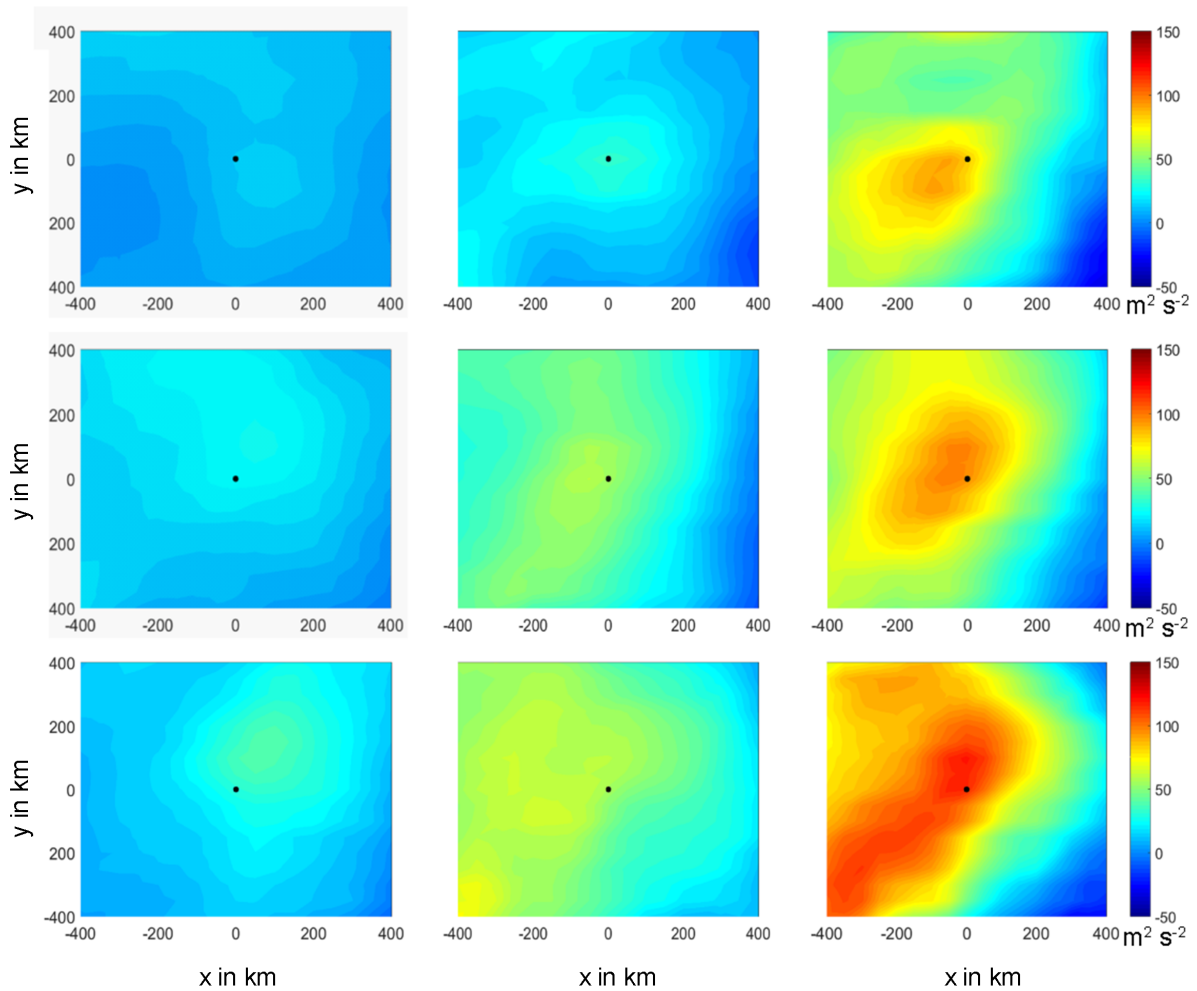


Figure 9. Same as Fig. 7, but for 0–3 km SRH.

Overall, the scatter plots presented in Figure 10 show a much clearer separation between the events when SRH is considered (Fig. 10a) instead of DLS (Fig. 10b). About 50% of the high-intensity events have values of $100 \text{ m}^2 \text{ s}^{-2}$ or greater for SRH, while only 3% of the less-intensity events display these values. Furthermore, most of the latter events have values between -50 and $50 \text{ m}^2 \text{ s}^{-2}$. It can also be seen that SLI for all events in these two categories vary between 0 and -10 K , with only a few exceptions having positive values. Approximately 70% of the high-intensity events have values of -2.5 K or less. Unlike DLS (Fig. 10b), splitting the events into two different categories is not possible. Even if most of the high-intensity events form in an environment with DLS of at least 15 m s^{-1} (approx. 60% of these events), there are still many low-intensity events for larger DLS values.

In addition, we quantitatively evaluated the prediction skill of certain parameter combinations based on categorical verification. Highest skill scores are given for SRH in combination with a thermodynamic proxy such as SLI or 700–500 hPa lapse rate.

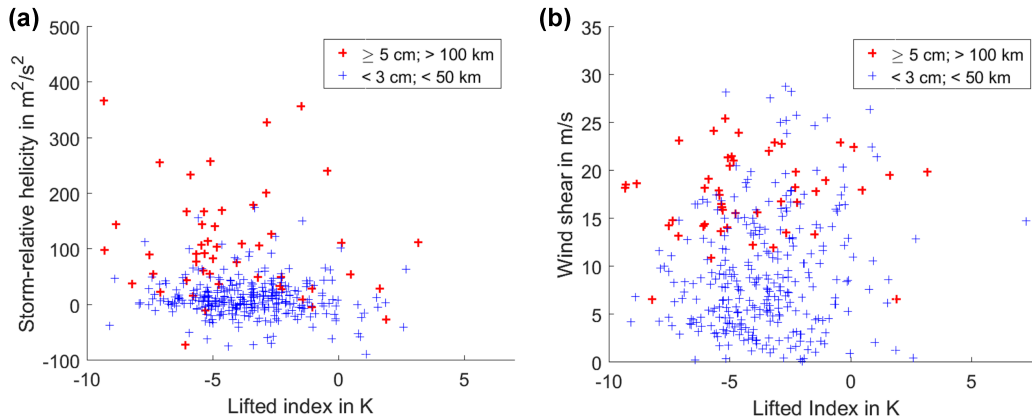


Figure 10. Scatter plots between (a) SLI and SRH and (b) DLS for two different categories of track length and hail diameter.

For example, the Heidke Skill Score (HSS), which is particularly suitable for predicting rare events, is 0.50 for the SRH–SLI combination (0.50 for lapse rate instead of SLI), but only 0.27 for the DLS–SLI combination (0.33 for lapse rate). This finding suggests considering SRH at least in addition to DLS when predicting the severity of SCS.

5.3 Effects of model resolution on convective parameters

Subgrid-scale spatial variations of the environmental conditions, for example as a result of diabatic heating or temperature and moisture advection (Markowski and Richardson, 2010), cannot be expected to be reproduced by the coarse ERA-Interim reanalysis. For this reason, we additionally considered the high-resolution coastDat-3 reanalysis. Due to the hourly resolved model fields, the maximum time difference between the HS events and the environments is 30 min. The purpose is not to reproduce the above analyses, but to investigate exemplarily the influence of the model resolution on the results. Since SLI and SRH are not available or quantifiable from coastDat-3, we used LR as thermodynamical proxy and DLS as dynamical proxy (cf. Sect. 2.2). Because the two proxies do not show significant differences between the nine intensity categories (cf. Figs. 7 to 9), we discuss only the most severe HS category with $L \geq 100$ km and $D \geq 5$ cm.

As shown in Figure 11, the higher model resolution (right column) has little influence on the spatial distribution of the environmental parameters even though coastDat-3 composites show a much larger spatial variability compared to ERA-Interim. In case of LR, the maximum is located southwest, in case of DLS it is located northeast of the HS events as was already found in the above analyses. Also the distance between the maxima and the events remains almost the same. The coastDat-3 values around the maxima show a slight increase of approximately 10% for both parameters. In the vicinity of the HS centers, the increase is only marginal, but larger for LR compared to DLS. In particular the LR increase is a consequence of the higher temporal resolution of coastDat-3 leading to an improved representation

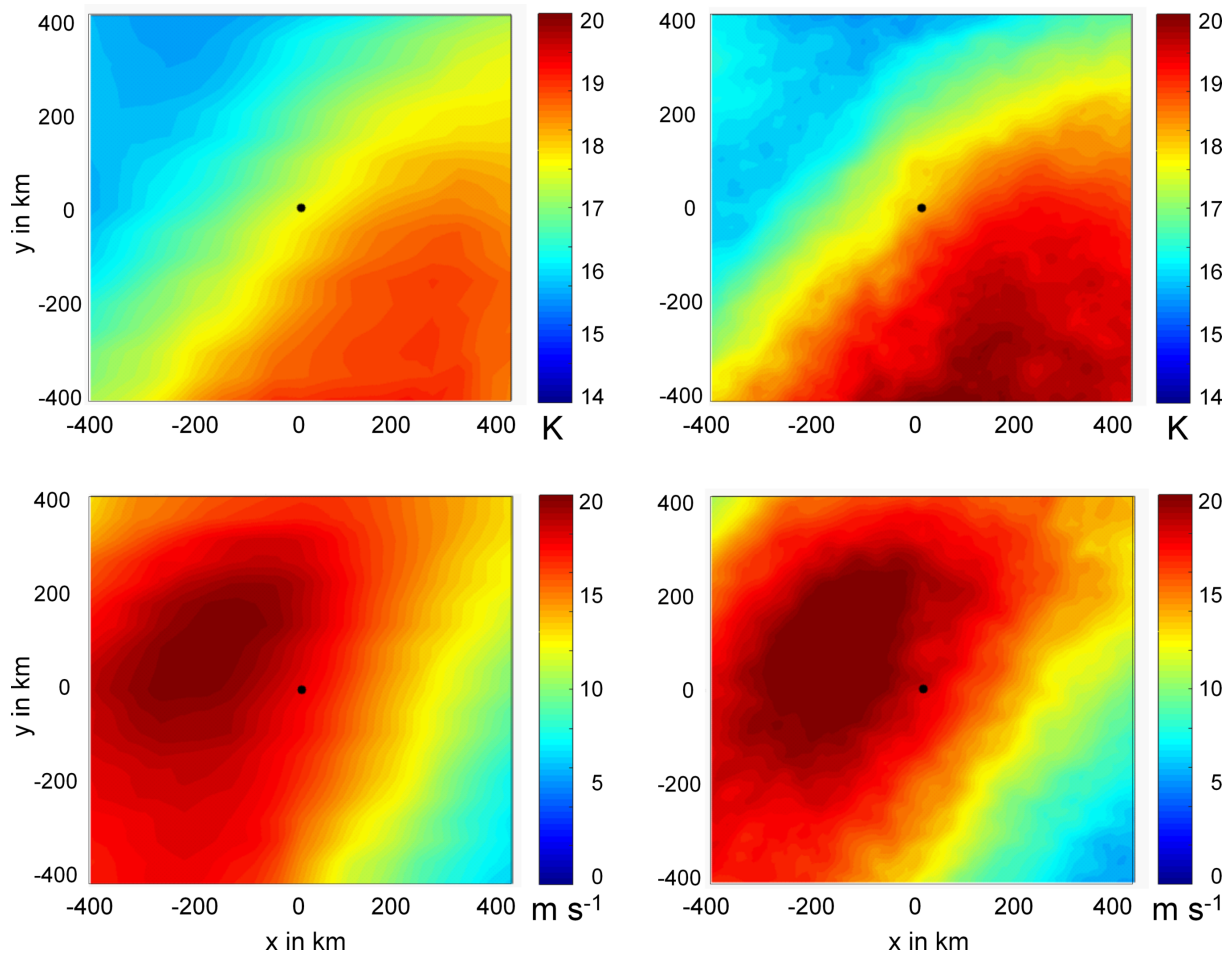


Figure 11. Composites of LR (top row) and DLS (bottom row) for hail diameters $D \geq 5$ cm and track lengths $L \geq 100$ km based on ERA-Interim (left column) and coastDat-3 (right column) reanalyses.

430 **of the diurnal temperature and moisture cycles. Note that this finding does not only apply to LR, but also to other thermodynamic quantities such as the precipitable water (not shown).**

5.4 Frontal vs. non-frontal HS tracks

As already discussed in Section 4.2, the characteristics of HS tracks having a front nearby substantially differ from non-frontal events, especially with regard to the maximum hail size and the track lengths (cf. Fig. 5). This suggests that prevailing environmental conditions may likewise differ for the two kinds of events. Therefore, we further subdivided the HS sample into
 435 frontal and non-frontal types. To ensure that enough events enter the subsamples, we made a further separation by considering **only two length classes ($L < 75$ km and ≥ 75 km) and two diameter classes ($D < 3$ cm and ≥ 3 cm; the former not shown).**

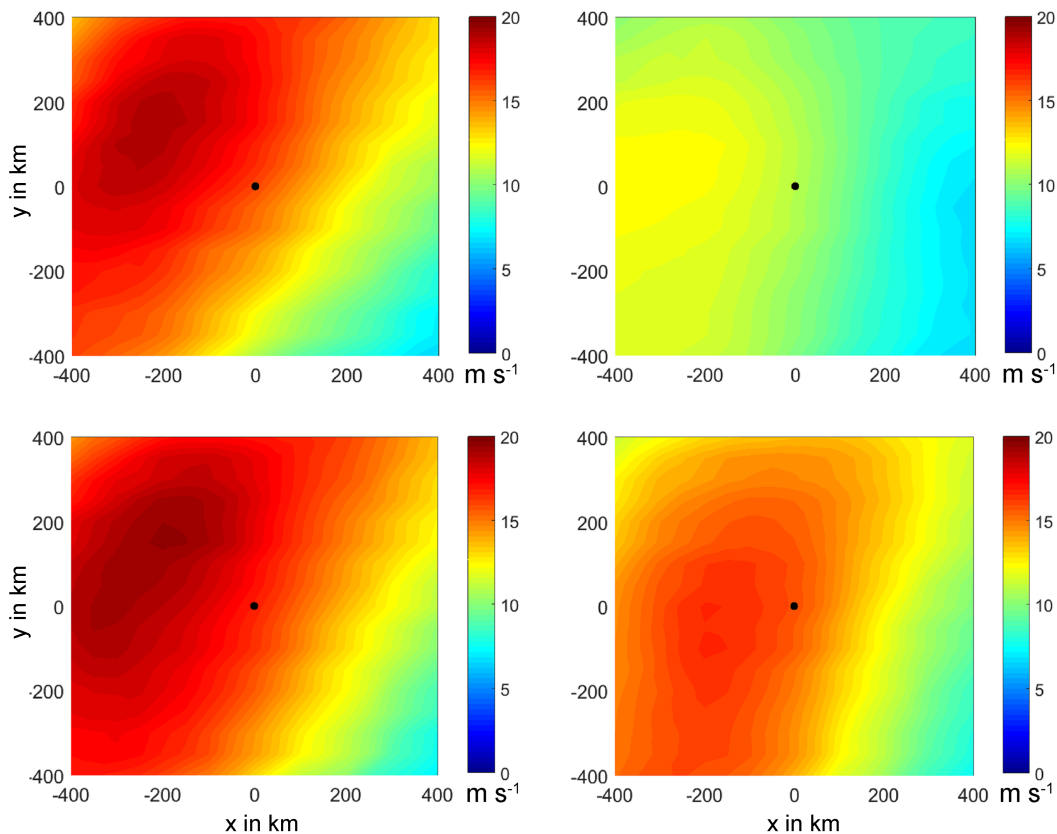


Figure 12. Composites of DLS for maximum observed hail diameters $D \geq 3$ cm and track lengths of $L < 75$ km (top row) and $L \geq 75$ km (bottom row) for frontal (left column) and non-frontal (right column) HS events.

Whereas the mean SLI composites are almost similar for frontal / non-frontal events (not shown), DLS shows significant differences between the four classes (Fig. 12). Overall, DLS reaches higher values with larger gradients for frontal compared to non-frontal events (Fig. 12, left vs. right column). However, when considering additionally the track lengths, much larger differences in DLS can be found, but only for non-frontal events (Fig. 12, right column). While short non-frontal tracks form at a DLS of 10.9 m s^{-1} on average, long tracks require medium sheared environments, here with values of 15.9 m s^{-1} . A similar result is obtained for small hail sizes ($D < 3$ cm) with DLS even rising from 9.0 to 16.7 m s^{-1} (not shown). Furthermore, while the DLS maximum for non-frontal events is located to the west of the center, it is more northwest for frontal events at a distance of about 200 km. Since almost all synoptic fronts in Europe propagate in a west-to-east direction, this location is a clear indication that frontal HS events preferably develop in prefrontal environments (and not postfrontal).

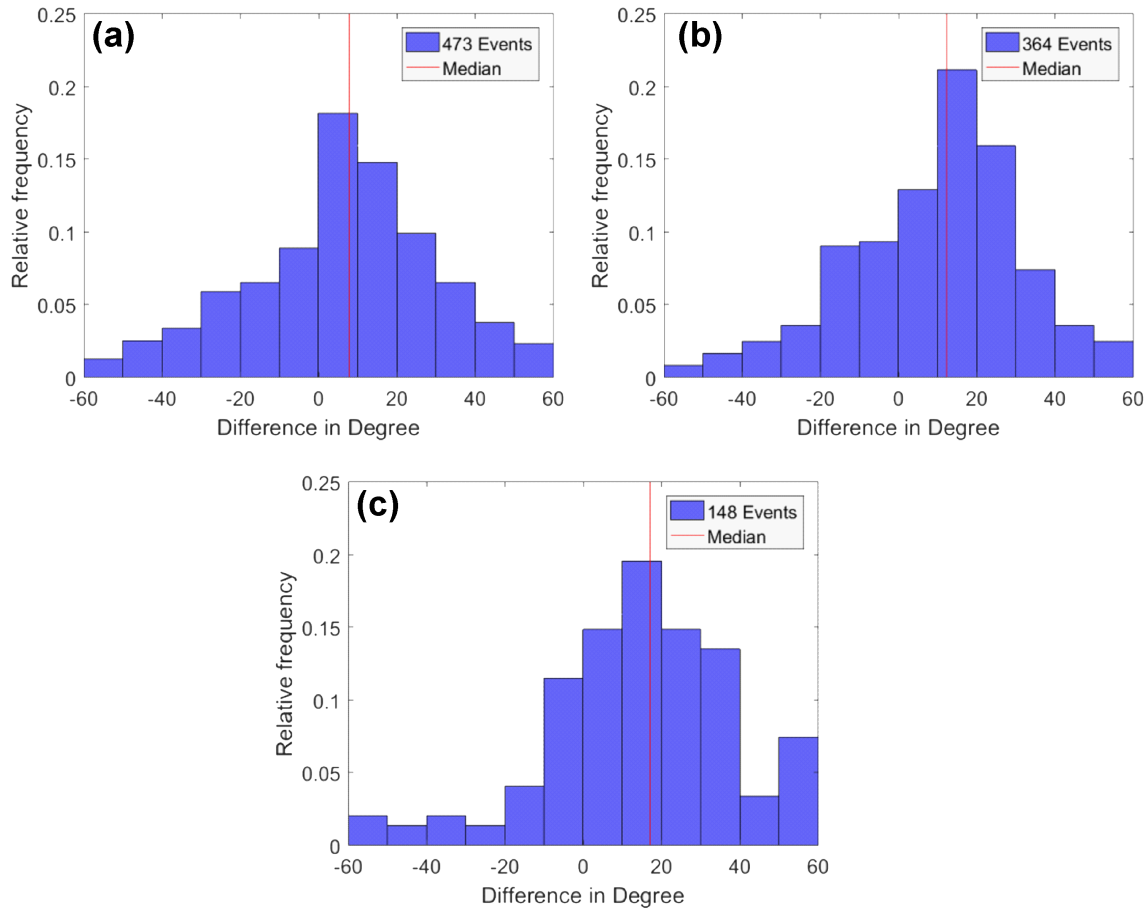


Figure 13. Histograms of HS events showing the relative frequency of the differences in the propagation direction between the storm motion vectors \mathbf{c} and the wind in 500 hPa from ERA-Interim at the location and time of the HS events for three different diameter categories: (a) $D < 3$ cm, (b) $3 \leq D < 5$ cm, and (c) $D \geq 5$ cm. Median values are indicated by the red line.

5.5 Differences in wind direction

450 **It is well-known that supercells due to specific conditions, such as a strong and spatially extended updraft, high amounts of supercooled liquid water, or their longevity, are capable to produce the largest hailstones (Foote, 1984; Markowski and Richardson, 2010; Dennis and Kumjian, 2017). The propagation of these highly organized convective systems can substantially deviate from the horizontal wind at mid-tropospheric levels, mainly because of the dynamics of the cold pools and induced vertical pressure deviations (Markowski and Richardson, 2010).**

In the last step, therefore, we want to investigate whether our samples show a relation between the storm motion relative to the mean wind and the hail size. The storm motion vector \mathbf{c} follows from the radar tracking of the individual HS events, the

455 wind direction is estimated from the 500 hPa mean wind from ERA-interim (3×3 grid point around the HS centers). **The cell tracking algorithm (cf. Sect. 2.4) yields very reliable shift vectors of individual hailstorms. The wind field in 500 hPa, on the other hand, is mainly determined by the setting of the synoptic systems and only marginally affected by local-scale flow deviations. Positive differences in the analyses indicate right-moving storms, negative values left-moving storms.**

Most of the events with smaller hail ($D < 3$ cm) propagate approximately parallel to the wind vectors in 500 hPa; the mean
460 difference between the tracks and the wind vectors is only 8° (Fig. 13a). About 13% of all HS events have a deviation between 30 and 60° to the right, while only 6% of the events show deviations to the left for this interval (-30 to -60°). Hail events with maximum diameters between 3 and 4.5 cm show a deviation of the propagation direction preferably to the right of the wind vectors (Fig. 13b). 23% of all HS events propagate with the wind in 500 hPa (decreasing by 8% compared to small hail), while 38% of the tracks show a deviation between 10 and 30° .

465 HS events of the largest hail class not only show an increased spread of the propagation deviation but also the entire histogram is shifted to more right-movers (median: 17° ; Fig. 13c). An angle difference between 10 and 30° is observed for 35% of all events. The largest difference to the other hail size classes is the comparatively high number of HS events between 30 and 60° (21%). In contrast, 27% of the events propagate with the wind in 500 hPa, and only 10% have a negative deviation to the left of the wind in 500 hPa. In summary, the larger the hailstone diameters, the stronger is the deviation of the cell's propagation
470 direction from the flow at 500 hPa.

6 Discussion

Severe convective storms, chiefly hailstorms, are high-frequent perils that, due to their local-scale nature, affect only small areas (Changnon, 1977). Their reconstruction requires high-resolution observational data such as radar reflectivity used in our study. The results of the analyses show high spatial variability of both SCS and HS events, with a gradual increase with growing
475 distance from the ocean and several hotspots, mainly over and downstream of mountain ranges. For example, as shown by Kunz and Puskeiler (2010), these hotspots are connected to flow convergence at lower layers in the low Froude number regime, when the flow tends to go around rather than over the mountains. Overall, the spatial distribution of SCS/HS events agrees with other studies on that topic considering different data sets such 3D radar reflectivity (Kunz and Kugel, 2015; Puskeiler et al., 2016; Lukach et al., 2017), a combination of radar data with weather stations (Junghänel et al., 2016), or overshooting top
480 detections from satellite (Bedka, 2011; Punge et al., 2017). This applies also to the detected seasonal and diurnal cycles (Nisi et al., 2016; Punge and Kunz, 2016; Nisi et al., 2018). The good quantitative and qualitative agreement is a strong indication of the reliability of our methods and results.

All composites of environment parameters created for radar-derived HS tracks show a similar spatial pattern: whereas the thermodynamic proxies such as SLI have their highest values at some ten kilometers up to 100 km southeast of the center of
485 the HS events, the maxima of the dynamic proxies (DLS and SRH) are found to the northwest at a distance of 100 to 200 km. This applies to all intensity classes and to all proxies originally considered in our study (also for KO-Index and lapse rate, but not for PW, where the maximum is located north of the events).

In total 651 of all 985 HS events have a southwest to the northeast propagation direction, reflecting the mean flow direction at mid-troposphere levels. On average, HS events usually occur downstream of the eastern flank of a mid-troposphere trough, where southerly to southwesterly winds are frequently associated with the advection of unstable, warm, and moist air masses from the Mediterranean (Graf et al., 2011; Wapler and James, 2015; Piper et al., 2019). This constellation is often referred to as “Spanish Plume” (Morris, 1986). The trough, on the other hand, creates an environment with increased wind shear and large-scale lifting. The axis of the trough is usually located several hundreds of kilometers upstream of the HS events, which explains why the highest shear is found on the western flank at larger distances. Furthermore, as convection initiation requires an additional lifting mechanism to overcome the convective inhibition in the planetary boundary layer, the area downstream of a trough is an ideal location for the development of (organized) convection as shown, for example, by Wapler and James (2015), Piper et al. (2019), or Mohr et al. (2020).

The separation of the environmental composites into different classes of hail diameter and track length yields several interesting results. Thermal instability, as expressed for example by SLI, increases slightly (smaller values of SLI) from small hail with shorter tracks to large hail with longer tracks, as might be expected. While the strongest decrease is found for increasing hail sizes, the composites are only marginally sensitive to variations in the track length. By contrast, the separation for DLS and SRH is much stronger, particularly for the track lengths. This dependence of the track lengths to DLS or SRH can be explained plausibly by the storm’s organization. Low-to-medium sheared environments ($\leq 10 \text{ m s}^{-1}$) permit single cells to develop (Markowski and Richardson, 2010), which are not able to produce large hail. For organized convective storms such as multicells, supercells, or MCS, substantial shear ($> 10 \text{ m s}^{-1}$) is required, which spatially separates the updraft from the downdraft. Supercell thunderstorms, bearing the largest hailstones, preferably develop in environments with DLS exceeding 18 m s^{-1} (Weisman and Klemp, 1982; Markowski and Richardson, 2010). High-resolution model simulations by Dennis and Kumjian (2017) show that increased DLS upstream elongates the storm’s updraft downshear, providing an increased volume of the hailstone growth region, an increased hailstone residence time within the updraft, and a larger region for potential hail embryos. Altogether, these mechanisms lead to increased hail masses and, thus, increased hail diameters, even though the average value of DLS for our event set is at the lower end of the typical value range for multicellular convection.

From the comparison of the two reanalyses, we conclude that ERA-Interim with a comparatively coarse spatial and temporal resolution is suitable to estimate environmental conditions. A higher model resolution is mainly important for estimating thermodynamical parameters, especially those depending on the diurnal temperature cycle. Because the dynamical environment is not directly connected to the diurnal temperature cycle and therefore does not change much during the day, DLS or SRH, for which our results suggest the closest relation to track length and hail size diameter, can be reliably estimated from low-resolution global models such as ERA-Interim.

The hypothesis that supercells preferably enter the subsample of long tracks / large hail is also supported by the findings of the differences between the propagation vector of cells determined by the tracking algorithm and the mean wind at 500 hPa from ERA-Interim reanalysis. The larger the hailstones, the larger is the relative share of events with a deviation mostly to the right of the ambient wind. Because of vertical dynamic pressure perturbations, supercells tend to deviate substantially from the mean wind direction (Markowski and Richardson, 2010). So-called right-moving supercells, usually persisting after cell

splitting (Klemp, 1987) because of positive linear dynamic forcing, may deviate from the mean wind direction by angles of up to 30°. Such deviations have already been observed for single supercells in Germany (Kunz et al., 2018). In contrast, multicell
525 thunderstorms or MCS bearing smaller hailstones show fewer deviations from mid-tropospheric winds.

When a synoptic cold front is involved, the pre-convective environment can substantially change on short time scales because of four independent effects (Giaiotti et al., 2003; Markowski and Richardson, 2010; Schemm et al., 2016): (i) lapse rate increase by cold air advection aloft; (ii) vertical lifting by frontal cross-circulations, which simultaneously increases CAPE and reduces CIN; (iii) **along-front advection of moisture at lower levels leading to an increase in CAPE**; and (vi)
530 enhanced curvature of the hodograph related to the thermal wind equation **and, thus, enhanced vertical wind shear**. The latter, not directly connected to a front, potentially occurs several hundred kilometers upstream. **All the above-listed factors create an environment that favors the development of organized and more persistent thunderstorms, such as multicells and supercells. Therefore, hail events associated with cold fronts are likely to have different properties than non-frontal events. We found, for example, frontal HS events to produce larger hail and longer tracks compared to non-frontal HS**
535 **events on average. Furthermore, the tracks are strongly coupled to the (typically eastward) propagation of the fronts.**

Frontal detection in ERA-Interim is based on some specific criteria such as temperature gradient, minimum length or propagation speed to consider only significant synoptic fronts. The use of fixed thresholds for these parameters may bring in some bias in the analyses. Especially over and downstream of larger mountain ranges, such as the Massif Central, the Black Forest or the Alps, fronts can be significantly fragmented or distorted (Löffler-Mang et al., 1996;
540 **Dickinson and Knight, 1999) and thus be eliminated by the detection criteria. When interpreting the results of the relationship between fronts and SCS/HS events, it is important to be aware of this limitation. Our purpose was to use an objective identification of fronts, which is valid for the whole study area, and to consider only significant fronts.**

The share of frontal SCS (and HS) to all events substantially varies among the regions. For example, whereas only a limited number of SCS in southern Germany have a front nearby, almost half of the events over northern Germany are front-related.
545 By combining radar-based hail events for Switzerland between 2002 and 2013 with cold front detections (Schemm et al., 2015) based on COSMO analysis, Schemm et al. (2016) found that locally up to 45% of all hail events in northeastern and southern Switzerland are associated with a cold front. This is similar to our study region, where we identified values of up to 50% locally.

Over complex terrain, it can be assumed that moisture flux convergence at low levels caused by flow deviations at obstacles
550 and local wind systems is the most important trigger mechanism for convection initiation (Weckwerth and Parsons, 2005; Barthlott et al., 2011; Trefalt et al., 2018). In contrast over mainly flat terrain such as in northern Germany, a front is often required as a trigger for convection. Instability and vertical wind shear are two additional effects that partly determine the probability of frontal SCS. These two quantities on average are highest in the southern parts of France and Germany, where frontal SCS are not very frequent. Thus, we conclude that the share of frontal SCS to all events is the result of the interaction
555 of various influencing factors, mainly of thermal instability and lifting mechanisms to initiate convection.

When a front is nearby, HS events tend to develop east of the maximum of wind shear and northwest of the most unstable stratification. In contrast, non-frontal HS events frequently occur in proximity to the highest wind shear and most unstable

conditions. In low-sheared environments, hailstorms capable of producing hail larger than 3 cm develop only when the air mass is highly unstable. Higher instability, in general, enables stronger updrafts that are required for the development of larger hailstones. For frontal HS events, the stratification remains almost the same, but with the highest instability located more to the southeast of the events. This region of highest instability, however, is characterized by lower shear. At the same time, assuming a trough prevailing at the western side of the HS events, large-scale descent associated with high-pressure systems tend to suppress convection initiation (Piper and Kunz, 2017). This relation also explains why the dynamical and thermodynamical conditions in terms of DLS and SLI prevailing during HS events for the different classes are consistent among themselves.

565 7 Conclusions

In our study, we have reconstructed a large number of past severe convective storms and investigated prevailing environmental conditions over a 10-year period in central Europe. The combination of SCS tracks derived from 2D radar data with hail reports from ESWD gave additional information on the hailstone size of a storm, but also ensured that the resulting subsample consisted of hailstorms solely. The resulting HS subsample allowed us to investigate prevailing environmental conditions from reanalysis as a function of hail size and track length. In addition, we have investigated how and through which mechanisms synoptic cold fronts modify the characteristics and the frequency of SCS and HS events. Our study is the first of its kind that relies on both hail size and track length, a combination essential for the damage potential of severe hailstorms.

The main conclusions from our research are the following:

- 575 – ~~The probability of occurrence of both SCS and HS events shows a distinct seasonality with most of the events occurring between May and July, and a pronounced diurnal cycle with a maximum in the afternoon between 15 and 18 LT, where storms with large hail ($D \geq 5$ cm) are most likely to occur.~~
- Approximately one-quarter of all SCS across the investigation area is connected to a front, being usually pre-frontal events. Over complex terrain, such as in southern Germany, the share of frontal SCS is low (partly below 10%), while over flat terrain a front is more often required (up to 50% of all events) to trigger convection.
- 580 – Frontal HS events on average produce larger hailstones and have longer tracks. These events preferably develop in a high-shear environment related to the cold front.
- Dynamical proxies such as DLS or SRH become important when separating between hailstorms of different intensity classes with respect to hail diameter and length (or likewise duration). Thermodynamic proxies such as SLI or lapse rate show only small differences around the event's centers between the different classes.
- 585 – SRH (0-3 km) as dynamical proxy performs better compared to DLS when separating HS events according to hail size and track length.

- The larger the hail size, the larger the deviation between track direction and direction of the mean wind at 500 hPa. Most of the large hail events ($D \geq 5$ cm) propagate to the right of the mean wind, suggesting an increased probability of right-moving supercells in that subsample of HS events.

590 A potential weakness of our study is that it relies on eye-witness reports (ESWD), which are biased towards denser populated regions and towards daytime (Groenemeijer et al., 2017). This constraint reduces not only the size of the HS sample but also creates a spatial bias as can be seen in the substantially lower number of HS events in France than in Germany. Furthermore, the estimation of the largest diameter for hailstones that may substantially deviate from a sphere creates additional uncertainty.

595 Despite the different sources of uncertainty and the limited representativity of the reports for several regions, the comparatively large sample including approximately 1000 events enables reliable statistical analyses when aggregated over the whole investigation area. Furthermore, ESWD reports are the only dataset that gives additional information about hail diameter. Insurance loss data used in several hail-related studies (e.g., Vinet, 2001; Schuster et al., 2006; Kunz, 2007) or data from hailpad networks (e.g., Dessens and Fraile, 1994; Sánchez et al., 2017) cannot be applied because of the large spread inherent in the damage-to-diameter relation or the limited regions gauged. In the future, ground-truth observations collected through
600 crowdsourcing via specific platforms such as the European Weather Observer App (EWOBS; Groenemeijer et al., 2017) or the MeteoSwiss App (Trefalt et al., 2018; Barras et al., 2019) might overcome the underreporting of hail events.

In our study, we have taken the HS events as basis of the analysis and then examined prevailing environmental conditions. From a forecasting perspective, however, the reverse questions is actually of great relevance: What is the probability of a severe footprint (length and hail diameter) under the current (or predicted) environmental conditions. This question, however, could not be evaluated quantitatively or probabilistically as the hail reports archived by the ESWD are incomplete, especially over France. One possibility would be to consider only the expected lifetime (or length) of a storm cell in the prediction scheme and to ignore the hail diameter - even if this quantity is most important for the damage.

605

Nevertheless, the main findings and conclusions of our study can be considered in several ways. Above all, the results can
610 (and should) be considered in the forecasting of SCS for lead times between 1 and 12 hours. This time range is of considerable importance for many users as well as for issuing warnings of SCS associated with high impact weather phenomena such as hail, heavy rainfall, or severe wind gusts. In the hierarchy of prediction models, this time range is covered by nowcasting tools and very short-range forecasts (Nisi et al., 2014; James et al., 2018). Hence, convective indices, particularly SRH or DLS, might be employed in both systems. Our results can help to distinguish between less severe and more severe convection. When
615 focusing on the most severe storms, the magnitude and temporal evolution of SRH/DLS, and whether a front is nearby should be considered. Finally, because there is evidence of an increase in the number of extremely strong weather fronts during the summer over Europe (Schemm et al., 2017), our findings have implications for explaining trends and regional-scale variability of front-related SCS and HS.

Data availability. 2D radar data for Germany can be downloaded via the DWD ftp Server. Tracks of SCS were calculated from DWD radar data. The track data are not freely available, but may be provided upon request. ESWD reports are available via www.eswd.eu. ERA-Interim can be downloaded from the ECMWF server, coastDat-3 reanalysis can be requested from the Helmholtz Zentrum Geesthacht (HZG). Front analyses can be provided upon request.

Author contributions. MK designed the research and wrote most parts of the paper. JW conducted the analyses of the environmental conditions and wrote the corresponding sections. EF performed the SCS/hail track analyses, while SB combined the tracks with ESWD data and frontal detections. SeS provided the data of synoptic fronts and wrote the corresponding section. SM and SeS edited the paper and provided substantial comments and constructive suggestions for scientific clarification and further improvements.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The authors thank the German Weather Service (DWD) and Meteo France for providing radar data, ESSL for making available archived observations, Siemens AG (S. Thern) for providing lightning data, and HZG (B. Geyer) for providing coastDat-3 hindcasts. ERA-Interim was downloaded from the ECMWF web server. Data is stored at the Research Data Archive at the Karlsruhe Institute of Technology (KIT) and is available upon request to M. Kunz. The authors also gratefully acknowledge the two anonymous reviewers for their helpful comments and suggestions.

References

- Aran, M., Pena, J., and Torà, M.: Atmospheric circulation patterns associated with hail events in Lleida (Catalonia), *Atmos. Res.*, 100, 428–438, 2011.
- 635
- Barras, H., Hering, A., Martynov, A., Noti, P.-A., Germann, U., and Martius, O.: Experiences with > 50 000 crowd-sourced hail reports in Switzerland, *Bull. Amer. Meteor. Soc.*, 100, 1429–1440, 2019.
- Barthlott, C., Burton, R., Kirshbaum, D., Hanley, K., Richard, E., Chaboureau, J.-P., Trentmann, J., Kern, B., Bauer, H.-S., Schwitall, T., Keil, C., Seity, Y., Gadian, A., Blyth, A., Mobbs, S., Flamant, C., and Handwerker, J.: Initiation of deep convection at marginal instability in an ensemble of mesoscale models: A case study from COPS, *Quart. J. Roy. Meteor. Soc.*, 137, 118–136, 2011.
- 640
- Bedka, K. M.: Overshooting cloud top detections using MSG SEVIRI Infrared brightness temperatures and their relationship to severe weather over Europe, *Atmos. Res.*, 99, 175–189, 2011.
- Berthet, C., Dessens, J., and Sanchez, J. L.: Regional and yearly variations of hail frequency and intensity in France, *Atmos. Res.*, 100, 391–400, 2011.
- 645
- Bunkers, M. J., Klimowski, B. A., Zeitler, J. W., Thompson, R. L., and Weisman, M. L.: Predicting supercell motion using a new hodograph technique, *Wea. Forecasting*, 15, 61–79, 2000.
- Changnon, S. A.: The scales of hail, *J. Appl. Meteor.*, 16, 626–648, 1977.
- Chronis, T., Carey, L. D., Schultz, C. J., Schultz, E. V., Calhoun, K. M., and Goodman, S. J.: Exploring lightning jump characteristics, *Wea. Forecasting*, 30, 23–37, 2015.
- 650
- Dee, D. P., Uppala, S., Simmons, A., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M., Balsamo, G., Bauer, P., et al.: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system, *Quart. J. Roy. Meteor. Soc.*, 137, 553–597, 2011.
- Dennis, E. J. and Kumjian, M. R.: The impact of vertical wind shear on hail growth in simulated supercells, *J. Atmos. Sci.*, 74, 641–663, 2017.
- Dessens, J. and Fraile, R.: Hailstone size distributions in southwestern France, *Atmos. Res.*, 33, 57–73, 1994.
- 655
- Dickinson, M. J. and Knight, D. J.: Frontal interaction with mesoscale topography, *J. Atmos. Sci.*, 56, 3544–3559, 1999.
- Dotzek, N., Groenemeijer, P., Feuerstein, B., and Holzer, A. M.: Overview of ESSL’s severe convective storms research using the European Severe Weather Database ESWD, *Atmos. Res.*, 93, 575–586, 2009.
- Emanuel, K. A.: Frontal circulations in the presence of small moist symmetric stability, *J. Atmos. Sci.*, 42, 1062–1071, 1985.
- Fluck, E.: Hail potential over Western Europe, Ph.D. thesis, Karlsruhe Institute of Technology (KIT), 2017.
- 660
- Foote, G. B.: A study of hail growth utilizing observed storm conditions, *J. Clim. Appl. Met.*, 23, 84–101, 1984.
- Galway, J. G.: The lifted index as a predictor of latent instability, *Bull. Amer. Meteor. Soc.*, 37, 528–529, 1956.
- García-Ortega, E., López, L., and Sánchez, J.: Atmospheric patterns associated with hailstorm days in the Ebro Valley, Spain, *Atmos. Res.*, 100, 401–427, 2011.
- Gaiotti, D., Nordio, S., and Stel, F.: The climatology of hail in the plain of Friuli Venezia Giulia, *Atmos. Res.*, 67–68, 247–259, 2003.
- 665
- Graf, M. A., Sprenger, M., and Moore, R. W.: Central European tornado environments as viewed from a potential vorticity and Lagrangian perspective, *Atmos. Res.*, 101, 31–45, 2011.
- Groenemeijer, P., Púčik, T., Holzer, A. M., Antonescu, B., Riemann-Campe, K., Schultz, D. M., Kühne, T., Feuerstein, B., Brooks, H. E., Doswell III, C. A., et al.: Severe convective storms in Europe: Ten years of research and education at the European Severe Storms Laboratory, *Bull. Amer. Meteor. Soc.*, 98, 2641–2651, 2017.

- 670 Groenemeijer, P. H. and van Delden, A.: Sounding-derived parameters associated with large hail and tornadoes in the Netherlands, *Atmos. Res.*, 83, 473–487, 2007.
- Handwerker, J.: Cell tracking with TRACE3D - a new algorithm, *Atmos. Res.*, 61, 15–34, 2002.
- Heinselman, P. L. and Ryzhkov, A. V.: Validation of polarimetric hail detection, *Wea. Forecasting*, 21, 839–850, 2006.
- Hermida, L., López, L., Merino, A., Berthet, C., García-Ortega, E., Sánchez, J. L., and Dessens, J.: Hailfall in southwest France: Relationship
675 with precipitation, trends and wavelet analysis, *Atmos. Res.*, 156, 174–188, 2015.
- Hewson, T. D.: Objective fronts, *Meteor. Appl.*, 5, 37–65, 1998.
- Hohl, R., Schiesser, H. H., and Knepper, I.: The use of weather radars to estimate hail damage to automobiles: An exploratory study in Switzerland, *Atmos. Res.*, 61, 215–238, 2002.
- Holleman, I., Wessels, H. R. A., Onvlee, J. R. A., and Barlag, S. J. M.: Development of a hail-detection-product, *Phys. Chem. Earth (B)*, 25,
680 1293–1297, 2000.
- James, P. M., Reichert, B. K., and Heizenreder, D.: NowCastMIX: Automatic integrated warnings for severe convection on nowcasting time scales at the German Weather Service, *Wea. Forecasting*, 33, 1413–1433, 2018.
- Junghänel, T., Brendel, C., Winterrath, T., and Walter, A.: Towards a radar-and observation-based hail climatology for Germany, *Meteor. Z.*, 25, 435–445, 2016.
- 685 Kapsch, M.-L., Kunz, M., Vitolo, R., and Economou, T.: Long-term trends of hail-related weather types in an ensemble of regional climate models using a Bayesian approach, *J. Geophys. Res.*, 117, D15 107, 2012.
- Klemp, J. B.: Dynamics of tornadic thunderstorms, *Ann. Rev. Fluid Mech.*, 19, 369–402, 1987.
- Kunz, M.: The skill of convective parameters and indices to predict isolated and severe thunderstorms, *Nat. Hazards Earth Syst. Sci.*, 7, 327–342, 2007.
- 690 Kunz, M. and Kugel, P.: Detection of hail signatures from 3D C-Band radar reflectivity, *Atmos. Res.*, 153, 565–577, 2015.
- Kunz, M. and Puskeiler, M.: High-resolution assessment of the hail hazard over complex terrain from radar and insurance data, *Meteor. Z.*, 19, 427–439, 2010.
- Kunz, M., Blahak, U., Fluck, E., Punge, H., Schmidberger, M., Handwerker, J., Mohr, M., Mühr, B., and Bedka, K.: The severe hailstorm in SW Germany on 28 July 2013: Characteristics, impacts, and meteorological conditions, *Quart. J. Roy. Meteor. Soc.*, 144, 231–250, 2018.
- 695 Löffler-Mang, M., Kunz, M., and Kossmann, M.: A nocturnal cold front over complex terrain and its influence on the ozone concentrations at the ground, *Meteor. Z.*, 5, 308–317, 1996.
- Lukach, M., Foresti, L., Giot, O., and Delobbe, L.: Estimating the occurrence and severity of hail based on 10 years of observations from weather radar in Belgium, *Meteor. Appl.*, 24, 250–259, 2017.
- Madonna, E., Ginsbourger, D., and Martius, O.: A Poisson regression approach to model monthly hail occurrence in Northern Switzerland
700 using large-scale environmental variables, *Atmos. Res.*, 203, 261–274, 2018.
- Manzato, A.: A climatology of instability indices derived from Friuli Venezia Giulia soundings, using three different methods, *Atmos. Res.*, 67, 417–454, 2003.
- Manzato, A.: The use of sounding-derived indices for a neural network short-term thunderstorm forecast, *Wea. Forecasting*, 20, 896–917, 2005.
- 705 Markowski, P. and Richardson, Y.: *Mesoscale meteorology in midlatitudes*, John Wiley & Sons, Chichester, United Kingdom, 407 pp., 2010.
- Mason, B. J.: *The physics of clouds*, Oxford University Press, Clarendon Press, 671 pp., 1971.

- Mathias, L., Ermert, V., Kelemen, F. D., Ludwig, P., and Pinto, J. G.: Synoptic analysis and hindcast of an intense bow echo in Western Europe: The 9 June 2014 storm, *Wea. Forecasting*, 32, 1121–1141, 2017.
- Merino, A., Wu, X., Gascón, E., Berthet, C., García-Ortega, E., and Dessens, J.: Hailstorms in southwestern France: Incidence and atmospheric characterization, *Atmos. Res.*, 140, 61–75, 2014.
- 710 Mohr, S. and Kunz, M.: Recent trends and variabilities of convective parameters relevant for hail events in Germany and Europe, *Atmos. Res.*, 123, 211–228, 2013.
- Mohr, S., Wandel, J., Lenggenhager, S., and Martius, O.: Relationship between atmospheric blocking and warm-season thunderstorms over western and central Europe, *Quart. J. Roy. Meteor. Soc.*, 145, 3040–3056, 2019.
- 715 Mohr, S., Wilhelm, J., Wandel, J., Kunz, M., Portmann, R., Punge, H., Schmidberger, M., and Grams, C.: The role of large-scale dynamics in an exceptional sequence of severe thunderstorms in Europe May/June 2018, *Weather Clim. Dynam. Discuss*, 2020.
- Morris, R. M.: The Spanish plume-testing the forecaster’s nerve, *Meteorol. Mag.*, 115, 349–357, 1986.
- Mroz, K., Battaglia, A., Lang, T. J., Cecil, D. J., Tanelli, S., and Tridon, F.: Hail-detection algorithm for the GPM core observatory satellite sensors, *J. Appl. Meteor. Climatol.*, 56, 1939–1957, 2017.
- 720 MunichRe: NatCatSERVICE; natcatservice.munichre.com; accessed 04 May 2020, 2020.
- Ni, X., Liu, C., Cecil, D. J., and Zhang, Q.: On the detection of hail using satellite passive microwave radiometers and precipitation radar, *J. Appl. Meteor. Climatol.*, 56, 2693–2709, 2017.
- Nisi, L., Ambrosetti, P., and Clementi, L.: Nowcasting severe convection in the Alpine region: The COALITION approach, *Quart. J. Roy. Meteor. Soc.*, 140, 1684–1699, 2014.
- 725 Nisi, L., Martius, O., Hering, A., Kunz, M., and Germann, U.: Spatial and temporal distribution of hailstorms in the Alpine region: A long-term, high resolution, radar-based analysis, *Quart. J. Roy. Meteor. Soc.*, 142, 1590–1604, 2016.
- Nisi, L., Hering, A., Germann, U., and Martius, O.: A 15-year hail streak climatology for the Alpine region, *Quart. J. Roy. Meteor. Soc.*, 144, 1429–1449, 2018.
- Piper, D. and Kunz, M.: Spatio-temporal variability of lightning activity in Europe and the relation to the North Atlantic Oscillation teleconnection pattern, *Nat. Hazards Earth Syst. Sci.*, 17, 1319–1336, 2017.
- 730 Piper, D., Kunz, M., Allen, J., and Mohr, S.: Investigation of the temporal variability of thunderstorms in central and western Europe and the relation to large-scale flow and teleconnection patterns, *Quart. J. Roy. Meteor. Soc.*, pp. 1–23, 2019.
- Půčik, T., Groenemeijer, P., Rýva, D., and Kolář, M.: Proximity soundings of severe and non-severe thunderstorms in Central Europe, *Mon. Wea. Rev.*, 143, 4805–4821, 2015.
- 735 Půčik, T., Castellano, C., Groenemeijer, P., Kühne, T., Rädler, A. T., Antonescu, B., and Faust, E.: Large hail incidence and its economic and societal impacts across Europe, *Mon. Wea. Rev.*, 147, 3901–3916, 2019.
- Punge, H. and Kunz, M.: Hail observations and hailstorm characteristics in Europe: A review, *Atmos. Res.*, 176–177, 159–184, 2016.
- Punge, H., Bedka, K., Kunz, M., and Reinbold, A.: Hail frequency estimation across Europe based on a combination of overshooting top detections and the ERA-INTERIM reanalysis, *Atmos. Res.*, 198, 34–43, 2017.
- 740 Puskeiler, M., Kunz, M., and Schmidberger, M.: Hail statistics for Germany derived from single-polarization radar data, *Atmos. Res.*, 178–179, 459–470, 2016.
- Renard, R. J. and Clarke, L. C.: Experiments in numerical objective frontal analysis, *Mon. Wea. Rev.*, 93, 547–556, 1965.
- Rockel, B., Will, A., and Hense, A.: The Regional Climate Model COSMO-CLM (CCLM), *Meteor. Z.*, 17, 347–348, 2008.

- Ryzhkov, A. V. and Zrnic, D. S.: Polarimetric Characteristics of Deep Convective Storms, Springer *Atm. Sci.*, Springer, Cham. In: Radar polarimetry for weather observations, eBook, 2019.
- 745 Ryzhkov, A. V., Kumjian, M. R., Ganson, S. M., and Khain, A. P.: Polarimetric radar characteristics of melting hail. Part I: Theoretical simulations using spectral microphysical modeling, *J. Appl. Meteor. Climatol.*, 52, 2849–2870, 2013.
- Sánchez, J., Merino, A., Melcón, P., García-Ortega, E., Fernández-González, S., Berthet, C., and Dessens, J.: Are meteorological conditions favoring hail precipitation change in Southern Europe? Analysis of the period 1948–2015, *Atmos. Res.*, 198, 1–10, 2017.
- 750 Sánchez, J. L., Marcos, J. L., Dessens, J., López, L., Bustos, C., and García-Ortega, E.: Assessing sounding-derived parameters as storm predictors in different latitudes, *Atmos. Res.*, 93, 446–456, 2009.
- Schemm, S., Rudeva, I., and Simmonds, I.: Extratropical fronts in the lower troposphere—global perspectives obtained from two automated methods, *Quart. J. Roy. Meteor. Soc.*, 141, 1686–1698, 2015.
- Schemm, S., Nisi, L., Martinov, A., Leuenberger, D., and Martius, O.: On the link between cold fronts and hail in Switzerland, *Atmos. Sci. Lett.*, 17, 315–325, 2016.
- 755 Schemm, S., Sprenger, M., Martius, O., Wernli, H., and Zimmer, M.: Increase in the number of extremely strong fronts over Europe? A study based on ERA-Interim reanalysis (1979–2014), *Geophys. Res. Lett.*, 44, 553–561, 2017.
- Schuster, S. S., Blong, R. J., and McAneney, K. J.: Relationship between radar-derived hail kinetic energy and damage to insured buildings for severe hailstorms in Eastern Australia, *Atmos. Res.*, 81, 215–235, 2006.
- 760 Smart, J. R. and Alberty, R. L.: The NEXRAD hail algorithm applied to Colorado thunderstorms, in: Preprints, 14th Conf. on Severe Local Storms, Indianapolis, IN, USA, Amer. Meteor. Soc., 244–247, 1985.
- SwissRe: Sigma - Natural catastrophes and man-made disasters in 2014: convective and winter storms generate most losses, Tech. rep., Swiss Re Economic Research and Consulting, Switzerland, 2015.
- Thompson, R. L., Mead, C. M., and Edwards, R.: Effective storm-relative helicity and bulk shear in supercell thunderstorm environments, 765 *Wea. Forecasting*, 22, 102–115, 2007.
- Trapp, R. J.: Mesoscale-convective processes in the atmosphere, Cambridge, New York, USA, 2013.
- Trefalt, S., Martynov, A., Barras, H., Besic, N., Hering, A. M., Lenggenhager, S., Noti, P., Röthlisberger, M., Schemm, S., Germann, U., et al.: A severe hail storm in complex topography in Switzerland - Observations and processes, *Atmos. Res.*, 209, 76–94, 2018.
- Vinet, F.: Climatology of hail in France, *Atmos. Res.*, 56, 309–323, 2001.
- 770 Waldvogel, A., Federer, B., and Grimm, P.: Criteria for the detection of hail cells, *J. Appl. Meteor.*, 18, 1521–1525, 1979.
- Wandel, J.: Analyse der Umgebungsbedingungen während Hagelereignissen, Master’s thesis, Institute of Meteorology and Climate Research, Karlsruhe Institute of Technology (KIT), 101 pp., 2017.
- Wapler, K.: High-resolution climatology of lightning characteristics within Central Europe, *Meteorol. Atmos. Phys.*, 122, 175–184, 2013.
- Wapler, K.: The life-cycle of hailstorms: Lightning, radar reflectivity and rotation characteristics, *Atmos. Res.*, 193, 60–72, 2017.
- 775 Wapler, K. and James, P.: Thunderstorm occurrence and characteristics in Central Europe under different synoptic conditions, *Atmos. Res.*, 158, 231–244, 2015.
- Weckwerth, T. and Parsons, D. B.: A review of convection initiation and motivation for IHOP_2002, *Mon. Wea. Rev.*, 134, 5–22, 2005.
- Weisman, M. L. and Klemp, J. B.: The dependence of numerically simulated convective storms on vertical wind shear and buoyancy, *Mon. Wea. Rev.*, 110, 504–520, 1982.
- 780 Witt, A., Eilts, M. D., Stumpf, G. J., Johnson, J. T., Mitchell, E. D. W., and Thomas, K. W.: An enhanced hail detection algorithm for the WSR-88D, *Wea. Forecasting*, 13, 286–303, 1998.