Review for NHESS-2020-138, Revision 1

General Comments

The authors have made significant revisions to the paper based on the reviewers' comments, including provision of technical details on the French and German radar datasets and an assessment of the Froude number on hail days over the Massif Central using reanalysis data. I would like to thank them for these efforts. However, some of the new material was very difficult to understand (particularly the details of the radar tracking algorithm), while other parts have raised further questions that need to be addressed. I have a large number of additional comments; however, I expect most of these to be fairly simple to address. As such I am recommending further minor revisions.

The authors would like to deeply thank the reviewer for the further time and effort spent on the revision of the manuscript. We appreciate very much the detailed and helpful comments provided by the reviewer that helped for the understanding of the article. In the final version, we will address each comment and questions raised by the reviewer.

Specific Comments

L19-20: Hail can also be produced by single-cell "pulse" storms; however, large hail is almost always associated with organised convection, particularly supercells (Smith et al. 2012; Wapler et al. 2016)

We added this statement to the final version of the manuscript.

L23-24: A decrease in θ e with height indicates that the atmosphere is potentially unstable. According to Markowski and Richardson (2010), potential instability is not generally considered to play a role in the destabilisation of the atmosphere that precede convective initiation. The key instability for deep, moist convection is conditional instability, which exists when θ e * (the equivalent potential temperature that the environment would have if it were saturated) decreases with height. However, I don't think you need to get into these definitions; just note that conditional instability is needed for lifted parcels of air to become positively buoyant.

We changed lines 23-24 according to the reviewer's suggestion.

L28-30: Other studies you might mention here include Brooks et al. (2003), Johnson and Sugden (2014), and Taszarek et al. (2017).

We included these additional references to the final version.

L31-39: Since the Spanish Plume isn't mentioned again I don't think you need to discuss it in detail here.

We deleted the sentences related to the Spanish Plume.

L67-69: I know I requested you add the definition of overshooting tops (OTs), but it rather

breaks up the flow of this sentence. Also, this definition is relevant to the visual identification of an OT, but in satellite imagery these features are typically identified based on infrared brightness temperature signatures. As such I would suggest getting rid of the definition and instead including something like the following "...such as overshooting tops (an indicator of strong convective updrafts) in satellite imagery…"
We removed the definition of OTs and replaced it with "an indicator of strong convective updrafts" as suggested.

L70-71: The statement "the link between the observed quantities and hail occurrence at the surface is less reliable than using radar measurements" doesn't really make sense and in any case only really applies to satellite measurements, not model data. I would recommend making a more general statement in this sentence about these approaches, then noting the specific limitations of satellite- and model-based hail proxies in the subsequent sentences. The key benefit of radar proxies over overshooting tops is that the former is a more direct measure of hail within a storm (presence of large reflectivities), whereas the latter just indicates the presence of a strong updraft that might produce large hail.
We deleted the sentence "the link between the observed quantities and hail occurrence at the surface is less reliable than using radar measurements" and replaced it with the suggestion mentioned by the reviewer.

L73-75: I think there is an opportunity here to emphasise the unique nature of the hail climatology you present and, in doing so, better link these two paragraphs. You could note that the advantage of satellite- and model-based hail proxies is that they can cover a wide geographic area, whereas radar-based climatologies are typically limited to a single country or region. However, yours is, I believe, the first study to combine radar observations from multiple countries.
The authors followed this suggestion and made a transition between the two paragraphs.

L96-98: Does this mean that, for a given time, a CAPPI was taken from each radar and these were composited onto a common grid by taking the maximum value at every grid point? If so, at what altitude are the CAPPIs?
For each single radar, the MaxCAPPI is achieved by projecting the vertical maximum of the 3D radar data on a horizontal plane.

L101-113: I don't think you need to mention specific projects such as Panthère or RHyTMME, though you could add the relevant references (Tabary, 2007; Beck and Bousquet, 2013) to the end of the opening sentence of this paragraph. All you really need to note here is how many radars were present in the network and when these were replaced or upgraded.
We fully agree with the reviewer and removed the project names as well as their descriptions.

L117-119: It looks like complete coverage of Luxembourg is only provided by one of the German radars. Perhaps this discussion of the inclusion of Belgium and Luxembourg should be moved to the opening paragraph of section 2.1, since it relies on both French and

German data.
The German radar composite did not include Luxembourg. We moved the discussion after the history of the French radar composite.

L122-123: The dual-pol variables aren't used so no need to mention them here.
We removed the sentence with dual-pol variables.

L129-130: It's not obvious to me how a VPR (vertical profile of reflectivity) could be used to account for beam broadening with range. My understanding is that VPR corrections are used to extrapolate measurements taken aloft to the surface. This issue is worse at long range because the beam height increases (due to both the non-zero elevation angle and the curvature of the Earth); however, this has nothing to do with beam broadening, which tends to lead to an increase in partial beam filling with range.
Many thanks for pointing out a mistake. We mixed up two sentences. The VPR is used to correct the overestimation of reflectivity due to the bright band and to take the underestimation of reflectivity above the 0°C isotherm into account (Tabary et al., 2013); but the VPR is not used to correct beam widening. We changed the sentence in the manuscript.

L135-141: Is this discussion of QPE and associated quality control relevant since you are using reflectivity data only? Or do you take the corrected rain rates and convert them back to reflectivities? If not then all this detail can probably go.
We agree with the reviewer and deleted this part.

L154-155: As noted in my original review, these regions are covered, according to Fig. 2, but only by a single radar. In your reply, you say "Looking at long-term radar composites, the far north place near the Danish border and the southeastern part of Bavaria have no reflectivity values; we rather see some values near the location of the radar stations." This leaves two possibilities. Either the German radar composite excludes pixels that are covered by only a single site or both these regions happen to be affected by beam blockage. It would be good to know what the reason is.
We think that the southeastern part of Bavaria is affected by beam blockage, whereas the Danish border region is excluded from the German radar composite.

L160: Is orographic shading actually corrected for (e.g. by interpolating from higher tilts) or are pixels that are affected by it simply masked or labelled as low quality?
From our knowledge, the orographic shading is not corrected, but rather labelled as low quality in each single radar scan.

L160-161: Again is the conversion from Z to R relevant, since you're only using reflectivity?
We deleted "transformation of reflectivity Z to rain rate R".

L169-170: How do you get a "terrain-following near-ground reflectivity"? Are VPRs used to extrapolate values to the surface? Also it doesn't make sense to say CAPPI here since CAPPI stands for "constant-altitude plan position indicator". Here the "constant altitude"

typically refers to above radar level (ARL) or above sea level (ASL), rather than above ground level (AGL), which is what you'd have if the data really are "terrain-following".
Thanks for these precisions. We rephrased this sentence.

L204: You should say 0.25° resolution rather than 31 km as that is what you used (as I can see from Fig. 4). I think 31 km refers to the underlying Gaussian grid from which the 0.25° products are derived.
We agree with this statement and changed 31km for 0.25°.

L216-217: I'm not sure what you mean here; please rephrase.
We rephrased the sentence.

L224: How are you defining "high reflectivity" here?
Here, "high reflectivity" means a reflectivity value above 55 dBZ.

L236-239: I would suggest using Z MAX and Z RC to indicate the maximum reflectivity in an ROIP and the threshold for RCs, respectively.
We followed the suggestions of the reviewer and included the changes in the final paper version.

L239-240: Presumably this means that if an RC's maximum reflectivity briefly (say for one or two time steps) drops below 55 dBZ it will be treated as two separate convective cells, is that correct? If so, this is an important limitation and should be noted.
No, otherwise this would mean that the hailstreak lengths tend to be significantly underestimated.

L241-242: Change "surface area" to just "area" as the former implies a 3D object. Also, rather than "radar bins" I presume you mean pixels? In which case, this quantity should have units of km 2 , both here and in Table 1. I would suggest adding symbols to represent these variables; maybe A RC for the minimum RC area and A hail for the minimum area with reflectivities $\geq$ 55 dBZ. Finally, can you provide any justification for the inclusion of these two criteria? It makes sense to filter out very small cells as these could be spurious, but is there much sensitivity to the specific choice of thresholds?
We replaced "surface area" to just "area". In the same way, we changed "radar bins" to "pixels". We added km^2 units to "pixels" in the article and also included this unit to Table 1. We will not add further symbols here as we only mention once the RC minimum area and the minimum area with reflectivity $\geq$ 55 dBZ.
After testing few thresholds for the RC minimum area, including the overlap of tracks on hail observations, the threshold of 5 km2 was the best fit for our study. Area inferior to 5 km2 leaded to too many non-convective cells as well as too many cells splittings, whereas RC with an area superior to 5 km2 gave us fewer tracks. We chose a minimum of 3 pixels with Z$\geq$ 55 dBZ in the RC, so that more than half of the RC is covered by non-spurious data.

L251: FAR = b / (a + b) is the false alarm *ratio* , whereas the false alarm *rate* (also known as

the probability of false detection) is defined as F = b / (b + d). It seems this mistake was present in Puskeiler et al. (2016) but must have been missed by that paper's reviewers. You could maybe add a note to this effect: "...the false alarm ratio (FAR; incorrectly labelled the false alarm *rate* in Puskeiler et al. 2016) was 0.4..."
Many thanks for notifying us about this wrong definition. We wouldn't add "...the false alarm ratio (FAR; incorrectly labelled the false alarm *rate* in Puskeiler et al. 2016) was 0.4..." explicitly in the article. As noted in Wilks (2011), we quote: "The FAR has been called the false alarm **rate** (Barnes et al. 2009 sketch a history of the confusion) [..]".

L253-254: Change "The algorithm assigns the RC of the previous radar composite to the actual composite" to "The algorithm associates RCs between consecutive radar composites".
Changed.

L255-256: What specifically are the similarity criteria?
We reformulated this sentence.

L257: I'm guessing that the "velocity factor" is intended to account for uncertainties in the motion estimates. Is there any reason for using 0.6?
Right, the velocity factor is intended to account for uncertainties in the motion estimates. After testing several factors, 0.6 was the best fit in our study. The velocity factor of 0.6 was also used in Handwerker (2002).

L258-260: The right-moving storm tends to be favoured in the northern hemisphere, whereas the left-moving storm is favoured in the southern hemisphere (due to mirrored shear profiles). However, I'm not sure that it's fair to say that "in most cases, the left-moving cell weakens very quickly". It may sometimes evolve into more of a multicell structure, while in unidirectional shear profiles both storms may persist as supercells. Furthermore, it's important to note that cell splitting may also occur due to changes in storm intensity that cause a single RC to break up (or vice versa in the case of cell merging).
We rephrased the sentence including "left-moving cell" and added a comment about the cell splitting.

L262-268: I'm afraid I really don't follow this explanation at all. It seems to imply that single cell associations where there is a significant change in cell area can be labeled as splits or mergers, but that doesn't make sense. A split would normally be defined where a cell at time t can be associated with two or more cells at time $t + \Delta t$. In this case the choice has to be made as to whether all or only one (e.g. the largest or most intense) of the "child" cells inherit the history of the "parent" cell. Similarly, a merger is defined where multiple cells at time t are associated with a single cell at time $t + \Delta t$. In this case, a choice has to be made how to assign a history to the child cell. I suggest you completely rework this description (possibly adding a schematic) to make it easier to understand.

We apologize for the unclear explanations, and replaced/rephrased this part with the reviewer's suggestions. For the schematic, we kindly invite the reviewer to look at Fig. 5 in Handwerker (2002).

L276-277: Motion vectors include both speed and direction, so I don't think you need to say "as well as the track direction of the convective cells".
We deleted this part from the sentence.

L279: If I am understanding this correctly, the shift vector is a spatial increment rather than a velocity. In this case you shouldn't use du and dv, but rather dx and dy. If it *is* a velocity vector then the components should be u and v.
No, the shift vector is a velocity with the unit m/s.

L281-287: Again, I found this explanation really hard to follow. I tried reading the equivalent explanation in Puskeiler et al. (2016), but that is equally perplexing. In my head, the way this type of procedure would work would be to take a cell at two consecutive times and then shift the early cell forward in time and the later cell backward in time and average the two. This would be done for multiple intermediate time steps in order to create a smooth track. However, it sounds like the procedure used here is considerably more complicated. Again, I think this needs a complete rewrite in order to make it comprehensible.
We agree with the reviewer that our explanations weren't clear enough. In fact the algorithm does not store each cell, but only record the track and the maximum reflectivity values. Thus, we decided to completely rewrite this part and to use some of the terms evoked by the reviewer. Equation 3 was removed to avoid confusions and to simplify the explanations.

L318: Presumably by "hail days" you mean hail days within the subdomain shown in Fig. 4 as opposed to anywhere in your analysis domain. Please clarify in the text.
Thanks. Changed.

L320: Again, you should say 0.25° rather than 31 km.
We changed 31km for 0.25°.

L322: Here and elsewhere in this section you should change "hail days" to "hail days per year"
Good point, thanks for that. We changed all the relevant "hail days" to "hail days per year".

L326-327: Is this flow convergence perhaps the signature of a surface cold front or pressure trough? The circulation associated with a front/trough might favour the development of severe storms (through the associated generation of mid-level instability and advection of low-level moisture), but flow interactions with the Massif Central could still act as a focussing/initiation mechanism.
Without further investigations, it is still too early to affirm if the flow convergence is a signature of cold front or pressure trough.

L337: Surely the winds should be averaged over the lowest 1200 m to be consistent with the definition of H?
Decisive for flow effects evolving over complex terrain is not only the terrain height, but also the layers above. Therefore, the vertical (density-weighted) integration differs from the characteristic mountain height H. In the article, H is still set to 1300m. We changed the altitude of the Livradois mountains in the text.

L338-341: The definition of the Brunt–Väisälä frequency is pretty standard so I don't think you need to cite Huschke (1959). However, I would recommend that you modify the equation so that it is an expression for N rather than N 2 . Also, you can just use θ v for the virtual temperature (you don't need to refer to it as "ambient").
We deleted the citation of Huschke (1959) and changed θ va for θ v in the Brunt Vaisala equation. We also modified the equation and used N rather than N^2.

L345-346: You might want to note that this deflection of the flow is unlikely to show up in Fig. 4 due to the fairly coarse resolution of the ERA5 reanalysis (the Massif Central will be much lower and smoother in the IFS than is shown in the figure).
We added this thoughtful statement to the manuscript.

L350-353: This sentence is overly long and should be rephrased. You could potentially get rid of the bit starting "by referring to"
We rephrased the sentence and followed the reviewer suggestion.

L390-395: The Froude number that you obtain is actually a bit larger than that estimated by Kunz and Puskeiler (2010), even though they used a smaller value of H (1000 m). Presumably then, your values of U are larger and/or your values of N are smaller. Can you comment on these differences?
There can be few reasons why the Froude Number we found is higher than in Kunz and Puskeiler (2010). First, the time period investigated is not the same: We calculated the Froude Number for 2005 until 2014, whereas Kunz and Puskeiler (2010) focused on 1997 up to 2007. Then, in this article, U is coming from era5 reanalysis whereas Kunz and Puskeiler (2010) used soundings data. After that, the region considered in this article does not include for example the Vosges mountains or the entire Rhine Valley. This might also explain the differences in U.

L411: I would suggest changing "is partly caused by" to "show an association with" as causality has not been firmly established.
We followed the reviewer suggestion.

L415: I've checked an Puskeiler et al. (2016) don't actually examine interannual variability in hail frequency. Nisi et al. (2018) do, but they find much higher interannual variability compared to this study. Perhaps this reflects the much larger study domain considered here.
Thanks for that. We deleted the reference Puskeiler et al. (2016).

L430-431: You don't show maps for 2012 so this sentence can probably be deleted.

Sentence deleted.

L438-440: I'm not sure this last sentence is needed. Given the size of your study domain it is hardly surprising that years with a below-average number of hail days overall could still have a few localised high-impact events. Perhaps you could simply rephrase what you have to make this point. Also 2013 is only slightly below average and the average is arguably dragged up by the anomalously high number of hail days in 2006.
We deleted the last sentence of the paragraph.

L451: What you show in Fig. 7 isn't a 10-day moving average as the averaging windows don't overlap. Instead you could say that you calculate the average number of hail days for consecutive 10-day periods. Alternatively, given the short length of your climatology, you might consider just plotting the relative frequency of hail in each month. This might also make for an easier comparison with previous studies.
The sentence "10-day moving average" was replaced by "average number of hail days for consecutive 10-day periods" as suggested. This replacement was also done everywhere else. Concerning the relative frequency, we understand the concern of the reviewer. Note that many authors focusing on hailstorms in Europe do use the number of hail days per year or the total number of hail days per grid points. (Puskeiler, 2013; Puskeiler et al. 2016; Lindloff, 2003 ; Kunz and Puskeiler, 2010; Lukach and Delobbe, 2017 among others).

L468-469: Can it really be argued that these are distinct maxima? This could just be an artefact of the relatively short length of your climatology.
Note that this statement is only correct for the period 2005 to 2014.

L473-474: It doesn't make sense to describe this as "a right skewed distribution" since you're not really talking about a distribution but a time series.
We removed "a right skewed distribution" from the manuscript.

L479: Which of the aforementioned studies used hail pad data? What data sources did the other studies consider? And why would the hail pads being clustered "near the subdomains influenced by maritime air mass" lead to an earlier seasonal peak?
All of the aforementioned studies used hailpad data. The hailpad network of ANELFA is located in Southwestern France, a region particularly influenced by maritime air masses.
We deleted "Furthermore, the scattered network of hail pads is denser near the subdomains influenced by maritime air mass" from the manuscript.

L484: It's not really right to say "confirmed" here as the Lukach and Delobbe (2013) study obviously came before yours and showed an earlier maximum. To me the seasonal cycle for subdomain BEL looks pretty flat, but this is because you have plotted the number of hail days rather than the relative frequency.
We replaced "confirmed" by "also found".

L486: Since you only consider the first time that CCTA2D detects the cell this analysis

pertains more to the development of hail storms, rather than their overall diurnal cycle. As stated in my original review, it would make more sense to consider all times when a storm exceeded 55 dBZ as this would account for storms that persist for multiple hours. This would also make comparisons with previous studies easier, since I imagine most of these considered all hours with hailstorms, rather than just the hour in which storms developed.
We replaced "diurnal cycle of hailstorms" by "development of hail storms". We agree with the reviewer that to consider all times would permit to reconstruct the full diurnal cycle. Nevertheless, the main focus here was to compare when the first signatures of cells appear between all regions (When does the first signatures appear in mountainous area compared to regions near coastlines?). Note that the first signatures of cells detected by the algorithm allowed to study the synoptic environments favorable for hail development (Fluck, 2018).

L492: There's no need to keep reminding the reader where each region is located as this has already been stated and is shown in Fig. 3. The same comment applies to L495-496 and L498.
We deleted the region locations.

L495-496: The peak at 16 LT is only slightly above the values for the adjacent hours. Given the relatively short length of your climatology I'm not sure you can read too much into this difference.
We added "slightly" to the sentence.

L510-511: The peak at 18 LT is much later than what you and most other studies find. Can this difference be explained?
We replaced the sentence with the peak at 18 LT with "In this study we found a peak of hail around 16~LT in subdomain SWF, while Malafre, 2009 established a peak later in the afternoon, around 18~LT". Note that only the hail seasons of years 2004 and 2005 in the Ebro Valley were considered in the study of Malafre, 2009. The different regions and time period analyzed in Malafre (2009) and our study might explain the shift in time of the hail maximum.

L512-531: I'm really not convinced that this section adds much, if anything, to the manuscript. The distribution of track onset locations seems pretty consistent with the overall distribution of hailstorms shown in Fig. 3, with fewer points overnight and during the morning and more points during the afternoon and evening (as one would expect from Fig. 8). Unless you can quantitatively show that some regions show a *disproportionately* high/low onset frequency for a given time (i.e. many more/less onsets than one would expect based on overall hailstorm frequency) I would suggest getting rid of section 4.4 altogether.
We agree with the reviewer and deleted section 4.4 altogether.

L513-515: Get rid of "and to distinguish between mechanisms triggering nighttime events and convection being triggered within the boundary layer occurring preferably in the afternoon and early evening" - it makes the sentence overly long and isn't needed.
We actually think that this sentence explains why we investigated the spatial distribution of hailstorms.

L518: Hours less than 10 should be written as 00, 03, etc. The same comment applies to L519, L522, and 530.
We added two digits to the hours less than 10.

L540: I would say "a length less than 10 km". There's no need to include the symbol "L" if You're not going to include it in an equation.
We removed "L" from the sentence.

L541-542: Figure 10 only shows track lengths up to 100 km so perhaps the values you quote here should correspond to track lengths of 20–100 km and > 100 km, rather than 20–150 km and > 150 km.
We agree and quote the lengths according to the reviewer suggestion.

L559-561: In my original review I suggested that it might make more sense for orientation to be computed either as the angle of a line connecting the first and last points in the track or by fitting a line of best fit to all points in the track. In your response you argued that these methods "fit for straight (e.g. undeviated) swaths only". However, your method still does not account for curved storm tracks since it only considers a single pair of points in the track (before and after the centre point). Furthermore, it is likely to be more sensitive to sudden changes in cell direction associated with splits/mergers or changes in cell area. Assuming you have the start and end positions of the cell track I would suggest using these to compute the orientation as it is simply but consistent with how track length is defined.
#In the end, only the center of a swath was available for plotting. Sorry about that. We do like the ideas about the computation of the orientation very much, and might consider them in future investigations.

L563: Technically a west-to-southwest direction would be from 225 to 270°. As such I would just say "from between 200 and 260°."
We deleted "a west-to-southwest direction" from the sentence.

L576-578: This sentence needs reworking. First, you should change "none or only several" to "only a few". Second, "along the European coastlines" isn't very specific and is arguably repetition since Brittany and north Germany could be classed as "along the European coastlines". Third, I would say "farther inland" rather than "far off the coasts" as the latter implies offshore. Finally, you should quote hail frequencies for both coastal and inland regions in days per year.
We considered most of these suggestions in the conclusion.

L580-581: "The high spatial variability in the number of radar-derived hail days and the increasing number around orographic structures…" – I'm not sure what you mean by this; please rephrase.
We removed the first part of the sentence and replaced the second part by "The large number of track onsets around orographic structures..".

L583: Is the diurnal cycle of hail that different between different regions? It's hard to say from Fig. 8 because it plots the absolute number of hail days rather than the relative frequency.
We replaced "significant" by "some".

L603-604: X-band radars are actually less suitable for hail detection because their signal is strongly attenuated by large precipitation particles. You might instead mention the use of dual-polarisation measurements, which can provide more accurate detection of hail (e.g. Heinselman and Ryzhkov, 2006).
We deleted "the recently installed X-band radars in the French Alps" and replaced it with "the use of dual-polarisation measurements" and added the reference indicated by the reviewer.

Figure 1: In my previous review I requested that you use different line thicknesses or colours for country and region/state borders, so that readers less familiar with European geography can distinguish between the two; however, this change does not appear to have been made.
We included the wrong figure in the revised version. Sorry about that. Country borders appear now with a thicker line in the final version of the manuscript.

Figure 2: Since the X-band radars aren't included in your analysis they probably should be removed from this figure. You might also consider using different symbols for those radars that were replaced or upgraded to dual-polarisation during the study period.
We removed the X-band radars from the figure. We did not use different symbols for the dual-polarisation radars for visibility purposes.

Figure 3: As noted in my original review, a colorbar would be more appropriate than the individual blocks with value ranges, since these imply gaps (e.g. between 0.6 and 0.7 day per year). I would also suggest using an increment of 0.5 days per year as this is much more intuitive (0.5 days per year = once every two years). One more thing. In transforming the projection of this plot, the boxes defining the different regions appear to have become distorted, such that their edges don't properly line up. Can this be rectified?
Thanks for these suggestions and for pointing out some misunderstandings. There are actually no gaps between the value ranges. For example the first interval ranges from 0 up to 0.699 hail days, even "0-6" is labelled in the legend. To avoid confusions, we added an additional digit to the upper limit of each interval in the legend.
We have tried to plot this figure in many ways (with/without colorbars or individual blocks; and different increments of hail days). The authors have agreed that the increment of 0.6 hail days fits the best for the visualization of the regions affected by hail.
We fixed the projection issue of the boxes.

Figure 4: It looks like you have two coastlines in this map - can you get rid of the coarse one? Again, I would suggest using an increment of 0.5 days per year for the hail frequency contours (values of 2.5, 3.0, and 3.5 days per year). Finally, could you add a box showing the area for which the Froude number was evaluated?

Good points! We deleted the coarse coastline. Contours with an increment of 0.5 days sometimes overlapped on each other. Therefore we decided not to touch the contours.
Furthermore, we wanted to highlight the exact location of the high hail frequency (3.9) in the Massif Central region.
Finally, concerning the box showing the area for which the Froude number was evaluated: The upper left corner of the box would overlap the legend and "break up" the wind vectors flow. To avoid any confusion, we decided not to represent the box in the Figure.

Figure 6: Again, use a colorbar rather than the individual blocks for each colour range (with increments of 1 day), if possible.
#Same comment as for Figure 3.

Figure 7: The shading under the curves isn't needed so I suggest getting rid of it. The same comment applies to Fig. 8. I would also shift the curves so that points are centred on each 10-day window, rather than at the end of the window.
We removed the shading under the curves for Figures 7 and 8. For figure 7, we centered the curves on each 10-days time window.

Figure 8: The results in this and the previous figure might be better presented as relative frequencies for each month/hour rather than the absolute number of hail days, as this will allow you to better infer differences in the seasonal and diurnal cycles of hail between the five regions. Just a suggestion.
We appreciate the suggestion very much, but we will keep the absolute number of hail days for now.

Figure 9: Hours less than 10 should be written as 00, 03, etc.
We rewrote the hours less than 10 with two digits.

Figure 11: My first impression was that this figure, and the ones before and after, show relative frequency of the y axis. However, for this one the values for the first four bars alone add up to more than 100 %. Is this an error?
Thanks a lot for this remark. Indeed there was an error in Figure 11 where the year 2014 was counted twice. We fixed that. Figures 10 and 12 are fine (we checked them again).

Technical Corrections

L4: Change "reflectivity radar data and lightning data" to "radar reflectivity and lightning data".
L8: Change "or" to "and".
L20: It should be "mesoscale convective *systems*" not "mesoscale convective *storms*".
L22: "subtle" not "subtile".
L30: It should be "aside from" rather than "aside of".
L94-95: "that requested some computation adjustment into the national radar composite" - I'm not sure what you mean here; please rephrase.

We change it for “that requested some additional time to calibrate each X-band radar and to implement their data into the national radar composite”.

L95: Change “up to nowadays” to “onward”. Also, should it be “2015 onward” rather than
“2014 onward”, since data were available for 2014?
L219: Change “Despite” to “Although” and “has included” to “includes”.
L240: When referring to RCs you should use the pronoun “an” rather than “a” since the initialism is read as *arr-see* . There are many other occurrences in section 3.3 where this needs to be corrected.
L245: Change “BZ” to “dBZ”.
L260: Change “after they have been splitted” to “after they have split” or “following the split”.
L290: It should be “Kunz et al. (2020)” not “(Kunz et al., 2020)”.
L293-294: I know what you’re trying to say here but “prevail” isn’t the right word. The key point is that there are far fewer hail reports in France, Belgium, and Luxembourg than in Germany.
L318: Get rid of “(in terms of speed and direction)”.
L323-324: Get rid of “named” in both sets of parentheses.
L325: Change “northern” to “southerly” (wind directions refer to the direction from which the
wind is blowing).
L355: “require s”
L357: Get rid of “referred to as”. The same change should be applied on L377, 387, L401, and L404.
L363: “1424 m” (get rid of the period).
L364: Changed “mentioned as” to “labelled”. The same change should be applied on L388.
L365-366: You should put “e.g.” at the start of the list of references rather than having “among others” at the end.
L384: This shouldn’t be a new paragraph.
L427: This shouldn’t be a new paragraph.
L437: Change “entire” to “all of”.
L458: Change “Pyrenean” to “Pyrenees”.
L472: Get rid of “located in the very southwest of France” (this is obvious from the name).
L483: Change “upper western part of Belgium” to “much of Belgium”.
L499: Change “reminds” to “remains”.
L542: Get rid of the comma after “MCSs”.
L548: This shouldn’t be a new paragraph.

We are very thankful for the technical corrections that have all been included in the final version of the manuscript.

References

Brooks, H.E., Lee, J.W. and Craven, J.P. (2003) The spatial distribution of severe thunderstorm and tornado environments from global reanalysis data. Atmospheric Research, 67, 73‑94.
Heinselman, P.L. and Ryzhkov, A.V. (2006) Validation of polarimetric hail detection. Weather and Forecasting, 21, 839‑850.
Johnson, A.W. and Sugden, K.E. (2014) Evaluation of sounding-derived thermodynamic and wind-related parameters associated with large hail events. Electronic Journal of Severe Storms Meteorology, 9, 1‑42.
Smith, B.T., Thompson, R.L., Grams, J.S., Broyles, C., and Brooks, H.E. (2012) Convective modes for significant severe thunderstorms in the contiguous United States. Part I: Storm classification and climatology. Weather and Forecasting, 27, 1114‑1135.
Taszarek, M., Brooks, H.E. and Czernecki, B. (2017) Sounding-derived parameters associated with convective hazards in Europe. Monthly Weather Review, 145, 1511‑1528.
Wapler, K., Hengstebeck, T., and Groenemeijer, P. (2016) Mesocyclones in Central Europe as seen by radar. Atmospheric Research, 168, 112‑120.

The authors appreciate the additional references that were included in the final version.