Responses to Reviewer 1

Recommendation: Major Revisions

Summary

This study investigates the occurrence of a supercell in Bulgaria first by providing an analysis of the event itself and then by examining the predictability of the event by analyzing forecasts from the ECMWF as well as convective-allowing WRF ensemble. Other than supercells apparently being a rare event for Bulgaria, there is nothing particularly novel about this study. The overview of the event is basic. It consists of an analysis of the environmental CAPE and shear as well as radar imagery that shows a weak Doppler couplet, a hook echo, and a weak V-notch signature. There is also an analysis of lightning flash rates and how the number of flashes corresponds to changes in storm intensity. This was interesting, and admittedly outside my area of expertise, but from the numerous citations it seems like many previous studies have already shown that relationship.

We would first like to thank the reviewer for providing feedback on this article and hope that their specific comments have been addressed in a satisfactory manner. One of the driving motivations behind conducting this study is the continuously growing number of severe convection events in the southeastern parts of Europe, evidence for which can be found in the recently published climatology Taszarek et al. [7]. Understanding the dynamics and predictability of severe weather outbreaks is crucial for the adaptation of our country to an ever-changing climate and building a ‘weather-ready’ nation.

Furthermore, despite the absence of a formal supercell climatology, unofficial statistics suggest that supercell occurrence in Bulgaria is not very frequent, with annual numbers ranging only between 1 and 5. The relatively recent operational use of Doppler radars in Bulgaria means that isolated supercell events, such as the one presented in this paper, provide unique opportunities to gain insights into their morphology and evolution. Amongst other things, the detection of a V-notch signature for the first time since Doppler radars became operational in the Hail Suppression Agency (HSA) is an important motivation to analyse the 15 May 2018 case in particular.
Due to large computational expenses, the majority of previous work has relied on deterministic forecasts to study the dynamics and predictability of supercells [3,8,9]. Nevertheless, Miglietta et al. [8] concludes that “an ensemble approach (even a ‘poor man’ ensemble, as the one shown here) appears absolutely necessary to provide some indication on the risk of localized severe convective weather” on account of the considerable discrepancies between their supercell simulations. In direct response to these conclusions, our work utilizes a large (51-member) high-resolution (1-km) real-time convection-allowing ensemble system to study the predictability of the 15 May 2018 supercell case. Employing such a model configuration is perhaps one the most novel aspect of this work; operational CAM systems usually consist of only 10 or 20 ensemble members and can only afford to run at horizontal grid spacings of 3-km [10,11]. In addition, they usually have to rely on initial and lateral boundary conditions from several different modelling systems, which can introduce different systematic biases (refer to Section 3.4.3 in the manuscript) and complicate the overall interpretation of the forecast results. Instead, the WRF ensemble configuration employed in this work relies on a single modelling framework, which ensures that the initial condition perturbations are consistent with the underlying model dynamics. The findings of our study confirm the utility of a large CAM ensemble, especially in the presence of model errors. In particular, there exists a small subset of ensemble members in which the evolution of the simulated supercell is closer to the verifying radar observations. Back-of-the-envelope calculations show that that the likelihood of missing out on these members in a small ensemble is high. Furthermore, the behaviour of the ensemble results also allows us to critically comment on the adverse impact that model errors have on the quality of the ensemble forecasts.

Note that most of the aforementioned remarks have been added to the updated version of Section 1.

The second part of the paper focuses on determining the predictability of the event. The predictability of the larger scale environment was based solely on a combined cape-shear parameter from the ECMWF model. Short-term predictability was analyzed by investigating simulation results from a convective allowing WRF ensemble. The WRF simulations seem to have been judged almost entirely on updraft helicity. The analysis of both the ECMWF and WRF simulations is too simplistic and I have provided more detail in my Major Comments.

We thank the reviewer for these remarks and have addressed them in the
specific comments below.

The paper concludes with an analysis of the errors associated with WRF simulations, however this analysis seems to be almost entirely conjecture. The only analysis of the WRF simulations presented consists of UH probabilities (Figs. 11 – 14, Fig. 17), a panel plot of simulated radar reflectivity from a single ensemble member (Fig. 16), and a plot of maximum UH from the ensemble members (Fig. 15). The authors’ claim to have performed a “detailed analysis of the CAM simulations” (lines 608-609) and conclude that the main sources of error include things such as spurious convection limiting the supply of buoyant air to the supercell and errors in the supercell track. None of this has been explicitly shown in the paper.

The determination of the model errors in our study has been justified in the specific comments below.

Major Comments

1. The introduction/literature review needs to be cleaned up a bit. Some of the statements are vague and I believe some key citations are missing (see specific comments below).

We appreciate the reviewer’s comments regarding Section 1 and have made our best to incorporate their suggestions.

2. Validation of the long-range ECMWF forecasts is questionable due to the use of the CAPE-Shear parameter. I recommend either replacing or simply supplementing this discussion with plots of CAPE and shear individually (similar to Figure 2). You’ve mentioned that you’re following the ingredients-based approach outlined by Doswell. In order for a supercell to occur, a certain amount of wind shear must be present, along with some degree of instability. A composite parameter like cape-shear doesn’t tell you how much wind shear is present (or forecasted to be present). A large cape-shear value may indicate large instability and shear that may be too weak to support the supercell mode.

The ECMWF medium-range forecast is evaluated by looking at an appropriate ensemble product for forecasting severe weather - the Extreme Forecast Index (EFI). The EFI compares the forecast Cumulative Distribution Function (CDF) with the model climate CDF in order to provide information about the extremity of the forecast. Defining particular thresholds for convective parameters is a bit risky and we avoid this by using the EFI. Nevertheless, we agree with the reviewer that the composite
nature of the CAPE-shear parameter means that it could reach high values with high CAPE and somewhat weaker deep-layer shear. In order to highlight that wind shear played a significant role in this case, we compared both operationally available convective EFI parameters: the CAPE EFI and CAPE-shear EFI. This comparison reveals that the former is much lower than the latter, indicating that there was a substantial deep-layer shear contribution to the CAPE-shear parameter. Note that the probability of deep-layer shear exceeding 20 ms\(^{-1}\) has been also added as part of Figure 10.


3. The WRF ensemble discussion is also suspect due to the very low updraft helicity threshold. The 10 m\(^2\)s\(^{-2}\) updraft helicity threshold is an incredibly low value for your 1-km grid spacing configuration. Numerous studies have suggested much higher values (e.g., Kain et al. 2008; Sobash et al. 2011; Naylor et al. 2012).

We thank the reviewer for pointing out that our UH threshold values that are lower in comparison to those used in previous CAM studies. The reason for choosing 10 m\(^2\)s\(^{-2}\) in the original analysis is that it provided a good subjective depiction of where severe convective activity is likely to occur. The new version of the manuscript uses a UH threshold value of 50 m\(^2\)s\(^{-2}\) in order to be consistent with the Kain et al. (2008) study. Importantly, the interpretation of the 3 WRF simulations with the new UH threshold (cf. Figures 11-14) remains the same and does not change the conclusions regarding the predictability of the event in the WRF ensemble.

4. Based on the results shown, it is not clear that any of the simulations produced a supercell as claimed. The model is producing convection in the proper location, but I don’t see evidence that it could be interpreted as a supercell. Potential remedies include recreating Figs. 11-15 with a more appropriate threshold value, or creating an updraft helicity swath from the best performing member.

As explained in major point 3, the UH threshold used in the new NEP calculations has been increased to 50 m\(^2\)s\(^{-2}\). Although the NEP UH values in the updated Figures 11-14 are lower compared to those based on the old threshold, they still indicate the presence of supercells in our CAM simulations.

Further evidence for the existence of supercells can be found by following the study of Naylor et al. (2012), who found that a UH threshold suitable for the detection of supercells on a 1-km model grid is 180 m\(^2\)s\(^{-2}\). In accordance to these findings, we use
the ensemble maximum of 2-5 km UH and filter any values that fall below the aforementioned threshold value. Figure R4 displays the resulting filtered fields at 12:15 UTC and shows numerous convective cells in which this threshold value has been exceeded. In fact, the UH values in some of the convective cells go above 350 m$^2$s$^{-2}$ and suggests the presence of strongly rotating storm updrafts. One of these well-pronounced supercells belongs to the best performing member 30.

Third evidence for the presence of supercells in our WRF ensemble can be gathered by a subjective assessment of storm motion in different ensemble members. A careful look at the WRF ensemble reveals several instances in which the storms deviate from the mean N-NE 0-6 km wind. Once again, the most pronounced deviation in storm motion happens to member 30. The propagation of the supercell in this forecast is very close to the right-mover Bunker’s vector calculated from ECMWF’s ensemble mean forecast (discussed starting from L503 in the manuscript).

![Figure R4](image)

**Figure R4.** Ensemble maximum of hourly maximum 2-5 km UH between 12:45 UTC from WRF.15may00utc.

5. There is really no supporting evidence for your three main sources of error in the CAM simulations.

We would thank the reviewer for raising the important point of determining model errors in our CAM simulations. In our conclusions, we have identified three main sources of errors: (i) timing of CI; (ii) spurious convective activity and (iii) deviation of the simulated supercells from the observed track. Determination of these errors is based on comparing the CAM forecasts to the available observations:
The timing of CI is one of the primary differences between the three WRF ensemble presented in our study. Although all of them tend to initialize convection too early, the timing error is especially pronounced in the WRF.14may00utc and WRF.14may12utc experiments. Nevertheless, the WRF.15may00utc forecasts also suffer from a similar problem. For example, the relatively high NEP UH values in Figure 11c indicate the presence of rotating updrafts in some of the ensemble members, while in reality the supercell has just started to form (Figure 11d).

Evidence for spurious convective activity is most readily found in Figure 16, which shows results from the best performing member 30. Although the initial evolution of the supercell is simulated fairly accurately, the presence of additional convective activity to the southwest of the supercell leads to its premature decay. Note that such excessive convective activity can be found in most of the other ensemble members and is a common feature of many CAM systems (e.g., [12]).

There are numerous places where errors in the track of the simulated supercells show up, e.g. the comparison of NEP UH and radar observations in Figures 13c and 13d and the deviation of most supercells from the one simulated in member 30 (cf. Figure R2)\(^1\).

It is worth pointing out that the aforementioned list of model errors is by no means exhaustive. In particular, there may be errors in the model environment that would require more comprehensive analysis tools to detect. Our intention in this article was to simply highlight the importance of model errors in reducing the predictability of the event and also comment on how they might adversely affect the probability of detection in a typical CAM-based ensemble system.

Specific Comments

Lines 34-36. The first few sentences of a paper usually introduce the topic as well as provide information as to why this is an important topic to study. Certainly, severe thunderstorms are worth studying because of their societal impacts, but “Sometimes they are associated with loss of life and major damage” is a particularly vague opening statement. Are there any estimates as to the estimated economic cost per year? How many fatalities or injuries have occurred over the last 10,20, or 30 years?

\(^1\) Here we assume that the supercell track in member 30 is a good representation of the truth.
We thank the reviewer for these questions. Currently, there are no official studies indicating the number of losses/injuries/fatalities due to convective storms over the whole European continent. If we look at the Munich Re’s NatCatSERVICE, a database for analysing and evaluating natural catastrophes (https://www.munichre.com/en/reinsurance/business/non-life/natcatservice/index.html), we will see that the overall losses due to convective storms over West Europe are approaching 50 bn US dollars in the past 20 years with more than 500 fatalities in their database (Figure R5).

**Figure R5.** Overall and insured losses in US$ for convective storm events in West Europe 1998 – 2018.
Figure R6. Percentage distribution for convective storm events in West Europe 1998 – 2018.

In Bulgaria, which is the geographical area of our study, the mean annual number of lightning-induced fires with different range of damages is 63 during the last 10 years (period 2009-2018) according to annual reports of Fire Safety and Civil Protection General Directorate (https://www.mvr.bg/gdpbzn/info-center/%D1%81%D0%BF%D1%80%D0%B0%D0%B2%D0%BE%D1%87%D0%BD%D0%B0%8D0%BD%0-%D0%B8%D0%BD%D1%84%D0%BE%D1%80%D0%BC%D0%B0%D1%86%D0%B8%D1%8F/%D1%81%D1%82%D0%B0%D1%82%D0%B8%D1%81%D1%82%D0%B8%D0%BA%D0%B0). They present about 4-8% of all fires on the territory of the country. The number of reported lightning victims for the same period was 15, while the number of those killed due to floods caused by severe convective storms was 42. Only in 2014, 21 people lost their life in events connected to severe convective storms and torrential precipitation.

One of the primary objectives of the European Severe Storms Laboratory (ESSL) is to collect reports of severe convective events over the whole European continent and to store them in the European Severe Weather Database (ESWD). From our communication with ESSL, we have understood that they would like to carry out and eventually publish a study assessing the impact of severe thunderstorms across Europe. We are looking forward to such kind of study. For now, we’ll refrain from citing any data for losses and injuries caused by severe convection in the current article.
Lines 40-41 – “violent” is not particularly descriptive. Other forms of convection can be violent as well. Multicellular convection can produce features such as bow echoes that can cause substantial damage and would be considered “violent”. We agree with the reviewer and have modified the discussion on L38-L40 correspondingly.

Lines 44-47- Recommend citing Weisman and Klemp 1982 (The Dependence of Numerically Simulated Convective Storms on Vertical Wind Shear and Buoyancy). You cite their 1984 paper, but that deals specifically with directional wind shear. Also recommend switching the last two sentences in this paragraph. Per the reviewer’s suggestion, we have incorporated the suggested changes.

Lines 51-52: “the presence of a strong, single, quasi-steady updraft with a lifetime of several hours and that exhibits rotation in response to the supercell’s mesocyclone”. I don’t quite follow this statement. The mesocyclone is the rotating updraft. This statement makes it sound like you consider them to be two separate things. We agree the reviewer and have deleted the second part of the aforementioned sentence to avoid confusion. The new sentence now reads: “The bounded weak echo region (BWER) signature, for instance, appears at mid-levels above the edge of the low-level reflectivity gradient and indicates the presence of a strong quasi-steady updraft with a lifetime of several hours [13,14]”.

Line 52: You introduce the term “V-notch” but do not describe what it is or how it can be used in a nowcasting scenario. We thank the reviewer for pointing this out. The new version of the manuscript includes a definition of the V-notch signature and makes reference to Dotzek et al. [13] to justify its use in nowcasting applications.

Line 77-78: what is “enough” moisture? We understand the reviewer’s concern and have replaced enough with sufficient.

Line 83: 0-6 km wind shear has been used in severe weather applications for decades. There are many studies that pre-date Brooks et al. (2003) that would be an acceptable citation. The following reference has been added to complement the Brooks et al. (2003)
reference:

Lines 86-88: I don’t see any mention in the Brooks et al. paper about a parameter that is the product of CAPE and shear. The other study you mention uses the product of maximum updraft velocity predicted from parcel theory calculations (which is two times the square root of CAPE) and shear. The text in the updated manuscript has been changed to replace “product” with “combination”. Another sentence (L120-L123) has been also added to clarify the statement.

Line 166: Is this a typo? Why just the square root of CAPE instead of the square root of 2*CAPE, which represents the estimated maximum updraft velocity from parcel theory? In addition, what type of CAPE are you using? Is it surface based or mixed layer? If it is mixed layer, what depth are you using to represent the mixed layer? Does it use virtual temperature of just temperature?

ECMWF implemented a CAPE-shear parameter which is defined according to the formula presented in this paper. It does not include 2 under the square root sign since the relative nature of the EFI metric makes the inclusion of this constant irrelevant. We see the point made by the reviewer; however, omitting or including 2 under the square-root sign would not change the interpretation of our results.

The CAPE-shear parameter uses the CAPE output from ECMWF’s Integrated Forecasting System (IFS). It’s the most-unstable CAPE found from parcels lifted from all the model levels up to 350 hPa. Surface-based parcels are excluded. Instead, mixed layer parameters are used to compute CAPE in the lowest 60 hPa. This CAPE output was implemented in the IFS in 2004. Due to computational constraints at that time, the CAPE calculations were based on the equivalent potential temperature, which is a conserved quantity during a pseudo-adiabatic ascent. Nevertheless, there is a consensus that virtual temperature should be used in the CAPE formulation and a decision whether to implement this new CAPE calculation will be taken after the completion of several evaluation tests. The ultimate goal at ECMWF is to make the CAPE computation in the IFS model consistent with other NWP models and allow a direct intercomparison between them.

Line 204: Your description of updraft helicity is flawed. It is not just a measure of rotation. It represents the covariance of rotation and vertical velocity. A large UH could result from strong updrafts and weak rotation, or medium updrafts and strong rotation.
We thank the reviewer for pointing out the relationship between vertical velocity \( w \) and the vertical component of the three-dimensional vorticity vector \( \zeta \) in the definition of updraft helicity. As pointed out by Kain et al. (2008), updraft helicity is the vertical component of helicity and is an integral measure of the potential for an updraft to rotate. Given that helicity is a dot product between the three-dimensional velocity \( \mathbf{V} \) and vorticity \( \mathbf{\omega} \), helicity obtains its maximum values when (i) the two vectors are aligned and (ii) have large magnitude. The definition of UH has been updated and now reads: “UH measures the potential of storm updrafts to rotate and is especially suitable for convection-allowing models where the vertical velocity \( w \) is explicitly resolved”.

Lines 627-635, lines 597-599: You imply that supercells are rare in Bulgaria, but how rare? Are we talking about a once per year event or once per decade? Unfortunately, the relatively recent operational use of remote-sensing observations in Bulgaria has prevented us from generating official statistics concerning the number of supercells. The need for a comprehensive thunderstorm climatology has been highlighted in the updated version of Section 5. Nevertheless, preliminary analysis indicates that the annual number of supercells ranges from 1 to 5.

**Figure Comments:**

What are the background lines in your maps? Country borders? Other political boundaries? These should be pointed out in the captions.

Thank you for the suggestion. The updated figure captions clarify that the background lines correspond to country borders.

All of your maps need a distance scale or lat/lon values on the axes. Currently, there’s no way to get a sense of scale.

Following the reviewer’s suggestion, we have added parallels and meridians to all WRF figures.

Some of the figures (figure 11, for example) are confusing because in addition to not having any sort of distance scale, some of the panels are showing different geographic extent.

We understand the reviewer’s concern and have made changes to facilitate an easy comparison between model simulations and observations:
• The country border of Bulgaria has been highlighted by a heavy black contour.
• The concentric radar rings have been labelled more clearly to indicate the distance from Bardarski Geran radar and now give a better sense of the scale of the observed supercell.

Figure 16 - The differences in scale found in the various figure panels makes it difficult to compare the radar observations and numerical simulations. The radar images for Figure 16 are cropped to closely match the geographic extent of the d03 WRF domain. Similar to the NEP UH figures, we have also made the country border thicker and hope that the new version of Figure 16 allows for a better comparison between model and observations.
References


