
This work focuses on environmental evaluation of supercell thunderstorms across Spain. Authors combine supercells producing and non-producing hail with proximal ERA5 environments and discuss several convective parameters and accompanying synoptic-scale patterns. Authors divide results into the early and mature stage of the supercell thunderstorm. While this is an interesting work addressing a niche of very much needed studies focusing on supercell thunderstorms across Europe, it requires major revisions to meet certain publishing quality. My most important concern is that authors have a tendency to make statements that are not scientifically relevant while in other instances speculative. My feeling is that authors sometimes 'overdo' interpretation of their results and do not entirely understand the subtle difference between the mesoscale environment derived from a coarse-grid ERA5 and local storm-scale features such as convective updraft that can be only resolved with high-resolution convective-allowing models. A good example is L387 where authors write „Omega vertical velocity reveals that the SP-HAIL’s updraft is higher“ or „sounding composites show large wind values in upper-levels, which may favor wind divergence at the upper troposphere and deep-moist convection„. Manuscript contains several such statements that need to be revised before the article can be accepted. There are also awkward sentence constructions (e.g. „a mechanical trigger to force the mechanism that initiates convection„) that in certain places make it difficult to understand the meaning of the sentence. Thus, further language proof-reading of the manuscript is required.

I also have a feeling that analysis of differences between t0 and tc is a redundant part of this study as it doesn’t introduce important findings. This is especially strange given that authors try to find differences in small details over small distances between t0 and tc, but at the same time they average their profiles to 9x9 grids and do not benefit from 0.25 deg resolution of ERA5. Trying to evaluate subtle differences among closely located t0 and tc for large synoptic-scale features at figures 3, 4 and 5 is even less scientifically relevant. While I like the concept of dividing supercells into hail producing and non-producing events, I am just skeptical whether division into t0 and tc is worth all the attention authors devote in this study. This is not a major issue and I leave the decision regarding
incusion/exclusion of this part to authors. At the end of the day it is their decision what and how they want to present in their work. However, for future studies with this dataset, instead of ERA5 with 9x9 grid averaging, a convective-allowing high-resolution simulation would be likely more appropriate to evaluate different stages of the supercell lifecycle at t0 and tc and investigate the influence of ambient environment and local orographical features.

In this work, I think that division of results into high-CAPE and high-shear events would be probably more interesting and scientifically important in the context of other similar work that has been done for Europe (compared to t0 and tc approach). It is well known that European severe storms are mostly driven by strong kinematics and in lower degree by high instability, which is also a case for supercells.

**Minor comments:**

L15: Suggest changing to „the synoptic configurations and proximity atmospheric profiles related to the supercell events“.

L19-L21: Awkward sentence construction, please rewrite for clarity. Perhaps splitting this sentence into two can help.

L27: Suggest changing „life” to „lifecycle”

L35: I am not entirely sure I can agree with this sentence and the phrase „easily detected”. Supercell detection in Europe is generally not easy if high-quality Doppler radar data is not available (like in the U.S.). I am also not sure how a mesocyclone (which is a core definition of the supercell thunderstorm) can be detected by lightning data. In the majority of instances we can only suspect that supercell thunderstorms developed based on its morphological features, but only a small fraction of these events can be captured by nearby Doppler radar velocity products that provide ultimate confirmation of the mesocyclone. I suggest authors reword and soften this sentence or remove it.

L43: No need to use „observational” ahead of „reports”. Authors may consider using „severe weather reports” instead.

L45: Is there any scientific proof that they are indeed smaller? Authors speculate that it is due to orography and land-sea interactions, but is it really the case? What about big supercells in Nebraska or Southeastern U.S. along the coast of GOM? Is there any scientific proof showing that orography acts to reduce size of the supercells? Perhaps weaker supercells in Spain are rather due to smaller CAPE and WS / less favorable wind
profile compared to their U.S. equivalents. I suggest rewording.

L59: Which „other regions of the world“? Please be more specific.

L95-101: Did authors also use surface data in addition to pressure levels, and eliminate all pressure levels falling below orography for the purposes of parameter calculations? This information should be included in this paragraph. Also, which software was used to calculate convective parameters. SHARPy, MetPy, other, or your own scripts? Did you also consider that some of the proximity profiles may be contaminated by the convection ongoing in ERA5? Did you use convective precipitation threshold equalling 0mm to eliminate such profiles? This might be an approach worth considering in potential future studies to make sure evaluated profiles are pre-convective.

108-111: I am not sure if that was a good idea. In this way authors do not benefit from the superior (compared to other reanalyses) resolution of ERA5. This averaging can have an impact on areas with complex orography and result in the loss of important details. Did authors try to reproduce their results without a 9x9 grid averaging approach? Were these results much different?

L124: What authors mean by „The 2-meter temperature (T2M) and dew-point (DWPT) are computed“. In which aspect T2M and DWPT required computations? To avoid using a word „computed“ authors can reword into something like „We selected the 2-meter temperature (…)“.

L126: What depth was used for calculating mixed-layer?

L149: I am not sure if I understand what authors mean by „This variable is much more interesting than MUCIN, as there are other buoyancy terms which can be evaluated“.

L205: Change „moist“ to „moisture“.

L244: Why do authors think that 90th percentile of MU_CAPE would indicate „largest and severe supercells“? Suggest rewording to „of the supercells developing in highly unstable environments“. Instead of providing mean skew-t profiles divided into t0 and tc it could be potentially interesting to provide also mean skew-t profiles for 90th percentile of WS events as high WS is a major contributor to severe storms in Europe compared to instability that is often limited.
L256: What exactly „helps to organize convection“? Please rewrite for clarity.

L259: „Also, the sounding composites show large wind values in upper-levels (< 400 hPa), which may favor wind divergence at the upper troposphere and deep-moist convection“ – wind values from single profile cannot be used to determine upper-tropospheric divergence and deep-moist convection. It is a spatial pattern of the pressure field that allows to determine divergence and potential areas for the large-scale lift that may trigger deep moist convection. Please rewrite.

L260: „The evolution from t0 to tc depicts a reduction in WS for SP- HAIL, which is mainly denoted in the wind speed and not in the rotation“ – I do not understand what authors mean by „and not in the rotation“. The degree of veering in the vertical wind profile?

L272: „These differences are mainly originated in the low-level wind flows.” - awkward sentence construction, please rewrite for clarity.

L280: CAPE can be a useful predictor but only with the combination of vertical wind shear. Over the tropics there is plenty of CAPE but rarely any supercell or large hail due to weak WS.

L286: How CAPE can be dependent on the orography? Please be more specific. Over northern Great Plains CAPE can reach as high as 9000 J/kg over higher elevation in Nebraska.

L288: Larger compared to what?

L311-L312: I believe this sentence is inaccurate. It is not only an ERA5-related issue but nearly every reanalysis (or NWP dataset) and is related to limited vertical resolution of available levels. applied convective parameterizations and convective contamination. Given that authors used less numerous pressure levels (instead of more frequent sigma levels), CIN values are expected to be less accurate as well. However, as shown in other studies, compared to other reanalyses ERA5 still performs better for CIN (e.g. table 2 in https://doi.org/10.1175/JCLI-D-20-0484.1). I suggest to soften this sentence and reword it to something like: „It is well known that due to limited vertical resolution reanalyses do not represent capping inversions very well“.

L321 Airmass advections from NW Africa and development of elevated mixed-layers can be also another reason for higher CIN across Spain and W part of Mediterranean compared to other parts of Europe.
L325: “a mechanical trigger to force the mechanism that initiates convection” – awkward sentence construction, please rewrite for clarity.

L357: Helicity or rather storm-relative helicity?

L359: Period missing before “Environments”. Also, this sentence has an awkward construction, please rewrite for clarity.

L387: “Omega vertical velocity reveals that the SP-HAIL’s updraft is higher” – ERA5 omega vertical velocity derived from 0.25 deg grid and averaged by authors to 9x9 matrix surely does not tell anything about local storm-scale convective updraft.

Figure 3, 4 and 5: Text that is at the top of each figure and x and y axis is too small and impossible to read.

Caption to figure 6. 90th percentile of what? Please be more specific in the figure caption.