REVIEWER 2

My background is in trajectory calculations in the upper troposphere and lower stratosphere.

Based a method reported by Kalmus et al. (2019), this paper proposes an interesting extension to improve predictability of strong convective events from daily sun-synchronous satellite profiles by coupling with forecasts of adiabatic forward-trajectories. The paper assesses a proxy setup against ERA5 reanalysis data for the year 2019 over a part of the United States. Broader implications of the work are realised in the discussion. The results and discussion are well-written and scientifically reasonable, however I found the introduction somewhat hard to follow and some of the methods unclear. Overall, the paper is suitable for publication with ACP, however there are some specific comments that need addressing, listed below:

Thank you for taking the time to review our paper. The perspective of a trajectory calculations expert was very helpful for us.

In particular, your review encouraged further testing of some comments we made on physical processes behind atmospheric evolution. We found support for some statements but also removed one that we could not convincingly demonstrate. The investigation was an interesting case study so we describe it in a response comment below, but do not include it in the paper since it doesn't affect any primary conclusions.

We also thank you for your very specific statements about which parts of the paper were unclear. In at least one case we had assumed knowledge of products that is not widespread, so the changes made in response to your comments should really help make our work more accessible. Importantly, we now explicitly state details that could be important context for readers to understand potential limitations.

The authors support open data & transparency and uploaded the data to the JPL Open Repository. The repository is new and the data may not be immediately downloadable, but the files will be here: https://dataverse.jpl.nasa.gov/dataset.xhtml?persistentId=doi:10.48577/jpl.EESTWM

Our point-by-point response follows. Your review text is **bold black**, our commentary is magenta (this colour!) and any quoted text that is in the main paper is in quotation marks and is "dark green".

L64-65: Whether ERA5 is a reasonable representation of AIRS should be described. For instance, does ERA5 assimilate AIRS retrievals?

The introduction has been extended in response to other changes, we think a discussion of ERA5's suitability fits best in the Methods Sec. 2.1 so have inserted:

"ERA5 is an ideal data source since it provides the necessary fields with horizontal resolution similar to that of AIRS and vertical resolution similar to that of our output ERA5-FCST. This work aims to evaluate the trajectory-enhancement method for adding time resolution to LEO IR products, so we are not concerned about small differences between AIRS and ERA5."

This paragraph attempts to cover the parts of ERA5 that are relevant to our goals. We will evaluate AIRS-FCST against ERA5, radiosondes and MRMS in future but our requirements here are just (i) a sensible atmospheric state and (ii) information at fine-enough resolution to capture the structures we're looking at. Combined with the new supplementary analysis (see reviewer 1 response or later), we hope the justification is now sufficient.

L187-188: The WRF 27km datset should be introduced appropriately. From looking around, it appears to be a series of daily forecasts and not an analysis dataset (which would prompt other questions), but the text should explain this.

We briefly summarise this important detail, but do so earlier than where you point to in your comment. We think it makes mores sense in Sec. 2.1 where the data are introduced:

"The WRF simulations are forecasts for each UTC-defined day (Ngan and Stein, 2017)."

L197-200: I'm a bit confused about the local enhancement metrics, mainly their names dMU_CAPE and dMU_CIN. How are they calculated for ERA5? You do not calculate the most unstable parcel for ERA5, so is dCAPE a more precise label in that case? Or are you using MU_CAPE from ERA5-FCST for the baseline in all cases? Please explain in the text somewhere, and consider the labelling convention.

We have added discussion text along with Supplementary Figures 1—3. The associated text goes into more detail - ERA5 CAPE is for an MU parcel but uses a slightly different definition.

We changed the main text description to:

"The ERA5 CAPE is converted to an enhanced version using only the same footprints as in ERA5-FCST, and subtracting the ERA5 sample's daily median. For simplicity, we refer to this as ERA5 dMU_CAPE since ERA5 product CAPE is derived from its MU parcel (for details, see discussion associated with Supplementary Figures 1—3 and Supplementary Table 1)."

L391 and Fig 12 caption: What is meant by time-matched? I guess you mean some sort of integration over all timesteps shown in Figs 10 and 11, but it isn't explained in the text. As part of this, I am not quite sure how you show ERA5-overpass in Figs 10 and 11 for each hour, surely you have only one set of ERA5-overpass tp values (are these the same as ERA5-FCST at t0?), I guess it is the histogram bins that are changing with hour. Again, this should be clarified in the text.

The final paragraphs of Sec. 2.3.3 are now greatly expanded:

"Using dMU_CAPE as an example, data from all retrieval days are concatenated, resulting in N values of dMU_CAPE per UTC hour, where the N values have unique combinations of date, latitude and longitude. Thresholds in dMU_CAPE are then calculated from the percentiles of an Nx6 data array, where the x6 refers to each of the forecast UTC hours. Each location in is assigned to a percentile bin, and the associated ERA5 tp mean and frequency with which tp>4 mm hr-1 are calculated within each bin. This calculation is referred to as "matched time" and represents the performance of ERA5-FCST.

For comparison, we also calculate "overpass time" statistics in which the dMU_CAPE values are simply the N ERA5-overpass values of dMU_CAPE repeated for each UTChour. The bin edges for this sample are calculated for the same percentiles, and then the time-varying ERA5 tp statistics are calculated for this new bin assignment. In this calculation, each physical location has a single dMU_CAPE value for all six UTChours, but contributes up to six values of tp to the calculation. This represents the case of using the same nearest neighbour AIRS sounding for every forecast hour.."

Text in the Fig. 12 and 13 captions added:

"See Sec. 2.3.3 for the definition of overpass time versus matched time calculations."

L465-467: Do you have references to support these two sentences?

We have cited the AIRS performance test and validation report (Yue et al., 2020).

Data and code availability: In line with the ACP data policy, the data underlying the results in this paper should be FAIR. Please provide a link to a copy of the ERA5-FCST data you generated, and if possible the analysis scripts too. Either a preliminary link for reviewers to consider, or a link to a FAIR-aligned reliable public data repository.

We have obtained clearance to upload the ERA5-FCST data and the time-matched ERA5 fields to the JPL Open Repository (JOR). The dataset is now referenced in the data availability statement via its link, there is also an assigned DOI.

https://dataverse.jpl.nasa.gov/dataset.xhtml?persistentId=doi:10.48577/jpl.EESTWM

JPL is committed to NASA's Transform to Open Science (TOPS) and the JOR is a new part of this with the rules established only in November 2022. For now we do not believe adding the analysis code would provide enormous added value in exchange for the burden required to obtain clearance, but we fully intend to expand code availability for our work as JPL procedures evolve.

Specific comments that do not need addressing:

The terms pre-convective and pro-convective are similar and risk confusing the reader, I would suggest rephrasing one or the other.

We have replaced "pro-convective" with "conditions favourable for convection" or similar throughout.

L240-243: Could the use of maximum unstable CAPE for ERA5-FCST be causing the higher values relative to ERA5 CAPE? If it was possible to calculate MU-CAPE for ERA5, would that also be higher? Or, how does MML CAPE for ERA5-FCST compare with ERA5 CAPE?

This point also touches on the CAPE calculation issues discussed above, we have added the following to the text, which points to the same supplementary material that discusses CAPE calculations.

"The marginally higher mean CAPE in ERA5-overpass compared with ERA5 in Figure 2(b) may be due to differences in the ERA5 computational approach or in vertical resolution (see text associated with Supplementary Figures 1—3…)."

The supplementary analysis can't distinguish the reason because we do not have the original 137 ERA5 model level output, from which ERA5 product CAPE is calculated. We put plenty of effort into the supplementary analysis and trying to understand this point, and think we've established that our results will not be greatly affected by associated issues.

L283-284: The argument in support of ERA5-FCST would be clearer if you could show the changes to CAPE at the different levels. Does the CAPE gained around ~800hPa outweigh losses to T-q biases near the surface?

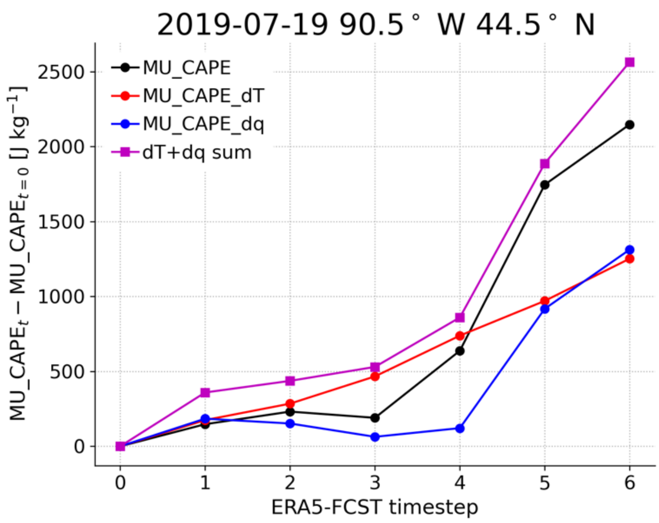
This is a thoughtful comment, we investigated further and removed the claim since it is not crucial to our conclusions and this example is actually rather messy. We include some analysis below that we ultimately decided was too much to put into the main paper since its inclusion would distract without greatly adding to achieving the paper's purposes.

We generally use the MU parcel, so CAPE by *P* layers may not be informative in the Fig. 6 case. In ERA5-FCST the MU parcel moves from ~950 hPa to ~850 hPa, so for the latter period *P*>850 hPa changes are not relevant.

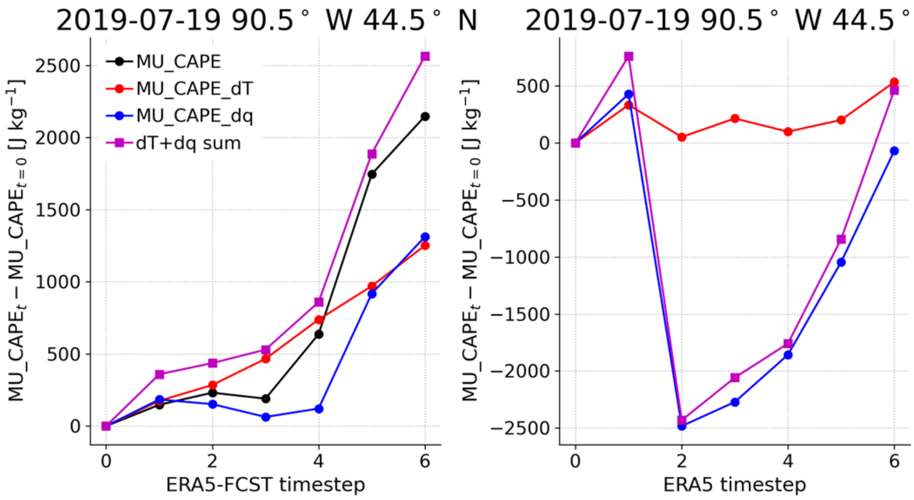
It's also challenging to attribute CAPE differences to any individual level or process because it doesn't decompose linearly, e.g. changes higher up can affect the (Tv-Tv,env) integral *and* the LFC/LNB integration limits.

We recalculated MU_CAPE at the Fig. 6 location with either *T*(*z*) fixed at its t=0 value (MU_CAPE_dq) or *q*(*z*) similarly fixed (MU_CAPE_dT). Changes relative to overpass MU_CAPE are plotted below. The magenta line is the sum of the red and blue lines, while the black line is the calculation from the ERA5-FCST profiles reported in the paper with changes in *both* T and q. Magenta follows black, so the overall time progression seems to be captured, but the decomposition is not exact since the black and magenta lines do not lie on top of each other.

The change from hours 0—6 seems to be nearly 50:50 *T* and *q*, but the strong jump from hours 4—6 is mostly moisture driven. The *q* changes switch the MU parcel from ~950 hPa to ~850 hPa, and the moister parcel saturates lower such that LFC drops ~70 hPa towards the surface, further boosting CAPE.



That story largely backs up the original phrasing. However, the direct ERA5 profiles behave *very* differently, see right panel below:

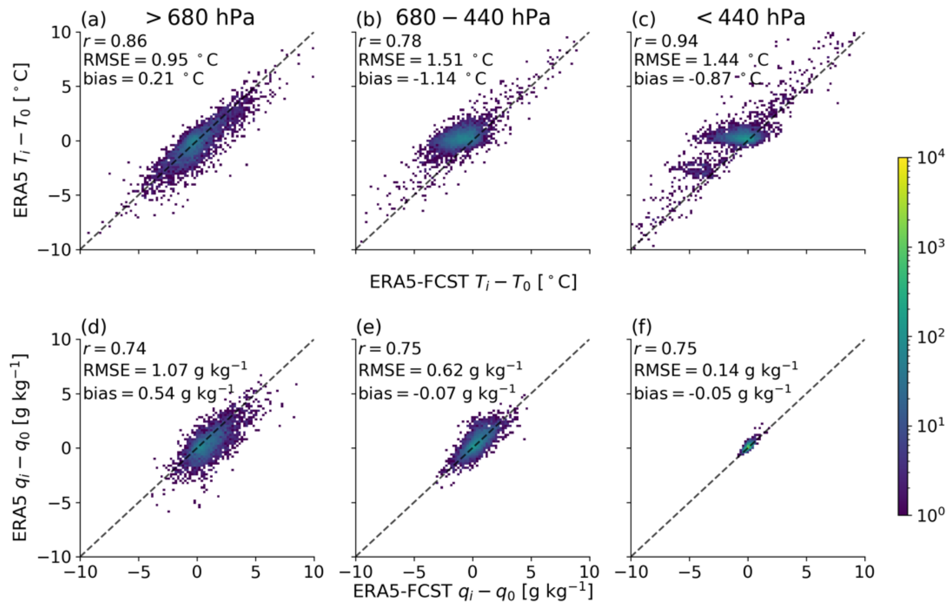


Clearly our argument does not apply to MU_CAPE in ERA5! The MU parcel stays nearer the surface, and therefore dries in ERA5. MU parcel drying combined with warming aloft decreases CAPE, including via raising the LFC by about 100 hPa.

This example is related to the points we touch on later (e.g. subgrid convection warming higher levels and reducing CAPE, albeit potentially non-locally) but we argue that this detailed description is just too much for the current paper. The purpose of Figs. 6 and 7 is to show the reader our process so that they can understand and interpret Fig. 8, which we consider our most-important result alongside Figs 12 & 13.

Fig 8: The points made in the text are strong ones, but there might be a more appropriate figure. If the motivation of the eventual ERA5-FCST product is to predict hazardous convective weather, then Fig 7 is more meaningful than Fig 8, where the biases are more apparent, however it is shows a single event only. Might it be more insightful (as supplementary material) to show a version of Fig 8 restricted to periods of high-CAPE or high precipitation?

Your suggestion re: figure 8 is a neat way to gain insight into part of what's going on, so we produced a supplementary version for grid cells where ERA5 *tp*>1.8 mm hr^{-1} (~99th percentile). The features seem consistent with our statements throughout the paper, e.g. >50 % of variance is still captured. Secondly, we assert that ERA5-FCST gets a mid/upper level cold bias when convection happens, since convection in ERA5 causes relative warming in that product. We would expect the cold bias to get more negative during convection and that happens, the 680—440 hPa layer bias is -0.19 °C in the full sample (Fig. 8) but -1.14 °C in the precipitating columns (below).



We realise that there is an important subtlety in our study that was not adequately explained in the initial draft, since both you & reviewer 1 picked up on it. You both question: why are we looking at all atmospheres and not just convective cases?

Basically, we're less interested in questions "when it's raining heavily, is CAPE higher?" We're more interested in "when CAPE is higher, is it raining heavily?" This is because our product will provide CAPE (and other thermodynamic properties) rather than convective metrics. And for our planned users, we have one of two cases:

Nowcasting, e.g. NUCAPS-FCST – forecasters will have the thermodynamic fields and have the job of inferring things such as convective risk. *They do not know with certainty whether convection will happen*, so results based only on cases where convection happens would be difficult for them to use.

Climate, e.g. AIRS-FCST – for 2002—2013 we don't have a spatially complete, quality controlled & quantitative dataset of convective risk. For 2014—2020ish MRMS will give us that "climate quality" data, with consistent processing, sampling etc. Before MRMS we will therefore *only* have thermodynamics.

A new paragraph in Section 4 discusses our future goals, using some Bayesian terminology to be precise and concise:

"The 2002—recent AIRS-FCST record of thermodynamics will be used with the MRMS surface radar (2014—recent) to relate the derived thermodynamics to convection. In a Bayesian sense, AIRS-FCST will provide P(thermodynamics) and to obtain our target of P(convection) we aim to derive P(convection|thermodynamics) using the combination of AIRS-FCST and MRMS. The proposed analysis is subtly different from previous work such as Kalmus et al. (2019), which studied thermodynamics in convective versus non-convective atmospheres and so reported results relevant to the inverse problem of P(thermodynamics|convection). We also emphasise that while the present study considered CAPE and CIN, this is a proof of concept that only considered a subset of potential thermodynamic properties."

Technical corrections:

L64, L78, L94, L144, L269: Tidy citation parentheses.

Thanks for your careful reading –correction can trigger the reference plugin to force update to all references and add red “track changes” everywhere. If the paper is accepted, we will ensure that parentheses are correctly used for citations during typesetting.

L205: Should there be a word before CAPE? low/high?

We changed the sentence in response to the comment below.

L205-207. This is an important sentence but is quite hard to follow. Consider rephrasing.

We have split this into two sentences and hope it is now clearer:

“We expect that precipitation should be consistently heavier and more frequent in areas of high CAPE and/or low CIN. If high-CAPE and low-CIN conditions in ERA5-FCST are more predictive of precipitation than those conditions in ERA5-overpass, then this is good evidence of the utility of trajectory-enhancement.”

L219: A word seems to be missing from this sentence.

We chose to delete a word instead.

L239: Reference order for Fig 3b and 3c needs correcting.

Done.

Fig 8 caption: Please mention the time period being calculated over.

Done.

Data and code availability: Please include SHARPpy and the WRF27km dataset.

Done.

Supplementary figs: These refer to ERA5-AIRS-FCST, is that ERA5-FCST in the main text? Please check for consistency.

Thanks for your careful attention, this was a good catch – it was the term used in an early draft. The figures and text have been changed to ERA5-FCST.