Comments on “On the relationship between static stability and anvil clouds” by Zhenquan Wang

This paper presents a statistical analysis of the relationship between high cloud fraction and upper-level static stability, as measured by two different metrics. The main conclusion is that the newly proposed stability metric (EAS) has a stronger relationship to ice cloud fraction than a previously used metric (UTS). This result is supported by the analysis, and the author suggests it is useful for studies that use “cloud-controlling factor” analysis. The analysis itself seems sound, although there are some aspects that can be clarified.

I find the motivation of this study and the usefulness of its main result to be unclear. There are also some significant flaws in the presentation of previous work in this area, and as a result the physical interpretations are not always sound. At the end of the day, this paper provides an incremental improvement to how one specific cloud-controlling factor might be calculated.

Paper Organization and Writing

- In my opinion, the use of the term “anvil clouds” is misleading. The analysis includes all high ice clouds. In the tropics, many of these may be convectively generated anvils, but a large portion are thin cirrus clouds formed in situ. In the midlatitudes and subtropics, depending on the time of year, the upper-level ice clouds are dominated by midlatitude weather systems. From what I can tell, there is no attempt in the paper to distinguish anvil clouds from other ice clouds. “High clouds” or “High ice clouds” may be more appropriate.

- I found the Introduction to be a bit off-topic at times. There is ample discussion of low clouds, which are not the subject of this paper, yet little discussion of *why* there is a relationship between stability and high clouds in the first place. In addition, some previous work on anvil clouds is misrepresented (details below). I suggest the author refocus the Introduction on the use of stability as a high-cloud controlling factor. The existing discussion of this topic, and the discussion of LRT, are most interesting and useful.

- There are many spelling, syntax, and vocabulary errors in this paper. I have listed a few of them in the line comments. I recommend that the author pursue professional editing help if such services are available at the author’s institution.

General comments on the study

- **Conflation of various relationships & hypotheses regarding anvil clouds and static stability.** The author touches on a few different ideas involving anvil clouds and stability. As I see it, these can be separated into
    - **(1)** The use of an upper-level stability metric in cloud-controlling factor analysis of high cloud amount, such as in Li et al (2014).

(2) The relationship between stability, clear-sky convergence, and anvil cloud fraction as laid out in Zelinka & Hartmann (2010, doi:10.1029/2010JD013817) and reformulated into the Stability Iris hypothesis by Bony et al (2016).

(3) A process-level relationship between environmental stability and the evolution (i.e., lifetime) of detrained anvil clouds.

Of course, there is overlap between these ideas. But this study seems most relevant to (1), as it presents a new metric of upper-level stability that it claims to be more tightly linked to high cloud area. Ideas (2) and (3) are primarily used as motivation or to briefly speculate some reasons for the study results, but they are not always invoked correctly.

With regard to (2): this relationship between stability, clear-sky convergence (i.e. cloudy-sky divergence), and anvil cloud fraction is an equilibrium argument. The Stability Iris idea is about differences in stability between different climate states, not about differences in stability between cloudy and clear regions. In fact, the hypothesis assumes weak temperature gradients (i.e., minimal spatial variation in stability). Fig 3 is used to argue that low stability -> divergence -> anvil clouds, while high stability -> convergence -> no clouds. This may be true, but it does not seem like a new finding. Deep convection will not penetrate strongly stratified areas of the Tropics, so it is preferentially occurs where there is lower stability…the upper-level divergence at low stability is required by mass continuity…and the clouds will of course be found where convection is. These spatial contrasts between cloudy and clear-sky regions are *not* the stability changes addressed by the Stability Iris ideas, and I don't see a clear connection to climate feedbacks (the connection is not necessary, but the authors use it as a motivation)

With regard to (3), there seems to be a misunderstanding of some previous research about anvil cloud evolution. The author suggests throughout the paper that greater ambient stability is associated with shorter anvil cloud lifetime, but this is not what the cited papers show. The most appropriate citation by the author may be Lilly (1988) (line 60). In Lilly's model, the stratification of the environment acts to flatten and spread the neutrally buoyant anvil—but this is a simple theoretical model that, to my knowledge, has never been tested. The most problematic misunderstanding is on lines 58-61, which cites some previous work about the vertical gradients of diabatic *heating* within anvils. But these hypotheses also assume weak temperature gradients, so the heating is balanced by vertical motion. This may lead to in-cloud mixing but does not actually affect the vertical temperature gradient. Moreover, the vertical gradient in heating has not been shown to be the main cause of anvil spreading and thinning. Wall et al (2020) and Gasparini et al (2022) suggested that diabatic heating and lofting of the entire layer, as opposed to the vertical gradient, is very important to the anvil life cycle. While this question is still unsettled, the in-cloud mixing played a lesser role in these studies.

- **Use of ground-based radar.** The MMCR is not the ideal choice of instrument for detecting cloud top height (CTH). The radar is not very sensitive to small ice crystals, which tend to dominate the upper parts of anvil cirrus. This can make a big difference in retrieved CTH and cloud fraction statistics; see the Key Figures on this page: https://climatedataguide.ucar.edu/climate-data/combined-cloudsat-spaceborne-radar-and-calipso-spaceborne-lidar-cloud-fraction-dataset. As a result, the typical CTHs shown in Fig 2 are 1-2 km lower than previous studies using spaceborne lidar, e.g., Berry & Mace (2014; doi: 10.1002/2014JD021458), Hartmann & Berry (2017; doi:10.1002/2017JD026460), Dessler et al (2006; doi:10.1029/2005JD006705), among many others. The same bias can be seen when compared to the DARDAR cloud fraction in Fig 6, where CTHs look close to 15 km.

In addition, the MMCR can become attenuated by precipitation and optically thick clouds (Hollars et al 2004, doi:10.1016/j.atmosres.2004.03.015), which would also bias CTH low in the case of precipitating convection.

This issue does not necessarily invalidate the rest of the analysis, since this is not a study of CTH itself. But it means that UTS, which measures stability between ~13.5-16.5 km, does coincide rather closely with true CTH. This weakens the author's argument that EAS is favorable over UTS because of its closer proximity to the anvil. Nevertheless, the EAS-HCC relationship still seems to be better than UTS-HCC, so the main conclusion is not affected.

- Please specify how the moist adiabatic $d\theta/dz$ is calculated. Does the author takethe observed pressure and temperature at some vertical level and use it calculate the moist adiabatic lapse rate? Or is it found by launching a moist adiabatic parcel profile from some assumed surface conditions? These two methods would give different results. Also, does the calculation use the saturation vapor pressure over liquid throughout the entire troposphere, or is there a transition to ice saturation at cold temperatures?

Line Comments

- Line 28: I found the wording "Cloud responses to the environmental changes have not been correctly simulated in models" to be a bit odd. We do not know if cloud responses have been correctly simulated, since we do not know the ground truth in future climate scenarios. I'd suggest more precise language for the first sentence of the paper, i.e., something about uncertainty.
- Line 35 "to the mass" -> "on the mass"
- Line 36-37: "are the first-order cloud-controlling factor in all scales"….this strikes me as an enormous claim that I have never heard before. *All* scales, even microphysical? Again, more precise language is needed here.
- Line 44-45 "convection-to-radiation transition" – again, not very precise language. I know what the authors mean here, but many readers may not.
- Line 49: "moist adiabat" -> "moist adiabatic ascent"
- Line 51: "(neutrally buoyant for the air-parcel ascending)" what do the authors mean by this? Is the author trying to say that ascending parcels are neutrally buoyancy? If so, that is not truly the case, just an approximation used in theoretical studies of the lapse rate, e.g. Singh & O'Gorman 2013.

- Line 59: “latent heat release and radiation and interactive with the anvil clouds” -> fix grammar
- Line 104: “with the distance less than 250” – specify that this is the clear-sky distance.
- Line 126: Lidar is sensitive to small ice particles in addition to liquid. As written, it sounds like radar detects only ice and lidar detects only liquid. DARDAR is an ice-only product that still relies heavily on the lidar.
- Line 133: shouldn’t the vertical resolution be in hPa, not meters?
- Line 137: Sentence starting with “The lowest half level…” is unclear. In addition, I think the same symbol is being used to indicate a hyphen as well as a negative sign, which is confusing in this case
- Line 144: “existence” -> “significance”
- Line 146: “The number of independent samples is determined…by the distance between independent samples.” Something is off here—you can’t use the number of independent samples to determine the number of independent samples. And is this distance referring to spatial or temporal distance? If it is temporal distance, shouldn’t the autocorrelation be used to determine the number of independent samples?
- Line 154: caption references a dashed red line but there is no dashed red line
- Line 164: absorption *of* longwave radiation *by* water vapor.
- Line 175: “is close to the anvil top of the maximum convective outflows”…is the author referencing the anvil top height or the height of max outflow? As they noted earlier, these heights are different.
- Fig 3: where are the results of the statistical significance tests for the correlations?
- Fig 3: Do pabels b, c, e, f show results just for the Manus location, or for the whole 60S-60N study region? It is sometimes hard to follow which instruments are being used in each of the figures (e.g., is ice cloud fraction still from MMCR or now from MODIS). It would be helpful to specify in the caption.
- Line 245: “approves” is not the right word here. Perhaps something like “supports the idea that”
- Line 254: what does the author mean by “in which only the correlation at the 95% significance level is counted”? Is the daily mean UTS and HCC being computed for the entire HCC regions, and then the correlations are computed using single values for the entire region? Or, are the daily mean correlations being computed for each grid point, and then the R values averaged across all grid points in the region? If it is the latter, I think all R values should be included in the averaging, not just those that meet the significance test.
- Line 276: could it be that large UTS forces convection to detrain at lower altitudes, producing more clouds in the 11-13 km range?

- Line 280: why would a more unstable environment sustain anvil clouds over time?
- Line 286: is this relationship indeed linear? This is hard to tell from the color scale used in Figs 6-7.
- Line 293: “approved” -> showed
- Line 298: “approves” -> shows
- Line 306: servers -> serves