

Using EUREC4A/ATOMIC field campaign data to improve trade-wind regimes in the Community Atmosphere Model

Graap & Zarzycki (2023)

The study seeks to assess improvements to the prediction of tropical shallow cumulus regimes by modifying CLUBB to allow for counter-gradient momentum fluxes (/prognostic momentum fluxes). Overall, it is important work and I found the paper well written and fairly straightforward to follow. I have a few comments that require revisions before the article should be published, but I consider them minor.

- Lines 30-33: “Changes in low cloud fractions... The Hadley cell” – it’s unclear how the Hadley cell links to the opening/topic sentence of this paragraph. Suggest adding language to more cleanly transition between the two sentences.
- Lines 149-152: Was the CMIP6 version using 58L CAM, with the refined resolution in the BL, and with the SE grid? My impression is that this might be different from the original CESM2 release version.
- Lines 167-168: “C6 and C7 are also tunable constants, although they are left as 4 and 0.5, respectively, for all simulations here.” – are these the default values of C6 and C7 that CLUBB uses out of the box?
- Lines 236-238: “At each of these 10-meter levels, state variable values meant to represent model output are calculated...” – the phrasing here is a bit confusing. The state variables *are* the model output, no? Is the point ultimately that model output is interpolated as a linear vertical distance-weighted average for every 10-meter observation?
- Line 284: “This along with observations in our study being qualitatively similar to the LES-derived profiles in L19...” – Is this implying that the limited observations of $u'w'$ are in line with the LES profiles of L19? I’m confused by the use of “observations” here, which seems contrary to what was stated in the paragraph before.
- Lines 290-291: “In x101, $v'w'$ is also about half as negative at altitudes between 300 m and 2 km.” Would be helpful to note that this refers to Figure 1d, not 1c.
- Lines 302-310: I’m not sure what the discussion of the Ekman spiral in the atmosphere lends to this study in particular. Perhaps draw draw a clearer link or consider removing most of this?
- Figure 2: It seems that panels (a) and (b) are just repetition of Figure 1 (a) and (c); is there a way to combine them then to reduce redundancy? Would also be good to name in the caption which panels refer to which part (i.e., “Vertical profiles of means (a-c),...”) even though the axis labels are fairly clear.
- Lines 314-315: “It can be seen that although x101 has a stronger jet maximum than x001, it has a reduced maximum easterly bias when compared to x001 since its jet placement matches observations better.” – this feels redundant as well, since the smaller bias maximum was noted when discussing Fig 1. Combining Figs 1 & 2 might make this a bit easier to discuss with less repetition.

- Line 317: “near the jet maximum” – is this near the observed jet max, or the model simulated?
- Lines 317-318: “The remainder of the RMSE profiles are quite similar...” – they’re nearly identical for v , but for u it looks like the simulations are fairly different throughout the vertical; maybe a more nuanced statement is warranted?
- Figure 3: “The vertical axis is a rough estimate of the pressure level of the model output” – could you be more specific? Is this the hybrid coordinate pressure?
- Lines 335-336: “Most points with negative $Keff$ in x101 are above this threshold...” – how far above the threshold do these points typically lie? Is there a large spread in the value, and values are often much larger than the threshold, or are values often close to the value (and perhaps thus the findings are sensitive to the choice of cutoff)?
- Lines 347-348: “Confidence is added to this hypothesis by...” – I see the discussion of LES results with different forcing (i.e., Helfer et al.), but does this refer to other studies that use the EUREC4A/ATOMIC forcing? Would be good to discuss/cite those if so.
- Figure 4 and related discussion: Are these differences in theta and Q profiles statistically significant?
- Line 382: A better qualitative match, yes; is this not a better quantitative match as well?
- Lines 382-385: Suggest adding an in-text reference to panels of Fig 6 as they’re discussed. In terms of reducing theta/Q biases in the x200 runs, isn’t this to be expected when any run is tuned to better match the ERA5 results? It seems that potential bias reductions in these runs could be driven more by the tuning than by the formulation of L/prognostic momentum.
- Figure 7: It looks like the experimental L formulation has a substantial impact on u biases in the lowest 1 km. In 001 and 101, negative biases extend to the surface, but those seem to be removed with the L cases. Is there a reason for that? Is the L formulation most sensitive closest to the surface?
- Line 404: “Tropospheric” should probably be “troposphere”.
- Line 410: “between 200 and 2 km” – should read 200 m and 2 km.
- Figure 8: Missing legend
- Lines 424-425: “do demonstrate a likely connection between the prediction of upgradient fluxes and modifications to various terms in the vertical momentum flux budget” – Could you elaborate on this a bit more for clarity? How does this tie into the vertical momentum flux budget terms? It seems that this is just the prediction of upgradient occurrence in the figure.
- Fig 10: Would be helpful to have additional percentages labeled, not just 100% and “same” (and perhaps same should be written as 0%?). Overall the colorbar combined with the actual bias values in the boxes is a little confusing. It would seem for example that the bias in x101 for Mixing Ratio should be not quite the darkest red (it’s not a doubling of the bias), but It’s the same color as x201, which is more than a doubling of the bias...
- Lines 445-446: It’s worth noting that although “the greatest improvements are seen in u and U_h , there’s a stronger degradation in Q when you add in the experimental L

calculation. Would be a more balanced description of the results, at least; elaboration would be great.

- Fig 11: Please add additional colorbar markers, as for Fig 10.
- Lines 489-490: “One of the most notable...” – suggest adding a parenthetical reference to guide the reader exactly where to see this. So maybe at the end, add “(solid brown line in Fig. 12)”? Would help in additional sentences of this paragraph as well.
- Lines 495-496: “...could be leading to changes in atmospheric stability...” – any evidence that could be added to support this?
- Lines 528-536: “This study is a targeted regional investigation and as such, the improvements seen here cannot necessarily be generalized to the global climate system without further exploration...” – This is a really important caveat, and I appreciate the discussion surrounding it. The question arises then – why not use these simulations to evaluate global performance? You have the full global output, so could this dataset be a tool for exploring additional regions/field campaigns, and more generally for looking at global biases? It may be beyond the scope of this particular study, but is it something that’s targeted for future work or are the runs not suitable for that analysis?