Reviewer #2

By conducting a series of CAM hindcast simulations during the EUREC4A/ATOMIC field campaign in the Tropical Atlantic in early 2020, this study points out that the prognostic treatment of momentum flux improves the simulation of tropical trade-wind to a greater extent, compared with the default CAM-CLUBB. This study improves the CLUBB scheme by further implementing a generalized calculation of the turbulent length scale, which reduces model bias and RMSE relative to observation. Recently, more and more studies have noticed that momentum fluxes can be upgradient in shallow convection regimes, and it is challenging to parameterize them. In my opinion, this study is interesting. This paper also gives us some insights on how to further improve the prognostic treatment of momentum flux and simulation of trade-wind regimes in global models.

In addition, this paper is well-written. I therefore recommend the publication of this paper in GMD after some revisions.

Thank you for your review and thoughtful comments.

I noticed that Guo et al. (2021) applied additional damping of non-cloudy layers stable stratification to scaler flux and also w’3 in the stable layers, because of gravity-wave dispersion under cloudy conditions (please refer to eq23 and 24 in Guo et al. 2021). This might be of some help in the simulation of shallow convections and stratocumulus. The case of momentum flux may be similar to the scaler flux, should an additional clear sky damping be considered in Eq 6 as well? And, will it further improve the results?

This is a good thought and certainly something we’d like to explore in the future. For this exploration, we aimed to compare the potential changes in CLUBB with respect to the EUREC4A/ATOMIC field campaign observations, and therefore the exploration only included a potential subset of modifications to CLUBB (including the simple “taus” formulation in Guo et al., 2021). We are working with Vince Larson and other researchers at the University of Wisconsin-Milwaukee and the National Center for Atmospheric Research to explore the utility of additional damping mechanisms.

We have added a sentence to the discussion: “*While we only evaluated a subset of targeted dissipation permitted by this experimental length scale treatment, other possibilities, such as the additional damping of the third-order moment of vertical velocity in stable layers described in Guo et al. (2021), merit further study.*”

C6 is used in scale flux and momentum flux, which are very important for the return-to-isotropy terms. But, I also noted that C6=C6b= 1 in Guo et al. (2021). So, the Taus could have full control over the return-to-isotropy in the momentum flux equation. This study uses the default CAM setting of C6=4 (and I assume C6b=6), which would involve the skewness function in the calculation of the pressure term (Larson 2022), which overlaps with the role of the Tau scheme. It creates some difficulties in understanding the performance and improvement of parameterization. I suggest that the authors set C6=C6b=1 for the experiments to simplify the problem.

Thanks for raising this point. We actually set C6 = 2 as constant (i.e., no skewness function) in the experimental length scale runs. This is based on personal communication with CLUBB's developer Vince Larson. This has been noted in the text and the correspondence from Vince Larson is reproduced below for posterity. We have also updated the text to properly define C6 for the "O" (original length scale) runs (i.e., that we retain the CAM6 default settings).

"The value of 2 is a historical artifact, and it is preserved in our code so that we can recover [the Guo et al., 2021] tuning.

Someday we'll get rid of it, but the key point is not the particular value of C6, but that if you're tuning, you should tune the C_invrs_... parameters and not C6, C1, etc., because the latter parameters are redundant, and we don't want the space of tuning parameters to be too large."

The difference between the X101 and X204 horizontal momentum fluxes budgets is remarkable. The budgets of horizontal momentum fluxes in X204 qualitatively resemble those of BOMEX and RICO in LES (e.g. Figure 7 and 9 in Larson et al., 2019), with the buoyant term predominantly balancing the turbulence production term. I'm curious what causes this big change, is it due to the tuning listed in Table 1 or the model structure changes? This study does give us some explanations, but it would be better to add some more discussions and show more turbulent profiles that help the reader to better understand the improvement, such as w^2 , u^2 , v^2 and scaler fluxes. u^2 , v^2 would also benefit directly from the prognostic treatment of momentum flux.

We agree! Unfortunately, performing a much deeper dive than what we have shown here would be difficult without re-running many of the simulations with additional diagnostics. However, we do more closely link the results of Larson et al., 2019 and do also highlight that we should more closely look at these budget terms in future work to better understand why the set of tunings in PM-X appears to better produce LES profiles for test cases with similar dynamical setups to the EUREC4A/ATOMIC atmospheric environment.

The text in this section has been modified to read: *"Another notable difference is the changing of the sign of the buoyancy production term (solid blue lines) from weakly negative in PM-O to notable positive below 700 m, and negative above in PM-X, particularly in the zonal momentum budget. This is also qualitatively consistent with the BOMEX LES budgets in L19 (their Fig. 7) which lends credence to process-level improvement in the PM-X runs. We hypothesize that this may be related to increased stratification in the θ profile in PM-X making vertical transport or air parcels due to buoyancy more difficult in the lowest 700 m. In that case, improvement in the thermodynamic profile in PM-X could be leading to changes in atmospheric stability (e.g., note the differences in the change in θ with height in Fig. 6a), which in turn lead to changes in buoyant production of uh'w' which then feeds back to changes in the dynamic profiles. Since downgradient diffusion corresponds to a simple balance between turbulent production and return-to-isotropy, the fact that the buoyancy term is so large in PM-X could explain the enhanced upgradient fluxes in Fig. 9. We admit this is speculative, however, and experiments with more constrained model configurations (e.g., single column, nudged runs) and voluminous diagnostics would be helpful in providing deeper insight, including more detailed consideration* *of other turbulent quantities such as the vertical fluxes of temperature and moisture as well as variances (e.g., u'2 and v'2 would be directly affected by the additional of prognostic momentum to CLUBB)."*

Also, how were the parameters in x201-x204 determined (Table 1)?

These were defined using a Nelder-Mead optimization strategy described below. As suggested by Reviewer #1 we now only include one set of tuning parameters to simplify the paper and allow for a cleaner evaluation of the prognostic momentum and experimental length scale formulations. The original Table 1 has been removed and replaced with in-text notation of the tuning coefficients. The text now reads:

"We determine tuning coefficients for this configuration using a Nelder-Mead optimization (Nelder and Mead, 1965). Specifically, a set of very short (48-hour) hindcasts initialized on January 1st, 2012 is run, optimizing various tunable parameters in CLUBB to minimize the difference in the predicted wind field after 2 days when compared against ERA5 reanalysis at the same time. Optimization is completed relative to global ERA5 reanalysis data rather than the local EUREC4A/ATOMIC data to ensure a reasonable global simulation."

The names of the experiments in the main text make it a bit difficult for me to read, could you please give them more visual names?

The simulations have been renamed to include the prefix "ED" if they diagnose momentum fluxes via eddy diffusivity or "PM" if they apply the prognostic momentum formulation described in Section 2. We also include a suffix of either "O" if the model uses the original length scale formulation published by Golaz et al., 2002, or "X" if using the experimental "taus" formulation from Guo et al., 2021. See our new Table 1 and our response to Reviewer #1 regarding the naming convention, as well.

Figure 8, legend explanation?

We have added a legend to the center panel of Fig. 8 that defines the specific lines in these plots.

If possible, it would be better to add some LES results to the plots, like in Figures 1, 5, and 12?

We have set up a configuration of Cloud Model 1 (CM1) developed by George Bryan (NCAR) to run a "EUREC4A/ATOMIC" LES case. This LES case is built off the traditional BOMEX test case published in Siebesma et al., (2003, J. Atm. Sci.) but with updates to more accurately handle the large-scale environment observed during the EUREC4A/ATOMIC field campaign. Qualitatively these results agree with LES simulations published using the BOMEX test case, but the state field profiles we simulate here are better matched to the actual radiosonde observations. We describe the updates in the manuscript so that other researchers can reproduce these results and the domain-averaged model data (along with CM1 namelists and configuration files) are included in the project's GitHub repository archived at Zenodo. Like the CAM simulations, the LES simulation also shows a region of countergradient fluxes, as has previously been demonstrated in recent BOMEX LES simulations (e.g., Larson et al., 2019, Dixit et al., 2020).

Line 410, page 18, “between 200 and 2 km”, do you mean 200 meters?

Yes, you are correct -- we accidentally omitted the "m" label here. This is fixed.