

Response to comments for “Quantifying the Oscillatory Evolution of Simulated Boundary-Layer Cloud Fields Using Gaussian Process Regression”

Loren Oh*

February 12, 2025

We sincerely appreciate the constructive feedback for the submitted manuscript. We have made extensive changes to the manuscript based on the reviews, and performed additional model runs to reflect the points raised by the reviewers. The following document summarizes those points, and the changes we’ve made.

The line numbers in the responses refer to those in the revised manuscript.

Responses to Comments

Major Points

1a. I see spectral analysis as the gold standard for identifying periodic signals in a dataset - if the dataset has a signal at a certain wavelength, it **must** show up in a spectral analysis. I also disagree that spectral analysis is particularly sensitive to noise - noise usually shows up in a power spectrum as a lower limit to spectral power across a range of frequencies. It is true that this lower limit can be so large that it swallows your signal, but there are 2 peaks in Fig. 5 that rise above this floor. In this context it is hard to square the results of your GPM against what I see in your power spectrum or even via visual inspection of your time series.

Thanks for pointing this out. We realized that some of the steps we have taken were left unjustified in the original draft. One of the things we observed from the time-series is that it is very noisy and non-stationary, which makes DFT less informative. The relevant observation

*Corresponding author: loh@eoas.ubc.ca

at line 208 now reads: “We are interested in determining whether the fluctuations in the time-series of the cloud size distribution in Figure 4 are consistent with a periodic behaviour. The oscillatory evolution in b is not immediately obvious in Figure 4, and performing an Augmented Dickey-Fuller (ADF; Dickey and Fuller, 1979; Hamilton, 2020) test reveals that the time-series is non-stationary. That is, the observed time-series $b(t)$ cannot be assumed to be stationary, which would have a unit root. We would like to quantify the extent to which the time-series is consistent with earlier studies regarding oscillations in the cloud size distribution. In the following section, we follow Feingold et al. (2017) and use Fourier spectral analysis to identify the underlying periodic behaviour in the observed time-series”.

In the presence of noise and non-stationarity, it comes down to the problem of interpretation. On this front, we encountered three issues with both DFT and ACF, which is that the observed data is quite noisy (which may or may not behave as white noise), that the time-series is non-stationary and that the signals are found in the low-frequency regime.

To show that these factors make interpretation of the spectral analysis more difficult, we defined the *noise floor* to be the magnitude of the spectral power for Gaussian noise using a chi-squared distribution. We have revised the manuscript to include more details about this particular choice, and the relevant paragraph at line 226 now reads: “Figure 5 shows the estimate of the power spectral density using the periodogram (blue) and the 95% confidence interval (red) obtained from the noisy time-series $b(t)$, plotted as a function of period $T(k) = N/k$. The 95% confidence interval defines the threshold that separates oscillatory signals from noise, against the null hypothesis that all signals in the periodogram are Gaussian noise, and is based on a chi-squared distribution with 2 degrees of freedom (Panofsky and Brier, 1958)”.

The goal here was to see if the traditional spectral analysis could identify the 90-minute period with enough confidence (i.e. 95% confidence that the signal is not white noise). From what I can tell from the spectral analysis, the time-series indeed contains large variability for the spectral analysis to be effective.

1b. I’m pretty sure all you are doing is effectively band-pass filtering the data by looking only for periods between 5 and 150 minutes which throws out the large signal at 230 minutes and causes the secondary peak at 100 minutes in Fig 5 to rise to the level of statistical significance because so much of the rest of the power spectrum has been thrown away. I could be wrong about this, but I would need to see my hypothesis proven wrong before I’d be comfortable accepting this paper.

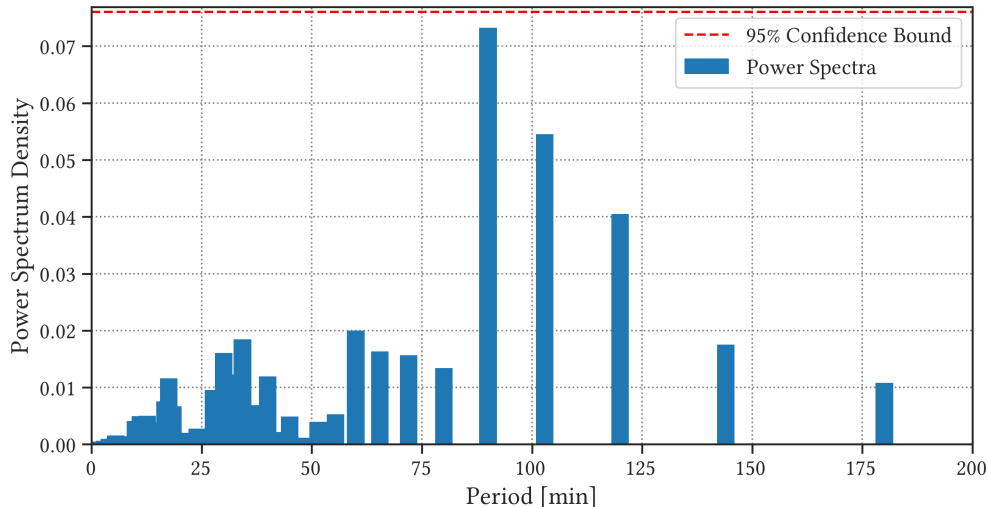


Figure 1: Power spectral density from FFT, with periods larger than 200 minutes filtered out.

We understand the concern, but the choice of the prior does not exclude the signal at longer periods, unlike the DFT. First of all, band-pass filtering the Fourier series still does not yield a strong enough signal given the noise and/or the variability within the periodicity (see Figure 1).

The choice of the prior (periodicity) in the GP regression method, on the other hand, is not meant to be (and is not) deterministic. In most cases, the GP model does not get stuck in local minima. We still tested a range of periods simply to demonstrate that our model does not get stuck in local minima; the GP model converges to the 90-minute periodicity given enough time.

We should definitely have explained why we ignored the 230-minute period. The relevant discussion at line 233 now reads: “There are two prominent periods on the periodogram, one at $T = 100$ minutes and the other at $T = 233$ minutes, but only the latter signal is above the 95% confidence interval. Both signals have periods longer than the 80-minute period observed by Feingold et al. (2017) and no significant signals can be found at shorter periods. The signals are found in the low-frequency regime, and because of that, the frequency bin width is too large to pinpoint the exact period from the periodogram, especially for $T = 233$ minutes. This period is also much longer than the expected oscillatory behaviour (Dagan et al., 2017; Yamaguchi et al., 2019), and is possible that it is a harmonic of a fundamental frequency hidden by the noise and non-stationarity.

We have also tested other methods to estimate power spectral density, such as the circular autocorrelation function (ACF; Parzen, 1962b), but (partial) autocorrelation function of the

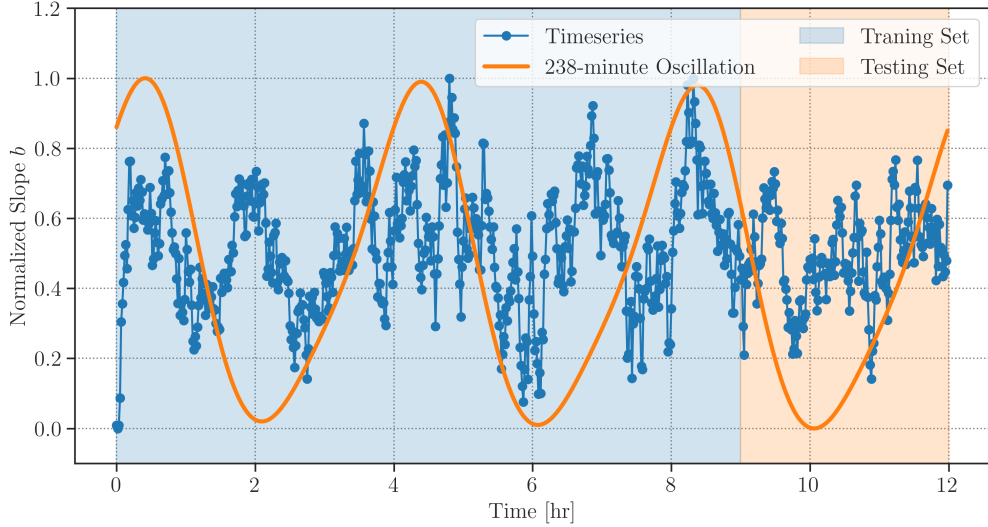


Figure 2: Normalized time-series of slope b of the cloud size distribution $C(a)$ (blue line), compared to the the mean posterior from the periodic GP model (orange line), representing the 238-minute oscillation.

observed time-series $b(t)$ decreases slowly over time, and no significant lag can be found. The presence of noise as well as the non-stationary nature of the time-series makes it difficult to examine the behaviour of the time-series”.

Still, we can definitely use our GP model to see what the low-frequency oscillation looks like. We used a 230-minute period as the prior, and let the GP model converge slightly (by stopping the gradient descent prematurely) to an oscillation with a period of 238 minutes. See Figure 2; it is clear that the longer period ignores much of the variability in the measured time-series. We have performed a number of trials with longer periods, but found no improvements. The manuscript now mentions this at line

1c. I think my simple story in 3b is getting lost in a maelstrom of math. Einstein once said that a model should be as simple as possible, but no simpler. In that context, why do you perform your analysis on the derivative of the signal rather than the signal itself (when there is no trend in the data to get rid of and subtracting the linear regression would be simpler if there were)? Why do you need GPM instead of applying a digital filter? I’d like to see clear justification for each step towards complexity you take

We believe that the reviewer meant to say point 1b.

This must be the third thing that had to be justified, but was missing from the original draft. Thank you for pointing this out again. As we have noted in 1a, the observed time-series is

both noisy and non-stationary. And we have found that taking the derivative of b normalizes and isolates the periodic component, which is indicative of oscillatory component. The relevant discussion starts at line 340, which reads: “The smoothly modelled distribution $\tilde{b}(t)$ from the mean GP posterior distribution $b(t)$ in Figure 6 corresponds well to the observed time-series, showing the oscillatory behaviour within the noisy observation with a period $T = 95$ minutes. In this particular case, the initial regression attempt yields a good estimate of the hyper-parameters for the observed time-series. However, in situations where a general, long-term trend breaks the quasi-stability assumption, additional steps can be taken in order to better isolate the oscillatory behaviour of the cloud field, which still remains noisy and non-stationary. The standard practice to account for the non-stationarity is to take the derivative of the time-series $\partial_t \tilde{b}(t) = \partial \tilde{b}(t) / \partial t$. If the oscillation is dominated by a single frequency, the frequency should also characterize the derivative of the oscillation. Given this, we build a GP regression model to estimate the period of the oscillation in $\partial_t \tilde{b}(t)$, which can be seen in Figure 7.

We have found that taking the derivative of $b(t)$ also successfully normalizes the observed time-series, which is useful for statistical analysis. Applying the ADF test (Dickey and Fuller, 1979; Hamilton, 2020) to $\partial_t \tilde{b}(t)$ also confirms that the resulting time-series is now stationary. As shown in Figure 7, the values of $\partial_t \tilde{b}(t)$ varies with zero mean with no obvious trend over time. There are small variations in the amplitude, but we can now drop the SE kernel to account for the variability in the y-axis, and only use the periodic kernel (i.e. $\hat{k} = k_{\text{per}}$) to estimate of the periodicity T . This reduction in the number of hyper-parameters is also necessary to perform Bayesian inference, which will be described in more detail in Section 2.7”.

2. I’m concerned that oscillations in cloud populations might be fairly local in nature such that averaging over a 45 km fetch averages out most of the variability you seek. Thankfully, you can easily address this question (which would be a nice addition to your paper)! In particular, can you repeat your analysis by only analyzing output over subsets of your domain? If so, does the peak frequency or spectral power change as a function of region size? Also, discussion of averaging over a large region occurs around line 330, but I’d like to see more discussion and I’d like to see it show up in the experimental design section.

We performed the analysis again over a (randomly chosen) quarter of the domain. Figure 3 shows the resulting power spectral density. As the reviewer suspected, the signals appear stronger on a smaller domain, which is expected as expanding the size of the domain would introduce more variability in our time-series. Still, the spectral analysis returned multiple

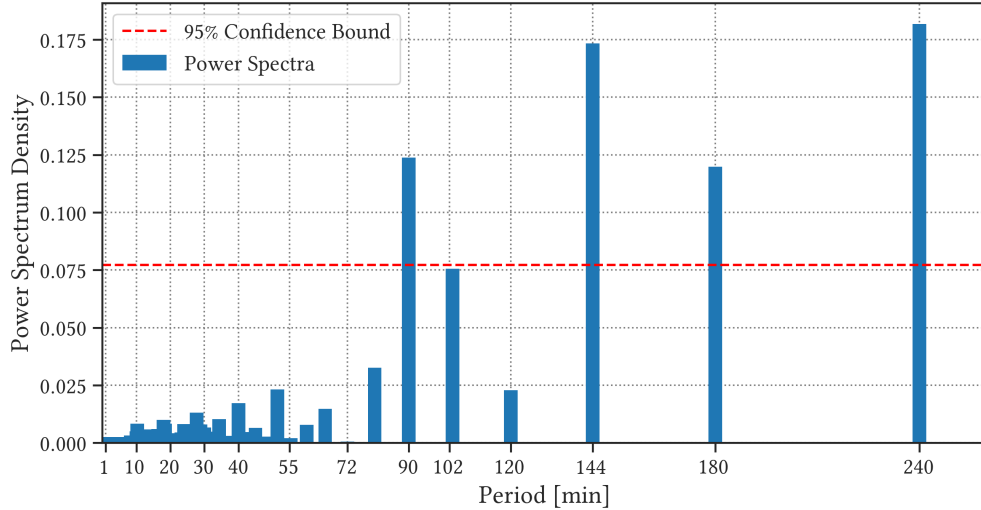


Figure 3: Power spectral density from FFT for a quarter of the domain.

signals. We can see additional signals at 144 and 180 minutes above the 95% confidence bound, which is interesting.

Still, the signals at 180 and 240 minutes are likely harmonics of higher-frequency oscillations, and the bin size for the DFT is too large to determine what the actual period should be. Given that, we used 90 and 144 minutes as priors for our GP model, and tested how accurate they were (see Figures 4 and 5). The former converged to a 95-minute period, which is expected, and the latter converged to a 144-minute period, which is less accurate. So it does not become significantly easier to identify the oscillatory evolution of the cloud field even on a smaller domain, as the variability is still too large for traditional methods.

3. When I look at Fig 9, I see 4 very clear waves with 95 min period while the rest of the timeseries doesn't fit this frequency very well at all. This leads me to suggest that your active/quiet hypothesis works well for some chunks of your simulation but something is interrupting the signal for other chunks. Can you watch a movie of your simulation or perform other analysis to figure out what is interrupting your behavior of interest? Finding conditions which disrupt these oscillations would be a very noteworthy contribution!

The revised manuscript introduces quite a bit of discussions on possible causes of those deviations. For example, line 419 reads: “Based on Figure 9 at around $t = 4$ hours, the cloud field is expected to go towards a phase of relatively weak convection where there is a relative abundance of smaller clouds. Small normalized values correspond to more negative slopes of the cloud size distribution. However, upon inspecting the evolution of the cloud

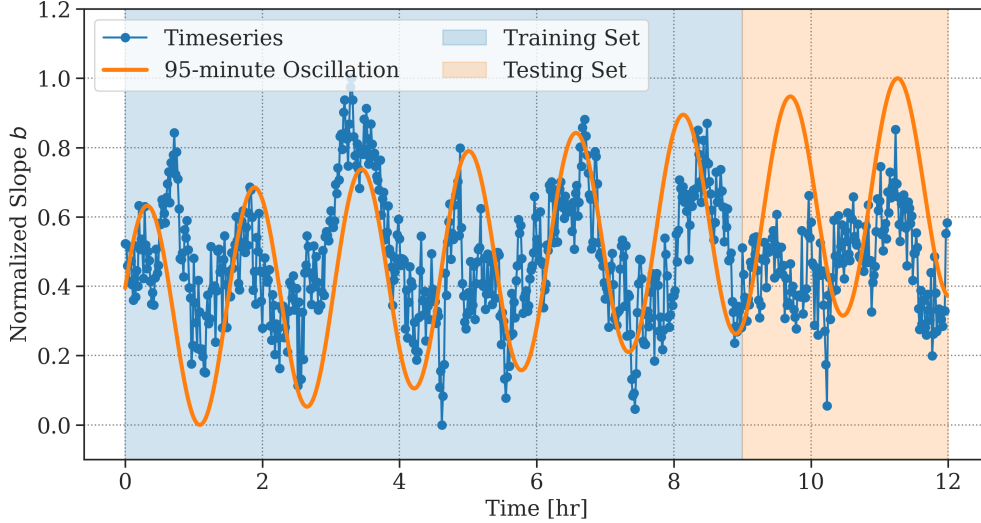


Figure 4: Normalized time-series of slope b of the cloud size distribution $C(a)$ (blue line), compared to the the mean posterior from the periodic GP model (orange line), representing the 95-minute oscillation.

field during this time, large structures form at the top of the cloud layer (light grey regions in Figure 11) as a result of strong convective activity, which persist until $t = 5$ hours into the simulation. Once the large, thin layer of clouds at the top dissipates, the deviation the observed time-series of b becomes much smaller. During this time, as shown in Figure 11, the cloud field is dominated by the growth of small clouds. However, a thin layer of clouds persist at the top of the cloud layer, which skews the slope b of the cloud size distribution.

On the other hand, around 9 hours into the simulation, the cloud field is expected to go towards a phase of relatively strong convection where there is a relative abundance of larger clouds. Visually inspecting the development of the cloud field (Figure 12) suggests that strong convective activities occur in groups, and clouds tend to merge and become much larger, which reduces the relative number of large clouds compared to small ones that are less likely to merge with other clouds or reach the cloud top layer. Because of this, the normalized value of b of the cloud size distribution becomes smaller, indicating a much steeper slope despite the strong convective activity”.

We have also added Figures 11 and 12, which are snapshots of the model run to give more insight into the state of the cloud field.

4. I'm also worried that the narrow dimension of your bowling-alley domain might be corrupting your results. It seems to me that the narrow horizontal direction would constrain the size of cloud radii. And since your whole paper

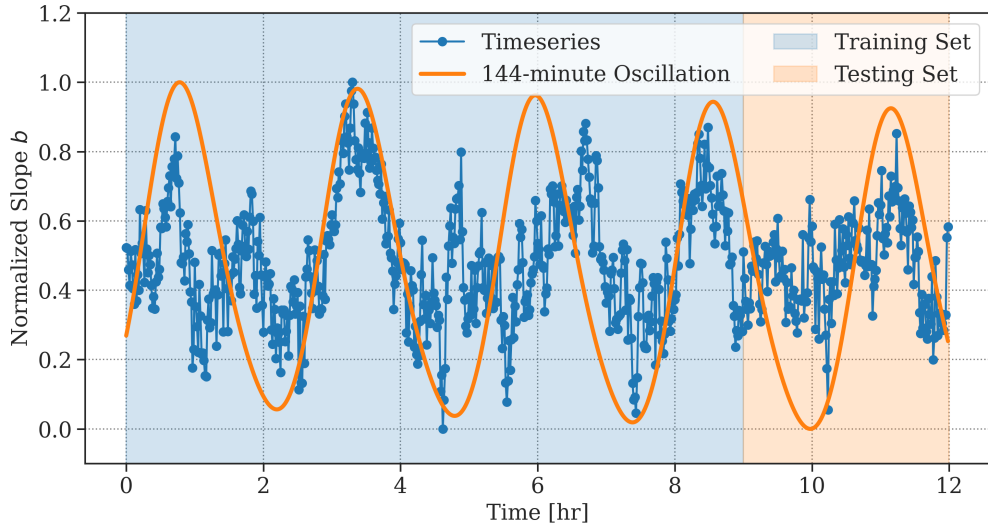


Figure 5: Normalized time-series of slope b of the cloud size distribution $C(a)$ (blue line), compared to the the mean posterior from the periodic GP model (orange line), representing the 144-minute oscillation.

centers around the cloud size distribution, locking the biggest clouds to the domain size seems like it could affect your conclusions. Can you reassure me that this isn't happening? Just providing citations of previous work showing this isn't a concern would be fine.

The model domain is definitely large enough for cloud growth, as if it were not the case, the clouds will definitely grow towards the size of the domain and it becomes obvious that the domain is too small. We have added snapshots from the model run, and while there are no previous studies specifically for the CGILS S6 case, doi.org/10.5194/gmd-6-1261-2013 could be a reference for this.

However, it should be mentioned that we have observed what appears to be the effect of spatial organization (due to the formation of cold pools). If we were to study how cold pools affect this oscillatory behaviour, the domain size might not be large enough. This response relates to comment 6, so we will mention the relevant corrections there.

5. around line 105: I don't understand whether you're just taking cloud cross sections at a single level (e.g. 1 km) or whether you consider each vertical level to be an independent cloud. The latter approach could be problematic because clouds at different levels wouldn't be independent samples which would mess up statistical analysis tests. This would also distort interpretation because what we think of as a distribution of cloud sizes could in fact just represent

the vertical structure of a single cloud. Can you clarify what you're doing and address potential concerns if you are using all levels?

We realized that there were some unnecessary discussions on cloud sampling, so we tried to make it clear. Line 128 now reads: "In order to obtain the cloud size distribution, individual clouds need to be sampled conditionally. In this study, horizontally contiguous regions (grid cells) containing condensed liquid water ($q_l > 0$) are considered to be the cloud region. The *size* of a cloud is then defined as the area of the horizontal cross-section, which is the number of grid cells containing condensed liquid water multiplied by the horizontal grid size".

Connected convective columns being statistically correlated is definitely a possible issue, which we were aware of. We initially used 60% of the random samples, but found that the statistical distributions appear to be the same regardless. We address this issue at line 130, which reads: "The size distribution contains multiple horizontal cross-sections of the same cloud. However, randomly sampling 60% of the cloud samples had no impact on the analysis, likely thanks to the sheer number of samples and the shallow nature of the modelled clouds in comparison to the size of the domain".

6. There's a lot of focus on matching the periodicity found in previous studies, but do we know that such periodicity should be spatially and temporally invariant? Recharge/discharge seems like it would be proportional to boundary layer depth and the vigor of turbulent/convective mixing. Timescales might also be different in the different sorts of cloud analyzed by these previous papers.

Thanks for the suggestion. This was brought up a couple times in the review, and we have added a bit of discussion on factors that can influence this oscillation, namely aerosol concentration and spatial organization as well as domain size. For example, line 592 now reads "We have also observed the spatial organization of shallow convective clouds that disrupts the oscillatory evolution (Figure 12), which can also be seen in studies modelling the boundary-layer cloud field over a smaller domain (Wang and Feingold, 2009). Reducing the size of the model domain may reduce the effect of spatial organization and make it easier to estimate the periodic behaviour of the cloud field using traditional methods. However, given that these factors can manifest even on a smaller domain, a robust method to estimate the periodicity of a noisy, non-stationary time-series is still useful, especially if no smallest, *optimal* domain size exists where the recharge-discharge cycle can be isolated".

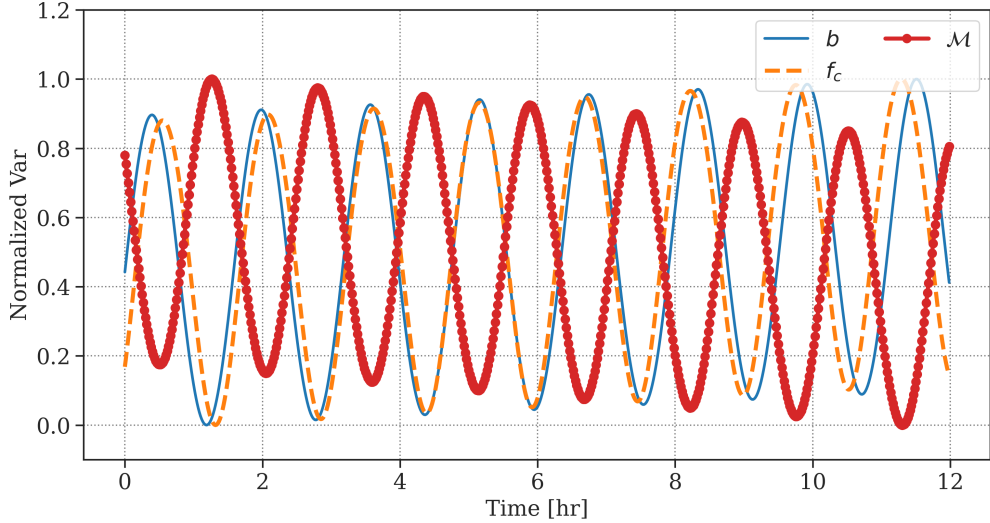


Figure 6: Mean posterior distributions from the GP regression for the slope b of the cloud size distribution (blue), cloud fraction f_c (orange) and average vertical mass flux \mathcal{M} .

Minor Points

1. It would be nice to see the timeseries of cloud size distribution, mass flux, and cloud fraction all on a single graphic. The degree of correlation between variables is difficult to see by flipping between graphics. A lead/lag correlation analysis would be a nice way to show the relationships between these fields.

Multiple people pointed this out, so to address this point we have come up with an additional figure (see Figure 6 here) showing the oscillations in the three variables.

This has been added as Figure 15 with relevant discussions.

2. A cartoon explaining the relationship between cloud size, mass flux, cloud fraction, and precipitation in terms of hypothesized regional life cycle might be useful for driving home what you're seeing.

This has more to do with the dynamics between mass flux and precipitation, and we thought it would be made redundant as it has been shown by Dagan et al. (2018). We would like to perform a follow-up study that looks into the changes in vertical mass flux distributions, and we will consider making a figure on the relationship once we know more about the mechanism.

3. I really like Figure 1! It does a nice job of grounding the paper in reality. It might be even more powerful if you showed snapshots like this for both a very active and very quiescent period.

Thank you. We added more discussions on factors that disrupt the oscillatory behaviour (based on Figure 9), and added two more snapshots (Figure 11 and 12) to visualize the state of the cloud field and gain more insight into what is causing the deviations.

4. The KDE gives you a non-parametric PDF which you can use to get the slope you need...so why do you introduce a power-law distribution? I suspect all you want is to compute the linear regression of the nonparametric KDE PDF in log-log space (after chopping off the nonlinear bit on the left-hand side of the curve in e.g. Fig 3). I suspect I'm complaining about your wording rather than what you've actually done.

There were requests of comparing our result with previous studies (specifically satellite observations). Line 197 reads: "Here, the slope is measured to be roughly $b \approx -0.65$. We have calculated b for non-normalized values of $C(a)$, and the time-series is found to vary roughly between -1.4 and -1.7 , which corresponds to a range of -0.7 to -0.85 based on the methods by Neggers et al. (2003), and -1.7 to -1.85 by Benner and Curry (1998). The measured slopes are slightly smaller in magnitude but comparable to the slope of $b = -1.7$ found in a large-eddy simulation (Neggers et al., 2003) and the slope of $b = -1.98$ from remote sensing observations (Cahalan and Joseph, 1989; Benner and Curry, 1998)".

5. I found Fig 3 to be a bit confusing. I would expect the derivative of the original plot to come below/after the plot it is the derivative of. It also took me a while to see the delta in front of the ylabel in panel a, which added to my confusion. Further, it would be reassuring to see several snapshots in time to make sure your log-log linearity assumption is justified in general instead of just in this test case. Alternatively, you could plot RMSE of fit (or something like that) for your linear regression as a function of time.

We found that the panel a in Figure 3 is actually never mentioned in writing. We apologize for the confusion. Line 184 now reads: "A decision tree regression algorithm (Breiman et al., 1984) is used to divide the distribution into two parts by limiting the maximum number of possible branches to two, corresponding to the portion of the distribution with relatively constant slope and the rest of the distribution. This is effectively done by fitting a simple piecewise-constant function $\hat{C}(a)$ to the derivative of $C(a)$, as shown in Figure 3a, where the breakpoint \hat{a} which minimizes the error between the distribution $C(a)$ and $\hat{C}(a)$. Here, the error is defined as the mean square error".

Typos

Thanks for catching these. We will fix them in the revised manuscript.