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

Loren Oh*

June 25th, 2024

We appreciate the constructive feedback for the submitted 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.

We agree that a true signal will show up in a spectral analysis, but we do not agree that noise would always behave so nicely, especially since we are dealing with a very noisy time-series where the noise occurs at different scales. Here, 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. 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

*Corresponding author: loh@eoas.ubc.ca

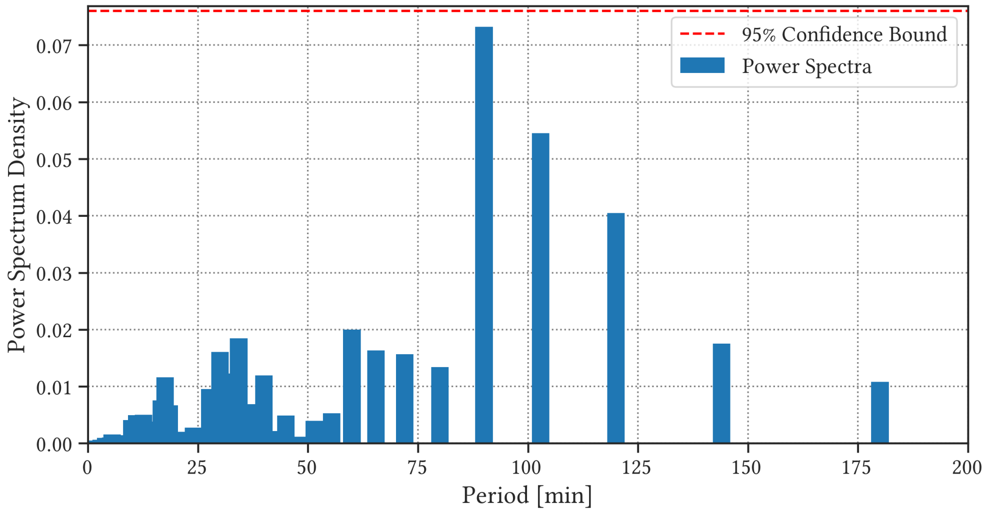

Figure 1: Power spectral density from FFT, with periods larger than 200 minutes filtered out.

can tell from the spectral analysis, the time-series indeed contains large variability for the spectral analysis to be effective (more on this later).

> 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.

We understand the concern, but the reviewer is conflating band-pass filtering for spectral analysis and gradient descent algorithm used by the GP regression method. 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 deterministic; the only reason we repeated the regression over a range of periods is to ensure that there are no local minima with steep gradient. In most cases the GP model converged nicely to the 90-minute periodicity. Furthermore, there are reasons to be skeptical of the 230-minute period – it is too long for either the lifetime of shallow clouds or cloud organization, it could be a harmonic of higher-frequency oscillations that are hidden in the noise, and the frequency bin of the DFT is too large at this scale.

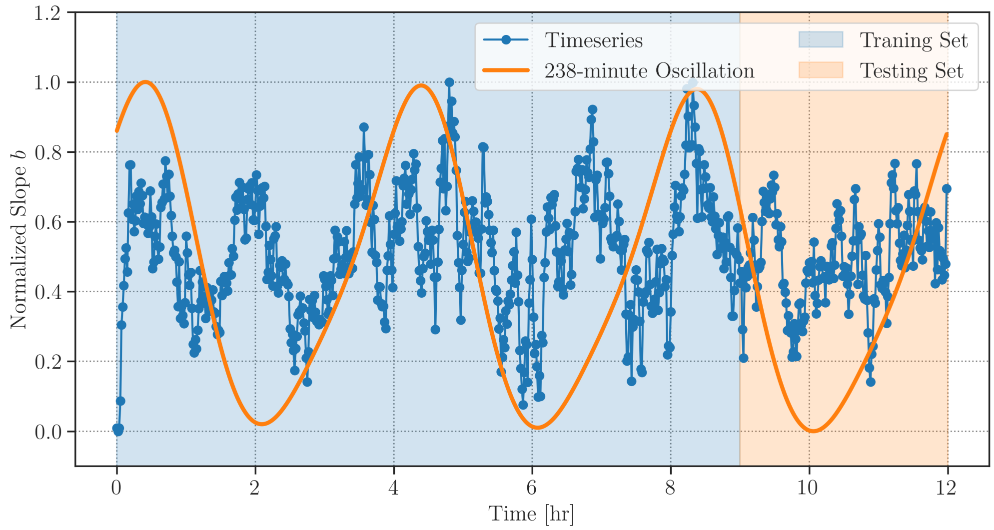

Figure 2: Normalized time-series of slope b of the cloud size distribution C(a) (blue line), compared to the the mean posterior from the periodic GP model (orange line), representing the 238-minute oscillation.

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 the model converged to an oscillation with a period of 238 minutes. See Figure 2; it is clear that the longer period ignores much of the measured time-series.

Still, we do need to acknowledge that the spectral analysis *did* find a signal at 90 minutes. We have added more details on this matter.

> 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

I think the reviewer meant to say point 1b (correct me if I am wrong).

Taking the derivative for the GP regression serves two purposes. First, as we have noted in the manuscript, it detrends the time-series, and not just a linear trend; oftentimes the trend is difficult to remove by a linear regression. Second, it normalizes/standardizes the time-series, which makes the GP model more stable. It is not particularly necessary for this case, and we do obtain the same result without taking the derivative (albeit with more steps), but we would say it is advisable to take the derivative of a noisy time-series.

We admit that some details have been left out when we were trying to reduce the length of the manuscript for submission. We have added these details to the steps we have taken for the GP regrssion model.

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

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.

We will also expand our discussion on the effect of domain size.

> 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!

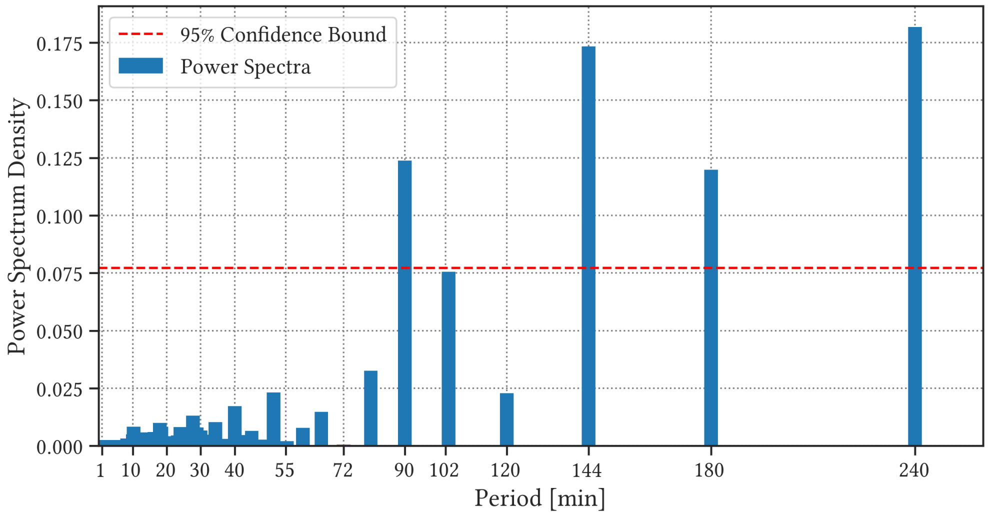

Figure 3: Power spectral density from FFT for a quarter of the domain.

We have definitely looked at the snapshots to see what is interrupting the oscillatory behaviour. At around 10 hours into the sampling period, multiple columns of clouds merge to form a large anvil-like structure near the top of the cloud layer. It is not clear if this is due to convective organization, as based on a visual inspection, clouds seem to simply follow the mean wind (along the longer side of the domain). We revised the manuscript to include more details about these observations.

> 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 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.

It is not. The width (or the shorter end) of the bowling-alley domain 12.8 km, which would be long enough for a square domain (12.8 km $\times$ 12.8 km). The width of the largest cloud is only roughly 1 km or so. Moreover, because of the way the LES model is set up (no large-scale forcing, etc), clouds will grow towards the size of the domain given the right conditions, in which case we would know that the size of the domain is not large enough. The simulation presented here showed no signs of that happening. There are no previous studies specifically for the CGILS S6 case, but doi.org/10.5194/gmd-6-1261-2013 should give you a good idea. This is based on a BOMEX case, and size distributions behave similarly in the two cases.

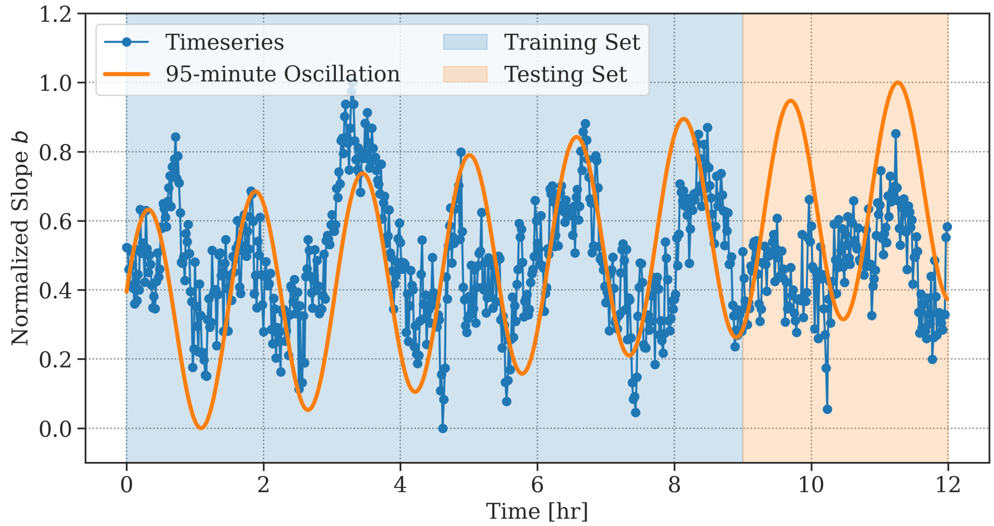

Figure 4: Normalized time-series of slope b of the cloud size distribution C(a) (blue line), compared to the the mean posterior from the periodic GP model (orange line), representing the 95-minute oscillation.

> 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?

As it is described in line 103, we take the horizontal cross-sections of all cloudy regions. We also sample 60% of these cloud samples to reduce the number of cloud samples (horizontal cross-sections) that are connected vertically. This detail must have been removed when we were trying to shorten the manuscript (as it was too long before the submission). Still, sampling all horizontal cross-sections does not change our result, and just makes the variability in the time-series somewhat larger.

> 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.

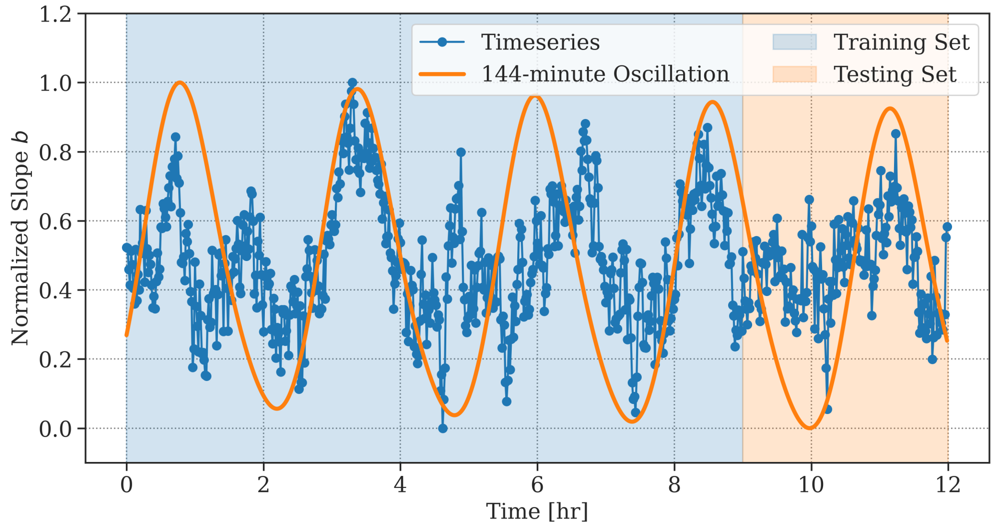

Figure 5: Normalized time-series of slope b of the cloud size distribution C(a) (blue line), compared to the the mean posterior from the periodic GP model (orange line), representing the 144-minute oscillation.

We tried to make no prior assumptions about the nature of the oscillatory behaviour, and we simply found the 90-minute periodicity to be consistent with previous studies.

We are working on a follow-up paper on the effect of the properties of the cloud field on this oscillatory behaviour, but we will add a discussion on possible factors that can affect this oscillation.

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.

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.

We will work on this, but it might be more appropriate for a follow-up paper looking at the physical mechanisms of the oscillatory behaviour.

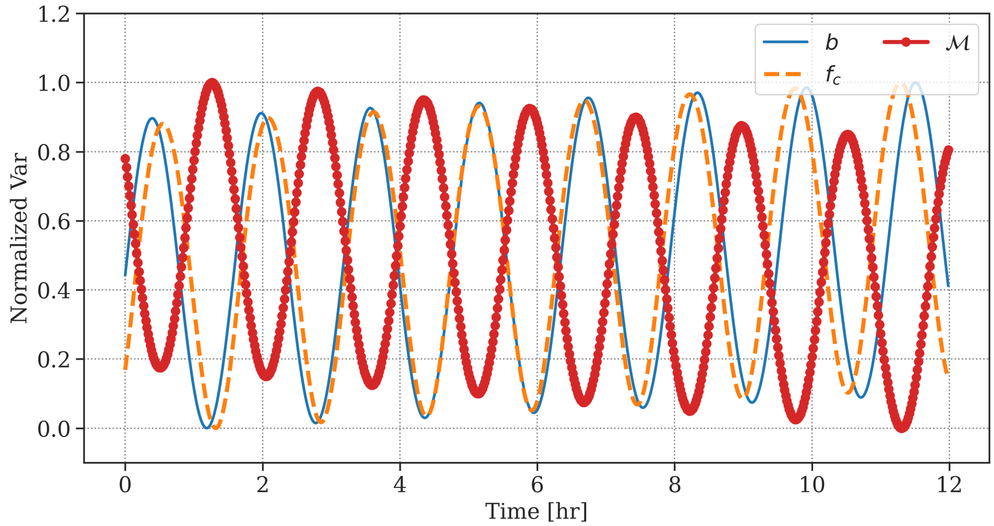

Figure 6: Mean posterior distributions from the GP regression for the slope b of the cloud size distribution (blue), cloud fraction f_c (orange) and average vertical mass flux $\mathcal{M}$.

> 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 definitely thought of that, and do show more snapshots when we present this particular research in person, but the manuscript was already too long.

> 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.

This definitely has to do with inertia (i.e. traditional literature) and we do not need it once we use KDE. I will clean up the section and make it clear that those are two different things.

> 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.

This might have to do with a typo in the paragraph that introduces Figure 3. We will re-write this paragraph to ensure the steps taken to obtain the figure. We will also calculate the RMSE for the linear regression when we calculate the slopes of the cloud size distributions.

Typos

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