Quantifying the Oscillatory Evolution of Simulated Boundary-Layer Cloud Fields Using Gaussian Process Regression

Gunho (Loren) Oh[1] and Philip H. Austin[1]

[1]Department of Earth, Ocean and Atmospheric Sciences, University of British Columbia, Vancouver, BC, Canada

Correspondence: Gunho (Loren) Oh (loh@eoas.ubc.ca)

Abstract. Average properties of the cloud field, such as cloud size distribution and cloud fraction, have previously been observed to ~~show periodic, oscillatory changes~~evolve periodically. Identifying this behaviour, however, remains difficult due to the intrinsic variability ~~of~~ within the boundary-layer cloud ~~distribution. We use the~~ field. We apply Gaussian Process (GP) ~~regression to identify this~~ machine-learning model to the regression of the oscillatory behaviour in the statistical distributions of individual cloud properties. Individual cloud samples are retrieved from high-resolution ~~LES model results~~large-eddy simulation, and the ~~distribution of cloud sizes is modelled as~~ cloud size distribution is modelled based on a power-law ~~distribution~~fit. We construct the time-series for the slope of the cloud size distribution b, a slope that is consistent with satellite observations of marine boundary-layer clouds, by observing the changes in the slope of the modelled cloud size distribution. Then, we build a GP model based on prior assumptions about the cloud field following observational studies: a boundary-layer cloud field goes through a phase of relatively strong convection where large clouds dominate, followed by a phase of relatively weak convection where precipitation ~~causes~~ leads to formation of cold pools and suppression of convective growth. The GP model successfully identifies ~~oscillatory motions~~the oscillatory behaviour from the noisy time-series, with a period of 95 ± 3.2 minutes. Furthermore, we examine the time-series of cloud fraction f_c and average vertical mass flux $\overline{M}$, whose periods were 93 ± 2.5 and 93 ± 3.7 minutes, respectively. The oscillations reveal the role of precipitation in governing convective activities through recharge-discharge cycles.

1 Introduction

Cumulus convection plays a central role in regulating the moisture and energy budget in the atmosphere. However, representing the effect of moist convection remains difficult for general circulation and weather prediction models ~~(Bony and Dufresne, 2005; Bony et al., 2006). The~~ (Bony and Dufresne, 2005; Bony et al., 2006); for example, the cloud radiative feedback remains the largest source of uncertainty in the ~~estimation~~ estimates of equilibrium climate sensitivity ~~(ECS) (Ceppi et al., 2017; Mauritsen and Roeckner, 2020; Zelinka et al., 2020), and convective parameterization of cumulus effects~~ (Ceppi et al., 2017; Mauritsen and Roeckner, 2020; Zelinka et al., 2020), and radiative effects of moist convection remains poorly constrained. ~~Large-scale climate model parameterization schemes employ different~~These models make widely differing assumptions about the dynamics and thermodynamics of boundary-layer clouds ~~(Ceppi et al., 2017; Myers and Norris, 2016; Lipat et al., 2018)~~(Ceppi et al., 2017; Lipat et al., 2018; Myers and Norris, 2016) that cannot be directly resolved.

The resolution required to accurately model shallow cumulus convection, on the order of 10 m in the boundary layer ~~(Sato et al., 2017, 2018)~~(Sato et al., 2017, 2018), still remains computationally prohibitive (*cf.* Figure 2 in Schneider et al., 2017), as short-term simulations of the global climate ~~on the scales~~ with spatial resolutions on the order of a few kilometres have only recently been ~~performed Stevens et al. (2019)~~introduced (Stevens et al., 2019). As such, the role of high-resolution ~~LES~~ large-eddy simulation (LES) models in improving our understanding of convective effects continues to be essential.

~~For an accurate~~ To improve the indirect representation of the effects of moist convection, ~~a better understanding of the complex~~ large-scale models of the atmosphere must account for the large variability in the dynamics and thermodynamics of the ~~cloud field is required. A radiative scheme in a global climate model~~ sub-grid scale cloud field. For example, a sub-grid scale radiation scheme that approximates the radiative effects of shallow convection would greatly benefit from having a better estimate about the geometrical structure and the distribution of clouds, which can be used to improve our estimate of the short-wave radiative effects of low clouds Ceppi et al. (2017). An important aspect of the cloud field in this context is the ~~probability~~ distribution of cloud sizes (Neggers et al., 2003), which has long been the main topic for observational studies, from aircraft measurements and satellite imagery ~~(Plank, 1969; Machado et al., 1992; Machado and Rossow, 1993; Wilcox and Ramanathan, 2001; Peters et al.,~~ (Benner and Curry, 1998; Berg and Stull, 2002; Machado et al., 1992; Machado and Rossow, 1993; Plank, 1969; Peters et al., 2009; Raga to numerical simulations ~~of cloud field (Brown et al., 2002; Neggers et al., 2003; Garrett et al., 2018).~~

~~An important insight from recent numerical studies of~~ (Brown et al., 2002; Garrett et al., 2018; Neggers et al., 2003) of the cloud field~~using LES models is that the evolution of the cloud field appears to oscillatory (Feingold et al., 2017; Koren and Feingold, 2011,~~ ~~. Convective self-organization drives this temporal oscillation in marine~~ .

~~Marine~~ boundary-layer clouds ~~. Evaporation of precipitation from large, well-developed clouds form cold pools below the cloud base, which promotes the~~ have been observed to organize into cellular patterns (Malkus and Riehl, 1964; Nair et al., 1998; Seifert and as a response to the formation of cold pools, formed by evaporative cooling due to precipitation (Zuidema et al., 2012). The cold pools promote the formation of negatively buoyant downdrafts that inhibit further growth of thermals ~~(Seifert and Heus, 2013; Seifert et al.,~~ ~~. The~~ (Seifert and Heus, 2013; Seifert et al., 2015; Seigel, 2014a). At the boundaries of these open cells, convective formation is promoted due to the moistening of downdrafts (Seifert and Heus, 2013) and mechanical lifting due to the convergence of cold pool outflows (Xue et al., 2008).

This dynamics between the formation of cold pools from precipitation and the subsequent formation of clouds has been observed to manifest as temporal oscillations; the cloud field goes through a ~~relatively weak convective phase~~phase of relatively weak convection, until multiple downdrafts from the cold pools collide into a convergence zone, where convective growth begins anew. For mesoscale marine boundary-layer stratocumulus clouds, the spatial organization of precipitation is found to be important in promoting subsequent cloud formation and the evolution of open cell convection (Feingold et al., 2010; Koren and Feingold, 20 . High-resolution ~~studies have confirmed~~ large-eddy simulations have shown that the formation of cold pools ~~as~~ is the main mechanism that drives organized ~~formation of~~ marine stratocumulus convection, which corresponds well to long-term satellite ~~measurements of cold pools (Zuidema et al., 2012; Seifert and Heus, 2013).~~ observations (Bretherton and Blossey, 2017; Seifert and Heus, .

60 The temporal oscillation has also been observed in modelling studies of precipitating cumulus convection (Dagan et al., 2018; Feingold et ... . For both shallow and deep clouds, the dominant mechanism that drives this oscillatory evolution is found to be the formation of cold pools due to evaporative cooling from precipitation in the sub-cloud layer (Seifert and Heus, 2013; Yano and Plant, 2012; Tompkins, 2... . This mechanism is referred to as the recharge-discharge cycle of thermodynamic instability by Dagan et al. (2018), motivated by Bladé and Hartmann (1993), where the evaporative cooling due to precipitation *charges* instability in the atmosphere, which

65 is *discharged* by convection.

Precipitation facilitates both the spatial organization and temporal oscillation of the cloud field, and is governed by a number of factors including cloud microphysics and cloud layer depth. Aerosols, acting as cloud condensation nuclei (CCN), can influence the cloud microphysics by enhancing the cloud droplet number concentration but suppress droplet growth (Twomey, 1974), and LES studies have shown that when the aerosol concentration is increased, the cloud layer deepens,

70 which then affects rain formation (Dagan et al., 2017; Seifert et al., 2015). Furthermore, modelling studies have shown that an increase in aerosol concentration influences both the amount and the timing of precipitation; in a polluted environment, the efficiency in precipitation formation is reduced, and as a result, rain formation is suppressed and delayed (Dagan et al., 2018; Seigel, 2014b; .

This study is motivated by these observations of the periodic evolution of the ~~observation that the~~ cloud size distribu-

75 tion~~evolves periodically, indicating that~~; if the cloud size distribution evolves periodically over time, the dynamic and thermo-dynamic properties of the cloud field, such as mass flux and cloud cover, ~~also oscillate~~ must also oscillate accordingly as the cloud field alternates between the two phases of strong and weak ~~convective phases~~convection. However, numerical algorithms used ~~for period detection~~to detect the period of such oscillation suffer from the presence of noise. Because of that, earlier studies concerning the oscillatory growth of cumulus field relied on manual inspections ~~(Dagan et al., 2018; Koren and Feingold, 2011)~~

80 (Dagan et al., 2018; Koren and Feingold, 2011) or spectral analysis (Feingold et al., 2017) although ~~it remains difficult to accurately determine~~ an accurate estimate of the period of this ~~oscillation.~~ oscillatory behaviour is yet to be found. Given that, we use the ~~Gaussian process~~ kernel-based machine-learning method using the Gaussian Process (GP) ~~regression method to estimate the periodicity in noisy~~ model and apply it to the regression of the periodicity in the observed time-series of ~~cloud size distributions. A Gaussian process~~ the cloud size distribution. GP is a flexible non-parametric Bayesian machine-learning

85 model that is well-suited to be used to infer useful information from a noisy dataset. ~~GP modelling has~~

GP models have been used for this specific purpose in astronomy (Angus et al., 2017) and medical studies ~~(Durrande et al., 2016; Cheng e... to infer~~ (Cheng et al., 2020; Durrande et al., 2016) to estimate the periodicity in a time-series with observational noise and extract an underlying trend in the observed data.

~~We~~ Likewise, we propose the use of a ~~Gaussian process~~ GP model to obtain the underlying trend of a time-series in the

90 presence of observational noise. We use LES model results to sample cloud size densities, construct a cloud size distribution, and follow the evolution of the cloud field properties. Based on the periodic behaviour estimated by the GP model, we reconstruct the periodic time-series $f(x)$ without noise, which is compared to the original time-series of cloud size distribution. Lastly, we repeat the GP regression for mass flux and cloud cover to test our hypothesis.

~~The large-eddy simulation is described~~ Details of the LES model run is given in Section 2.1. Section 2.2 illustrates how the
individual clouds are sampled from the model output. Section 2.3 gives the methods used to construct the time-series of the
slope b from the cloud size distribution. Section 2.4 examines the traditional Fourier spectral analysis to identify the oscillatory
behaviour within the time-series. A brief introduction of ~~Gaussian process regression~~ the GP regression method is given in
Section 2.5, as well as the construction of kernels, which is further explained in Section 2.6, where we present the method
used to estimate the periodicity from the time-series. The fully Bayesian model to estimate the uncertainty in the periodicity
estimate is given in Section 2.7. The results are discussed in Section 3 and summarized in Section 4.

2 Methods

2.1 Model Description

The System for Atmospheric Modeling ~~(SAM; Khairoutdinov and Randall (2003))~~ (SAM; Khairoutdinov and Randall, 2003)
version 6.11.8 was used to simulate ~~CGILS (the~~ CFMIP/GASS Inter-comparison of Large-Eddy and Single-Column Models ~~)~~
~~trade cumulus boundary-layer (case~~ case (CGILS; Blossey et al., 2013; Zhang et al., 2013). In this study, we use the large-scale
forcing and thermodynamic tendencies of the CGILS S6 ~~) (Blossey et al., 2013).~~ regime, representing marine sub-tropical
shallow cumulus convection.

The model grid size was set to 25 m in all directions over a 43.2 km $\times$ 12.8 km $\times$ 4.8 km model domain. This elongated,
bowling alley domain was employed to minimize the effect of periodic boundaries. Because the mean air flows along the
elongated axis, most of the convective activity occurs without being spatially recycled as clouds rarely veer from the mean
airflow. The temporal resolution of the model was 1 second, and the output was written every minute. We performed a total of
36 hours of simulation, although the first 24 hours were used for the model spin-up and therefore excluded from the analysis,
and the last 720 time steps for the remaining 12 hours of simulation were used for the analysis. A long spin-up time was
necessary for the boundary layer to reach (quasi-)stability across the domain. The cloud field was then sampled every minute.
Details of conditional cloud sampling are presented in the following section.

Doubly periodic domains were employed with a soft top to exclude gravity waves as the possible cause of oscillatory be-
haviour (Dagan et al., 2018). A two-moment microphysics scheme ~~(Morrison et al., 2005a, b) has been employed~~(Morrison et al., 2005a, b)
was used, although the cloud layer remained shallow within 2 km of the boundary layer and no ice formation was observed.
The ~~LES model~~ microphysics scheme was initialized with ~~the parameters~~ a cloud droplet number concentration (CDNC)
value of 120 cm^{-3}, which is typical for maritime convection (Rasmussen et al., 2002). Compared to other relevant studies of
aerosol-cloud interactions, this represents a relatively pristine environment (Dagan et al., 2018; Seigel, 2014b; Yamaguchi et al., 2019)
. Other parameters used in the LES model run are based on CGILS S6 control run ~~(Tan et al., 2016)~~(Blossey et al., 2013; Tan et al., 2016; Zh
. A diurnally averaged solar insolation was applied, and both short-wave and long-wave radiative effects were calculated using
the Rapid Radiative Transfer Model (RRTM) (Clough et al., 2005; Iacono et al., 2008).

To make accurate measurements of vertical mass flux, we implemented the tetrahedral interpolation scheme (Dawe and
Austin, 2011). Every time instantaneous cloud fields are being sampled, mass flux rates are integrated over a cloudy surface.
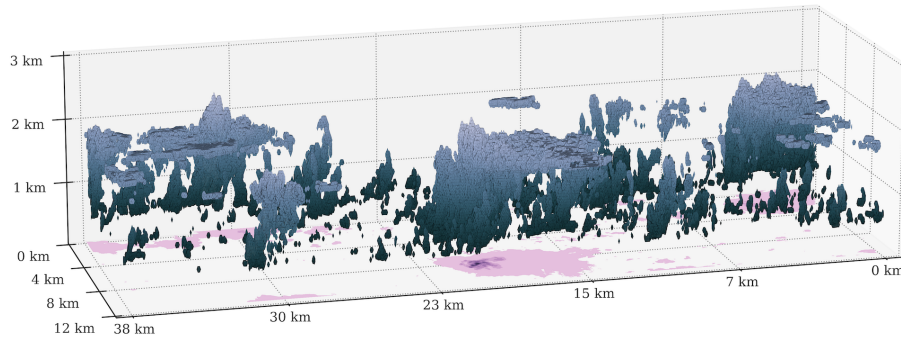
Figure 1. A 3-dimensional overview of the LES modelled cloud field from SAM model output, taken 24 hours into the simulation. The shaded regions indicate cloudy cells that contain condensed liquid water. Darker regions are low in altitude, and brighter regions are closer to the cloud top at ~~1.8~~ 2 km. Regions shaded in pink represent areas containing column-integrated precipitable water.

This is done by interpolating each grid cell by 48 tetrahedrons, and calculating the changes in the 3-dimensional cloud surface. Recent studies suggest a grid size of roughly 10 m to achieve meaningful statistical accuracy even for shallow convective clouds (Sato et al., 2017, 2018), but ~~running the LES model~~ the results from the LES model run with 12.5 m grids did not ~~yield significant improvements to mass flux rates~~ appear to cause significant changes in average vertical mass flux distributions compared to 25 m grids~~, likely thanks to the sub-grid interpolation scheme~~. Since computational resources are limited, and keeping a large domain size (and hence a large number of statistically independent cloud samples) and a longer simulation time is more important for the purpose of this study, we will only examine the results from the 25-m resolution LES run.

Figure 1 shows a three-dimensional snapshot from the LES model run. Clouds are generally scattered over the model domain, and the distribution of smaller clouds ~~appear~~ appears to be granular. However, areas of vigorous convective activities accompanying precipitation often form in clusters (right side of Figure 1) between areas that are either devoid of clouds or dominated by scattered small clouds (middle of Figure 1). Earlier studies of cloud patterns ~~(Seifert and Heus, 2013; Stevens et al., 2020)~~ (Seifert and Heus, 2013; Stevens et al., 2020) suggest such gravel-like patterns could be due to the formation of cold pools. The evaporative cooling due to precipitation can form cold pools, and such patterns can manifest as pronounced convective activities followed by weak, scattered cloudy regime on the leeward side. There are signs of cloud clustering in the simulated cloud field where strongest convective activities (right side of Figure 1, for example) mostly occur in clusters.

2.2 Cloud Sampling

In order to obtain the cloud size ~~distributions~~ distribution, individual clouds ~~were conditionally sampled~~ 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 ~~with~~ containing condensed liquid water multiplied by the horizontal grid size (25 m × 25 m). ~~It is also possible to take the square root of this size as a proxy for cloud radius R (Neggers et al., 2003; Dawe and Austin, 2013) and~~

~~find its correlation with dynamic properties of the cloud (especially entrainment and mass flux, for example), but in this paper, we are interested in the evolution of the cloud size distribution over time~~The size distribution contains multiple horizontal cross-sections of the same cloud. However, randomly sampling 60% of the cloud samples did not result in significant changes in the time-series, likely thanks to the sheer number of samples and the shallow nature of the modelled clouds in comparison to the size of the domain, and to the numerical methods used to reduce the statistical uncertainties.

An alternative method ~~to define the cloud region has been introduced by Neggers et al. (2003) , where the~~ by Neggers et al. (2003) is to take a vertical projection of each cloud volume ~~was taken to~~ upon a two-dimensional surface. This resembles two-dimensional images taken from a high altitude, which can be useful when comparing the LES output to satellite images, for example. However, the focus of the study is to evaluate the dynamic and thermodynamic properties of the clouds, ~~such as entrainment and mass flux, and~~ and simply taking the horizontal cross-section allows us to directly compare the distribution of cloud sizes to that of mass flux. There is another benefit to this approach: when the cloud field is projected onto a two-dimensional surface, the number of smaller clouds sampled from the cloud field increases. This is because smaller clouds are less likely to be projected onto each other. Hence, vertically projected cloud size distribution tends to overestimate the number of smaller clouds, while taking a horizontal cross section (e.g. Brown, 1999) gives a better representation of the realistic three-dimensional cloud size distribution.

Once individual cloud sizes are extracted from the LES output field, we ~~can~~ define the *cloud size distribution*. The cloud size distribution $C(a)$ is a cumulative distribution defined as an integral over the *cloud size density* $c(a)$. The cloud size density is a type of probability density function (PDF) that defines the probability of a cloud having a certain size.

Typically, the probability density function is calculated using a histogram with a discrete bin size for a ~~piecewise estimate. Here,~~ piece-wise estimate, where the density is defined as the frequency of clouds within ~~a discrete range~~each discrete *bin*. However, the choice of ~~this discrete~~ bin size has a large effect on the resulting distribution, and can yield different results with qualitatively independent features depending arbitrarily on the choice of bin size.

To alleviate these issues, we use the Kernel Density Estimator (KDE; Parzen, 1962a) to reliably estimate the distribution of cloud size density. A kernel density estimator $\hat{k}(x)$ at a point x given a set of n observations $X_1, \ldots, X_n$ is defined as

$$\hat{k}(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right) \tag{1}$$

which is governed by a *kernel function* K and its *bandwidth* h that control the amount of smoothing applied by the kernel.

We will use a Gaussian kernel $K(x)$ for this purpose, defined as

$$K(x) = \frac{1}{2\pi} e^{-x^2/2} \tag{2}$$

which will be used to smooth the cloud size density function. Each cloud size sample is added to a distribution not as a single point of observation, but a probability distribution based on a Gaussian distribution. It can be considered as an uncertainty in the measurement; ~~we associate each cloud size sample with an uncertainty from the observation,~~ that is, each cloud sample is considered to be a Gaussian probability distribution whose width is defined by the bandwidth h.
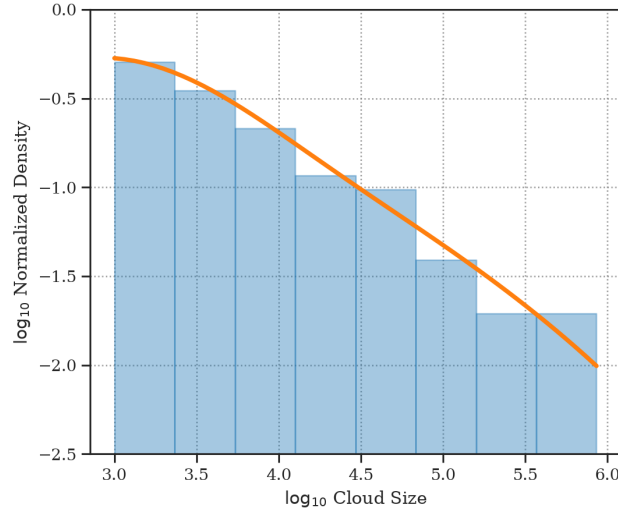
Figure 2. A comparison of the histogram (blue) and the kernel density estimate (KDE; orange line) showing $\log_{10}$ of normalized density based on the cloud size distribution over $\log_{10}$ of cloud size in m^2.

180   The KDE integrates these probability distributions of cloud sizes, which gives the cloud size distribution $C(a)$. Figure 2 shows the histogram as well as the cloud size distribution based on the KDE using cloud samples taken at 12 hours into the simulation.

     The cloud size distribution $C(a)$ in Figure 2 shows a decreasing slope with a scale break for the smallest clouds. The probability density for the smallest clouds is nearly constant. The distribution resembles that of Brown (1999) (*cf.* Figure 12 in

185 Neggers et al., 2003), but with a greater number of smaller clouds. This could be because smaller clouds appear near the cloud base, which is lower than the sampling height used in Brown (1999), reproduced in Neggers et al. (2003). Since the cloud samples in Figure 2 have been taken at all heights, the transition between the smaller and the larger clouds appears to be a lot less abrupt than previously observed.

2.3 Cloud Size Model

190 Given the cloud size distribution $C(a)$, a model of the probability density function needs to be constructed in order to study its temporal evolution. In this paper, we use the power-law distribution ~~(Cahalan and Joseph, 1989; Kuo et al., 1993; Benner and Curry, 1998;~~ (Benner and Curry, 1998; Cahalan and Joseph, 1989; Feingold et al., 2017; Kuo et al., 1993; Neggers et al., 2003; Zhao and Di Girolamo, , which has been widely used to represent the observed distribution of clouds. Here, the cloud size density $c(a)$ is defined as a function of cloud size, or

195    $c(a) = c_0\, a^b$ (3)

where a is the cross-sectional cloud area in m^2, and c_0 is the coefficient used for the power-law fit. Integrating the cloud size density $c(a)$ over all observed cloud sizes a yields the cloud size distribution $C(a)$.

From Figure 2, we observe that the cloud size distribution can roughly be divided into two parts, defined by a scale break; the cloud density appears to be relatively constant for the smallest clouds before it decreases linearly. This is a useful feature for the purpose of this paper. As smaller clouds are short-lived and their contribution to the upward mass flux M is very small, we are interested in the oscillation between the two phases of the cloud field where there are a relative abundance of intermediate-sized clouds, which contributes the most to the mass flux, and where there are a relative abundance of large clouds, which contributes the most to precipitation and formation of cold pools~~(Dagan et al., 2018)~~.

We isolate the (quasi-)linear portion by first taking the derivative of the cloud size distribution $C(a)$. 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 ~~into~~ 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 ~~$\hat{C}(a)$~~ $\bar{C}(a)$ to the derivative of $C(a)$, ~~with asingle~~ 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~~ $\bar{C}(a)$. For this purpose, we make use of the mean square error (MSE) defined as

$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N}\left(C(a_i) - \bar{C}(a_i)\right)^2 \tag{4}$$

for N samples in the cloud size distribution $C(a)$. As shown in Figure 3a, by minimizing MSE, the decision tree regression algorithm isolates the linear portion of the distribution (green region).

Figure 3b shows how the decision tree splits the distribution based on the derivative of $C(a)$. Once we isolate the relatively linear portion of the distribution, we use the Theil-Sen estimator (Theil, 1950; Sen, 1968) to perform a robust linear regression, which does well in the presence of outliers and small deviations from the linear trend, as seen in Figure 3a, where the slope represents the ratio between medium-sized and largest clouds. ~~Here, the slope is measured to be roughly $b \approx -0.50$. This is smaller than the values reported from previous observations of boundary-layer clouds (Cahalan and Joseph, 1989; Benner and Curry, 1998). The~~ The cloud size distribution $C(a)$ given in Figure 3~~is a (normalized)~~b represents a normalized probability density function, which will differ from the histograms obtained from observations. 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 method 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; .

We repeat ~~this process for the entire model run, which consists of~~ the calculation of the slope b for 720 time steps~~for~~, or the entire duration of the simulation excluding the spin-up time. The resulting time-series of the slope b for the cloud size distribution $C(a)$ can be seen in Figure 4. The use of a robust linear regressor along with a decision tree regression algorithm helps isolate the linear segment of the distribution, ~~representing~~ which better represents the slope of size distribution of the cloud field and reduces numerical uncertainties involved in the linear regression to obtain the slope b.
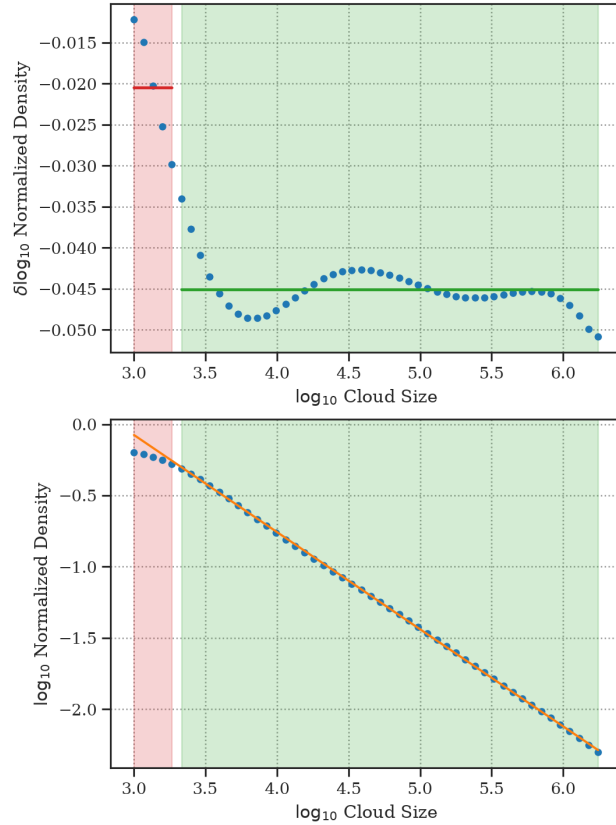
Figure 3. The decision tree regression algorithm is applied to the derivative of cloud size distribution (top), which is divided into a region of relatively constant slope (top; green region) and the rest of the distribution (top; red region). The green region is chosen by the algorithm, which is then used to separate the linear portion of $C(a)$ (bottom; green region) from the non-linear portion (bottom; red region). The green region, defined by the part of the distribution with constant slope (bottom; green region), is used to fit a linear curve (bottom; orange line) while the rest of the distribution is ignored (bottom; red region) in order to estimate the slope of cloud size distribution.

230   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 shows that the time-series is non-stationary. 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~~ perform Fourier spectral analysis to
235   identify the underlying periodic behaviour in the observed time-series.
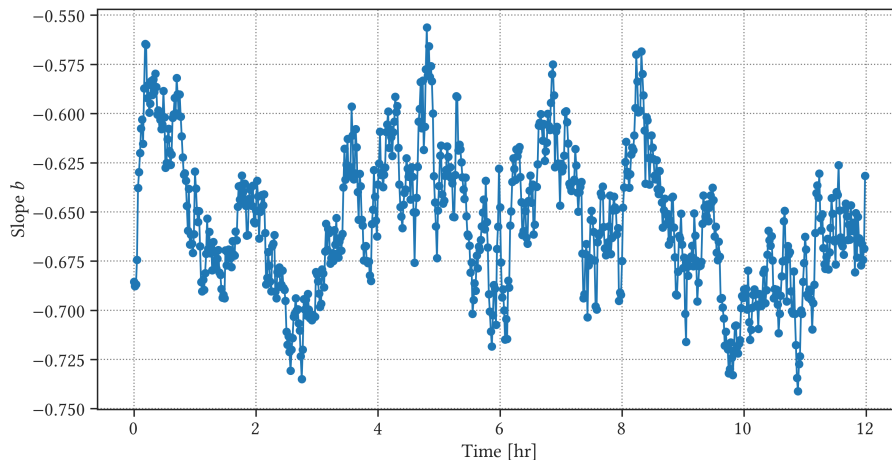
Figure 4. The time-series of slope b of the cloud size distribution $C(a)$ for the 12-hour simulation. The slope is calculated every minute (blue line) for the duration of the entire simulation, but the first 9 hours are used for training the GP model.

2.4 Fourier Spectral Analysis

Given the noisy, non-stationary time-series ~~in Figure 4, we perform a spectral analysis using~~ of the slope $b(t)$ (Figure 4), we tested the traditional spectral analysis based on discrete Fourier transform (DFT) to capture possible oscillatory ~~motions in the observed time-series~~ behaviour (Feingold et al., 2017). Here, given a discrete time-series $f_k = f(k/N)$ for $(k = 0, 1, \ldots, N-1)$, the corresponding DFT, $(\mathcal{F}(f_0), \mathcal{F}(f_1), \ldots, \mathcal{F}(f_{N-1}))$, can be obtained by

$$\mathcal{F}(f_k) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} f(n) \cdot \exp\left(-2\pi i \frac{kn}{N}\right) \tag{5}$$

where $k = 0, 1, \ldots, N - 1$.

Equation 5 translates the observed time-series f_k into a function of frequency k/N that is a linear combination of oscillatory, or sinusoidal, components. The strength of each component can then be observed by examining the *power spectral density* of the time-series. The oscillatory components showing the strongest signals represent the dominant modes of oscillation in the observed time-series.

The power spectral density for a DFT of on a discrete sequence can be estimated by a periodogram, defined as the squared modulus of DFT, or,

$$\mathcal{P}(f_k) = |\mathcal{F}(f_k)|^2 \tag{6}$$

$$= \frac{1}{N} \left(\sum_{n=0}^{N-1} f(n) \cdot \exp\left(-2\pi i \frac{kn}{N}\right)\right)^2 \tag{7}$$

where $k = 0, 1, \ldots, (N-1)/2$ for real-valued input sequence.
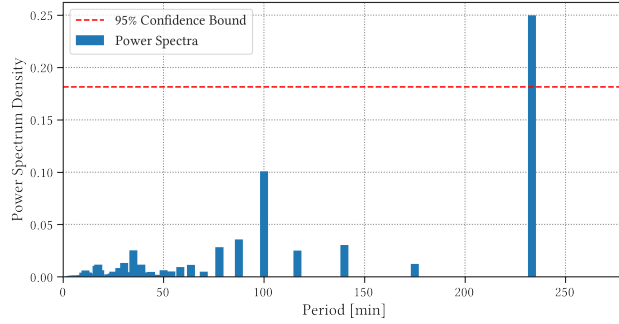
Figure 5. The power spectral density (blue) and 95% confidence interval for the time-series $b(t)$ based on the slope of the cloud size distribution $C(a)$, obtained from Fourier spectral analysis.

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 χ^2 with 2 degrees of freedom (Panofsky and Brier, 1958).

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 ~~oscillations~~ 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 periodigram, especially for $T = 233$ minutes. Each coefficient in the periodogram corresponds to period $T(k) = N/k$ for $k = 0, 1, \ldots, N-1$. The longer the period, the coarser the resolution, which greatly reduces the effectiveness of the periodogram in isolating oscillatory components from the noisy time-series. 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 noise and non-stationarity.

We have also tested other methods to estimate power spectral density, such as the circular autocorrelation function ~~(ACF; Fisher and Lee, , but the~~ (ACF; Parzen, 1962b), but (partial) autocorrelation function of the observed time-series $b(t)$ decreases slowly over time, and no significant lag can be found. The presence of noise ~~hinders their ability in detecting the underlying oscillatory behaviour and no modes were detected for periods shorter than 100 minutes~~ as well as the non-stationary nature of the time-series makes it difficult to examine the behaviour of the time-series.

2.5 Gaussian Process Regression

As shown in previous section, ~~it is possible that the internal fluctuations in cumulus convection and the noise from numerical instabilities and uncertainties~~ traditional methods struggle to identify the underlying oscillatory behaviour in the observed time-series. Internal fluctuations in the cloud field and the uncertainties from the numerical methods used in the estima-

tion ~~of the slope in~~ for the slope of the cloud size distribution $C(a)$ ~~mask an underlying oscillation in the time-series.~~

275 ~~We use~~ can mask the underlying trend in the observation. To handle the uncertainty, we make use of the Gaussian Process (GP) regression ~~, a Bayesian inference method which utilizes GP as a prior distribution(Rasmussen and Williams, 2006)~~ (Rasmussen and Williams, 2006). A *Gaussian process* is a set of random variables, any finite number of which have a joint Gaussian distribution. GP models can ~~be used to fit an oscillatory model that specifically includes the noise from the internal fluctuations of cumulus convection and numerical uncertainties. The GP model can~~ explicitly model and learn the noise level

280 directly from the ~~observation~~ data by introducing an explicit noise term that ~~represents numerical instability and~~ we will use to represent numerical uncertainties involved in sampling and constructing the time-series $b(t)$ of the cloud size distribution~~, which are difficult to quantify using traditional methods~~.

The time-series can be considered as a regression problem ~~over observations x~~ for a latent function f over observations $\mathbf{y} = y(\mathbf{x})$ made at points $\mathbf{x} = x_1,\ldots,x_n$ with noise ϵ_y, or

285
$$\mathbf{y} = y(\mathbf{x}) = f(\mathbf{x}) + \epsilon_y. \tag{8}$$

Assuming no prior knowledge about the noise ϵ_y ~~in the observation, we can formalize the independent identically-distributed uncertainty as a~~, we formalize this uncertainty as an additive independent identically-distributed Gaussian distribution with zero mean and ~~variance~~ a variance of σ_y^2 (i.e. $\mathcal{N}(0,\sigma_y^2)$). The subscript indicates that the uncertainty comes from the ~~measurement in~~ noisy observation y~~, which includes the estimation of the slope b given in previous sections. Likewise, we can assume that~~

290 ~~the measurement of input points (time x) comes with (Gaussian) uncertainty ϵ_x.~~

In this framework, ~~observational uncertainty needs to be formalized in terms of probability distributions, assuming that the data y is only available through observations, and inherently includes noise due to estimates involved in the calculation of b based on cloud size distribution $C(a)$. Incorporating this uncertainty to our regression model is necessary to perform Gaussian process (GP) regression (Rasmussen and Williams, 2006). A *Gaussian process* is a set of random variables, any~~

295 ~~finite number of which have a joint Gaussian distribution (Definition 2.1 in Rasmussen and Williams, 2006), which means that every Gaussian process can be uniquely identified~~ the prior $p(\mathbf{y}|\mathbf{f})$ can be formalized as a (Gaussian) probability distribution, defined uniquely by its mean $\mu(\mathbf{x})$ and ~~a~~ covariance function $k(\mathbf{x_i}, \mathbf{x_j})$~~, which can be used to formalize a Gaussian process as~~:

$$p(\underline{\mathbf{x}}\mathbf{y}|\mathbf{f}) \sim \underline{\mathrm{GP}}\mathcal{N}\big(\mu(\mathbf{x}),\mathbf{K}\big) \tag{9}$$

where $\mathbf{K}$ is a n × n matrix of covariances ($K_{ij} = k(\mathbf{x_i},\mathbf{x_j})$) of joint distribution $p(\mathbf{x})$ evaluated at two arbitrary points, which

300 can then be used ~~as~~ to define a Bayesian prior ~~to make predictions from noisy data. The prior distribution~~ that reflects our belief about how the model should behave, prior to observing any data points.

~~Here, we present a brief description for Gaussian process regression; see, for example, Rasmussen and Williams (2006) for a more detailed description for the GP regression method.~~

For a set of observations points $\mathbf{x} = x_1, \ldots, x_n$, we can write the covariance matrix $\mathbf{K}(\mathbf{x}, \mathbf{x})$ as

$$\mathbf{K}(\mathbf{x},\mathbf{x}) = \begin{pmatrix} k(x_1,x_1) & k(x_1,x_2) & \cdots & k(x_1,x_n) \\ k(x_2,x_1) & k(x_2,x_2) & \cdots & k(x_2,x_n) \\ \vdots & \vdots & \ddots & \vdots \\ k(x_n,x_1) & k(x_n x_2) & \cdots & k(x_n,x_n) \end{pmatrix} \tag{10}$$

and because the noise is assumed to be independent, we can modify the covariance function to include the noise term. Because the additive, independent noise term only applies to the diagonal elements of the covariance matrix,

$$\mathrm{cov}(\mathbf{x},\mathbf{x}) = \mathbf{K}(\mathbf{x},\mathbf{x}) + \sigma^2 \mathbf{I} \tag{11}$$

where σ^2 is the variance of our (independent, Gaussian) noise term.

~~Then we can consider the joint distribution of a set of observations $\mathbf{y}$ made at $\mathbf{x}$.~~ To make predictions, we need to obtain the posterior distribution by conditioning the prior distribution using the observations. To this end, we set up a joint distribution between the observations made at training points $\mathbf{y} = f(\mathbf{x})$ and the function values ~~$\mathbf{f}_*$ taken at test points $\mathbf{x}_*$ as~~ at testing points $\mathbf{f}_* = f(\mathbf{x}_*)$ as

$$p\left(\begin{bmatrix} \mathbf{y} \\ \mathbf{f}_* \end{bmatrix}\right) = \mathcal{N}\left(\begin{bmatrix} \mu(\mathbf{x}) \\ \mu(\mathbf{x}_*) \end{bmatrix}, \begin{bmatrix} \mathbf{K}(\mathbf{x},\mathbf{x})+\sigma^2\mathbf{I} & \mathbf{K}(\mathbf{x},\mathbf{x}_*) \\ \mathbf{K}(\mathbf{x}_*,\mathbf{x}) & \mathbf{K}(\mathbf{x}_*,\mathbf{x}_*) \end{bmatrix}\right) \tag{12}$$

~~from which we can obtain the posterior distribution p_*~~ which can be conditioned by the observations $\mathbf{y}$ to yield the posterior distribution, which is also Gaussian, as

$$_* \triangleq p(\mathbf{f}_*|\mathbf{x}_*,\mathbf{y}) = \mathcal{N}(\mathbb{E}[\mathbf{f}_*], \mathbb{V}[\mathbf{f}_*]) \tag{13}$$

that is fully specified by ~~the posterior~~ mean and variance

$$\mathbb{E}[\mathbf{f}_*] = \boldsymbol{\mu}(\mathbf{x}_*) + \mathbf{K}(\mathbf{x}_*,\mathbf{x})\left(\mathbf{K}(\mathbf{x},\mathbf{x})+\sigma^2\mathbf{I}\right)^{-1}(\mathbf{y}(\mathbf{x}) - \boldsymbol{\mu}(\mathbf{x})) \tag{14}$$

$$\mathbb{V}[\mathbf{f}_*] = \mathbf{K}(\mathbf{x}_*,\mathbf{x}_*) - \mathbf{K}(\mathbf{x}_*,\mathbf{x})\left(\mathbf{K}(\mathbf{x},\mathbf{x})+\sigma^2\mathbf{I}\right)^{-1}\mathbf{K}(\mathbf{x},\mathbf{x}_*) \tag{15}$$

respectively.

A crucial step in Gaussian process regression is determining the covariance ~~$\mathbf{K}(\mathbf{x},\mathbf{x})$~~ $\mathbf{K}$, also called the *kernel* ~~function~~, which embodies the prior knowledge or our assumptions about the observed processes. A kernel function $k(x,x')$ determines how an arbitrary pair of sample points x and x' are ~~correlated~~ related to each other. Essentially, it reflects our belief about how the probability distribution (the target function $f(x)$ under the observed data, as seen in Equation 8) behaves.

The most widely used covariance function is the Square-Exponential (SE), defined as

$$k_{\mathrm{SE}} = \exp\left(-\frac{(x-x')^2}{2\lambda^2}\right) \tag{16}$$

where λ is typically defined as a length-scale, or a timescale for the time-series data. The timescale λ is one of the hyper-parameters in our GP model that determines how smooth the resulting process varies in time. The SE kernel is most widely used as it gives enough freedom to model a wide range of timescales, and it has an added benefit of being infinitely smooth, which will be useful later.

For the purpose of this study, we will also be using a periodic kernel that assumes a sinusoidal process, which was originally defined by MacKay (1997) as

$$k_{\mathrm{per}} = \exp\left(-\frac{2\sin^2\left(\frac{\pi}{T}|x-x'|\right)}{\lambda^2}\frac{2\sin^2\left(\pi|x-x'|T^{-1}\right)}{\lambda^2}\right) \tag{17}$$

where λ is a timescale, and T is the period of oscillation.

The timescale ~~l~~ λ and the period T are the hyper-parameters, which control the shape of the posterior distribution. These hyper-parameters are initially unknown, although the prior distribution can be specified based on domain knowledge, initial observations, or some prior assumptions about the data. In order to determine a better estimate of the hyper-parameters, we need to perform inference based on the *marginal likelihood*, which is defined as the integral of the likelihood and the prior, or

$$p(\mathbf{y}|\mathbf{x}) = \int p(\mathbf{y}|\mathbf{f},\mathbf{x})p(\mathbf{f}|\mathbf{x})d\mathbf{f} \tag{18}$$

where the term *marginal* refers to the process of taking an integral over ~~$\mathbf{f}$~~ $\mathbf{y}$, or *marginalizing* over the ~~function values $\mathbf{f}$~~ observations, in order to obtain ~~$p(\mathbf{y}|X)$~~ $p(\mathbf{y}|\mathbf{x})$.

Note that under the Gaussian Process model, both the prior and the likelihood must also be Gaussian ~~(MacKay, 1997; Rasmussen and Wil~~ (MacKay, 1997; Rasmussen and Williams, 2006). The integral above reduces to

$$\log p(\mathbf{y}|\mathbf{x}) = -\frac{1}{2}\mathbf{y}^\top(\mathbf{K}+\sigma^2\mathbf{I})^{-1}\mathbf{y} - \frac{1}{2}\log|\mathbf{K}+\sigma^2\mathbf{I}| - \frac{n}{2}\log 2\pi \tag{19}$$

which defines the marginal log likelihood (MLL). A more detailed derivation can be found in Chapter 2 from Rasmussen and Williams (2006). By maximizing the marginal log likelihood, we can find the hyper-parameters that maximize MLL, which can best explain the observed data.

Based on the time-series of cloud size distribution in Figure 4, we can assume that a simple periodic kernel is not enough to model the complexity of the observed time-series. Therefore, we have added an SE kernel to the periodic kernel to model numerical instability and uncertainty ~~(MacKay, 1997; Rasmussen and Williams, 2006)~~ (MacKay, 1997; Rasmussen and Williams, 2006) which we can use to customize our GP model to better reflect the underlying assumptions about the observed dataset. In this case,

$$\hat{k} = k_{\mathrm{SE}} + k_{\mathrm{per}} \tag{20}$$

where the SE kernel has been added to account for the ~~non-uniformity in the oscillation~~non-stationarity. With no prior assumptions about the underlying dynamics, we have also tested different combinations for $\hat{k}$, such as adding two periodic kernels together in order to model two oscillations at different timescales (Dagan et al., 2018), but using a single periodic kernel for periodicity detection has been proven to be sufficient in our case.

The GP model uses a gradient descent algorithm to find the hyper-parameters that optimize the (log) marginal likelihood. However, there is no guarantee that it will reach the point that maximizes the marginal likelihood, which will define the best set of hyper-parameters. It is possible that even with multiple experiments, the GP model might be fitted to local maxima, as the gradient descent cannot evaluate the full posterior distribution (see Section 2.7 for more information). Fortunately, at least for the timescale in the periodic kernel k_{per}, we can define a reasonable range of values that can represent the oscillation in the cloud size distribution. ~~We~~ To demonstrate that the 95-minute period is not a local minimum, we have repeatedly trained our GP model with different initial periods, ranging from 5 minutes to 150 minutes at 5-minute intervals, and all of these tested periods converged to the 95-minute period after training. ~~The~~ We have also tested periods longer than 150 minutes, but the resulting posterior distribution ignores most of the variability in the time-series with $T > 150$ minutes. In most cases, the initial period only affected the number of steps that needed to be taken for the gradient descent algorithm to get to the 95-minute period; the further away from the 95-minute period, the longer it took for the algorithm to optimize the periodic kernel.

Given that, we chose an initial period of 90 minutes, which corresponds to the previously observed period of oscillation (Dagan et al., 2018). ~~For the length-scale l, the only constraint in our GP model was to keep it from getting too small, in order to incorporate some~~ We have set a lower bound of 15 minutes to timescale λ to avoid overfitting and to ensure that most of the variability in the observed time-series ~~into noise, and to avoid over-fitting the observational and numerical uncertainties~~ can be explained by the periodic kernel k_{per}. If the timescale λ is allowed to be small, the GP regression model will simply fit the observed data as close as possible regardless of the periodicity.

The GP models are implemented in GPytorch (Gardner et al., 2018) and the ~~corresponding~~ hyper-parameters are ~~fitted~~ trained using the Adam optimizer (Kingma and Ba, 2014). The resulting GP posterior distribution with the kernel $\hat{k}$ can be seen in Figure 7. The initial application of the GP model serves two purposes. First, by assuming a smooth variation in the slope b, we relegate the uncertainty involved in calculating the slope b from the cloud size distribution $C(a)$ to the noise term. Second, as we assume no prior knowledge about the periodic evolution of b, the hyper-parameters can be used as a preliminary estimate for the following procedure, which will be described in the next section.

2.6 Periodicity Detection

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 ~~inference does a good job in estimating~~ 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 ~~need~~ can to be taken in order to better isolate the oscillatory behaviour of the ~~convective process in such cases~~ cloud field, which still remains noisy and non-stationary.

~~One possible detrending method is to perform the regression on the~~ 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
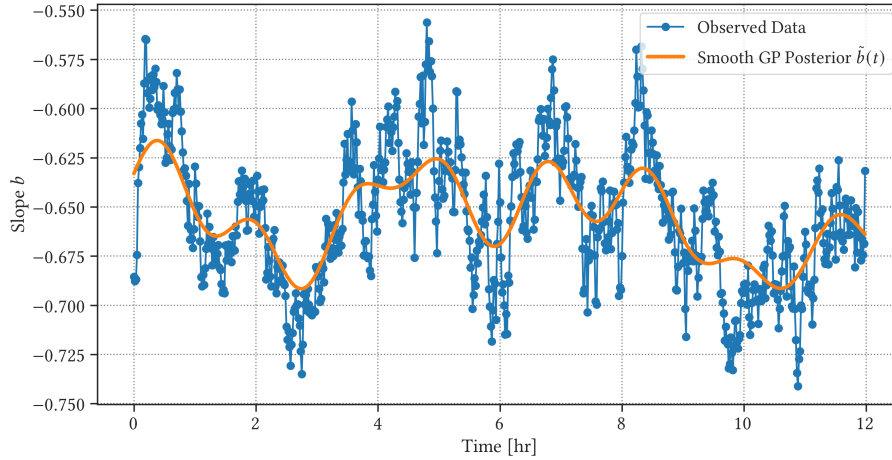
Figure 6. Mean posterior of the trained GP model (orange line) compared to the observed slope b of the cloud size distribution $C(a)$ (blue line) for the first 9 hours of simulation used for training.

the oscillation in $\partial_t \tilde{b}(t)$. ~~The main difference for this round of regression is that we no longer need the~~, 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 ~~needs to be used~~ (i.e. $\hat{k} = k_{\mathrm{per}}$) ~~as we are interested mainly in the~~ 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.

Figure 7 shows the target observation $\partial_t \tilde{b}(t)$ and randomly drawn samples from the posterior distribution. These samples are drawn from the posterior distribution and represent possible realizations of ~~the resulting distribution~~our GP model trained by the observations. The variability in the periodicity ~~, which is the main hyper-parameter of interest,~~ seems to be small relative to the uncertainties in the observed time-series.

The results of applying the periodic GP ~~regression~~ model can be seen in Figure 8, showing the mean posterior distribution for our Gaussian process based on a periodic kernel k_{per} with noise. Most of the variability comes from deviations in $\partial_t \tilde{b}(t)$, and the use of simple periodic kernel is sufficient to isolate the underlying oscillatory behaviour from the observation. The mean posterior distribution in Figure 8 based on the GP model again yields a period of $T \approx 95$ minutes, which is close to the 90-minute period observed by Dagan et al. (2018) in ~~trade~~ tropical marine shallow cumulus clouds under precipitating conditions.
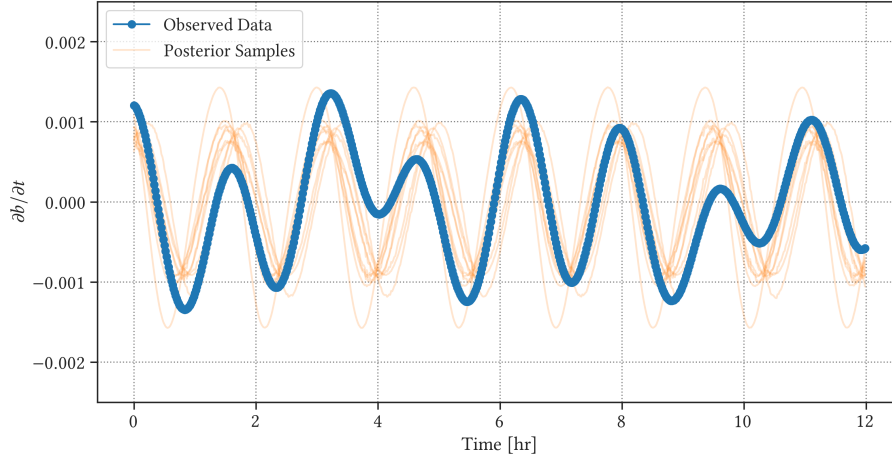
Figure 7. Samples from the posterior distribution of the trained GP model (orange line) compared to the observed time-series of $\partial_t \tilde{b}(t)$ (blue line) for the first 9 hours of simulation used for training.
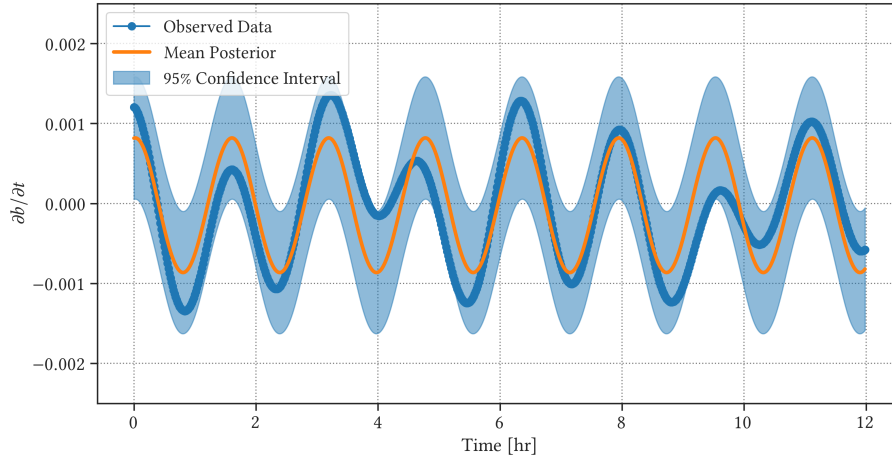
Figure 8. Mean posterior distribution of the ~~trained~~ periodic GP model (orange line) compared to the derivative of time-series $\tilde{b}(t)$ used as the observation (blue line). Shaded regions show (point-wise) 95% confidence interval.

Next, we compare the mean posterior distribution from the GP derivative model to $b(t)$. We integrate the mean posterior distribution $\partial_t \tilde{b}(t)$ in Figure 8 to obtain an estimate for $b(t)$ within a constant. For a direct comparison with $b(t)$ we applied the min-max normalization algorithm

$$\tilde{y} = \frac{y - \min(y)}{\max(y) - \min(y)} \tag{21}$$

415 where $\tilde{y}$ is the normalized time-series, and y is the original ~~function value~~observation. This can be used to perform a quick comparison between the integral of the mean posterior distribution $\tilde{b}(t)$ and the observed time-series $b(t)$, which is shown
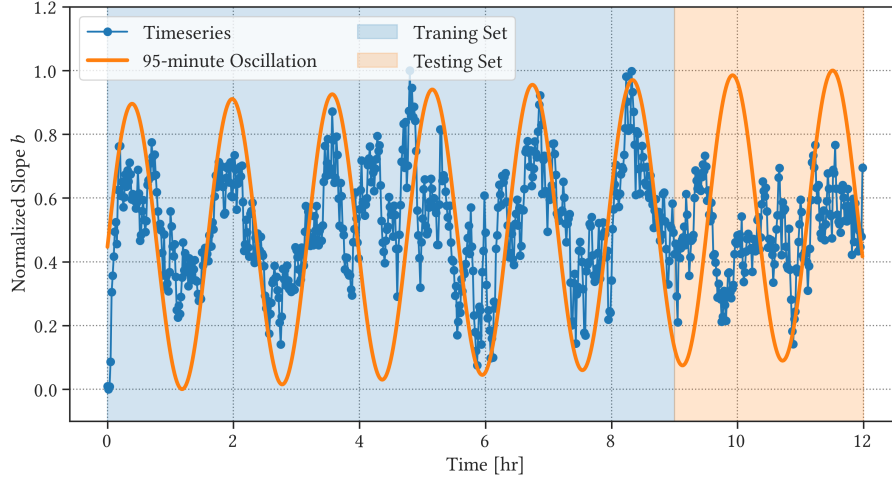
Figure 9. Normalized time-series of slope b of the cloud size distribution $C(a)$ (blue line; same as Figure 4), compared to the ~~the~~ mean posterior from the periodic GP model (orange line), representing 95-minute oscillation. The full time-series includes the 9-hour training set used for GP model fit (blue region) as well as the 3-hour testing set used for GP model evaluation (orange region).

in Figure 9. The ~~simulated~~ normalization, however, makes it more difficult to see that the time-series of the ~~mean posterior distribution $\tilde{b}(t)$ corresponds well to the observed changes in~~ slope $b(t)$ ~~, except at 4 and 10 hours from the beginning of the time-series. Small-scale fluctuations within the cloud domain could disrupt large-scale convective and precipitative behaviour, especially when the model domain is large. As the apparent oscillatory tendencies are dependent on domain size, it is possible that one side of~~ of the ~~domain can be out of phase with the rest of the cloud field, which would make it difficult to identify the oscillations in cloud size distribution~~ cloud size distribution has a negative sign (see Figure 6). A small value in the normalized slope $\tilde{b}$ indicates a more negative slope, or a steeper slope where there is a relative abundance of smaller clouds. On the other hand, a large value in $\tilde{b}$ represents a less negative slope where there is a relative abundance of larger clouds.

2.7 Fully Bayesian Gaussian Process

~~We extend the GP model using Bayesian inference on~~ While the GP regression model yields posterior distributions with confidence bounds, as shown in Figure 9, we cannot obtain a measure of uncertainty for the hyper-parameters ~~. Fully Bayesian GP models have the ability~~ because the gradient descent algorithm cannot evaluate the entire hyper-parameter space. We can instead use a fully Bayesian GP model to estimate the uncertainties in the hyper-parameters of the GP model by stochastically modelling the full posterior distribution. This is done by assuming a prior over model hyper-parameters $\boldsymbol{\theta} \sim p(\boldsymbol{\theta})$, also called the *hyper-prior*; that is, the prior over the hyper-parameters. We can then define joint posterior with the hyper-prior $p(\boldsymbol{\theta})$ as

$$p(\mathbf{f}, \boldsymbol{\theta}|\mathbf{y}) \propto p(\mathbf{y}|\mathbf{f})p(\mathbf{f}|\boldsymbol{\theta}, \mathbf{x})p(\boldsymbol{\theta}) \tag{22}$$
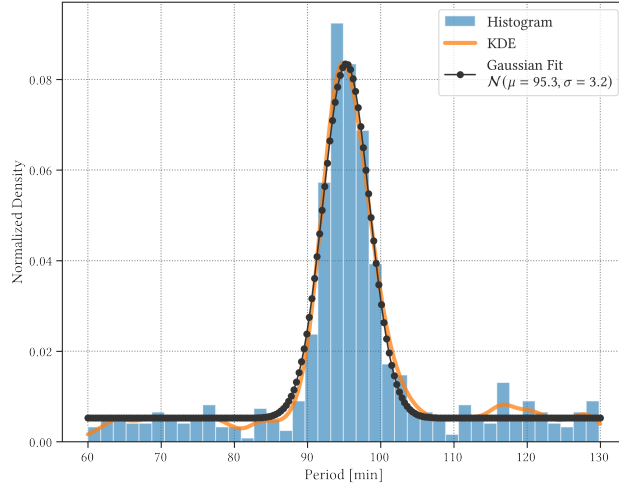
Figure 10. A histogram (blue) and the corresponding kernel density estimate (orange) of periods retrieved from Markov chain Monte Carlo (MCMC) experiments. Fitting the Gaussian distribution to KDE (black dots) infers ~~95 ± 3.6~~ 95 ± 3.2 minutes for the observed periodicity in the time-series.

where we have omitted input data $\mathbf{x}$ and $\mathbf{x}_*$ for the sake of simplicity. For a full, illustrative description of Bayesian model selection, refer to Chapter 5.2 in Rasmussen and Williams (2006).

435 Given the test input data $\mathbf{x}_*$, we retrieve the predictive posterior by integrating the joint posterior

$$p(\mathbf{f}_*|\mathbf{y}) = \iint p(\mathbf{f}_*|\mathbf{f},\boldsymbol{\theta})p(\mathbf{f},\boldsymbol{\theta}|\mathbf{y})\,\mathrm{d}\mathbf{f}\,\mathrm{d}\boldsymbol{\theta} \tag{23}$$

$$= \iint p(\mathbf{f}_*|\mathbf{f},\boldsymbol{\theta})p(\mathbf{f}|\boldsymbol{\theta},\mathbf{y})p(\boldsymbol{\theta}|\mathbf{y})\,\mathrm{d}\mathbf{f}\,\mathrm{d}\boldsymbol{\theta} \tag{24}$$

whose inner integral reduces to the standard GP posterior, which has the same structure as Equation 14. Using the same formalization in Section 2.5, the outer integral can be estimated as

440 $$p(\mathbf{f}_*|\mathbf{y}) = \int p(\mathbf{f}_*|\mathbf{y},\boldsymbol{\theta})p(\boldsymbol{\theta}|\mathbf{y})\,\mathrm{d}\boldsymbol{\theta} \tag{25}$$

$$\simeq \frac{1}{N}\sum_{k=1}^{N} p(\mathbf{f}_*|\mathbf{y},\boldsymbol{\theta}_k), \tag{26}$$

where $\boldsymbol{\theta}_k \sim p(\boldsymbol{\theta}|\mathbf{y})$. The integral $p(\boldsymbol{\theta}|\mathbf{y}) \sim p(\boldsymbol{\theta}|\mathbf{y})p(\boldsymbol{\theta})$ remains intractable. In order to estimate this integral and obtain the predictive posterior as described in Equation 25, we use a modern variant of Hamiltonian Monte Carlo (HMC) algorithm called *No-U-Turn-Sampler* (NUTS; Hoffman et al., 2014) in Pyro (Bingham et al., 2019).

445 The fully Bayesian GP model was applied to the time-series of $\partial_t \tilde{b}(t)$. We initially employed normal priors for characteristic ~~length-scale l~~ timescale λ and periodicity T for the periodic kernel, but the GP model quickly converged to the previously observed 95-minute period. To avoid over-fitting, we initialized the Bayesian GP model with an uninformative, uniform prior that allows any quasi-realistic values for the period of oscillation. Unfortunately, the characteristic ~~length-scale l~~ timescale λ

fails to converge with the uniform prior in the presence of periodicity, which ~~suggests~~ confirms that the Bayesian GP model

450 can explain the observed time-series solely with a periodic kernel k_{per}, and that the post-processing steps taken in Section 2.6 is useful for the Bayesian inference. The characteristic periodicity T was given a uniform prior $\mathcal{U}(30, 130)$, and Bayesian inference via MCMC has been performed with four chains generating 700 samples, including 300 warm-up samples.

The resulting histogram of the characteristic periods taken from 700 samples can be seen in Figure 10, which shows that the most probable period that can be inferred from the time-series $\partial_t \tilde{b}(t)$ is ~~95 ± 3.6~~ 95 ± 3.2 minutes. It is not surprising that the

455 fully Bayesian GP model converged to the same period from the previous section, and it confirms that non-Bayesian GP model is a reliable framework that can identify oscillatory motions in noisy time-series observations of moist convection.

3 Discussion

3.1 Cloud Size Distribution

~~The estimated period of the~~ We have estimated the periodicity in the evolution of the slope of the cloud size distribution; the

460 estimated periodicity in $\tilde{b}(t)$ ~~time-series oscillation is $T = 95 \pm 3.6$~~ is $T = 95 \pm 3.2$ minutes, which corresponds well to previous studies involving marine ~~boundary-layer clouds (Dagan et al., 2018; Koren and Feingold, 2013). though it~~ shallow cumulus clouds (Dagan et al., 2018). This periodicity is slightly longer than ~~the 80-minute period based on~~ 80 minutes estimated by a Fourier spectral analysis of ~~an~~ LES cloud field (Feingold et al., 2017). This could be due to the ~~high-frequency noise in the observed time-series, which~~ differences in precipitation strength, aerosol concentration, or domain size. In the bowling-alley

465 domain, clouds are allowed to develop along the direction of the mean wind, which corresponds to the elongated axis of 42 km. Given that, we expect the development of the cloud field to resemble the larger model run by Dagan et al. (2018). Also, the lower-frequency mode of oscillation found in these studies makes it difficult to ~~detect periodic behaviours at higher frequencies using the Fourier spectral analysis. Each coefficient in the periodogram corresponds to period $T(k) = N/k$ for $k = 0, 1, \ldots, N - 1$. The longer the period, the coarser the resolution, which greatly reduces the effectiveness of the~~

470 ~~periodogram in isolating oscillatory components from the noisy time-series.~~ accurately determine the periodicity, as described in Section 2.4.

The simulated time-series of the mean posterior distribution for $\tilde{b}(t)$ generally corresponds well to the observed changes in $b(t)$, except at 4 and 10 hours from the beginning of the time-series. Small-scale fluctuations within the cloud domain could disrupt large-scale convective and precipitative patterns, even with a large model domain. To gain more insight about the

475 internal variations within the cloud field, we have visualized the three-dimensional cloud field for the duration of the simulation. Figure 11 shows a snapshot of the modelled cloud field taken at 250 minutes into the simulation.

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 field during this time, large structures form at the top of

480 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
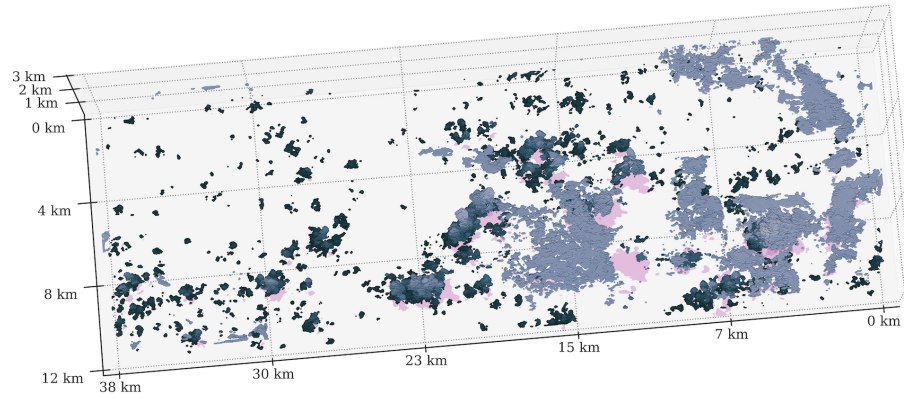
Figure 11. Same as Figure 1, but seen from the top of the atmosphere at 250 minutes into the simulation.
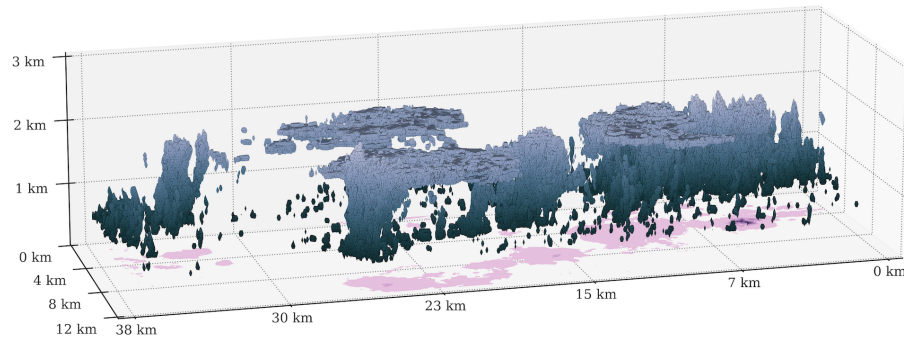
Figure 12. Same as Figure 1, taken at 600 minutes into the simulation.

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 persists and 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.

As shown in Figure 12, clouds form in groups and merge into large convective columns. The successive formation of clouds can also be seen, where clouds appear in a line following the advection of the large clouds (Moser and Lasher-Trapp, 2017). This can also be attributed to the effect of spatial organization, where the cloud field in Figure 12 can be divided into regions

of strong convection, followed by regions devoid of clouds. The development of large clouds can be seen, as expected by our GP model, but these clouds could have been merged into small regions due to spatial organization of the cloud field. Features of organized shallow convection have been observed in large-eddy simulations with small to moderate domain sizes (Seifert and Heus, 2013; Xue et al., 2008). However, such features are sporadic and the clouds become scattered again at $t = 11$ hours into the simulation.

We were not able to reliably isolate a high-frequency oscillation reported in previous studies, either for $T \approx 10$ minutes (Dagan et al., 2018) or for $T \approx 15$ minutes (Feingold et al., 2017) ~~used as a prior for our modified periodic kernel k_{per}~~with a modified prior distribution. This is likely due to the large noise in the time-series $b(t)$ (*cf.* Figure 4), as well as small variations in the amplitude of the oscillation. Over-fitting of noisy time-series becomes an issue in such cases as the GP model quickly adheres to random noise, or observational uncertainties, but not necessarily to high-frequency oscillation.

Given that the primary mode of oscillation found in this study has a period of 95 minutes, we suspect that the size of the domain used for the LES run (43.2 km $\times$ 12.8 km) might have made it more difficult to ~~detect an oscillatory~~ isolate the evolution of individual clouds (Dagan et al., 2018), as we expect the 15-minute period to be correlated to the convective timescale for individual clouds, rather than changes in the mean cloud field ~~(Feingold et al., 2017; Heus et al., 2009)~~ (Feingold et al., 2017; Heus et al., 2009). It is also possible that such convective oscillation exists, but due to the nature of non-linear oscillation at small scales, our attempt in resolving the time-series with a periodic kernel was not able to adequately capture the complex dynamics of non-linear oscillators ~~(Koren and Feingold, 2011; Seifert and Heus, 2013; Koren et al., 2017)~~ (Koren and Feingold, 2011; Koren et al., 2017; Seifert and Heus, 2013) at the scale of individual clouds.

~~The observed 95-minute periodicity in the cloud size distribution matches the observed oscillatory behaviour of a modelled cloud field (Dagan et al., 2018; Seifert et al., 2015) as well as satellite observations (Koren and Feingold, 2013) under precipitating conditions. Once convective clouds develop, precipitation and evaporative cooling due to rain generates cold pools, inhibiting immediate follow-up growth of cumulus clouds. Since the air from these cold pools are negatively buoyant, the downdrafts from the surrounding cold pools converge to form updrafts, promoting convective cloud formation (Dagan et al., 2018; Feingold et al., 2010).~~

Lastly, using fully Bayesian GP model to estimate the uncertainty in the characteristic periodicity T (Figure 10) confirms that without any prior knowledge, except with an assumption that a reasonable period value would lie somewhere between 30 and 130 minutes, we could reliably retrieve a time-series oscillation that is statistically significant, with a period of ~~$T = 95 \pm 3.6$~~ $T = 95 \pm 3.2$ minutes.

3.2 ~~Average~~ Cloud ~~Vertical Mass Flux~~fraction

As the oscillation is assumed to be driven mainly by spatial and temporal organization of precipitating clouds, we ~~assume~~ make an assumption that the same periodic evolution in the dynamic and thermodynamic properties of the mean cloud field, such as cloud cover f_c (Feingold et al., 2017) and vertical mass flux M, can also be observed. Given that, we ~~repeat~~ applied the GP regression method for the cloud fraction over the model domain f_c, which is simply the fraction of the domain that is covered by clouds. Cloud fraction is calculated by projecting the 3-dimensional clouds onto the ground, and calculating the
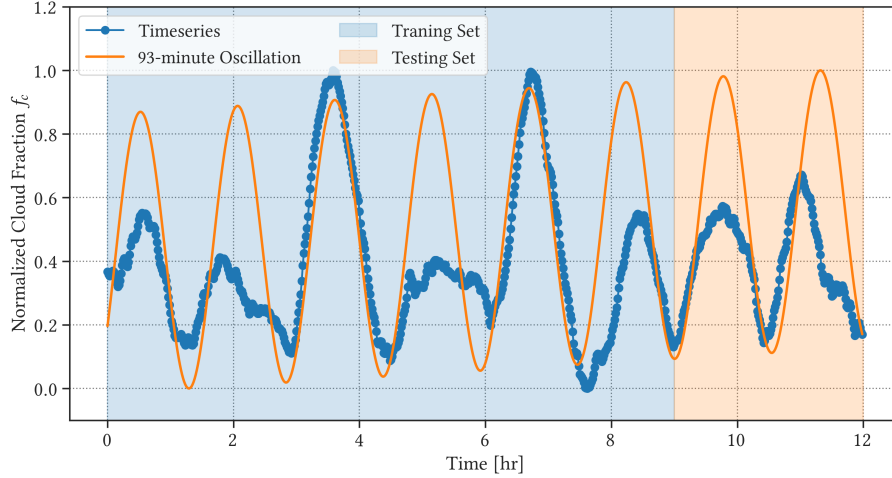
Figure 13. Same as Figure 9, but for cloud fraction f_c.

fraction of the 2-dimensional grid cells that are covered by cloudy region (with condensed liquid water, or $q_l > 0$). We repeat the calculation every minute to construct the cloud fraction time-series $f_c(t)$.

We applied the same numerical techniques presented in Section 2, including the uncertainty estimate for ~~the~~ ~~experiment for~~ ~~these two dynamic properties. Here,~~ the hyper-parameters using the fully Bayesian GP model (Section 2.7). The results of applying the GP regression method can be seen in Figure 13, where the period of oscillation estimated by the fully Bayesian GP model is shown to be $T_f = 93 \pm 3.7$ minutes.

The period lies well within the range of estimated periods for $\tilde{b}(t)$, even when the modelled oscillatory evolution $\tilde{b}(t)$ deviates from the observed values of $b(t)$. This is because merging and splitting of clouds do not significantly change the area that is covered by those clouds. Likewise, persistent clouds forming at the top of the cloud layer tend to form over shallower clouds near the bottom of the cloud layer, and they do not seem to affect the oscillatory evolution in f_c.

The proposed GP regression model can accurately estimate the periodic evolution of cloud fraction, but cloud fraction f_c cannot account for internal variations in the cloud size distribution. The estimated mean posterior distributions of b and f_c are well correlated, but the deviations noted in Section 3.1 are hidden in the time-series of f_c. Large values of f_c are normally associated with a relative abundance of large clouds, but the actual state of the cloud size distribution could differ; for example, the cloud field can be dominated by small clouds near the cloud base but covered by a thin layer of clouds at the top (Figure 11), in which case the radiative properties of the cloud field will be different from one that is dominated by large clouds, despite both having large cloud fraction.

3.3 Average Cloud Vertical Mass Flux
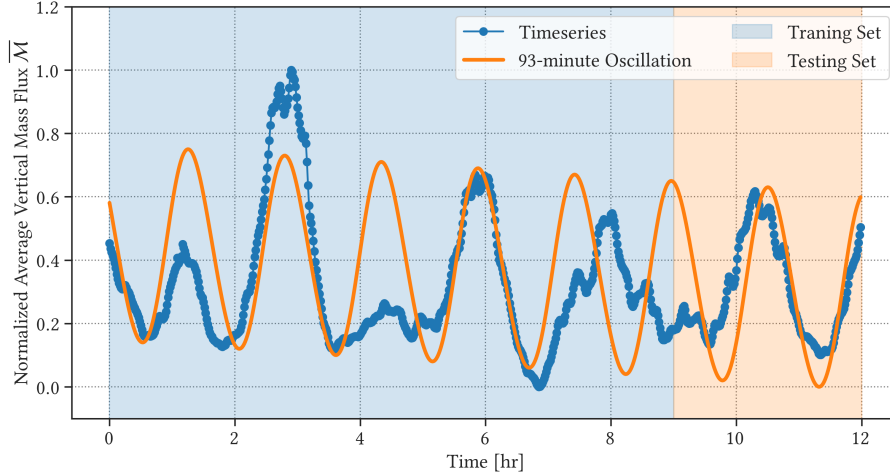
Figure 14. Same as Figure 9, but for average vertical mass flux $\overline{M}(t)$.

We have also applied the GP regression method to vertical mass flux M, which is defined at each grid cell as

$$M = \rho w \mathcal{A} \tag{27}$$

where ρ is air density [$\mathrm{kg\,m^3}$], w is vertical velocity of the air [$\mathrm{m\,s^{-1}}$], and $\mathcal{A}$ is the activity field which is 1 for the cloudy cell , (containing condensed liquid water, or $q_l > 0$), and 0 otherwise.

The average mass flux $\overline{M}(t)$ can then be obtained by calculating average mass flux across the cloud field and taking the average value over the vertical column as the vertical distribution of mass flux remains relatively consistent within the cloud layer. ~~This value will be used as a measure of strength of convection~~ We use this value as a proxy for convective vigour in the modelled cloud field. We calculate $\overline{M}$ at each time step to construct ~~mass-flux~~ the time-series of average mass flux $\overline{M}(t)$ for the simulated 12-hour period (Figure ~~13~~14).

~~We applied the same numerical techniques presented in Section 2, including the uncertainty estimate using the fully Bayesian GP model (Section 2.7). For the~~ For vertical mass-flux over the mean cloud field, the period is estimated to be $T_M = 93 \pm 2.5$ minutes, which ~~is well within the estimated range for the cloud size distribution. Figure 13~~ agrees well with the estimated ranges for both $\tilde{b}$ and f_c. Figure 14 shows both the normalized mass flux time-series $\overline{M}(t)$ as well as the integrated mean posterior of the periodic GP model. The time-series of $\overline{M}(t)$ tends to be much smoother than that of the cloud size distribution because the fluxes ~~are~~ have been averaged over the model domain.

The mean posterior ~~in Figure 13~~ distribution shown in Figure 14 follows the observed time-series quite closely for the training set, except ~~between 7 and 9 hours of~~ around $t = 4$ hours and $t = 9$ hours into the simulation. The ~~deviation occurs slightly earlier than the sudden decrease in $\tilde{b}(t)$~~ suppression of vertical mass flux around $t = 4$ hours is connected to the deviation in $\tilde{b}$ (Figure 9)~~, which explains the interaction between vertical mass flux and cloud size distribution. For example, weakening of~~ . Small clouds form across the model domain during this period, which corresponds to large values of average

565 vertical mass flux ~~due to the formation of cold pools would lead to a relative abundance of smaller clouds once larger clouds dissipate without the convective forcing to support them.~~

~~The resulting time-series $\overline{M}(t)$ corresponds well to that of $\tilde{b}(t)$, except that the slope of the cloud size distribution *decreases* with increasing mass flux. As the steeper slope in the cloud size distribution indicates that there is a relative abundance of smaller clouds , this seems to suggest that mass flux is the main indicator of the formation of smaller clouds . .~~

570 ## 3.4 ~~Cloud fraction~~

~~We repeated the experiment for the cloud fraction over the model domain f_c, which is simply the fraction of the domain that is covered by clouds. Cloud fraction is calculated by projecting the 3-dimensional clouds onto the ground, and calculating the fraction of~~ , but the a thin layer of clouds persists near the top and reduces overall values of $\overline{M}(t)$. On the other hand, the ~~2-dimensional grid cells that are covered by cloudy region (with condensed liquid water, or $q_t > 0$). We repeat the calculation~~

575 ~~every minute to construct the cloud fraction time-series $f_c(t)$.~~

~~Same as Figure 9, but for cloud fraction f_c.~~

~~The results of applying the GP models can be seen in Figure 14, where the period of oscillation estimated by the fully Bayesian GP model is shown to be $T_f = 93 \pm 3.7$ minutes. Much like the time-series of vertical mass flux~~ oscillatory behaviour in $\overline{M}(t)$ ~~, the period lies well within the range of estimated periods for both $\tilde{b}(t)$ and $\overline{M}(t)$, and the oscillatory behaviour for~~

580 ~~cloud fraction $f_c(t)$ matches that of $\tilde{b}$ as well as $\overline{M}(t)$, although the oscillation in average~~ is disrupted between $t = 7$ and $t = 10$ hours of simulation. As noted in Section 3.1, spatial organization of clouds could influence the deviation; during this time, clouds tend to form close to existing convective columns, which can be seen as regions of strong convective activities followed by regions devoid of clouds (*cf.* Figure 12). The formation of cold pools from convective precipitation will result in a convergence of descending moist air, which can disrupt the oscillatory evolution of the cloud field and force the formation of

585 large, organized groups of clouds.

In both cases, it would be useful to diagnose the changes in the vertical distribution of mass flux M to examine the factors contributing to internal fluctuations and gain more insight into how vertical distributions of moisture and momentum periodically re-arrange over time. This also allows us to examine how the clouds mix in terms of the entrainment and detrainment rates, where

590
$$\frac{\partial}{\partial z} M = e - d \tag{28}$$

for rates of entrainment e and detrainment d [$\text{kg m}^{-2}\,\text{s}^{-1}$], but a preliminary examination of domain-average vertical mass flux $\overline{M}$ ~~appears to be leading the oscillation in cloud fraction f_c by 45 to 50 minutes. This could explain the source of the oscillatory behaviour in mean cloud field properties, which will be discussed in the following section~~ profiles reveals that the variability within the cloud field is too large and more work needs to be done, which is outside the scope of this study.
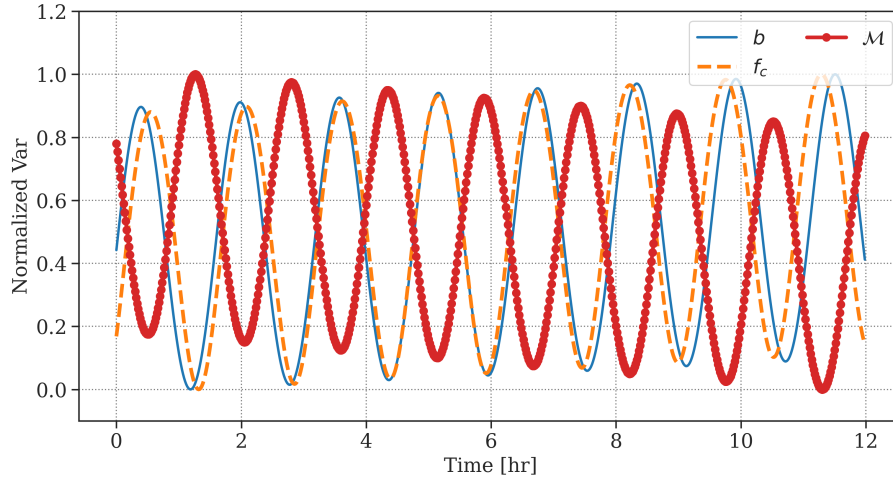
Figure 15. Normalized time-series of slope $\tilde{b}$ of the cloud size distribution $C(a)$ (blue), average vertical mass flux $\overline{M}(t)$ (red) cloud fraction f_c (orange), based on the periodic GP model.

3.4 Precipitation Flux

As shown in Figures 13 and 14, oscillations in $\tilde{b}$ and ~~$\overline{M}$~~ f_c are well-correlated, while the oscillation in ~~f_c~~ $\overline{M}$ seems to be lagging ~~by half a period, which can be seen from the fact that the peaks in f_c~~ That is, the peaks in $\overline{M}$ correspond roughly to the troughs in both $\tilde{b}$ and ~~$\overline{M}$. That is~~ f_c. To visualize the relationship, we plotted normalized mean posterior distributions from our GP model for the three variables in Figure 15.

Figure 15 shows the oscillatory evolutions of the mean posterior distributions for $\tilde{b}$, ~~a peak~~ f_c and $\overline{M}$, which gives more insight about how the marine boundary-layer cumulus clouds evolve over time in a high-resolution model. When there is a relative abundance of larger clouds, the normalized slope $\tilde{b}$ of the cloud size distribution and cloud fraction f_c become larger, which corresponds to a less negative (less steep) slope b of the cloud size distribution. Hence, the changes in cloud fraction f_c ~~, for example, occurs between 45 and 50 minutes following a peak in~~ is correlated to the number of large clouds; that is, the number of large clouds (mostly observed as anvil-like structures near the cloud layer top) determines how much of the model domain is covered by clouds.

On the other hand, peaks in mass flux $\overline{M}$, averaged across the model domain, is associated with smaller values of $\tilde{b}$. This corresponds to a more negative (steeper) slope b of the cloud size distribution, which suggests a relative abundance of smaller clouds, associated with smaller values of cloud fraction f_c. The formation of smaller clouds is connected to an average increase in vertical mass flux across the domain, while the formation of larger clouds is associated with weak overall mass flux. This is indicative of the recharge-discharge mechanism, where strong convective activities are associated with precipitation and the formation of cold pools, which suppresses cloud formation. As the larger clouds dissipate and surface fluxes introduce atmospheric instability, average vertical mass flux across the model domain increases and small clouds develop, which then
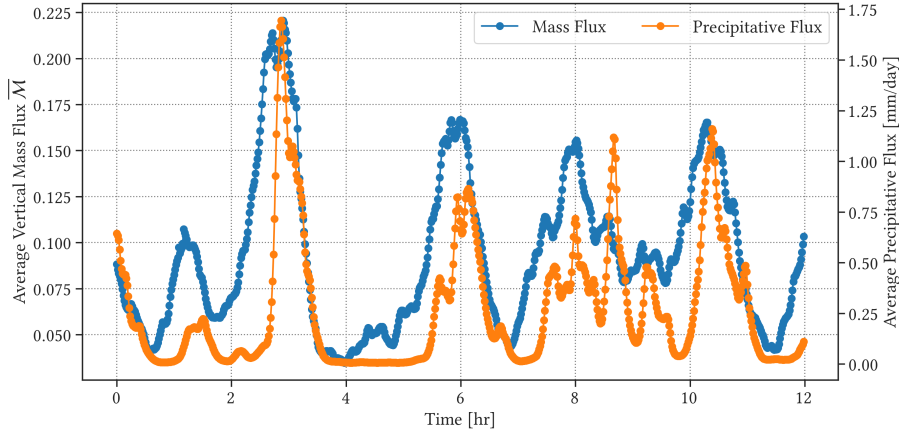
Figure 16. Normalized time-series of average vertical mass flux $\overline{M}(t)$ [kg m^{-2} s^{-1}] and precipitation flux [mm per day].

promotes the formation of larger clouds (Dagan et al., 2018). The increase in average vertical mass flux can also be attributed to the dynamic lifting at the boundaries of cold pools, where descending moist air from cold pools merge to promote the growth of large clouds.

~~This makes aggregation-separation assumption (Feingold et al., 2017) less likely as the underlying mechanism of the oscillatory behaviour shown above as the peak in mass flux does not match the peak in total cloud fraction. Precipitation and changes to the atmospheric instability during convection (Dagan et al., 2018), on the other hand, can better explain the source of the oscillatory behaviour.~~

~~Figure 15 shows the~~ To further examine the relationship between the strength of convection, inferred from average mass flux $\overline{M}$, and precipitation, we plotted the time-series of the average vertical mass flux $\overline{M}$ compared to the precipitation flux, averaged over the cloud layer in Figure 16. It is clear that ~~the peaks in $\overline{M}$ correspond well to those in precipitation flux, suggesting that strong convective formations induce precipitation. During this time,~~ periods of strong vertical mass flux are followed by vigorous precipitation, where latent heat gets released, warms the cloud layer, and ~~leads to vigorous formation of smaller clouds. The~~ promotes the formation of large clouds. As large clouds form and begin to precipitate, the sub-cloud layer cools due to evaporation, and warming of the cloud layer and cooling of the sub-cloud layer reduce ~~the~~ (discharge) atmospheric instability within the boundary layer, rapidly weakening convection and precipitation ~~. Comparing Figures 14 and 15 reveals that the peaks in f_c closely correspond to the troughs following the peaks in the two fluxes . Once a peak in cloud fraction is reached, the convective cycle starts anew as surface fluxes de-stabilize the boundary-layer (Dagan et al., 2018).~~ until surfaces fluxes re-introduce atmospheric instability.

~~The periodic behaviour between~~ Hence, the time-series of mass flux $\overline{M}$ and precipitation ~~is~~ in Figure 16 are consistent with the proposed recharge-discharge cycle ~~of atmospheric instability proposed by Dagan et al. (2018)~~ (Dagan et al., 2018), which is primarily driven by precipitation and latent heat release due to convection. That is, each recharge-discharge cycle consists of a 95-minute oscillation where atmospheric instability is *charged* by surface fluxes and *discharged* by convection/precipitation

over the 95-minute period. This period, however, can be disrupted by internal fluctuations and spatial organization of the cloud field (*cf.* Section 3.1). Furthermore, changes in aerosol concentrations will affect precipitation efficiency, where rain formation in polluted environments tends to be less efficient and therefore slower (Seigel, 2014b; Yamaguchi et al., 2019).

~~We have constructed a time-series for cloud size distribution $C(a)$ that reflects the changes in the relative abundance of large clouds in the cloud field, and used the process (GP)~~ To isolate the role of precipitation in driving the proposed recharge-discharge cycle, we performed a follow-up simulation of the CGILS case where rain formation in the cloud microphysics scheme is artificially suppressed. Likewise, we have tested the GP regression method to ~~model fluctuations within the~~ the mass flux time-series ~~. The~~ in shallow convection during BOMEX (Holland and Rasmusson, 1973) case with precipitation turned off as well. In both cases, the GP regression method described in Section 2 failed to converge towards a single periodicity and no prominent oscillatory behaviour could be found.

4   Conclusions

We performed a high-resolution, large-eddy simulation of tropical marine boundary-layer clouds and implemented numerical methods to analyze the temporal changes in the slope b of the cloud size distribution $C(a)$ ~~has been used as a proxy to examine~~ that reflects the state of ~~the boundary-layer~~ moist convection across the modelled cloud field. ~~LES-modelled~~ LES modelled boundary-layer clouds were conditionally sampled, and the probability distribution of cloud sizes was defined by the Kernel Density Estimator (KDE; Parzen, 1962a). The slope b of the cloud size distribution at each ~~timestep~~ time step was then obtained by a decision tree algorithm and a robust linear regression method. ~~The resulting~~ By applying these numerical methods for every model output, we constructed the time-series ~~contains internal fluctuations within the convective cloud field as well as numerical uncertainties introduced in constructing the~~ of the changes in the slope $b(t)$ for the cloud size distribution, ~~but is an accurate representation the distribution of cloud sizes in~~ which can be used as a proxy of the state of the modelled cloud field.

The numerical steps described in this study are taken primarily to reduce noise and numerical uncertainties involved in calculating the slope b. We have also ~~created the time-series based on the size distribution~~ tested the cloud size distributions of vertically projected ~~clouds (Neggers et al., 2003) using an outlier detection algorithm to reproduce the results in Neggers et al. (2003) and Feingold et al. (2017). This, however,~~ three-dimensional cloud field (Neggers et al., 2003; Feingold et al., 2017) using the histogram, but this resulted in a very noisy time-series and an unsuccessful attempt at estimating the underlying oscillation (*e.g.*. Figure 9 in Feingold et al., 2017). We have also examined other criteria from the literature ~~(e.g. Neggers et al., 2003; Feingold et al., 2017; D~~, but obtaining a stable time-series for the slope b remains difficult ~~, likely because of the large domain size compared to these studies~~.

Under the assumption that the noisy time-series $b(t)$ consists of a single periodic oscillation with large observational variability (Equation 20), we retrieve the underlying trend from the observation using a GP regression model, which is then used to estimate the period of oscillation by a GP model with a periodic kernel k_{per}. The kernel defines our ~~(a priori)~~ ~~belief about what~~ belief of the underlying behaviour ~~should look like~~, and the GP model optimizes the hyper-parameters, especially the period T

of the oscillation, that maximizes the likelihood of the posterior distribution against the observations. We further calculate the uncertainty of the estimated periodicity using a fully Bayesian GP inference. The smooth gradient $\partial_t \tilde{b}(t)$ of the slope of the cloud size distribution $C(a)$ is used as a proxy to determine the underlying behaviour of cloud size distribution, whose period is estimated to be $T = 95 \pm 3.2$ minutes.

We have also applied this technique to total cloud cover f_c and average vertical mass flux $\overline{M}$. Using the fully Bayesian GP model, we identified $T_M = 93 \pm 2.5$ minutes as the period of oscillation for average vertical mass flux, and $T_f = 93 \pm 3.7$ minutes for total cloud cover. Dynamic properties of the cloud field, therefore, are found to oscillate along with the cloud size distribution. The estimated periods agree well with the study involving a satellite observation of open cells (Koren and Feingold, 2013). There are other studies where the cloud field properties can be seen to show similar oscillatory behaviour, which can be identified by the GP regression method presented here; for example, the time-series of cloud cover and liquid-water path (LWP) (Figure 2 from Seifert and Heus, 2013) as well as mass flux (Figure 4 from Plant and Craig, 2008) also appear to oscillate in time (Figure 4).

The oscillation in the time-series of the precipitation flux (Figure 16) is also consistent with the 95-minute period found for the cloud size distribution and the mass flux, suggesting that the oscillation is primarily driven by the recharge-discharge cycle in atmospheric instability (Dagan et al., 2018). The atmospheric instability is weakened (discharged) when clouds grow large and begin to precipitate, as the upper boundary layer warms due to latent heat release and the lower layer cools due to evaporation. This stabilization of the boundary layer continues until cloud growth slows down and precipitation stops. After that, surface fluxes destabilize the boundary-layer from below and convective phase starts again due to atmospheric instability. This cycle takes roughly 95 minutes in the LES model run based on the CGILS S6 case. It should also be emphasized that the same periodic behaviour can be observed using the two-moment microphysics scheme, whereas the bin microphysics scheme has been used by Dagan et al. (2018). The thermodynamic processes that govern the recharge-discharge cycle seem to work consistently in multiple simulations of the boundary-layer atmosphere.

The oscillatory behaviour of precipitating marine boundary-layer clouds has been noted in the literature (Dagan et al., 2018; Feingold et al., 2017; Seigel, 2014b; Yamag . Given that the LES model run here represents sub-tropical shallow cumulus regime, it is not surprising that the 95-minute oscillation found by our GP model is consistent with previous modelling studies of boundary-layer clouds. We have initialized the model run with a pristine atmosphere (where CDNC was

set to 120 cm^{-3}), but studies have shown that changing the aerosol concentration can influence the precipitation efficiency, and therefore the period of oscillation (Dagan et al., 2018; Yamaguchi et al., 2019). As the atmosphere becomes more polluted, we expect this periodicity to increase. 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.

~~Moreover, one can focus on the capabilities of Gaussian process regression in handling noisydata, which is especially usefulfor observational data. The GP modelling allows the model to assume that the noise is in both the time step x and the observed data y, and that the sampled data do not lie on a discrete grid (Roberts et al., 2013).~~

Lastly, ~~it would be interesting to see if the GP modelling can be used to retrieve the high-frequency oscillation found in previous studies of modelled cumulus clouds (Feingold et al., 2017; Dagan et al., 2018; Moser and Lasher-Trapp, 2017). The procedure introduced here seems to be inadequate for this purpose, as the signal to noise ratio remains too low. Base on the literature, this is likely because this oscillation is based on the pulse-like development of individual clouds (Blyth and Latham, 1993; Zhao a , as opposed to an organizational change to the mean cloud field , which lends itself to a slightly different approach. Subsequent studies will focus on using the statistical tools to uncover the underlying dynamics of the cloud field based on the distribution of cloud properties~~ examining the changes in the vertical distribution of mass flux M can give more insight into how the cloud field mixes with the environment over time, especially because we have already examined the temporal changes in f_c, or $\partial\sigma/\partial t$ where $f_c = \sigma$ (Section 2.6). Given the results from the high-resolution large-eddy simulation, the mass continuity of the cloud field can be examined in terms of entrainment and detrainment rates, or

$$\rho\frac{\partial\sigma}{\partial t} + \frac{\partial M}{\partial z} = e - d \tag{29}$$

for vertical mass flux $M = \rho w \sigma$ and rates of entrainment e and detrainment d [kg m^{-2} s^{-1}]. We have observed that both cloud fraction and vertical mass flux oscillate over time, but the latter can also be seen to be lagging by half a period (Section 3.4). Given that, investigating how the vertical distribution of mass flux changes over time can provide valuable insights about the dynamics of shallow convection.

Code and data availability. The data from the LES model run are available upon request. The LES model used in this paper, as well as the exact set of case parameters, are available at https://github.com/lorenghoh/sam_loh. Jupyter notebooks including the numerical analysis are also available at https://github.com/lorenghoh/size_oscillation.

Author contributions. LO designed and ran the LES model run, carried out the numerical analysis and prepared this manuscript. PHA conceptualized this project and reviewed the manuscript.

735 *Competing interests.* The authors declare that they have no conflict of interest.

Acknowledgements. This work was partially supported by Korea Polar Research Institute (KOPRI) grant funded by the Ministry of Oceans and Fisheries (KOPRI PE24010). We would like to thank KOPRI for kindly providing the ~~hardware~~ computational resources required to perform various machine-learning tasks for this paper.

References

740   Angus, R., Morton, T., Aigrain, S., Foreman-Mackey, D., and Rajpaul, V.: Inferring probabilistic stellar rotation periods using Gaussian processes, arXiv, pp. 2094–2108, 2017.

Benner, T. C. and Curry, J. A.: Characteristics of small tropical cumulus clouds and their impact on the environment, J. Geophys. Res., 103, 28 753–28 767, 1998.

Berg, L. K. and Stull, R. B.: Accuracy of Point and Line Measures of Boundary Layer Cloud Amount, J. Appl. Meteor., 41, 640–650, 2002.

745   Bingham, E., Chen, J. P., Jankowiak, M., Obermeyer, F., Pradhan, N., Karaletsos, T., Singh, R., Szerlip, P., Horsfall, P., and Goodman, N. D.: Pyro: Deep universal probabilistic programming, The Journal of Machine Learning Research, 20, 973–978, 2019.

Bladé, I. and Hartmann, D. L.: Tropical intraseasonal oscillations in a simple nonlinear model, Journal of Atmospheric Sciences, 50, 2922–2939, 1993.

Blossey, P. N., Bretherton, C. S., Zhang, M., Cheng, A., Endo, S., Heus, T., Liu, Y., Lock, A. P., de Roode, S. R., and Xu, K.-M.: Marine low
750   cloud sensitivity to an idealized climate change: The CGILS LES intercomparison, J. Adv. Model. Earth Syst., 5, 234–258, 2013.

Blyth, A. M. and Latham, J.: Development of ice and precipitation in New Mexican summertime cumulus clouds, Q. J. R. Meteorol. Soc., 119, 91–120, 1993.

Bony, S. and Dufresne, J.-L.: Marine boundary layer clouds at the heart of tropical cloud feedback uncertainties in climate models, Geophys. Res. Lett., 32, 2055, 2005.

755   Bony, S., Colman, R., Kattsov, V. M., Allan, R. P., Bretherton, C. S., Dufresne, J.-L., Hall, A., Hallegatte, S., Holland, M. M., Ingram, W., Randall, D. A., Soden, B. J., Tselioudis, G., and Webb, M. J.: How Well Do We Understand and Evaluate Climate Change Feedback Processes?, J. Climate, 19, 3445–3482, 2006.

Breiman, L., Friedman, J., Stone, C. J., and Olshen, R. A.: Classification and Regression Trees, Taylor & Francis, 1984.

Bretherton, C. and Blossey, P.: Understanding mesoscale aggregation of shallow cumulus convection using large-eddy simulation, Journal of
760   Advances in Modeling Earth Systems, 9, 2798–2821, 2017.

Brown, A. R.: The sensitivity of large-eddy simulations of shallow cumulus convection to resolution and subgrid model, Q. J. R. Meteorol. Soc., 125, 469–482, 1999.

Brown, A. R., Cederwall, R. T., Chlond, A., Duynkerke, P. G., Golaz, J. C., Khairoutdinov, M., Lewellen, D. C., Lock, A. P., MacVean, M. K., Moeng, C. H., Neggers, R. A. J., Siebesma, A. P., and Stevens, B.: Large-eddy simulation of the diurnal cycle of shallow cumulus
765   convection over land, Q. J. R. Meteorol. Soc., 128, 1075–1093, 2002.

Cahalan, R. F. and Joseph, J. H.: Fractal Statistics of Cloud Fields, Mon. Wea. Rev., 117, 261–272, 1989.

Ceppi, P., Brient, F., Zelinka, M. D., and Hartmann, D. L.: Cloud feedback mechanisms and their representation in global climate models, WIREs Clim. Change, 8, e465, 2017.

Cheng, L.-F., Dumitrascu, B., Darnell, G., Chivers, C., Draugelis, M., Li, K., and Engelhardt, B. E.: Sparse multi-output Gaussian processes
770   for online medical time series prediction, BMC Med. Inform. Decis. Mak., 20, 1351–23, 2020.

Clough, S. A., Shephard, M. W., Mlawer, E. J., Delamere, J. S., Iacono, M. J., Cady-Pereira, K., Boukabara, S., and Brown, P. D.: Atmospheric radiative transfer modeling: a summary of the AER codes, J. Quant. Spectrosc. Radiat. Transfer, 91, 233–244, 2005.

Dagan, G., Koren, I., Altaratz, O., and Heiblum, R. H.: Time-dependent, non-monotonic response of warm convective cloud fields to changes in aerosol loading, Atmospheric Chemistry and Physics, 17, 7435–7444, 2017.

775 Dagan, G., Koren, I., Kostinski, A., and Altaratz, O.: Organization and Oscillations in Simulated Shallow Convective Clouds, J. Adv. Model. Earth Syst., 10, 2287–2299, 2018.

Dawe, J. T. and Austin, P. H.: Interpolation of LES Cloud Surfaces for Use in Direct Calculations of Entrainment and Detrainment, Mon. Wea. Rev., 139, 444–456, 2011.

Dawe, J. T. and Austin, P. H.: Direct entrainment and detrainment rate distributions of individual shallow cumulus clouds in an LES, Atmos.
780   Chem. Phys., 13, 7795–7811, 2013.

Dickey, D. A. and Fuller, W. A.: Distribution of the estimators for autoregressive time series with a unit root, Journal of the American statistical association, 74, 427–431, 1979.

Durrande, N., Hensman, J., Rattray, M., and Lawrence, N. D.: Detecting periodicities with Gaussian processes, PeerJ Comput. Sci., 2, e50, 2016.

785 Feingold, G., Koren, I., Wang, H., Xue, H., and Brewer, W. A.: Precipitation-generated oscillations in open cellular cloud fields, Nature, 466, 849–852, 2010.

Feingold, G., Balsells, J., Glassmeier, F., Yamaguchi, T., Kazil, J., and McComiskey, A.: Analysis of albedo versus cloud fraction relationships in liquid water clouds using heuristic models and large eddy simulation, J. Geophys. Res. Atmos., 122, 7086–7102, 2017.

Fisher, N. I. and Lee, A. J.: A correlation coefficient for circular data, Biometrika, 70, 327–332, 1983.

790 Gardner, J. R., Pleiss, G., Bindel, D., Weinberger, K. Q., and Wilson, A. G.: Gpytorch: Blackbox matrix-matrix gaussian process inference with gpu acceleration, arXiv preprint arXiv:1809.11165, 2018.

Garrett, T. J., Glenn, I. B., and Krueger, S. K.: Thermodynamic Constraints on the Size Distributions of Tropical Clouds, J. Geophys. Res. Atmos., 123, 8832–8849, 2018.

Hamilton, J. D.: Time series analysis, Princeton university press, 2020.

795 Heus, T., Jonker, H. J. J., Van den Akker, H. E. A., Griffith, E. J., Koutek, M., and Post, F. H.: A statistical approach to the life cycle analysis of cumulus clouds selected in a virtual reality environment, J. Geophys. Res., 114, 97, 2009.

Hoffman, M. D., Gelman, A., et al.: The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo., J. Mach. Learn. Res., 15, 1593–1623, 2014.

Holland, J. Z. and Rasmusson, E. M.: Measurements of Atmospheric Mass, Energy, and Momentum Budgets Over a 500-Kilometer Square
800   of Tropical Ocean, Mon. Wea. Rev., 101, 44–55, 1973.

Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., and Collins, W. D.: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models, J. Geophys. Res., 113, 233, 2008.

Khairoutdinov, M. F. and Randall, D. A.: Cloud Resolving Modeling of the ARM Summer 1997 IOP: Model Formulation, Results, Uncertainties, and Sensitivities, J. Atmos. Sci., 60, 607–625, 2003.

805 Kingma, D. P. and Ba, J.: Adam: A method for stochastic optimization, arXiv preprint arXiv:1412.6980, 2014.

Koren, I. and Feingold, G.: Aerosol–cloud–precipitation system as a predator-prey problem, Proceedings of the National Academy of Sciences of the United States of America, 108, 12 227–12 232, 2011.

Koren, I. and Feingold, G.: Adaptive behavior of marine cellular clouds, Sci. Rep., 3, 938, 2013.

Koren, I., Tziperman, E., and Feingold, G.: Exploring the nonlinear cloud and rain equation, Chaos, 27, 013 107, 2017.

810 Kuo, K. S., Welch, R. M., Weger, R. C., Engelstad, M. A., and Sengupta, S. K.: The three-dimensional structure of cumulus clouds over the ocean: 1. Structural analysis, J. Geophys. Res., 98, 20 685, 1993.

Lipat, B. R., Voigt, A., Tselioudis, G., and Polvani, L. M.: Model Uncertainty in Cloud–Circulation Coupling, and Cloud-Radiative Response to Increasing CO 2, Linked to Biases in Climatological Circulation, J. Climate, 31, 10 013–10 020, 2018.

Machado, L. and Rossow, W. B.: Structural characteristics and radiative properties of tropical cloud clusters, Mon. Wea. Rev., 121, 3234–3260, 1993.

Machado, L. T., Desbois, M., and Duvel, J.-P.: Structural characteristics of deep convective systems over tropical Africa and the Atlantic Ocean, Mon. Wea. Rev., 120, 392–406, 1992.

MacKay, D. J. C.: Gaussian Processes - A Replacement for Supervised Neural Networks?, Tech. rep., 1997.

Malkus, J. S. and Riehl, H.: Cloud structure and distributions over the tropical Pacific Ocean 1, Tellus, 16, 275–287, 1964.

Mauritsen, T. and Roeckner, E.: Tuning the MPI-ESM1. 2 global climate model to improve the match with instrumental record warming by lowering its climate sensitivity, Journal of advances in modeling earth systems, 12, e2019MS002 037, 2020.

Morrison, H., Curry, J. A., and Khvorostyanov, V. I.: A New Double-Moment Microphysics Parameterization for Application in Cloud and Climate Models. Part I: Description, J. Atmos. Sci., 62, 1665–1677, 2005a.

Morrison, H., Curry, J. A., Shupe, M. D., and Zuidema, P.: A New Double-Moment Microphysics Parameterization for Application in Cloud and Climate Models. Part II: Single-Column Modeling of Arctic Clouds, J. Atmos. Sci., 62, 1678–1693, 2005b.

Moser, D. H. and Lasher-Trapp, S.: The Influence of Successive Thermals on Entrainment and Dilution in a Simulated Cumulus Congestus, J. Atmos. Sci., 74, 375–392, 2017.

Myers, T. A. and Norris, J. R.: Reducing the uncertainty in subtropical cloud feedback, Geophys. Res. Lett., 43, 2144–2148, 2016.

Nair, U., Weger, R., Kuo, K., and Welch, R.: Clustering, randomness, and regularity in cloud fields: 5. The nature of regular cumulus cloud fields, Journal of Geophysical Research: Atmospheres, 103, 11 363–11 380, 1998.

Neggers, R. A. J., Jonker, H. J. J., and Siebesma, A. P.: Size Statistics of Cumulus Cloud Populations in Large-Eddy Simulations, J. Atmos. Sci., 60, 1060–1074, 2003.

Panofsky, H. A. and Brier, G. W.: Some applications of statistics to meteorology, Mineral Industries Extension Services, College of Mineral Industries . . . , 1958.

Parzen, E.: On Estimation of a Probability Density Function and Mode, Ann. Math. Statist., 33, 1065–1076, 1962a.

Parzen, E.: On estimation of a probability density function and mode, The annals of mathematical statistics, 33, 1065–1076, 1962b.

Peters, O., Neelin, J. D., and Nesbitt, S. W.: Mesoscale convective systems and critical clusters, J. Atmos. Sci., 66, 2913–2924, 2009.

Plank, V. G.: The size distribution of cumulus clouds in representative Florida populations, J. Appl. Meteor., 8, 46–67, 1969.

Plant, R. S. and Craig, G. C.: A Stochastic Parameterization for Deep Convection Based on Equilibrium Statistics, J. Atmos. Sci., 65, 87–105, 2008.

Raga, G. B., Jensen, J. B., and Baker, M. B.: Characteristics of Cumulus Band Clouds off the Coast of Hawaii, J. Atmos. Sci., 47, 338–356, 1990.

Rasmussen, C. E. and Williams, C.: Gaussian Processes for Machine Learning, MIT Press, 2006.

Rasmussen, R. M., Geresdi, I., Thompson, G., Manning, K., and Karplus, E.: Freezing drizzle formation in stably stratified layer clouds: The role of radiative cooling of cloud droplets, cloud condensation nuclei, and ice initiation, Journal of the atmospheric sciences, 59, 837–860, 2002.

Roberts, S., Osborne, M., Ebden, M., Reece, S., Gibson, N., and Aigrain, S.: Gaussian processes for time-series modelling, Phil. Trans. R. Soc. A., 371, 20110 550, 2013.

Rodts, S. M. A., Duynkerke, P. G., and Jonker, H. J. J.: Size Distributions and Dynamical Properties of Shallow Cumulus Clouds from Aircraft Observations and Satellite Data, J. Atmos. Sci., 60, 1895–1912, 2003.

Sato, Y., Shima, S.-i., and Tomita, H.: A grid refinement study of trade wind cumuli simulated by a Lagrangian cloud microphysical model: the super-droplet method, Atmos. Sci. Lett., 18, 359–365, 2017.

Sato, Y., Shima, S.-i., and Tomita, H.: Numerical Convergence of Shallow Convection Cloud Field Simulations: Comparison Between Double-Moment Eulerian and Particle-Based Lagrangian Microphysics Coupled to the Same Dynamical Core, J. Adv. Model. Earth Syst., 10, 1495–1512, 2018.

Schneider, T., Teixeira, J., Bretherton, C. S., Brient, F., Pressel, K. G., Schär, C., and Siebesma, A. P.: Climate goals and computing the future of clouds, Nat. Clim. Change., 7, 3–5, 2017.

Seifert, A. and Heus, T.: Large-eddy simulation of organized precipitating trade wind cumulus clouds, Atmos. Chem. Phys., 13, 5631–5645, 2013.

Seifert, A., Heus, T., Pincus, R., and Stevens, B.: Large-eddy simulation of the transient and near-equilibrium behavior of precipitating shallow convection, J. Adv. Model. Earth Syst., 7, 1918–1937, 2015.

Seigel, R. B.: Shallow Cumulus Mixing and Subcloud-Layer Responses to Variations in Aerosol Loading, J. Atmos. Sci., 71, 2581–2603, 2014a.

Seigel, R. B.: Shallow cumulus mixing and subcloud-layer responses to variations in aerosol loading, Journal of the Atmospheric Sciences, 71, 2581–2603, 2014b.

Sen, P. K.: Estimates of the regression coefficient based on Kendall's tau, J. Am. Stat. Assoc., 63, 1379–1389, 1968.

Stevens, B., Satoh, M., Auger, L., Biercamp, J., Bretherton, C. S., Chen, X., Düben, P., Judt, F., Khairoutdinov, M., Klocke, D., Kodama, C., Kornblueh, L., Lin, S.-J., Neumann, P., Putman, W. M., Röber, N., Shibuya, R., Vanniere, B., Vidale, P. L., Wedi, N., and Zhou, L.: DYAMOND: the DYnamics of the Atmospheric general circulation Modeled On Non-hydrostatic Domains, Prog. Earth Planet. Sci., 6, 151–17, 2019.

Stevens, B., Bony, S., Brogniez, H., Hentgen, L., Hohenegger, C., Kiemle, C., L'Ecuyer, T. S., Naumann, A. K., Schulz, H., Siebesma, P. A., et al.: Sugar, gravel, fish and flowers: Mesoscale cloud patterns in the trade winds, Quarterly Journal of the Royal Meteorological Society, 146, 141–152, 2020.

Tan, Z., Schneider, T., Teixeira, J., and Pressel, K. G.: Large-eddy simulation of subtropical cloud-topped boundary layers: 1. A forcing framework with closed surface energy balance, J. Adv. Model. Earth Syst., 8, 1565–1585, 2016.

Theil, H.: A rank-invariant method of linear and polynomial regression analysis, 3; confidence regions for the parameters of polynomial regression equations, Indag. Math., 1, 467–482, 1950.

Tompkins, A. M.: Organization of tropical convection in low vertical wind shears: The role of water vapor, Journal of the atmospheric sciences, 58, 529–545, 2001.

Twomey, S.: Pollution and the planetary albedo, Atmospheric Environment (1967), 8, 1251–1256, 1974.

Wang, H. and Feingold, G.: Modeling mesoscale cellular structures and drizzle in marine stratocumulus. Part I: Impact of drizzle on the formation and evolution of open cells, Journal of the Atmospheric Sciences, 66, 3237–3256, 2009.

Wilcox, E. M. and Ramanathan, V.: Scale dependence of the thermodynamic forcing of tropical monsoon clouds: Results from TRMM observations, J. Climate, 14, 1511–1524, 2001.

Xue, H., Feingold, G., and Stevens, B.: Aerosol effects on clouds, precipitation, and the organization of shallow cumulus convection, Journal of the Atmospheric Sciences, 65, 392–406, 2008.

Yamaguchi, T. and Feingold, G.: On the relationship between open cellular convective cloud patterns and the spatial distribution of precipitation, Atmospheric Chemistry and Physics, 15, 1237–1251, 2015.

Yamaguchi, T., Feingold, G., and Kazil, J.: Aerosol-cloud interactions in trade wind cumulus clouds and the role of vertical wind shear, Journal of Geophysical Research: Atmospheres, 124, 12 244–12 261, 2019.

Yano, J.-I. and Plant, R.: Finite departure from convective quasi-equilibrium: Periodic cycle and discharge–recharge mechanism, Quarterly Journal of the Royal Meteorological Society, 138, 626–637, 2012.

Zelinka, M. D., Myers, T. A., McCoy, D. T., Po-Chedley, S., Caldwell, P. M., Ceppi, P., Klein, S. A., and Taylor, K. E.: Causes of higher climate sensitivity in CMIP6 models, Geophysical Research Letters, 47, e2019GL085 782, 2020.

Zhang, M., Bretherton, C. S., Blossey, P. N., Austin, P. H., Bacmeister, J. T., Bony, S., Brient, F., Cheedela, S. K., Cheng, A., Del Genio, A. D., et al.: CGILS: Results from the first phase of an international project to understand the physical mechanisms of low cloud feedbacks in single column models, Journal of Advances in Modeling Earth Systems, 5, 826–842, 2013.

Zhao, G. and Di Girolamo, L.: Statistics on the macrophysical properties of trade wind cumuli over the tropical western Atlantic, J. Geophys. Res., 112, 847, 2007.

Zhao, M. and Austin, P. H.: Life Cycle of Numerically Simulated Shallow Cumulus Clouds. Part I: Transport, J. Atmos. Sci., 62, 1269–1290, 2005a.

Zhao, M. and Austin, P. H.: Life Cycle of Numerically Simulated Shallow Cumulus Clouds. Part II: Mixing Dynamics, J. Atmos. Sci., 62, 1291–1310, 2005b.

Zuidema, P., Li, Z., Hill, R. J., Bariteau, L., Rilling, B., Fairall, C., Brewer, W. A., Albrecht, B., and Hare, J.: On trade wind cumulus cold pools, Journal of the Atmospheric Sciences, 69, 258–280, 2012.