https://doi.org/10.5194/egusphere-2026-2871
Preprint. Discussion started: 18 June 2026

Technical note: Evaluation of conceptual predator-prey models for the quantitative modeling of precipitating open-cell stratocumulus via feature-based Bayesian inversion of a suite of Large eddy simulations

Rebecca Gjini[1], Matthias Morzfeld[1], Franziska Glassmeier[2, 3], and Graham Feingold[4]

[1]Cecil H. and Ida M. Green Institute of Geophysics and Planetary Physics, Scripps Institution of Oceanography, University of California, San Diego, CA
[2]Max Planck Institute for Meteorology, Hamburg, Germany
[3]Delft University of Technology, Delft, Netherlands
[4]National Oceanic and Atmospheric Administration, Chemical Sciences Laboratory, Boulder, CO

Correspondence: Rebecca Gjini (rgjini@ucsd.edu)

Abstract. We consider two very different types of models of precipitating open-cell stratocumulus clouds. The first model type is a computationally expensive large eddy simulation (LES), that resolves convection and clouds at high temporal and spatial resolutions. The second model type is the nonlinear cloud and rain (C&R) equation, a scalar delay differential equation (DDE) that interprets interactions of precipitation and clouds phenomenologically by predator (rain) and prey (cloud) dynamics. We evaluate the extent to which one may use the C&R equation as a quantitative tool for representing selected aspects of an LES. Specifically, we estimate parameters of the C&R equation from a suite of LES via feature-based Bayesian inversions and track the evolution of posterior distributions over C&R model parameters under changing meteorological conditions in the LES. Our inversions show that the C&R equation can be calibrated to generate limit cycles that are quantitatively compatible with cycles of cloud growth and decay across a wide spectrum of meteorological conditions. The successful inversions reiterate the robustness of the predator-prey analogy to the dynamics of precipitating open-cell stratocumulus. When we interpret the inversions jointly, however, we observe counterintuitive and partially nonphysical shifts and changes in the posterior distributions over the C&R model parameters. Our evaluation study thus highlights the challenges one faces when mapping LES dynamics to a scalar DDE, which can stem either from structural inadequacies in the DDE model, or from the specific feature-based inversion framework, or a mixture of both.

1 Introduction

One of the largest sources of uncertainties in computations and projections of Earth's energy budget is our incomplete understanding of clouds and aerosols (Boucher et al., 2013; Schneider et al., 2017; Sherwood et al., 2020; Forster et al., 2007; IPCC, 2021). Stratocumulus clouds play a particularly important role here because they cover large stretches of subtropical ocean (Wood, 2012), with cloud fields on the scale of thousands of km and annually averaged cloud cover of 34%. Stratocumulus clouds provide a net cooling effect on the planet (L'Ecuyer et al., 2019; Hartmann and Doelling, 1991), but stratocumulus cloud

https://doi.org/10.5194/egusphere-2026-2871
Preprint. Discussion started: 18 June 2026

cover may change as Earth's climate changes, with commensurate effects on our future climate (Liu et al., 2023). Importantly, stratocumulus clouds occur in two distinct spatial patterns: Open- and closed-cell (Agee, 1984; Wood and Hartmann, 2006; Glassmeier and Feingold, 2017). Closed-cell stratocumulus are characterized by honeycomb-like cloudy structures, bounded by narrow cloud-free regions. Open-cell stratocumulus are the reverse, i.e., narrow cloud walls bound honeycomb-like cloud-free regions. Aerosols can determine whether stratocumulus appear as bright, reflective scenes (i.e., closed-cell) or low cloud fraction scenes revealing the underlying dark ocean (i.e., open-cell) (see, e.g., Stevens et al., 2005; Wang and Feingold, 2009; Yamaguchi et al., 2017).

Mathematical and computational models are, in addition to observations, one of the tools we have to understand stratocumulus clouds. Such models exist at different levels of accuracy, complexity and computational costs. The end-member of the model hierarchy with the highest level of detail is a large eddy simulation (LES). An LES is a 3D, time-evolving simulation that resolves convection, clouds, and the influence of aerosol particles on cloud microphysical and macro-physical structure at high spatial (order of 100 meters) and temporal (order of seconds) resolutions. Running an LES requires solving a set of partial differential equations numerically. In addition, inputs including the boundary conditions and a large number of initial conditions need to be specified and, hence, the meteorological conditions that are being modeled. Due to their computational complexity LES are run on high-performance computing (HPC) systems and simulate spatial domains on the order of 50 km (although stratocumulus cloud decks can be much larger).

The other end-member of the model hierarchy are delay-differential equation (DDE) models, designed to capture emergent behavior of stratocumulus, via predator-prey dynamics, with rain acting as a predator of clouds (Koren and Feingold, 2011; Feingold and Koren, 2013; Koren et al., 2017). The delay in the differential equation model appears naturally as a result of rain formation physics. DDE models are, in comparison to a LES, simple to use, but they are designed to capture only a few important physical processes that govern emergent behavior of the atmosphere. As such, DDE models are an important conceptual tool to understand the main processes driving atmospheric dynamics and cloud microphysics.

In this model evaluation study, we assess the extent to which the C&R equation can quantitatively describe selected aspects of large eddy simulations. Specifically, we continue the work started by Lunderman et al. (2020), where a DDE model is "calibrated" to output variables of an LES, via feature-based Bayesian inversion (see, e.g., Morzfeld et al., 2018). The basic idea of Lunderman et al. (2020) is as follows. We select features of an LES and then search for a DDE model that exhibits similar features as the LES. Since DDE models represent precipitating open-cell stratocumulus by (limit) cycles of cloud growth, followed by decay due to rain, we, naturally, select cycles of cloud growth and decay as the LES feature. Lunderman et al. (2020) showed that feature-based inversion can be used to calibrate a DDE model to exhibit cycles of cloud growth and decay that are similar to the cycles extracted from an LES.

To evaluate and stress-test the calibration strategy, we consider a suite of 14 LES that simulate strongly precipitating open-cell stratocumulus under different boundary and initial conditions (different meteorological conditions); these 14 LES are a carefully selected subset of the set of 127 stratocumulus LES described by Hoffmann et al. (2023). We apply feature-based inversion with two variants of the nonlinear cloud and rain (C&R) equation of Koren et al. (2017) (see Section 2), for a total

https://doi.org/10.5194/egusphere-2026-2871
Preprint. Discussion started: 18 June 2026

of 28 inversions. In the process, we adapt the techniques described by Lunderman et al. (2020) to be more robustly applicable and with fewer tunable inversion parameters.

With the inversions in place, we can track the evolution of the posterior distributions over the C&R equations' parameters under changing meteorological conditions, represented by the suite of LES. Our inversions confirm and extend the results of Lunderman et al. (2020) in that the DDE models can be calibrated to the 14 precipitating LES with internally consistent and physically meaningful dependencies across model parameters. When we consider the 14 inversions (for each model) jointly, however, the posterior distributions over DDE model parameters shift and change in surprising and partially nonphysical ways. We discuss possible causes of this unexpected behavior and its implications on the calibration of DDE models to LES.

The rest of this paper is organized as follows. In Section 2, we review the nonlinear cloud and rain equation and introduce a new variation for which rain generation is dominated by accretion, rather than autoconversion. The LES we invert are described in Section 3. We briefly review the basic ideas of feature-based inversion in Section 4 and then discuss the details of our inversions in Section 5. We present and discuss our results in Section 6 and end the paper with a summary of conclusions in Section 7.

2 Predator-prey models for precipitating stratocumulus

Emergent behavior of precipitating stratocumulus cloud systems can be described via predator-prey interactions (Koren and Feingold, 2011; Feingold and Koren, 2013; Koren et al., 2017). Cloud droplets are the prey population and rain drops are the predator population. As clouds deepen, the number and size of cloud droplets increases. Eventually, the droplets are large enough to generate rain, which converts cloud water droplets into rain drops. As more and more cloud water is consumed by rain, the source of rain is reduced, eventually shutting it off completely. The process repeats once clouds have had enough time to re-develop.

There are several variants of predator-prey models for stratocumulus clouds. The first model, proposed by Koren and Feingold (2011), describes the time evolution and interaction of cloud depth (H in m) and droplet concentration (N in m^{-3}) via a set of two coupled DDEs. This first DDE model has no spatial scale but subsequently Feingold and Koren (2013) coupled several $H - N$ oscillators to obtain a system of DDEs with a spatial dimension. Koren et al. (2017) consider a single DDE, called the "nonlinear cloud and rain (C&R) equation," derived from the DDE of Koren and Feingold (2011). For all DDE models, closed-cell stratocumulus are represented by a steady state, for which the clouds rain at the same rate as they replenish; open-cell stratocumulus correspond to limit cycles (nonlinear oscillations). The bifurcation analysis of Koren et al. (2017) shows how the model's parameters define the transition from closed- to open-cell in the C&R equation. Further analysis and extensions of predator-prey dynamics and nonlinear DDEs have been studied by Chekroun et al. (2020) and Chekroun et al. (2022).

https://doi.org/10.5194/egusphere-2026-2871
Preprint. Discussion started: 18 June 2026

2.1 Nonlinear cloud and rain equation (autoconversion)

For the rest of this paper we focus on, and subsequently modify, the C&R equation of Koren et al. (2017). We summarize the derivation and some of the stability analysis here so our paper can be read independently and to motivate our modifications of this model further below.

Let H be the cloud depth with balance equation

$$\frac{\mathrm{d}H}{\mathrm{d}t} = \frac{H_0 - H}{\tau} + \dot{H}_r(t-T), \tag{1}$$

where t is time. The first term represents an exponential decay of cloud depth to a cloud carrying capacity (H_0) at a characteristic time scale τ (recharge). This exponential decay is caused by a dynamical forcing, including latent heating associated with phase changes of water, radiative flux divergence, entrainment-mixing, atmospheric instability and mesoscale forcing (Koren and Feingold, 2011). The cloud carrying capacity H_0, thus, represents the environmental potential for cloud development in the absence of sinks. The sink $\dot{H}_r$ describes the loss of cloud depth due to rain, i.e., as a result of converting small cloud droplets to raindrops. Since rain production depends on the past cloud state, the sink term is delayed (with microphysical delay T).

The loss term can be connected to the rain rate R by

$$\dot{H}_r = -\frac{R}{c_1 H}, \tag{2}$$

where $c_1 \simeq 2 \cdot 10^{-6}\,\mathrm{mm\,m^{-2}}$ (Koren and Feingold, 2011). If rain generation is dominated by autoconversion, the rain rate is, to first order

$$R = \frac{\alpha H^3}{N}, \tag{3}$$

where N is the (cloud-mean) droplet concentration (in $\mathrm{cm^{-3}}$) and $\alpha \simeq 2\,\mathrm{mm\,m^{-6}\,day^{-1}}$. Combining the above equations leads to the C&R equation [1]

$$\frac{\mathrm{d}H}{\mathrm{d}t} = \frac{H_0 - H}{\tau} - \frac{\alpha}{c_1 N} H^2(t-T). \tag{4}$$

In non-dimensional form, the model is

$$\frac{\mathrm{d}}{\mathrm{d}s} h = 1 - h - \frac{1}{\mu} h^2 (s - D), \tag{5}$$

where $h = H/H_0$, $s = t/\tau$, $D = T/\tau$, and $\mu = c_1 N/(\alpha H_0 \tau)$.

Koren et al. (2017) study the dynamical behavior of the C&R equation . For closed-cell stratocumulus, cloud depth is constant in time, so that $h(s) = h(s-D) = \bar{h}$, for large enough non-dimensional time s. We can find the steady state solution of the non-dimensional model by solving the quadratic equation

$$\bar{h}^2 + \mu\bar{h} - \mu = 0, \tag{6}$$

[1] Note that our formulation is slightly different from the formulation of Koren et al. (2017) because we separate α and c_1, whereas Koren et al. (2017) absorb c_1 in their definition of α. We also note that Koren et al. (2017) use the square root of the droplet concentration instead of the droplet concentration.

https://doi.org/10.5194/egusphere-2026-2871
Preprint. Discussion started: 18 June 2026

for the steady-state solution

$$\bar{h} = \sqrt{\frac{\mu^2}{4} + \mu} - \frac{\mu}{2}. \tag{7}$$

Note that we discard the negative root because cloud depth is non-negative.

To study the dynamical behavior and possible bifurcations, Koren et al. (2017) define δ to be a small perturbation of the steady state, i.e., $h(t) = \bar{h} + \delta(t)$. Linearizing around $\bar{h}$ and neglecting nonlinear terms in δ gives

$$\frac{\mathrm{d}\delta}{\mathrm{d}s} = -\delta - \frac{2\bar{h}}{\mu}\delta(s - D). \tag{8}$$

Koren et al. (2017) use the Ansatz $\delta = e^{\gamma s}$ to obtain a transcendental equation for the exponent

$$\gamma = -1 - \frac{2\bar{h}}{\mu} e^{-\gamma D}, \tag{9}$$

that has the closed form solution

$$\gamma = \frac{1}{D}\mathcal{W}\left(-\frac{2\bar{h}}{\mu} D e^{D}\right) - 1, \tag{10}$$

where $\mathcal{W}(\cdot)$ is the Lambert $\mathcal{W}$ function. Following the arguments of Koren et al. (2017), γ is real and negative as long as the argument of the Lambert $\mathcal{W}$ function is negative and larger than $-e^{-1}$:

$$\xi = -\frac{2\bar{h}}{\mu} D e^{D} > -e^{-1}. \tag{11}$$

Thus, as long as

$$\frac{2\bar{h}}{\mu} \leq \frac{1}{D e^{D+1}}, \tag{12}$$

the solution of the C&R equation will approach $\bar{h}$ exponentially without oscillations (overdamped). Note that the steady state solution $\bar{h}$ depends on the environmental parameters summarized in the non-dimensional parameter μ, but not the delay D, so that the condition in (12) separates the contributions from the delay and the environmental model parameters.

When $\xi < -e^{-1}$, the exponent γ in (10) is complex, and the solutions decay exponentially, but with oscillations (underdamped), as long as the real part of γ is less than zero, $\mathrm{Re}(\gamma) < 0$. The transition from real exponential decay (overdamped) to complex exponential decay, characterized by damped oscillations (underdamped) occurs when $\mathrm{Im}(\gamma) > 0$. When the real part of γ is non-negative,

$$\mathrm{Re}(\gamma) \geq 0, \tag{13}$$

the C&R equation exhibits limit cycles and, for large values of D (at fixed but small μ), the model generates a period doubling route to chaos (Koren et al., 2017). The chaotic regime, however, is characterized by negative cloud depth, which is unphysical.

https://doi.org/10.5194/egusphere-2026-2871
Preprint. Discussion started: 18 June 2026

2.2 Nonlinear cloud and rain equation (accretion)

We modify the C&R equation to model strongly precipitating stratocumulus for which rain generation is dominated by accretion, rather than autoconversion. In this case, the rain rate is independent of the droplet concentration N and, to first order, we can assume that

$$R = \beta H^4 (t - T), \tag{14}$$

where $\beta \simeq 0.7\,\mathrm{m\,day^{-1}\,km^{-4}}$. With this rain rate, the C&R equation becomes

$$\frac{\mathrm{d}H}{\mathrm{d}t} = \frac{H_0 - H}{\tau} - \frac{\beta}{c_1} H^3 (t - T). \tag{15}$$

In non-dimensional form, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}s} h = 1 - h - \frac{1}{\eta} h^3 (s - D), \tag{16}$$

where $h = H/H_0$, $s = t/\tau$, $D = T/\tau$ as before, but where $\eta = c_1/(\beta H_0^2 \tau)$.

We can mimic the stability and bifurcation analysis of Koren et al. (2017) for the C&R equation with accretion. The steady state cloud depth is the solution of the cubic equation

$$\bar{h}^3 + \eta \bar{h} + \eta = 0, \tag{17}$$

which has only one real and positive root that can be found numerically (the other two roots are a complex conjugate pair). As in the case of autoconversion, the steady state cloud depth depends only on environmental parameters (summarized in the non-dimensional parameter η), and not the delay. Figure 1(a) shows the steady state cloud depth for both C&R equations and we note that both models behave similarly, although the steady state cloud depth of the accretion model increases faster with its non-dimensional parameter.

Linearizing around $\bar{h}$ and neglecting nonlinear terms in the perturbations yields a differential equation for the perturbation δ:

$$\frac{\mathrm{d}\delta}{\mathrm{d}s} = -\delta - \frac{3\bar{h}^2}{\eta} \delta (s - D). \tag{18}$$

The same exponential Ansatz as before ($\delta = e^{\gamma s}$) gives a transcendental equation for the exponent γ:

$$\gamma = -1 - \frac{3\bar{h}^2}{\eta} e^{-\gamma D}, \tag{19}$$

which has the closed form solution

$$\gamma = \frac{1}{D} \mathcal{W}\left(-\frac{3\bar{h}^2}{\eta} D e^D \right) - 1. \tag{20}$$

The stability and oscillatory behavior of the solutions of the C&R equation with accretion depend, as before, on the argument of the Lambert $\mathcal{W}$ function

$$\xi = -\frac{3\bar{h}^2}{\eta} D e^D. \tag{21}$$

https://doi.org/10.5194/egusphere-2026-2871
Preprint. Discussion started: 18 June 2026

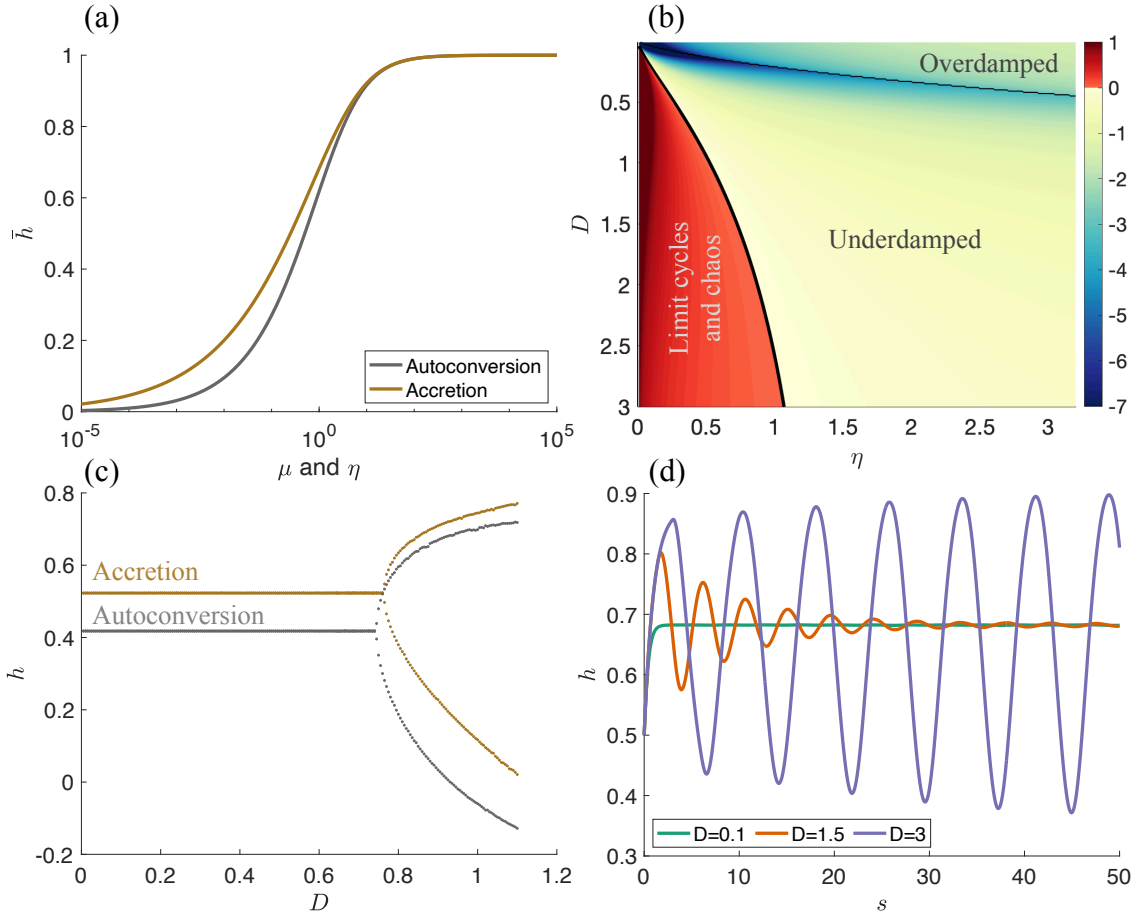

Figure 1. (a) Steady state cloud depth as a function of μ or η for the C&R equations with rain generation dominated by autoconversion (grey, μ) and accretion (brown, η). (b) Real part of the exponent γ that defines the stability of the C&R equation (with rain generation dominated by accretion). Highlighted are regions in parameter space for which the C&R equation exhibits real exponential decay (overdamped), damped oscillations (underdamped) and limit cycles or chaos. This panel should be compared to Figure 1a of Koren et al. (2017), which shows equivalent information for the C&R equation with rain generation dominated by auto-conversion. (c) System evolution from a stable fixed point to a limit cycle state for the C&R equation with rain generation dominated by accretion (grey) or autoconversion (brown). Shown are maximum and minimum values of h after a spin-up period for $\mu = \eta = 0.3$ (accretion/autoconversion) as a function of the delay D (compare to Figure 2(a) of Koren et al. (2017)). (d) Three solutions of the C&R equation (accretion) with initial condition $h(s) = 0.5$ for $s \leq 0$ and $\eta = 1$. For $D = 0.1$, the system exhibits real exponential decay (overdamped); for $D = 1.5$, we observe damped oscillations (underdamped); and for $D = 3$ the system exhibits a limit cycle.

If $\xi > -e^{-1}$, i.e., if

$$\frac{3\bar{h}^2}{\eta} < \frac{1}{De^{D+1}}, \tag{22}$$

https://doi.org/10.5194/egusphere-2026-2871
Preprint. Discussion started: 18 June 2026

the exponent γ is real and negative so that the solutions decay to $\bar{h}$ without oscillations (overdamped). When $\xi < -e^{-1}$, the exponent γ is complex and the solutions decay exponentially, but with oscillations, as long as $\mathrm{Re}(\gamma) < 0$ (underdamped). The transition from overdamped to underdamped behavior occurs when $\mathrm{Im}(\gamma) > 0$. The C&R equation with accretion exhibits limit cycles if $\mathrm{Re}(\gamma) \geq 0$. Figure 1(b) shows the real part of the exponent γ as a function of the non-dimensional parameters η and D and should be compared to Figure 1(a) of Koren et al. (2017), which shows equivalent information for the C&R equation with rain generation dominated by auto-conversion (see also Section 2.1). We highlighted the regions in parameter space for overdamped, underdamped and limit cycle states. Figure 1(c) shows the bifurcation diagram of the C&R equation with accretion for $\eta = 0.3$ and we can see the transition from a stable fixed point to a limit cycle state at $D \approx 0.76$. We do not study the behavior for larger delays D, because larger delays lead to negative cloud depth (as in the original C&R equation) and therefore are unphysical. We would not be surprised, however, if this modified C&R equation also follows a period-doubling route to chaos.

In summary, and perhaps not surprisingly, the small modification of the C&R equation to represent rain processes dominated by accretion leads to similar dynamics as the C&R equation with autoconversion, and we conclude that our new C&R equation may be used alongside the original C&R equation to study open- and closed-cell precipitating stratocumulus. In reality, autoconversion dominates the initiation of precipitation, while accretion dominates when precipitation is more fully developed. This distinction emphasizes the importance for us to consider both rain processes.

3 LES of precipitating open-cell stratocumulus

We consider a total of 127 LES described by Hoffmann et al. (2023). The LES differ in their initial/boundary conditions. For example, the temperature, moisture, boundary layer height, moisture and temperature jumps, aerosol levels, and surface flux variations within the atmosphere vary from one simulation to the next. As such, the suite of LES covers a wide range of meteorology and microphysical conditions.

The LES numerically solve the anelastic Navier-Stokes equations on an Eulerian spatial grid, resolving convection and clouds. Cloud microphysical processes are simulated using a bin-emulating, two-moment bulk scheme (Hoffmann et al., 2023), including the formation of droplets on suspended particles (suspended aerosol particles that act as condensation nuclei), their growth by diffusion and coalescence, and their removal by rain. Each LES covers a 48 km $\times$ 48 km $\times$ 2.5 km domain. Simulations are of 12 hour duration (nighttime). All LES use the same grids in space (200 m $\times$ 200 m $\times$ 10 m) and time (1 second). Model output is saved every 2 minutes. All LES are initiated in a quiescent state and turbulence develops during the first 200 minutes. This portion of the simulation is removed as "model spin-up." At some point after the spin-up, clouds thicken sufficiently, and/or have low enough drop concentrations that they can form rain and transition to the open-cell configuration.

We focus on LES with precipitating open-cell stratocumulus, which we identify as follows. First, we require that precipitation occurs and impose that:

(i) Surface precipitation reaches at least 0.001 mm/day;

(ii) the buildup of surface precipitation does not exceed 10 mm/day during the first two hours of the simulation.

https://doi.org/10.5194/egusphere-2026-2871
Preprint. Discussion started: 18 June 2026

Criterion (ii) filters for cases where significant rain formation would occur during spin-up but is prevented for technical reasons, giving rise to a nonphysically strong rain-out just after the spin-up phase.

If criteria (i) and (ii) are satisfied, we identify the stable formation of an open-cell regime by finding the onset of oscillation in cloud fraction. Specifically, we first de-trend cloud cover fraction because it gradually decreases over time (for each LES). After subtracting the mean, we find the onset of oscillation by determining when the cloud cover time series changes sign. We then shorten the cloud cover time series up to this point and repeat the process, which converges if the simulation develops a stable open-cell regime. If we detect convergence, the LES is accepted as a simulation of a precipitating open-cell stratocumulus and we remove an additional spin-up to the open-cell regime (determined via the above iterative process). All subsequent analysis is performed after the spin-up process to the open-cell regime, including the development of turbulence (first 200 mins for each LES) and the additional spin-up to the open-cell configuration (variable across the LES).

With our selection criteria, we obtain a suite of 16 LES of precipitating open-cell stratocumulus. One of these simulations (3-011), however, is characterized by a small moisture difference between the cloud and the free troposphere. The small moisture difference prevents drying of the boundary layer through cloud-top mixing and thus produces more rain when compared to the other 15 LES, leading to clouds that are more akin to a fog than a precipitating open-cell stratocumulus. Thus, while LES 3-011 (barely) cleared our criteria for precipitating open-cell stratocumulus, we exclude it from subsequent analysis because it is essentially a simulation of a different type of cloud.

To further understand the scenarios modeled by the suite of LES, we compute a typical cloud depth H_{75} and a drop effective radius r_e. Typical cloud depth is defined as the 75th percentile of the cloud depth distribution of an LES. Following Wood (2006), the drop effective radius is defined by

$$r_e = \left(\rho_w \frac{3}{4\pi} \cdot \frac{H_{75}\Gamma}{N}\right)^{1/3}, \tag{23}$$

where, $\rho_w = 1$ is the density of water (in $\mathrm{g/cm^3}$), Γ is the rate at which liquid water content increases with respect to height (in $\mathrm{cm^2/g}$), and N is the spatial and temporal average of droplet concentration weighted by cloud fraction (in $\mathrm{cm^{-3}}$). Figure 2(a) shows typical cloud depth H_{75} as a function of effective radius r_e. We observe a linear relationship between typical cloud depth H_{75} and effective radius r_e, so that all-else-equal, as expected, deeper clouds generate larger drops. The exception is LES 3-012, which is an extremely polluted simulation, leading to a relatively high droplet concentration ($> 60\ \mathrm{cm^{-3}}$), with an abnormally high initial liquid water path (LWP) that affects the effective radius.

In general, deeper clouds tend to rain, and rain leads to washout of aerosols (scavenging). The washout of aerosols decreases the droplet concentration, meaning deeper precipitating clouds have smaller droplet concentrations. Figure 2(b) illustrates this effect and shows typical cloud depth H_{75} as a function of droplet concentration N. We note again that LES 3-012 is different from the other 14 LES, because it is characterized by a large cloud depth and a large droplet concentration. Thus, although LES 3-012 made it formally through our selection criteria for precipitating open-cell stratocumulus, we remove it going forward and consider a suite of 14 LES, each modeling precipitating open-cell stratocumulus for which deeper clouds are associated with larger droplets and smaller droplet concentrations. This suite of 14 LES will be used to evaluate the extent to which one may use the C&R equation as a quantitative tool for representing selected aspects of an LES.

https://doi.org/10.5194/egusphere-2026-2871
Preprint. Discussion started: 18 June 2026

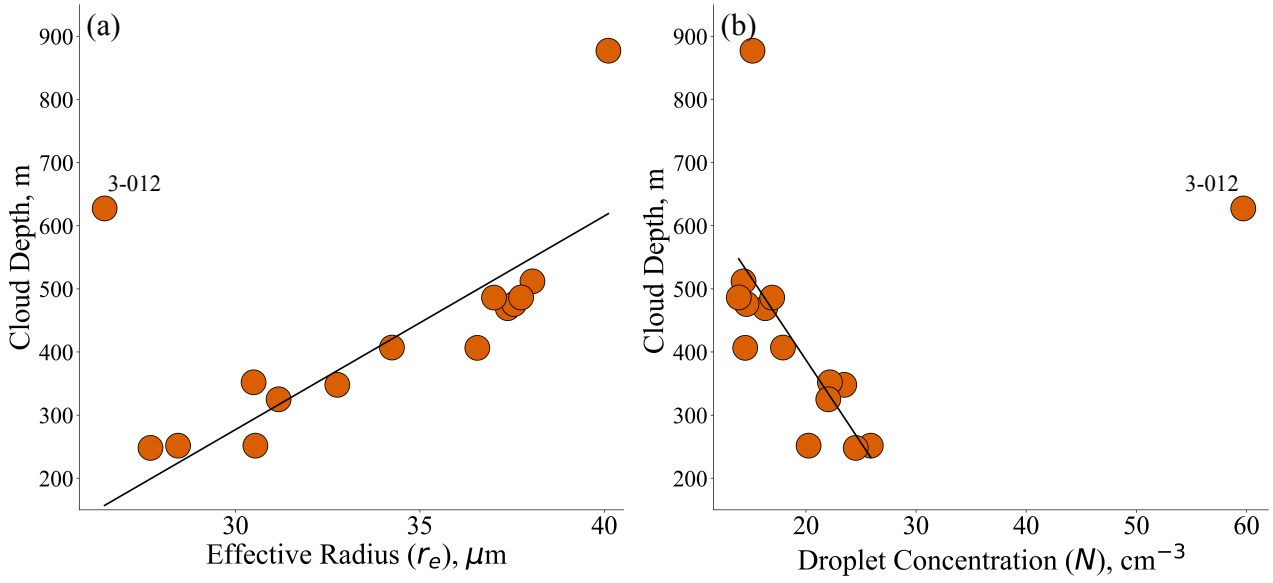

Figure 2. (a) Typical cloud depth H_{75} as a function of effective radius (r_e) and associated least squares fit (black line, $R^2 = 0.72$). (b) Typical cloud depth H_{75} as a function of droplet concentration N and associated least squares fit (black line, $R^2 = 0.48$). LES 3-012 was removed for the least squares fits.

4 Review of feature-based Bayesian inversion

Below, we will use feature-based Bayesian inversion to compute model parameters of the C&R equations from LES outputs, but we first briefly review the basics of this method. We start with the "traditional" (not feature-based) Bayesian inversion, which amounts to computing a Bayesian posterior distribution that describes the probability of model parameters θ given data y (see, e.,g., Sanz-Alonso et al., 2023). Model parameters may be boundary or initial conditions, or parameters of the C&R equations, such as H_0, τ, T, N, and β. The posterior distribution is the product of a prior and a likelihood

$$p(\theta|y) \propto p_0(\theta)\, p_l(y|\theta)\,. \tag{24}$$

The prior, $p_0(\cdot)$, contains information about the model parameters independent of the data y, e.g., positivity constraints or parameter bounds. The likelihood, $p_l(y|\theta)$ connects the model with the data. The model is abstractly written as $\mathcal{M}(\theta)$, so that the model parameters θ are inputs and the output of the model can be compared to the data:

$$y = \mathcal{M}(\theta) + \varepsilon. \tag{25}$$

Here, ε is a random variable with known statistics, e.g., Gaussian, that models discrepancies between the model $\mathcal{M}(\cdot)$ and the data y. Since the statistics of ε are known, equation (25) defines the likelihood $p_l(y|\theta)$. With prior and likelihood in hand, the posterior is proportional to their product and one can use optimization or sampling techniques to obtain a numerical approximation of the posterior distribution and, hence, a numerical solution of the Bayesian inverse problem. Below, we use

https://doi.org/10.5194/egusphere-2026-2871
Preprint. Discussion started: 18 June 2026

Markov chain Monte Carlo (MCMC) to draw samples from a posterior distribution. In short, the basic idea of MCMC is to set up a Markov chain whose stationary distribution is equal to the posterior distribution. Running the Markov chain sufficiently long then yields (correlated) samples of the posterior distribution. We use a tried-and-true MCMC sampler that is known to be robustly applicable (see Section 5.3).

Feature-based inversion is a variation of Bayesian inversion with more general interpretations of "data" y and model $\mathcal{M}(\cdot)$ (Morzfeld et al., 2018; Maclean et al., 2017; Haario et al., 2015; Hakkarainen et al., 2012). Specifically, y could be a vector that represents selected aspects of a larger data set, and the model $\mathcal{M}(\cdot)$ may be a model for only some selected aspects, or *features* of the larger data set. Below we will explain in detail how we construct a feature model $\mathcal{M}(\cdot)$, based on the C&R equations, and how we extract features from the LES. For other successful feature-based inversions in the context of chaotic systems, see, e.g., Kazarnikov and Haario (2020); Springer et al. (2021, 2019); Maraia et al. (2021); Srivastava et al. (2023). Numerically, feature-based inversions can be performed with the same MCMC methods that apply to other Bayesian inversion problems.

5 Feature-based inversion of a suite of LES

We now apply feature-based inversion to compute parameters of the C&R equations based on output variables of the suite of 14 LES of precipitating open-cell stratocumulus. In short, feature-based inversion means that we define a prior distribution to enforce constraints on the parameters of the C&R equations (Section 5.1), and a feature-based likelihood to connect an LES with the C&R equations (Section 5.2). As a feature, we select cycles of cloud growth and decay that we extract from an LES and then set up a feature-based likelihood to compare the LES cycles to limit cycles of the C&R equations. Jointly, the prior and likelihood define a feature-based posterior distribution which we sample via MCMC (Section 5.3).

We will consider both versions of C&R equations and estimate all of their parameters. For the C&R equation with rain generation dominated by autoconversion, we estimate droplet concentration N, along with the model's environmental parameters (H_0 and τ) and the microphysical delay T, summarized in the parameter vector

$$\theta_{\text{acon}} = \Big(H_0, \quad \tau, \quad T, \quad N \Big). \tag{26}$$

Note that here we deviate from the approach described by Lunderman et al. (2020), where N was retrieved directly from the LES and where the scaling factor α was treated as a model parameter. We decided to fix α and invert for the droplet concentration because Koren et al. (2017) defined the scaling factor α to be fixed and suggest that N is a model parameter, and because inverting for the droplet concentration is physically more meaningful.

For the C&R equation with rain generation dominated by accretion, the rain rate is independent of the droplet concentration N and we estimate the scaling factor β instead, leading to the parameter vector

$$\theta_{\text{accr}} = \Big(H_0, \quad \tau, \quad T, \quad \beta \Big). \tag{27}$$

Selecting β as a model parameter, rather than a fixed constant, is reasonable as β now controls the rain generation in the C&R equation equation with accretion instead of the droplet concentration N. The feature-based inversions will explore a wide range of model parameters with bounds listed in Table 1.

https://doi.org/10.5194/egusphere-2026-2871
Preprint. Discussion started: 18 June 2026

Parameter	Description and units
$0 < H_0 \leq 4$	Cloud depth (km)
$0 < \tau \leq 720$	Recharge timescale (min)
$0 < T \leq 720$	Microphysical delay (min)
$0 < N < 100$	Droplet concentration (cm^{-3}, autoconversion)
$0 < \beta \leq 4$	Scaling factor ($\mathrm{day}^{-1}\,\mathrm{m\,km}^{-4}$, accretion)

Table 1. Model parameters of the C&R equations and their bounds.

As already noted, we consider two variants of C&R equations and estimate their parameters from 14 LES. This means that we perform a total of 28 inversions (two models, 14 data sets). In terms of distributions, we have two prior distributions for the two variants of the C&R equation and 14 likelihoods defined by the 14 LES. Below, we largely follow the inversion process outlined in Lunderman et al. (2020), but the details of our implementation are more flexible and automated, minimizing the number of tunable inversion parameters, which allows us to process and invert the suite of 14 LES for both variants of the C&R equation.

5.1 Prior distribution

Recall that the prior distribution imposes constraints on model parameters that are independent of the data. In the context of inverting LES features to compute parameters of C&R equations, this means that the prior enforces the following three constraints (Lunderman et al., 2020):

1. The parameters are within the bounds listed in Table 1.
2. The recharge time scale τ is larger than the microphysical delay T.
3. The parameters lead to limit cycles with positive cloud depth.

The 3rd condition can be imposed using the results of the stability analysis presented earlier. Specifically, the C&R equation with rain generation dominated by autoconversion exhibits limit cycles if the real part of γ in (10) is positive. Similarly, the C&R equation with rain generation dominated by accretion exhibits limit cycles if the real part of γ in (20) is positive. Thus, we impose condition (3) by first checking if $\mathrm{Re}(\gamma) > 0$ and, if it is, we numerically integrate the C&R equation. If we encounter negative cloud depth, we discard the parameter set – because the limit cycle is not physically meaningful. Note that condition (3), and in particular the requirement that $\mathrm{Re}(\gamma) > 0$ couples all four parameters of the C&R equation (see the definitions of γ in (10) and (20)).

We visualize the prior distribution and, hence a priori dependencies between model parameters via a triangle plot, which contains histograms of all one and two-dimensional marginals of a multivariate probability distribution. Figure 3(a) shows the triangle plot of the prior for the C&R equation with rain generation dominated by autoconversion. To understand the triangle

https://doi.org/10.5194/egusphere-2026-2871
Preprint. Discussion started: 18 June 2026

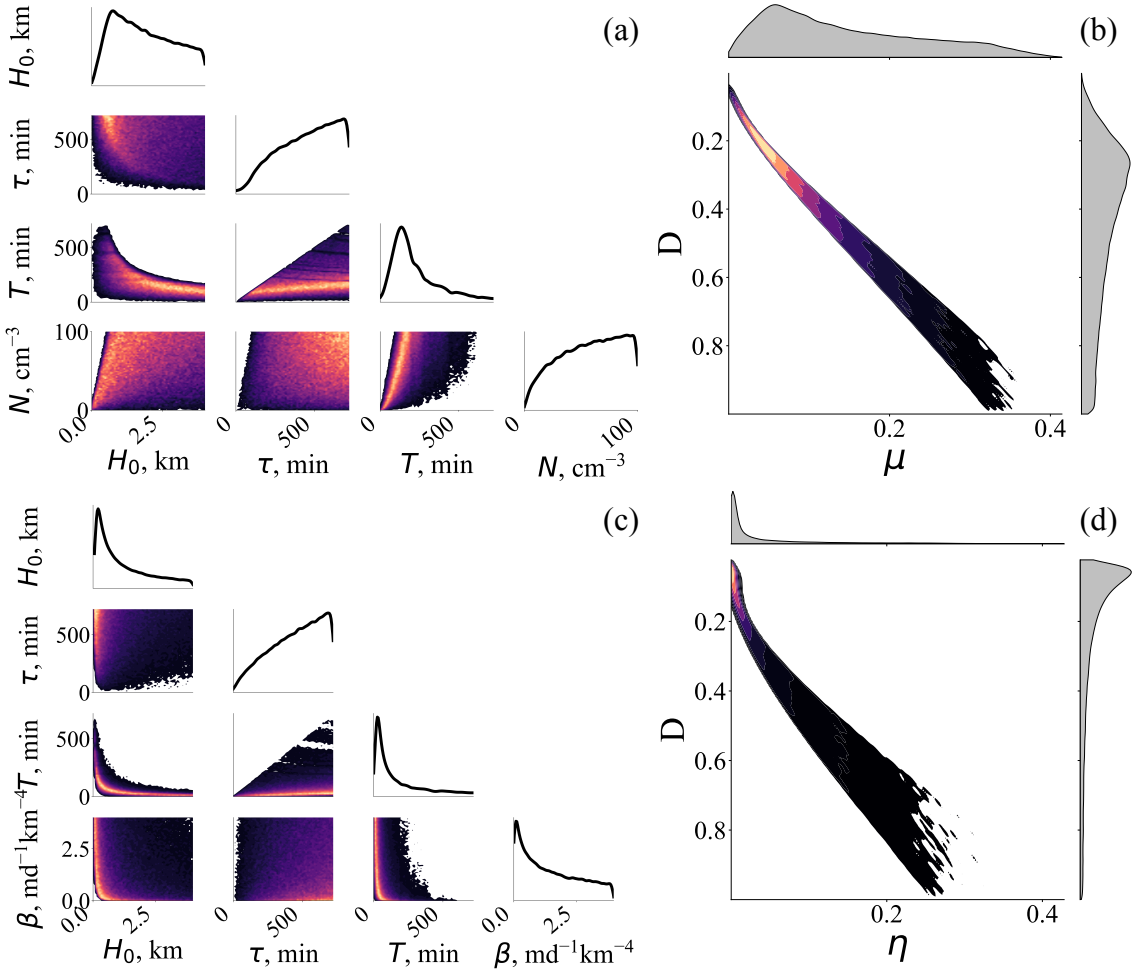

Figure 3. (a) Triangle plot of the prior distribution over the parameters of the C&R equation with rain generation dominated by autoconversion. Each parameter's histogram is plotted on the diagonal and the lower triangle contains histograms of every combination of any two parameters. Warm colors (yellow and pink) represent regions of high prior probability and cold colors (black and purple) represent lower prior probability. White regions contain few or no samples and, therefore, correspond to regions of very low prior probability. (b) Histogram of the two non-dimensional parameters of the C&R equation with rain generation dominated by autoconversion. The top and right sides also show histograms of μ and D. (c) Triangle plot of the prior distribution over the parameters of the C&R equation with rain generation dominated by accretion. (d) Histogram of the two non-dimensional parameters of the C&R equation with rain generation dominated by accretion. The top and right sides also show histograms of η and D.

plot, we note that peaks in the diagonal panels define regions of high prior probability for each model parameter and that ridges of warm colors in the off-diagonal panels indicate high prior probability among two parameters. We can thus conclude that the three basic assumptions that define the prior (parameter bounds, $\tau > T$, limit cycles with positive cloud depth) render some parameter values more likely than others (peaks in diagonal panels) and also impose correlations/dependencies among the

https://doi.org/10.5194/egusphere-2026-2871
Preprint. Discussion started: 18 June 2026

model parameters (ridges of warm colors in off-diagonal panels). We note in particular strong dependencies between droplet concentration N and the delay T, between N and the cloud carrying capacity H_0, as well as a strong relationship between H_0 and the delay T.

There is also a preferred linear relationship between the time scale τ and the delay T, but there are only weak dependencies between N and τ or H_0 and τ. The prior for the C&R equation with rain generation dominated by accretion is very similar, as shown in Figure 3(c). The diagonal and off-diagonal plots corresponding to the model parameters H_0, τ and T are qualitatively similar to the prior for the C&R equation with rain generation dominated by autoconversion. The prior over the scaling factor β of the C&R equation with rain generation dominated by accretion explores a wide range of values, but favors smaller values of β. We also note that there are strong correlations between the scaling factor β and the cloud carrying capacity H_0 and the delay T, while we do not detect an obvious relation between β and the recharge timescale τ.

One may also collapse all physical (dimensional) parameters into the non-dimensional parameters D and μ (or D and η, for accretion) and consider the prior in non-dimensional parameter space. We show a histogram for the prior of the C&R equation with rain dominated by autoconversion in Figure 3(b), with an equivalent figure for the C&R equation with accretion shown Figure 3(d). We note a strong correlation between μ and D (or η and D for accretion), but this correlation is contaminated by spurious correlations caused by the common denominator τ (see, e.g., Davis, 2002). We also see how the prior constraint of limit cycles with positive cloud depth (i.e. behavior associated with precipitating open-cell stratocumulus clouds) cuts across the much larger limit cycle/chaotic regime shown as a red region in Figure 1(b) for the C&R equation with rain generation dominated by accretion (or Figure 1(a) of Koren et al. (2017) for the C&R equation with rain generation dominated by autoconversion). Comparing the prior in non-dimensional and dimensional parameter spaces shows how tight constraints on μ and D (or η and D) map to a wide range of dimensional parameters, with preferred regions of high prior probability. As noted by Koren et al. (2017), it may be difficult to go from the non-dimensional parameters to physical parameters and, for that reason, we will invert for the physical parameters directly.

5.2 Feature-based likelihood

Defining the feature-based likelihood requires that we specify an *LES-feature* and a corresponding *model feature*, along with an error model that describes discrepancies between the C&R equation and the LES.

5.2.1 LES-feature

Following Lunderman et al. (2020), the LES-feature is a typical cycle of cloud growth and decay, defined as follows. We apply a spatial smoothing to the LES cloud depth field to reduce the effects of noise, while retaining the main aspects of the cellular cloud structure. After spatial smoothing, we obtain, for each LES, 900 time series of cloud depth (one time series per spatially smoothed grid point). Each time series is smoothed in time and we then determine local extrema of the temporally smoothed time series via finite differencing. Two consecutive local minima define one cycle and each cycle is stored (without temporal smoothing). We only consider cycles that start below, grow above, and end below the typical cloud depth H_{75} (defined in Section 3) to ensure that the cycle formation is from local cloud growth and decay, not from nearby clouds. The LES-feature is

https://doi.org/10.5194/egusphere-2026-2871
Preprint. Discussion started: 18 June 2026

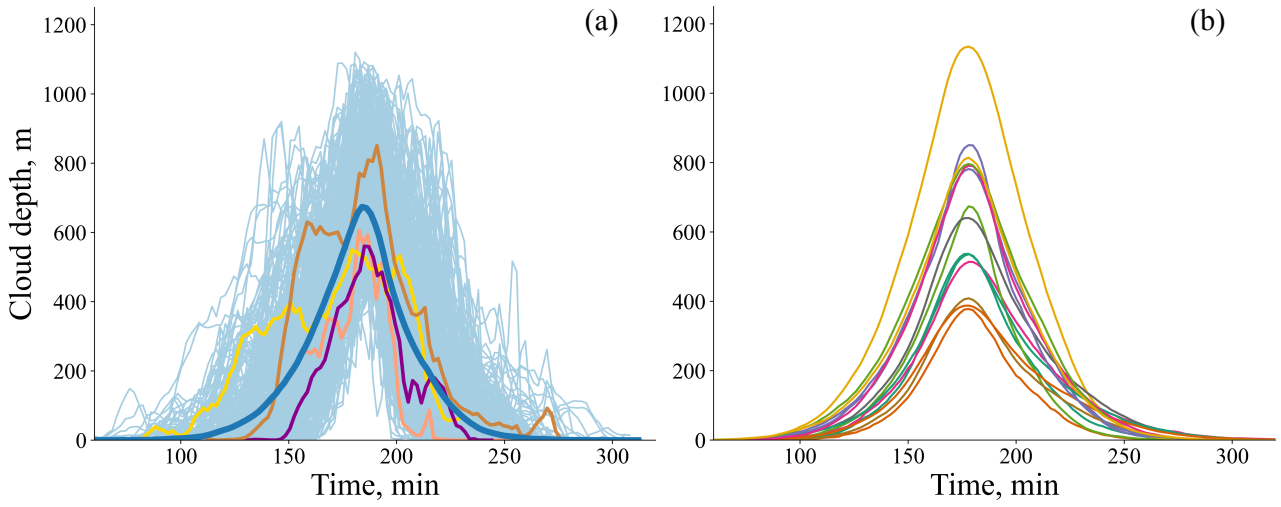

Figure 4. Illustration of the LES-feature. (a) Cycles of growth and decay of one LES (aligned at their maxima). The light blue lines represent all cycles extracted from an LES. The purple, orange, yellow, and brown lines highlight a few examples. The dark blue line is the LES-feature (the average of all LES cycles). (b) LES-features of the 14 LES we consider.

the average of all cycles extracted from an LES. We note that the cycles have different periods, which we address by padding shorter cycles with zeros (following, again, Lunderman et al. (2020)).

An example of an LES-feature is shown in Figure 4(a) (adapted from Lunderman et al. (2020)). In Figure 4(a), the light blue lines represent all cycles we identified within an LES. The dark blue line is the average of all cycles (the LES-feature). The yellow, pink, purple, and brown cycles are examples of LES cycles. Figure 4(b) shows the LES-features of the suite of 14 LES we consider. We note a large spread in both period and amplitude of the LES-features across the suite. The large spread arises because the suite of LES covers a wide range of meteorological conditions of precipitating open-cell stratocumulus and the LES feature we define is sensitive to these changing conditions.

5.2.2 Model-feature

The model-feature ($\mathcal{M}(\theta)$ in (25)) is a limit cycle of the KTF model. We compute limit cycles iteratively as described by Lunderman et al. (2020). We first simulate the KTF model for a duration of one-day and compare the last two cycles. If the cycles are different, as described by a normed error exceeding a threshold, we continue the simulation for another day and compare the last two cycles of the extended simulation. This process is repeated until we reach convergence and, hence, can extract a limit cycle. The feature-model $\mathcal{M}(\cdot)$ is thus the computer code that executes the above steps to extract a limit cycle from a C&R equation.

https://doi.org/10.5194/egusphere-2026-2871
Preprint. Discussion started: 18 June 2026

5.2.3 Modeling uncertainties in the feature

With model and LES-features in hand, we still need to define the random variable ε that describes discrepancies between the limit cycles of a C&R equation and an LES. As is common, we assume Gaussian errors with mean zero and covariance matrix R. With these choices, we obtain a feature-based likelihood of the form

$$p_l(y|\theta) \propto \exp\left(-\frac{1}{2}\left\|R^{-1/2}(\mathcal{M}(\theta)-y)\right\|^2\right), \tag{28}$$

where vertical bars denote the two-norm and $R^{1/2}$ is a matrix square root of the symmetric positive definite matrix R. We compute the covariance matrix R directly from the variations in the cycles extracted from an LES (light blue lines in Figure 4(a)). However, the variation is low towards the beginning and end of the average cycle, which is an artifact of padding the LES cycles when averaging. We address this issue by inflating the covariance diagonal by 10% of the maximum variance, σ^2_{max} (i.e. we add $0.1 \times \sigma^2_{\text{max}}$ to each value along the covariance diagonal).

Finally, we note that the construction of the feature-based likelihood depends on a few ad hoc choices, e.g., with respect to the spatial and temporal smoothing of the LES cloud depth fields and also on how the cycles of different periods are handled. Lunderman et al. (2020) present a thorough study of the sensitivity of the inversion results to such choices and conclude that the various choices have only minor effects on the inversion results. Going forward, we thus only consider the "default" setup described by Lunderman et al. (2020).

5.3 Numerical solution via MCMC

We numerically solve the feature-based inverse problems via MCMC, using the affine invariant ensemble sampler of Goodman and Weare (2010) in its robust and well-tested python implementation called *emcee* (Foreman-Mackey et al., 2013). Emcee is an ensemble sampler, which means that it runs several interacting Markov chains. The use of an ensemble is important here, because it ensures affine-invariance (Goodman and Weare, 2010), which is critical for problems with strong correlations across model parameters. Specifically, we use emcee with a 20-member ensemble, each taking 10^5 steps, to generate 2×10^6 posterior samples.

To ensure accurate sampling, we need to ensure that the Markov chain is long enough to account for two effects: (i) a transient period, or burn-in, during which the Markov chain transitions from an arbitrary initial state to its stationary distribution; and (ii) how quickly the chain de-correlates, because samples would ideally be uncorrelated. To measure the de-correlation, we compute the integrated auto correlation time (IACT) of the chains for each parameter and obtain maximum values of about 10^3 (meaning only every 1000 samples are effectively independent). A good guideline is to remove about 10 IACTs worth of samples as burn-in, but we are overly cautious and consider the first 25,000 steps as burn-in (and hence these steps are removed). We thus have 20 chains, each producing $7.5 \cdot 10^4$ samples after burn-in. Dividing by the largest IACT means that we have about $20 \cdot 7.5 \cdot 10^4 \cdot 10^{-3} = 1500$ effectively independent posterior samples.

https://doi.org/10.5194/egusphere-2026-2871
Preprint. Discussion started: 18 June 2026

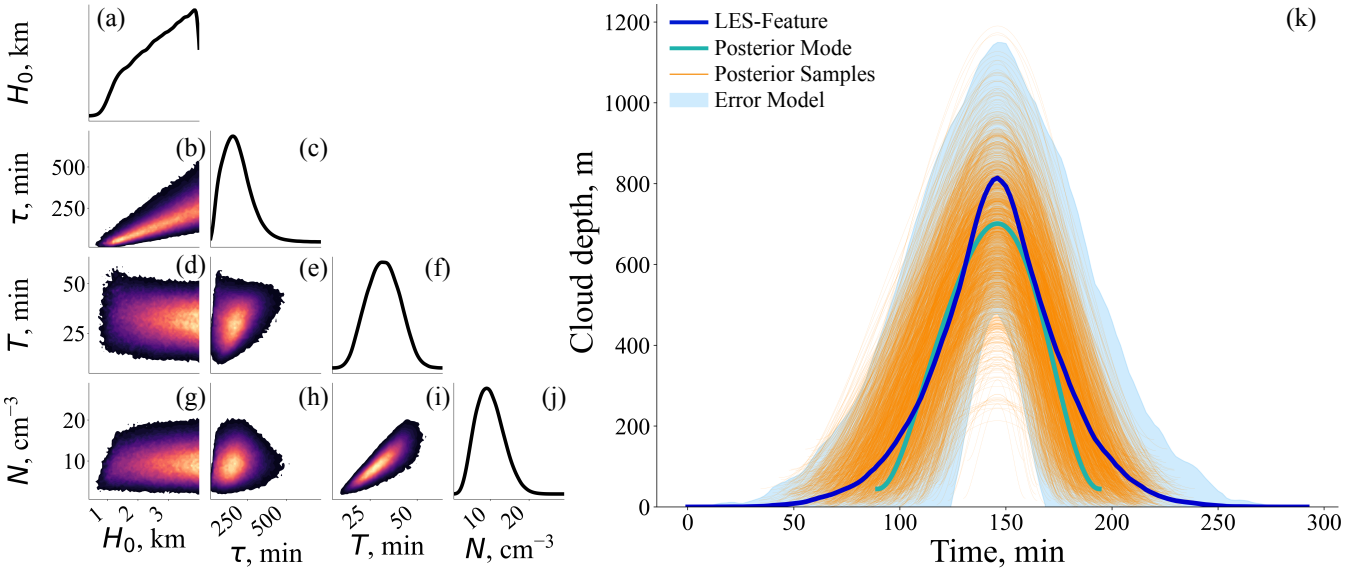

Figure 5. Feature-based inversion results for LES-4-023 with the C&R equation with rain generated by autoconversion. (a)-(j) Triangle plot of the feature-based posterior (based on 1.5×10^6 MCMC samples). Each parameter's histogram is plotted on the diagonal and the lower triangle contains histograms of every combination of any two parameters. Warm colors (yellow and pink) represent regions of high posterior probability and cold colors (black and purple) represent lower posterior probability. White regions contain few or no samples and, therefore, correspond to regions of very low posterior probability. (k) LES-feature (dark blue), error model (light blue), and limit cycles of the C&R equation (autoconversion) corresponding to draws from the posterior (orange). The green line is the limit cycle corresponding to the MAP estimate.

6 Results and discussion

We interpret the inversion results and report how the 14 LES create dependencies among model parameters of the C&R equations. First we consider how any one of the LES imposes parameter dependencies and then trace how variations in meteorological variables of the suite of LES causes variation in the 14 posterior distributions for a C&R equation.

6.1 Variation in the parameters of C&R equations

We describe how the parameters of the C&R equations vary and co-vary within a posterior distribution corresponding to one of the LES.

We first consider the C&R equation with rain generation dominated by autoconversion and show in Figures 5(a)-(j) a triangle plot of one of the 14 posterior distributions, which are all qualitatively similar (see Figure A1, which shows equivalent plots for all 14 posterior distributions associated with this C&R equation). Figure 5(k) illustrates how the posterior distribution over model parameters maps to cycles of cloud growth and decay. Specifically, the figure contains the LES-feature (dark blue) and the expected variation around the LES-feature (light blue), along with a few limit cycles of the C&R equation with parameters

https://doi.org/10.5194/egusphere-2026-2871
Preprint. Discussion started: 18 June 2026

drawn from the posterior distribution (orange). The green line in Figure 5(k) corresponds to the limit cycle produced by the maximum a posteriori (MAP) estimate. The MAP estimate is the set of parameters that gives the largest posterior probability value. We note that most limit cycles lie close to the LES-feature and within the assumed errors. Again, this is true for all 14 inversions we performed.

Studying the triangle plot of the posterior distribution in more detail reveals how the various model parameters vary and co-vary:

1. The delay T is well-constrained by the LES-feature, resulting in a tight, nearly Gaussian posterior distribution with a small variance (Figure 5(f)).

2. The marginal distributions of the cloud carrying capacity (H_0) and recharge timescale (τ) have a larger variance and more non-Gaussian features (Figures 5(a) and (c)). Compared to the prior (see Figure 3), the variance in the recharge timescale τ is reduced, suggesting that the LES constrains this parameter beyond the prior assumptions (as already reported by Lunderman et al. (2020)). The constraint on the cloud carrying capacity is less strong, but the inversion has "shifted" from a prior favoring relatively small values of H_0 to a posterior that tends to favor larger cloud carrying capacities.

3. The marginal posterior distribution of the droplet concentration N in the C&R equation is tightly constrained and variance in N is reduced compared to the prior, again indicating that the LES imposes constraints on the droplet concentration of the C&R equation.

4. For all 14 inversions, we find that the recharge timescale τ increases with cloud carrying capacity H_0 (Figure 5(b)), which is expected, because deeper clouds are characterized by a longer recharge time.

5. All 14 inversions exhibit a strong correlation between the droplet concentration N and the microphysical delay T (Figure 5(i)), with larger drop concentrations typically resulting in larger delays. This is expected because, microphysically, a larger droplet concentration delays the formation of rain via collision-coalescence (the sink term in the C&R equation).

6. The 14 inversions are consistent in that other pairs of model parameters (H_0-T, H_0-N, τ-T, τ-N, Figures 5(d), (e), (g), (h)) do not exhibit obvious relationships. The marginals of the posterior distributions are either nearly uniform (H_0-T, H_0-N, Figures 5(d), (g)) or "blobs" of high posterior probability (τ-T, τ-N, Figures 5(h), (e)).

We now consider the C&R equation with rain generation dominated by accretion. A representative triangle plot of a posterior distribution, along with how well the model fits the LES data is shown in Figure 6 and all 14 posterior distributions associated with that model are shown in Figure A2. The two versions of the C&R equation share the parameters H_0, τ and T and we see from Figures 5 and 6 that these three parameters vary and co-vary similarly in both models. This is reassuring, and we focus our attention on how the scaling factor β varies and covaries with other model parameters. First, we note that there is no obvious relationship between the scaling factor β and the cloud carrying capacity H_0 or the recharge timescale τ (Figures 6(g) and (h), similar to how N covaries with H_0 and τ in the C&R equation with rain generation dominated by autoconversion).

https://doi.org/10.5194/egusphere-2026-2871
Preprint. Discussion started: 18 June 2026

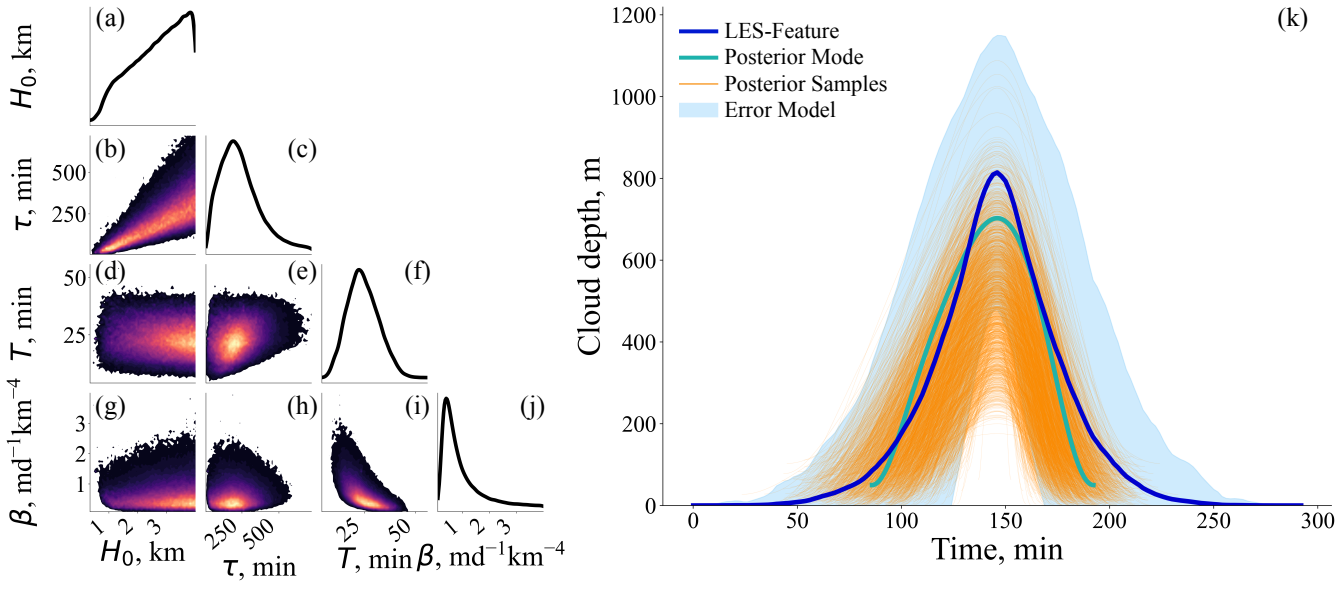

Figure 6. Feature-based inversion results for LES-4-023 with the C&R equation with rain generation dominated by accretion. (a)-(j) Triangle plot of the feature-based posterior (based on 1.5×10^6 MCMC samples). Each parameter's histogram is plotted on the diagonal and the lower triangle contains histograms of every combination of any two parameters. Warm colors (yellow and pink) represent regions of high posterior probability and cold colors (black and purple) represent lower posterior probability. White regions contain few or no samples and, therefore, correspond to regions of very low posterior probability. (k) LES-feature (dark blue), error model (light blue), and limit cycles of the C&R equation (accretion) corresponding to draws from the posterior (orange). The green line is the limit cycle corresponding to the MAP estimate.

Second, we note in Figure 6(i) that, in each LES, the scaling factor β decreases with the microphysical delay T. This is to be expected because a more efficient rain process (larger β) is equivalent to a smaller microphysical delay.

Finally, we note that the scaling factor is well constrained if it takes on small values, but the variance of this model parameter is larger when the parameter β itself is larger (see Figure A2). Physically, large β values would be associated with conditions of co-occurring large cloud water (H) and large rain water (R). Generally speaking, the value of β is uncertain (Geoffroy et al., 2008) and therefore it is not surprising that it might vary significantly in heavily raining clouds.

In summary, the variation in the model parameters driven by any *one* of the LES is as expected and highlights the strengths of this simplified model – the model parameters vary and covary in a physically meaningful way, while generating limit cycles that are comparable with the cycles of cloud growth and decay in the LES. The above holds true for both variants of the C&R equation (autoconversion or accretion) and for all 14 LES, highlighting the robustness of the C&R equation and the underlying predator-prey analogy.

https://doi.org/10.5194/egusphere-2026-2871
Preprint. Discussion started: 18 June 2026

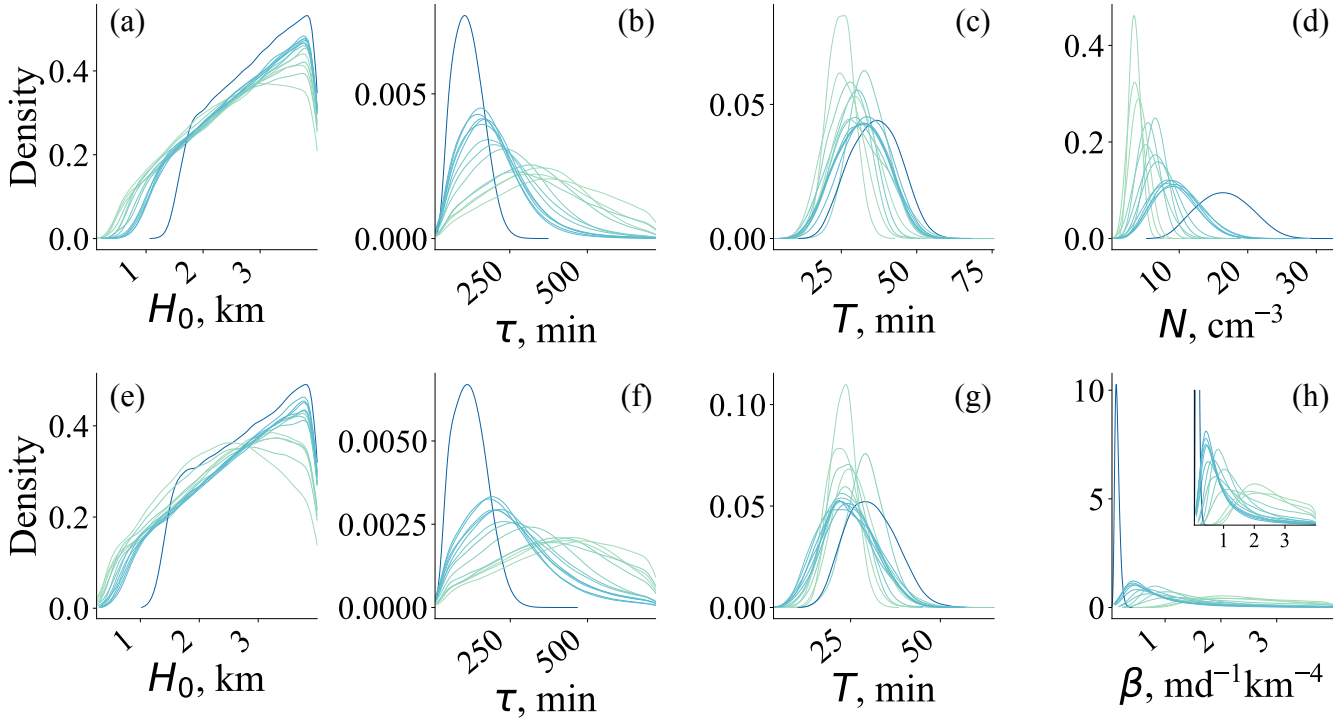

Figure 7. Marginal posterior distributions of the C&R equations with rain generation dominated by autoconversion (a-d) or accretion (e-h). The posterior distributions are colored by the typical cloud depth of the LES that is being inverted, with darker colors corresponding to a higher typical cloud depth.

6.2 Variation of model parameters across the suite of LES

We now consider the inversions jointly, i.e., we study how the posterior distributions over model parameters change across the suite of 14 LES. This approach allows us to investigate if the model parameters represent changes in the micro- and macro-physics, as represented by the suite of LES, in a physically meaningful way.

We show all 28 posterior distributions in the appendix (Figure A3) and note that there is a large overlap between the 14 posterior distributions of each model and that all 14 posterior distributions are *qualitatively* similar for each model (and in line with what was already reported in Lunderman et al. (2020)). The discussion and interpretation of the results, however, can be streamlined by considering only the one-dimensional marginals over the four model parameters shown in Figure 7.

First, we note that the posterior distributions of the microphysical delay T exhibit a large overlap across the 14 inversion (for both variants of the C&R equation) and each distribution is tightly constrained (Figure 7(c) and (g)). Thus, our inversions indicate that the microphysical delay does not vary much across the suite of LES. Estimating the delay in the LES more directly by the lag between cloud water path (CWP) and rain water path (RWP) indeed confirms this result. The delays of the 14 LES are between 9 minutes and 16 minutes (average delay across the 14 LES is 11 minutes with a standard deviation

https://doi.org/10.5194/egusphere-2026-2871
Preprint. Discussion started: 18 June 2026

of 2 minutes). The C&R equations have posterior means ranging from 24-37 minutes (autoconversion) and 22-31 minutes (accretion). The posterior means of the C&R equation (rain generation dominated by autoconversion) average to 31 minutes with a standard deviation of 3 minutes. The posterior means of the C&R equation (rain generation dominated by accretion) average to 25 minutes with a standard deviation of 2 minutes. Our results suggest that the inversion results across the 14 LES are self-consistent and physically meaningful for the delay parameter T for both variants of the C&R equation.

The posterior distributions of the cloud carrying capacity H_0 also show a large overlap (Figure 7(a) and (e)). Comparing the posterior to the prior, cloud carrying capacity is perhaps the parameter that is least constrained by the LES, indicating a limited sensitivity of the limit cycles of the C&R equation to H_0. We do find, however, that the posterior average of the cloud carrying capacity correlates well with the typical cloud depth H_{75} of the LES. There are positive linear relationships between H_{75} and the posterior averages of H_0 for both C&R equations, with R^2 values of 0.88 and 0.75 for rain generated by autoconversion and accretion, respectively. This reiterates that the inversion is self-consistent and physically meaningful for the cloud carrying capacity H_0.

Considering the recharge time scale τ, we note that the distributions shift towards smaller values as the typical cloud depth H_{75} of the LES increases (Figure 7(b) and (f)). This is counterintuitive because deeper clouds typically tend to take *longer* to recharge, i.e., we would expect that the distribution over the recharge time scale τ should shift towards *larger* values with increasing typical cloud depth. On the other hand, one may also argue that deeper clouds occur in a more dynamic environment, where updrafts are stronger and clouds grow faster, giving rise to a faster recharge.

The distributions over the droplet concentration N (C&R equation with rain generation dominated by autoconverion) cover the range of values we computed from the LES (Figures 7(d) and 2(b)), however the values of the model parameter N tend to be smaller than in the LES. Moreover, we note that the distributions over N tend to shift towards larger values when the typical cloud depth increases, but we would expect the opposite. In the LES, typical cloud depth *decreases* with droplet concentration N because of the washout of aerosols (Figure 2b), but in the C&R equation (autoconversion), cloud depth, as represented by the cloud carrying capacity, increases on average with droplet concentration. This issue further causes the droplet concentration of the C&R equation (autoconversion) to correlate negatively with the droplet concentrations of the LES. Physically meaningful changes in the posterior distributions over model parameters within the suite of LES, however, would require that an increase in the (average) droplet concentration in an LES causes a larger droplet concentration in the C&R equation.

The distributions over the scaling factor β of the C&R equation with rain generation dominated by accretion (Figure 7(h)) tend to cover a very large range of values, especially if β itself is large. This happens for LES with thinner clouds with a smaller typical cloud depth. For deeper clouds (larger typical cloud depth), β is somewhat within the range of the nominal value ($0.7\,\mathrm{m\,day^{-1}\,km^{-4}}$). The fact that β tends to decrease with increasing cloud depth is the opposite of what we would expect based on the LES, where deeper clouds generate rain more efficiently (see Figure 2). The scaling factor β controls the sink term in the C&R equation and, hence, a larger β corresponds to more efficient rain generation. We would thus expect that β should increase with typical cloud depth, but our inversions suggest the opposite.

https://doi.org/10.5194/egusphere-2026-2871
Preprint. Discussion started: 18 June 2026

6.3 Synthesis

Our feature-based inversions are successful in the sense that we can always find model parameters that lead to C&R equations that exhibit limit cycles similar to the cycles of cloud growth and decay of an LES. The fact that the C&R equation can quantitatively emulate limit cycles across a wide variety of individual LES runs (Figures A1 and A2) confirms the robustness of the predator-prey analogy for stratocumulus dynamics.

When we consider all inversions jointly, however, the variations and shifts in the posterior distributions over the model parameters are partially nonphysical. For both C&R equations, our results suggest that the recharge time scale *decreases* for deeper clouds, which is the opposite of what happens physically. For the C&R equation with rain generation dominated by autoconversion, we find that cloud carrying capacity correlates positively with droplet concentration, while the suite of LES is characterized by the opposite trend (deeper clouds have smaller droplet concentrations (i.e., larger droplets), Figure 2(a) and (b)). For the C&R equation with rain generation dominated by accretion, we find that rain generation becomes less efficient for deeper clouds, which is also the opposite of what happens in the LES (Figure 2(a)).

The nonphysical variation of the posterior distributions over C&R equation parameters can occur due the following reasons. First, the C&R equations are parsimonious, conceptual models with few parameters – both variants of the C&R equation have at most four free parameters. It may not be possible to fully capture the variation of meteorological conditions in the suite of LES with this model. If the model is indeed structurally deficient, then model parameters tend to "compensate" for model error during an inversion. As a result, the model parameters may need to be interpreted as "effective parameters" that depend on the large-scale meteorological forcing. Second, we use simplified assumptions about the rain rate and assume that the rain rate is either dominated by autoconversion (sensitive to N) or accretion (insensitive to N). The actual physics modeled by the suite of LES is more complex, with precipitation initiated by autoconversion but later dominated by accretion. Lastly, the inversion setup may be flawed. While we detect variation in the LES features, the associated errors are also large (large error covariance matrix R). We follow Lunderman et al. (2020) to construct the observation error covariance directly from the variations in the LES cycles, meaning we are reasonably capturing all the various cycle variations in each LES. However, these variations may overpower the variations in the LES cycles due to the changing meteorology and, then, the inversion cannot detect the variations in the meteorological conditions. A robust investigation of our inversion approach was also done in Lunderman et al. (2020), but it is possible that a different approach to inversion *can* incorporate the cloud physics across LES that is currently not being captured by the parameter variations.

What we encounter here may be interpreted as an occurrence of "Simpson's paradox" (Simpson, 1951). Broadly, Simpson's paradox is a statistical phenomenon where trends appear within groups but disappear or reverse when the groups are considered jointly. In the case of inverting a suite of LES for parameters of C&R equations, we find that, for each inversion, deeper clouds (larger H_0) take longer to recharge (larger τ), see Figures A1 and A2. Considering the 14 inversions (for each C&R equation) jointly, however, reverses this trend, i.e., distributions over τ tend to smaller values as H_0 (or H_{75} in the LES) increases. Similarly, we find for each inversion with the C&R equation with rain generation dominated by autoconversion that cloud carrying capacity H_0 correlates positively with droplet concentration, but the trend reverses when the inversions are considered

https://doi.org/10.5194/egusphere-2026-2871
Preprint. Discussion started: 18 June 2026

jointly across the suite of 14 LES (with similar results for the C&R equation with rain generation dominated by accretion, see just above).

7 Conclusions

Summary and conclusion We consider the nonlinear cloud and rain (C&R) equation for precipitating open-cell stratocumulus clouds (Koren et al., 2017). The model was originally proposed as a phenomenological model but it has since also been used for quantitative modeling of selected aspects of a large eddy simulation (LES) (Lunderman et al., 2020).

We explore the potential of the C&R equation as a quantitative tool for understanding precipitating open-cell stratocumulus for a range of conditions. To that extent, we consider two variants of the C&R equation. In one variant, rain generation is dominated by autoconversion (small cloud droplets form larger rain droplets through collisions and coalescence). In the other variant, rain generation is dominated by accretion (larger rain drops grow by collecting smaller cloud droplets as the rain drops fall through the cloud). Along with the C&R equations, we consider a suite of 14 LES of precipitating open-cell stratocumulus cloud systems for which (i) deeper clouds generate rain more efficiently; (ii) larger droplet concentrations lead to less efficient rain processes; and (iii) the microphysical delay does not vary much across all 14 LES. These 14 LES are a carefully selected subset of the LES described by Hoffmann et al. (2023).

We set up and perform a feature-based inversion that is capable of inverting a suite of 14 LES to estimate model parameters of both C&R equations. The feature-based inversions show that both C&R equations are capable of self-consistently producing limit cycles that match the cycles of cloud growth and decay observed in the 14 LES.

When we consider the 14 LES and corresponding C&R model parameters jointly, however, the results are less consistent. For both C&R equations and all 14 LES, we find that the delay of the C&R equations does not vary much across the 14 LES, which we also observe when we estimate the delay directly from the LES. Moreover, the cloud carrying capacity parameter of the C&R equation correlates well with a typical cloud depth of the 14 LES. These two aspects of the "across-LES" inversions are physically meaningful and self-consistent. The cloud carrying capacity parameter however correlates positively with droplet concentration, while the suite of LES is characterized by the opposite trend (deeper clouds have larger droplets). Moreover, a recharge time scale parameter of the C&R equations decreases for deeper clouds, while deeper clouds typically take longer to recharge.

The nonphysical behavior of some of the model parameters may be an instance of Simpson's paradox, where we observe a trend in subgroups, but the trend may disappear or reverse if the subgroups are considered jointly. Here, the subgroups are the 14 LES we invert – taken separately, the C&R equations emulate cycles of cloud growth and decay within the LES well and the model parameters vary self-consistently and meaningfully. When we consider all 14 LES and implied variations in model parameters jointly, however, some physically meaningful trends disappear.

The nonphysical variations may be due to structural inadequacies of the C&R equations, the inversion process, or both. In light of imperfections in the model *and* the inversion process, one may view the C&R equation parameters as "effective parameters," rather than physical constants. Effective here means that the parameters absorb structural model errors: The

https://doi.org/10.5194/egusphere-2026-2871
Preprint. Discussion started: 18 June 2026

inversions prioritize matching the data, but they do not account for the physical meaning of the model parameters. Instead, the inversions essentially use the model parameters as "knobs to turn" to match the data. For example, the C&R equations may not be able to faithfully represent a complex dry air entrainment feedback, but the inversion can make up for it by increasing an "effective" recharge time to force the model to match the data, even if physical intuition suggests that the recharge time should decrease. Interpreting the model parameters as effective parameters also connects to Simpson's Paradox: The effective parameters retain physical and intuitive behavior for any single LES (subgroup) because the model error is constant. Model error, however, is likely not constant across the suite of LES and different errors across the various LES can explain the counter-intuitive evolution of some of the model parameters. Consequently, the effective nature of the model parameters implies that their values are conditioned not only on the microphysics but also on the meteorological conditions. The emergence of Simpson's paradox in our results highlights that while effective parameters can faithfully emulate cycles of cloud growth and decay in specified atmospheric conditions, they may lack the dynamical completeness required to maintain physical scaling across different atmospheric conditions.

Code availability. The software and post-processed LES data used to generate the results in this paper are uploaded to both Github and Zenodo (Gjini, 2026).

Appendix A: Additional figures

https://doi.org/10.5194/egusphere-2026-2871
Preprint. Discussion started: 18 June 2026

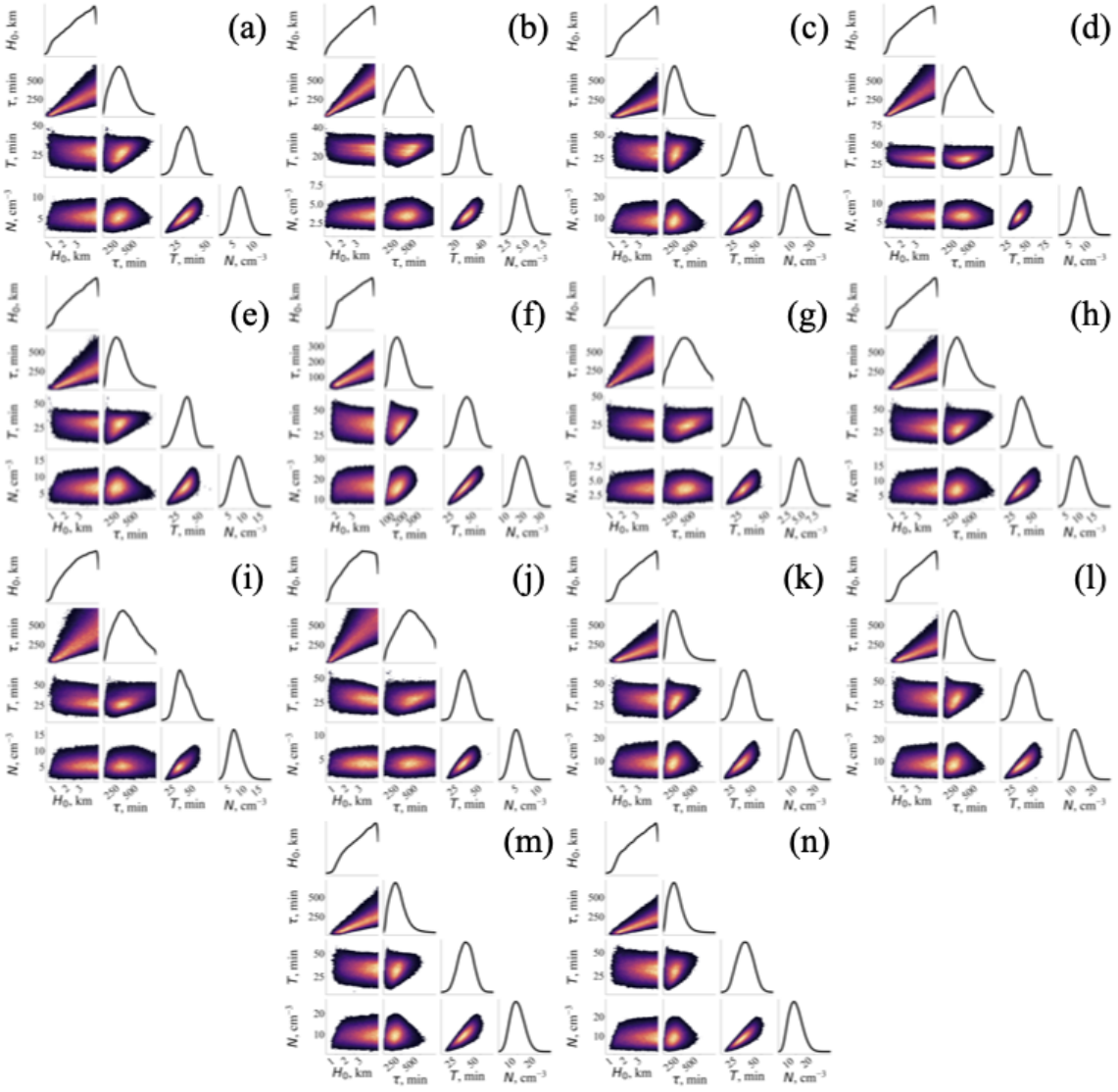

Figure A1. Feature-based inversion results for all 14 LES with the C&R equation equation with rain generation dominated by autoconversion. (a)-(n) Triangle plots of the feature-based posteriors (based on 1.5×10^6 MCMC samples). Each parameter's histogram is plotted on the diagonals and the lower triangle contains histograms of every combination of any two parameters. Warm colors (yellow and pink) represent regions of high posterior probability and cold colors (black and purple) represent lower posterior probability. White regions contain few or no samples and, therefore, correspond to regions of very low posterior probability.

https://doi.org/10.5194/egusphere-2026-2871
Preprint. Discussion started: 18 June 2026

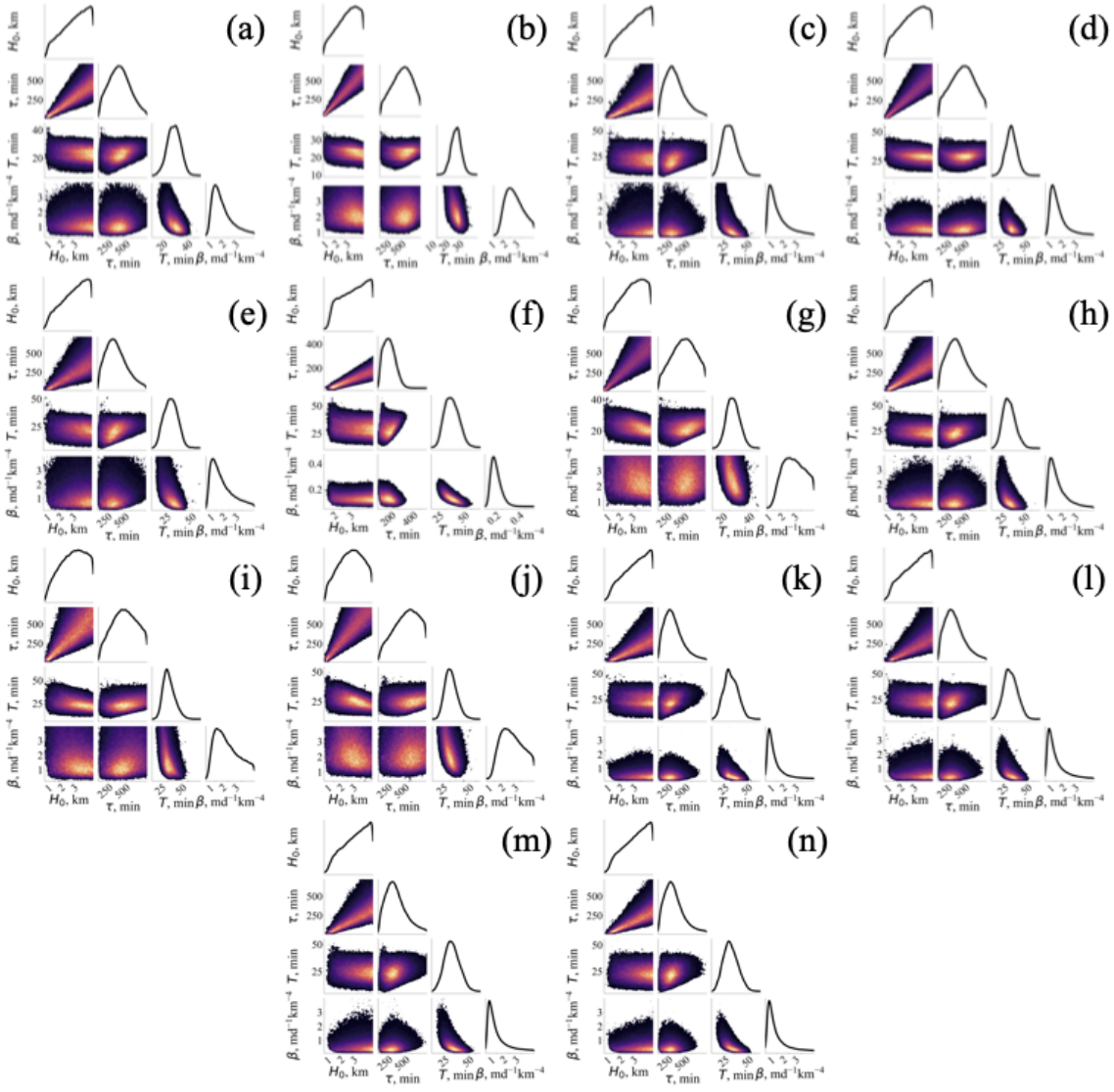

Figure A2. Feature-based inversion results for all 14 LES with the C&R equation with rain generation dominated by accretion. (a)-(n) Triangle plots of the feature-based posteriors (based on 1.5×10^6 MCMC samples). Each parameter's histogram is plotted on the diagonals and the lower triangle contains histograms of every combination of any two parameters. Warm colors (yellow and pink) represent regions of high posterior probability and cold colors (black and purple) represent lower posterior probability. White regions contain few or no samples and, therefore, correspond to regions of very low posterior probability.

https://doi.org/10.5194/egusphere-2026-2871
Preprint. Discussion started: 18 June 2026

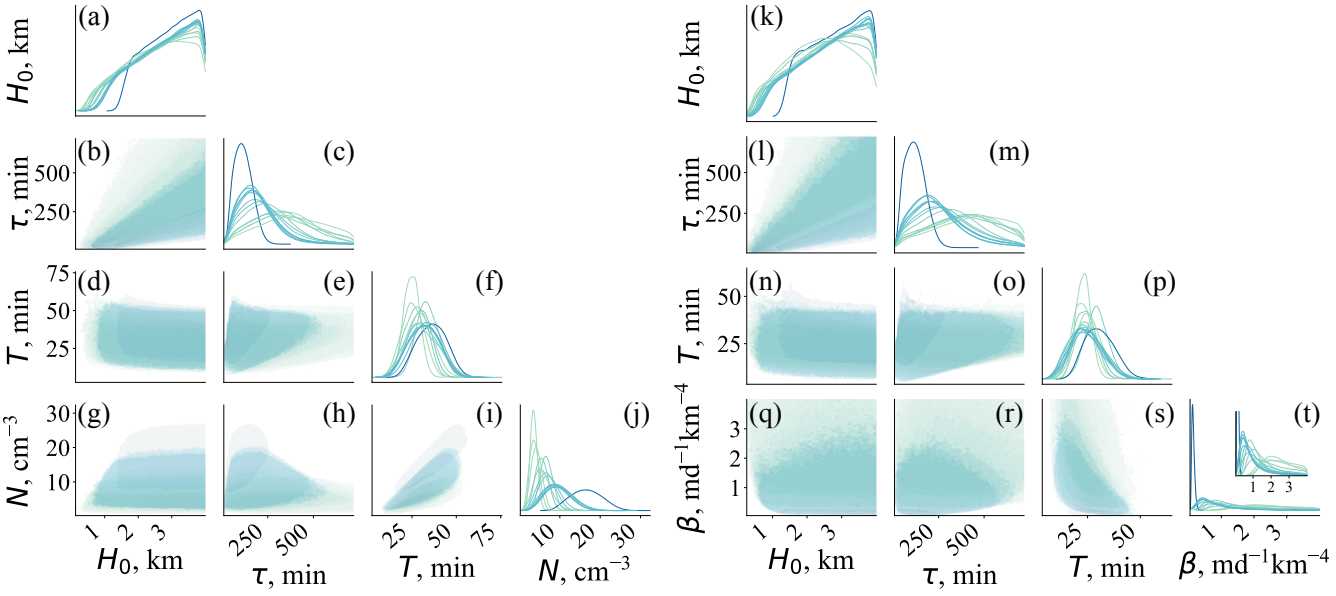

Figure A3. Triangle plots of all 28 feature-based posteriors. (a)-(j) 14 posterior distributions for the C&R equation with rain generation dominated by autoconversion. (k)-(t) 14 posterior distributions for the C&R equation with rain generation dominated by accretion. The color transitions from a light green to a dark blue as the typical cloud depth of the LES (H_{75}) increases. The transparency allows us to identify overlap among the posterior distributions. The filled circles correspond to the posterior means. (Note: (t) - Inset with truncated y-axis included in top-right corner.)

https://doi.org/10.5194/egusphere-2026-2871
Preprint. Discussion started: 18 June 2026

Author contributions. MM conceptualized the project. RG, MM, FG, and GF curated the data. RG, MM, FG, and GF analyzed the results. MM acquired funding. RG, MM, FG, and GF conducted the investigation. RG, MM, FG, and GF developed the methodology. MM, FG, and GF managed and coordinated responsibilities. MM, FG, and GF provided resources for this study. RG and MM developed the software. RG, MM, FG, and GF validated and created visualizations for the results. RG and MM wrote the manuscript draft; FG and GF reviewed and edited the manuscript.

Competing interests. At least one of the (co-)authors is a member of the editorial board of Atmospheric Chemistry and Physics.

Acknowledgements. RG and MM are supported by the US Office of Naval Research (ONR), grant N00014-21-1-2309.

https://doi.org/10.5194/egusphere-2026-2871
Preprint. Discussion started: 18 June 2026

References

Agee, E. M.: Observations from space and thermal convection: A historical perspective, Bulletin of the American Meteorological Society, 65, 938–949, 1984.

Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., Kerminen, V.-M., Kondo, Y., Liao, H., Lohmann, U., et al.: Clouds and aerosols, in: Climate change 2013: The physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change, pp. 571–657, Cambridge University Press, 2013.

Chekroun, M. D., Koren, L., and Liu, H.: Efficient reduction for diagnosing Hopf bifurcation in delay differential systems: Applications to cloud-rain models, Chaos: An Interdisciplinary Journal of Nonlinear Science, 30, 053 130, https://doi.org/10.1063/5.0004697, 2020.

Chekroun, M. D., Koren, I., Liu, H., and Liu, H.: Generic generation of noise-driven chaos in stochastic time delay systems: Bridging the gap with high-end simulations, Science Advances, 8, eabq7137, 2022.

Davis, J. C.: Statistics and Data Analysis in Geology, John Wiley & Sons, Inc., 2002.

Feingold, G. and Koren, I.: A model of coupled oscillators applied to the aerosol–cloud–precipitation system, Nonlinear Processes in Geophysics, 20, 1011–1021, 2013.

Foreman-Mackey, D., Hogg, D. W., Lang, D., and Goodman, J.: Emcee: the MCMC hammer, Publications of the Astronomical Society of the Pacific, 125, 306, 2013.

Forster, P., Ramaswamy, V., Artaxo, P., Berntsen, T., Betts, R., Fahey, D., Haywood, J., Lean, J., Lowe, D., Myhre, G., Nganga, J., Prinn, R., Raga, G., Schulz, M., and Dorland, R. V.: Changes in Atmospheric Constituents and in Radiative Forcing, in: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, solomon, S. and D. Qin and M. Manning and Z. Chen and M. Marquis and K.B. Averyt and M.Tignor and H.L. Miller (eds.), 2007.

Geoffroy, O., Brenguier, J.-L., and Sandu, I.: Relationship between drizzle rate, liquid water path and droplet concentration at the scale of a stratocumulus cloud system, Atmospheric Chemistry and Physics, 8, 4641–4654, https://doi.org/10.5194/acp-8-4641-2008, 2008.

Gjini, R.: rgjini/nonlinear-CandR-inversion: v0.1, https://doi.org/10.5281/zenodo.20263104, 2026.

Glassmeier, F. and Feingold, G.: Network approach to patterns in stratocumulus clouds, Proceedings of the National Academy of Sciences, 114, 10 578–10 583, 2017.

Goodman, J. and Weare, J.: Ensemble samplers with affine invariance, Communications in Applied Mathematics and Computational Science, 5, 65–80, 2010.

Haario, H., Kalachev, L., and Hakkarainen, J.: Generalized correlation integral vectors: A distance concept for chaotic dynamical systems, Chaos: An Interdisciplinary Journal of Nonlinear Science, 25, 063 102, https://doi.org/10.1063/1.4921939, 2015.

Hakkarainen, J., Ilin, A., Solonen, A., Laine, M., Haario, H., Tamminen, J., Oja, E., and Järvinen, H.: On closure parameter estimation in chaotic systems, Nonlinear Processes in Geophysics, 19, 127–143, https://doi.org/10.5194/npg-19-127-2012, 2012.

Hartmann, D. L. and Doelling, D. R.: On the net radiative effectiveness of clouds, Journal of Geophysical Research, 96, 869–891, 1991.

Hoffmann, F., Glassmeier, F., Yamaguchi, T., and Feingold, G.: On the Roles of Precipitation and Entrainment in Stratocumulus Transitions between Mesoscale States, Journal of the Atmospheric Sciences, 80, 2791 – 2803, https://doi.org/10.1175/JAS-D-22-0268.1, 2023.

IPCC: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, vol. In Press, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, https://doi.org/10.1017/9781009157896, 2021.

https://doi.org/10.5194/egusphere-2026-2871
Preprint. Discussion started: 18 June 2026

Kazarnikov, A. and Haario, H.: Statistical approach for parameter identification by Turing patterns, Journal of Theoretical Biology, 501, 110 319, 2020.

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

Koren, I., Tziperman, E., and Feingold, G.: Exploring the nonlinear cloud and rain equation, Chaos: An Interdisciplinary Journal of Nonlinear Science, 27, 013 107, 2017.

Liu, H., Koren, I., Altaratz, O., and Chekroun, M. D.: Opposing trends of cloud coverage over land and ocean under global warming, Atmospheric Chemistry and Physics, 23, 6559–6569, https://doi.org/10.5194/acp-23-6559-2023, 2023.

Lunderman, S., Morzfeld, M., Glassmeier, F., and Feingold, G.: Estimating parameters of the nonlinear cloud and rain equation from a large-eddy simulation, Physica D: Nonlinear Phenomena, 410, 132 500, 2020.

L'Ecuyer, T. S., Hang, Y., Matus, A. V., and Wang, Z.: Reassessing the Effect of Cloud Type on Earth's Energy Balance in the Age of Active Spaceborne Observations. Part I: Top of Atmosphere and Surface, Journal of Climate, 32, 6197 – 6217, 2019.

Maclean, J., Santitissadeekorn, N., and Jones, C. K.: A coherent structure approach for parameter estimation in Lagrangian Data Assimilation, Physica D: Nonlinear Phenomena, 360, 36–45, 2017.

Maraia, R., Springer, S., Haario, H., Hakkarainen, J., and Saksman, E.: PARAMETER ESTIMATION OF STOCHASTIC CHAOTIC SYSTEMS, International Journal for Uncertainty Quantification, 11, 49–62, 2021.

Morzfeld, M., Adams, J., Lunderman, S., and Orozco, R.: Feature-based data assimilation in geophysics, Nonlinear Processes in Geophysics, 25, 355–374, 2018.

Sanz-Alonso, D., Stuart, A. M., and Taeb, A.: Inverse Problems and Data Assimilation, Cambridge University Press, 2023.

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, Nature Climate Change, 7, 3–5, 2017.

Sherwood, S. C., Webb, M. J., Annan, J. D., Armour, K. C., Forster, P. M., Hargreaves, J. C., Hegerl, G., Klein, S. A., Marvel, K. D., Rohling, E. J., Watanabe, M., Andrews, T., Braconnot, P., Bretherton, C. S., Foster, G. L., Hausfather, Z., von der Heydt, A. S., Knutti, R., Mauritsen, T., Norris, J. R., Proistosescu, C., Rugenstein, M., Schmidt, G. A., Tokarska, K. B., and Zelinka, M. D.: An Assessment of Earth's Climate Sensitivity Using Multiple Lines of Evidence, Reviews of geophysics (Washington, D.C. : 1985), 58, https://doi.org/https://doi.org/10.1029/2019RG000678, 2020.

Simpson, E. H.: The Interpretation of Interaction in Contingency Tables, Journal of the Royal Statistical Society: Series B (Methodological), 13, 238–241, https://doi.org/10.1111/j.2517-6161.1951.tb00088.x, 1951.

Springer, S., Haario, H., Shemyakin, V., Kalachev, L., and Shchepakin, D.: Robust parameter estimation of chaotic systems, Inverse Problems and Imaging, 13, 1189–1212, https://doi.org/10.3934/ipi.2019053, 2019.

Springer, S., Haario, H., Susiluoto, J., Bibov, A., Davis, A., and Marzouk, Y.: Efficient Bayesian inference for large chaotic dynamical systems, Geoscientific Model Development, 14, 4319–4333, https://doi.org/10.5194/gmd-14-4319-2021, 2021.

Srivastava, A., Kang, W., and Tartakovsky, D. M.: Feature-informed data assimilation, Journal of Computational Physics, 494, 112 499, https://doi.org/https://doi.org/10.1016/j.jcp.2023.112499, 2023.

Stevens, B., Vali, G., Comstock, K., Wood, R., van Zanten, M. C., Austin, P. H., Bretherton, C. S., and Lenschow, D. H.: POCKETS OF OPEN CELLS AND DRIZZLE IN MARINE STRATOCUMULUS, Bulletin of the American Meteorological Society, 86, 51 – 58, https://doi.org/10.1175/BAMS-86-1-51, 2005.

https://doi.org/10.5194/egusphere-2026-2871
Preprint. Discussion started: 18 June 2026

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, https://doi.org/10.1175/2009JAS3022.1, 2009.

Wood, R.: Relationships between optical depth, liquid water path, droplet concentration, and effective radius in adiabatic layer cloud, University of Washington, 3, 2006.

Wood, R.: Stratocumulus Clouds, Monthly Weather Review, 140, 2373 – 2423, 2012.

Wood, R. and Hartmann, D. L.: Spatial variability of liquid water path in marine low cloud: The importance of mesoscale cellular convection, Journal of Climate, 19, 1748–1764, 2006.

Yamaguchi, T., Feingold, G., and Kazil, J.: Stratocumulus to Cumulus Transition by Drizzle, Journal of Advances in Modeling Earth Systems, 9, 2333–2349, https://doi.org/https://doi.org/10.1002/2017MS001104, 2017.