Capturing Aerosol-Cloud-Precipitation Interactions

A Physics-Informed Sparse Regression Approach for a Coupled Multiscale System With Time Delay

Cheng, Meiling; Glassmeier, Franziska

DOI
10.1029/2024JD043226

Publication date
2025

Document Version
Final published version

Published in
Journal of Geophysical Research: Atmospheres

Citation (APA)
Cheng, M., & Glassmeier, F. (2025). Capturing Aerosol-Cloud-Precipitation Interactions: A Physics-Informed Sparse Regression Approach for a Coupled Multiscale System With Time Delay. *Journal of Geophysical Research: Atmospheres*, *130*(12), Article e2024JD043226. https://doi.org/10.1029/2024JD043226

JGR Atmospheres

RESEARCH ARTICLE
10.1029/2024JD043226

Special Collection:
Analyzing Big Data for Understanding Climate Variability, Natural Phenomena and Rapid Environmental Changes

Capturing Aerosol-Cloud-Precipitation Interactions: A Physics-Informed Sparse Regression Approach for a Coupled Multiscale System With Time Delay

Meiling Cheng[1,2] and **Franziska Glassmeier**[1]

[1]Department of Geoscience and Remote Sensing, Delft University of Technology, Delft, Netherlands, [2]Department of Geoscience and Engineering, Delft University of Technology, Delft, Netherlands

Abstract Aerosols exert a net cooling effect on the climate system by reflecting solar radiation, both directly and indirectly through their role in cloud formation, known as aerosol-cloud interactions. The multiscale nature of aerosol-cloud interactions, and especially their mesoscale adjustments and associated challenges for their representation in climate models, makes the aerosol forcing a key uncertainty of climate projections. Here we show that a physics-informed data-driven approach in the form of delay differential equations (DDEs) for coupled cloud-rain dynamics of mesoscale adjustments can combine the interpretability of conceptual models with the quantitative reliability of large-eddy simulations (LESs). Applied to a conceptual model that describes the coupled system as a predator-prey relationship between cloud depth H and cloud droplet number concentration N, the proposed approach faithfully reconstructs the known DDEs when providing information about the microscale physics in the form of an assumed rain-formation function. We further apply our approach to approximate governing DDEs for the complex aerosol-cloud adjustments modeled by LESs. Capturing the governing cloud-rain dynamics as coupled DDEs also requires providing macroscale physics, which translates into separating the rain and nonrain regimes and assumptions about their asymptotic behavior. These governing equations offer a quantitative pathway for predicting the emergent behaviors of aerosol-cloud-precipitation interactions.

Plain Language Summary Aerosol-cloud interactions are a major source of uncertainty in climate predictions. Their complex, multiscale nature makes it difficult to accurately represent them in climate models, limiting our understanding of their role in the climate system. To address this, we propose a machine learning-based approach that simplifies these interactions. Our method uses a physics-informed technique to represent aerosol-cloud-precipitation interactions (ACPIs) as coupled delay differential equations (DDEs). We test this approach with two toy models that represent ACPIs as predator-prey systems, incorporating microscale physics through a simplified rain-formation function. We further apply our method to analyze how clouds respond to aerosol changes using data from large-eddy simulations (LES). The results show our method is effective in reconstructing coupled DDEs and uncovering physical principles from LES data. This work advances our understanding of ACPIs and their role in climate systems.

Key Points:

- We pragmatically adapt data-driven discovery of governing equations for mesoscale aerosol-cloud-precipitation interactions (ACPIs) as a coupled multiscale system with time delay
- Rediscovery of predator-prey-like delay differential equations as a conceptual model for ACPI requires information about microscale rain-formation physics
- Approximation of mesoscale ACPI from large-eddy simulations requires information about macroscale physics in asymptotic behavior in addition to microphysics

Correspondence to:
M. Cheng,
m.cheng-1@tudelft.nl

Citation:
Cheng, M., & Glassmeier, F. (2025). Capturing aerosol-cloud-precipitation interactions: A physics-informed sparse regression approach for a coupled multiscale system with time delay. *Journal of Geophysical Research: Atmospheres*, *130*, e2024JD043226. https://doi.org/10.1029/2024JD043226

Received 31 DEC 2024
Accepted 6 JUN 2025

Author Contributions:
Conceptualization: Meiling Cheng, Franziska Glassmeier
Data curation: Meiling Cheng, Franziska Glassmeier
Funding acquisition: Franziska Glassmeier
Investigation: Meiling Cheng
Methodology: Meiling Cheng

1. Introduction

Aerosol-cloud interactions have been recognized as one of the largest uncertainties in climate projections by Intergovernmental Panel on Climate Change assessments (Forster et al., 2021). Aerosol-cloud interactions are multiscale, nonlinear processes, where aerosol perturbations lead to changes in cloud formation and the microscopic composition of a cloud, which in turn affect rain formation and cloud dissipation. These cloud-scale responses trigger mesoscale adjustments, that is, responses on time scales of several hours and spatial scales of cloud fields (Alinaghi et al., 2024; Glassmeier et al., 2021a; Savic-Jovcic & Stevens, 2008; Seifert et al., 2015; Wang & Feingold, 2009; Yamaguchi et al., 2019). This complexity makes aerosol-cloud interactions a challenge for climate modeling and projections (Bellouin et al., 2020).

Aerosol-cloud interactions have been numerically captured with a broad spectrum of approaches that range from conceptual models on one end (Glassmeier & Feingold, 2017; Koren et al., 2017; Koren & Feingold, 2011) to purely data-driven approaches (Alfonso & Zamora, 2021; Gettelman et al., 2021; Glassmeier et al., 2019, 2021a; Gryspeerdt et al., 2022; Lamb et al., 2024; Silva et al., 2021) on the opposing end. Conceptual models prioritize interpretability, whereas many data-driven approaches amount to black-box modeling, that is, quantitatively

Software: Meiling Cheng
Supervision: Franziska Glassmeier
Validation: Meiling Cheng
Visualization: Meiling Cheng
Writing – original draft: Meiling Cheng
Writing – review & editing: Meiling Cheng, Franziska Glassmeier

optimized. Most classical parameterizations of cloud-rain dynamics influenced by mesoscale adjustments in models (e.g., Ekman et al. (2023); Flossmann and Wobrock (2010); Gao et al. (2016); Gettelman et al. (2015); Terai et al. (2020); Zhang et al. (2021)) can be situated in between these two extremes. For aerosol-cloud-precipitation interactions (ACPIs) in stratocumulus clouds, one notable conceptual approach employed to characterize the ACPI system is found in the work of Koren and Feingold (2011), who proposed a predator-prey model (KF model hereafter) in the form of two delay differential equations (DDEs). By setting key parameters (notably, time delay, cloud depth, and cloud droplet concentration), their DDEs were able to qualitatively replicate the emergent mesoscale behaviors of aerosol-stratocumulus interactions. The KF model distinguished distinct states of the system, including nonprecipitating, weakly precipitating, and strongly precipitating clouds. Using a simplified version of the KF model by Koren et al. (2017) (KTF model hereafter), Lunderman et al. (2020) illustrated how conceptual models are often the starting point for more quantitative descriptions and parameterizations by estimating the parameters of the KTF model. On the other end of ACPI descriptions, Glassmeier et al. (2021a) and Gryspeerdt et al. (2022) provided data-driven descriptions of aerosol-stratocumulus interactions on the mesoscale.

To combine the strengths of both conceptual models and data-driven approaches, physics-informed tools have emerged as a promising alternative. By integrating physical principles into data-driven frameworks, these tools enhance predictive accuracy while preserving interpretability, offering a balanced and robust approach to understanding complex systems. In this context, the sparse identification of nonlinear dynamic (SINDy) framework (Brunton et al., 2016) has received considerable attention in recent years, and variants of SINDy algorithms have successfully discovered many dynamical systems. The SINDy algorithm assumes that only a few terms control a physical system to find the optimal balance between model accuracy and complexity, so that the discovered equations are sparse in the space of possible functions that follow physical laws (Brunton et al., 2016). This assumption contributes to the success of SINDy in determining the underlying governing equations of nonlinear dynamical systems from data. Examples include sparse identification of nonlinear dynamics for model predictive control (Kaiser et al., 2018) and for boundary value problems (Shea et al., 2021), and parallel implicit sparse identification of nonlinear dynamics to identify implicit dynamics and rational nonlinearities (Kaheman et al., 2020). Recently, Sandoz et al. (2023) discovered DDEs from experimental data of gene expressions.

Against this background, it is interesting to explore whether the SINDy approach can discover the DDEs of ACPIs from data. The special challenge of ACPIs lies in that they are governed by delay-coupled interactions between precipitation, clouds, and aerosols. Furthermore, the sensitivity of clouds and precipitation to aerosol perturbations exhibits regime dependency, notably on nonprecipitating and precipitating regimes, as, for example, elucidated by Savic-Jovcic and Stevens (2008), Stevens and Feingold (2009), and Wang and Feingold (2009). Recently, analytical progress in capturing this complexity in DDEs of Koren and Feingold (2011) has been made by means of reduced-order methods (Chekroun et al., 2020, 2022; Liu, Pei, & Qi, 2022; Liu, Pei, Wang, & Qi, 2022). Specifically, the studies of Chekroun et al. (2020, 2022) applied Galerkin-Koornwinder techniques to approximate DDEs as systems of ODEs. This complexity sets ACPI dynamical systems apart from the more general DDE investigations conducted in Sandoz et al. (2023).

In this work, we propose a pragmatic approach to adapt the SINDy framework for the coupled time-delayed nature of ACPI dynamics. It combines the advantages of conceptual models (interpretability), machine learning (efficiency), and detailed cloud models (accuracy). In the remainder of the manuscript, Section 2 describes the proposed coupled time-delay SINDy model, including model architecture and optimization process. We then test the proposed method by discovering the known KTF and KF models from data generated by them (Section 3). Subsequently, we apply the approach to a large ensemble of LESs (Section 4). Discussion and conclusions are provided in Section 5.

2. Adaptation of SINDy Framework to Coupled DDE

Considering a coupled DDE system of the following forms:

$$\frac{dx(t)}{dt} = f(x(t), x(t-T), y(t), y(t-T)), \tag{1}$$

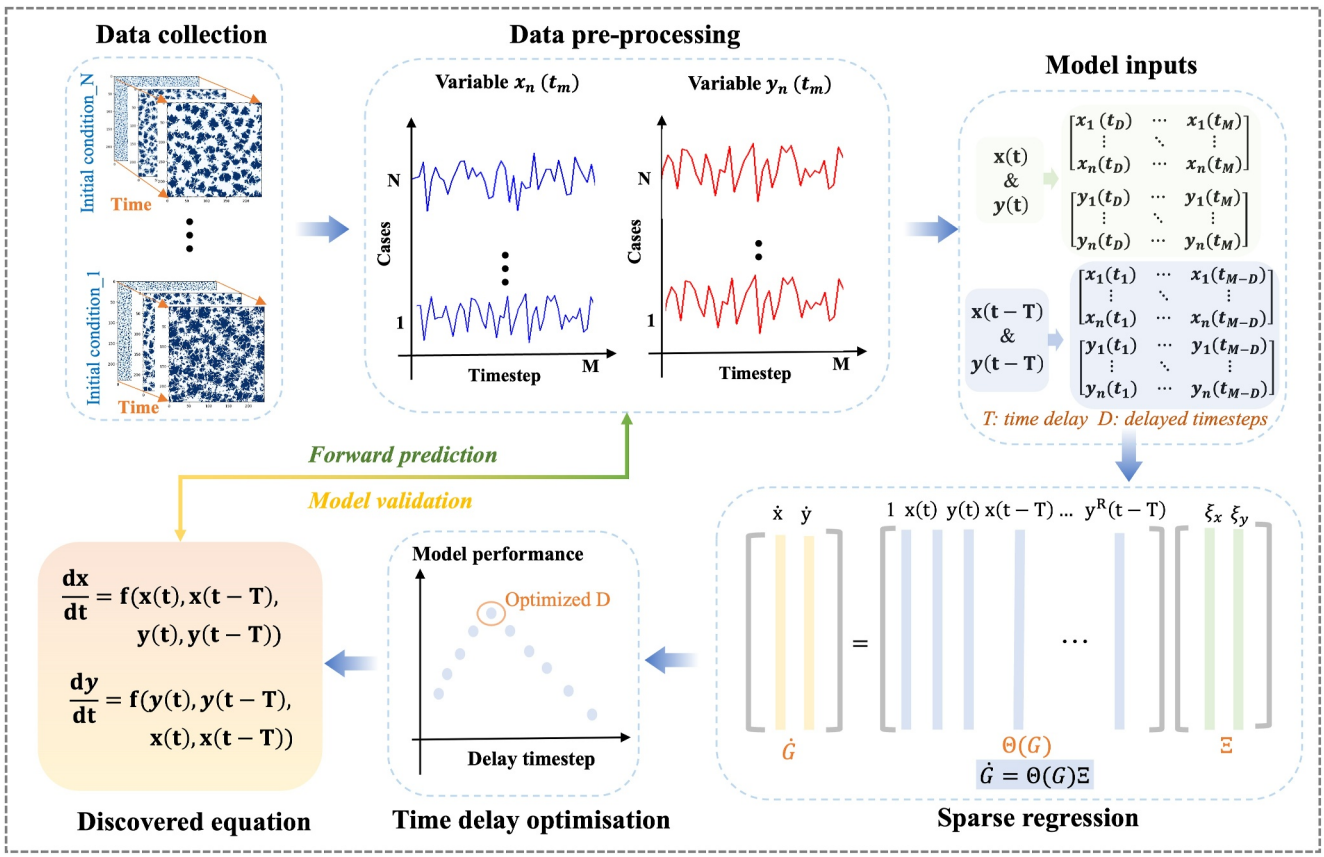

Figure 1. Schematic of the coupled time-delay sparse identification of nonlinear dynamic model to discover the governing delay-differential equations. Data are collected from simulations on different initial conditions, and the spatial average values of variables are derived by preprocessing. The feature library is constructed with variables $x(t)$ and $y(t)$ and the corresponding time-delayed terms $x(t-T)$ and $y(t-T)$. This nonlinear feature library is used to construct the sparse regression model for finding the fewest terms and best time delay to obtain the governing equations.

$$\frac{dy(t)}{dt} = f(y(t), y(t-T), x(t), x(t-T)), \tag{2}$$

where x and y are physical state vectors (i.e., time series data sets), and T denotes the time delay. For simplification, we use G to represent the variables x and y ($G = (x, y)$) so that the total data set can be summarized as follows:

$$\mathbf{G} = \begin{bmatrix} G_1(t_0) & \cdots & G_1(t_M) \\ \vdots & \ddots & \vdots \\ G_n(t_0) & \cdots & G_n(t_M) \end{bmatrix}, \tag{3}$$

where n represents the number of state variables (i.e., the number of cases with different initial conditions for simulations in Figure 1). M denotes the number of time steps, and the time domain is $t \in [t_{-D}, t_{M-D}]$ for G(t). The corresponding time-delayed state vector $G(t-T)$ (where $G(t-T) = (x(t-T), y(t-T))$) is defined for $(t-T) \in [t_{-D}, t_{M-D}]$.

The SINDy method approaches the coupled DDE in Equation 1 and Equation 2 based on the assumption that the dynamic models f could be represented by a set of nonlinear functions Θ weighted by coefficients Ξ as follows:

$$\frac{dG(t)}{dt} = \Theta(G(t), G(t-T))\Xi_G, \tag{4}$$

where $\Theta(G(t), G(t-T))$ denotes candidate nonlinear functions, and Ξ_G is the coefficient ($\Xi_G = \left[\xi_1, \cdots, \xi_n\right]^T$). Considering the library consisting of polynomials up to order R, the library matrix $\Theta(P)$ ($P = (G(t), G(t-T))$) becomes as follows:

$$\Theta(P) = \left[1 \; P \; P^2 \; \cdots \; P^R\right]. \tag{5}$$

2.1. Model Optimization

In Equation 4, coefficients Ξ_G can be solved by regressions. Each entry in Ξ_G corresponds to a term in the discovered dynamical system, while only a few terms in the regression are essential. Thus, it is important to employ sparse regression to identify a sparse matrix of coefficients Ξ_G signifying the fewest active terms from the library that result in a good model fit. The sparse regression problem can be solved by classical regression methods, for example, the sequential thresholded ridge regression (STRidge) method. The STRidge method rapidly converges to a sparse solution in a small number of iterations (Brunton et al., 2016; Rudy et al., 2019). By using the STRidge method, the loss function is defined as follows:

$$\varepsilon = \arg\min_{\Xi_G} \left\| \dot{G} - \Theta(G)\,\Xi_G \right\|_2^2 + \lambda \|\Xi_G\|_2^2, \tag{6}$$

where $\dot{G}$ represents the time derivative of $G(t)$ computed by the finite differentiation method, and λ is a hyperparameter to weight the sparsity constraint. At each iteration, all coefficients whose magnitudes fall below λ are removed, and the model is refit to the remaining terms until only those exceeding the threshold remain. This constraint ensures that the identified model is parsimonious and physically interpretable, which helps to reduce the risk of overfitting the data. As different λ values are likely to produce different coefficient estimates and, ultimately, different discovered models, we adopt the cross-validation method as recommended by Naozuka et al. (2022). The cross-validation method defines a prior set of λ values, computes the cross-validation error for each defined value, and selects the value of λ with the smallest cross-validation error.

As $\Theta(G)$ contains the time-delayed variable $G(t-T)$ in Equation 6, we pragmatically employ an additional optimization procedure to optimize the delay time step D. We consider a range of delay time steps $D \in [0, 1, 2, \cdots]$ and compute the reconstruction error $\varepsilon(D)$ for a fixed D as follows:

$$\varepsilon(D) = \arg\min_{\Xi_G} \left\| \dot{G} - \Theta(G(t), G(t - D\Delta t))\Xi_G \right\| + \lambda \|\Xi_G\|_2^2. \tag{7}$$

The optimal delay time is estimated as the solution of the following equation:

$$D^* = \arg\min_D \varepsilon(D). \tag{8}$$

The model optimization process is trained using the stochastic gradient descent by performing alternative updates of each component as Algorithm 1.

Algorithm 1. Parameter optimization in the coupled time-delay SINDy model.

- Obtain the time series inputs $\mathbf{G} = [G_1(t) \cdots G_n(t)]$.
- Initialize delay time step D and sparsity scale λ.

for D = 1 to N_d (the maximum value of time delay) **do**

  for k = 1 to N_k (iterations for updating the coefficients in cross-validations) **do**

- Obtain the derivatives $\dot{\mathbf{G}}$ of $\mathbf{G}(t)$ by the finite difference method.

Table
Continued

$$\varepsilon(D) = \arg\min_{\Xi_G} \|\dot{\mathbf{G}} - \Theta(\mathbf{G}(t), \mathbf{G}(t - D\Delta t))\Xi_G\| + \lambda\|\Xi_G\|_2^2.$$

end for

- Obtain the optimal hyperparameter λ, and the corresponding model score.

$$\arg\min_D \varepsilon(D).$$

end for

- Obtain the optimal delay time step D, and the corresponding coefficient matrix Ξ_G in Equation 4.

2.2. Model Selection Using Akaike Information Criterion

Accurate identification of the model in Equation 6 depends on selecting an appropriate value for λ. However, the "optimal" λ_a determined from cross-validation in Section 2.1 yields models with differing numbers of terms that exhibit comparable prediction performance. To automate the model selection procedure, we follow Mangan et al. (2017) in utilizing the Akaike information criterion (AIC) as the statistical metric to trade off model fit with sparsity. For least squares fitting, the AIC value for each candidate model $f_{\Xi,\lambda}$, that is, the library functions with coefficients Ξ that remain for sparsity scale λ, is calculated as follows:

$$\mathrm{AIC}(f_{\Xi,\lambda}) = 2k + m \ln \frac{\mathrm{RSS}(f_{\Xi,\lambda})}{m}, \tag{9}$$

where k is the number of free parameters (terms) and m the number of observations. The residual sum of squares is given by $\mathrm{RSS} = \sum_{i=1}^{m}(y_i - f_{\Xi,\lambda}(x_i))^2$, where x_i and y_i represent the independent state and observed state, respectively.

The model selection procedure is shown in Algorithm 2. Both Algorithm 1 and Algorithm 2—adapted from prior work (e.g., Brunton et al. (2016); Rudy et al. (2019); Mangan et al. (2017))—are included with methodological extensions for delay time step optimization and AIC-based model selection, respectively, to support transparency and reproducibility.

Algorithm 2. Model selection in the coupled time-delay SINDy model.

- Obtain the time series inputs $\mathbf{G} = [G_1(t) \cdots G_n(t)]$.
- Obtain the optimal delay time step D from Algorithm 1.
- Define the refined λ scales around the optimal hyperparameter λ_a from Algorithm 1.

for k = 1 to N_λ **do**

- Generate libraries Ξ_k that contain m candidate terms.
- Obtain Model (k) from SINDy $(\dot{\mathbf{G}}, \Xi_k, \lambda_k)$.
- Numerically integrate the DDE system from Model (k) and obtain $f_{\Xi_k,\lambda_k}(t)$.
- Compute AIC(k, $\mathbf{G}(t)$,$f_{\Xi_k,\lambda_k}(t)$).
- Rank models by AIC score.

end for

- Return model with lowest AIC score.

Table 1
The Definition of Variables Used in KTF and KF Equations

Variable	Definition
H	Cloud depth.
H_0	Maximum cloud depth/cloud carrying capacity.
N	Cloud droplet concentration.
N_0	Maximum background aerosol concentration.
α	A scaling constant links cloud depth, droplet concentration, and rain rate.
T	Delayed time converting cloud water into rain by stochastic microphysical collection processes.
τ (τ_1)	Cloud recovery time when H would reach the maximal potential in the absence of rain-related losses.
τ_2	Cloud recovery time when N would reach the maximal potential in the absence of rain-related losses.
c_1	A temperature-dependent constant.
c_2	Constant.
R	Rain rate.

3. Reconstruction of Cloud and Rain Equations

In this section, we evaluate the performance of the proposed approach in reconstructing two known cloud-rain models, which employ a single DDE (KTF model, Koren et al. (2017)) and a set of two DDEs (KF model, Koren and Feingold (2011)) to capture the interplay between cloud and rain dynamics, respectively.

3.1. Reconstruction of KTF Equation

Koren et al. (2017) proposed a DDE that captures the nonlinear effect of depletion by rain on cloud depth H:

$$\frac{dH}{dt} = \frac{H_0 - H}{\tau} - \frac{\alpha}{\sqrt{N}} H^2(t - T), \tag{10}$$

where variables are defined and summarized in Table 1.

The evolution of cloud depth in the KTF equation derives from a predator-prey analogy. The second term on the right-hand side of Equation 10 corresponds to rain as a "predator" that consumes the cloud ("prey") with a delay T that is associated with the time required for cloud droplets to coalesce to form larger raindrops. Equation 10 features a Hopf bifurcation that produces different dynamic regimes (Koren et al., 2017), namely, a steady state where rain consumes cloud at an exact rate, and oscillations when stronger rain depletes the cloud, which dissipates until the cloud thickens enough to reform rain again. In this validation experiment to estimate the nonlinear sensitivity of the time-delay SINDy model with respect to cloud-rain states, three scenarios have been set up to represent the cloud dynamics: (a) steady state, (b) oscillation, and (c) oscillation to steady state, with different initial parameter settings in Table A1. We present the full details of data set generation in the three scenarios in Appendix A1.

Time delay discovery: Figures 2a–2c present the relationship between time delay and model performance, measured by the coefficient of determination $\left(r^2\right)$ in different scenarios. Without time delay ($D = 0$), Equation 10 cannot capture oscillations, which is reflected in the weak correlation ($r^2 < 0.35$) for scenarios 2 and 3 in this case and the magenta line in Figure 3. The optimal delayed time steps, at which the models achieve their strongest correlation ($r^2 > 0.99$), are found to be 50, 250, and 150, which correspond to 5, 25, and 15 min of time delay T ($T = D\Delta t$, $\Delta t = 0.1$minute), respectively, as expected from Table A1. It is noted that the strength of the correlation, as measured by r^2 in the three scenarios, exhibits a pattern of decrease, increase, and then decrease again as the time delay increases. It is important to note that the strong correlation ($r^2 > 0.99$) in Figure 2 indicates agreement between the derivatives of the true and learned models during training, but this does not necessarily

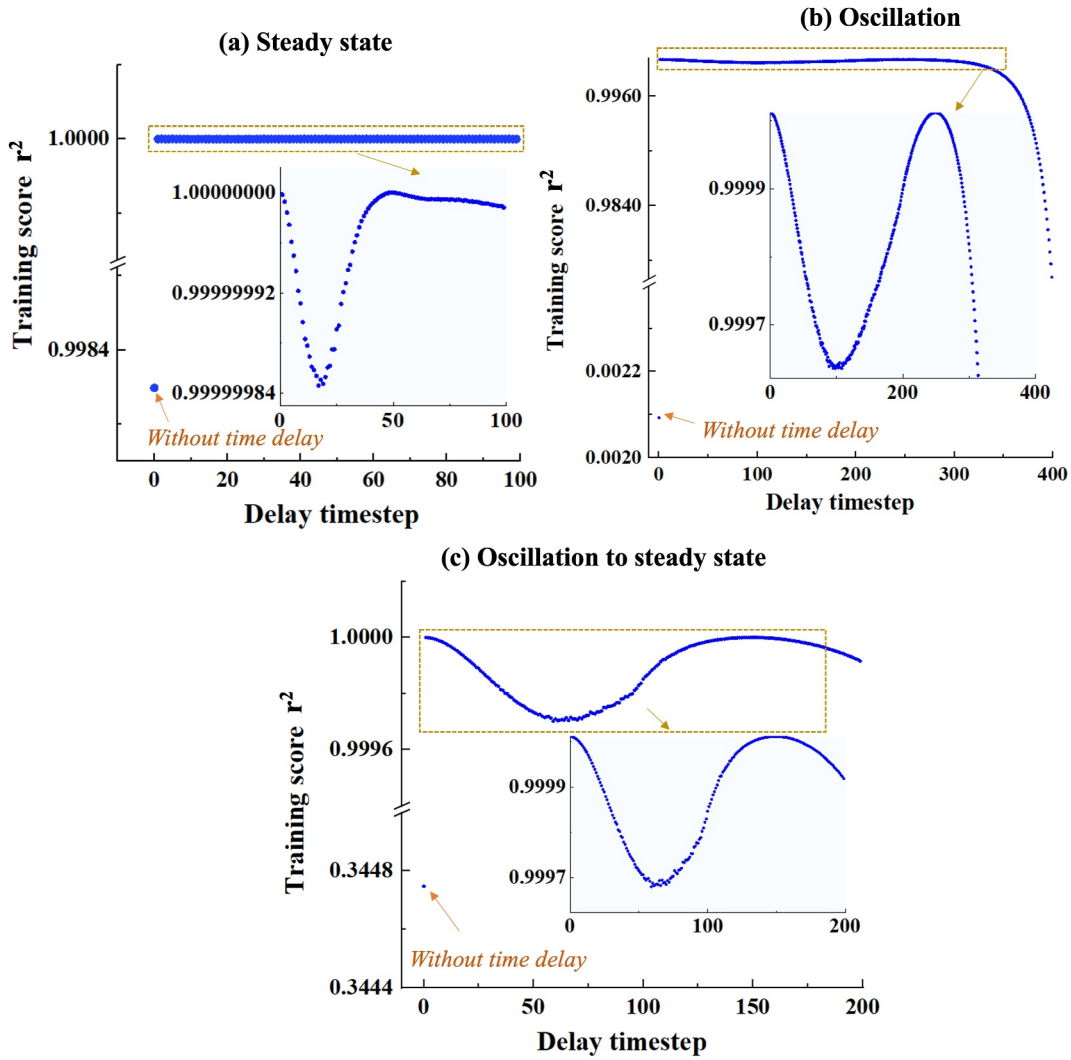

Figure 2. Model training performance with different delay time steps in the three scenarios: (a) steady state, (b) oscillation, and (c) oscillation to steady state.

guarantee accurate forward predictions. Even minor discrepancies in the learned dynamics—despite near-perfect derivative fits—can lead to substantial divergence over time due to the accumulation of integration errors (Brunton et al., 2016). Therefore, while r^2 values in Figure 2 remain high (very close to 1), minor fluctuations of r^2 values reflect variations in the coefficients of the discovered equations due to different delay configurations. These variations, in turn, lead to differences in forward predictions, as illustrated in Figure 3. The model prediction is only apparently correlated with the observations (refer to the orange line with $D = 20$) during the first decrease pattern of r^2 values. As the time delay continues to increase, the models exhibit improved performance, as reflected from $D = 50$ (green line) to 250 (pink line). However, further increasing the time delay leads to a decline in the r^2 values beyond the optimal point, and the predictions fail to capture the dynamics, as shown in the yellow line (time delay $D = 400$) in Figure 3.

Model candidate importance: Table 2 shows the discovered equations with different items in the scenario of oscillation. We choose polynomials to two orders, that is, $\Theta(H) = [1, H(t), H(t-T), H(t)^2, H(t)H(t-T), H(t-T)^2]$ as the library functions. With a relatively small threshold λ in Equation 6, there are many terms in the discovered equations (see the discovered equation with full terms). When increasing the threshold, the discovered terms are reduced (from full terms, 4 terms, to 2 terms). The optimal discovered equation has 4 terms, which achieves the balance between model accuracy and equation complexity to avoid overfitting. In accordance with our

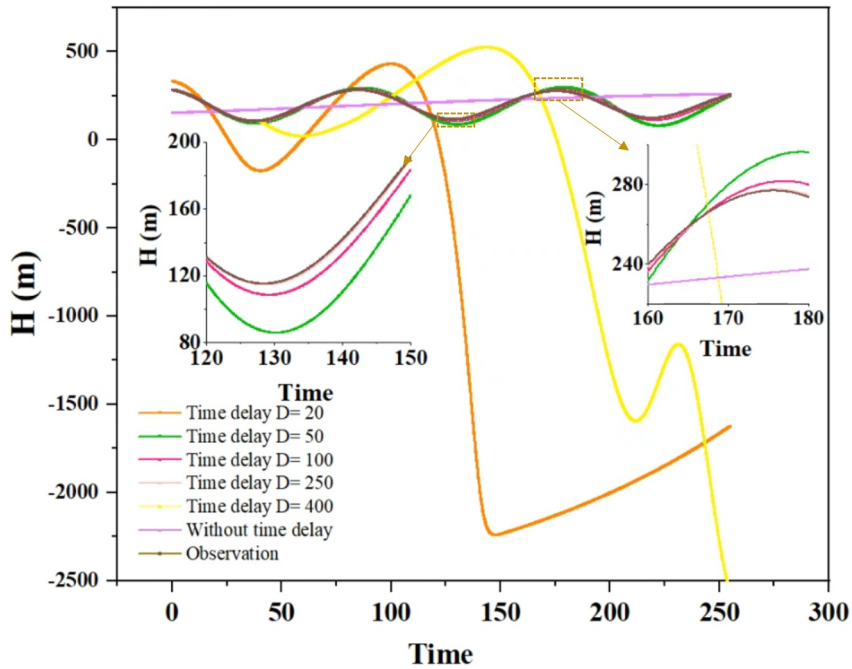

Figure 3. Model predictive performance in an oscillation test case with different time delays of $D = 20, 50, 100, 200, 250, 400,$ and 0 (without time delay).

expectations, we observe that the discovered equation with 3 terms, including $H(t-T)^2$ term (orange line), achieves at least a similar or even better performance in model prediction than that of the discovered model with 4 terms (green line) in Figure 4. It is not easy to set the threshold λ value because the terms $H(t-T)$ and $H(t-T)^2$ both feature coefficients of order 10^{-4}, but the model can better capture high peaks without the $H(t-T)$ term. In addition, in the comparison of 2 and 3 terms with different candidate terms, it is found that although the coefficient of $H(t-T)^2$ is of order 10^{-4}, which is the same as the coefficient of $H(t-T)$ (blue line), its presence is crucial. This finding highlights the significance of small coefficients; however, detecting them proves challenging when employing a singular threshold, λ, and simultaneously disregarding items with similar magnitudes of small coefficients. This aspect warrants future study, for example, using ensemble SINDy to reduce coefficient differences (Fasel et al., 2022).

Overall, the proposed approach can provide an excellent estimation of the optimal delay time and identify the correct items in the library. The summary of discovered equations in three scenarios is shown in Table 3, and model predictive performance is further provided in Appendix A2.

Table 2
Discover Equations and Akaike Information Criterion (AIC) Score for Different Values of λ and the Corresponding Numbers of Terms in the Scenario of Oscillation (With the Optimal Time Delay $T = 25$ Minutes)

	1	H	$H(t-T)$	H^2	$HH(t-T)$	$H(t-T)^2$	
Terms	Coefficients						AIC score
KTF	10	-1.25×10^{-2}	0	0	0	-1.73×10^{-4}	
Full terms	9.95	-1.23×10^{-2}	-3.37×10^{-4}	1.40×10^{-7}	7.50×10^{-7}	-1.73×10^{-4}	
4 terms	9.95	-1.23×10^{-2}	-1.22×10^{-4}	0	0	-1.73×10^{-4}	−154.82
3 terms-1	9.95	-1.23×10^{-2}	0	0	0	-1.73×10^{-4}	−129.29
3 terms-2	1.54	-1.12×10^{-2}	-6.59×10^{-2}	0	0	0	15,213.35
2 terms	0.17	-4.4×10^{-4}	0	0	0	0	25,847.48

Note. Full terms represent the 6 library functions $\Theta(H) = [1, H(t), H(t-T), H(t)^2, H(t)H(t-T), H(t-T)^2])$.

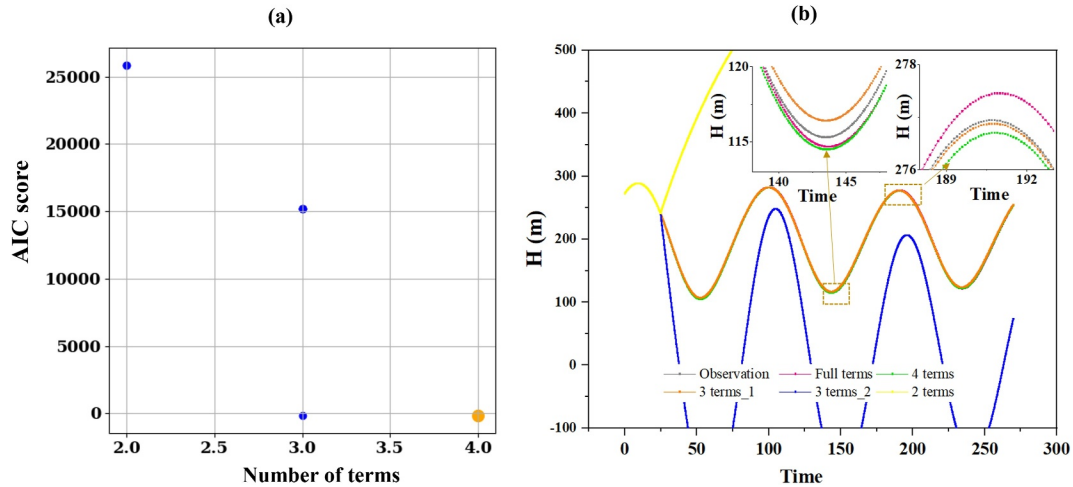

Figure 4. KTF model predictive performance with different terms in the library functions: (a) Akaike information criterion criteria for all possible models (blue dots) and the best one found by sparse identification of nonlinear dynamic (orange dot) shown in Table 2 (red circle). (b) Shows model predictive performance compared to observation values, where the full terms include 6 library functions, 4 terms exclude the terms $H(t)^2$ and $H(t)H(t-T)$, the 3 terms-1 further exclude the $H(t-T)$ terms, whereas the 3 terms-2 exclude the $H(t-T)^2$ term in the discovered equations, and the 2 terms only have the H term and constant term.

3.2. Reconstruction of KF Equations

To further investigate the proposed tool in discovering coupled DDEs, we focus on the KF equations proposed by Koren and Feingold (2011), which use the coupled predator-prey models with delays to mimic the interplay between aerosol, cloud, and rain as follows:

$$\frac{dH}{dt} = \frac{H_0 - H}{\tau_1} - \frac{\alpha H^2(t-T)}{c_1 N(t-T)}, \tag{11}$$

$$\frac{dN}{dt} = \frac{N_0 - N}{\tau_2} - c_2 \alpha H^3(t-T), \tag{12}$$

where the rain rate R is a function of H and N:

$$R(t) = \frac{\alpha H^3(t-T)}{N(t-T)}. \tag{13}$$

We obtain the different dynamic regimes of the cloud-rain system by varying the delay time (see Table 3). To test the performance of the sparse regression part of our approach, it is therefore sufficient to focus on one dynamic regime, for which we choose an oscillation case. Full details of data set generation are given in Appendix B.

Library function setting: We approach the rediscovery of Equation 11 and Equation 12 by considering them as implicit differential equations. This amounts to generalizing the library $\Theta(G)$ in Equation 4 to include basic functions of $((H(t), H(t-T), N(t), N(t-T))$ and $(\dot{H}, \dot{N})$ ($\dot{H}$ denotes $\frac{dH}{dt}$). The implicit DDEs reformulate the

Table 3
Coefficient Comparison Between the Original and Discovered KTF Equations in Three Scenarios

Scenario	Original equation	Discovered equation
Scenario 1	$\dot{H} = 10 - 1.25 \times 10^{-2}H - 1.73 \times 10^{-4}H(t-5)^2$	$\dot{H} = 9.9 - 1.03 \times 10^{-2}H - 1.79 \times 10^{-4}H(t-5)^2$
Scenario 2	$\dot{H} = 10 - 1.25 \times 10^{-2}H - 1.73 \times 10^{-4}H(t-15)^2$	$\dot{H} = 9.9 - 1.24 \times 10^{-2}H - 1.73 \times 10^{-4}H(t-15)^2$
Scenario 3	$\dot{H} = 10 - 1.25 \times 10^{-2}H - 1.73 \times 10^{-4}H(t-25)^2$	$\dot{H} = 10 - 1.24 \times 10^{-2}H - 1.73 \times 10^{-4}H(t-25)^2$

Table 4
Coefficient Comparison in the Original and Discovered KF Equations

Original equation	Discovered equation
$\dot{H} = 8.83 - 1.67 \times 10^{-2}H - 6.96 \times 10^{-4}\frac{H^2(t-25.4)}{N(t-25.4)}$	$\dot{H} = 8.85 - 1.67 \times 10^{-2}H - 6.95 \times 10^{-4}\frac{H^2(t-25.4)}{N(t-25.4)}$
$\dot{N} = 3 - 1.67 \times 10^{-2}N - 4.17 \times 10^{-8}H^3(t-25.4)$	$\dot{N} = 3.01 - 1.67 \times 10^{-2}N - 4.17 \times 10^{-8}H^3(t-25.4)$

SINDy problems in an implicit form (Mangan et al., 2016): $\Theta\left(G(t), G(t-T), \frac{dG(t)}{dt}\right)\Xi_G = 0$ (different from Equation 4). However, it is worth mentioning that finding a matrix Ξ_G whose columns ξ_i multiply the library functions to obtain zero vectors is challenging. As discussed by Wu (2023), the success of sparse identification hinges significantly on the quality of the constructed library functions. In our case, this translates to constraining our library based on physical knowledge. We specifically provide microscale physical knowledge in the form of rain formation relationships by adding $\frac{H^2(t-T)}{N(t-T)}$ and $H^3(t-T)$ as library functions.

Reconstructed model predictive performance: Table 4 shows the reconstructed coefficients in equations, which closely match those of the original KF equations. The relative errors of the reconstructed model over time are further elucidated in Figures 5a and 5b. It can be seen that the relative errors of cloud depth H and cloud droplet concentration N are between -0.1% to 0.15% and between -0.2% to 0.6%, respectively. Notably, these errors exhibit a pattern of being positive around the high peaks and negative around the low peaks of cloud depth (H) or cloud droplet concentration (N). These biases arise from discrepancies between the reconstructed and original coefficients within the coupled DDEs, as illustrated in Table 4. When H and N values exceed a certain value, the predictive derivatives are positive, and then H and N increase (gradually overestimated in high peaks). At some point, the predictive derivatives become negative, thus making H and N decrease (gradually underestimating the low peaks). It is essential to emphasize that the proposed model demonstrates the capability to discern small coefficients, such as the $H(t-T)^3$ term with values on the order of 10^{-8} within the discovered models. While minor errors persist, these promising results underscore the robustness of the proposed coupled time-delay SINDy model in accurately identifying coefficients of coupled DDEs within the KF structure.

4. Discovering ACPI Equations From Large-Eddy Simulations

We have demonstrated the capability of the proposed tool in reconstructing DDEs that capture ACPI on the mesoscale from two idealized models if library functions are constrained by our microscale knowledge about rain formation. In this section, we will demonstrate the versatility of our coupled time-delay SINDy model in approximating a predator-prey-like model derived from large-eddy simulations (LESs) of cloud systems.

4.1. Data Sets

The LESs were carried out using the System for Atmospheric Modeling (Khairoutdinov & Randall, 2003). The model domain area is 48 km × 48 km at a horizontal resolution of 200 m and a vertical resolution of 10 m. The simulation duration is 12 h, with a time step of 1 s. For more details on the model setup, see the study by

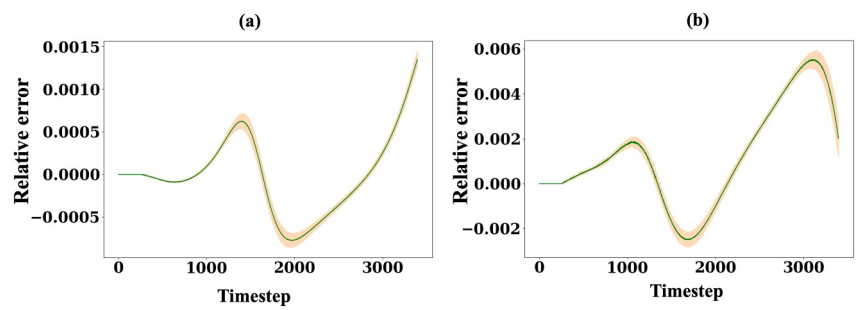

Figure 5. Model predictive performance assessment: (a) and (b) represent the relative errors of cloud depth H and cloud droplet concentration N at different time steps throughout all test cases, respectively, and the shaded regions represent standard deviations.

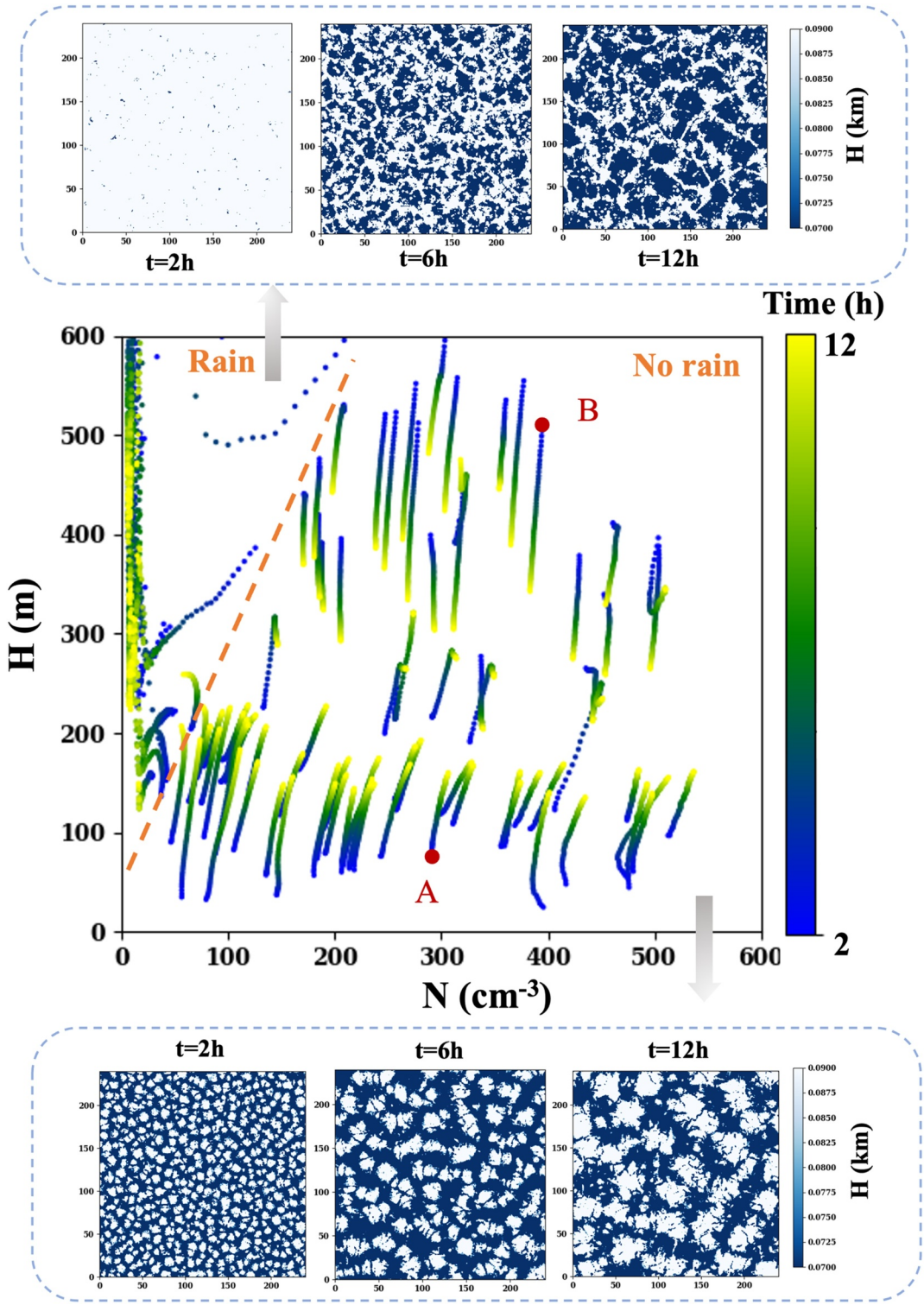

Figure 6. A total of 107 large-eddy simulation simulations with varying initial conditions in an N-H space (H and N are the spatial-averaged values of cloud fields in each trajectory), colored by the temporal scale from 2 to 12 hr. The dashed orange line corresponds to an adiabatic volume-mean droplet radius at the cloud top of about 12 μm, which separates the regimes of no-rain (droplet sizes are too small for precipitation formation) and rain states (droplet size crosses the threshold value for precipitation formation). The top and bottom snapshots illustrate the open-cell and closed-cell cloud depth fields in temporal evolution in no-rain and rain regimes. Red letters A and B indicate trajectories discussed in the main text.

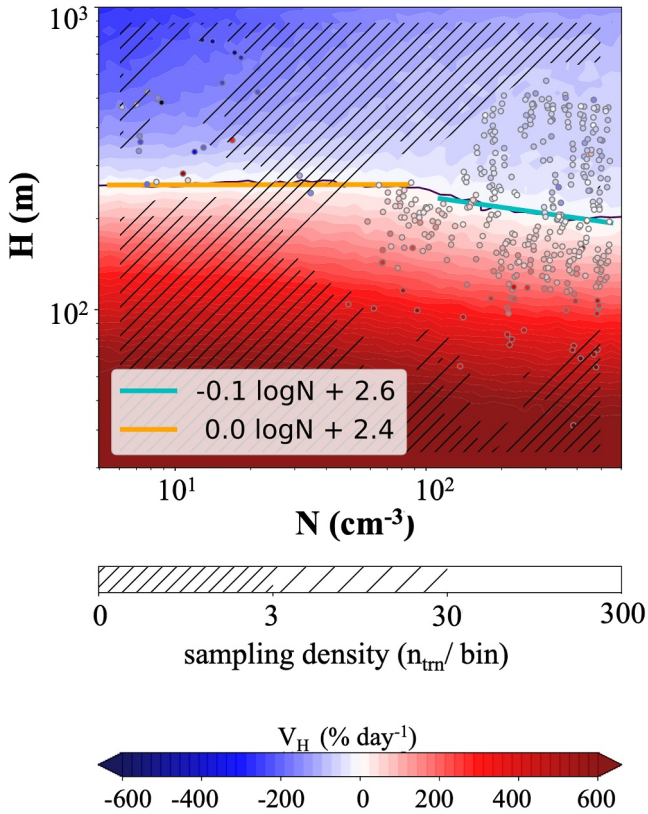

Figure 7. Emulated N-H surface of H-component $V_H = d \ln H/dt$ of the flow field as a function in cloud depth H and cloud droplet concentration N, details of surface interpolation referring to Glassmeier et al. (2019, 2021a). The blue and yellow lines indicate the location of steady-state H in the no-rain and rain regimes, respectively.

Yamaguchi et al. (2017). By varying initial conditions (e.g., aerosol concentration in a well-mixed boundary layer) in the model simulations of Glassmeier et al. (2021a), we obtained a total of 107 simulations (Figure 6). It is worth mentioning that the total of 107 simulations could be separated into two regimes: no-rain and rain, as shown in Figure 6. The separation is based on an adiabatic volume-mean droplet radius at the cloud top of about 12 μm; for details, see the study by Glassmeier et al. (2021a). We remove simulations with cloud base at 5m or lower for a notable part of the time series. After this filtering, 81 simulations remain. We discard the first 2 hr of spin-up and use data sets from 2 to 12 hr, at output intervals of 10 min. In the no-rain regime, there are 77 simulations associated with 60 time steps for model training, whereas for the rain regime, there are only 4 simulations with 30 time steps. During the model validation process, we utilized the data from the initial time delay period as inputs for the derived equations and conducted forward predictions, which were subsequently compared with the remaining time steps of each simulation in both regimes.

4.2. LES-Derived Equations

It is challenging to directly discover the governing equations from LESs, as the LES simulations resolve the complex governing equations to faithfully represent the formation of stratocumulus clouds and their interaction between aerosols and rain. We therefore adapt the predator-prey model to LES data sets. Based on previous studies pointing out the distinction between clouds in no-rain and rain regimes (Chen et al., 2014; Koren & Feingold, 2011; Liu, Pei, Wang, & Qi, 2022; Terai et al., 2020), we approximate the predator-prey-like models in the two separate regimes. Unlike the estimation of parameter value distributions in a KTF model from LES simulations (Lunderman et al., 2020), we approach the discovery of governing equations from LES data sets by approximating a KF framework, which includes the coupled dynamical system between aerosols, precipitation, and clouds. However, it is noted that the parameters in the KF equation shown in Table B1 are unknown (except for $\Delta t = 10$ minutes) for LES simulations. Among these parameters (see Table 1), we assume the typical time scales τ_1, τ_2, and the coeffcients α, c_1, and c_2 to be only dependent on the overall regime such that they are treated as constants within the two regimes, while H_0 and N_0 vary across all simulations. According to the definitions of H_0 and N_0, we make the following assumptions to constrain them:

1. In the no-rain regime, H_0 represents the maximum cloud depth that is reached asymptotically in the absence of a rain sink. This value corresponds to the steady state where cloud deepening due to radiative and surface-flux cooling (exemplified by trajectory A) and cloud thinning due to entrainment drying (exemplified by trajectory B) approximately balance each other in the central region of Figure 6 (Hoffmann et al., 2020). To approximate the steady state in N-H space, we extract N-H tendencies from all LES data sets and then interpolate the tendencies using Gaussian-process emulation (Glassmeier et al., 2019, 2021a) to obtain the tendency surface (Figure 7). We assume a log-log-linear relationship between H and N (blue line) to obtain the steady-state function as follows:

$$\log H_0 = -0.1 \log N_{\text{initial}} + 2.6, \tag{14}$$

when additionally assuming the absence of a tendency in N such that the starting value N_{initial} of the N-evolution is maintained.

2. In the no-rain regime, N_0 is the aerosol concentration obtained asymptotically in response to the aerosol flux from the surface. As the simulations maintain a slow but steady increase in N, we set $N_0 = 1000\ \text{cm}^{-3}$ for all LES simulations.

Table 5
Approximated Coefficients in the Discovered Equations in Two Regimes

No-rain	$\dot{H} = 1.08 \times 10^{-2}$	$H_0 - 8.92 \times 10^{-3}H - 5.05 \times 10^{-4}\frac{H^2(t-270)}{N(t-270)}$
Rain	$\dot{H} = 0.34$	$H_0 + 9.97 \times 10^{-3}H - 5.46 \times 10^{-3}\frac{H^2(t-20)}{N(t-20)}$
No-rain	$\dot{N} = 2.92 \times 10^{-4}$	$N_0 - 5.92 \times 10^{-5}N - 3.82 \times 10^{-9}H^3(t-270)$
No-rain	$\dot{N} = -2.61 \times 10^{-2}$	$N_0 + 1.86 \times 10^{-2}N - 1.42 \times 10^{-7}H^3(t-20)$

3. In the rain regime, where rain consumes aerosols and droplets and dissipates clouds, H_0 and N_0 each balance at small values. We choose $H_0 = 10$ m and $N_0 = 10$ cm^{-3}.

In addition to these asymptotic constraints, we also keep the microscopic constraints (emphasizing that these constraints originate from small-scale physical processes—rain formation, aerosol, and cloud microphysics) implemented through the library function settings in Section 3.2 and obtain the approximated coefficients summarized in Table 5 for the coupled DDEs for the LES data set. For the H-evolution, coefficients in Table 5 clearly reflect the role of rain in the two cases. In the no-rain case, the main sink of H is the H-term in the library, which decreases large H and corresponds to evaporation drying, because the H^2/N-term, which corresponds to rain, has a small coefficient. In the rain case, the H^2/N-term is the dominant and only sink because the H-term has a positive coefficient. The role of rain for the N-evolution is similar, where it is, however, noteworthy that the source term decreases with decreasing N due to the positive coefficient of that term. Physically, this would correspond to a complete rainout and aerosol scavenging. The optimal time delays T are 270 and 20 min for no-rain and rain regimes, respectively. The delay in the rain regime is consistent with the time required for rain formation through collision and coalescence. We interpret the time delay of 270 min in the no-rain regime as a low probability of having an influence of the nonlinear rain term, that is, collision-coalescence rarely affects cloud evolution in this regime. Additionally, note that the rain term is relatively unimportant in the no-rain regime as discussed above.

Figure 8 presents the comparison of the cloud depth H and the cloud droplet concentration N obtained from the LES-derived equations (discovered models) and the original LES data sets. These equations well capture the tendency of the trajectories from LES simulations in the no-rain and rain regimes. In the no-rain regime, N-evolution is negligible compared to the equilibration of H to its steady state (Hoffmann et al., 2024). In contrast, the rain regime features a rapid co-decrease of cloud depth H and cloud droplet concentration N. Figures 9 and 10 present the prediction accuracy of derived equations by correlation coefficients and relative errors, respectively. It can be noted that correlation coefficients of H and N in two regimes are beyond 0.98 in Figure 9, which indicates a good match of model predictions from LES-derived equations with the LES data sets. Furthermore, Figure 10 depicts the relative errors at different time steps in two regimes. It can be seen that in no-rain and rain regimes, the relative errors are between −10% to 10% and between −5% to 5% for cloud depth H, respectively, which for cloud droplet concentration N are between −5% to 5% and between −15% to 10%, respectively. In the no-rain regime, the relative errors are 0 from 0 to 27 time steps because the predictions are the observations in the initial delayed period. After the delayed period, the relative errors of derivatives are gradually increased due to the error accumulation in the predictive period (28 − 60 time steps). We can understand the systematic error that leads to the accumulations as follows: Predicted tendencies are "slower" than the LES data sets in high-H trajectories that evolve toward lower H ("downward" in Figure 8), which is caused by the slight underestimation of model predictive $\dot{H}$ in each time step. Similar but opposite behavior is found for low-H trajectories. Mean relative errors in Figure 10 are close to 0, which indicates the discovered equations capture ACPIs in the LES data set reasonably well. Thus, the proposed coupled time-delay SINDy model presented here is useful for characterizing ACPIs and for deepening our understanding of cloud microphysics on the mesoscale.

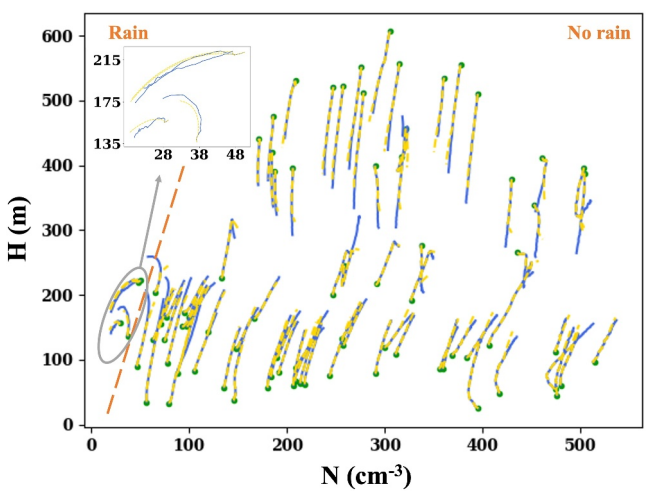

Figure 8. Comparison between the model predictions based on the discovered equations in Table 5 and the large-eddy simulation (LES) data from LES data sets with varying initial conditions in an N-H space. The blue and yellow lines represent LES and predictions, respectively, and the green points are the initial values in each simulation. The dashed orange line mentioned in Figure 6, corresponds to an adiabatic volume-mean droplet radius at the cloud top of about 12 μm, which separates the regimes of no-rain and rain states.

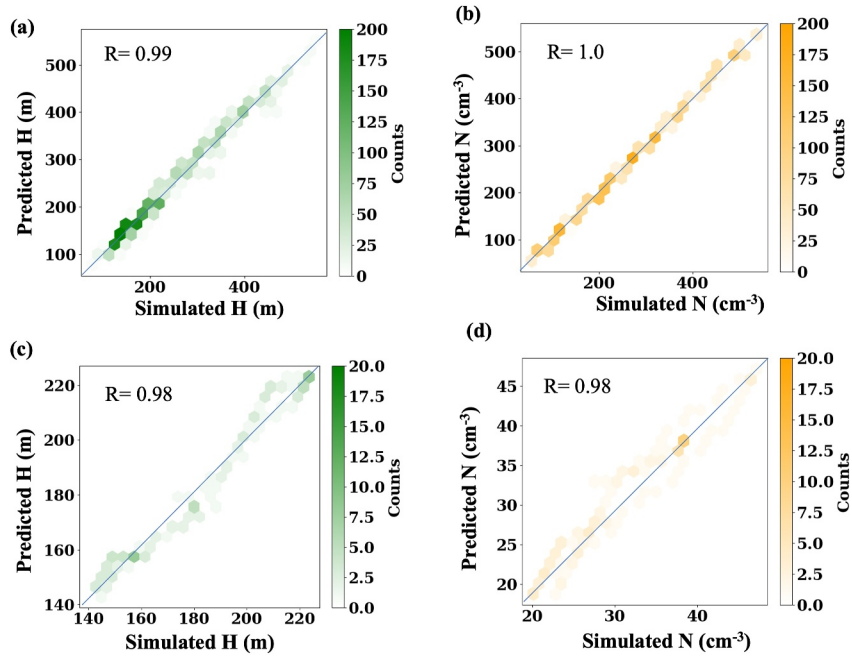

Figure 9. Correlation coefficients between model predictions in Figure 8 and large-eddy simulation data: (a) cloud depth H and (b) cloud droplet concentration N in the no-rain regime, and (c) cloud depth H and (d) cloud droplet concentration N in the rain regime. The points are excluded in the time delay period of 270 and 20 min in the no-rain and rain regimes, respectively.

4.3. Extrapolation and Asymptotic Behavior of Discovered ACPI Equations

For further analysis of the governing equations discovered from the LES data set, we extend the predictions in Figure 8 to 1,000 and 60 time steps in the no-rain and rain regimes, respectively, as shown in Figure 11. In the no-rain regime, it depicts that the cloud droplet concentration N increases gradually, whereas cloud depth H exhibits a steady state. This behavior suggests that high aerosol concentrations suppress droplet growth, leading to more but smaller cloud droplets and preventing precipitation. In contrast, the cloud depth H and cloud droplet concentration N both rapidly decrease to 0 in the rain regime. With low cloud droplet concentration, these droplets efficiently collide and coalesce into raindrops, initiating precipitation and leading to a decrease in cloud depth over time. Although the derived equations in the two regimes here are only indicative of N-H coupled relationships and cannot represent the full complexity of ACPI in the LES, they provide a pathway to predicting the

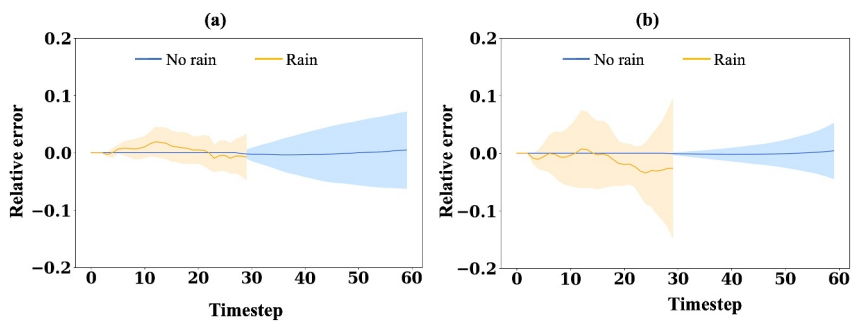

Figure 10. Relative errors between model predictions in Figure 8 and large-eddy simulation data at different time steps of (a) cloud depth H and (b) cloud droplet concentration N in the no-rain and rain regimes. The shaded regions represent standard deviations. In the no-rain regime, there are 77 simulations associated with 60 time steps for model comparison, whereas for the rain regime, there are 4 simulations and 30 time steps. In addition, the relative errors are set to 0 in the time delay period of 270 and 20 min in the no-rain and rain regimes, respectively.

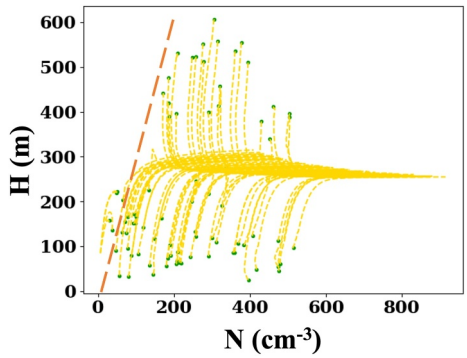

Figure 11. Time series extension of model predictions based on the discovered equations in Table 5 in an *H*-*N* space. The green points are the initial values of trajectories from large-eddy simulation data sets. The dashed orange line mentioned in Figure 6 corresponds to an adiabatic volume-mean droplet radius at the cloud top of about 12 μm, which separates the regimes of no-rain and rain states.

emergent behavior of ACPI on the mesoscale. They especially highlight the roles of different time scales in ACPI, notably the stark difference between the slow equilibration of *N* in the no-rain regime as opposed to its very rapid decline in the rain regime.

5. Discussion and Conclusions

We have presented a physics-informed sparse regression approach for ACPI as a coupled multiscale system with time delay (Section 2). We demonstrate our approach by rediscovering a single DDE (DDE) that conceptually captures ACPI (KTF model, Section 3.1). Application to a coupled set of DDEs for ACPI highlights the need for informing equation discovery by providing knowledge about the physics of rain formation (KF model, Section 3.2). Capturing ACPI as represented by LESs with our approach highlights the multiscale challenge of the system (Section 4) in that we require microscale rain formation as well as information about asymptotic regimes, whose quantitative details reflect the large-scale conditions. The LES-derived equations combine the physical understanding and interpretability of a conceptual model with the quantitative reliability of a full-complexity LES data set. In doing so, they not only provide a parameterization of ACPI, but are also an instructive demonstration of two key differences of conceptual models as compared to purely data-driven approaches. First, functional relationships like Equation 13 would be considered complex if part of a general function library but owe their simplicity to the fact that they have a direct physical interpretation. In other words, conceptual models are simple because they condense the complexity of physical processes into a few clear relationships, not because these relationships are mathematically simple. This can be considered a core challenge for the automatic discovery of conceptual models illustrated by the rain formation term, which cannot be discovered with a standard polynomial SINDy library. Second, conceptual models are characterized by a parametric flexibility. Together with their physics-based justification, this is why we trust that conceptual models can be extrapolated. This point closely mirrors discussions about physics-based versus empirical parameterizations, notably in application to weather forecast and climate projection (Lohmann et al., 2007), which depend on subgrid-scale schemes to represent unresolved small-scale processes due to the coarse resolution. As demonstrated by Koren et al. (2017) for the KTF model, governing equations might scale with such parameters. We had to add such scaling as asymptotic information capturing large-scale physics to the purely data-driven equation discovery.

A key future direction is to leverage LES-derived equations to improve subgrid parameterizations in climate models. LES-derived equations, by capturing essential turbulent dynamics at subgrid scales, offer physically grounded relationships that can help refine these schemes. One challenge lies in the limited diversity, spatial extent, and duration of most LES data sets. LES is typically performed over small domains and short periods, often focused on specific regimes (e.g., stratocumulus or shallow cumulus), limiting the generalizability of derived parameterizations. In contrast, climate models must represent a broad and continuous spectrum of atmospheric states across the globe. Applying LES-informed parameterizations without sufficient regime diversity and contextual representativeness may lead to reduced robustness and performance in out-of-sample conditions. To address this gap, more comprehensive LES ensembles—such as the Botany ensemble (Jansson et al., 2023), which systematically varies large-scale forcings—are needed to capture a wider range of cloud regimes. In turn, hybrid approaches may also help to better interpret observations: Alternative functional forms for the rain rate in Equation 13 are also consistent with observations (Geoffroy et al., 2008). An LES data set with a suitable sampling, as discussed above, might be leveraged to find the optimal formulation.

In conclusion, we find the sparse regression-based discovery of governing equations a useful tool to capture ACPI dynamics on the mesoscale, where we so far lack physical understanding. Equation discovery thus fills the gap between the microscale and the large-scale if informed by our physical understanding of these two neighboring scales.

Appendix A: KTF Equation

A1. Data Generation

For a given set of parameters and initial conditions, we solve Equation 10 numerically by a fourth-order Runge-Kutta method with time step Δt. In each scenario, we create 130 time series by prescribing that the cloud depth H (t) in the initial time delay period ($t \in [-T, 0]$) is constant, where $H(t) \in [50, 180]$ with an interval of 1 (Figure A1). We use 104 (80%) time series in each scenario for time-delay SINDy model training and validation, and the remaining 20% time series for testing. The parameter setting for three scenarios is shown in Table A1.

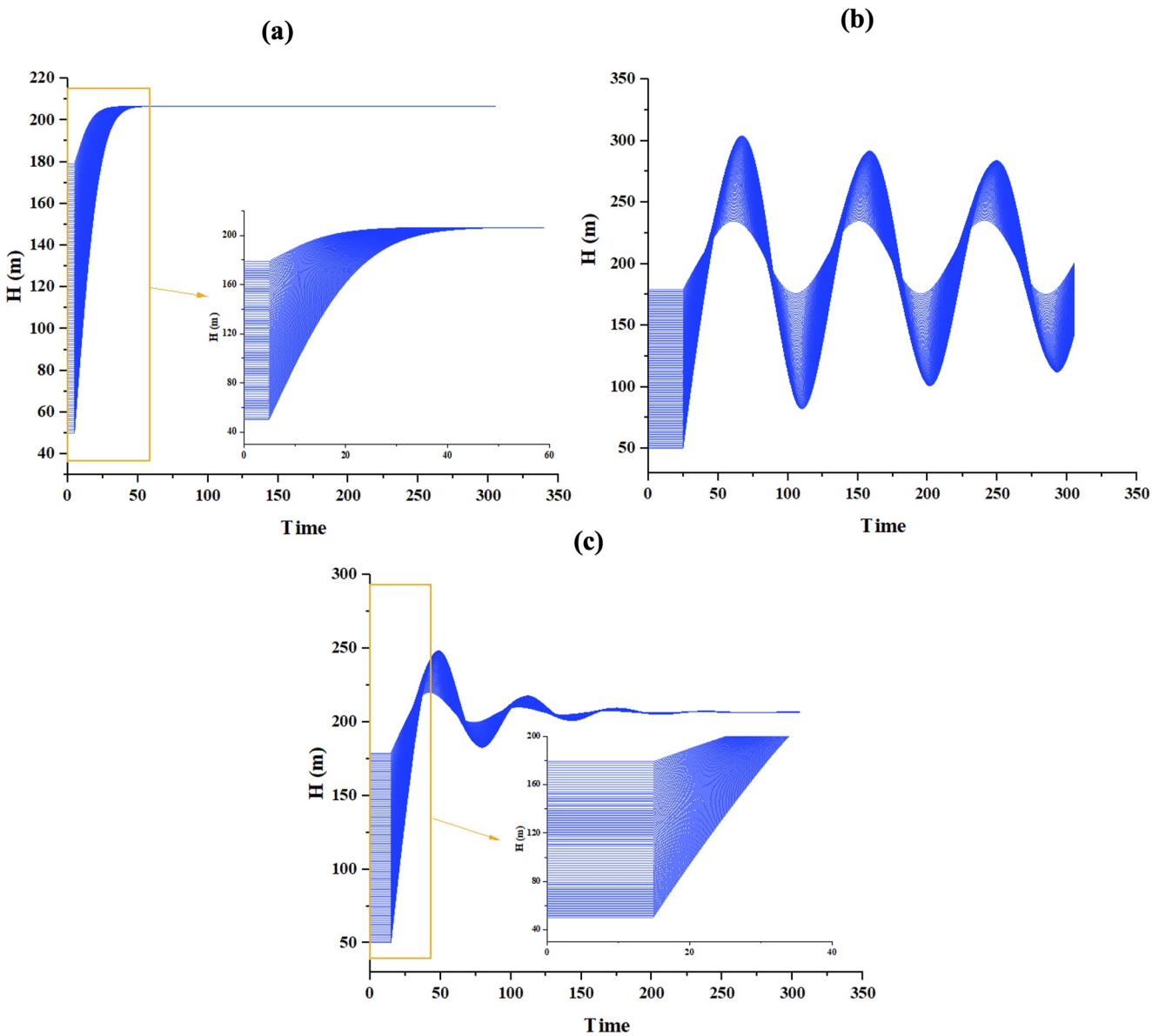

Figure A1. Data sets in the three scenarios generated by KTF equations with parameter settings in Table A1: (a) steady state, (b) oscillation, and (c) oscillation to steady state. Here we assume the cloud depth H(t) in the initial time delay period is constant (as amplified in the yellow rectangular region), and $H(t) \in [50, 180]$ ($t \in [-T, 0]$) with a total of 130 time series of each scenario.

Table A1
Parameter Setting in the KTF Model to Describe Cloud-Rain States

Parameters	Steady state	Oscillation	Oscillation to steady state
H_0(m)	800	800	800
$T(s)$	5	25	15
$N\left(\mathrm{cm}^{-3}\right)$*	16	16	16
$\tau(s)$	80	80	80
$\alpha\left(\mathrm{day}^{-1}\mathrm{m}^{-2.5}\right)$	1,000	1,000	1,000
$\Delta t(s)$	0.1	0.1	0.1

Note. The values of N * refer to the study by Koren et al. (2017).

A2. Discovered Model Prediction

The discovered equations in three scenarios are summarized in Table 3. In order to further evaluate the predictive performance of the discovered equations, we compare all of the test trajectories in three different scenarios (Figure A2). The predictions generated by the discovered equations (middle column) exhibit a remarkably good agreement with the observations obtained from the KTF models (right column). The relative errors (right column) of cloud depth H and cloud droplet concentration N are between -0.05% and 0.15%, indicating the high accuracy of the predictions. The large values of relative errors occur in the low and high peaks of cloud depth H in the

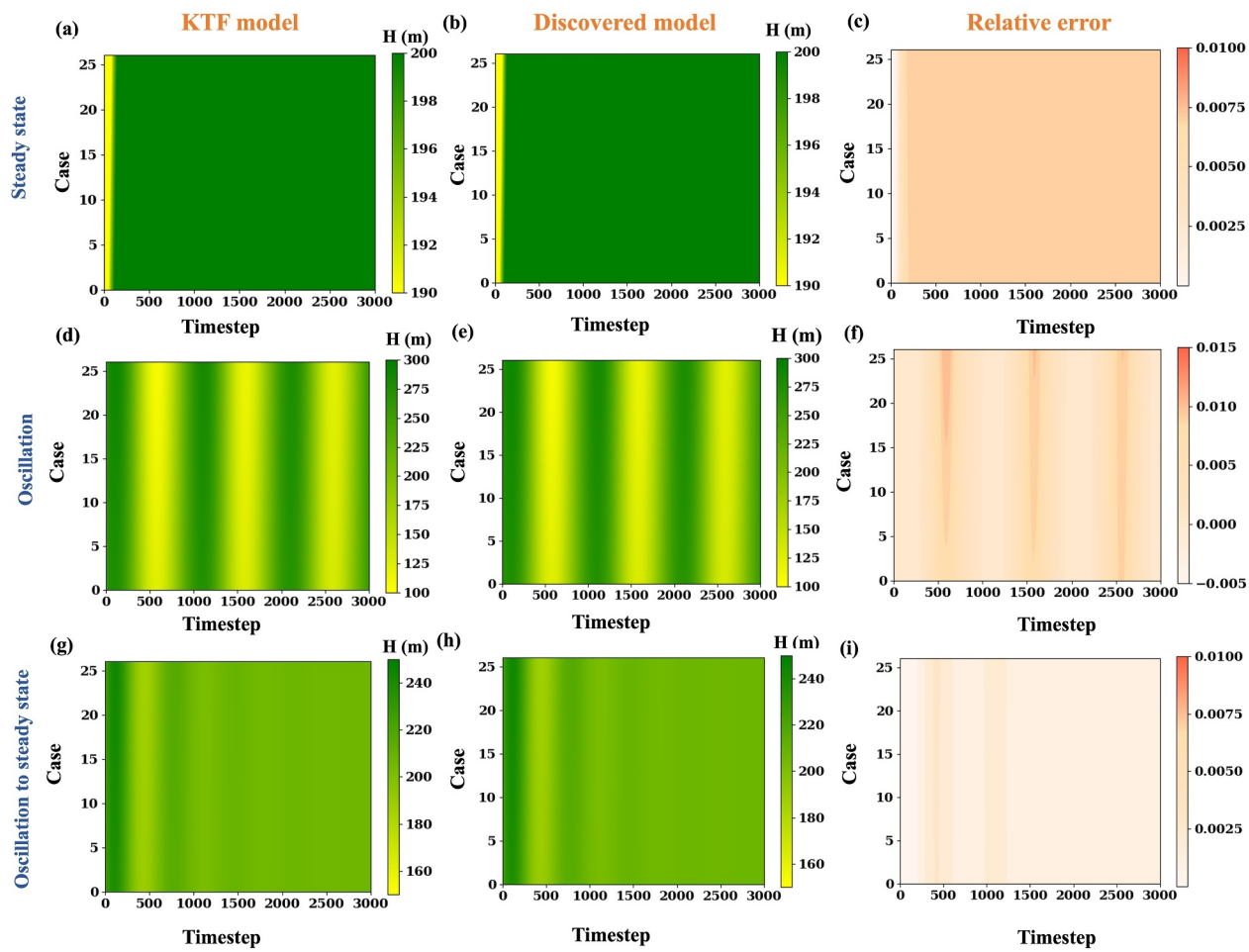

Figure A2. Trajectory comparisons between the KTF equation (left column) and the discovered models with 3 terms shown in Table 3 (middle column) across the remaining 20% test cases. The right column displays the relative errors of model predictions.

scenario of oscillation and oscillation to steady state, whereas for the scenario of steady state, the predictions are generally slightly higher than the observations (from the KTF model). These biases are caused by the differences between the reconstructed and original coefficients in the DDEs, as shown in Figure 4 (orange line with 3 terms). Overall, the promising result shows that the proposed coupled time-delay SINDy model is robust to discovering the correct underlying KTF structure.

Appendix B: Data Generation of KF Equation

For a given set of parameters in Table B1, we solve Equation 11 and Equation 12 numerically by a fourth-order Runge-Kutta method with time step Δt. We create 500 time series (Figure B1) by a combination of input

Table B1
Parameter Setting in the KF Model to Describe ACPIs

Parameters	H_0	N_0	τ_1*	$\tau_2(s)$*	α*	c_1*	c_2	Δt
Unit	m	cm^{-3}	min	min	$day^{-1}m^{-2.5}$	mm/m^2	m^{-1}	min
Value	530	180	60	60	2	2×10^{-6}	3×10^{4}	0.1

Note. The values of * refer to the study by Koren and Feingold (2011).

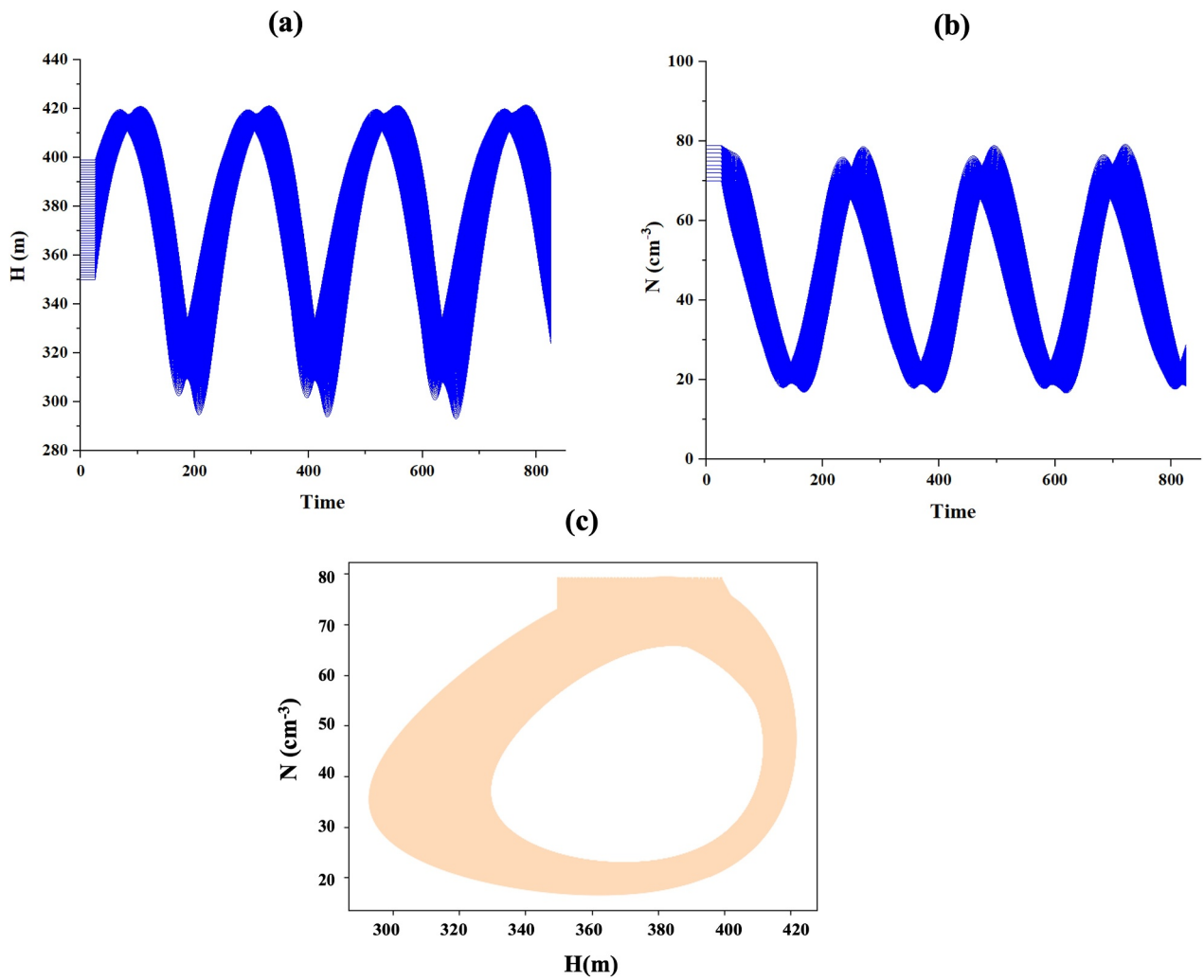

Figure B1. Data sets generated by the KF equation with parameter settings in Table B1: (a) time series of cloud depth H, (b) time series of cloud droplet concentration N, and (c) H-N cycles.

parameters and initial values in the time delay period, where the cloud depth H(t) and N(t) in the initial time delay period ($t \in [-T, 0]$) are constant, and $H(t) \in [350, 400]$ with an interval of 1 and $N(t) \in [70, 80]$ with an interval of 1. We use 400 (80%) time series of cloud depth H and cloud droplet concentration N for coupled time-delay SINDy training, and the remaining 20% time series are used for testing.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

The data from LESs in this paper are available in the study by Glassmeier et al. (2021b).

Acknowledgments
We thank Takanobu Yamaguchi and Graham Feingold for providing the LES data set and are grateful to Mickaël D. Chekroun, Graham Feingold, and Georg Gottwald for helpful discussion. This research has been supported by the Branco Weiss Fellowship Society in Science, administered by the ETH Zürich. FG further acknowledges funding by the European Union (ERC, MesoClou, 101117462). Views and opinions expressed are, however, those of the author (s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them. ChatGPT was used for grammar checks in parts of the manuscript.

References

Alfonso, L., & Zamora, J. M. (2021). A two-moment machine learning parameterization of the autoconversion process. *Atmospheric Research*, *249*, 105269. https://doi.org/10.1016/j.atmosres.2020.105269

Alinaghi, P., Jansson, F., Blázquez, D. A., & Glassmeier, F. (2024). *Cold pools mediate mesoscale adjustments of trade-cumulus fields to changes in cloud-droplet number concentration* (Vol. 2024, pp. 1–26). EGUsphere.

Bellouin, N., Quaas, J., Gryspeerdt, E., Kinne, S., Stier, P., Watson-Parris, D., et al. (2020). Bounding global aerosol radiative forcing of climate change. *Reviews of Geophysics*, *58*(1), e2019RG000660. https://doi.org/10.1029/2019rg000660

Brunton, S. L., Proctor, J. L., & Kutz, J. N. (2016). Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences*, *113*(15), 3932–3937. https://doi.org/10.1073/pnas.1517384113

Chekroun, M. D., Koren, I., & Liu, H. (2020). Efficient reduction for diagnosing Hopf bifurcation in delay differential systems: Applications to cloud-rain models. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *30*(5). https://doi.org/10.1063/5.0004697

Chekroun, M. D., Koren, I., Liu, H., & Liu, H. (2022). Generic generation of noise-driven chaos in stochastic time delay systems: Bridging the gap with high-end simulations. *Science Advances*, *8*(46), eabq7137. https://doi.org/10.1126/sciadv.abq7137

Chen, Y.-C., Christensen, M. W., Stephens, G. L., & Seinfeld, J. H. (2014). Satellite-based estimate of global aerosol–cloud radiative forcing by marine warm clouds. *Nature Geoscience*, *7*(9), 643–646. https://doi.org/10.1038/ngeo2214

Ekman, A. M., Nygren, E., Pérez, A. B., Schwarz, M., Svensson, G., & Bellouin, N. (2023). Influence of horizontal resolution and complexity of aerosol-cloud interactions on marine stratocumulus and stratocumulus-to-cumulus transition in HadGEM3-GC3. 1. *Quarterly Journal of the Royal Meteorological Society*, *149*(755), 2049–2066. https://doi.org/10.1002/qj.4494

Fasel, U., Kutz, J. N., Brunton, B. W., & Brunton, S. L. (2022). Ensemble-SINDy: Robust sparse model discovery in the low-data, high-noise limit, with active learning and control. *Proceedings of the Royal Society A*, *478*(2260), 20210904. https://doi.org/10.1098/rspa.2021.0904

Flossmann, A. I., & Wobrock, W. (2010). A review of our understanding of the aerosol–cloud interaction from the perspective of a bin resolved cloud scale modelling. *Atmospheric Research*, *97*(4), 478–497. https://doi.org/10.1016/j.atmosres.2010.05.008

Forster, P., Storelvmo, T., Armour, K., Collins, W., Dufresne, J.-L., Frame, D., et al. (2021). Chapter 7: The earth's energy budget, climate feedbacks, and climate sensitivity. In *Climate change 2021: The physical science basis. Contribution of working group I to the sixth assessment report of the intergovernmental Panel on climate change* (pp. 923–1054). Cambridge University Press.

Gao, W., Fan, J., Easter, R. C., Yang, Q., Zhao, C., & Ghan, S. J. (2016). Coupling spectral-bin cloud microphysics with the mosaic aerosol model in WRF-Chem: Methodology and results for marine stratocumulus clouds. *Journal of Advances in Modeling Earth Systems*, *8*(3), 1289–1309. https://doi.org/10.1002/2016ms000676

Geoffroy, O., Brenguier, J.-L., & Sandu, I. (2008). Relationship between drizzle rate, liquid water path and droplet concentration at the scale of a stratocumulus cloud system. *Atmospheric Chemistry and Physics*, *8*(16), 4641–4654. https://doi.org/10.5194/acp-8-4641-2008

Gettelman, A., Gagne, D. J., Chen, C.-C., Christensen, M., Lebo, Z., Morrison, H., & Gantos, G. (2021). Machine learning the warm rain process. *Journal of Advances in Modeling Earth Systems*, *13*(2), e2020MS002268. https://doi.org/10.1029/2020ms002268

Gettelman, A., Morrison, H., Santos, S., Bogenschutz, P., & Caldwell, P. (2015). Advanced two-moment bulk microphysics for global models. Part II: Global model solutions and aerosol–cloud interactions. *Journal of Climate*, *28*(3), 1288–1307. https://doi.org/10.1175/jcli-d-14-00103.1

Glassmeier, F., & Feingold, G. (2017). Network approach to patterns in stratocumulus clouds. *Proceedings of the National Academy of Sciences*, *114*(40), 10578–10583. https://doi.org/10.1073/pnas.1706495114

Glassmeier, F., Hoffmann, F., Johnson, J. S., Yamaguchi, T., Carslaw, K. S., & Feingold, G. (2019). An emulator approach to stratocumulus susceptibility. *Atmospheric Chemistry and Physics*, *19*(15), 10191–10203. https://doi.org/10.5194/acp-19-10191-2019

Glassmeier, F., Hoffmann, F., Johnson, J. S., Yamaguchi, T., Carslaw, K. S., & Feingold, G. (2021a). Aerosol-cloud-climate cooling overestimated by ship-track data. *Science*, *371*(6528), 485–489. https://doi.org/10.1126/science.abd3980

Glassmeier, F., Hoffmann, F., Johnson, J. S., Yamaguchi, T., Carslaw, K. S., & Feingold, G. (2021b). Aerosol-cloud-climate cooling overestimated by ship-track data [Dataset]. *NOAA*. Retrieved from https://csl.noaa.gov/groups/csl9/datasets/data/2020-Glassmeier-etal/

Gryspeerdt, E., Glassmeier, F., Feingold, G., Hoffmann, F., & Murray-Watson, R. J. (2022). Observing short-timescale cloud development to constrain aerosol–cloud interactions. *Atmospheric Chemistry and Physics*, *22*(17), 11727–11738. https://doi.org/10.5194/acp-22-11727-2022

Hoffmann, F., Glassmeier, F., & Feingold, G. (2024). The impact of aerosol on cloud water: A heuristic perspective. *Atmospheric Chemistry and Physics*, *24*(23), 13403–13412. https://doi.org/10.5194/acp-24-13403-2024

Hoffmann, F., Glassmeier, F., Yamaguchi, T., & Feingold, G. (2020). Liquid water path steady states in stratocumulus: Insights from process-level emulation and mixed-layer theory. *Journal of the Atmospheric Sciences*, *77*(6), 2203–2215. https://doi.org/10.1175/jas-d-19-0241.1

Jansson, F., Janssens, M., Grönqvist, J. H., Siebesma, A. P., Glassmeier, F., Attema, J., et al. (2023). Cloud botany: Shallow cumulus clouds in an ensemble of idealized large-domain large-eddy simulations of the trades. *Journal of Advances in Modeling Earth Systems*, *15*(11), e2023MS003796. https://doi.org/10.1029/2023ms003796

Kaheman, K., Kutz, J. N., & Brunton, S. L. (2020). Sindy-pi: A robust algorithm for parallel implicit sparse identification of nonlinear dynamics. *Proceedings of the Royal Society A*, *476*(2242), 20200279. https://doi.org/10.1098/rspa.2020.0279

Kaiser, E., Kutz, J. N., & Brunton, S. L. (2018). Sparse identification of nonlinear dynamics for model predictive control in the low-data limit. *Proceedings of the Royal Society A*, *474*(2219), 20180335. https://doi.org/10.1098/rspa.2018.0335

Khairoutdinov, M. F., & Randall, D. A. (2003). Cloud resolving modeling of the arm summer 1997 Iop: Model formulation, results, uncertainties, and sensitivities. *Journal of the Atmospheric Sciences*, *60*(4), 607–625. https://doi.org/10.1175/1520-0469(2003)060<0607:crmota>2.0.co;2

Koren, I., & Feingold, G. (2011). Aerosol–cloud–precipitation system as a predator-prey problem. *Proceedings of the National Academy of Sciences*, *108*(30), 12227–12232. https://doi.org/10.1073/pnas.1101777108

Koren, I., Tziperman, E., & Feingold, G. (2017). Exploring the nonlinear cloud and rain equation. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *27*(1), 013107. https://doi.org/10.1063/1.4973593

Lamb, K. D., van Lier-Walqui, M., Santos, S., & Morrison, H. (2024). Reduced-order modeling for linearized representations of microphysical process rates. *Journal of Advances in Modeling Earth Systems*, *16*(7), e2023MS003918. https://doi.org/10.1029/2023ms003918

Liu, X., Pei, L., & Qi, S. (2022a). Complex dynamics in the improved Koren–Feingold cloud–rain system. *International Journal of Non-Linear Mechanics*, *147*, 104210. https://doi.org/10.1016/j.ijnonlinmec.2022.104210

Liu, X., Pei, L., Wang, W., & Qi, S. (2022b). Double Hopf bifurcation and stability of Koren–Feingold cloud–rain system with rain production delay. *The European Physical Journal Special Topics*, *231*(11–12), 2103–2114. https://doi.org/10.1140/epjs/s11734-021-00412-6

Lohmann, U., Quaas, J., Kinne, S., & Feichter, J. (2007). Different approaches for constraining global climate models of the anthropogenic indirect aerosol effect. *Bulletin of the American Meteorological Society*, *88*(2), 243–250. https://doi.org/10.1175/bams-88-2-243

Lunderman, S., Morzfeld, M., Glassmeier, F., & Feingold, G. (2020). Estimating parameters of the nonlinear cloud and rain equation from a large-eddy simulation. *Physica D: Nonlinear Phenomena*, *410*, 132500. https://doi.org/10.1016/j.physd.2020.132500

Mangan, N. M., Brunton, S. L., Proctor, J. L., & Kutz, J. N. (2016). Inferring biological networks by sparse identification of nonlinear dynamics. *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, *2*(1), 52–63. https://doi.org/10.1109/tmbmc.2016.2633265

Mangan, N. M., Kutz, J. N., Brunton, S. L., & Proctor, J. L. (2017). Model selection for dynamical systems via sparse regression and information criteria. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *473*(2204), 20170009. https://doi.org/10.1098/rspa.2017.0009

Naozuka, G. T., Rocha, H. L., Silva, R. S., & Almeida, R. C. (2022). SINDy-SA framework: Enhancing nonlinear system identification with sensitivity analysis. *Nonlinear Dynamics*, *110*(3), 1–21. https://doi.org/10.1007/s11071-022-07755-2

Rudy, S., Alla, A., Brunton, S. L., & Kutz, J. N. (2019). Data-driven identification of parametric partial differential equations. *SIAM Journal on Applied Dynamical Systems*, *18*(2), 643–660. https://doi.org/10.1137/18m1191944

Sandoz, A., Ducret, V., Gottwald, G. A., Vilmart, G., & Perron, K. (2023). SINDy for delay-differential equations: Application to model bacterial zinc response. *Proceedings of the Royal Society A*, *479*(2269), 20220556. https://doi.org/10.1098/rspa.2022.0556

Savic-Jovcic, V., & Stevens, B. (2008). The structure and mesoscale organization of precipitating stratocumulus. *Journal of the Atmospheric Sciences*, *65*(5), 1587–1605. https://doi.org/10.1175/2007jas2456.1

Seifert, A., Heus, T., Pincus, R., & Stevens, B. (2015). Large-eddy simulation of the transient and near-equilibrium behavior of precipitating shallow convection. *Journal of Advances in Modeling Earth Systems*, *7*(4), 1918–1937. https://doi.org/10.1002/2015ms000489

Shea, D. E., Brunton, S. L., & Kutz, J. N. (2021). SINDy-BVP: Sparse identification of nonlinear dynamics for boundary value problems. *Physical Review Research*, *3*(2), 023255. https://doi.org/10.1103/physrevresearch.3.023255

Silva, S. J., Ma, P.-L., Hardin, J. C., & Rothenberg, D. (2021). Physically regularized machine learning emulators of aerosol activation. *Geoscientific Model Development*, *14*(5), 3067–3077. https://doi.org/10.5194/gmd-14-3067-2021

Stevens, B., & Feingold, G. (2009). Untangling aerosol effects on clouds and precipitation in a buffered system. *Nature*, *461*(7264), 607–613. https://doi.org/10.1038/nature08281

Terai, C., Pritchard, M., Blossey, P., & Bretherton, C. (2020). The impact of resolving subkilometer processes on aerosol-cloud interactions of low-level clouds in global model simulations. *Journal of Advances in Modeling Earth Systems*, *12*(11), e2020MS002274. https://doi.org/10.1029/2020ms002274

Wang, H., & Feingold, G. (2009). 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*(11), 3237–3256. https://doi.org/10.1175/2009jas3022.1

Wu, Y. (2023). Reconstruction of delay differential equations via learning parameterized dictionary. *Physica D: Nonlinear Phenomena*, *446*, 133647. https://doi.org/10.1016/j.physd.2023.133647

Yamaguchi, T., Feingold, G., & Kazil, J. (2017). Stratocumulus to cumulus transition by drizzle. *Journal of Advances in Modeling Earth Systems*, *9*(6), 2333–2349. https://doi.org/10.1002/2017ms001104

Yamaguchi, T., Feingold, G., & Kazil, J. (2019). Aerosol-cloud interactions in trade wind cumulus clouds and the role of vertical wind shear. *Journal of Geophysical Research: Atmospheres*, *124*(22), 12244–12261. https://doi.org/10.1029/2019jd031073

Zhang, Y., Fan, J., Li, Z., & Rosenfeld, D. (2021). Impacts of cloud microphysics parameterizations on simulated aerosol–cloud interactions for deep convective clouds over houston. *Atmospheric Chemistry and Physics*, *21*(4), 2363–2381. https://doi.org/10.5194/acp-21-2363-2021