Discovering How Ice Crystals Grow Using Neural ODE's and Symbolic Regression

Kara D. Lamb
Department of Earth and Env. Engineering
Columbia University
New York, NY
kl3231@columbia.edu

Jerry Y. Harrington
Department of Meteorology
Pennsylvania State University
State College, PA
jyh10@psu.edu

Abstract

Depositional ice growth is an important process for atmospheric cloud formation, but the physics of ice growth in atmospheric conditions is still poorly understood. One major challenge in constraining depositional ice growth models against observations is that the early growth rates of ice crystals cannot be directly observed, and proposed models require assumptions about the functional dependence of physical processes that are still highly uncertain. Neural ordinary differential equations (NODE's) are a recently developed machine learning method that can be used to learn the derivative of a hidden state. Here we explore how NODE's can be used to evaluate model structural uncertainty in depositional ice growth models by optimizing against experimental measurements. We find a functional form for the depositional ice growth model that best fits 290 mass time series of ice crystals grown in a levitation diffusion chamber. We use symbolic regression to derive a closed-form equation for the function learned by the NODE model.

1 Introduction

Scientific machine learning for discovering unknown physics directly from observations has demonstrated significant promise in recent years [1, 2, 3]. Approaches such as physics-informed neural networks (PINN's) [4] can be leveraged to integrate observational data with known governing physical laws, even in cases with partially unknown physics. PINN's learn the solution of ODE's or PDE's using neural networks, with the loss function formulated to include data mismatch and gradients of known ODE's or PDE's. Neural ODE's (NODE's) integrate neural networks parameterizing a hidden state with numerical ODE solvers to learn a continuous depth model for a physical system [5].

The growth of ice crystals from vapor is an important microphysical process impacting cloud formation. However, significant gaps in our understanding of depositional ice growth, particularly the early growth of ice crystals in clouds, continue to limit our ability to accurately model these processes in atmospheric models [6, 7]. Early growth is particularly hard to constrain because this growth connects freshly nucleated particles to larger faceted crystals. These small ice crystals (radii < 50 μm) undergo various surface transformations as habit forms develop, affecting the growth rates.

Structural and parametric uncertainty in ice growth models has been difficult to address [8]. Structural uncertainty refers to uncertainty in the functional dependence of a physical model, while parametric uncertainty refers to uncertainty in parameter values. While depositional ice growth is typically represented in cloud and climate models with capacitance theory [9], the early stages of ice growth leading to the formation and growth of facets is determined by surface attachment kinetics, which are poorly constrained by current theories, observations (those made in clouds), and lab measurements. Measurements from different experimental studies sometimes show large disagreement [10, 11], although recent work has suggested these studies can be reconciled with a saturation and temperature dependent functional form for surface attachment kinetics modeled as a deposition coefficient [12, 13].

38th Conference on Neural Information Processing Systems (NeurIPS 2024).

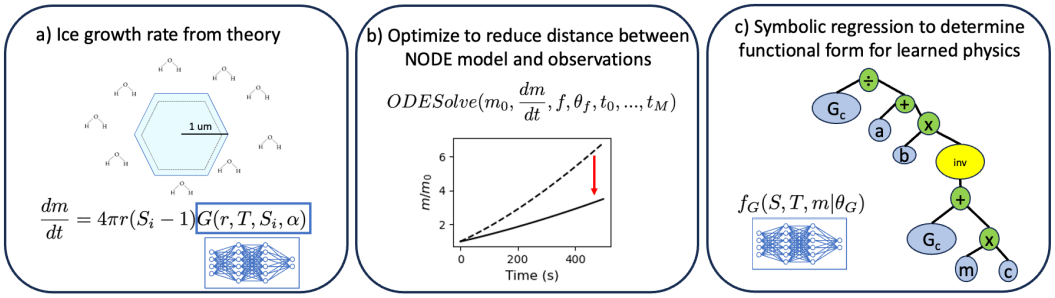

Figure 1: Overview of methodology for learning unknown physics in depositional ice growth models.

Here we use scientific machine learning to reduce structural uncertainty in depositional ice growth models. The contributions of our paper are as follows:

- We develop a methodology to learn the functional dependence for an ordinary differential equation (ODE) by optimizing a NODE across multiple time series simultaneously. We compare strongly-constrained and weakly-constrained NODE models, based on the amount of prior physical knowledge that is included.
- We apply this method to experimental observations to learn a mathematical model for the depositional ice growth rate and use symbolic regression to discover a closed form expression for the non-linear relationship learned by the neural network.

1.1 Data sets

Here we explore how scientific machine learning can be applied to observations of ice crystals grown in a levitation diffusion chamber at temperatures between 205 - 240 K and saturation with respect to ice between 1.0 and 1.8 [14, 15, 16]. During experiments, single ice crystals with initial radii between 6 and 12 μm are nucleated, levitated, and grown from vapor in constant saturation S_i and temperature T conditions. Observations consist of 290 time series of the mass ratio (m/m_0) of individual ice crystals, where m is the mass of an ice crystal, and m_0 is its initial mass. The crystals remain relatively small, with equivalent spherical radii < 60 μm. Since mass ratios have varying durations (depending on how efficiently the ice crystals grow), we interpolate observations to 1 Hz and crop all data sets to a maximum time-length of 500 seconds.

In addition to the experimentally observed ice crystal mass ratios, we create a synthetic data set with a known functional dependence for the depositional growth rate to validate the performance of our equation discovery method. This synthetic data set consists of 290 mass ratios of ice crystals growing in the same saturation and temperature conditions and with the same initial mass, but with an assumed functional form for the depositional ice growth model based on Nelson and Baker, 1996 [17]. Realistic measurement noise is added to the synthetic time series by using the trailing eigenmodes of a singular spectrum analysis decomposition applied to the observed time series. Additional details about experimental and synthetic data sets are given in Appendix Sections A.2 and A.3.

2 Methodology

In this section, we describe the problem and the machine learning methodology (Figure 1) to discover a mathematical model for depositional ice growth. The depositional ice growth rate for a single ice crystal growing from vapor has traditionally been modeled using the capacitance ice growth model [9], which is sometimes modified to include surface kinetics [12]. This model is an ODE,

$$\frac{dm}{dt} = 4\pi r(S_i - 1)G(r, T, S_i, \alpha) \tag{1}$$

where m is the mass of the crystal, r is the radius of the ice crystal, S_i is the supersaturation with respect to ice far from the crystal, and the function $G(r, T, S_i, \alpha)$ represents the combined effects of vapor and thermal diffusivity to the surface of the ice crystal. $G(r, T, S_i, \alpha)$ is a function of r, the ambient temperature T, S_i, and surface attachment kinetics typically parameterized by a deposition coefficient α. The deposition coefficient α has previously been parameterized as a saturation and temperature dependent function in Nelson and Baker, 1996 [17]. Predicting ice growth in the

levitation diffusion chamber amounts to solving an initial value problem, where the observed mass as a function of time is given by

$$m(t) = \int_0^t \frac{dm}{dt} dt + m_0 \tag{2}$$

If we knew the functional dependence of Eq 1, then this problem would be straightforward to solve.

Since the functional dependence for $G(r, T, S_i, \alpha)$ is uncertain, however, we use a NODE model to solve Eq 2. We assume two cases for prior physical knowledge based on past literature (Figure 1a). The first model makes a stronger assumption about the amount of prior physical knowledge to include, and the second model learns a greater part of the ice growth model. In the first case, we use a capacitance model for ice growth that assumes an unknown function for the deposition coefficient α, which we refer to as the "strongly-constrained NODE" model. In the second case, we fit the ratio of G relative to the expression for a spherical ice crystal assuming the continuum limit G_c (Eq. A9), which we refer to as the "weakly-constrained NODE" model. In each case, we use prior physical knowledge from Eq 1 and replace only the uncertain portion of the model with a neural network. For the strongly-constrained NODE model, we assume the functional form for $G(r, T, S_i, \alpha)$ is given by Eq. A3, and we assume α is a function of temperature and supersaturation,

$$\alpha = f_\alpha(S_i, T|\theta_\alpha) \tag{3}$$

where $f_\alpha(S_i, T|\theta_\alpha)$ represents a neural network that takes as input the saturation and temperature and given the weights of the neural network θ_α predicts a value for α. For the weakly constrained NODE model, we assume the ratio between G and G_c is given by a function of the temperature, supersaturation, and mass of the ice crystal,

$$\frac{G}{G_c} = f_G(S_i, T, m|\theta_G). \tag{4}$$

Parameterizing kinetics with a modified G is advantageous because it is more general, and it has some measurement [15] and theoretical [18] backing. For both the strongly and weakly constrained models, we use an MLP to parameterize the functional dependence, with 3 linear layers with 50 neurons in each layer, and ReLU activation functions after the first and second layers. Following the third linear layer, we use a sigmoid activation function, as both functional forms are expected to be constrained to a range of values ($0 \leq \alpha \leq 1$ and $0 < \frac{G}{G_c} \leq 2$); without this constraint, the NODE model is significantly harder to optimize against observations. We then integrate Eq. 1 using an ODE solver (RK4, implemented in PyTorch with the torchdiffeq library [5]). Further details of these models are provided in Section A.1.

We optimize the NODE model against the observational time series to reduce the distance between the model and the observed time series (Figure 1b). Since we assume all 290 experimental mass time series can be modeled with the same physical model, we optimize the NODE model across all time series simultaneously. In cases where the time series are less than 500 s, we mask experimental data such that it is not included in the calculation of the loss function. We minimize L2 loss between the observed time series and the model,

$$\mathcal{L} = \sum_{j=0}^{290} \sum_{i=0}^{T_j} \left(\frac{m_j(t)}{m_{j,0}} - \frac{\hat{m}_j(t)}{m_{j,0}} \right)^2 \tag{5}$$

where j is the experiment number, and T_j is the length of the jth time series, and $\hat{m}_j(t)$ indicates the prediction from the NODE model. To minimize the loss function, we use the AdamW optimizer [19], with a base learning rate of 0.01, which we train for 1000 epochs, with cosine decay [20].

Once the NODE model has been optimized against observations, we use symbolic regression to derive functional forms for the trained neural networks (Figure 1c). Symbolic regression uses genetic algorithms to search for equations that optimize accuracy while limiting model complexity. Here we use the symbolic regression library PySR [21]. We use the observed values for S_i and T, and the initial mass m_0 for each of the 290 experiments. These values are used as inputs for the trained neural networks to determine the corresponding values for α for the strongly-constrained NODE model (Eq. 3) and for G for the weakly-constrained NODE model (Eq. 4). To find a symbolic expression for the function α in terms of S_i and T, we use the binary operators for summation, multiplication, exponentiation, division, and subtraction, and the unary operators log, exp, $1/x$, $square$, $cube$, and $tanh$, and run PySR for 100 generations. To fit a symbolic expression to G, we use the binary

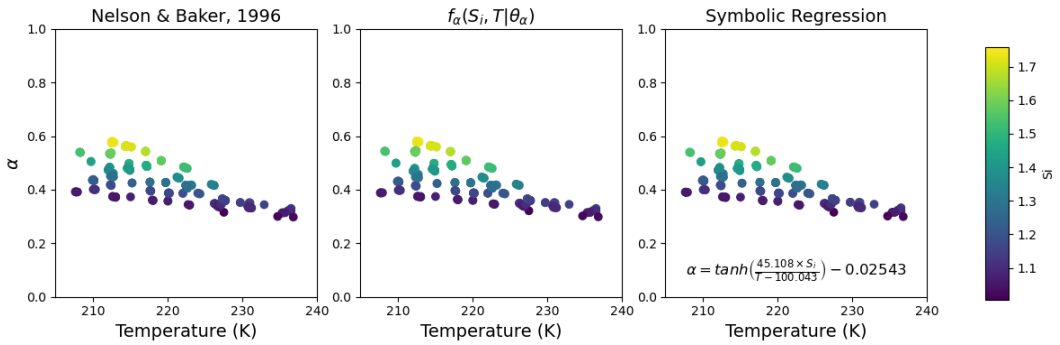

Figure 2: Functional dependence of α learned from synthetic data sets. a) Saturation and temperature dependence of α parameterization from Nelson and Baker, 1996 [17]. b) Saturation and temperature dependence from the trained NN for the synthetic data sets. c) Predictions from an expression learned by symbolic regression from the trained NN for the synthetic data sets.

Table 1: Comparison of model performance on observational data. MSE loss between model and observations is evaluated across all 290 time series.

Ice growth model	MSE Loss	# Exps. for which model performs best
Nelson and Baker, 1996 [17]	40742	47
Strongly-constrained NODE model	28153	89
Weakly-constrained NODE model	**15249**	**154**

operators for summation, multiplication, exponentiation, division, and subtraction, and the unary operators $1/x$, $square$, and $cube$, and run PySR for 100 generations. We use m, r, T, S_i, and G_c as input features, and target the prediction of G; this approach led to more accurate fits to the trained neural network than finding a symbolic expression for the ratio G/G_c. G_c is given by Eq. A9, and does not depend on unknown or unobserved physics, such as the shape of the growing ice crystals.

3 Results

We first test our method on synthetic data sets with a known functional form for the depositional coefficient α based on [17]. After optimizing the strongly-constrained NODE model against the synthetic data sets, we find that the model is able to reproduce the time series very accurately (Figure A5 shows examples of a subset of time series compared to NODE model fits). We compare the predicted values of α from the trained neural network α_{NN} to those used to generate the synthetic data set α_{Nelson}. The strongly-constrained NODE model is able to learn the functional dependence for α that closely matches the one used to generate the synthetic data sets (Figure 2, left and middle panels). In addition, when we use symbolic regression to learn a functional dependence for α from the trained neural network, it closely matches the true dependence (Figure 2, right panel).

We next use our approach on the experimentally observed time series. Since we have no ground-truth for the depositional ice growth model for the observational data sets, we compare the performance of the weakly and strongly-constrained NODE models optimized against the experimental data sets with the model from Nelson and Baker, 1996 [17] (Table 1). We evaluate the MSE loss between models and observations across all 290 experiments (Eq. 5), as well as evaluating which model performs best on the majority of individual experiments (determined by the lowest MSE loss for the individual experiment). According to both metrics, the weakly-constrained NODE model outperforms the Nelson and Baker, 1996 model and the strongly-constrained NODE model. The weakly-constrained NODE model performs best on 154 out of 290 experiments.

In comparing the strongly-constrained NODE model optimized against the experimental data sets with the observed mass ratios, in some cases, we find significant deviations between the observed mass ratios and predicted mass ratios (Figure A6). The predicted values of α from the trained neural network also differ significantly from those predicted by the functional form given in [17]. As the strongly-constrained NODE model is not able to learn a functional dependence for α that reproduces the mass ratio time series for the majority of the experiments, this suggests that a model including the

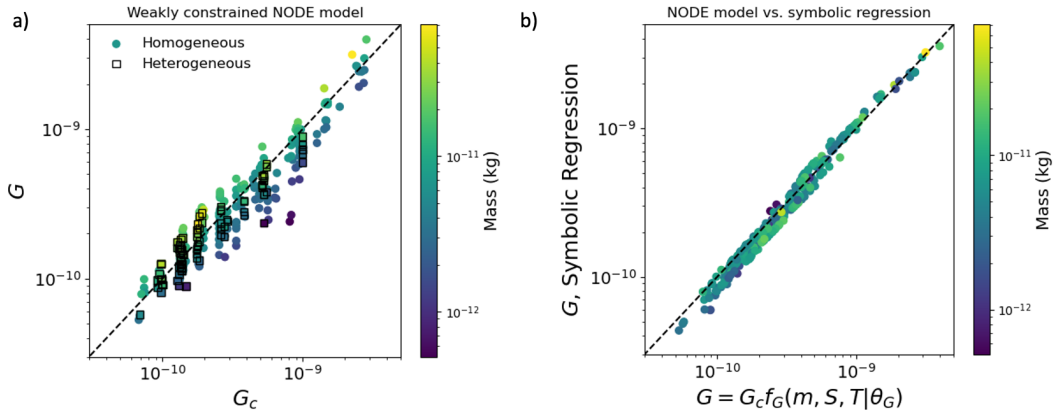

Figure 3: Unknown physics learned by weakly-constrained NODE model from observations. a) The transfer coefficient G learned by the weakly-constrained NODE model for the 290 experiments compared with G_c (the transfer coefficient for the continuum case, assuming a spherical ice crystal). b) Predictions from the trained neural network for G compared with the predictions from an expression learned by symbolic regression from the trained neural networks (Eq. 6) for the 290 experiments.

deposition coefficient α may not be the optimal functional relationship to describe the depositional ice growth process in these experiments. By contrast, when we optimize the weakly-constrained NODE model against the experimental data sets, this generally reduces deviations between the mass ratio time series and the predicted values for the mass ratios (Figure A6) compared to the strongly-constrained NODE model. In addition, we can use the trained neural network to evaluate the learned functional dependence of G compared with G_c (Figure 3a), and our model learns a consistent functional dependence across the 290 experiments. While G is generally within a factor of 1.2 of G_c for most experiments, the learned functional dependence indicates there is an additional dependence on m that is not accounted for by G_c.

As described in the previous section, we use symbolic regression to determine a functional dependence for G learned by the neural network. The learned expression for G is

$$G = \frac{G_c}{a + b(G_c + cm)^{-1}} \tag{6}$$

where $a = 0.6517688$, $b = 2.98707 \times 10^{-9}$, and $c = 1000$ are constants, m is the mass of the ice crystal in kg, and G_c is given by Eq. A9. The predictions for this functional dependence closely match the predictions from the neural network (Figure 3b).

4 Conclusions and Outlook

In this paper we have demonstrated the application of NODE's and symbolic regression to learn a mathematical model directly from multiple experimental time series. We validate our approach to learning the unknown part of a partially known ODE using a synthetic data set with a known functional dependence for the deposition coefficient α. We find a weakly-constrained NODE mode is able to more accurately reproduce the majority of our experimental time series, suggesting that prior parameterizations for surface kinetic effects as a deposition coefficient function may be too restrictive. We instead propose a new functional dependence for G in Eq. 1 that more accurately reproduces the experimental observations, and that can be expressed as a closed-form function using symbolic regression (Eq. 6). Future work will focus on simultaneously learning parametric and structural uncertainty in ice growth models [22]. When we included 17 additional experiments from [14] in our analysis, which had higher uncertainty on S_i (since these came from earlier experiments, when the levitation diffusion chamber was not fully characterized), the NODE model struggled to learn a consistent ice growth model across all experiments. Thus, future extensions of the equation discovery approach we have developed here should also focus on robustly identifying outliers when constraining ice growth models. This research contributes to the development of improved model parameterizations for cloud processes, as processes related to ice formation in clouds are a significant source of uncertainty in current climate models [8, 23].

References

[1] Michael Schmidt and Hod Lipson. Distilling free-form natural laws from experimental data. *science*, 324(5923):81–85, 2009.

[2] Miles Cranmer, Alvaro Sanchez Gonzalez, Peter Battaglia, Rui Xu, Kyle Cranmer, David Spergel, and Shirley Ho. Discovering symbolic models from deep learning with inductive biases. *Advances in neural information processing systems*, 33:17429–17442, 2020.

[3] Silviu-Marian Udrescu and Max Tegmark. Ai feynman: A physics-inspired method for symbolic regression. *Science Advances*, 6(16):eaay2631, 2020.

[4] M. Raissi, P. Perdikaris, and G. E. Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, February 2019.

[5] Ricky TQ Chen, Yulia Rubanova, Jesse Bettencourt, and David K Duvenaud. Neural ordinary differential equations. *Advances in neural information processing systems*, 31, 2018.

[6] Alexander Avramov and Jerry Y Harrington. Influence of parameterized ice habit on simulated mixed phase arctic clouds. *Journal of Geophysical Research: Atmospheres*, 115(D3), 2010.

[7] Sylvia C Sullivan and Aiko Voigt. Ice microphysical processes exert a strong control on the simulated radiative energy budget in the tropics. *Communications Earth & Environment*, 2(1):1–8, 2021.

[8] Hugh Morrison, Marcus Van Lier-Walqui, Ann M. Fridlind, Wojciech W. Grabowski, Jerry Y. Harrington, Corinna Hoose, Alexei Korolev, Matthew R. Kumjian, Jason A. Milbrandt, Hanna Pawlowska, Derek J. Posselt, Olivier P. Prat, Karly J. Reimel, Shin-Ichiro Shima, Bastiaan Van Diedenhoven, and Lulin Xue. Confronting the Challenge of Modeling Cloud and Precipitation Microphysics. *Journal of Advances in Modeling Earth Systems*, 12(8), August 2020.

[9] Hans R Pruppacher, James D Klett, and Pao K Wang. Microphysics of clouds and precipitation, 1998.

[10] Nathan Magee, Alfred M Moyle, and Dennis Lamb. Experimental determination of the deposition coefficient of small cirrus-like ice crystals near- 50 ° c. *Geophysical research letters*, 33(17), 2006.

[11] J. Skrotzki, P. Connolly, M. Schnaiter, H. Saathoff, O. Möhler, R. Wagner, M. Niemand, V. Ebert, and T. Leisner. The accommodation coefficient of water molecules on ice – cirrus cloud studies at the AIDA simulation chamber. *Atmospheric Chemistry and Physics*, 13(8):4451–4466, April 2013.

[12] Jerry Y. Harrington, Alfred Moyle, Lavender Elle Hanson, and Hugh Morrison. On Calculating Deposition Coefficients and Aspect-Ratio Evolution in Approximate Models of Ice Crystal Vapor Growth. *Journal of the Atmospheric Sciences*, 76(6):1609–1625, June 2019.

[13] Kara D Lamb, Jerry Y Harrington, Benjamin W Clouser, Elisabeth J Moyer, Laszlo Sarkozy, Volker Ebert, Ottmar Möhler, and Harald Saathoff. Re-evaluating cloud chamber constraints on depositional ice growth in cirrus clouds–part 1: Model description and sensitivity tests. *Atmospheric Chemistry and Physics*, 23(11):6043–6064, 2023.

[14] Alexander Harrison, Alfred M Moyle, Marcus Hanson, and Jerry Y Harrington. Levitation diffusion chamber measurements of the mass growth of small ice crystals from vapor. *Journal of the Atmospheric Sciences*, 73(7):2743–2758, 2016.

[15] Gwenore F. Pokrifka, Alfred M. Moyle, Lavender Elle Hanson, and Jerry Y. Harrington. Estimating Surface Attachment Kinetic and Growth Transition Influences on Vapor-Grown Ice Crystals. *Journal of the Atmospheric Sciences*, 77(7):2393–2410, July 2020.

[16] Gwenore F Pokrifka, Alfred M Moyle, and Jerry Y Harrington. Effective density derived from laboratory measurements of the vapor growth rates of small ice crystals at- 65 to- 40 c. *Journal of the Atmospheric Sciences*, 80(2):501–517, 2023.

[17] JT Nelson and MB Baker. New theoretical framework for studies of vapor growth and sublimation of small ice crystals in the atmosphere. *Journal of Geophysical Research: Atmospheres*, 101(D3):7033–7047, 1996.

[18] Jerry Y Harrington, Dennis Lamb, and Robert Carver. Parameterization of surface kinetic effects for bulk microphysical models: Influences on simulated cirrus dynamics and structure. *Journal of Geophysical Research: Atmospheres*, 114(D6), 2009.

[19] I Loshchilov. Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101*, 2017.

[20] Ilya Loshchilov and Frank Hutter. Sgdr: Stochastic gradient descent with warm restarts. *arXiv preprint arXiv:1608.03983*, 2016.

[21] Miles Cranmer. Interpretable machine learning for science with pysr and symbolicregression. jl. *arXiv preprint arXiv:2305.01582*, 2023.

[22] Zongren Zou, Xuhui Meng, Apostolos F Psaros, and George E Karniadakis. Neuraluq: A comprehensive library for uncertainty quantification in neural differential equations and operators. *SIAM Review*, 66(1):161–190, 2024.

[23] Margaret L Duffy, Brian Medeiros, Andrew Gettelman, and Trude Eidhammer. Perturbing parameters to understand cloud contributions to climate change. *Journal of Climate*, 37(1):213–227, 2024.

[24] Daniel M Murphy and Thomas Koop. Review of the vapour pressures of ice and supercooled water for atmospheric applications. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, 131(608):1539–1565, 2005.

[25] Neil J Bacon, Marcia B Baker, and Brian D Swanson. Initial stages in the morphological evolution of vapour-grown ice crystals: A laboratory investigation. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, 129(591):1903–1927, 2003.

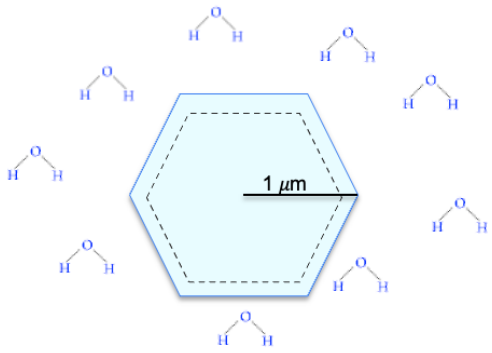

Figure A1: Depositional ice growth is an important process for ice formation in atmospheric clouds. Ice crystals grow via direct deposition of water molecules from the vapor phase onto the ice surface.

A Appendix

A.1 Depositional Ice Growth Model

The depositional ice growth rate for a single ice crystal growing from vapor (Figure A1) in the atmosphere has traditionally been modeled using the capacitance ice growth model, which describes the growth of an ice crystal from the vapor phase [9]. The capacitance growth model, including modifications for surface kinetic effects, can be written as an ordinary differential equation,

$$\frac{dm}{dt} = 4\pi C(S_i - 1)G(r, T, S_i, \alpha) \tag{A1}$$

where C is the capacitance, S_i is the ice supersaturation far from the crystal, and G represents the combined effects of vapor and thermal diffusivity to the surface of the ice crystal. G is a function of temperature, pressure, and the modified diffusivity D_v^*. The capacitance is a function that depends on the geometry of the ice crystal; here we assume that $C = r$. The modified diffusivity is a function of the diffusivity of water molecules in air D_v, the mean speed of molecules in the vapor $\bar{v}$, the molecular jump distance Δ_r, a particle length scale (typically assumed to be proportional to the radius of the ice crystal, r) and surface attachment kinetics typically parameterized by a deposition coefficient α. Surface attachment kinetics refer to the effects of individual molecules attaching to the surface of a growing crystal, which lead to differences in thermal and vapor diffusivity relative to the continuum case (where the effects of individual molecules can be ignored). The deposition coefficient has previously been parameterized as a saturation and temperature dependent function in [17] as

$$\alpha = \left(\frac{s_{local}}{s_{crit}}\right)^m tanh\left[\left(\frac{s_{crit}}{s_{local}}\right)^m\right] \tag{A2}$$

where s_{local} is the supersaturation immediately above the crystal surface, s_{crit} is the temperature-dependent critical supersaturation, and m is a parameter which relates to the surface of the ice crystals. A value of $m = 1$ is used to represent spiral dislocation growth, while a value of $m > 10$ represents ledge nucleation. This parameterization is used to represent complex surface processes that are not fully represented by current theory.

The function G is given by

$$G = \left[\frac{\mathcal{R}T_\infty}{e_{sat,i}(T_\infty)D_v^* M_w} + \frac{L_s}{k_a^* T_\infty}\left(\frac{L_s M_w}{\mathcal{R}T_\infty} - 1\right)\right]^{-1} \tag{A3}$$

Due to surface attachment effects, the diffusivity in Eq. A3 is assumed to be modified from the continuum case, with the modified diffusivity expressed as

$$D_v^* = \frac{D_v}{\frac{r}{r+\Delta_r} + \frac{D_v}{r\alpha}\left(\frac{w\pi M_w}{RT_a}\right)^{1/2}} \tag{A4}$$

Table 2: Values of constants used in depositional ice growth model.

Name	Symbol	Value
Density of water	ρ_w	1000 kg/m^3
Density of ice	ρ_i	910 kg/m^3
Molecular mass of water	M_w	18 $\times 10^{-3}$ kg
Thermal deposition coefficient	α_T	1
Latent heat of vaporization	L_v	2.5$\times 10^6$ J/kg
Latent heat of sublimation	L_s	2.837$\times 10^6$ J/kg
Acceleration due to gravity	g	9.81 m/s^2
Universal gas constant	$\mathcal{R}$	8.3144521 J/mole/K
Individual gas constant of air	R_a	287.05 J/kg/K
Individual gas constant of water vapor	R_v	461.51 J/kg/K
Specific heat capacity	c_p	1005
Mean free path of water molecules in air	λ_a	8×10^{-8} m
Vapor jump length	Δ_v	1.3 λ_a
Thermal jump length	Δ_T	2.16×10^{-7} m

where D_v is the diffusivity of water molecules in air, r is the radius of the ice crystal, Δ_r is the vapor jump length, α is the deposition coefficient, w is the molecular speed of water in air, M_w is the molecular weight of water, R is the universal gas constant, and T_a is the air temperature.

The modified thermal conductivity is given by

$$k_a^* = \frac{k_a}{\frac{a}{a+\Delta_T} + \frac{k_a}{a\alpha_T \rho c_p}\left(\frac{2\pi M_a}{\mathcal{R}T_a}\right)^{1/2}} \tag{A5}$$

Eq. A1 relies on a number of constants, which we summarize in Table 2. In addition, we assume that the temperature dependence of saturation vapor pressure with respect to ice (in Pa) is given by [24],

$$e_{sat,i}(T) = exp\left(a_0 - \frac{a_1}{T} + a_2 Log(T) - a_3 T\right) \tag{A6}$$

where $a_0 = 9.550426$, $a_1 = 5723.265$, $a_2 = 3.53068$, and $a_3 = 0.00728332$.

and the diffusivity of water vapor in air (in m^2/s) is given by

$$D_v = 2.11 \times 10^{-5}\left(\frac{T}{T_0}\right)^{1.94}\left(\frac{p_0}{p}\right) \tag{A7}$$

where $T_0 = 273.15$ K and $p_0 = 101325$ Pa [9]. (Valid for -40 to 40 $^\circ C$).

The thermal conductivity of air in Joules is given by

$$k_a = 4.187 \times 10^{-3}(5.69 + 0.017(T - 273.15)) \tag{A8}$$

In the continuum limit for a spherical ice crystal, Eq. A3 reduces to

$$G_c = \left[\frac{\mathcal{R}T_\infty}{e_{sat,i}(T_\infty)D_v M_w} + \frac{L_s}{k_a T_\infty}\left(\frac{L_s M_w}{\mathcal{R}T_\infty} - 1\right)\right]^{-1} \tag{A9}$$

.

A.2 Levitation diffusion chamber data sets

Experimental data was taken in the Button Electrode Levitation (BEL) thermal gradient diffusion chamber and were described in detail in [14, 15, 16]. Charged droplets are initially levitated between two ice-coated parallel plates, with the bottom plate having an opposing direct current voltage, and the top place an alternating current to stabilize particles horizontally. Due to differences in temperature between the warmer, upper plate and colder, lower plate, a supersaturation gradient exists in the chamber as a function of height. After the ice crystal nucleates, the voltage of the bottom plate is automatically adjusted to maintain constant levitation of the ice crystal, with the measured voltage

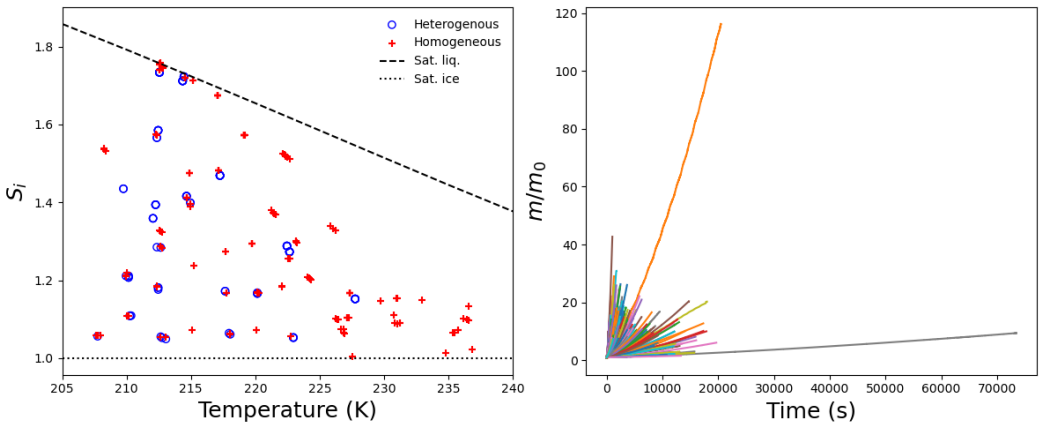

Figure A2: Overview of experimental data sets used in this analysis. Left: Saturation with respect to ice and temperature at which experiments included in this analysis were performed. The dashed line shows the temperature dependence of saturation with respect to liquid water, while the dotted line show saturation with respect to ice. Symbols indicate whether ice crystals were nucleated heterogeneously or homogeneously. **Right:** Mass ratios of all ice crystals as a function of time.

being directly proportional to the ice crystal's mass as a function of time. The initial radius of the particle r_0 can be estimated from Mie theory using the diffraction pattern generated from the scattering of light from a Helium Neon laser, thus determining the initial crystal mass m_0. Deriving the mass as a function of time from the voltage measurement requires the assumption that ice is initially spherical, when it is likely poly-crystalline when it is initially nucleated [25]. Since the uncertainty in derived mass is largely dominated by uncertainty in the initial size of the particle[14], we optimize the NODE model against m/m_0 rather than m.

Time series of the voltage of the lower plate are recorded at a 1 Hz frequency during the course of the experiments. Data sets consist of the measured temperature, pressure, supersaturation, initial ice particle radius observed from light scattering, and the voltage of the lower plate (proportional to the ice particle mass). Temperature uncertainty in the chamber is on the order of $<$1% [14]. Supersaturation uncertainty is estimated to be on the order of 10% [15]. Because supersaturation uncertainty is due to the location of the ice crystal in the chamber, it is likely to show a bias in one direction (typically low) although best estimates for the true value of the supersaturation are used in these data sets [14, 15, 16].

A.3 Synthetic data sets

To evaluate the method of using physics-informed neural networks to discover the functional dependence of ice growth from the mass ratios, we create synthetic data sets with a known depositional growth models. To simulate synthetic data, we start with the observed S_i, T, and m_0 from the 290 experiments that are shown in Figure A2. Given these initial conditions for m_0, S_i, T, we assume depositional ice growth is described by Eq. A1, with the deposition coefficient function given by Nelson and Baker [17] (Eq. A2). We use an ODE solver to integrate Eq. A1 to predict the evolution of the mass ratio for the same length as the experimentally measured time series (up to 500 s). Since the measured mass ratios in the levitation diffusion chamber have high frequency noise, we also use singular spectrum analysis (SSA) to determine realistic noise for the observed mass ratios. SSA decomposes time series into a sum of temporal principal components which account for a decreasing fraction of the variance in the original time series. Here we use a window size of 60 s and assume trailing eigenmodes represent the measurement uncertainty on $\frac{m}{m_0}$. An example of the simulated observations for one time series is given in Figure A3, showing the ODE solution and the SSA reconstruction from the trailing eigenmodes used to estimate measurement uncertainty. All 290 generated synthetic mass time series are shown in Figure A4, which also shows the values for α predicted by Eq. A2 used in generating the synthetic data sets.

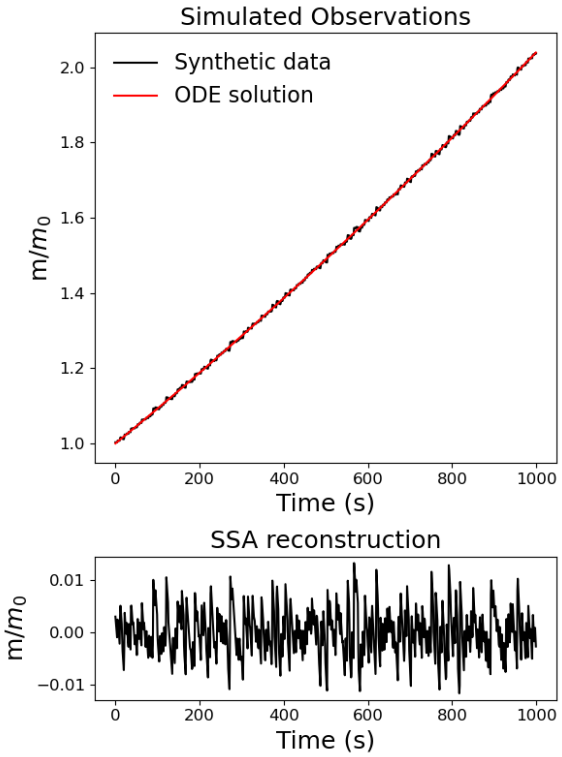

Figure A3: Example of synthetic time series for mass ratios, with measurement uncertainty derived from trailing eigenmodes of SSA analysis for observed time series.

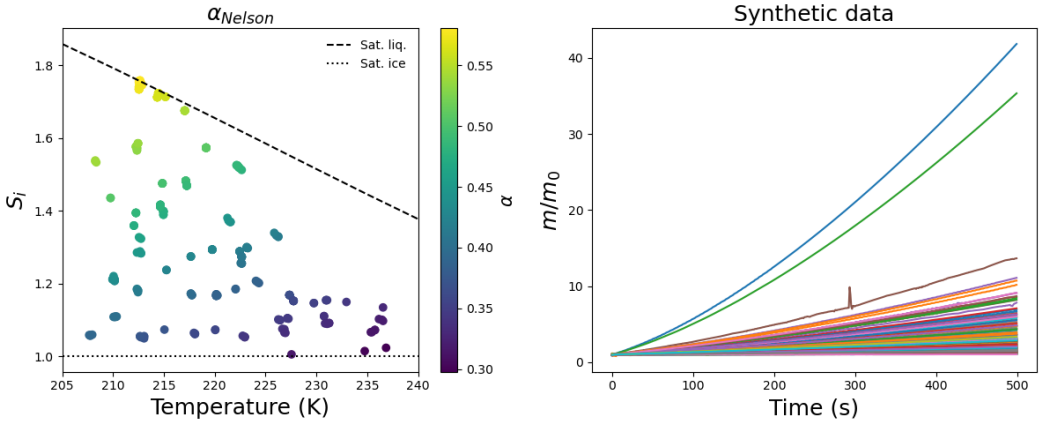

Figure A4: Synthetic data sets. Left: Values for α predicted by Eq. A2 for the S_i and T conditions for the 290 experimental data sets (used as initial conditions in creating the synthetic mass time series). **Right:** Times series of mass ratios for all of the synthetic data.

A.4 NODE Model Predictions

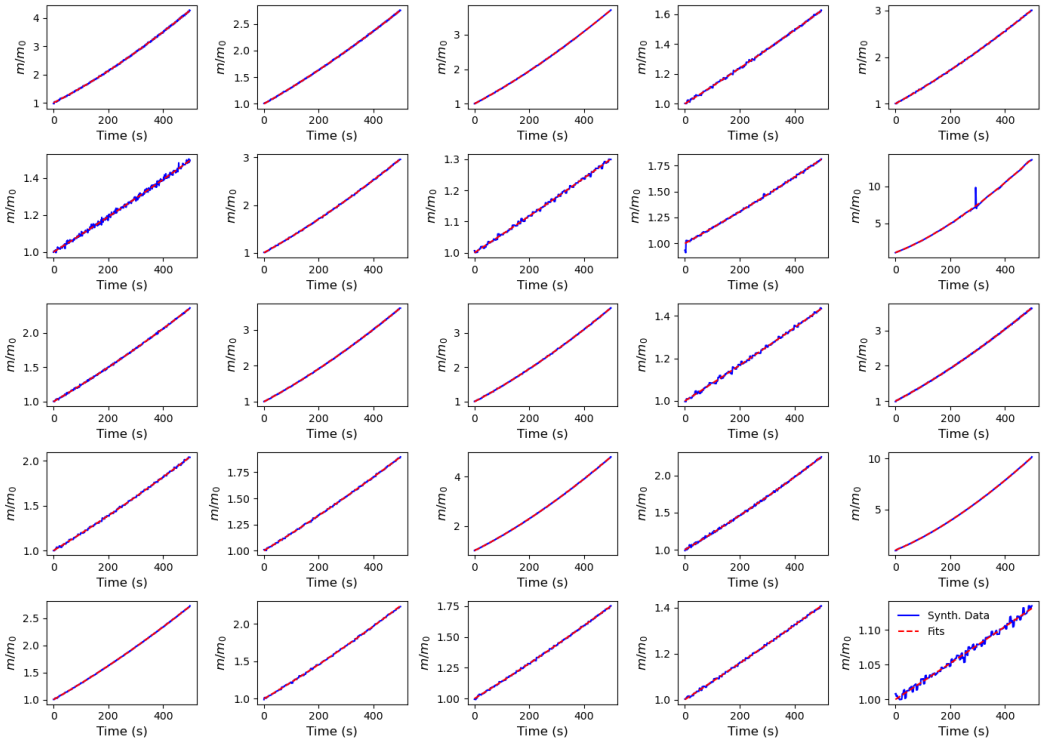

Figure A5: Model predictions for individual experiments for the synthetic data sets. Synthetic mass ratios and strongly-constrained NODE model predictions for a subset of the time series.

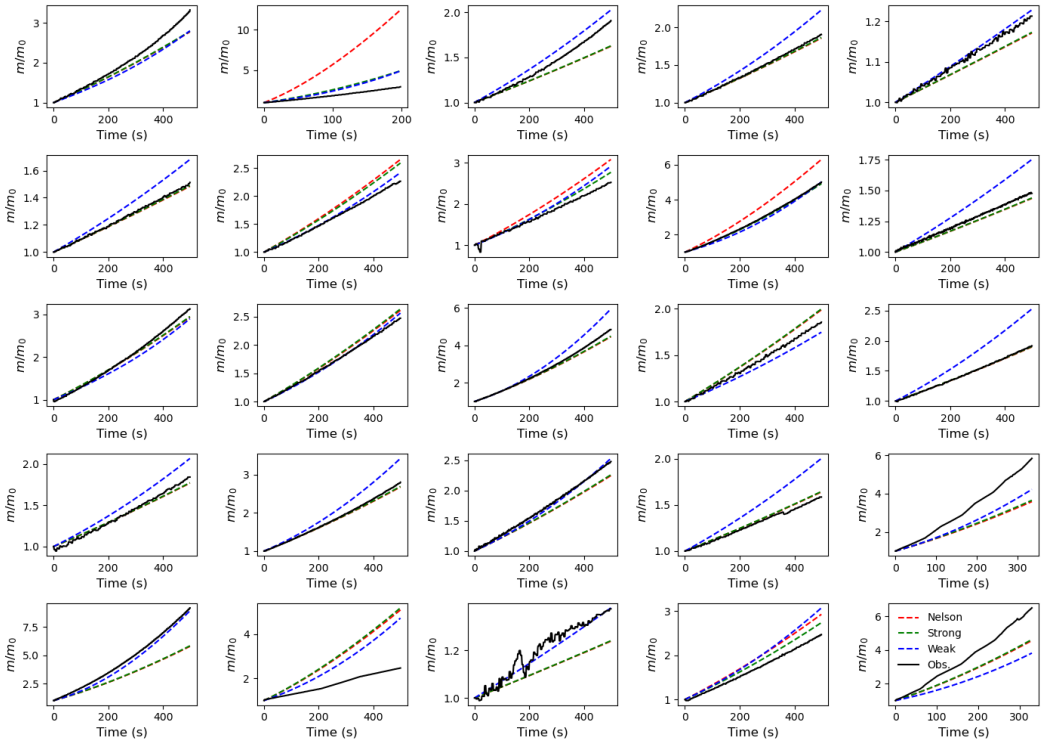

Figure A6: Model predictions for individual experiments for the real data sets. Nelson and Baker model predictions, strongly-constrained NODE model predictions, and weakly-constrained NODE model predictions compared to observations for a subset of the time series.