

Operator learning for models of tear film breakup

QINYING CHEN*, ARNAB ROY*, AND TOBIN A. DRISCOLL

Department of Mathematical Sciences, University of Delaware, Newark, 19716, DE, USA

*Corresponding author: Arnab Roy, arnabroy@udel.edu

[Received on Date Month Year; revised on Date Month Year; accepted on Date Month Year]

Tear film (TF) breakup is a key driver of understanding dry eye disease, yet estimating TF thickness and osmolarity from fluorescence (FL) imaging typically requires solving computationally expensive inverse problems. We propose an operator learning framework that replaces traditional inverse solvers with neural operators trained on simulated TF dynamics. This approach offers a scalable path toward rapid, data-driven analysis of tear film dynamics.

Keywords: Tear film, dry eye disease, fluorescence imaging, scientific machine learning, operator learning

MSC: 92C35

1. Introduction

A thin layer of liquid known as the tear film (TF) lies across the surface of the cornea of a human eye [Doane, 1980]. This dynamic film provides lubrication, protects against microbial invasion, maintains a smooth optical interface, and delivers essential oxygen and nutrients to the cornea [Willcox et al., 2017]. The TF has three layers: a lipid layer that is 20 to 100 nm thick [Braun et al., 2015], an aqueous layer consisting mainly of water [Holly, 1973] that is a few microns thick, and a half-micron thick mucin layer called the glycocalyx that is on the ocular surface [King-Smith et al., 2004]. The lipid layer slows evaporation of water from the TF [Mishima and Maurice, 1961], and a healthy glycocalyx facilitates the fluid movements of the ocular surface [Gipson, 2004]. The lacrimal gland supplies the majority of the water from the aqueous layer near the temporal canthus [Dartt, 2009]. Osmosis supplies water from the ocular epithelia [Braun, 2012].

When the tear film becomes unstable and breaks up, a process known as tear breakup (TBU), it can leave patches of the ocular surface exposed, triggering irritation, inflammation, and visual disturbances [King-Smith et al., 2018]. Persistent or frequent tear breakup is recognized as a major factor in the onset and progression of dry eye disease (DED), a multifactorial condition affecting millions of individuals worldwide that compromises both the comfort and visual function of the eye [Lemp, 2007, Nelson et al., 2017, Stapleton et al., 2017]. One common subtype, evaporative dry eye (EDE), is primarily driven by excessive water loss from the tear film, often due to dysfunction in the lipid layer that normally retards evaporation [Lin and Yiu, 2014, Baudouin et al., 2013]. TBU happens when a dry spot appears on the eye [Norm, 1970] and is often evaporation-driven [Lemp, 2007, Willcox et al., 2017].

Fluorescein (FL) imaging is one of the most widely used techniques for studying tear-film dynamics and tear-film breakup. Following instillation of sodium fluorescein, the aqueous layer of the tear film emits fluorescence whose intensity depends on both tear-film thickness and fluorescein concentration [Webber and Jones, 1986, Nichols et al., 2005]. FL imaging has been used to measure tear-film breakup times, visualize the formation and progression of breakup regions, and investigate tear-film thinning

* These authors contributed equally to this work.

dynamics [Norn, 1970, Begley et al., 2013, Li et al., 2014]. The interpretation of FL images, however, is not straightforward, because the observed intensity depends on both tear-film thickness and fluorescein concentration. In particular, the relationship between fluorescence intensity and tear-film thickness varies between the dilute and self-quenching concentration regimes [Webber and Jones, 1986, King-Smith et al., 2013, Braun et al., 2014]. Previous modeling studies incorporated fluorescein transport and intensity calculations to connect FL image data with underlying tear-film dynamics. Using combined experimental and modeling approaches, these studies showed that evaporation, osmolarity changes, and tangential flow can all influence the appearance of breakup regions in FL images and that advection of fluorescein may complicate direct interpretation of intensity measurements [Braun et al., 2015, 2018, Luke et al., 2020].

Measurement of TF dynamics in situ is challenging. Evaporation has been identified as a primary driver of the dynamics, with models incorporating spatially varying lipid layers, osmolarity transport in the aqueous layer, and osmosis across the cornea [Peng et al., 2014, Braun et al., 2015, 2018]. These studies showed that osmolarity diffusion prevents osmosis from halting thinning, a phenomenon absent in spatially uniform models Braun [2012]. Simpler models later included fluorescein transport, enabling direct comparison with imaging experiments, and one-dimensional partial differential equation (PDE) frameworks have combined evaporation and fluorescein dynamics to estimate otherwise unmeasurable parameters [Zhong et al., 2019, Luke et al., 2020, 2021a]. More recently, spatially independent ODE models have been calibrated to small TBU spots and streaks and coupled with neural-network based data extraction to expand datasets [Luke et al., 2021b, Driscoll et al., 2023]. Driscoll et al. combined reduced ODE models with a convolutional-neural-network-based image analysis pipeline that automatically identified multiple tear-breakup regions in fluorescein video sequences and extracted their intensity time series. This automation substantially increased the number of analyzable breakup events, enabling model fitting and parameter estimation for hundreds of TBU instances across multiple subjects.

Additional theoretical and experimental studies have investigated mechanisms underlying thin-film rupture and instability. Ji and Witelski [2017] analyzed a lubrication-theory thin-film model with modified evaporative loss and used self-similar asymptotic analysis together with high-resolution adaptive finite-difference simulations to characterize evaporation-driven finite-time rupture singularities. Shi et al. [2022] combined lubrication-theory modeling and linear stability analysis to show how interactions between Marangoni and capillary flows can induce symmetry breaking and instability in surfactant-laden thin films.

While the results of Driscoll et al. [2023] are an important first step toward a complete understanding of TF dynamics, they are limited on the one hand by the relative crudeness of the mathematical model and on the other hand by the need to computationally solve an inverse problem for each prospective TBU location. While fully 2D PDE model simulation has been established [Chen et al., 2024] and simple inverse problems for the 2D model are underway for isolated TBU cases [Chen and Driscoll], solving the 2D inverse problem at large scale is currently computationally expensive.

Recent advances in operator learning have produced a variety of architectures for approximating mappings between function spaces. [Tancik et al., 2020] showed that standard multilayer perceptrons exhibit spectral bias toward low-frequency functions and proposed Fourier feature mappings to improve the learning of high-frequency components, significantly enhancing the performance of coordinate-based neural networks on low-dimensional regression tasks. Li et al. [2021] introduced the physics-informed neural operator (PINO), which incorporates PDE constraints into neural operator training to improve accuracy and generalization, particularly when training data are limited. Wang et al. [2021b] analyzed the training dynamics of physics-informed neural networks (PINNs) for multiscale PDEs and proposed Fourier feature embeddings to improve their representation of high-frequency solutions. Chen

and Wu [2023] combined Fourier neural operators with neural ODEs to develop continuous spatial-temporal operator learning that is invariant to both spatial and temporal discretizations and improves long-term prediction stability. Lanthaler [2023] established a rigorous theoretical foundation for PCA-Net by proving universal approximation results and deriving complexity bounds for PCA-based operator learning in infinite-dimensional spaces.

Operator learning has also been successfully applied to a wide range of scientific computing problems. Lu et al. [2021] introduced DeepONet and demonstrated its effectiveness for learning solution operators of nonlinear ODEs and PDEs. Kovachki et al. [2023] showed that neural operators accurately approximate solution operators for benchmark problems including Burgers' equation, Darcy flow, and the Navier–Stokes equations with substantial computational speedups over traditional numerical solvers. Lu et al. [2022] further developed Multifidelity DeepONet, which combines low- and high-fidelity data to efficiently learn nonlinear operators and demonstrated its effectiveness on several forward and inverse PDE problems.

Motivated by these recent advances, we investigate operator learning techniques to ODE and 1D PDE models of TBU dynamics. In operator learning, one trains a deep neural network or other learning architecture to learn the mapping implied by the solution of a differential equation. This could mean the mapping from parameters to solutions, for instance, but our goal is to learn the mapping from observations, in the form of synthetic FL imaging data, to outcomes, in the form of TF thickness and osmolarity functions, without the need to simulate the physics or solve an inverse problem. If successful, this process provides a fast computational method that could be used in real time to predict characteristics of tear film flows from FL imaging. At the modeling level, we can compare the results of predictors trained on ODE and PDE models to clarify how often and under what circumstances they make substantially different predictions.

In section 2 we describe the ODE and PDE models used to simulate TF dynamics during TBU. In section 3 we describe three operator learning methods that are used and tested. In section 4 we present the results of numerical experiments with the two simulation models and operator learning methods, and in section 5 we make concluding remarks and observations.

2. Mathematical models

2.1. One-dimensional PDE

The PDE system is derived from Zhong et al. [2019] and describes the situation depicted in Figure 1 of a radially symmetric glob of lipid on top of the aqueous layer at a later time $t > 0$. The dashed line in the picture represents the tear film at $t = 0$. We assume there is nonzero constant evaporation under the glob and zero evaporation outside the glob. The Marangoni effect, driven by tangential flow, arises in regions where thicker globs of the lipid layer are present. These areas, with higher surfactant concentration, reduce the aqueous/air surface tension and trigger fluid motion along the surface, leading to rapid film breakup. We expect that diffusion, advection, evaporation, osmosis, and Marangoni flow together influence the TF thickness which would result in TBU.

The independent and dependent variables are shown in Table 1. Quantities named without primes are nondimensionalized according to

$$\begin{aligned} r' &= \ell r, & z' &= h'_0 z, & t' &= t'_s t, & h' &= h'_0 h, & u'_r &= U u_r, \\ p' &= \frac{\mu U}{\ell \epsilon^2} p, & \Gamma' &= \Gamma_0 \Gamma, & J' &= \epsilon \rho U J, & c' &= c_0 c, & f' &= f_{\text{cr}} f, & f'_0 &= f_{\text{cr}} f_0 \end{aligned}$$

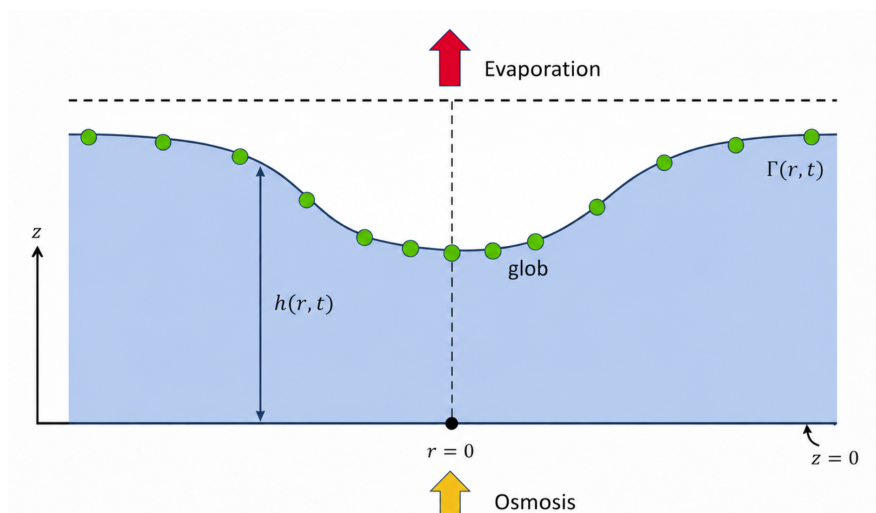


FIG. 1. Mechanisms in the tear film at a later time $t > 0$. Evaporation occurs at the top, osmosis at the bottom; The green globs represent localized excess lipids at the air–tear interface. Γ represents the surfactant concentration, and h represents the tear film thickness. Here, $r = 0$ denotes the axis of symmetry, and $z = 0$ corresponds to the corneal surface.

Variable	Meaning
r	Spot radius
z	Depth dimension
t	Time
$\Gamma(r, t)$	Lipid surfactant concentration
$u_r(r, h, t)$	Horizontal surface fluid velocity
$h(r, t)$	TF thickness
$p(r, t)$	Pressure
$J(r)$	Evaporation rate
$c(r, t)$	Osmolarity
$f(r, t)$	Fluorescein concentration

TABLE 1 *Variables in the 1D glob model.*

where primes denote dimensional quantities that are eventually replaced by nondimensionalized versions. The relevant physical parameters are given in Table 2, and the nondimensional parameters are shown in Table 3.

Parameter	Description	Value	Reference
μ	Viscosity	1.3e-3 Pas	Tiffany [1991]
σ_0	Surface tension	0.045 N/m	Nagyová and Tiffany [1999]
ρ	Density	1000 kg/m ³	Water
A^*	Hamaker constant	$6\pi \cdot 3.5 \times 10^{-19}$ J	Braun et al. [2018]
h_0	Initial TF thickness	2–8 μ m	Calculated
$(\Delta\sigma)_0$	Change in surface tension	2–60 μ N/m	Calculated
t_s	Observation timescale	10–60 s	Calculated
v	Thinning rate	0–40 μ m/min	Calculated
f_0	Initial FL concentration	0.08–0.2 %	Calculated
R_I	Glob radius	0.02–0.2 mm	Calculated
V_w	Molar volume of water	1.8e-5 m ³ /mol	Water
D_f	Diffusivity of fluorescein	0.39e-9 m ² /s	Casalini et al. [2011]
D_o	Diffusivity of salt	1.6e-9 m ² /s	Riquelme et al. [2007]
c_0	Isotonic osmolarity	300 Osm/m ³	Lemp et al. [2011]
P_0	Permeability of cornea	12.1 μ m/s	Braun et al. [2015]
ℓ	Characteristic length, $(t_s \sigma_0 h_0^3 / \mu)^{1/4}$	0.138–0.412 mm	Calculated
U	Characteristic velocity, $(\sigma_0 h_0^3 / \mu t_s^3)^{1/4}$	0.056–0.099 mm/s	Calculated
ε_f	Napierian extinction coefficient	1.75e7 L / m mol	Mota et al. [1991]
f_{cr}	Critical FL concentration	0.2%	Webber and Jones [1986]

TABLE 2 *Physical parameters (dimensional) used in the governing equations.*

Parameter	Description	Expression
ε	Aspect ratio	h'_0 / ℓ
M	Marangoni number	$\varepsilon(\Delta\sigma)_0 [t_s^3 / (\sigma_0 \mu^3 h_0^3)]^{1/4}$
A	Nondimensional Hamaker constant	$A^* / [\varepsilon(\Delta\sigma)_0 h'_0 \ell]$
P_c	Permeability of cornea	$(P_0 V_w c_0) / (\varepsilon U)$
Pe_f	Péclet number for FL diffusion	$U \ell / D_f$
Pe_c	Péclet number for salt ion diffusion	$U \ell / D_o$
Pe_s	Péclet number for surface diffusion	$\varepsilon(\Delta\sigma)_0 \ell / (\mu D_s)$
ϕ	Nondimensional Napierian extinction coefficient	$\varepsilon f_{cr} h'_0$

TABLE 3 *Nondimensional parameters used in the governing equations.*

The resulting nondimensional PDE system for the dependent variables h , p , c , Γ and f is

$$\partial_t h = -J + P_c(c - 1) - \frac{1}{r} \partial_r(rh\bar{u}), \quad (2.1a)$$

$$p = -\frac{1}{r} \partial_r(r \partial_r h) - Ah^{-3}, \quad (2.1b)$$

$$\partial_t \Gamma = \left[\text{Pe}_s^{-1} \left(\frac{1}{r} \partial_r(r \partial_r \Gamma) \right) - \frac{1}{r} \partial_r(r u_r \Gamma) \right] B, \quad (2.1c)$$

$$h(\partial_t c + \bar{u} \partial_r c) = \text{Pe}_c^{-1} \left(\frac{1}{r} \partial_r(r h \partial_r c) \right) + Jc - P_c(c - 1)c \quad (2.1d)$$

$$h(\partial_t f + \bar{u} \partial_r f) = \text{Pe}_f^{-1} \left(\frac{1}{r} \partial_r(r h \partial_r f) \right) + Jf - P_c(c - 1)f, \quad (2.1e)$$

where B is a smooth approximation to a transition step function,

$$B(r, R_I) = \frac{1}{2} + \frac{1}{2} \tanh\left(\frac{r - R_I}{0.1}\right),$$

and the horizontal surface TF fluid velocity $u_r(r, h, t)$ and average horizontal TF fluid velocity $\bar{u}$ throughout the film are given by

$$u_r(r, h, t) = -\frac{\frac{1}{2}h^2(\partial_r p)B + M\partial_r \Gamma h B}{B + (1 - B)h}, \quad (2.2)$$

$$\bar{u} = -\frac{\frac{1}{3}h^2(\partial_r p)[B + \frac{1}{4}h(1 - B)] + \frac{1}{2}M\partial_r \Gamma h B}{B + (1 - B)h}. \quad (2.3)$$

The governing equations are obtained from the incompressible Navier–Stokes equations under the lubrication approximation, using the small aspect ratio of the tear film to reduce the momentum equations to a balance between viscous stresses, capillary pressure gradients, and Marangoni stresses. The pressure is independent of depth at leading order and is determined by the curvature of the free surface,

$$p = -\frac{1}{r}\partial_r(r\partial_r h),$$

Conservation of mass over the film thickness gives the evolution equation

$$\partial_t h = -J + P_c(c - 1) - \frac{1}{r}\partial_r(rh\bar{u}),$$

in which $-J$ represents loss of water by evaporation, $P_c(c - 1)$ represents osmotic flow from the corneal surface into the tear film when the osmolarity is above isotonic, and the divergence term describes redistribution of tear fluid by the depth-averaged radial velocity $\bar{u}$. Thus, localized evaporation and divergent Marangoni-driven flow thin the film, while osmosis and capillary-driven flow may partially oppose thinning. In order to simulate the effect of the lipid glob on evaporation, we define the spatial dependence of the evaporative flux by

$$J(r) = (1 - B(r))v.$$

The depth-averaged velocity $\bar{u}$ and the surface velocity u_r are obtained by solving the leading-order radial momentum equation with no slip at the corneal surface and a mixed tangential stress condition at the tear–air interface. The pressure-gradient terms in these velocities represent capillary-pressure-driven flow, while the terms proportional to $M\partial_r \Gamma$ represent Marangoni flow generated by gradients in surfactant concentration. The surfactant equation

$$\partial_t \Gamma = \left[\text{Pe}_s^{-1} \frac{1}{r} \partial_r(r\partial_r \Gamma) - \frac{1}{r} \partial_r(r u_r \Gamma) \right] B$$

describes surface diffusion and advection of lipid surfactant outside the glob region. The factor $B(r)$ smoothly switches the surfactant dynamics and surface mobility between the immobile glob-covered region and the mobile exposed region. Finally, the osmolarity and fluorescein equations are depth-averaged advection–diffusion equations. In both equations, the left-hand side represents local change

and advection by $\bar{u}$, the first term on the right-hand side represents diffusion, the Jc or Jf term represents concentration increase due to evaporative water loss, and the $-P_c(c-1)c$ or $-P_c(c-1)f$ term represents dilution by osmotic inflow.

Since we are not interested in boundary effects such as eyelids, we assume periodic conditions on all the dependent variables and that all the dependent variables are initially uniform:

$$h(r,0) = c(r,0) = 1, \quad f(r,0) = f_0, \quad (2.4)$$

where f_0 is the FL concentration normalized to the critical concentration f_{cr} . Once h and f are known, we compute the FL intensity via [Webber and Jones, 1986, Braun et al., 2014]

$$I = I_0 \frac{1 - \exp(-\phi f h)}{1 + f^2}, \quad (2.5)$$

where I_0 is a normalization coefficient so that $I(0) = 1$ and ϕ is the nondimensional Napierian extinction coefficient.

The values of h'_0 , t'_s , and f'_0 are determined for each FL trial by the methods described in Wu et al. [2015], and we consider them as being provided to the model. The glob radius R'_l , surface tension change $(\Delta\sigma)'_0$, and nominal thinning rate v' are additional parameters to the model. Hence, there is a 6-dimensional parameter space affecting the output of the model.

For discretization, we use the method of lines with a Fourier spectral collocation method [Trefethen, 2000] in space on a uniform periodic grid and exploit symmetry by solving over $r \in (0, \pi]$. The resulting discretization of spatial terms in (2.1a)–(2.1e) creates a differential–algebraic system (DAE) that is solved in Julia using the Rodas5P solver, a 5th order A-stable stiffly stable Rosenbrock method in the **DifferentialEquations** package [Rackauckas and Nie, 2017]. Computations were performed using Julia version 1.10.5 on a Windows laptop with a 12th Gen Intel Core i7-12700H 2.30GHz processor and 32GB RAM.

2.2. ODE

To simplify the model (2.1), we follow the procedure in Driscoll et al. [2023] to remove the spatial dependence and model only at the glob center. In order to replace the tangential flow caused by surface tension gradients that are no longer possible, we impose a time-dependent shear stress term

$$g(t) = b'_1 e^{-b'_2 t}, \quad (2.6)$$

for dimensional parameters b'_1 , the maximum shear, and $b'_2 \geq 0$, the decay constant. When $b'_1 > 0$, the film thins as though there is outward flow from the center of the glob, while $b'_1 < 0$ causes thickening from an inward flow. Because there can no longer be diffusion, the only difference between osmolarity $c(t)$ and fluorescein concentration $f(t)$ is a constant of proportionality, so f is dropped as a dependent variable. In addition, we assume a constant evaporation rate J'_e . Then we nondimensionalize the parameters and variables according to the following relations:

$$J_e = \frac{J'_e t'_s}{h'_0}, \quad f_0 = \frac{f'_0}{f_{\text{cr}}}, \quad b_1 = b'_1 t'_s, \quad b_2 = b'_2 t'_s$$

where $f_{\text{cr}} = 0.2$ % is the critical fluorescence concentration.

The resulting system, in nondimensional form, is

$$\frac{dh}{dt} = P_c(c - 1) - g(t)h - J, \quad (2.7a)$$

$$h \frac{dc}{dt} = -g(t)hc - \frac{dh}{dt}c, \quad (2.7b)$$

$$h(0) = 1, \quad (2.7c)$$

$$c(0) = 1. \quad (2.7d)$$

This model was labeled as Model D in Driscoll et al. [2023]. Once h and c are computed, $f(t)$ is found via $f(t) = f_0 c(t)$. The FL intensity I is again given by (2.5).

The dimensional parameters for (2.7) again include the initial thickness h'_0 , the trial time t'_s , the initial fluorescein concentration, f'_0 , and the thinning rate J'_e , as in the PDE model. However, the glob radius and surface tension parameters are replaced by the constants b'_1 and b'_2 in (2.6).

3. Operator learning learners

Each of the mathematical models in section 2 effectively maps a parameter space to physically relevant outputs for nondimensional (i.e., relative) thickness h , osmolarity c , and the observed FL intensity I . In practice, our goal is to map an observed time series for $I(t)$ to the thickness and osmolarity functions. Using a mathematical model, this can be treated as an inverse problem, using an observation to find the best-fit parameters of the model, whose solution provides h and c .

Instead, we can use operator learning [Chen and Wu, 2023] to train a machine learning method to produce the $I \mapsto (h, c)$ mapping directly, without solving or even referring to a mathematical model. This approach offers the possibility of using model solutions solely to produce training data. After the training is complete, the learner can make predictions very quickly. In addition, by training learners on different mathematical models, we can assess to what extent, and for which observations, one model produces results similar to the other.

There are many forms of operator learning in the literature, using different ways to represent the input and output spaces as well as the connection between them. We report results for three approaches: Fourier feature networks (FFN), dense neural networks between spaces that are compressed via principal components analysis (Dense-PCA), and dense networks with inputs that enhance PCA inputs with externally known parameters (Dense-PCAX).

3.1. Generation of training/testing data

The ODE and PDE models of section 2 were used to generate two synthetic datasets for training and testing the ML learners. Each begins by sampling parameters using six-dimensional Halton sequences within the ranges given in Table 4. Because it is not straightforward to *a priori* characterize which parameters produce physically and physiologically reasonable results, we reject simulations that have any of the following characteristics:

1. **Upper bound on thickness.** Instances where $h(t) > 1.1$ at any time were discarded, as tear film thickness is generally expected to decrease at the start of simulation time.
2. **Lower bound on thickness.** Solutions where $h(t) < 0.2$ were excluded, since different dynamics will hold when the thickness of the film is comparable to the glycocalyx thickness.

ODE		PDE	
Parameter	Range	Parameter	Range
h'_0	$[2\mu\text{m}, 8\mu\text{m}]$	h'_0	$[2\mu\text{m}, 8\mu\text{m}]$
f'_0	$[0.08\%, 0.2\%]$	f'_0	$[0.08\%, 0.2\%]$
t'_s	$[10\text{s}, 60\text{s}]$	t'_s	$[10\text{s}, 60\text{s}]$
v'	$[0\mu\text{m}/\text{min}, 40\mu\text{m}/\text{min}]$	J'_e	$[0\mu\text{m}/\text{min}, 40\mu\text{m}/\text{min}]$
R'_I	$[0.02\text{mm}, 0.2\text{mm}]$	b_1	$[-0.2\text{s}^{-1}, 2.0\text{s}^{-1}]$
$(\Delta\sigma)'_0$	$[2\mu\text{N}/\text{m}, 60\mu\text{N}/\text{m}]$	b_2	$[0\text{s}^{-1}, 2.0\text{s}^{-1}]$

TABLE 4 *Ranges for dimensional parameters used to create synthetic datasets for ML training. See section 2 for their roles in the mathematical models.*

3. **Nonphysical growth.** Instances with $h(1) > 1.5 \min h(t)$ were removed to rule out nonphysical regrowth after an initial thinning phase. Here $t = 1$ is the nondimensional final time for all cases.
4. **Increasing intensity.** We collect data only with decreasing fluorescence intensity over time, so instances where $I(t)$ increases were discarded.

The surviving cases constitute a parameter set $\Omega \subset \mathbf{R}^6$.

Each element of Ω is associated with resulting solutions for c , h , and I . For the ODE, each variable is a function of time only and was sampled at $N = 601$ equally spaced times; for the PDE, each was sampled at N equally spaced times and $r = 0$, which is the center of the glob. These discretizations produced the realization sets $\mathcal{C}$, $\mathcal{H}$, and $\mathcal{I}$, all subsets of $\mathbf{R}^N$. The elements of $\mathcal{I}$ were used as input feature vectors mapping to corresponding output feature vectors from either $\mathcal{C}$ or $\mathcal{H}$. (It would also be possible to concatenate corresponding elements of $\mathcal{C}$ and $\mathcal{H}$ and train a single learner to produce both outputs simultaneously, but that possibility was not explored for this work.) For the PDE model, $|\mathcal{I}| = 28698$, while for the ODE model, $|\mathcal{I}| = 64522$.

In order to improve the robustness of predictions for noisy experimental observations, we augmented the input set with copies that were randomly perturbed by multiplicative noise. To make perturbations similar to those observed in the experimental dataset, the noise for a signal vector I is a Gaussian-smoothed ($\sigma = 9$) white noise of standard deviation $0.25|I_1 - I_N|$. Two perturbed copies of each input vector (and clean copies of the corresponding outputs) were added to $\mathcal{I}$, which was then split into training and testing subsets $\mathcal{I}_{\text{tr}}$ (75%) and $\mathcal{I}_{\text{te}}$ (25%), respectively, along with the corresponding outputs.

3.2. Fourier feature network (FFN)

A Fourier feature network (FFN) uses an input layer of randomized linear combinations of Fourier modes of the input features to reduce dimensionality while retaining the ability to learn functions with fine-scale variations [Wang et al., 2021a]. Our FFN comprises an input layer, three hidden layers with ReLU activation, and an output layer:

$$\hat{y} = f(\gamma(I)) = W_4 \sigma(W_3 \sigma(W_2 \sigma(W_1 \gamma(I) + v_1) + v_2) + v_3) + v_4.$$

Here, $\gamma(I)$ denotes the Fourier feature mapping, defined as

$$\gamma(I) = [\sin(2\pi\sigma_j I), \cos(2\pi\sigma_j I)]_{j=1}^3,$$

where I is the N -dimensional input vector and $\sigma_j \in \{0.5, 1, 2\}$ represents frequency scales selected to capture varying resolution details. This makes $\gamma(I)$ an array of length $6N$. The hidden layers are dense networks with weight matrices W_i and bias vectors v_i for $i = 1, 2, 3$, and W_4 and v_4 characterize the dense output layer to produce a time series prediction of length N for either c or h .

The first weight matrix W_1 has shape $300 \times 6N$ and the bias vector v_1 has length 300. The next two hidden layers have weight matrices of sizes 200×300 and 200×200 , with corresponding bias vectors. After each of the first three layers, the activation function $\sigma = \text{ReLU}$ is applied. The last weight matrix W_4 of the output layer has size $N \times 200$ to produce a time series of either c or h .

Training was performed by solving the optimization problem

$$\min_{\theta} \mathcal{L}(\theta) = \frac{1}{N} \sum_{t=1}^N (y_t - \hat{y}_t(\theta))^2 + \frac{\alpha}{N-1} \sum_{t=1}^{N-1} (\hat{y}_{t+1}(\theta) - \hat{y}_t(\theta))^2, \quad (3.1)$$

where θ denotes all trainable network parameters (weights and biases), and $\hat{y}(\theta)$ is the network prediction. The second term suppresses high-frequency oscillations, since $\hat{y}$ represents a time series from a dynamical system with smooth solutions. We chose $\alpha = 1.6$ as the regularization parameter.

3.3. Dense neural network with PCA (Dense-PCA)

Another way to encode the input and output time series is to apply PCA to them, connecting the coefficient spaces by a fully-connected neural network [Yadav, 2019]. Reducing dimensionality reduces the number of training parameters and may help focus attention on the most salient features of the signals.

The fully connected neural networks used three hidden layers of 300 neurons with ReLU activation:

$$x = P(I - \mu) \quad (3.2)$$

$$\hat{y} = v + Q[W_4 \sigma(W_3 \sigma(W_2 \sigma(W_1 x + v_1) + v_2) + v_3) + v_4], \quad (3.3)$$

where (P, μ) and (Q, v) denote PCA compression and reconstruction, respectively, and the W_i and v_i are the weights and biases of the dense hidden layers, determined through minimizing the MSE loss between the training samples and their corresponding predictions.

3.4. Dense network on PCA augmented with external parameters (Dense-PCAX)

In Driscoll et al. [2023], the values of the initial thickness h'_0 , initial fluorescein concentration f'_0 , and observation time length t'_s are all determined by processes external to the inverse problem for the ODE, which optimized for the remaining parameters describing the evaporation rate and flow.

Our ODE and PDE models produce training data by varying over both the shared dimensional parameters and the model-specific parameters without making a distinction. In order to determine whether forecasting ability improves when given access to exactly the same data as the ODE inverse problem in Driscoll et al. [2023], we ran a variant of Dense-PCA in which the PCA component vector x in (3.2) was concatenated with the values of h'_0 , f'_0 , and t'_s , which are included as inputs to the learner.

4. Results

As described in section 3, the two mathematical models (Equation 2.1 and Equation 2.7) each generated a synthetic dataset for three ML learners. Selected samples from these datasets are shown in Figure 2.

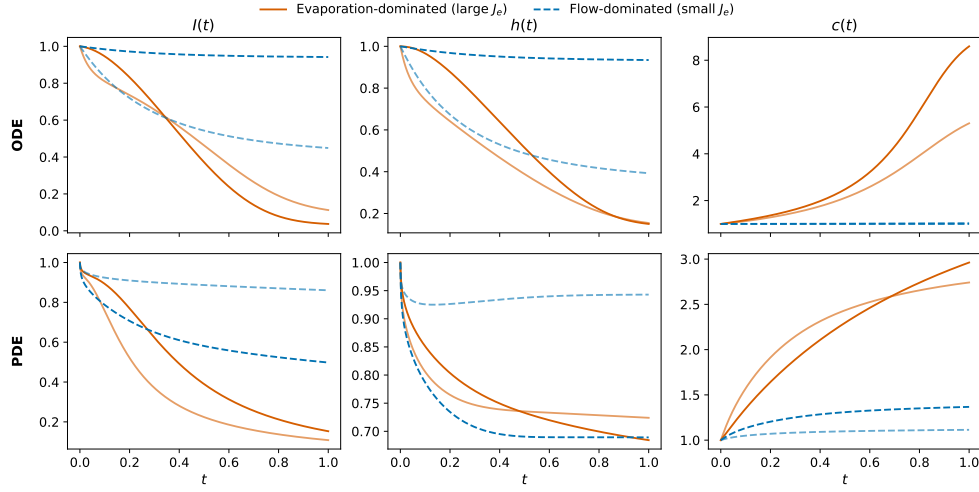


FIG. 2. Selected samples from the synthetic datasets generated by the ODE and PDE models. Each row shows a different sample, with the left column showing the fluorescein intensity $I(t)$, the middle column showing the thickness $h(t)$, and the right column showing the osmolarity $c(t)$. The top three rows are from the ODE model, while the bottom three rows are from the PDE model.

Learner	Input	Output	Hidden layers
FFN	601	601	300, 200, 200
Dense-PCA	4	4 (ODE) / 13 (PDE, c) / 8 (PDE, h)	300, 300, 300
Dense-PCAX	7	4 (ODE) / 13 (PDE, c) / 8 (PDE, h)	300, 300, 300

TABLE 5 Neural network layer sizes used for our results.

For the input training set $\mathcal{S}_{\text{tr}}$, 4 PCA components were sufficient to achieve reconstruction mean squared error (MSE) of approximately 1.4×10^{-4} . For the output training sets, the number of components needed for comparable reconstruction varied by mathematical model and output variable. The sizes of the layers in the ML learners are summarized in Table 5. Because these models are all small by modern standards, the learners were trained for a relatively long time (1000 epochs using the Adam optimizer with a learning rate of 0.001) in order to reach optimal weights.

Figure 3 shows example predictions of the three learners trained on ODE data, while Figure 4 shows them for learners trained on PDE data. These figures show only unperturbed intensity inputs; noisy corruptions of them are also part of the testing set. Across all the cases, the FFN predictions show a great deal of high-frequency oscillation, which is not characteristic of the true solutions, but these predictions can accurately follow the true curves. The PCA-based learners always produce smooth predictions.

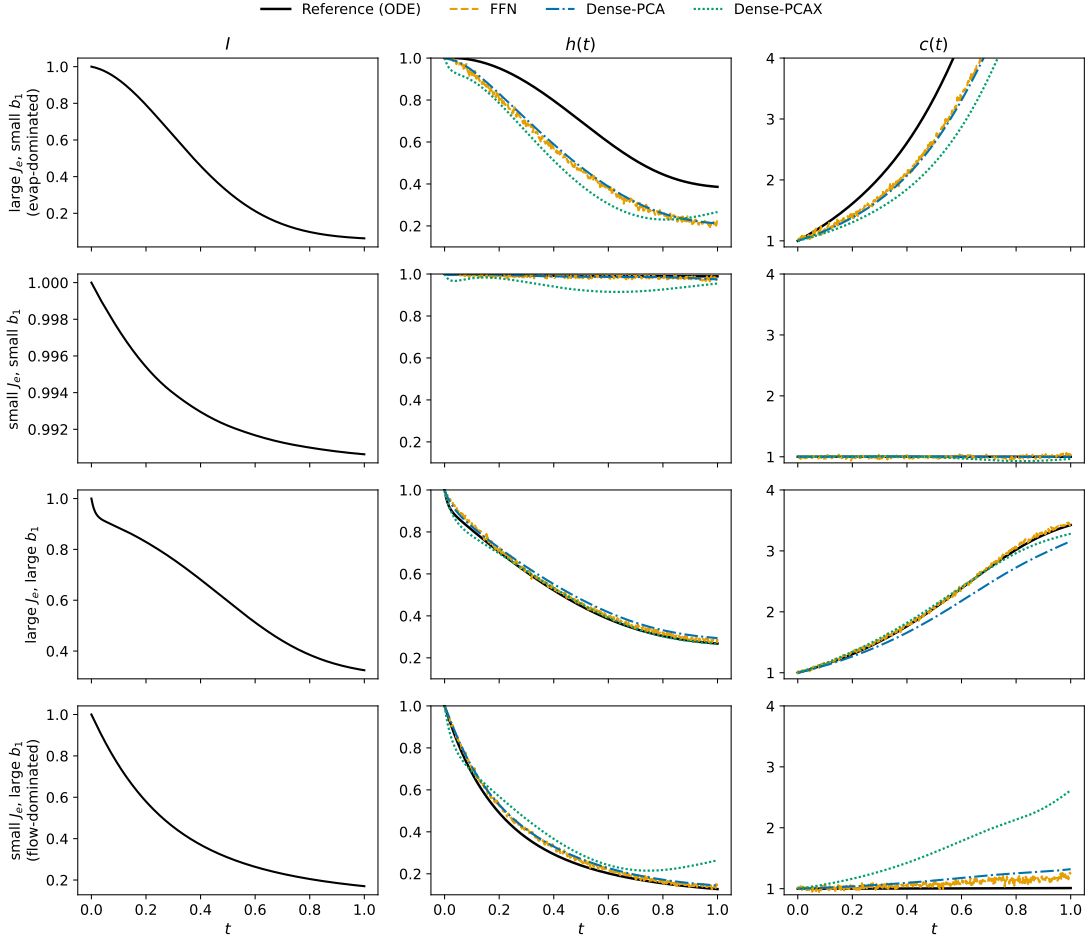


FIG. 3. Example predictions on synthetic data by the three operator learners trained on ODE data.

To summarize our results over the full training sets, we report distributions of relative root-mean-square error (rRMSE), defined for a prediction $\hat{y}$ of time series vector y by

$$\text{RMSE}_y = \left[\frac{1}{N} \sum_{k=1}^N (y_k - \hat{y}_k)^2 \right]^{1/2},$$

$$\text{rRMSE}_y = \frac{\text{RMSE}_y}{\left(\frac{1}{N} \sum_{k=1}^N y_k^2 \right)^{1/2}}.$$

4.1. Synthetic data

Figure 5 shows the distributions of rRMSE of both h and c for all three ODE-trained learners on the synthetic test set. Thickness h was predicted to at least one digit of accuracy in nearly all cases, as was c

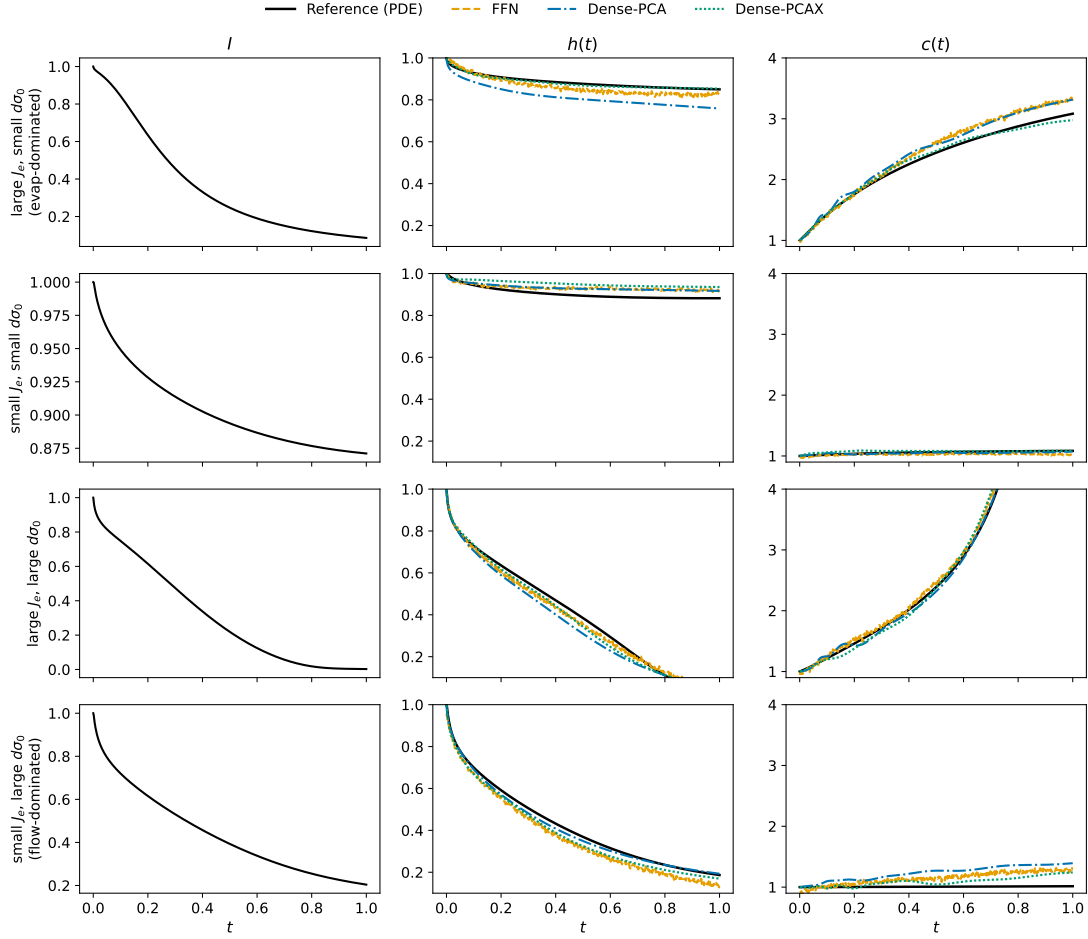


FIG. 4. Example predictions on synthetic data by the three operator learners trained on PDE data.

for FFN and Dense-PCA. The differences between FFN and Dense-PCA are fairly small; while Dense-PCA has a better mean value for predicting osmolarity, the difference is small compared to the MAD spreads. The Dense-PCAX predictions are more centrally concentrated but substantially less accurate in mean for both quantities, even relative to the spreads.

Figure 6 compares all the individual ODE-trained predictions and true values of the final-time thickness $h(1)$ and osmolarity $c(1)$. For the prediction of $h(1)$, all three learners perform well, with Dense-PCAX exhibiting the tightest clustering about the diagonal, followed closely by FFN, while Dense-PCA shows slightly greater scatter. For the prediction of $c(1)$, FFN and Dense-PCA remain accurate across most of the range, whereas the poorer performance of Dense-PCAX is primarily due to underestimation when $c(1) > 8$.

Figure 7 shows the rRMSE values of osmolarity and thickness of the PDE-trained learners on the PDE test data. As with the ODE learners, the FFN and Dense-PCA distributions are similar to each

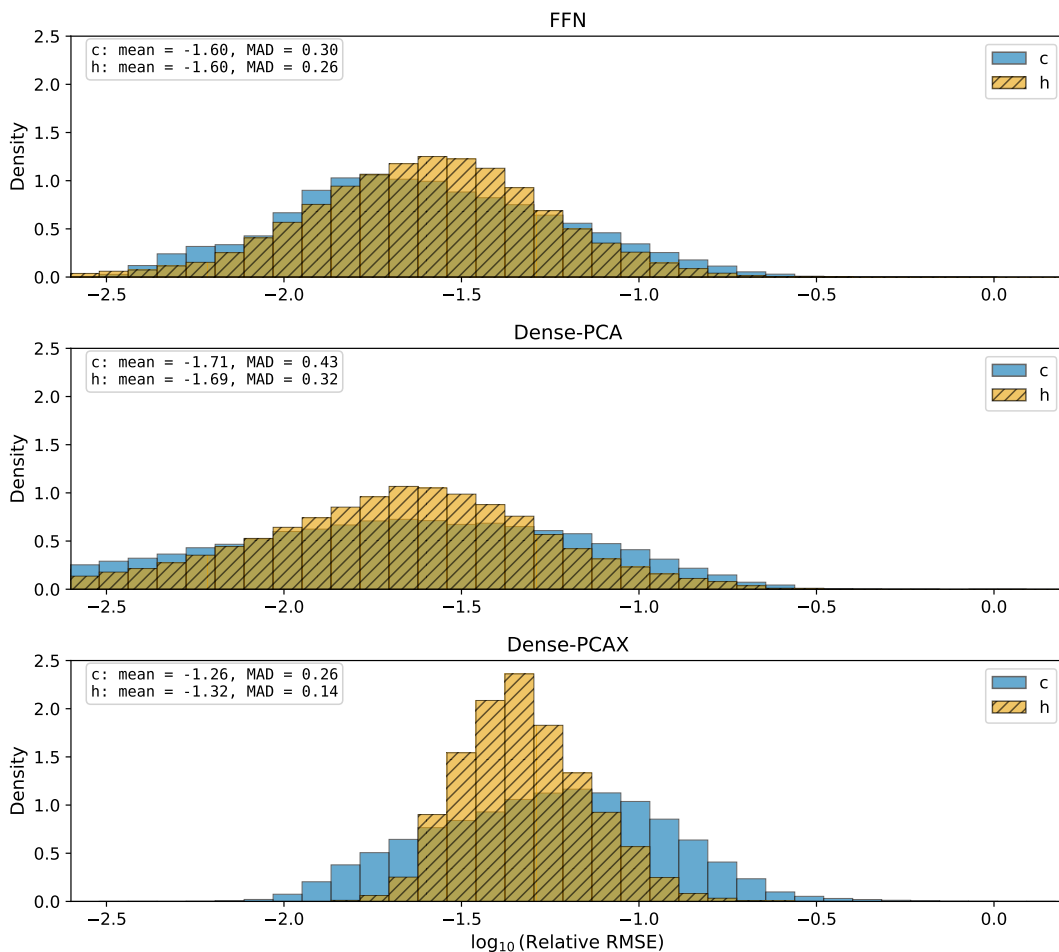


FIG. 5. Relative RMSE for synthetic testing of ML learners trained by ODE data. The mean and mean absolute deviation (MAD) is shown for each distribution. Top to bottom: FFN, Dense-PCA, Dense-PCAX.

other. Here, however, the Dense-PCAX learners do a bit better than the others on average, though by less than the MAD spread. Figure 8 shows analogous results for the synthetic PDE test set. Again, all three learners accurately predict the final thickness, although Dense-PCA is slightly more scattered than FFN and Dense-PCAX. For the final osmolarity, Dense-PCA shows the greatest variance, while FFN and Dense-PCAX remain more tightly clustered around the diagonal. The prediction task is most difficult for all three learners when $c(1)$ is large.

4.2. Experimental data

The experimental data used in this study were collected at Indiana University as part of a study approved by the Biomedical Institutional Review Board [Awisi-Gyau et al., 2019]. Fluorescein imaging data were obtained from 25 participants. Prior to enrollment, all subjects completed a screening procedure

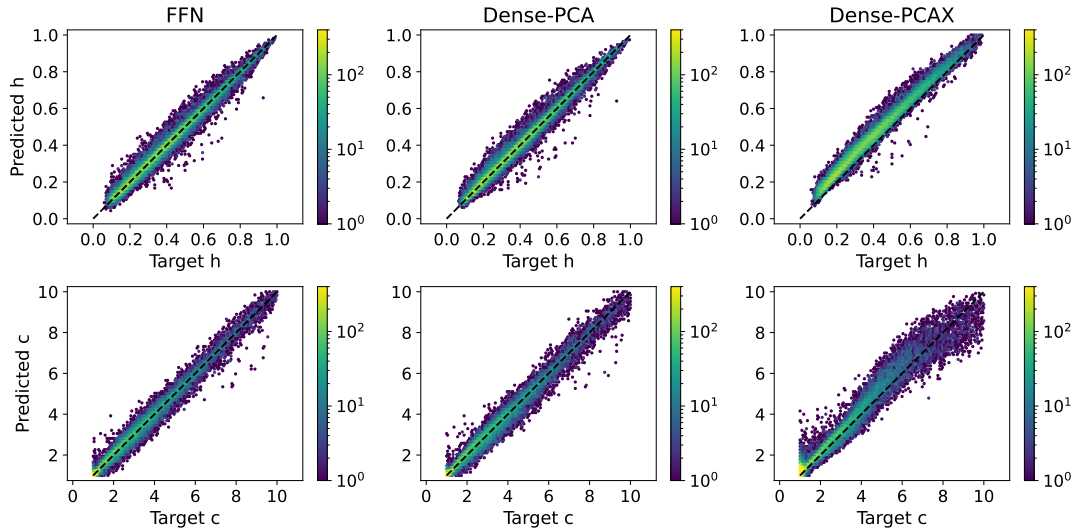


FIG. 6. Comparisons of true and predicted final values of thickness (first row) and osmolarity (second row) for ODE-trained learners on all the synthetic ODE testing data. The color map shows the number of predictions in each bin.

and subsequently attended two study visits, during which ten imaging trials were performed at each visit. Individuals who wore contact lenses or had received a clinical diagnosis of dry eye disease were excluded from participation.

Each trial consisted of a sequence of ocular surface images acquired at intervals of either 0.2 or 0.25 seconds, depending on the camera frame rate. At the start of each trial, a 2 microliter drop of 2% sodium fluorescein solution was instilled into the eye [Carlson et al., 2016]. The tear film was illuminated using a cobalt blue excitation filter (494 nm), and the emitted fluorescence was captured through a Wratten no. 8 yellow barrier filter positioned along the imaging axis. Under this illumination, fluorescein within the tear film emitted green light with a wavelength of approximately 521 nm, which was recorded by the imaging system [Lakowicz, 2006].

To begin the recording, subjects were instructed to blink three times in order to distribute the fluorescein uniformly across the tear film. During these initial blinks, the illumination intensity was kept low, and a custom MATLAB algorithm Wu et al. [2015] was used to estimate the initial fluorescein concentration, which was assumed to be spatially uniform over the corneal surface. Following the third blink, the illumination was increased to a predetermined high-intensity setting, and subjects were asked to keep their eyes open for as long as possible. Data collection continued until the next blink, which marked the end of the trial.

Driscoll et al. [2023] fit 467 intensity time series of presumed TBU to the ODE model using traditional inverse problem optimization, thereby producing solutions h and c consistent with the observed intensity. Our ML models are meant to accomplish the same predictions without reference to the mathematical models once trained. Note that only cases for which a good ODE fit was found were reported in that work, introducing a form of bias, and that the possible existence of fits to I nearly as good with different predictions of h and c is unknown. Moreover, while the ODE-trained predictors have learned the dynamics from the same mathematical models as in Driscoll et al. [2023], the PDE-trained predictors presumably have a more physically realistic picture of the dynamics. Thus, while we

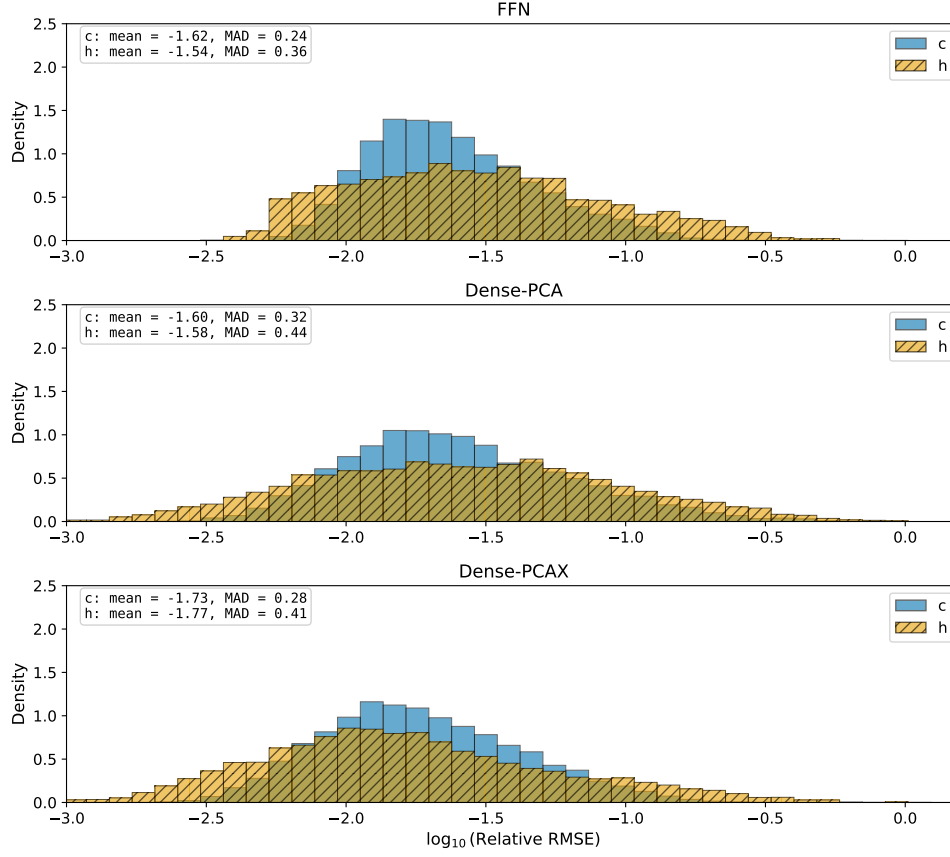


FIG. 7. Relative RMSE for synthetic testing of ML learners trained by PDE data. The mean and mean absolute deviation (MAD) is shown for each distribution. Top to bottom: FFN, Dense-PCA, Dense-PCAX.

continue to report “errors” in the following results with respect to Driscoll et al. [2023] as reference values, they might better be described as “discrepancies” for the PDE predictions.

Figure 9 shows example predictions of all the learners for two of the experimental intensity inputs. Compared to predictions on synthetic data as seen in Figure 3 and Figure 4, the predictions show more variance and less accuracy, particularly for thickness prediction, though conclusions cannot be drawn from only two examples.

Figure 10 shows the rRMSE distributions of predictions by the ODE-trained learners on the experimental data. For these predictions, the Dense-PCAX is a little better than the others on average for osmolarity, but by an amount less than the MAD. The thickness distribution has less spread but is no better on average. Figure 11 shows the distributions for PDE-trained learners. While their shapes vary, the quantitative aggregate differences are not large.

Figure 12 shows the dependence of rRMSE on the ODE fit parameters b_1 and J_e , which are proxies for shear flow rate and evaporation rate. The ODE-trained predictors show the greatest discrepancies in osmolarity for large $b_1 > 0$ and small J_e —that is, flow-dominated thinning—while the greatest

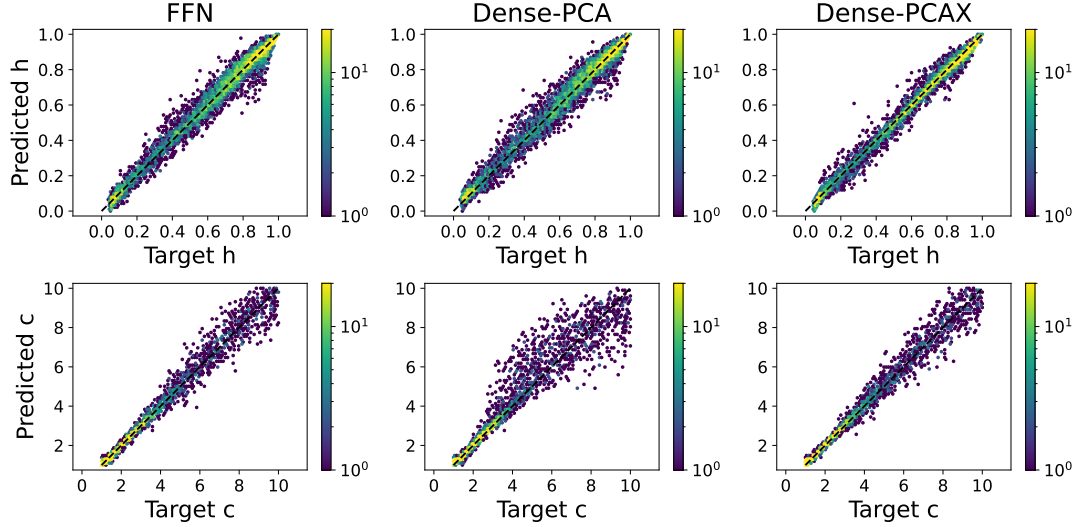


FIG. 8. Comparisons of true and predicted final values of thickness (first row) and osmolarity (second row) for PDE-trained learners on all the synthetic PDE testing data. The color map shows the number of predictions in each bin.

discrepancies in thickness are for large J_e and $b_1 < 0$, which indicates strong evaporation with resultant inward flow. For the PDE-trained predictor, osmolarity shows the same pattern as for the ODE-trained learner, while thickness discrepancy is largest for moderate outward flow over a wide range of evaporation rates.

Finally, Figure 13 shows the predictions of minimum thickness and maximum osmolarity for the reference data together with the predictions from the best-performing learner for each test instance, selected separately from the ODE-trained and PDE-trained models. Among the ODE-trained learners, Dense-PCA was selected most frequently. Among the PDE-trained learners, Dense-PCAX was selected most frequently for predicting maximum osmolarity, whereas Dense-PCA was selected most frequently for predicting minimum thickness. The distribution of minimum thickness is similar for the ODE predictor and reference, but the reference predicts far more cases of maximum osmolarity staying close to 1. The PDE predictor is much more likely to predict minimum thickness greater than 0.75 and a moderately larger maximum osmolarity on average.

5. Discussion

For this study, we used operator learning for the map between time series of the FL intensity and the film thickness and osmolarity. These outputs have clear physical significance, while the parameters of a particular mathematical model have a more indirect interpretation. However, one could instead train for the functional mapping to model parameter values and then simulate the model to obtain the outputs, as in a traditional inverse method, if the parameter values themselves are of interest.

In a broad sense, the primary task of learning the operator is to distinguish between thinning due to evaporation and thinning due to flow. Evaporative thinning tends to be slow—modeled mathematically as linear decrease—while thinning due to shear flow is exponential in nature, but the rates need not

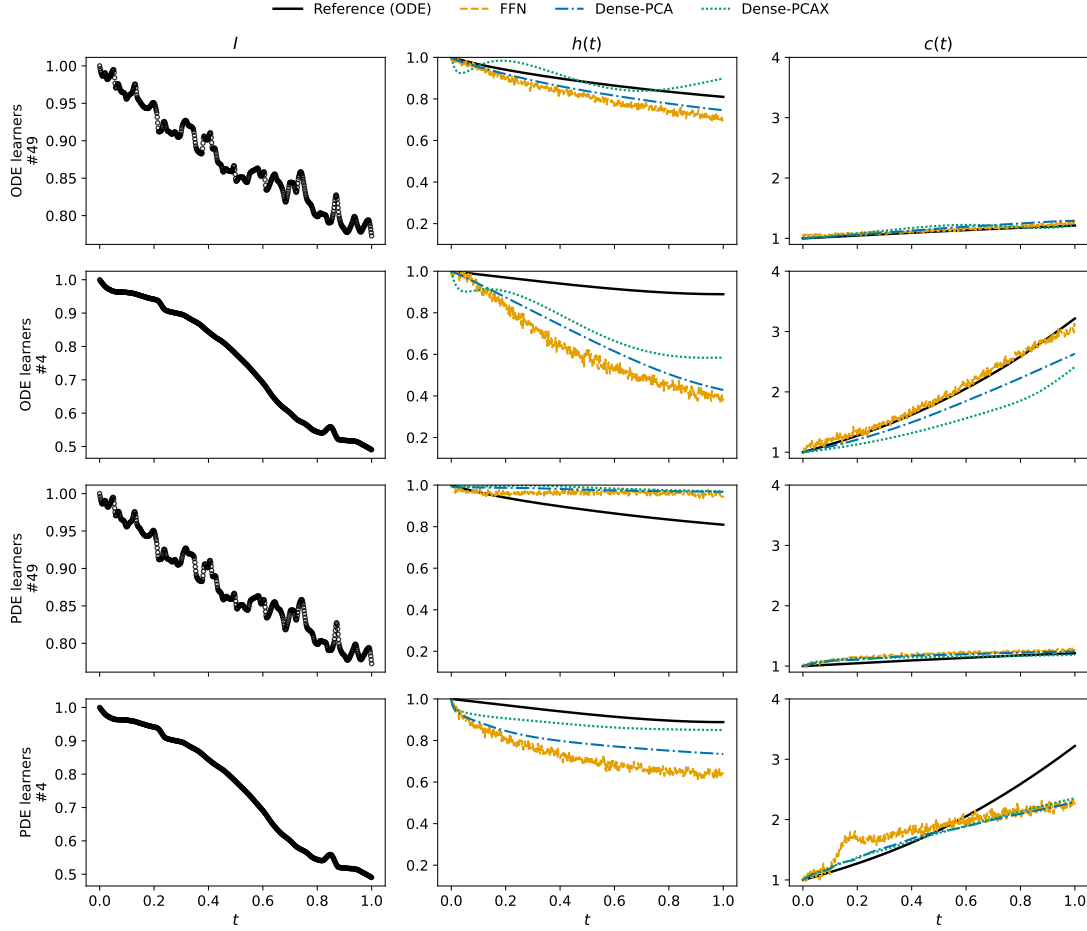


FIG. 9. Example predictions on two time series from the experimental data by the ODE (top two rows) and PDE (bottom two rows) learners.

be clearly distinct in every case. Distinguishing between these effects is complicated by the interaction between h and c in producing the FL intensity I .

Multiple operator learning methods were trained on datasets comprising simulations of ODE models and of PDE models. Performance, as measured on synthetic test data, did not depend strongly on the particular learning method or architecture used: 1–2 digits of rRMSE accuracy were found in most cases. Given the weak dependence on the learning method, we speculate that the accuracy of the predictions is limited mainly by the inherent identifiability of the dependent quantities from FL data presented in this form.

Compared to the inverse problem predictions of Driscoll et al. [2023] on experimental data, the ML predictions based on the same ODE model are much less likely to predict minor change in osmolarity simultaneously with substantial thinning, and large osmolarity simultaneously with little thinning.

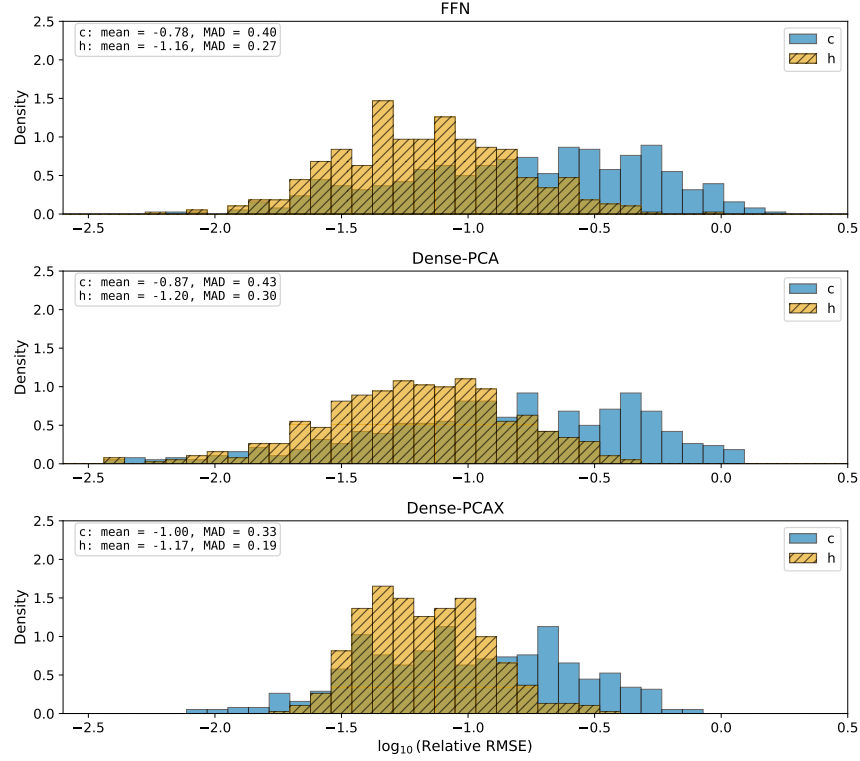


FIG. 10. Relative RMSE for ODE-trained learners on the experimental data.

Differences in osmolarity predictions appear most often in what [Driscoll et al. \[2023\]](#) describes as flow-dominated, zero-evaporation instances, while differences in thickness predictions also occur for strongly evaporative instances. One possibility is that the standard ODE inverse problem can be more easily induced to produce solution types that play a small role in the full solution set. Since the underlying physical realities are unknown, we cannot determine whether such results represent appropriate inverse problem responses or a manifestation of overfitting in the standard inverse problem.

On experimental data, the PDE-trained learners predict much less thinning on average than either of the ODE-based predictors do over a wide range of evaporation rates. This suggests a fundamental difference between the ODE and PDE models of the phenomenon. While the PDE predictions have even fewer cases than the ODE learners of osmolarity remaining very close to 1, they do make more predictions between 1.25 and 1.75 and thus are distributed more like the inverse problem reference results. It might be informative to compare PDE operator learning to traditional inverse-problem methods for the same PDE on the data, but that project is beyond the scope of this study.

It is plausible that PDE-trained learners would perform better if the training data and observations included information from more than a single location on the cornea. This would substantially increase the size of inputs and outputs to the learners, increasing training times. We also note that the learners in this work have no constraints from physics or dynamics other than what is implied by the training data. Methods that can incorporate such information, such as physics-informed neural

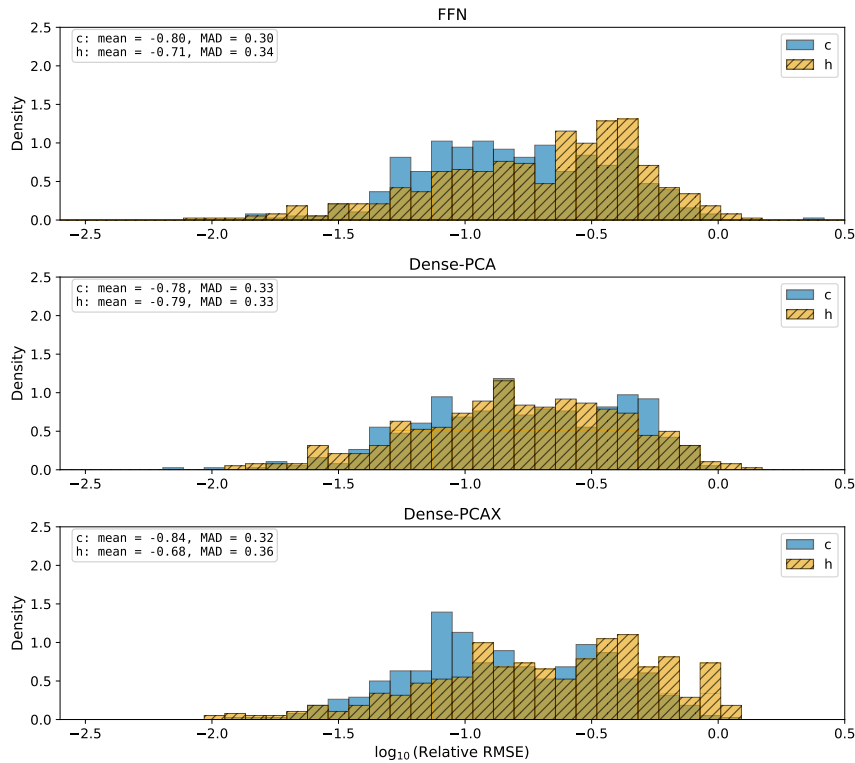


FIG. 11. Relative RMSE for PDE-trained learners on the experimental data.

nets [Cuomo et al., 2022, Li et al., 2023], neural ODEs [Chen et al., 2019, Kidger, 2022], and universal differential equations [Rackauckas et al., 2020] might be expected to perform better. Finally, models and experimental observations that included additional variables might show better results. In particular, thermal imaging is experimentally practical and could give a strong signal through evaporative cooling that is independent of flow.

Conflict of Interest. The authors declare no competing interests.

REFERENCES

- D. Awisi-Gyau, C. G. Begley, P. Situ, and T. L. Simpson. Changes in corneal detection thresholds after repeated tear film instability. *Invest. Ophthalmol. Vis. Sci.*, 60(13):4234–4240, 2019.
- C. Baudouin, P. Aragona, E. M. Messmer, A. Tomlinson, M. Calonge, K. G. Boboridis, Y. A. Akova, G. Geerling, M. Labetoulle, and M. Rolando. Role of hyperosmolarity in the pathogenesis and management of dry eye disease: Proceedings of the OCEAN group meeting. *Ocul. Surf.*, 11:246–258, 2013.
- C. G. Begley, T. L. Simpson, H. Liu, E. Salvo, Z. Wu, A. Bradley, and P. Situ. Quantitative analysis of tear film fluorescence and discomfort during tear film instability and thinning. *Invest. Ophthalmol. Vis. Sci.*, 54(4): 2645–2653, 2013.
- R. J. Braun. Dynamics of the tear film. *Annu. Rev. Fluid Mech.*, 44(1):267–297, 2012.

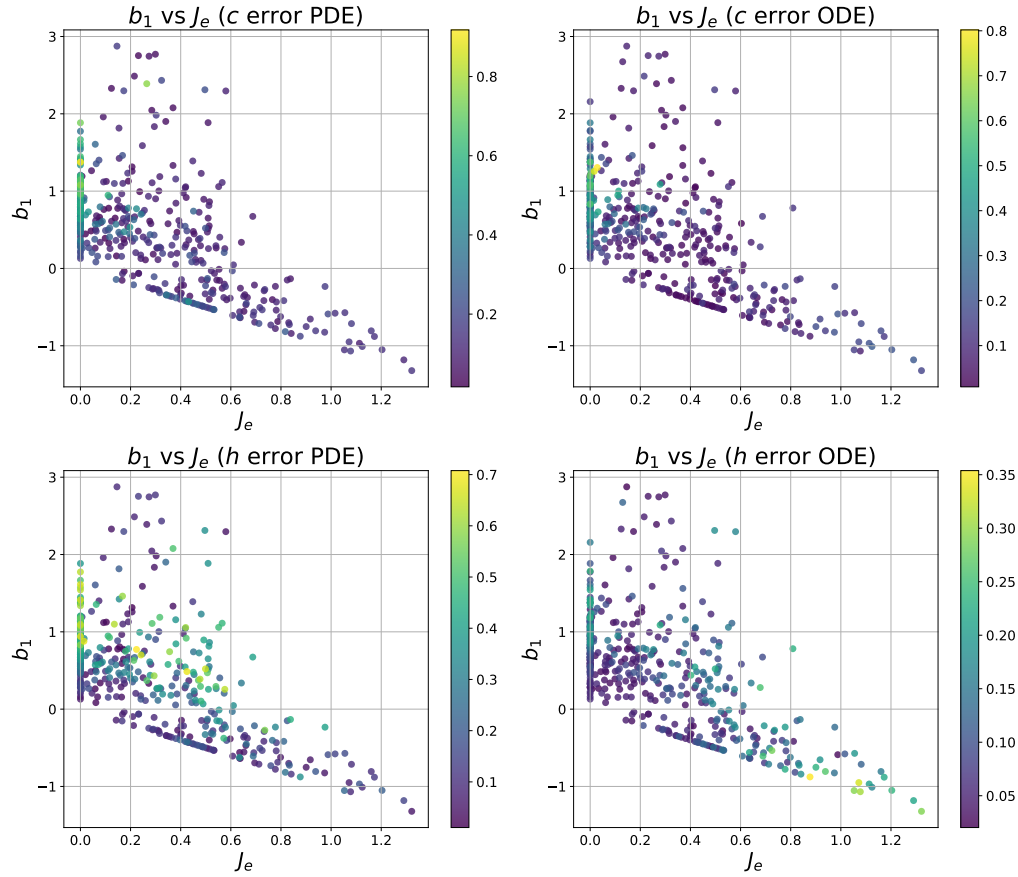


FIG. 12. By rows: osmolarity error to thickness error. By columns: PDE Dense-PCAX learner to ODE Dense-PCAX learner. Color indicates the error comparing to the calculated values, shown as functions of the nondimensional ODE parameters J_e , b_1 .

- R. J. Braun, N. R. Gewecke, C. G. Begley, P. E. King-Smith, and J. I. Siddique. A model for tear film thinning with osmolarity and fluorescein. *Invest. Ophthalmol. Vis. Sci.*, 55(2):1133–1142, 2014.
- R. J. Braun, P. E. King-Smith, C. G. Begley, L. Li, and N. R. Gewecke. Dynamics and function of the tear film in relation to the blink cycle. *Prog. Ret. Eye Res.*, 45:132–164, 2015.
- R. J. Braun, T. A. Driscoll, C. G. Begley, P. E. King-Smith, and J. I. Siddique. On tear film breakup (TBU): Dynamics and imaging. *Math. Med. Biol.: A Journal of the IMA*, 35(2):145–180, June 2018.
- N. B. Carlson, D. Kurtz, and C. Hines. *Clinical procedures for ocular examination*. McGraw-Hill Education, 2016.
- T. Casalini, M. Salvalaglio, G. Perale, M. Masi, and C. Cavallotti. Diffusion and aggregation of sodium fluorescein in aqueous solutions. *J. Phys. Chem. B*, 115(44):12896–12904, 2011.
- C. Chen and J.-L. Wu. Operator learning for continuous spatial-temporal model with a hybrid optimization scheme, Nov. 2023.
- Q. Chen and T. A. Driscoll. Parameter estimation for an evaporation-driven tear film model in two space dimensions. Under review.
- Q. Chen, T. A. Driscoll, and R. J. Braun. Evaporation-driven tear film thinning and breakup in two space dimensions. *Journal of Engineering Mathematics*, 149(1):5, Dec. 2024. ISSN 0022-0833, 1573-2703. doi:

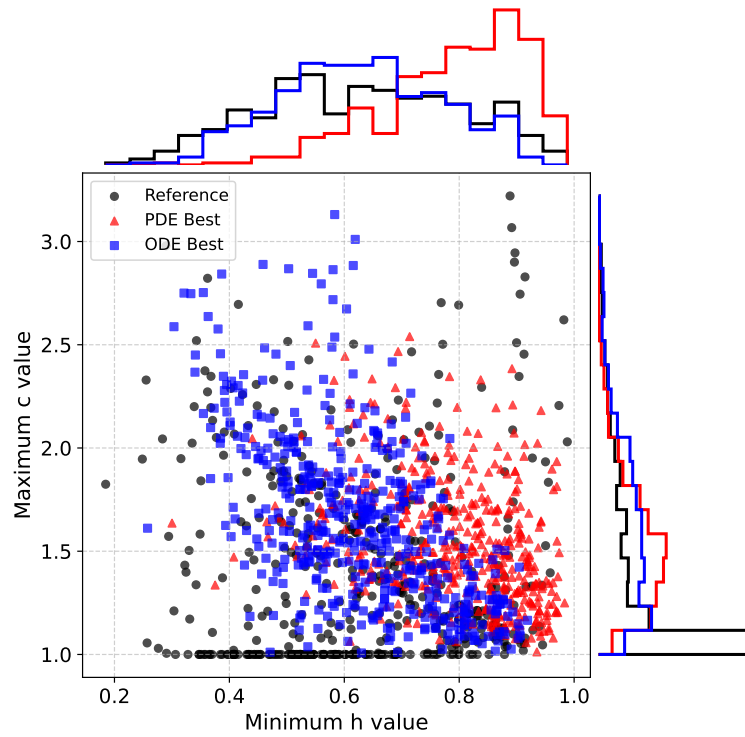


FIG. 13. Predictions of maximum osmolarity and minimum thickness for Driscoll et al. [2023] (black circles) and two ML learners (red triangles for PDE training, blue squares for ODE training). The top and right histograms show the marginal distributions of thickness and osmolarity, respectively.

- 10.1007/s10665-024-10407-6.
- R. T. Q. Chen, Y. Rubanova, J. Bettencourt, and D. Duvenaud. Neural ordinary differential equations, Dec. 2019.
- S. Cuomo, V. S. Di Cola, F. Giampaolo, G. Rozza, M. Raissi, and F. Piccialli. Scientific machine learning through physics-informed neural networks: Where we are and what's next. *Journal of Scientific Computing*, 92(3):88, Sept. 2022. ISSN 0885-7474, 1573-7691. doi: 10.1007/s10915-022-01939-z.
- D. A. Dartt. Neural regulation of lacrimal gland secretory processes: Relevance in dry eye diseases. *Prog. Ret. Eye Res.*, 28(3):155–177, 2009.
- M. G. Doane. Interaction of eyelids and tears in corneal wetting and the dynamics of the normal human eyeblink. *Am. J. Ophthalmol.*, 89:507–516, 1980.
- T. A. Driscoll, R. J. Braun, R. A. Luke, D. Sinopoli, A. Phatak, J. Dorsch, C. G. Begley, and D. Awisi-Gyau. Fitting ODE models of tear film breakup. *Model. Artif. Intell. Ophthalmol.*, 5:1–36, 2023.
- I. K. Gipson. Distribution of mucins at the ocular surface. *Exp. Eye Res.*, 78(3):379–388, 2004.
- F. J. Holly. Formation and rupture of the tear film. *Exp. Eye Res.*, 15(5):515–525, 1973.
- H. Ji and T. P. Witelski. Finite-time thin film rupture driven by modified evaporative loss. *Physica D: Nonlinear Phenomena*, 342:1–15, 2017.

- P. Kidger. On neural differential equations, Feb. 2022.
- P. King-Smith, K. Reuter, R. Braun, J. Nichols, and K. Nichols. Tear film breakup and structure studied by simultaneous video recording of fluorescence and tear film lipid layer, TFLL, images. *Invest. Ophthalmol. Vis. Sci.*, 54:4900–4909, 2013.
- P. E. King-Smith, B. A. Fink, R. M. Hill, K. W. Koelling, and J. M. Tiffany. The thickness of the tear film. *Curr. Eye Res.*, 29(4-5):357–368, 2004.
- P. E. King-Smith, C. G. Begley, and R. J. Braun. Mechanisms, imaging and structure of tear film breakup. *The ocular surface*, 16(1):4–30, 2018.
- N. Kovachki, Z. Li, B. Liu, K. Azizzadenesheli, K. Bhattacharya, A. Stuart, and A. Anandkumar. Neural operator: Learning maps between function spaces. *Journal of Machine Learning Research*, 24(89):1–97, 2023.
- J. Lakowicz. *Principles of Fluorescence Spectroscopy*, volume 1. 01 2006. ISBN 978-0-387-31278-1. doi: 10.1007/978-0-387-46312-4.
- S. Lanthaler. Operator learning with PCA-Net: upper and lower complexity bounds. *Journal of Machine Learning Research*, 24(318):1–67, 2023. URL <http://jmlr.org/papers/v24/23-0478.html>.
- M. A. Lemp, A. J. Bron, C. Baudouin, J. M. Benítez Del Castillo, D. Geffen, J. Tauber, G. N. Foulks, J. S. Pepose, and B. D. Sullivan. Tear osmolarity in the diagnosis and management of dry eye disease. *Am.J. Ophthalmol.*, 151(5):792–798.e1, 2011.
- MA. Lemp. The definition and classification of dry eye disease: Report of the Definition and Classification Subcommittee of the International Dry Eye WorkShop (2007). *Ocul. Surf.*, 5(2):75–92, Apr. 2007.
- L. Li, R. J. Braun, K. L. Maki, W. D. Henshaw, and P. E. King-Smith. Tear film dynamics with evaporation, wetting, and time-dependent flux boundary condition on an eye-shaped domain. *Phys. Fluids*, 26(5):052101, 2014.
- Z. Li, N. Kovachki, K. Azizzadenesheli, B. Liu, K. Bhattacharya, A. Stuart, and A. Anandkumar. Fourier neural operator for parametric partial differential equations, May 2021.
- Z. Li, H. Zheng, N. Kovachki, D. Jin, H. Chen, B. Liu, K. Azizzadenesheli, and A. Anandkumar. Physics-informed neural operator for learning partial differential equations, July 2023.
- H. Lin and S. C. Yiu. Dry eye disease: A review of diagnostic approaches and treatments. *Saudi J. Ophthalmol.*, 28(3):173–181, 2014.
- L. Lu, P. Jin, G. Pang, Z. Zhang, and G. E. Karniadakis. Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators. *Nature Machine Intelligence*, 3(3):218–229, 2021. doi: 10.1038/s42256-021-00302-5.
- L. Lu, R. Pestourie, S. G. Johnson, and G. Romano. Multifidelity deep neural operators for efficient learning of partial differential equations with application to fast inverse design of nanoscale heat transport. *Physical Review Research*, 4(4):043210, 2022. doi: 10.1103/PhysRevResearch.4.043210.
- R. A. Luke, R. J. Braun, T. A. Driscoll, C. G. Begley, and D. Awisi-Gyau. Parameter estimation for evaporation-driven tear film thinning. *Bull. Math. Biol.*, 82(6):71, June 2020.
- R. A. Luke, R. J. Braun, T. A. Driscoll, D. Awisi-Gyau, and C. G. Begley. Parameter estimation for mixed-mechanism tear film thinning. *Bull. Math. Biol.*, 83(5):56, 2021a.
- R. A. Luke, R. J. Braun, T. A. Driscoll, D. Sinopoli, V. Yawatkar, L. You, A. Phatak, and C. G. Begley. Fitting simplified models to machine learning-identified tear film breakup. *Invest. Ophthalmol. Vis. Sci.*, 62(8):1315–1315, 2021b.
- S. Mishima and D. M. Maurice. The oily layer of the tear film and evaporation from the corneal surface. *Exp. Eye Res.*, 1:39–45, 1961.
- M. C. Mota, P. Carvalho, J. Ramalho, and E. Leite. Spectrophotometric analysis of sodium fluorescein aqueous solutions. Determination of molar absorption coefficient. *Internat. Ophthalmol.*, 15(5):321–326, 1991.
- B. Nagyová and J. M. Tiffany. Components responsible for the surface tension of human tears. *Current Eye Research*, 19(1):4–11, 1999.
- J. D. Nelson, J. P. Craig, E. K. Akpek, D. T. Azar, C. Belmonte, A. J. Bron, J. A. Clayton, M. Dogru, H. S. Dua, G. N. Foulks, J. A. P. Gomes, K. M. Hammitt, J. Holopainen, L. Jones, C.-K. Joo, Z. Liu, J. J. Nichols, K. K. Nichols, G. D. Novack, V. Sangwan, F. Stapleton, A. Tomlinson, K. Tsubota, M. D. P. Willcox, J. S. Wolffsohn,

- and D. A. Sullivan. TFOS DEWS II Introduction. *Ocul. Surf.*, 15(3):269–275, 2017.
- J. J. Nichols, G. L. Mitchell, and P. E. King-Smith. Thinning rate of the precorneal and prelens tear films. *Invest. Ophthalmol. Vis. Sci.*, 46(7):2353, 2005. ISSN 1552-5783.
- M. S. Norn. Micropunctate fluorescein vital staining of the cornea. *Acta Ophthalmol.*, 48(1):108–118, 1970.
- C.-C. Peng, C. Cerretani, R. J. Braun, and C. J. Radke. Evaporation-driven instability of the precorneal tear film. *Adv. Coll. Interface Sci.*, 206:250–264, 2014.
- C. Rackauckas and Q. Nie. Differentialequations.jl—a performant and feature-rich ecosystem for solving differential equations in Julia. *Journal of Open Research Software*, 5(1):15, 2017.
- C. Rackauckas, Y. Ma, J. Martensen, C. Warner, K. Zubov, R. Supekar, D. Skinner, and A. Ramadhan. Universal differential equations for scientific machine learning. *arXiv:2001.04385 [cs, math, q-bio, stat]*, Jan. 2020.
- R. Riquelme, I. Lira, C. Pérez-López, J. A. Rayas, and R. Rodríguez-Vera. Interferometric measurement of a diffusion coefficient: Comparison of two methods and uncertainty analysis. *J. Phys. D: Appl. Phys.*, 40(9): 2769–2776, 2007.
- X. Shi, G. G. Fuller, and E. S. G. Shaqfeh. Instability and symmetry breaking of surfactant films over an air bubble. *J. Fluid Mech.*, 953:A26, 2022.
- F. Stapleton, M. Alves, V. Y. Bunya, I. Jalbert, K. Lekhanont, F. Malet, K.-S. Na, D. Schaumberg, M. Uchino, J. Vehof, E. Viso, S. Vitale, and L. Jones. TFOS DEWS-II epidemiology report. *Ocul. Surf.*, 15(3):334–365, 2017.
- M. Tancik, P. Srinivasan, B. Mildenhall, S. Fridovich-Keil, N. Raghavan, U. Singhal, R. Ramamoorthi, J. Barron, and R. Ng. Fourier features let networks learn high frequency functions in low dimensional domains. In *Advances in Neural Information Processing Systems*, volume 33, pages 7537–7547, Red Hook, NY, USA, 2020. Curran Associates, Inc.
- J. M. Tiffany. The viscosity of human tears. *Internat. Ophthalmol.*, 15(6):371–376, 1991.
- L. N. Trefethen. *Spectral Methods in Matlab*. Software, Environments, Tools. SIAM, Philadelphia, 2000. ISBN 978-0-89871-465-4.
- S. Wang, H. Wang, and P. Perdikaris. On the eigenvector bias of Fourier feature networks: From regression to solving multi-scale pdes with physics-informed neural networks. *Comput. Methods Appl. Mech. Eng.*, 384: 113938, 2021a.
- S. Wang, H. Wang, and P. Perdikaris. On the eigenvector bias of Fourier feature networks: From regression to solving multi-scale PDEs with physics-informed neural networks, Sept. 2021b. URL <https://doi.org/10.1016/j.cma.2021.113938>.
- W. R. Webber and D. P. Jones. Continuous fluorophotometric method of measuring tear turnover rate in humans and analysis of factors affecting accuracy. *Med. Biol. Eng. Comput.*, 24(4):386–392, July 1986.
- M. D. P. Willcox, P. Argüeso, G. A. Georgiev, J. M. Holopainen, G. W. Laurie, T. J. Millar, E. B. Papas, J. P. Rolland, T. A. Schmidt, U. Stahl, T. Suarez, L. N. Subbaraman, O. Ö. Uçakhan, and L. Jones. TFOS DEWS II Tear Film Report. *Ocul. Surf.*, 15(3):366–403, 2017.
- Z. Wu, C. G. Begley, N. Port, A. Bradley, R. Braun, and E. King-Smith. The effects of increasing ocular surface stimulation on blinking and tear secretion. *Invest. Ophthalmol. Vis. Sci.*, 56(8):4211–4220, 2015.
- S. Yadav. Combining PCA with neural networks: Improving model efficiency and interpretability. *Int. J. Innov. Res. Eng. Multidiscip. Phys. Sci.*, 7:N/A, 2019.
- L. Zhong, R. J. Braun, C. G. Begley, and P. E. King-Smith. Dynamics of fluorescent imaging for rapid tear thinning. *Bull. Math. Biol.*, 81(1):39–80, 2019.