



## Ensemble machine learning predicts displacement of cantilevered fibers impacted by falling drops



Panporn Orkweha<sup>a</sup>, Alexis Downing<sup>b</sup>, Amy P. Lebanoff<sup>c</sup>, Sharare Zehtabian<sup>d</sup>, S. Safa Bacanli<sup>d</sup>, Damla Turgut<sup>d</sup>, Andrew K. Dickerson<sup>c,\*</sup>

<sup>a</sup> Mechatronics and Robotics Engineering, Southern Illinois Univ. Edwardsville, United States of America

<sup>b</sup> Computer Engineering, Florida Polytechnic Univ., United States of America

<sup>c</sup> Mechanical and Aerospace Engineering, Univ. of Central Florida, United States of America

<sup>d</sup> Computer Science, Univ. of Central Florida, United States of America

### ARTICLE INFO

#### Article history:

Received 9 July 2020

Received in revised form 10 November 2020

Accepted 12 February 2021

Available online xxxx

#### Keywords:

Impact force

Ohnesorge number

Drop fracture

Flexible fiber

Wetting

Predictive modeling

### ABSTRACT

In this study, we consider a simple system of water drops impacting cantilevered, circular fibers with velocity 1.0–2.4 m/s. The dynamics of the system are complicated because the motion of the drop and deflecting fiber at impact are highly coupled and the outcome is influenced by several continuous variables. The unpredictable dynamics of the system call for the use of computational tools that can reveal relationships between variables and make predictions about the physical outcomes for parameter values for which there is no experimental data. This study considers three fibers of contrasting properties, with hydrophilic and hydrophobic wetting conditions, exposed to a range of falling drop diameters and velocities. We predict maximal fiber deflection due to the impacting drop using an ensemble machine learning algorithm combining three base algorithms: a random forest regressor, a gradient boosting regressor, and a multi-layer perceptron. We train and test our algorithm with experimental datasets comprising 405 total trials using three different fibers and five input variables per fiber. The approach allows the determination of relative dominance of the input features in the prediction, reveals shortcomings of traditional scaling approaches for momentum transfer, and shows that drop momentum plays only a minor role in fiber deflection. Finally, our approach provides another example for the application of machine learning to characterize complex and coupled systems in fluid mechanics.

© 2021 Elsevier Ltd. All rights reserved.

## 1. Introduction

The phenomena of water drops impacting cantilevered structures is ubiquitous in nature. Examples include blades of grass, mammal fur (Dickerson et al., 2012a), insect wings (Dickerson et al., 2012b, 2014; Dickerson and Hu, 2014), laminar leaves (Gart et al., 2015), and slender pine needles (Lebanoff and Dickerson, 2020). These ephemeral interactions are rife with dynamic complexity as fluid and solid motion are highly coupled and reliant upon a plethora of variables with a wide range of values. The characterization of these types of interactions has applications in fog harvesting (Andrews et al., 2011), inkjet printing (Zhang et al., 2015), and energy harvesting (Ilyas and Swingler, 2015; Hobbs and Hu, 2012; Kim et al., 2011). In this study we experimentally investigate the impact of water drops onto slender, cantilevered fibers, inspired

\* Corresponding author.

E-mail address: [dickerson@ucf.edu](mailto:dickerson@ucf.edu) (A.K. Dickerson).

by rainfall on suspended pine needles. Unraveling the physics of such interactions will lead to a deeper understanding of how pines and other conifers cope with the physical stresses of rainfall, reduce the kinetic energy of raindrops landing on the forest floor, and shed residual water that would otherwise hinder photosynthesis (Lebanoff and Dickerson, 2020; Jenks and A., 1999; Smith and McClean, 1989). Further applications include the development of novel sensors that can characterize and capture mass from passing drops, and flexible surfaces which can self-dry.

Although experiments of drop impact on cantilevered fibers have been performed previously using highly wetting silicon oils (Comtet et al., 2016; Dressaire et al., 2016), these studies considered a smaller number of parameters than we consider here. Furthermore, our experimental drop velocity  $V$  is generally higher than that of previous studies. A drop's impact onto a thin, flexible fiber causes a large angle of displacement and one of three types of collisions: drop capture (entire drop remains attached to fiber post-collision) (Comtet et al., 2016; Dressaire et al., 2016; Kim and Kim, 2016), impacts where the drop remains one body, and drop fragmentation. Impacting drops excite the fiber's natural frequency, with vertical displacement  $\delta$  determined by momentum transfer and fiber properties (Comtet et al., 2016; Dressaire et al., 2016). For large drop viscosity, impacts are inelastic and the initial speed of the fiber is proportional to  $V$ . In contrast, at low viscosity the contact time between the drop and fiber varies inversely with drop speed (Comtet et al., 2016). Drop capture is dependent on the fiber's ability to slow the drop's kinetic energy over a sufficient time to prevent drop breakup, and is thus favored by slower drops and more flexible fibers (Dressaire et al., 2016; Comtet et al., 2016). To capture a drop, the fiber must not begin its rebound, or reversal of motion, before the elongated drop begins to recoil.

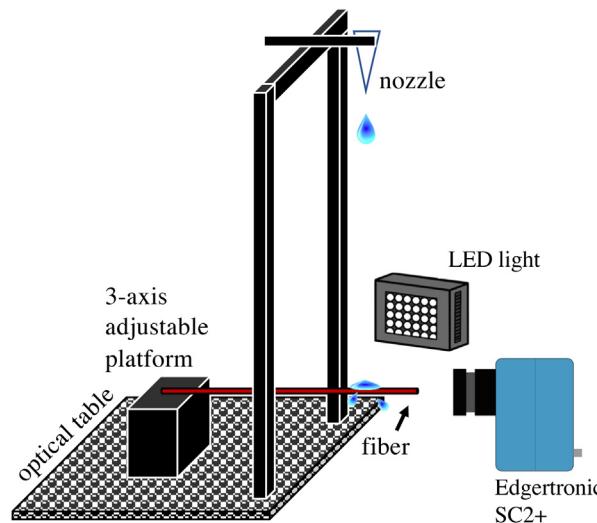
The ability to predict the maximal amplitude of the undamped vertical displacement of the fiber tip  $\delta_0$  for a given set of initial parameters is a difficult task. Fiber density  $\rho_f$  and linear density  $\rho'_f$ , surface wettability  $W$ , bending stiffness  $K = EI$ , cantilevered length  $\ell$ , drop diameter  $D$ , and velocity  $V$  are a few of the parameters which can vary in a natural system and those we tested in this study. Here,  $E$  is elastic modulus,  $I = \pi d^4/64$  is the area moment of inertia and  $d$  is fiber diameter. The above variables can occupy a continuous range of values, and thus the parameter space has an infinite number of combinations. Even if these select parameters could take only five values each, the experimental space would contain  $6^5 = 7776$  possible combinations. Hence, the behavior of this simple system is an excellent candidate to explore with machine learning.

Despite the widespread success of machine learning for studying complex problems, there have been only limited attempts to apply machine learning to fluid physics. A recent examples is drop ejection from vibrating cantilevers (Alam et al., 2020b,a), nearly the inverse of the problem at hand. More common, yet still seldom used in fluid mechanics, are more traditional predictive tools such as Kalman filters to estimate parameters associated with hydraulic fracturing (Abhinav et al., 2017; Krymskaya et al., 2009; Anonson et al., 2009), multiphase flow (Gu and Oliver, 2007; Jansen et al., 2008; Lorentzen et al., 2003), multi-particle tracking (Volk et al., 2008; Adrian, 2005; Takehara et al., 2000), and empirical models guided by large datasets (McMillan et al., 2017). Interactions like those in this study, highly coupled in solid-fluid domains, stand to capitalize on machine learning algorithms with dimensionality-reduction techniques such as orthogonal decomposition (Kutz, 2017; Tracey et al., 2015). It is intractable to derive the governing equations of the highly coupled physics analytically, and thus machine learning provides a powerful tool to predict the outcomes of many parametric combinations. *Ensemble learning* is a class of machine learning algorithms where multiple competing learning algorithms are combined to train a predictive model, and typically outperforms any of the constituent learning algorithms alone (Minku et al., 2010; Kuncheva and Whitaker, 2003; Brown et al., 2005; Melville and Mooney, 2005) by leveraging the best traits of multiple approaches to generate a prediction.

To determine how system properties determine maximal displacement amplitude  $\delta_0$  of our fibers, we perform 405 experimental trials with three different fibers, each with constant  $E$  and  $d$ . Both controlled and random experimental combinations of the remaining property values are performed. For each trial the temporal displacement  $\delta$  of fibers is recorded to produce the outcome of interest  $\delta_0$  for training and testing an ensemble machine learning algorithm. In Section 2, we describe our experimental approach and in Section 3, we describe the algorithms which comprise our ensemble approach. In Section 4, we discuss scaling maximum fiber displacement. In Section 5, we discuss algorithmic predictions and compare against experimental data, pointing out the variables which dominate the physical response of our system. Conclusions of our study are given in Section 6, and information on how to obtain our code and raw files is provided.

## 2. Experimental approach

In this section, we describe experimental methods that are detached from computation but designed with subsequent computation planned. We fix one end of the fiber onto a 3-axis adjustable platform while leaving the other free, as shown by a schematic of our experimental apparatus in Fig. 1. Fiber length is customizable by relocating the fiber on the stage and then adjusting the stage to realign the fiber to be struck by the drop at the free end tip. A nozzle of adjustable height and supplied by a water-filled syringe generates drops. Drop size is modulated by 21 nozzles of various diameter and length, which produce drop diameters ranging from 2.4–3.8 mm. Drops emitted from any single nozzle have variance about the mean diameter of 1%–17%. Impact locations on the fibers are chosen such that the drop impacts as close to the tip as possible without the spreading drop leaving the cantilever tip in the axial direction. A recent study (Lebanoff and Dickerson, 2020) showed drop spreading on a fiber is generally 1.5D–2.5D for the impact conditions considered in this study. Therefore for the longest 33-mm fibers, drops may impact 5%–14% the fiber length from the fiber tip. The



**Fig. 1.** Experimental apparatus.

**Table 1**  
Experimental fiber properties. For all measurements  $N = 3$ .

Fiber	$d$ [mm]	$\rho_f$ (g/cm <sup>3</sup> )	$\rho'_f$ (mg/cm)	$E$ [MPa]	$K$ [N m <sup>2</sup> ]	$\theta_e$ [°] ( $W = 0$ )	$\theta_e$ [°] ( $W = 1$ )
A	$0.94 \pm 0.02$	$1.29 \pm 0.04$	$8.92 \pm 0.55$	$6.86 \pm 1.95$	$3 \times 10^{-7}$	$95.0 \pm 3.7$	$156.0 \pm 8.3$
B	$0.80 \pm 0.01$	$1.08 \pm 0.02$	$5.41 \pm 0.05$	$2745 \pm 681$	$5.5 \times 10^{-5}$	$78.3 \pm 6.6$	$161.7 \pm 0.9$
C	$0.61 \pm 0.01$	$1.08 \pm 0.03$	$3.13 \pm 0.06$	$3334 \pm 62$	$2.3 \times 10^{-5}$	$78.3 \pm 11.0$	$160.7 \pm 1.7$

shortest fibers are shorter than drop spreading distances under certain impact conditions, in which case those trials are not considered for analysis. An Edgertronic SC2+ films impacts at 3000–6000 fps by viewing the impact down the fiber axis. Such a viewing angle permits the measurement of fiber displacement and ensures that only centric impacts are recorded. We do not consider impacts where the drop centerline falls outside the fiber diameter, thus permitting a maximum offset from perfectly centric impacts of  $\sim 0.5$  mm. Trials outside this tolerance were discarded. Positioning the camera normal to the fiber axis allows the capture of images used for visualization purposes only, as seen in photographs below, and not used for data collection. The scene is lit by a GS Vitec MultiLED LT lamp. We use three hydrophilic fibers sourced from a Koosh® ball (Fiber A), and ukulele strings (Fibers B,C), selected based on their range of moduli and fiber diameter. Fiber properties are tabulated in Table 1. Moduli of elasticity is measured with MTS Criterion 40 tensile tester fitted with a 2 kN load cell. Contact angles are measured optically for sessile drops on untreated fibers and contact angles found with the Contact Angle plugin for ImageJ. Unaltered fibers, condition  $W = 0$ , are sprayed with NeverWet® to increase their contact angle and make them hydrophobic, condition  $W = 1$ . Treated fiber contact angles are measured with Open Source Physics Tracker (OSPT) while partly submerged in distilled water

The Navier–Stokes and elasticity equations govern a thin flexible fiber's displacement in relation to aforementioned system properties, but three-dimensional flow inside and on the surface of impacting drops is not resolved, preventing an analytical solution to these coupled equations. Thus, *a priori* determination of parameter influence and dominance is not possible. However, we wish to limit the number of tests while retaining significant relationships between input parameters and output (displacement). As such, we split the experiment into two sections. The first uses random parameter values and the other using controlled parameter values, as outlined in Tables 2 and 3 respectively. In random parameter experiments, we generate 279 random parameter combinations (93 combinations per fiber) within a range specified in Table 2. Random parameter combinations are chosen by a random number generator at [www.random.org](http://www.random.org). In the controlled parameter experiment, we perform a full factorial experiment onto Fiber A only, producing 126 total combinations. These parameter combinations are meant to elucidate drop velocity  $V$  influence on Fiber A's displacement in particular. We thus reduce the number of nozzles to three, which are henceforth labeled 0, 1, and 2, producing approximately 2.9-mm, 3.3-mm, and 3.6-mm diameter drops respectively. As discussed in Section 5, we observe an uncertain relationship between  $V$  and  $\delta$  using random parameter experiment with Fiber A.

Experimental videos are processed with OSPT to extract drop size and kinematics, and to track time-varying cantilever displacement at the tip. The displacement minima of the highly under-damped vibration curve are fit with  $\delta(t) = \delta_0 e^b$ , where  $\delta_0$  represents the undamped maximum displacement which we report. Drop velocity was calculated by tracking drop position over 10 frames pre-impact and fitting these points with a line. The velocity gain under gravity of  $\sim 3.3$  cm/s within 10 frames, or 0.0033 s at 3000 fps, is insignificant. Drop diameter was measured with the circle fitter feature in OSPT.

**Table 2**

Range of values for each parameter in random parameter experiment.

Random Parameter	Range
Fiber length, $\ell$ (mm)	0–33
Drop height (mm)	55–305
Nozzle	0, 1, 2, . . . , 20
Wettability, $W$	0, 1

**Table 3**

Range of parameter values for combinations in controlled parameter experiment using Fiber A only.

Controlled Parameter	Range
Fiber length, $\ell$ (mm)	10, 20, 30
Drop height (mm)	88, 119, 150, 181, . . . , 305
Nozzle	0, 1, 2
Wettability, $W$	0, 1

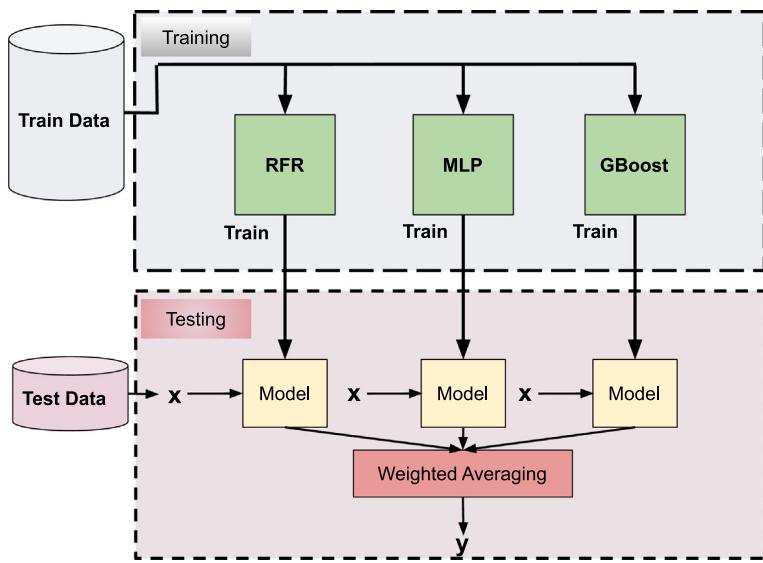
### 3. Algorithmic approach

In this section, we briefly describe ensemble learning and the base learners comprising our algorithm. A *learner* or trained model is the output of a *base learning algorithm*. A base learning algorithm takes experimental training data as input and creates a trained model. *Training* refers to the procedure that changes the parameters of the model until it achieves a minimum (or a small) error. Our goal is to predict the fiber's maximum displacement based on inputs such as fiber length, drop diameter and velocity, fiber wettability, fiber diameter, and modulus. *Ensemble learning* is a machine learning technique in which multiple base learning algorithms are combined to form a committee and achieve a more robust prediction than can be accomplished by using a single model (Dietterich, 1997). The use of a set of models (an ensemble) does not guarantee optimum performance but lowers the risk of a poor prediction (Wu and Xu, 2018). In the ensemble learning process, many independent models are generated, a step referred to as ensemble generation (Mendes-Moreira et al., 2012). The predictions of base learners are combined to form the output prediction.

Since the output of the prediction model is a continuous variable,  $\delta_0$  in our case, the prediction problem in this work is a regression problem, mandating the aforementioned committee is formed of regressors. We have implemented three base learners as the ensemble committee: the Random Forest Regressor (RFR), the Multi-Layer Perceptron (MLP), and the Gradient Boosting Regressor (GBoost). These are well-known ensemble learners, which are frequently used in regression problems. We use the Scikit-learn library, a free machine learning library for Python programmers, for the implementation (Pedregosa et al., 2011). Each algorithm has a number of configurable parameters which we list in the Online Supplement.

Below we briefly describe each base learner:

- The Random Forest Regressor (RFR) is itself an ensemble learning algorithm which is founded on a bagging (**bootstrap aggregating**) technique (Breiman, 2001). Each model is trained using a different training set sampled with replacements from the training data in a process called bootstrapping. In a random forest regressor, each model is a decision tree, with inner nodes corresponding to decisions and leaves corresponding to the final prediction values. In machine learning, a data point is represented as a vector called a *feature vector* in which each vector element corresponds to a measurement value of an input variable. Given a feature vector, the decision tree starts from a root node and follows a path in the tree based on the feature values and conditions of the nodes until it arrives at a leaf node. After all decision tree models are trained, the final prediction value becomes the average of the predictions of the individual decision trees. RFRs have been shown to improve on the overall performance of the individual decision tree models by providing lower prediction error (Segal, 2004; Zhang and Ma, 2012).
- The Multi-Layer Perceptron (MLP) is a class of artificial neural networks that learns a function  $f(X) : \mathbb{R}^m \rightarrow \mathbb{R}^0$  from a training dataset in which  $X : x_1, x_2, \dots, x_m$  is the input and a set of features given to the network. Having a target value  $y$ , the network learns  $f(X)$  such that  $f(X) \approx y$ . An MLP is composed of an input layer, one or more hidden layers, and an output layer. The input layer feeds data into one or more hidden layers. Each hidden layer transforms input data through an activation function to produce input for the next layer. Finally, the transformed data arrives at an output layer where it goes through a final transformation. Training an MLP involves adjusting the parameters of the model such that the prediction error is minimized.
- The Gradient Boosting Regressor (GBoost) is an ensemble learning algorithm similar to RFR. However, unlike RFR, which uses bagging, GBoost uses boosting. The key difference between bagging and boosting is that boosting trains its base learners sequentially to build one decision tree at a time, adaptively rather than independently, and calculates the average of tree outputs at the end. Each base learner is thus dependent on the previously trained base learner. The term adaptive here refers to the data passed to the next base learner in sequence. The idea of the boosting method is to pass along difficult-to-process data to train the subsequent model. In this way, the algorithm 'focuses' on training itself with data in which trends are difficult to determine and ultimately perform well with any input.



**Fig. 2.** Ensemble learning: Weighted averaging over Random Forrest Regressor (RFR), Multi-Layer Perceptron (MLP) and Gradient Boosting Regressor (GBoost).

Implementing both bagging and boosting techniques in our approach supports the accuracy of our prediction model and provides a varying sampling method. Our method of ensemble learning is schematized in Fig. 2. The prediction of each base learner is multiplied by an assigned weight before a summation forms the final prediction (Zhang and Ma, 2012). Weights are determined by the non-negative least squares (NNLS) method which assigns weights to each base learner such that the final prediction has the lowest prediction error possible given each base learner's prediction accuracy. No matter the distribution, the weights always sum to 1. We implement the NNLS algorithm created by Lawson and Hanson (Lawson and Hanson, 1995) based on the following optimization problem:

$$\begin{aligned} & \text{minimize } \|\mathbf{A}\omega - \mathbf{y}\|^2, \\ & \text{subject to } \omega \geq 0 \end{aligned} \quad (1)$$

where  $\mathbf{A}$  is the matrix of predictions obtained by base learner models, vector  $\mathbf{y}$  is the corresponding target value, and  $\omega$  is a vector of the weights for our three base learner models obtained by solving the optimization problem in Eq. (1). Predicted values  $\mathbf{Pr}$  from each base learner are multiplied by a respective weight in the regression to predict the actual outcome. The ensemble learner makes final predictions by summing the product of each base learner's prediction and its respective weight vector,

$$\mathbf{Pr}(\text{ensemble}) = \omega_{\text{RFR}}\mathbf{Pr}(\text{RFR}) + \omega_{\text{MLP}}\mathbf{Pr}(\text{MLP}) + \omega_{\text{GBoost}}\mathbf{Pr}(\text{GBoost}) \quad (2)$$

The NNLS method is implemented after training each base learner by 10-fold cross-validation with shuffle to acquire 10 sets of predictions from each base learner. k-fold cross-validation is a technique in which we train the algorithm k times repeatedly with a fraction  $1/k$  of training examples omitted for testing purposes, resulting in lower variance of the estimate (Bengio and Grandvalet, 2004).

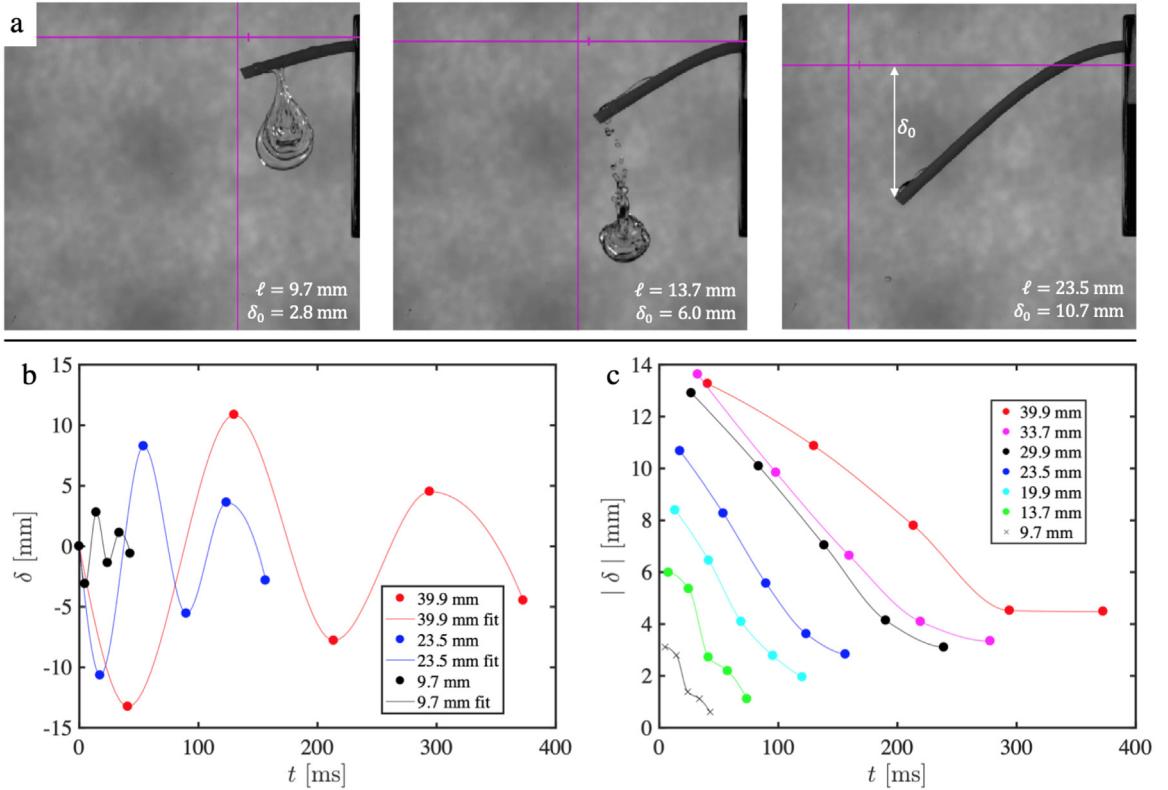
#### 4. Scaling maximum fiber displacement

We consider fiber flexibility and drop momentum to form scaling relations by which to compare machine learning predictions and variable scoring in Section 5. In the simplest case, a cantilevered fiber at mechanical equilibrium displaces under a static force, here taken to be the force of an impacting drop,  $\delta_0 \sim F\ell^3/K$ . We scale impact force  $F \sim \mu DV$  as done previously (Comtet et al., 2016) for low-viscosity liquids, where  $\mu = 0.89$  cP is drop viscosity. Thus,  $\delta_0 \sim \mu\ell^3 DV/Ed^4$ , but these assumptions omit  $\rho$  and  $\rho_f$ , parameters important to momentum transfer.

Instead consider fiber momentum generated by an impacting drop, which under first mode vibration, can be written as (Soto et al., 2014)

$$P_f = \int_0^\ell \rho_f \frac{\pi d^2}{4} 2\pi f_0 \delta_0 \frac{x^2}{\ell^2} \cos(2\pi f_0 t) dx = \frac{2\pi}{3} M f_0 \delta_0 \cos(2\pi f_0 t) \sim \rho_f \ell d^2 f_0 \delta_0, \quad (3)$$

where  $M = \rho_f \ell \pi d^2 / 4$  is the fiber mass and  $f_0 = (K/M\ell^3)^{1/2}$  is the first resonance frequency. The Reynolds number  $Re = \rho DV/\mu = 3380 - 9600$  and  $D > d$ , thus drops flow around fibers over the course of impact rather than produce



**Fig. 3.** Experimental displacement of Fiber A by changing length.  $We = 184$  for all panels. (a) Fibers of various length impacted by a 2-mm diameter drop traveling at 2590 mm/s. Cross hairs denote fiber tip location pre-impact. (b) Temporal displacement curves of fibers at three select lengths. (c) Temporal displacement magnitude or vibration peaks and troughs across experimental fiber length.

inelastic collisions, as pictured in Fig. 3a. Despite circumferential flow we reasonably expect  $P_f \sim mV$ , where  $m$  is spherical drop mass. Thus,  $\rho_f \ell d^2 f_0 \delta_0 \sim \rho D^3 V$ , which can be written, omitting constants,

$$\frac{\delta_0}{\ell} \sim \frac{\rho}{\rho_f^{\frac{1}{2}}} \frac{D^3}{d^3} \frac{V}{E^{\frac{1}{2}}} \sim \frac{\rho}{\rho_f^{\frac{1}{2}}} \frac{D^3}{d^2} \frac{V}{E^{\frac{1}{2}}}. \quad (4)$$

The work of Comtet et al. (2016) for drop impact onto flexible fibers provides yet another approach to predict  $\delta_0$  in terms of input parameters, which may be compared to Eq. (4). Our drops have Ohnesorge numbers  $Oh = 3\mu/\sqrt{\rho\sigma D/2} = 0.007 - 0.009$  and impact velocities in excess of 1000 mm/s, and therefore all drops fracture at impact (Comtet et al., 2016; Dressaire et al., 2016). Nevertheless, momentum transfer between the drop and fiber can be expressed,

$$M_{\text{eff}} V_{\text{fiber}} \sim F \Delta t, \quad (5)$$

where  $M_{\text{eff}} = \rho_f \pi d^2 \ell / 12$  is effective fiber mass with a clamped boundary,  $V_{\text{fiber}} = \dot{\delta}(0)$  is the initial fiber tip speed, and  $\Delta t \sim D/V$  is the interaction time. Rewriting Eq. (5) we find  $V_{\text{fiber}} = \mu D^2 / M_{\text{eff}}$  is independent of  $V$ . Maximum tip displacement  $\delta_0 = V_{\text{fiber}} \tau_b / 4$  where the natural bending time of the fiber (Comtet et al., 2016; Dressaire et al., 2016)

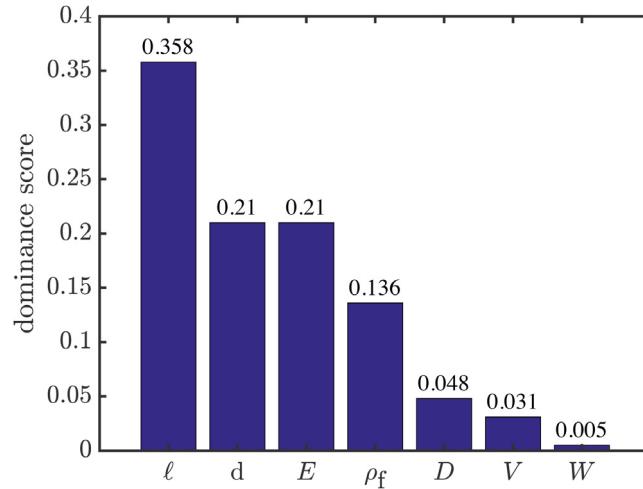
$$\tau_b = \frac{2\pi}{\beta^2} \ell^2 \sqrt{\frac{\rho_f \pi d^2}{4K}}, \quad (6)$$

and  $\beta = 1.875$  for first mode vibration. Higher modes may be present for the longest fibers, but the timescale of first mode excitation remains dominant. Neglecting higher-order vibration we can write

$$\delta_0 \sim \frac{\mu D^2}{M_{\text{eff}} \beta^2} \ell^2 \sqrt{\frac{\rho_f \pi d^2}{4K}}. \quad (7)$$

Rewriting Eq. (7) in terms of the same variables provided in Eq. (4) and omitting all constants,

$$\frac{\delta_0}{\ell} \sim \frac{\mu}{(\rho_f E)^{\frac{1}{2}}} \frac{D^2}{d^3} \sim \frac{\mu}{(\rho_f E)^{\frac{1}{2}}} \frac{D^2}{d^2}. \quad (8)$$



**Fig. 4.** Relative parameter dominance scores for producing cantilever displacement, assigned by our ensemble algorithm trained with the random parameter dataset.

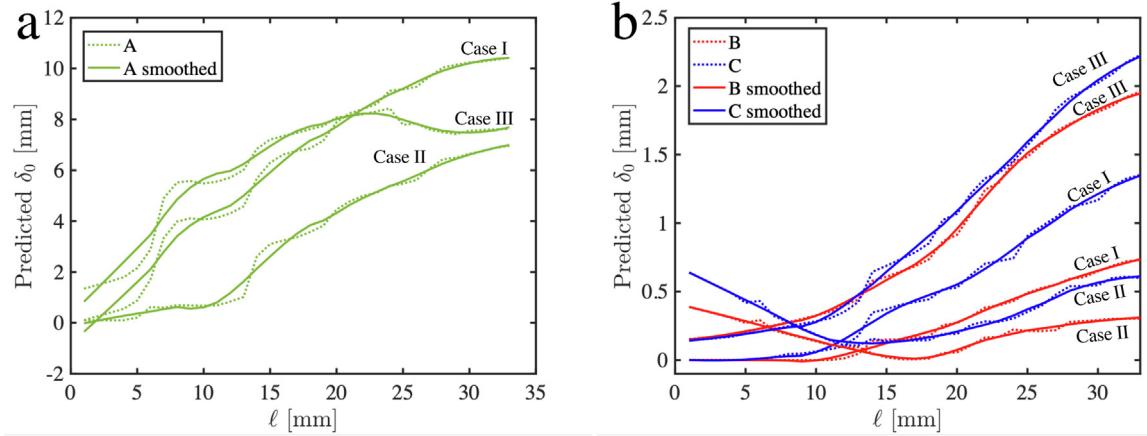
We note Eq. (8) may also be derived by a combination of Eqs. (3) and (5), and provides a non-intuitive relation for  $\delta_0$  as Eq. (8) is independent of drop momentum. We show below in Section 5 the seemingly nonphysical relation of Eq. (8) highlights the relative insignificance of drop velocity for producing  $\delta_0$  in our experiments. However, it is not obvious at this juncture if either the scalings of Eq. (4) or Eq. (8) provide a physically relevant characterization of the system, as both relations are derived using physical intuition. We must thus interpret their applicability alongside other approaches, and in this case, machine learning predictions. At a minimum, the above relations of Eqs. (4) and (8) provide an assumed dominance of each system parameter on fiber displacement and their contradictions exacerbate the need for a more sophisticated approach.

## 5. Results and discussion

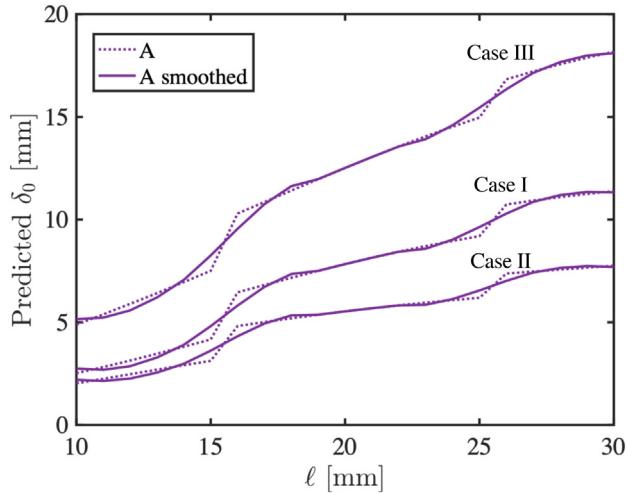
We impact cantilevered fibers with falling water drops to record their maximum displacement at impact using two parameter schemes outlined in Section 2. Inputs to each trial,  $W$ ,  $\ell$ ,  $D$ ,  $V$ ,  $d$ ,  $E$ ,  $\rho_f$ , and output  $\delta_0$  is passed to the ensemble algorithm described in Section 3 for training and comparison against algorithmic predictions. The trained algorithm assigns a dominance score to each input variable that represents the relative influence of each parameter on generating fiber displacement, as shown in Fig. 4. Fiber length  $\ell$  is the most critical factor governing displacement. Fiber diameter  $d$  and modulus  $E$  are likewise relatively dominant, contributing to fiber stiffness  $K$ . Holistically, the scores of Fig. 4 demonstrate fiber properties are more important to maximal displacement than drop properties, a relationship captured by Eq. (8) more so than Eq. (4). The mismatch between scaling relations and algorithmic scoring is perhaps due to the nature of drop flow and breakup at impact which is not well-represented in the equations. The least significant initial condition is fiber wettability  $W$ .

In this section, we present  $\delta_0$  predictions alongside select experimental results for changes in  $\ell$ ,  $D$ ,  $V$ , and  $W$ , and implicitly changes in  $K$  by using three fibers. Unless otherwise specified, predictions of  $\delta_0$  from our algorithm, described in Section 2, are made by holding all but one independent variable constant. We thus predict the behavior of  $\delta_0$  with respect to a single variable. Predictions are made using three cases. In Case I, variables are fixed at their mean values:  $\bar{\ell} = 17$  mm,  $\bar{D} = 3.1$  mm, and  $\bar{V} = 2020$  mm/s. In Case II, we fix variables at  $0.5\bar{\ell}$ ,  $0.5\bar{D}$ , and  $0.5\bar{V}$ . In Case III, we fix variables at  $1.5\bar{\ell}$ ,  $1.5\bar{D}$ , and  $1.5\bar{V}$ .  $E$  and  $d$  values are fixed according to individual fiber properties (Table 1) and  $W = 0$  (untreated fibers) for all three cases. In Case I, predictions are made within the experimental range of data used for training. Cases II and III provide predictions outside of experimental range, in some instances, and thus give insight into how the algorithm performs in extrapolation. Specifically, drop diameter and velocity lies outside of experimental ranges for Cases II and III.

Ensemble predictions change with every iteration and thus the prediction plots below are representative of the general predictive behavior and not exactly reproducible by subsequent executions of our algorithm. In the prediction plots that follow (Figs. 5–8), we plot algorithm output in dashes and smoothed predictions with solid curves for enhanced trend visualization. Prediction outputs can be noisy and contain artifacts, and thus smoothing is done post-prediction with a second-order Savitzky–Golay filter (Krishnan and Seelamantula, 2012; Watson et al., 2018, 2019, 2020; Smith et al., 2018) spanning 33% of the prediction dataset. Future researchers may choose different smoothing parameters.



**Fig. 5.** Prediction of  $\delta_0$  versus changes in length for (a) Fiber A, and (b) Fibers B and C garnered by training the ensemble learning algorithm with data from the random parameter experiment.

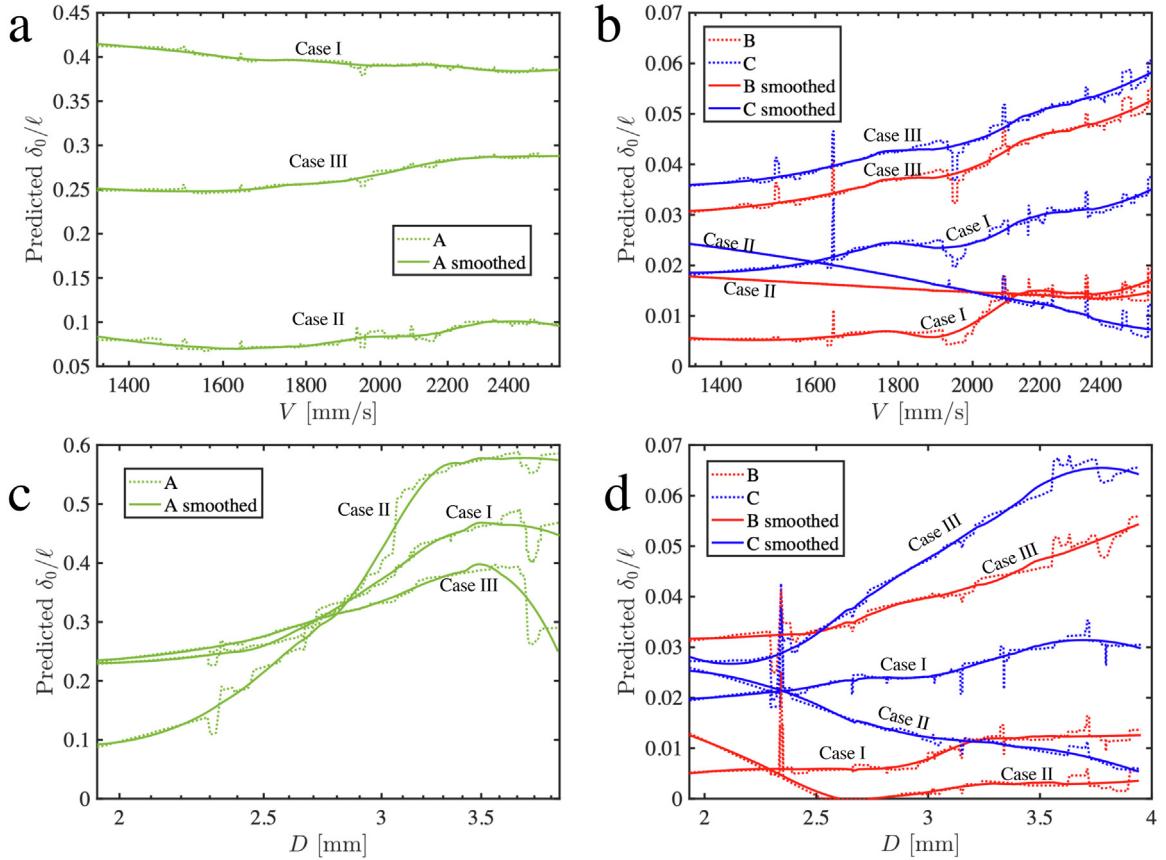


**Fig. 6.** Prediction of  $\delta_0$  versus changes in length for Fiber A garnered by training the ensemble learning algorithm with data from the controlled parameter experiment.

### 5.1. Fiber length

As expected, longer fibers experience greater values of  $\delta_0$  as shown in the experimental images of Fiber A in Fig. 3a. The drop moves lower in frame as the fibers become longer because the longest fibers continue downward motion after contact with the drop ceases. The cross-hairs in Fig. 3a denote the location of the tip pre-impact. Changing the length of the cantilevered fiber while keeping the drop characteristics unchanged yields the temporal displacement curves of Fig. 3b. Longer fibers experience lower vibration frequencies and remain in motion for an extended time dictated by inertia and various sources of damping (Gart et al., 2015; Alam et al., 2020a). By taking the absolute values of  $\delta$  at displacement peaks and troughs, we produce Fig. 3c, which compares displacement decay with time across length of Fiber A.

We plot predicted Fiber A  $\delta_0$  values against length in Fig. 5a and note the similarity in displacement magnitude to the experimental values of Fig. 3b,c, which were taken separately from our random and controlled parameter datasets. While the exact axial curvature of our bending fibers was not measured, we estimate bending strain at the fiber base  $\epsilon \approx 3\delta_0 d/2\ell^2 < 0.13$  for the greatest deflections of rubber Fiber A, a value generated using Euler–Bernoulli assumptions and one that decreases rapidly down the fiber axis. Curve-fitting the smoothed points of Case I in Fig. 5a reveals a best fit  $\delta_0 \sim \ell^{0.87}$  ( $R^2 = 0.98$ ), which is in reasonable agreement with Eqs. (4) and (8), and  $\delta_0 \sim \ell$  fits with  $R^2 = 0.97$ . It is expected based on fixed variable values that  $\delta_0$  for Case II lies below that of Case I for the entire range of length. The dip in the prediction for Case III below that of Case I at approximately  $\ell = 20$  mm is peculiar and likely a result of extrapolation error. Qualitative improvements in predictions for Fiber A are seen by training the algorithm with the controlled parameter dataset, as described in Section 2. The controlled parameter dataset yields predictions of  $\delta_0$  versus  $\ell$



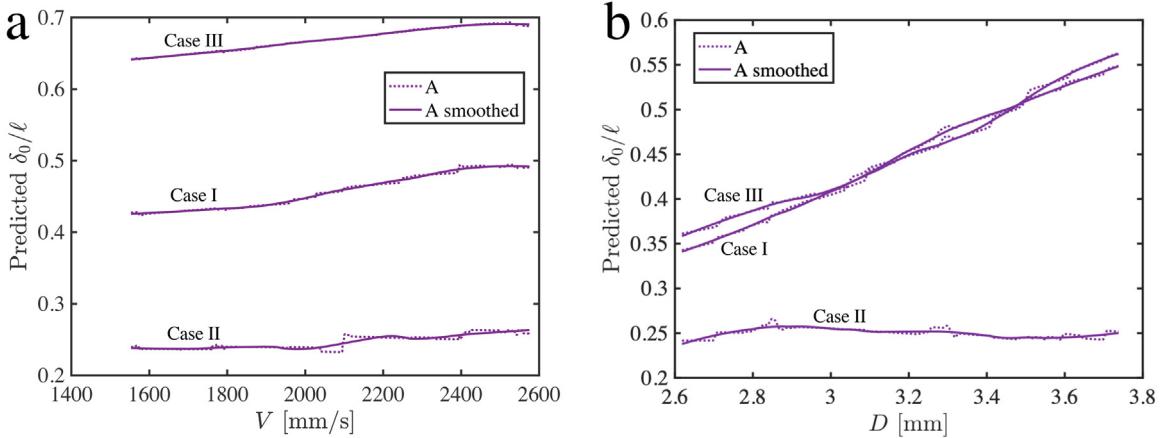
**Fig. 7.** Prediction of  $\delta_0/\ell$  by varying (a-b)  $V$  and (c-d)  $D$ , garnered by training the ensemble learning algorithm with data from the random parameter experiment.

in which curves for the three cases do not intersect and are ordered on the plot as expected, as seen in Fig. 6. Curve-fitting the smoothed points of Case I in Fig. 6 reveals a best fit  $\delta_0 \sim \ell^{1.28}$  ( $R^2 = 0.96$ ), which is again in reasonable agreement with Eqs. (4) and (8), and  $\delta_0 \sim \ell$  fits with  $R^2 = 0.93$ .

We measure, and likewise predict less displacement for Fibers B and C, shown in Fig. 5b, which are comparably stiffer (Table 1). We note the broad shape of the curves for B and C are different than that of A. This is perhaps a result of the considerable sag experienced by Fiber A (Fig. 3a) under its own weight, ensuring impacts onto the longest Fiber A lengths are oblique. Curve-fitting smoothed data of Case I with a first-degree polynomial for Fibers B and C at lengths 10-mm and beyond yields  $R^2 = 0.995$  for each. The abrupt change in slope at approximately 10-mm for Case I is due to our inability to detect displacement at small  $\ell$ . Experimentally, we find this lower limit of detection to be 2 mm, 14 mm, and 6 mm for Fibers A, B, and C respectively, corresponding to the relative values of bending stiffness  $K$  in Table 1. Predicted  $\delta_0$  is as expected for Case III with the exception of  $\delta_0 \rightarrow 0$  as  $\ell \rightarrow 0$ . Predictions for Case II are as expected above  $\ell \simeq 15$  mm, while the rise in  $\delta_0$  as  $\ell \rightarrow 0$  is likely due to extrapolation error.

## 5.2. Drop size and velocity

At a length  $\bar{\ell}$  the time to achieve  $\delta_0$ ,  $\tau_0 = \tau_b/4$ , is approximately 7 ms for Fiber A and < 1 ms for Fibers B and C. The passing time for the drop  $\Delta t < \tau_0$  for Fiber A and  $\Delta t > \tau_0$  for Fibers B and C across all drop speeds tested. Using our randomized parameter experiment to train our algorithm, predictions for  $\delta_0$  versus  $V$ , holding all other parameters constant, is shown in Fig. 7a,b. The relative strength of  $\ell$  over drop momentum to induce displacement, according to Fig. 4 can be observed in the range of  $\delta_0/\ell$  across the domain of Fig. 7a,b. The slight downward trend for Case I in Fig. 7a is a typical output of the algorithm for Fiber A and likely nonphysical because it contradicts the relations given by static assumptions and Eqs. (4) and (8), and is in stark contrast to the predictions of Fibers B and C (Fig. 7b) for Case I. Furthermore, the algorithm predicts Case III, which has a larger drop and longer fiber will deflect less than in Case I. These predictions thus mandate another approach to algorithm training for Fiber A. Training the algorithm with systematically varied data yields the positive correlation in Fig. 8a, as expected from scaling. Moreover, predicted  $\delta_0/\ell$  values for Case III are greater than those of Case I, as expected.



**Fig. 8.** Prediction of  $\delta_0/\ell$  by varying (a)  $V$  and (b)  $D$  garnered by training the ensemble learning algorithm with data from the controlled parameter experiment.

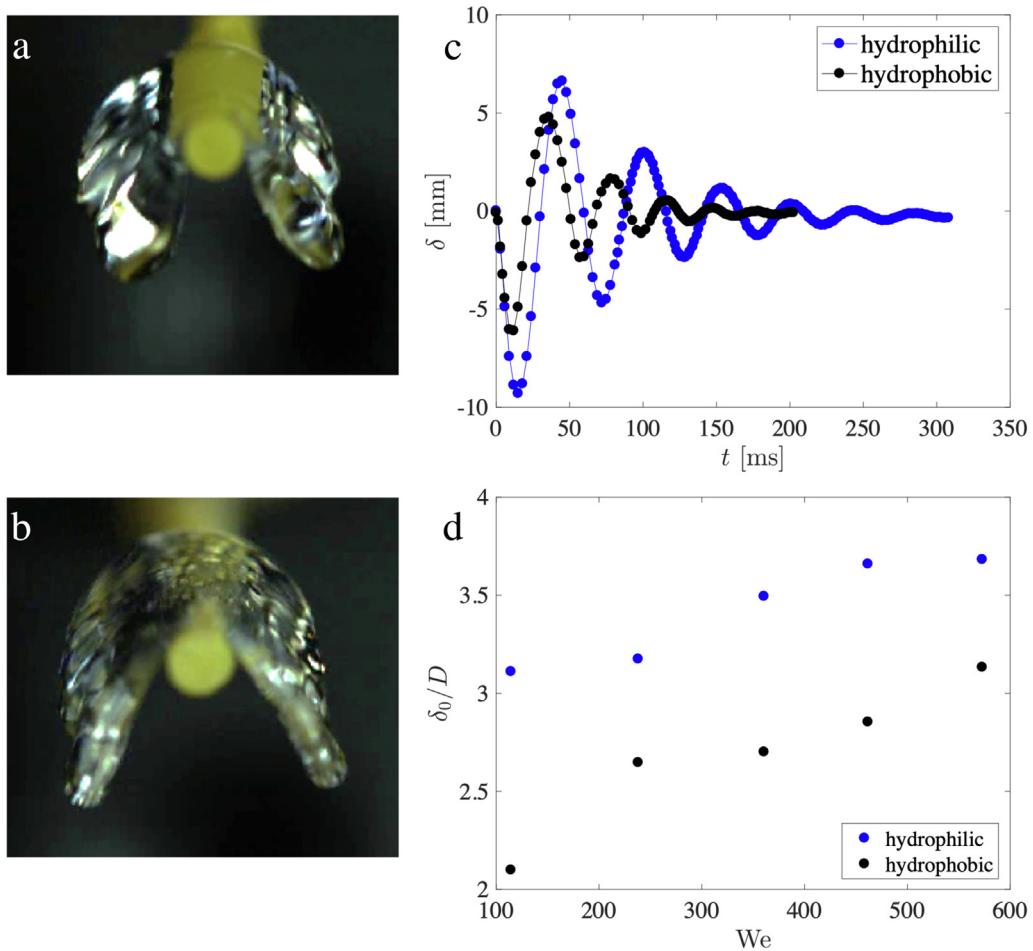
Changes in drop velocity has a lesser effect on the stiffer Fibers B and C,  $\Delta(\delta_0/\ell) \approx 0.015$  for  $\Delta V = 1200$  mm/s, compared to  $\Delta(\delta_0/\ell) \approx 0.05$  for Fiber A. The misbehavior of Case II predictions in Fig. 7b is associated with limitations in experimental observations at short fiber lengths for Fibers B and C. Case II specifies a simulation length of 8.5 mm and deflections become immeasurably small at 14 mm and 6 mm for Fibers B and C, respectively. Curve-fitting is not done for the smoothed curves in Figs. 7a,b and 8a because of the disparity in the range of independent versus dependent variables, yielding minute pre-factors. While our algorithm does predict increased deflection with  $V$  in Figs. 7b and 8a, the correlation is weak (Fig. 4) across a large range of drop velocity, confirming the non-intuitive absence of  $V$  in Eq. (8) is physically relevant.

Now, holding all other parameters constant, we predict how  $\delta_0/\ell$  changes with  $D$  in Fig. 7c-d. The algorithm predicts a positive correlation between  $\delta_0/\ell$  and  $D$  under Case I for all the fibers. Larger drops have more momentum than smaller drops and are able to exert force to the fiber for a longer  $\Delta t$ , creating greater displacement and fiber momentum. Larger drops likewise carry greater elongation times, which at  $\text{Oh} < 1$  is given by (Comtet et al., 2016) the inertia-capillary timescale  $\tau_{ic} \sim \sqrt{\rho D^3/8\sigma}$ . As drop size increases, however, momentum transfer is less efficient as the drops flow more freely around small targets (Dickerson et al., 2014). We rationalize that our predictions do not behave according to Eqs. (4) and (8) by noting that larger drops are more susceptible to deformation (Dickerson et al., 2014; Yarin, 2006; Lejeune and Gilet, 2019), and the smallest drops produce immeasurable deflection. Our predictions likewise show peaks and valleys in the prediction of  $\delta_0$  across the range of  $D$ . It is not clear if these changes in curvature are merely artifacts which can be rectified by additional training data and more aggressive smoothing, or if physical relevance to drop breakup and flow lies therein. Case III produces greater  $\delta_0/\ell$  predictions than Case I in Fig. 7d, indicating longer fibers experience proportionally greater deflections than shorter fibers and support dominance scores in Fig. 4. As with Fig. 7a, the predictions of Fig. 7c for Cases II and III appear non-physical. The prediction of  $\delta_0/\ell$  in Fig. 8b provides more intuitive predictions, but not ones that do not follow Eqs. (4) and (8), as predictions are linear in all cases, potentially highlighting deformation effects are not captured in our scaling relations.

### 5.3. Fiber wettability

The surface chemistry of a cantilevered fiber plays a role in the force it experiences by a passing drop. For this study, we consider fibers of two conditions, untreated ( $W = 0$ ) and treated with NeverWet<sup>®</sup> ( $W = 1$ ) that increases the contact angle by approximately 60°–80°. An untreated and treated Fiber A struck by a drop is shown in Figs. 9a and 9b, respectively. The hydrophobic coating increases the transverse spread of the bifurcating drop and thins the resulting symmetric lobes. One might expect this increase in lobe spreading would require a greater impact force imparted to the fiber, as increased spreading angles correspond to greater momentum transition from vertical to horizontal. Furthermore, drops more prone to breakup impart less momentum to cantilevered fibers (Comtet et al., 2016). As seen in Fig. 9c-d, hydrophilic fibers experience greater  $\delta_0$  across impact Weber number  $\text{We} = \rho V^2 D/\sigma = 100\text{--}600$ , where water density  $\rho = 1$  g/mL and surface tension  $\sigma = 72.9$  dyne/cm. The physics underlying this observation are not well-understood for such ephemeral and coupled three-dimensional flows.

We measure the drops' contact time with Fiber A from the controlled parameter trial for both wetting conditions. Contact time is the time interval from first contact to the moment when the passing drop is fully separated from the fiber or the portion of mass left attached to the fiber (Lebanoff and Dickerson, 2020; Kim and Kim, 2016). Contact time is plotted against Weber number in Fig. 10a. For a fixed fiber length, contact time for hydrophobic ( $W = 0$ ) fibers is



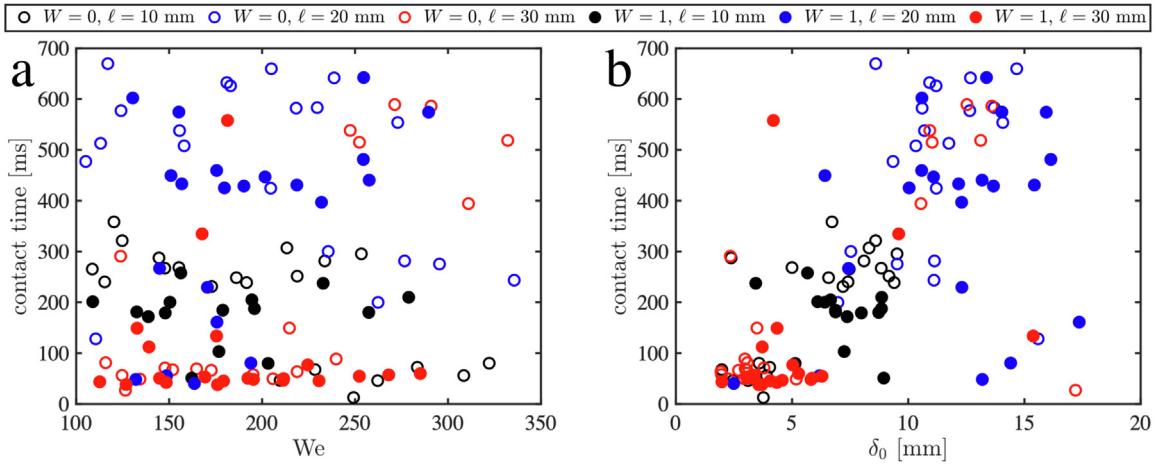
**Fig. 9.** (a) Drop impact on hydrophilic Fiber A. (b) Drop impact on hydrophobic Fiber A. (c) Temporal displacement of hydrophilic and hydrophobic Fiber A. The unaltered fiber experiences greater displacement. (d) Non-dimensionalized displacement  $\delta_0/D$  of Fiber A versus Weber number. Hydrophilic fibers experience greater displacement across the entire range of experimental Weber number.

generally longer than hydrophobic fibers ( $W = 1$ ) and independent of Weber number. Since we expect contact time to be a function of drop velocity, its independence of Weber number supports the notion that drop momentum is not dominant in creating deflection (Eq. (8) and Fig. 4). The hydrophilic fiber remains in contact with the drop for greater duration, and thus the drops “pull” downward even after the bulk of their volume is below the fiber. In contrast, drops impacting the hydrophobic fiber are only able to “push”. Thus, we see a generally longer contact time is associated with greater  $\delta_0$  as shown in Fig. 10b. For a fixed value of  $\delta_0$ , Fiber A with  $\ell = 20$  mm has the longest contact time (Fig. 10b), and thus suggests that our data spans a local contact time maximum. At  $\ell = 30$  mm, Fiber A’s inertia dictates a decrease in contact time over our range of experimental Weber number (Fig. 10a).

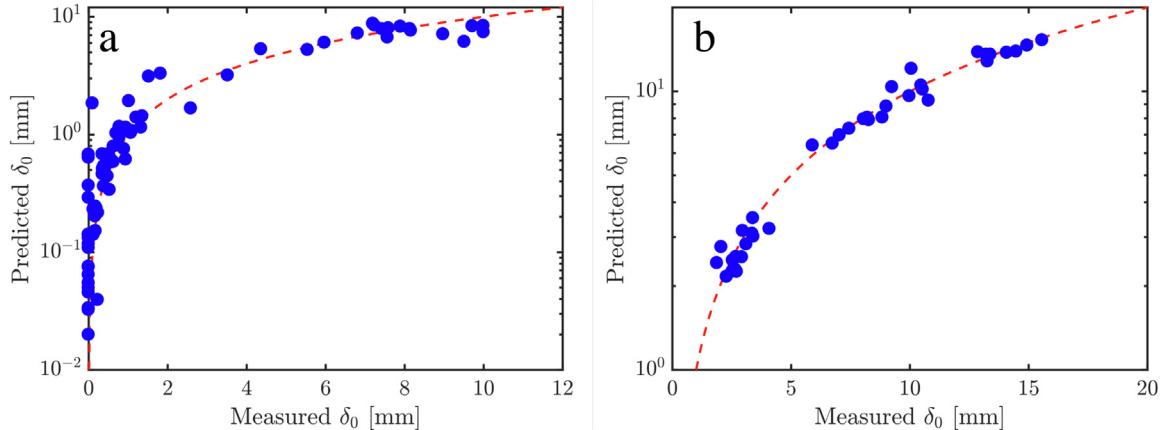
The systematic data collection of the controlled parameter experiment for Fiber A permits a view into the repeatability of impact outcomes. Despite having no two trials with the exact same drop size and velocity, we do record a number of trials that are similar. We show average  $D$ ,  $V$ , and  $\delta_0$  for Fiber A at three lengths in Table S1 or the Online Supplement. Standard deviation values for  $\delta_0$  are typically less than 10% for velocity variations less than 7%. Both wetting conditions show comparable levels of repeatability.

#### 5.4. Ensemble performance

A discussion of ensemble performance must be bifurcated to consider random parameter and controlled parameter datasets separately. A portion of experimental data is withheld from training for the purpose of testing, as discussed in Section 3. We plot predicted  $\delta_0$  against measured  $\delta_0$  in Fig. 11 using a semi-logarithmic scale. A perfectly performing algorithm tested against data devoid of error would produce predictions falling on the curves of slope equal to unity. However, experimental noise and numerical randomness produce variance in prediction. The algorithm performed



**Fig. 10.** Contact time of the falling drop with Fiber A versus (a) Weber number and (b)  $\delta_0$ .



**Fig. 11.** Algorithm prediction performance is seen graphically by comparing predicted versus measured  $\delta_0$  for (a) random and (b) controlled datasets. The dashed line represents exact agreement of predictions and measurements.

**Table 4**  
Algorithm performance scores with 95% confidence interval.

Performance scores	Controlled parameter			
	RFR	MLP	GBoost	Ensemble
$R^2$	0.96	0.97	0.97	<b><math>0.98 \pm 0.001</math></b>
RMSE	0.72	0.58	0.66	<b><math>0.57 \pm 0.02</math></b>

marginally better for the controlled (Fig. 11b) than random parameter dataset (Fig. 11a), which tends to under-predict  $\delta_0$ . The  $R^2$  score of the ensemble algorithm using the random dataset is  $0.94 \pm 0.01$  and using the controlled data is  $0.98 \pm 0.001$ . Individual algorithmic results are tabulated in Table 4, where we show  $R^2$  and root-mean-square-error (RMSE) scores using the controlled dataset. The ensemble of the three algorithms outperforms the individual performances. The use of ensemble learning reduces the probability that results are biased by the choice of a single algorithm.

## 6. Conclusion and future work

In this study, we observe drop impact onto flexible, cantilevered fibers and record their maximal displacement. This simple system has complex, highly coupled dynamics in fluid and solid domains making the outcome of fiber displacement difficult to predict. Using experimental data, we train an ensemble machine learning algorithm to predict fiber displacement across a range of six experimental variables: drop size and velocity, fiber diameter, modulus, length,

and wettability. Our ensemble approach consists of three base learners as the ensemble committee: the Random Forest Regressor, the Multi-Layer Perceptron, and the Gradient Boosting Regressor. The weight of the prediction produced by each base learner was determined by non-negative least squares. Changes in fiber length have the greatest influence on fiber displacement while fiber wettability has the least. Scaling relations developed by momentum transfer do not well capture the relationship between variables in the experimental system in this range of Ohnesorge number and drop velocity. Fiber properties rather than drop properties dominate maximal fiber displacement following drop impact. Our results indicate the physics of this seemingly simple system remain under-characterized and traditional scaling techniques are sufficient only over a small range of dynamical conditions. Greater experimental data would enhance predictive performance and future directions of research include more granularity in the experimental range of surface chemistry, drops of different fluid properties, and decoupling fiber diameter from bending stiffness. Though beyond the scope of the current study, predictive tools such as the one developed and applied here may drive additional experiments near inexplicable predictions to refine predictions or discover new physics at true local minima/maxima.

## CRediT authorship contribution statement

**Panporn Orkweha:** Methodology, Software, Formal analysis, Investigation, Data curation, Writing - original draft. **Alexis Downing:** Methodology, Formal Analysis, Investigation, Data curation, Writing - original draft. **Amy P. Lebanoff:** Methodology, Formal analysis, Investigation, Data curation. **Sharare Zehtabian:** Methodology, Software, Data curation, Writing - original draft, Writing - review & editing. **S. Safa Bacanli:** Methodology, Software, Data curation, Writing - original draft, Writing - review & editing. **Damla Turgut:** Conceptualization, Methodology, Software, Data curation, Writing - original draft, Writing - review & editing, Supervision, Project administration, Funding acquisition. **Andrew K. Dickerson:** Conceptualization, Methodology, Validation, Formal analysis, Data curation, Writing - original draft, Writing - review & editing, Visualization, Supervision, Project administration.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data access

Code, raw videos and .xls files containing experimental measurements are available in perpetuity via Open Science Framework: <https://osf.io/t7e6k/>.

## Acknowledgment

We would like to thank the National Science Foundation, United States of America (CNS-1852002 and CBET-1941341) for support.

## Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.jfluidstructs.2021.103253>.

## References

Aanonsen, S.I., Nævdal, G., Oliver, D.S., Reynolds, A.C., Vallès, B., 2009. The ensemble Kalman filter in reservoir engineering—a review. *SPE J.* 14, 393–412.

Abhinav, N., Prashanth, S., Joseph, S.K., 2017. Temporal clustering for order reduction of nonlinear parabolic PDE systems with time-dependent spatial domains: Application to a hydraulic fracturing process. *AIChE J.* 63, 3818–3831. <http://dx.doi.org/10.1002/aic.15733>.

Adrian, R.J., 2005. Twenty years of particle image velocimetry. *Exp. Fluids* 39, 159–169.

Alam, M.E., Kauffman, J.L., Dickerson, A.K., 2020a. Drop ejection from vibrating damped, damped wings. *Soft Matter* 16, 1931–1940.

Alam, M.E., Wu, D., Dickerson, A.K., 2020b. Predictive modelling of drop ejection from damped, damped wings by machine learning. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* 476, 20200467.

Andrews, H., Eccles, E., Schofield, W., Badyal, J., 2011. Three-dimensional hierarchical structures for fog harvesting. *Langmuir* 27, 3798–3802.

Bengio, Y., Grandvalet, Y., 2004. No unbiased estimator of the variance of k-fold cross-validation. *J. Mach. Learn. Res.* 5, 1089–1105.

Breiman, L., 2001. Random forests. *Mach. Learn.* 45, 5–32.

Brown, G., Wyatt, J., Harris, R., Yao, X., 2005. Diversity creation methods: a survey and categorisation. *Inf. Fusion* 6, 5–20.

Comtet, J., Keshavarz, B., Bush, J.W., 2016. Drop impact and capture on a thin flexible fiber. *Soft Matter* 12, 149–156.

Dickerson, A.K., Hu, D.L., 2014. Mosquitoes actively remove drops deposited by fog and dew. *Integr. Comp. Biol.* 1–6.

Dickerson, A., Mills, Z., Hu, D., 2012a. Wet mammals shake at tuned frequencies to dry. *J. R. Soc. Interface* 9, 3208–3218.

Dickerson, A.K., Shankles, P.G., Hu, D.L., 2014. Raindrops push and splash flying insects. *Phys. Fluids* 26, 027104.

Dickerson, A.K., Shankles, P.G., Madhavan, N.M., Hu, D.L., 2012b. Mosquitoes survive raindrop collisions by virtue of their low mass. *Proc. Natl. Acad. Sci.* 109, 9822–9827.

Dietterich, T.G., 1997. Machine-learning research. *AI Mag.* 18, 97.

Dressaire, E., Sauret, A., Boulogne, F., Stone, H.A., 2016. Drop impact on a flexible fiber. *Soft Matter* 12, 200–208.

Gart, S., Mates, J.E., Megaridis, C.M., Jung, S., 2015. Droplet impacting a cantilever: A leaf-raindrop system. *Phys. Rev. A* 3, 044019.

Gu, Y., Oliver, D.S., 2007. An iterative ensemble Kalman filter for multiphase fluid flow data assimilation. *SPE J.* 12, 438–446.

Hobbs, W.B., Hu, D.L., 2012. Tree-inspired piezoelectric energy harvesting. *J. Fluids Struct.* 28, 103–114.

Ilyas, M.A., Swingler, J., 2015. Piezoelectric energy harvesting from raindrop impacts. *Energy* 90, 796–806.

Jansen, J.-D., Bosgra, O.H., Van den Hof, P.M., 2008. Model-based control of multiphase flow in subsurface oil reservoirs. *J. Process Control* 18, 846–855.

Jenks, M.A., A., E., 1999. Plant epicuticular waxes: function, production, and genetics. *Hortic. Rev.* 23, 1–68.

Kim, S.-G., Kim, W., 2016. Drop impact on a fiber. *Phys. Fluids* 28, 042001.

Kim, H.S., Kim, J.-H., Kim, J., 2011. A review of piezoelectric energy harvesting based on vibration. *Int. J. Precis. Eng. Manuf.* 12, 1129–1141.

Krishnan, S.R., Seelamantula, C.S., 2012. On the selection of optimum savitzky-golay filters. *IEEE Trans. Signal Process.* 61, 380–391.

Krymskaya, M., Hanea, R., Verlaan, M., 2009. An iterative ensemble Kalman filter for reservoir engineering applications. *Comput. Geosci.* 13, 235–244.

Kuncheva, L.I., Whitaker, C.J., 2003. Measures of diversity in classifier ensembles and their relationship with the ensemble accuracy. *Mach. Learn.* 51, 181–207.

Kutz, J.N., 2017. Deep learning in fluid dynamics. *J. Fluid Mech.* 814, 1–4.

Lawson, C.L., Hanson, J., 1995. *Solving Least Squares Problems*. SIAM, Philadelphia, [rev. ed.] edition. isbn:0898713560 (pbk.).

Lebanoff, Amy P., Dickerson, Andrew K., 2020. Drop impact onto pine needle fibers with non-circular cross section. *Phys. Fluids* 32 (9), 092113. <http://dx.doi.org/10.1063/5.0019310>.

Lejeune, S., Gilet, T., 2019. Drop impact close to the edge of an inclined substrate: Liquid sheet formation and breakup. *Phys. Rev. Fluids* 4, 053601.

Lorentzen, R.J., Nævdal, G., Lage, A.C., 2003. Tuning of parameters in a two-phase flow model using an ensemble Kalman filter. *Int. J. Multiph. Flow* 29, 1283–1309.

McMillan, M., Liebens, J., Metcalf, C., 2017. Evaluating the bancs streambank erosion framework on the northern gulf of mexico coastal plain. *JAWRA J. Am. Water Resour. Assoc.* 53, 1393–1408.

Melville, P., Mooney, R.J., 2005. Creating diversity in ensembles using artificial data. *Inf. Fusion* 6, 99–111.

Mendes-Moreira, J., Soares, C., Jorge, A.M., Sousa, J.F.D., 2012. Ensemble approaches for regression: A survey. *ACM Comput. Surv.* 45, 1–40.

Minku, L.L., White, A.P., Yao, X., 2010. The impact of diversity on online ensemble learning in the presence of concept drift. *IEEE Trans. Knowl. Data Eng.* 22, 730–742.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al., 2011. Scikit-learn: Machine learning in python. *J. Mach. Learn. Res.* 12, 2825–2830.

Segal, M.R., 2004. Machine Learning Benchmarks and Random Forest Regression. Center for Bioinformatics & Molecular Biostatistics, University of California, San Francisco.

Smith, N.M., Clayton, G.V., Khan, H.A., Dickerson, A.K., 2018. Mosquitoes modulate leg dynamics at takeoff to accommodate surface roughness. *Bioinspiration Biomim.* 14, 016007.

Smith, W., McClean, T., 1989. Adaptive relationship between leaf water repellency, stomatal distribution, and gas exchange. *Am. J. Bot.* 76, 465–469.

Soto, D., De Lariviere, A.B., Boutillet, X., Clanet, C., Quéré, D., 2014. The force of impacting rain. *Soft Matter* 10, 4929–4934.

Takehara, K., Adrian, R., Etoh, G., Christensen, K., 2000. A Kalman tracker for super-resolution piv. *Exp. Fluids* 29, S034–S041.

Tracey, B.D., Duraisamy, K., Alonso, J.J., 2015. A machine learning strategy to assist turbulence model development. In: 53rd AIAA Aerospace Sciences Meeting, p. 1287.

Volk, R., Calzavarini, E., Verhille, G., Lohse, D., Mordant, N., Pinton, J.-F., Toschi, F., 2008. Acceleration of heavy and light particles in turbulence: comparison between experiments and direct numerical simulations. *Physica D* 237, 2084–2089.

Watson, D.A., Souchik, C.J., Weinberg, M.P., Bom, J.M., Dickerson, A.K., 2020. Making a splash with fabrics in hydrophilic sphere entry. *J. Fluids Struct.* 94, 102907.

Watson, D.A., Stephen, J.L., Dickerson, A.K., 2018. Jet amplification and cavity formation induced by penetrable fabrics in hydrophilic sphere entry. *Phys. Fluids* 30, 082109.

Watson, D.A., Stephen, J.L., Dickerson, A.K., 2019. Impacts of free-falling spheres on a deep liquid pool with altered fluid and impactor surface conditions. *J. Vis. Exp.*

Wu, D., Xu, C., 2018. (PDF) Predictive modeling of droplet formation processes in inkjet-based bioprinting. <http://dx.doi.org/10.1115/1.4040619>.

Yarin, A., 2006. Drop impact dynamics: splashing, spreading, receding, bouncing.... *Annu. Rev. Fluid Mech.* 38, 159–192.

Zhang, C., Ma, Y., 2012. Ensemble Machine Learning: Methods and Applications. Springer, <http://dx.doi.org/10.1007/978-1-4419-932-7>.

Zhang, L., Wu, J., Hediili, M.N., Yang, X., Wang, P., 2015. Inkjet printing for direct micropatterning of a superhydrophobic surface: toward biomimetic fog harvesting surfaces. *J. Mater. Chem. A* 3, 2844–2852.