

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/357255036>

Data-driven splashing threshold model for drop impact on dry smooth surfaces

Article in Physics of Fluids · December 2021

DOI: 10.1063/5.0076427

CITATIONS

16

READS

129

4 authors, including:



Maximilian Pierzyna
Delft University of Technology

6 PUBLICATIONS 20 CITATIONS

[SEE PROFILE](#)



David Burzynski
Coldsense Technologies

9 PUBLICATIONS 203 CITATIONS

[SEE PROFILE](#)



Richard Semaan
Technische Universität Braunschweig

74 PUBLICATIONS 642 CITATIONS

[SEE PROFILE](#)

Data-driven splashing threshold model for drop impact on dry smooth surfaces

Cite as: Phys. Fluids 33, 123317 (2021); <https://doi.org/10.1063/5.0076427>

Submitted: 25 October 2021 • Accepted: 07 December 2021 • Published Online: 22 December 2021

 Maximilian Pierzyna,  David A. Burzynski, Stephan E. Bansmer, et al.



[View Online](#)



[Export Citation](#)



[CrossMark](#)

Physics of Fluids

SPECIAL TOPIC: Flow and Acoustics of Unmanned Vehicles

[Submit Today!](#)



Data-driven splashing threshold model for drop impact on dry smooth surfaces

Cite as: Phys. Fluids **33**, 123317 (2021); doi: [10.1063/5.0076427](https://doi.org/10.1063/5.0076427)

Submitted: 25 October 2021 · Accepted: 7 December 2021 ·

Published Online: 22 December 2021



View Online



Export Citation



CrossMark

Maximilian Pierzyna,¹ David A. Burzynski,^{1,2} Stephan E. Bansmer,^{1,2} and Richard Semaan^{1,a)}

AFFILIATIONS

¹Institute of Fluid Mechanics, Technische Universität Braunschweig, Hermann-Blenk-Str. 37, 38108 Braunschweig, Germany

²ColdSense Technologies GmbH, Hermann-Blenk-Str. 37, 38108 Braunschweig, Germany

^{a)}Author to whom correspondence should be addressed: r.semaan@tu-bs.de

ABSTRACT

We propose a data-driven threshold model to redefine the boundary between deposition and splashing for drop impact on dry smooth surfaces. The starting point is the collection and digitization of multiple experimental sources with varying impact conditions. The model is based on the theory of Riboux and Gordillo [Riboux and Gordillo, “Experiments of drops impacting a smooth solid surface: A model of the critical impact speed for drop splashing,” *Phys. Rev. Lett.* **113**, 024507 (2014)] and is obtained by an uncertainty quantification analysis coupled with machine learning. The uncertainty quantification analysis elucidates the relevance of the impact condition uncertainties when estimating the splashing parameter. The proposed threshold model is trained using a support vector machine algorithm variant that includes uncertainty as a hyperparameter. This threshold model is generalized by complexity reduction and is eightfold cross-validated on the reference data. The results reveal a dependency of the splashing threshold on the impact velocity, the liquid viscosity, the surface tension, and the gas density. Detailed quantification of the prediction performance and a comparison with state-of-the-art models show that the proposed threshold model is the most accurate model to describe the boundaries between deposition and splashing for a wide range of impact conditions. The simplicity and accuracy of this model make it an alternative to existing approaches.

© 2021 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>). <https://doi.org/10.1063/5.0076427>

I. INTRODUCTION

Drop impact on dry smooth surfaces is a common phenomenon in nature and is relevant to many applications, such as coating processes, additive manufacturing, vehicle soiling, and aircraft icing.^{3–6} One critical aspect is the ability to predict the outcome of the impact for a wide range of impact conditions. The outcome of drop impact on a smooth dry surface can be categorized into deposition and splashing, as shown in Fig. 1. Splashing is a complex dynamic phenomenon that can be further classified into corona and prompt splash.⁷ These outcomes depend on multiple parameters, such as the drop diameter D_0 , impact velocity V_0 , the physical properties of the liquid, and the surrounding gas. Identifying the proper combination of impact conditions that lead to a transition from one outcome to the other is essential to understand the mechanism responsible for this phenomenon.

Several studies have been carried out over the last decades to identify the splash mechanism and to define the splashing threshold, as shown in the comprehensive reviews of Marengo *et al.*,⁸ Josserand and Thoroddsen,⁹ and Yarin *et al.*¹⁰ The most prominent models are the *K*-Parameter model proposed by Mundo *et al.*¹¹ the stress model

from Xu *et al.*,¹² and the more recent aerodynamic model of Riboux and Gordillo.¹ The *K*-Parameter is the most commonly used model today due to its simple formulation, which balances the viscous forces, inertial forces, and surface tension $K = \text{OhRe}^{5/4}$. Here, the Ohnesorge number $\text{Oh} = \mu_l/\sqrt{\rho_l D_0 \sigma}$ and the Reynolds number $\text{Re} = \rho_l D_0 V_0 / \mu_l$ are based on the drop liquid (subscript *l*). Xu *et al.*¹² interpret the splashing as the result of a weak shock in the surrounding gas (subscript *g*) caused by the acceleration of the spreading lamella. The stress on the lamella is estimated as $\Sigma = 0.5M^{-1}\text{WeRe}^{-1/2}\rho_g/v_l$, where *M* is the Mach number. Riboux and Gordillo—hereafter RG—developed a more physical theory that attributes the splashing to the aerodynamic lift force F_L acting on the spreading lamella. In this model, splashing is expected if the lift force—a combination of a lubrication force and a suction force exerted by the surrounding gas—is significantly larger than the capillary retraction force, i.e., $\beta = \sqrt{F_L/(2\sigma)} > 0.14$. This model has been successfully used to predict splashing for a large number of experiments.^{7,13,14}

The myriads of models and their various approaches illustrate the difficulties in accurately estimating the splashing threshold. This is

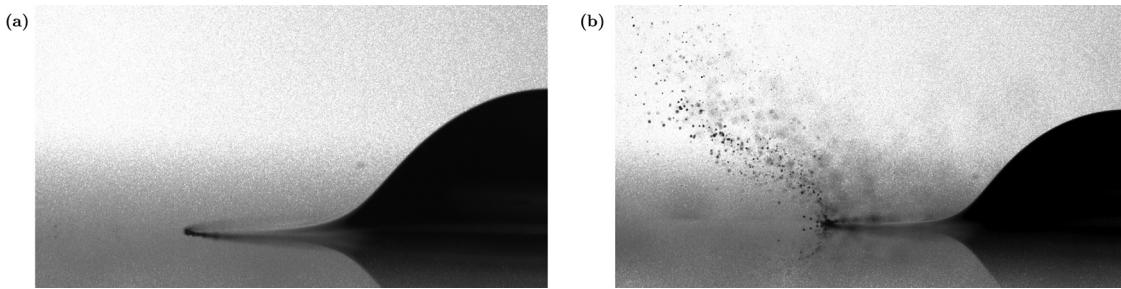


FIG. 1. Water drops with $D_0 \approx 3.7$ mm and velocity $V_0 \approx 10$ m/s impacting a dry smooth surface exhibiting (a) deposition when helium is the surrounding gas and (b) splashing when CO_2 is the surrounding gas. The ambient pressure for both cases is $p = 1043$ hPa. Reprinted with permission from D. A. Burzynski and S. E. Bansmer, “Role of surrounding gas in the outcome of droplet splashing,” Phys. Rev. Fluids 4, 073601 (2019). Copyright 2019 American Physical Society.

exasperated when attempting to predict the threshold at varying impact conditions. The lack of prediction accuracy for general impact conditions could be attributed to the lack of explicit dependencies on the impact and ambient conditions. While Riboux and Gordillo¹ proposed the constant splashing threshold $\beta = 0.14$ for their splashing model, recent studies reported a variable splashing threshold that shifts with the impact conditions.^{2,6,15–17} We note that the RG model does not include any explicit assumptions or limitations on the drop impact conditions. As such, the model failure under certain drop impact conditions cannot be attributed to any modeling assumption. An additional challenge for accurate prediction is the uncertainty of the measured variables and their propagation into the models. However, no comprehensive uncertainty analysis exists of any splashing model in the literature, which could elucidate how the accuracy of the models is affected by the experimental uncertainties.

In this study, we leverage the power of uncertainty quantification and machine learning to construct an accurate RG-based splashing threshold model that is applicable to a wide range of impact conditions. The approach is served by a large set of experimental studies with a wide range of conditions, a comprehensive uncertainty quantification analysis, and a sophisticated machine-learning algorithm. The uncertainty quantification reveals the various sensitivities of the measurement uncertainties on the splashing parameter. The modeling

approach elucidates the relevant physical properties that affect the splashing threshold. The reported model is simple, physically interpretable, and exhibits higher prediction accuracy than the three most-common models.

II. METHODOLOGY

The machine learning splashing threshold model is trained on an extensive database that uses the experimental uncertainties of the splashing parameter as regularization weights. The first step is the data collection and calculation of the splashing parameter using the RG theory (Sec. II A). The second step is to quantify the experimental data’s uncertainties and their propagation into the RG model (Sec. II B). Finally, all the data and their expanded uncertainty are incorporated in a support vector machine (SVM) algorithm to train the model (Sec. II C). Figure 2 summarizes the methodology, which is explained in detail in the current section. A code for the uncertainty propagation and the trained model is available at https://github.com/FlowModelingControl/data_driven_splashing_model.

A. Data collection

Machine learning modeling requires an extensive reference database uniformly sampled across the parameter space with an ensemble

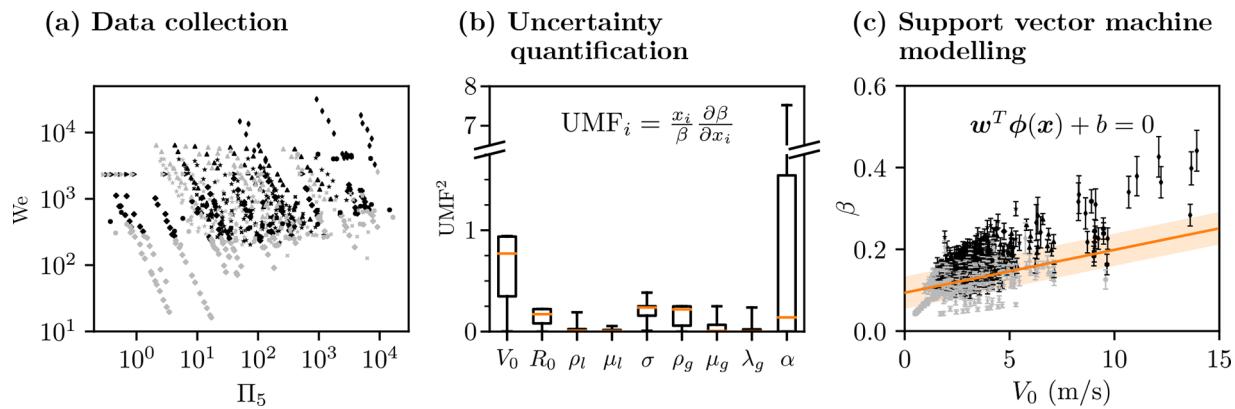


FIG. 2. Schematic of the data-driven three-step approach: (a) data are collected from existing studies. (b) The measurement uncertainties are quantified and propagated into the Riboux and Gordillo¹ model. (c) A soft-margin support vector machine (SVM) classification model is trained on the data using the uncertainties as regularization weights. The solid orange curve illustrates a candidate SVM model with its corresponding margin as a light-orange band. The symbols in (a) and (c) denote the various reference studies listed in Table I, where gray denotes deposition and black splashing upon impact.

size that scales with the number of dependent variables. This is to reduce sample selection bias, which occurs if the data are sampled under similar conditions that neglect a part of the feature space.¹⁸ The feature space of the current problem consists of the impact conditions $\mathbf{x} = (V_0, R_0, \rho_l, \mu_l, \sigma, \rho_g, \mu_g, \lambda_g)^T$, where V_0 is the impact velocity, $R_0 = D_0/2$ the drop radius, ρ and μ are the density and dynamic viscosity of the fluids (l for liquid, g for gas), σ is the surface tension, and λ_g the mean free path of the gas. In addition to the impact conditions, our modeling approach includes the RG splashing parameter β , which depends on the wedge angle α measured between the surface and the ejected lifted lamella at the instant of splashing.^{1,19} The angle depends on the wetting properties of the impact substrate but is typically assumed constant for impacts on solid, smooth, dry surfaces with an experimentally estimated value of $\alpha = 60^\circ$.^{1,15,19} A recent study, however, suggests that the angle could be subject to small variations even for solid, smooth, dry surfaces.²⁰ We treat this variation as an additional uncertainty that propagates into β , and account for it during our uncertainty analysis. However, since this study and the corresponding collected data only focus on solid, smooth, dry surfaces (with $\alpha \approx 60^\circ$ for all samples), we do not explicitly include α as a model variable of the splashing threshold.

The experimental data are selected from existing studies on drop impact on dry smooth surfaces with a mean surface roughness $R_a \approx \mathcal{O}(10 \text{ nm})$.^{2,7,21–24} In addition to requiring a smooth impact surface, three data selection criteria are employed: first, the boundary between deposition and splash on dry smooth surfaces has to be explicitly reported. Second, the experiments have to be conducted using cameras with a relatively high spatial resolution to ensure an acceptably low uncertainty. This criterion is essential for defining the splashing threshold correctly since the ejected secondary droplets during the lamella breakup can be very small.²⁵ Third, the study must solely report on spherical or quasi-spherical drops by either discarding non-spherical samples^{22–24} or ensuring spherical samples through a careful setup.^{1,12,21} This requirement is important, as non-spherical drops exhibit different splashing behavior than spherical ones.²² The

relevant studies that met these criteria are summarized in Table I and are presented in Fig. 2, where $\Pi_5 = \rho_g \sigma^2 D_0 / (V_0 \mu_l^3)$ is a non-dimensional parameter derived from the π -theorem dimensional analysis on the feature space \mathbf{x} , as detailed in Appendix A. The advantage of this parameter is that it incorporates the influence of the surrounding gas density and, therefore, provides better visualization of the range of impact conditions. A full overview of the non-dimensional range of the collected data is presented in Appendix B.

B. Uncertainty quantification analysis

The uncertainty quantification is central to the current work since it elucidates the effect of uncertainty propagation into the RG theory and highlights the quantities that strongly influence the uncertainty of the splashing parameter β . In this section, we present the uncertainty propagation method (Sec. II B 1), which yields the uncertainty magnification factors (UMFs) discussed in Sec. II B 2.

1. Taylor expansion method for uncertainty propagation

The uncertainties of the dependent variables are propagated to β using the Taylor expansion method²⁷ to the first order. This approach is more straightforward than Monte Carlo methods and has the additional benefit of delivering each variable's contribution to the result uncertainty, i.e., the threshold parameter. For more details about this method, the reader is referred to existing literature on the subject.^{27–29} The uncertainties of the impact conditions \mathbf{x} are propagated to the threshold parameter β through the Riboux and Gordillo¹ model, which reads

$$\beta = \left(\frac{K_l \mu_g V_t + 0.3 \rho_g H_t V_t^2}{2\sigma} \right)^{1/2}, \quad (1)$$

where $K_l \approx -(6/\tan^2 \alpha)[\ln(19.2\lambda_g/H_t) - \ln(1+19.2\lambda_g/H_t)]$, $V_t = 1/2\sqrt{3/t_e}V_0$, and $H_t = \sqrt{12t_e^{3/2}R_0}/\pi$. The wedge angle α in the K_l definition is measured between the surface and the lifted lamella at the

TABLE I. List of collected references, including the range of test conditions and their respective expanded uncertainties. Superscript (*) denotes the estimated uncertainty not reported by the reference authors. A general uncertainty of $\pm 1\%$ is assumed for all liquid and gas properties.

	Ref.	D_0 (mm)	V_0 (m/s)	ρ_l (kg/m ³)	μ_l (mPa s)	σ (mN/m)	ρ_g (kg/m ³)	μ_g (μPa s)
▲	Xu, Zhang, and Nagel ¹²	3.4 ± 0.1	(2.0 to 7.0) ± 0.05	780 to 790	0.58 to 2.0	21.7 to 22.7	0.01 to 6.6	15.8 to 25.6
✗	Vander Wal, Berger, and Mozes ²¹	$2.0 \pm 0.1^*$	(1.4 to 4.2) ± 0.15	684 to 1080	0.4 to 3.3	20.1 to 72.8	1.2	18.2
►	Xu ²⁶ (smooth surface impacts only)	3.1 ± 0.1	4.0 ± 0.05	820 to 950	1.0 to 18	17.4 to 21	0.01 to 6.6	15.8 to 25.6
◆	Palacios <i>et al.</i> ²²	(1.6 to 3.8) ± 0.1	(0.5 to 4.8) $\pm 2.5\%^*$	786 to 1175	0.5 to 20.4	20.5 to 70.8	1.2	18.2
●	Riboux and Gordillo ¹	(1.7 to 3.9) $\pm 0.1^*$	(1.5 to 4.2) $\pm 2.5\%^*$	789 to 1000	0.3 to 10	17.2 to 71.8	0.16 to 6.0	15.3 to 25.1
★	Stevens ²³	(1.6 to 4.0) ± 0.1	(0.1 to 5.5) $\pm 2.5\%^*$	750 to 1860	0.5 to 2.7	16 to 67	0.01 to 6.6	15.8 to 19.6
▼	Hao ²⁴	(2.3 to 3.8) ± 0.1	(1.0 to 5.1) $\pm 2.5\%^*$	791 to 999	1.0 to 1.8	22.9 to 72.9	1.2	18.2
+	De Goede <i>et al.</i> ¹⁵	1.8 ± 0.1	(2.0 to 4.5) $\pm 2.5\%^*$	789 to 997	0.9 to 2.9	21.8 to 72	1.2	18.2
●	Burzynski and Bansmer ²	3.7 ± 0.4	10 ± 0.6	998	1.0	72.8	0.2 to 3.2	12.7 to 21.7
◆	Burzynski, Roisman, and Bansmer ⁷	(2.0 to 4.0) ± 0.4	(3.0 to 26) ± 0.6	790 to 998	0.3 to 1.2	22.6 to 72.8	1.2	18.2

time of splashing. It is typically assumed constant at $\alpha = 60^\circ$.^{1,15,19} However, a recent study suggests that small variations of $\pm 3.6^\circ$ might occur,²⁰ which is the expanded uncertainty we assume for α . In the RG theory, t_e is defined as the ejection time of the lamella, which can be interpreted as the time when the ejection of secondary droplets begins. This time depends on the physical properties of the liquid and can be either calculated numerically for a general drop impact or estimated for high-viscosity liquids $\mu_l > 10 \text{ mPa s}$ as $t_e = 2\text{Re}_R^{-1/2}$.¹⁹ Note that Re_R is defined using the drop radius R_0 as a characteristic length and not the diameter.

We quantify the propagated relative uncertainty using the Taylor expansion method to the first order as³⁰

$$\left(\frac{u_\beta}{\beta}\right)^2 = \sum_{i=1}^I \left(\frac{x_i}{\beta} \frac{\partial \beta}{\partial x_i}\right)^2 \left(\frac{u_{x_i}}{x_i}\right)^2, \quad (2)$$

where u_{x_i} is the uncertainty of a *mean* measured variable x_i . Here, we assume statistical independence among the measured variables.

Equation (2) can be used to propagate either random or systematic uncertainty, which are then combined to form the combined standard uncertainty of the mean as $u = (\sigma_X^2 + \zeta^2)^{1/2}$. σ_X is the standard deviation of the mean, $\sigma_X = \sigma/\sqrt{N}$, where σ is the standard deviation, and N is the number of independent samples.³¹ ζ is the total systematic uncertainty of a measured variable x_i . We present all uncertainties with 95% confidence level, which is commonly used in the engineering sciences. We report the uncertainty using the combined *expanded* uncertainty $U = 2u$, where u is the combined standard uncertainty of the mean. The factor 2 is based on the large-sample record assumption for 95% confidence level.

The first requirement for the uncertainty propagation is the quantification of the individual random and systematic uncertainties. Unfortunately, from the employed reference data, only a few studies reported their uncertainties in detail. Notably lacking are the systematic uncertainties. For these data, we estimate the missing systematic uncertainties based on the setup and equipment used. Specifically, we assume an expanded uncertainty of $U_{D_0} = \pm 0.1 \text{ mm}$ for the drop diameter (corresponding to U_{D_0}/D_0 between $\pm 2.5\%$ and $\pm 6.25\%$) and a relative expanded uncertainty of $U_{V_0}/V_0 = \pm 2.5\%$ for the impact velocity. For the physical properties of the fluids, we assume an expanded relative uncertainty of $\pm 1\%$ of the reported values,^{32,33} and $U_\alpha/\alpha = \pm 6\%$ for the wedge angle.²⁰ All estimated uncertainties are denoted with (*) in Table I.

2. Uncertainty magnification factors

The uncertainty magnification factor (UMF) for a given measured variable x_i indicates the influence of the uncertainty in that variable on the uncertainty in the splashing parameter. A squared UMF value greater than 1 indicates that the influence of the variable's uncertainty is magnified as it propagates through the data reduction equation, i.e., the RG model, into β . A squared UMF value of less than 1 indicates that the influence of the uncertainty in the variable is diminished as it propagates through the data reduction equation into β . The UMFs are directly inferred from Eq. (2) as³⁰

$$\text{UMF}_i = \frac{x_i}{\beta} \frac{\partial \beta}{\partial x_i}, \quad (3)$$

The analytical expressions are listed in Appendix C. The individual UMF expressions for all nine measured variables are numerically evaluated using the existing 647 experimental samples and 100 000 synthetically generated random drop impacts from a domain 20% larger than the available feature space. The simulated impact conditions are randomly sampled from a multi-dimensional uniform probability distribution as proposed by McKay *et al.*³⁴ The exact bounds of the synthetic domain are presented in Table II of Appendix C. Specifically, 100 000 random vectors \mathbf{x}_i are randomly drawn from a multi-dimensional uniform distribution. The large sample size and the uniform distribution ensure that all features, i.e., all drop impact variables, are considered equally important avoiding one or more features dominating the results. Each random vector consists of the nine impact parameters, thus representing a simulated impact experiment. Note that the splashing outcome of the synthetic data are not simulated, since it is irrelevant to the uncertainty analysis. We emphasize that these additional synthetic samples are *only* utilized in the uncertainty propagation analysis to enable the examination of the RG model's sensitivities more accurately and over a broader range than would be possible solely based on the experimental data.

Three select UMF distributions are shown in Figs. 3(a)-3(c), while the remaining ones are presented in Appendix C. All distributions except $(\text{UMF}_\alpha)^2$ show UMF values less than one, indicating that the individual uncertainties are damped as they propagate through Eq. (1). The color bars correspond to various fluid properties. The strong uncertainty magnification of $(\text{UMF}_\alpha)^2$ shown in Fig. 3(a) exhibits a decaying trend with increasing impact velocity and increasing gas density. This is observed in both the simulated as well as the collected experimental (red markers) samples. U_α is magnified for almost all impacts with $V_0 < 7.5 \text{ m/s}$ in low-density gases, such as air, which includes a large subset of our data. This finding stresses the importance of accurately determining α to mitigate strong uncertainty magnification. To our knowledge, only De Goede *et al.*¹⁵ have explicitly measured α , albeit with a relative uncertainty of $U_\alpha = \pm 8\%$. Although $(\text{UMF}_{V_0})^2$ in panel (b) reveals that velocity uncertainties are not magnified, it still shows that U_{V_0} contributes to the propagated uncertainty. The observed ridge in the colored properties in panel (b) is due to the RG model discontinuity of the ejection time at liquid viscosity $\mu_l = 10 \text{ mPa s}$. Contrary to $(\text{UMF}_\alpha)^2$ in (a), a *decreasing* gas density ρ_g has a dampening effect on $(\text{UMF}_{V_0})^2$ (not shown). For a very low-impact velocity range $V_0 < 0.2 \text{ m/s}$ with high viscosities $\mu_l > 10 \text{ mPa s}$, $(\text{UMF}_{V_0})^2$ is larger than one, indicating uncertainty magnification. However, the splashing parameter in that range amounts to $\beta < 0.032$. This suggests a weak aerodynamic lift force compared to the capillary force so that splashing is not expected.³⁵ Hence, such cases are not practically relevant to the RG splashing model and are therefore neglected. For larger drop impact velocities, $(\text{UMF}_{V_0})^2$ gradually increases toward a value of 0.92, which suggests that its associated uncertainty becomes less damped with the increasing impact velocity. In other words, U_{V_0} becomes more relevant with increasing V_0 . Since the opposite is true for U_α , the propagated uncertainty U_β is therefore dominated by U_α at low velocities $V_0 < 10 \text{ m/s}$ and by U_{V_0} for higher velocity impacts.

The sensitivity analysis of all impact parameters demonstrates that the remaining UMFs are much smaller, as shown in Appendix C. One example of these weak uncertainty propagators is presented in Fig. 3(c) for the drop radius, which shows the weaker RG model

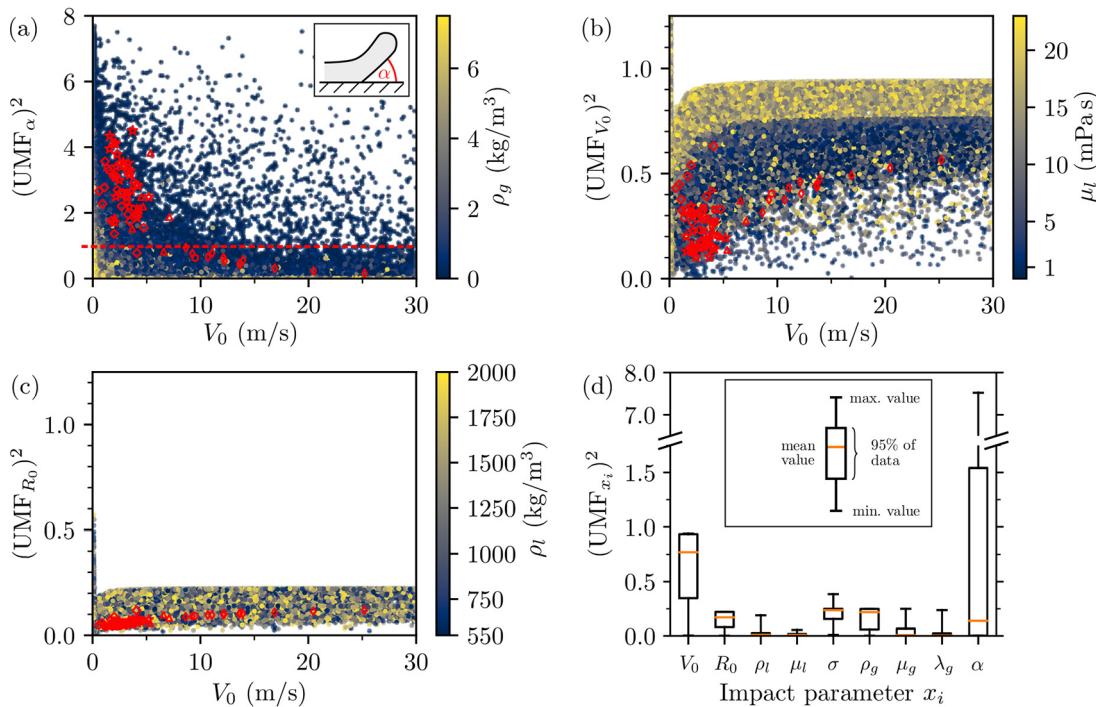


FIG. 3. Uncertainty magnification factor distributions of (a) $(UMF_\alpha)^2$ and (b) $(UMF_{V_0})^2$ over a range of impact velocities V_0 . Values above unity (dashed red line) in (a) indicate uncertainty magnification. The color-coding illustrates dependencies on ρ_g and μ_l in (a) and (b), respectively. The low sensitivity of U_{R_0} on β over the considered velocity range is illustrated in (c). A summary of all UMF distributions is presented as a boxplot in (d). Red symbols denote selected experimental samples.

sensitivity to R_0 uncertainties than for, e.g., V_0 uncertainties. The uncertainty propagation sensitivity analysis for all impact variables is summarized as boxplots in Fig. 3(d). The whiskers of $(UMF_{x_i})^2$ for each parameter x_i represent the minimum and maximum values of the data, whereas the orange lines represent the mean values over all drop impacts. As Fig. 3(d) shows, the mean UMF^2 values for ρ_b , μ_b , μ_g , and λ_g are close to zero. This suggests that the uncertainties of these parameters play almost no role in the propagated uncertainty into β .³⁶ On the other hand, V_0 , R_0 , σ , ρ_g , and α should be measured or estimated accurately since their uncertainties affect that of β .

C. Machine learning modeling

The collected data offer a unique opportunity to construct a generalized model for the splashing threshold using machine learning. We target an analytical, interpretable, and sparse model that extends the RG theory. Sparsity is desired as it reduces complexity and enables generalization of the constructed model. In other words, we seek a multi-dimensional hyperplane that best separates the splashing and deposition classes. The hyperplane is determined using a linear support vector machine (SVM) with a soft margin.^{37,38} The SVM decision function is expressed as

$$y(\mathbf{x}) = \mathbf{w}^T \phi(\mathbf{x}) + b, \quad (4)$$

where \mathbf{w} are weights, b is the offset of the linear separation plane, and ϕ is a function that transforms the impact conditions \mathbf{x} . This

transformation enables non-linear dependencies between the impact parameters and the splashing threshold. We employ dimensional variables rather than non-dimensional ones because the latter introduce correlations between the features, which complicates the model training process and typically yield less accurate models. Since we aim to extend and improve the RG model, we also include the experimentally calculated β as a feature into \mathbf{x} . This also ensures that our model captures all the physics implied in the splashing parameter. We note that the wedge angle α is not included as a feature, because it is not measured experimentally and is usually assumed constant. Moreover, preliminary analysis of the data encouraged us to employ $1/\mu_l$ instead of μ_l as a feature. This choice was subsequently justified by a preliminary feature elimination process which assigned higher importance to $1/\mu_l$ than μ_l .

The optimal splashing threshold is found when the soft margin between the plane and the closest experimental data vectors is maximized. The soft margin around the decision function is used to tolerate some classification errors ξ_i .³⁷ This is achieved by solving the expression

$$\min_{\mathbf{w}, b, \xi} \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \sum_{i=1}^n \xi_i, \quad (5)$$

where C is the constant regularization coefficient that controls the soft margin. This optimization problem is constrained by $y_i(\mathbf{w}^T \phi(\mathbf{x}_i) + b) \geq 1 - \xi_i$, $i = 1, \dots, n$ and $\xi_i \geq 0$, $i = 1, \dots, n$, where y_i are the observed impact outcomes (deposition or splashing) from the

experimental data. The splashing threshold is obtained by solving $y(\mathbf{x}) = 0$ for β using the optimization result.

In this study, we seek to include the propagated measurement uncertainties U_β in the modeling process. As such, we repurpose the SVM algorithm of Chang and Lin,³⁸ and employ a regularization vector $\mathbf{c} = C\omega$, $\omega \in \mathbb{R}^n$ instead of the scalar regularization coefficient C in Eq. (5). This method allows to control the importance of each drop impact measurement \mathbf{x}_i individually by assigning a sample weight $\omega_i \in [\gamma, 1]$, which we define as

$$\omega_i = 1 - \gamma \left(\frac{U_{\beta,i} - U_{\beta,\min}}{U_{\beta,\max} - U_{\beta,\min}} \right), \quad (6)$$

with

$$\gamma > 0, \quad U_{\beta,\min} = \min_{j=1,\dots,n} U_{\beta,j}, \quad \text{and} \quad U_{\beta,\max} = \max_{j=1,\dots,n} U_{\beta,j}. \quad (7)$$

The scaling assures that drop impacts with high measurement uncertainties receive a low sample weight, while experiments carried out with small uncertainty receive a larger weight. γ is introduced to avoid cases where the sample weights become $\omega = 0$, which would eliminate samples from the training data. In the current study, the scaling parameter is arbitrary set to $\gamma = 0.5$.

To decrease the risk of overfitting, the model coefficients are identified using eightfold cross-validation (CV) and the model complexity is systematically reduced. Cross-validation splits the training data into eight equally sized folds. The first fold is set aside as a validation set while the model is trained on the seven remaining folds. Due to the relatively low amount of available samples, we refrain setting aside separate testing data.³⁹ Moreover, we are primarily interested in a model that is applicable within the range of the collected data, which is sufficiently broad. A visual illustration of the model applicability range is provided in Fig. 8 of Appendix B. Extrapolation beyond this range would likely incur prediction uncertainty. The model performance is

then evaluated using the accuracy metric on the validation set,³⁸ which is defined as

$$\text{acc} = \frac{n_{\text{tp}} + n_{\text{tn}}}{n_{\text{tp}} + n_{\text{fp}} + n_{\text{tn}} + n_{\text{fn}}}, \quad (8)$$

where n_{tp} and n_{tn} are testing samples that are classified correctly as the true positives and the true negatives, respectively. The numbers of false positive and the false negative classifications are denoted by n_{fp} and n_{fn} , respectively. The cross-validation process is repeated until all folds are employed as a validation set once. This yields eight performance scores that are averaged to obtain a single accuracy score that estimates the out-of-sample error, i.e., the model prediction error on new data.⁴⁰

Model sparsification is achieved by recursive feature elimination, which ranks the relative importance of the features by comparing their squared weights $(w_i)^2$, $w_i \in \mathbf{w}$, and eliminates the one that does not affect the model outcome.⁴¹ The model is then reevaluated using eightfold CV with the reduced set of features. If the reduced model's performance does not change significantly, ranking and elimination are repeated, and a further simplified model is trained and assessed. The recursive feature elimination process on the proposed model is presented in Fig. 4. Figure 4(a) illustrates the relation between model complexity and accuracy. We quantify complexity by the number of model variables, whereas accuracy is quantified with Eq. (8). The sequential complexity reduction shown in panel (a) is performed based on the feature ranking shown in panel (b). Here, the relative importance of β is larger than 70% for all models, making it the most crucial parameter for the splash/no-splash classification. The significant contribution of β is expected, as the RG model already captures most of the drop impact physics. This result reinforces our approach, which retains the RG model and only corrects for the constant threshold assumption. This engenders the exclusion of β as a model-dependent variable. The remaining features affect the model accuracy in a typical manner, with a decrease in prediction accuracy with a decreasing number of dependent variables. The enlarged regions in panels (c) and

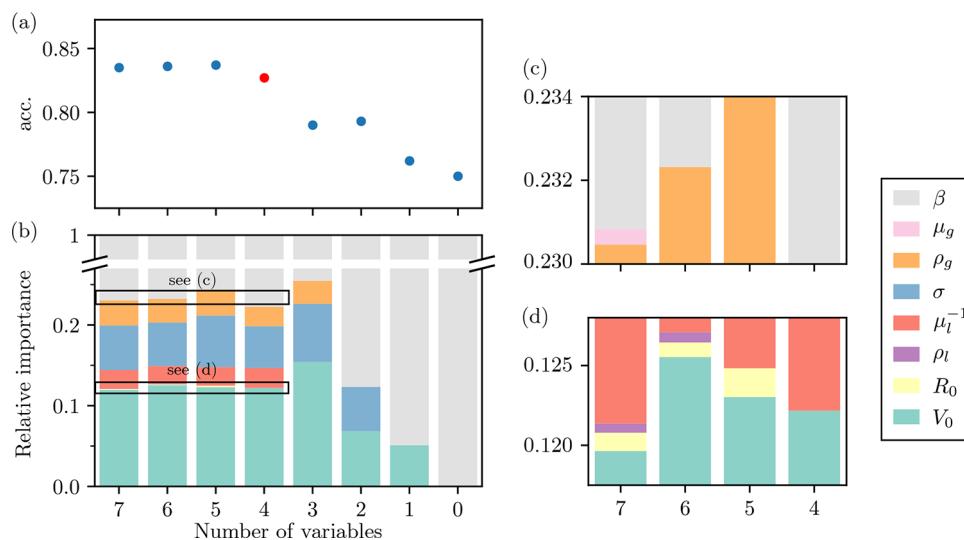


FIG. 4. Model complexity vs accuracy. (a) The evolution of SVM model performance with decreasing number of variables. Feature elimination is performed based on variable ranking shown in panel (b), with two close-ups in (c) and (d). Our model (marked in red) achieves a good compromise between accuracy and complexity.

(d) clearly show the negligible contribution of the first three eliminated variables (μ_g , ρ_b , R_0), which is reflected in the quasi-constant accuracy of the three most complex models. Further complexity reduction to four variables yields a slight decrease in accuracy, which is also manifested in the minor contribution of the eliminated variable R_0 [cf. Fig. 4(d)]. A notable drop in prediction accuracy is observed between four and three, followed by further degradation. Consequently, the model with four remaining features (V_0 , $1/\mu_l$, σ , ρ_g) is selected as the final generalized model. This model provides a good compromise between accuracy and complexity.

III. RESULTS

A. Uncertainty of the splashing parameter β

The detailed uncertainty analysis presented in the following illustrates how even when experimental data are generated with precision, the measurement uncertainties still play a notable role when estimating the impact outcome. This is highlighted in the resulting expanded uncertainty scatter distribution U_β shown in Fig. 5 along with the sample weights ω . The uncertainty scatter exhibits a distribution with an apparent increase with a higher β . There are notable differences among the various studies that employ different experimental setups. Figure 5(a) shows large propagated uncertainty of $U_\beta > 0.025$ for experiments carried out by Burzynski and Bansmer² (●) and

Burzynski *et al.*⁷ (◆), where high drop impact velocities of $V_0 > 10$ m/s are employed (cf. Table I). According to the findings in Sec. II B 2, the contribution of V_0 to the propagated uncertainties dominates in this velocity range. For datasets in the lower velocity range, the propagated uncertainty is dominated by U_x . Despite U_β 's relatively small magnitude, the relative uncertainty in this range is still driven upward $U_\beta/\beta \approx 10\%$ due to the strong magnification of U_x . As detailed in Sec. II C, we incorporate these propagated uncertainties into our modeling process through the sample weights ω , which are also shown in Fig. 5(a).

Figure 5(b) presents the calculated splashing parameter with associated error bars as a function of the impact velocity for a selected data subset. The error bars correspond to the 95% confidence interval of the propagated expanded uncertainty U_β . Multiple data points can be observed in the zoomed-in region near the splashing boundary with error bars crossing the threshold. These points are ambiguous with respect to the constant RG boundary. Such samples illustrate well how uncertainties near the splashing threshold can bias the results and, subsequently, the modeling. The results clearly demonstrate the importance of estimating the splashing parameter uncertainty and incorporating such information into the modeling process. Failure to account for uncertainties can lead to erroneous predictions.

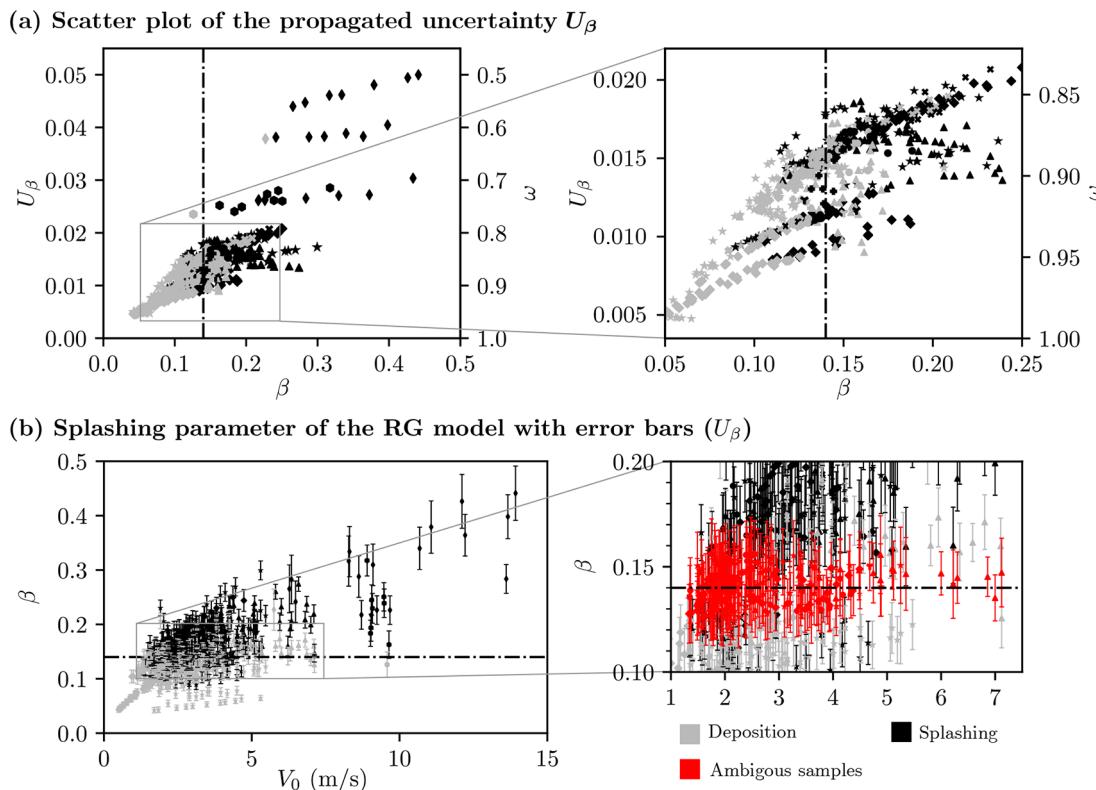


FIG. 5. Propagated uncertainties of the RG model for the collected experimental data (symbols cf. Table I). (a) Scatter plot showing the propagated uncertainties U_β and the corresponding sample weights ω over a range of splashing parameters. The dashed line denotes the constant threshold $\beta = 0.14$. (b) The calculated splashing parameters with associated error bars are presented for 95% confidence level as a function of the impact velocity for a selected subset. Red denotes ambiguous samples with error bars crossing the threshold.

B. The data-driven splashing threshold model

The proposed machine learning methodology enables constructing an accurate data-driven splashing threshold (DST) model applicable to a wide range of impact conditions. This DST model defines the boundary separating deposition from splashing in cases of drop impact on dry smooth surfaces as

$$\beta_{\text{DST}} = c_0 + c_1 V_0 + \frac{c_2}{\mu_l} + c_3 \sigma + c_4 \rho_g, \quad (9)$$

with $c_0 = 0.1086$, $c_1 = 9.5836 \times 10^{-3} \text{ (m/s)}^{-1}$, $c_2 = 1.7332 \times 10^{-5} \text{ Pa s}$, $c_3 = -7.0521 \times 10^{-1} \text{ (N/m)}^{-1}$, and $c_4 = 1.661 \times 10^{-2} \text{ (kg/m}^3)^{-1}$. The model represents a correction to the constant threshold $\beta = 0.14$ and retains the physics of the RG theory in the definition of the splashing parameter β . It can be applied over the entire range of training data listed in Table I and illustrated in a non-dimensional form in Appendix B. We note that using our threshold model outside the training range is possible but would likely incur larger prediction uncertainty, as typically associated with extrapolation.

The model shows a linear dependency between β and V_0 , σ , and ρ_g and an inversely proportional relation between β and μ_l . We note that the latter dependency is explicitly incorporated into the model by using $1/\mu_l$ instead of μ_l during the model training phase. This is justified by a preliminary analysis of feature importance as detailed in Sec. II C. Splashing is predicted for $\beta > \beta_{\text{DST}}$, where β and β_{DST} are individually calculated for each drop impact condition. The proposed threshold model exhibits only weak dependency on R_0 , ρ_b , μ_g , and λ_g as inferred by the recursive feature elimination process. The weak dependency of the splashing boundary on the gas viscosity μ_g and the mean free path length λ_g have been recently reported by Burzynski and Bansmer.² We emphasize that the eliminated variables only

suggest their irrelevance to the general drop impact conditions beyond the expression for β , which is a function of all drop impact variables. In other words, the RG theory accurately captures the eliminated variables' influence on the splashing parameter, rendering them superfluous to the correction function.

The model results are presented in Fig. 6 for the four dependent variables. Also shown are the collected experimental samples colored gray for deposition and black for splashing. The reference data clearly demonstrate that a constant splashing threshold does not correctly represent the boundary between deposition and splashing for a wide range of impact conditions. On the contrary, the proposed threshold model is capable of capturing the different β dependencies of each impact parameter. For example, Fig. 6(a) shows the linear increase in β_{DST} with impact velocity. Such an increasing splashing threshold has been reported in the literature, where the boundary for low-speed impacts is determined as $\beta \approx 0.11$,^{22,23} and as $\beta \approx 0.19$ for high-speed impacts.⁷ Similar linear behavior is observed with the gas density, albeit with a smaller slope, as shown in Fig. 6(b). The behavior is also in line with the observations made by Stevens,²³ Burzynski and Bansmer,² and the more recent experimental results of Ashida *et al.*,¹⁷ where splashing was observed even at low ambient pressure. Since the RG model requires the physical properties of the gas as input parameter and not the pressure, we calculate the gas density from the gas state at the given pressure as $\rho_g \propto p$ for the experiments conducted with variable pressure.

Contrary to the previous trends, the splashing threshold decreases linearly with increasing liquid surface tension. An increase in surface tension preventing splashing has been reported for low-speed impacts and observed especially in the latest stages of impact when the remaining liquid over the surface contracts avoiding its breakup.¹⁰ It is important to note that there are a few experimental

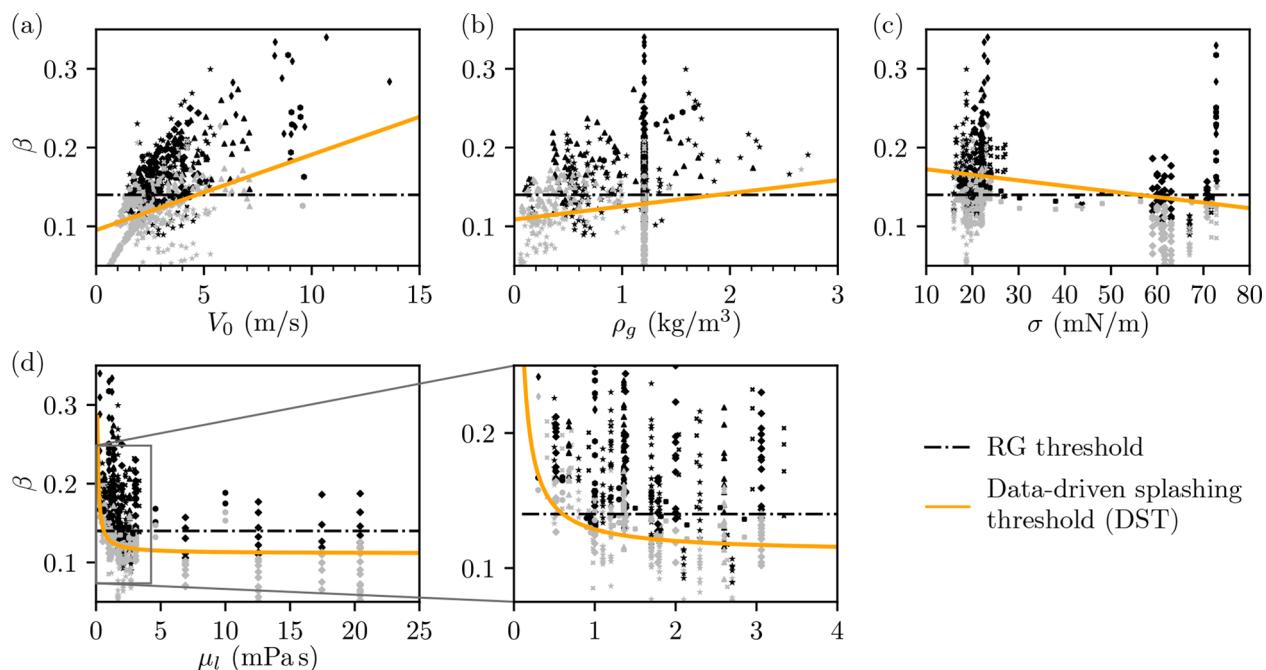


FIG. 6. The generalized data-driven splashing threshold distributions for the four dependent variables: (a) impact velocity V_0 , (b) gas density ρ_g , (c) liquid surface tension σ , and (d) liquid viscosity μ_l . Also shown are the collected experimental samples color-coded with gray for deposition and black for splashing with symbols according to Table I.

samples between $\sigma = 30$ and 60 mN/m to confidently infer any other possible relation between β_{DST} and σ than a linear one. More experimental studies are, therefore, needed in the missing range to reveal other possible relations. Villermaux⁴² remarked that the splashing may occur when the stress in the impacting drop overcomes surface tension or when instabilities on the lamella induce a drop shape during impact. The latter has been recently demonstrated by Burzynski *et al.*⁷ for high-speed impacts, where the Rayleigh–Taylor instability is shown as the responsible mechanism for prompt splash.

The splashing threshold dependency on the liquid viscosity exhibits a more complex behavior than the previous quantities. Figure 6(d) displays how the splashing threshold decreases rapidly with increasing liquid viscosity up to $\mu_l \approx 3 \text{ mPa s}$ and remains almost constant after that. This highlights how the threshold becomes less sensitive to viscosity variations in the high μ_l regime. This trend is corroborated by different studies for a wide range of viscosities,^{22,23,43} and reconfirms the recent conclusions made by Burzynski *et al.*,⁷ where liquid viscosity was reported to be the most relevant liquid parameter for splashing on smooth dry surfaces. All these dependencies demonstrate the need for an analytical model that depends on the impact conditions.

The DST model performance is assessed and compared against existing ones using the accuracy metric defined in Eq. (8). The proposed model yields $\text{acc} = 0.83$ which indicates good prediction accuracy. A comparison of the prediction performance of the current DST model and that of Riboux and Gordillo¹ using all the collected data samples is presented in Fig. 7. The green symbols in the $\text{We}-\Pi_5$ diagram denote correct classification achieved by both models, whereas the yellow symbols highlight samples that were wrongly classified by the constant threshold and are now adequately predicted. The table in Fig. 7 compares the accuracy of the most common models against the proposed one for all the collected data. The widely used K-Parameter proposed by Mundo *et al.*¹¹ correctly predicts the drop impact outcome for 57% of the samples. The low performance of this splashing model clearly suggests that it is not reliable to make accurate splashing predictions for a large range of impact conditions. The model of Xu *et al.*¹²

can correctly separate deposition from splashing in 74% of the cases, which is a notable improvement to the previous model. However, it incorrectly predicts splashing instead of deposition for 20% of the cases leading to overall lower confidence in the predicted results. The model of Riboux and Gordillo¹ is a further improvement to the first two models. The constant RG threshold correctly predicts the outcome in 76% of the studied cases. Our proposed data-driven splashing threshold (DST) model exhibits the highest prediction accuracy with 83%. It makes incorrect predictions of splashing instead of deposition in only 6% of the cases compared to 12% for the RG model. Further improvements to the model could be achieved by expanding the training dataset, particularly with samples that include measurement uncertainty. Based on its simplicity and accuracy, we demonstrate that the proposed model constitutes an excellent alternative to existing approaches.

IV. CONCLUSIONS

We propose a data-driven splashing threshold model for drop impact on dry smooth surfaces. We leverage existing experimental data, comprehensive uncertainty quantification, and machine learning techniques. The model can be thought of as an expansion to the theory of Riboux and Gordillo,¹ which defines a constant splashing parameter. The model is validated and compared against existing approaches showing clear superiority.

The uncertainty quantification analysis demonstrates that an accurate estimation of the splashing threshold requires consideration of the parameter uncertainty. Corroborating recent findings by Gordillo and Riboux,²⁰ particular attention should be paid to the wedge angle uncertainty that gets magnified for low-impact velocities $V_0 \leq 10 \text{ m/s}$ as it propagates through the RG model. Although the remaining measurement uncertainties are damped, their contribution to the prediction accuracy cannot be neglected, particularly the impact velocity, drop radius, surface tension, and gas density. We conclude that these variables should be carefully measured during experimental investigations.

The machine learning algorithm used to train the model is a support vector machine, where the splashing parameter uncertainty is used as regularization weight. The resultant threshold model is simple, interpretable, and retains the incorporated physics of the RG model. It is generalized by feature reduction through recursive feature elimination, which reduces the model's dependent variables to only four impact conditions: The impact velocity, liquid viscosity, surface tension, and the gas density. The model is validated through an eightfold cross-validation process. The achieved prediction accuracy is 83%, which is significantly better than the K-Parameter (57%), the Xu *et al.*¹² model (74%), and the constant threshold of Riboux and Gordillo¹ (76%).

The analysis of the collected experimental data and the developed model highlight the complex physical phenomenon and dependencies of the splashing parameter β on the impact conditions. The proposed generalized data-driven splashing threshold (DST) model is an accurate expansion to existing methods. It retains the physics-based foundation of the RG model and generalizes the threshold from a constant to an impact condition-dependent function as

$$\begin{cases} \beta \leq \beta_{\text{DST}}, & \text{deposition}, \\ \beta > \beta_{\text{DST}}, & \text{splashing}. \end{cases}$$

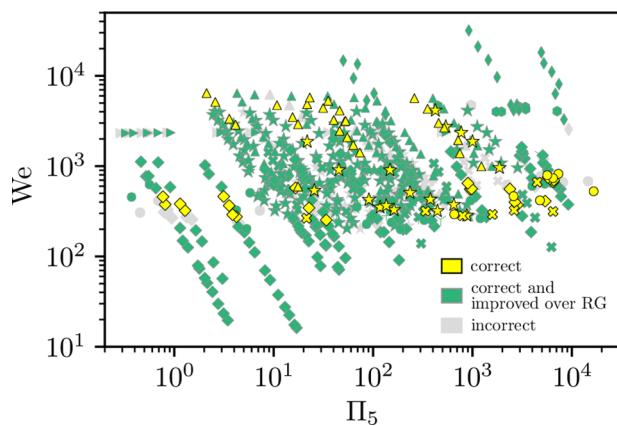


FIG. 7. Performance of the proposed data-driven splashing threshold (DST) model for a wide range of impact conditions. The green symbols mark correct classification achieved by the DST and the constant $\beta = 0.14$ model of RG. The yellow colors represent samples that were predicted correctly only by the DST model. The symbols denote the collected studies listed in Table I. The table summarizes the prediction accuracy of the different models.

We note that due to the limited availability of data, the current model is restricted to small quasi-spherical drops and impacts on dry, smooth surfaces. However, the proposed modeling methodology is flexible and easily expandable to include more dependencies when additional data become available. For example, it can be applied to other drop impact classification problems, such as the breakup regime classification when drops impact hot or superhydrophobic surfaces. Beyond drop impact research, the proposed approach opens new avenues in the general field of data-driven modeling, where uncertain data are used.

ACKNOWLEDGMENTS

This research was supported by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) through financing of Project Nos. BA 4953/3 and FOR 2895.

AUTHOR DECLARATIONS

Conflict of Interest

The authors report no conflict of interest.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request. The data that support the findings of this study are openly available in GitHub at https://github.com/FlowModelingControl/data_driven_splashing_model, Ref. 44.

APPENDIX A: DIMENSIONAL ANALYSIS OF THE DROP IMPACT CONDITIONS

The Buckingham π theorem describes a physical problem with n dimensioned parameters given in m base units with a minimal set of $n - m$ dimensionless parameters. We recall the governing parameters of the drop impact as V_0 in m/s, D_0 in m, ρ_l in kg/m³, μ_l in Pa s, σ in N/m, ρ_g in kg/m³, μ_g in Pa s, and λ_g in m. These $n = 8$ dimensioned variables are composed of the $m = 3$ base units N, m, and s yielding $n - m = 5$ dimensionless parameters. Denoting the unit of dimensioned parameter X with $[X]$, we can find the dimensionless expressions Π_i by solving

$$1 = [V_0]^\alpha \cdot [D_0]^\beta \cdot [\rho_l]^\gamma \cdot [\mu_l]^\delta \cdot [\sigma]^\epsilon \cdot [\rho_g]^\kappa \cdot [\mu_g]^\lambda \cdot [\lambda_g]^\phi \quad (\text{A1})$$

and

$$\Pi_i = V_0^\alpha \cdot D_0^\beta \cdot \rho_l^\gamma \cdot \mu_l^\delta \cdot \sigma^\epsilon \cdot \rho_g^\kappa \cdot \mu_g^\lambda \cdot \lambda_g^\phi. \quad (\text{A2})$$

This yields the well-known Reynolds number $\text{Re} = (\rho_l D_0 V_0)/\mu_l$, Weber number $\text{We} = (\rho_l D_0 V_0^2)/\sigma$, Stokes number $\text{St} = \mu_g/(\rho_l D_0 V_0)$, Knudsen number $\text{Kn} = \lambda_g/D_0$, and a fifth parameter $\Pi_5 = \rho_g \sigma^2 D_0 / (V_0 \mu_l^3) = (\rho_g/\rho_l)(\text{Re}^3/\text{We}^2)$. Π_5 balances inertial, viscous, and capillary effects and scales them with the density ratio between the surrounding gas and liquid. As such, presenting the results in the Π_5 -We plane offers better visualization of the spread of drop impact conditions.

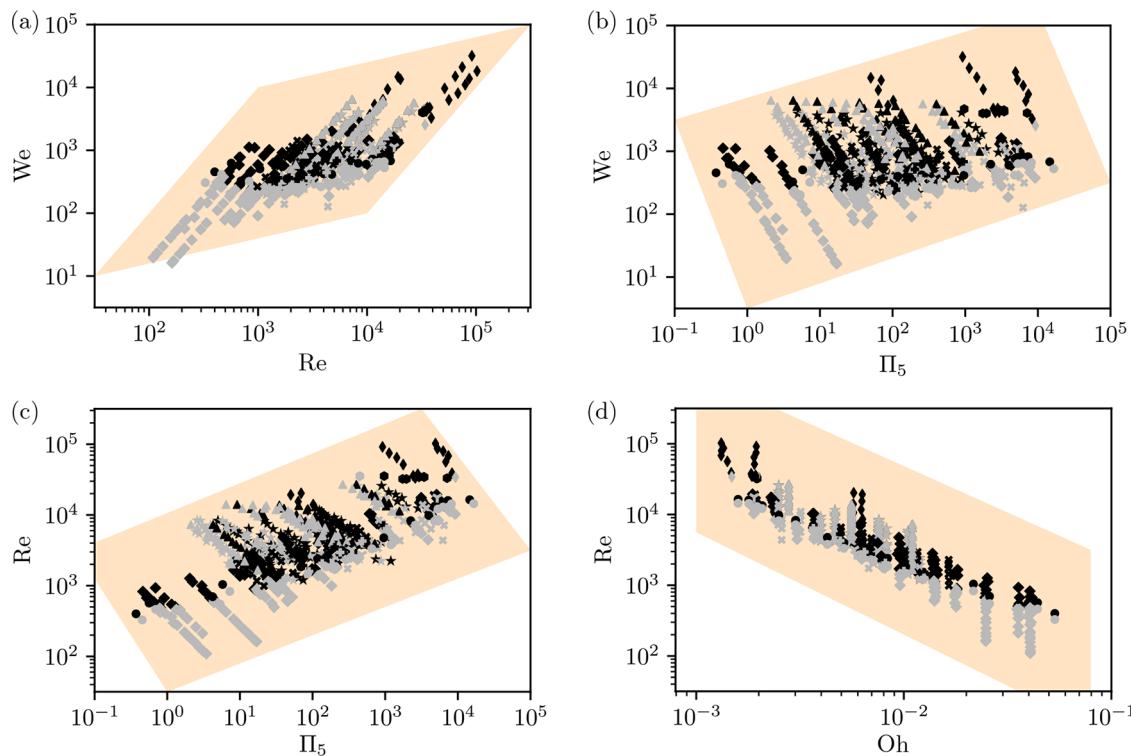


FIG. 8. Collected experimental data with symbols according to Table I plotted in four non-dimensional planes: (a) $\text{Re} - \text{We}$, (b) $\Pi_5 - \text{We}$, (c) $\Pi_5 - \text{Re}$, and (d) $\text{Oh} - \text{Re}$ planes. Gray points indicate deposition while black points splashing upon impact. The orange area corresponds to the approximate interpolation range of our DST model.

TABLE II. Bounds of the parameter space from which 100 000 random synthetic drop impact samples are drawn to analyze the UMF behavior of the RG model.

	V_0 (m/s)	R_0 (mm)	ρ_l (kg/m ³)	μ_l (mPa s)	σ (mN/m)	ρ_g (kg/m ³)	μ_g (μ Pa s)	λ_g (nm)	α (°)
Lower bound	0.1	0.7	566	≥ 0	10.3	≥ 0	11.4	≥ 0	56.4
Upper bound	30.0	2.2	1978	22.4	78.6	7.43	26.9	2784	63.6

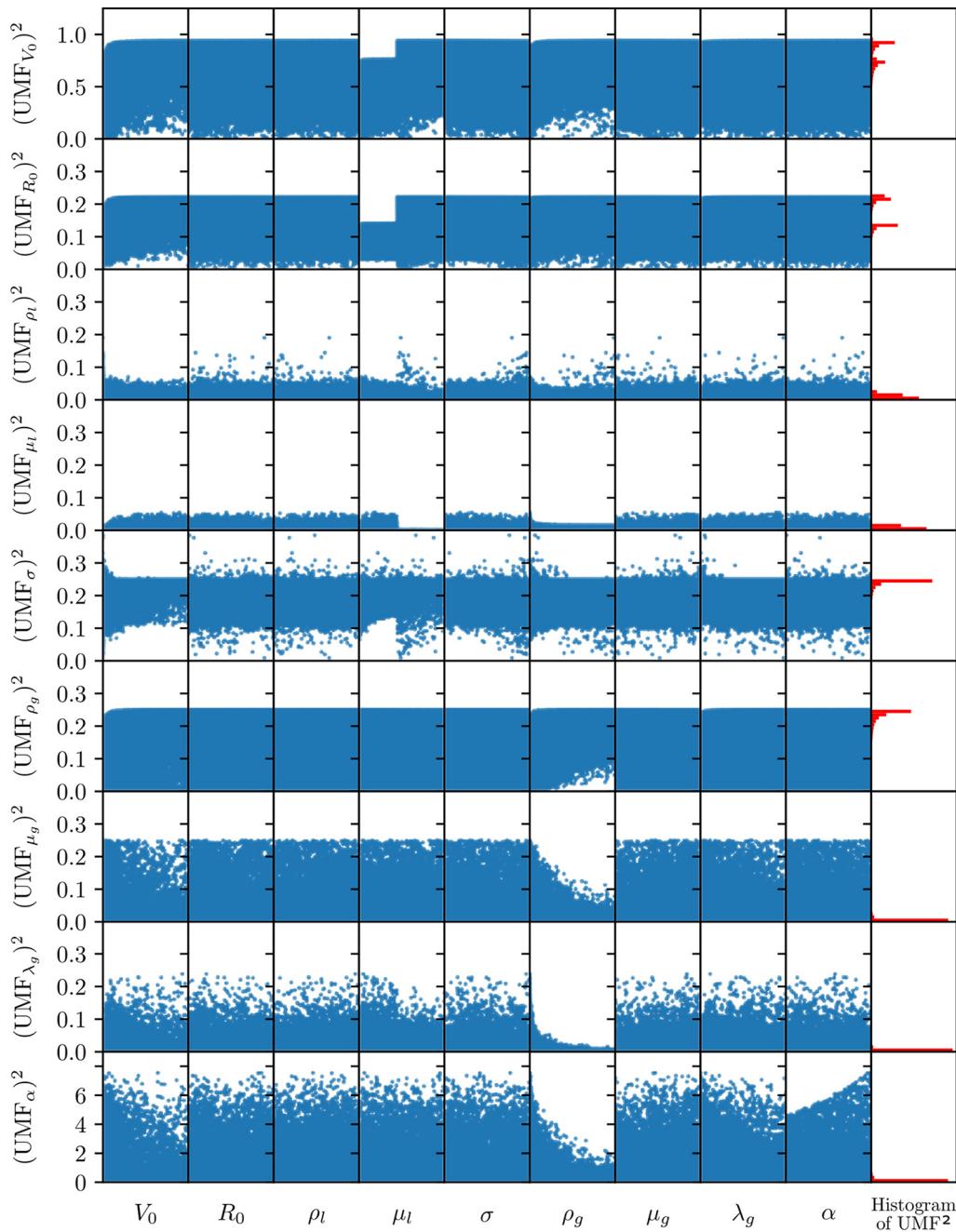


FIG. 9. Squared UMF distributions (rows) for all nine impact parameters (columns). The red histograms on the last column show frequency of respective squared UMFs across all nine parameters. Dominant peaks close to zero [e.g., $(\text{UMF}_{\mu_g})^2$ or $(\text{UMF}_{\lambda_g})^2$] indicate that corresponding measurement uncertainties can be ignored.

APPENDIX B: NON-DIMENSIONAL RANGE OF APPLICABILITY

Figure 8(a) shows the collected experimental data in the commonly used Re–We plane. The variation in gas density in the data are illustrated in the Π_5 –We plane and the Π_5 –Re plane in panels (b) and (c), respectively. Here, Re and We are based on the drop's diameter, and $\Pi_5 = \rho_g \sigma^2 D_0 / (V_0 \mu_l^3)$. The orange bounding boxes indicate the approximate interpolation range of the trained model and illustrate the wide range of impact conditions on which our threshold model [cf. Eq. (9)] is trained. For completeness, we also present the scatter distribution in the Oh–Re plane ($\text{Oh} = \sqrt{\text{We}/\text{Re}}$) in panel (d).

APPENDIX C: EVALUATION OF THE UNCERTAINTY MAGNIFICATION FACTORS

The uncertainty magnification factors (UMFs) of all nine input parameters of the RG splashing model are obtained according to Eq. (3). They are expressed as

$$\text{UMF}_{V_0} = \frac{3f_4 l + (1+f_1)[f_4(l-l_s) + f_2(l+2l_s)V_0]f_5}{2(1+f_1)(l+l_s)f_5 f_2 V_0}, \quad (\text{C1a})$$

$$\begin{aligned} \text{UMF}_{R_0} = & -\frac{l(V_0 - 3f_3)}{2(1+f_1)(l+l_s)f_5 V_0} \\ & + \frac{f_5[f_3(l+f_1l - l_s - f_1l_s) + (1+f_1)l_s V_0]}{2(1+f_1)(l+l_s)f_5 V_0}, \end{aligned} \quad (\text{C1b})$$

$$\text{UMF}_{\rho_l} = \frac{f_3[3l + (l+f_1l - l_s - f_1l_s)f_5]}{2(1+f_1)(l+l_s)f_5 V_0}, \quad (\text{C1c})$$

$$\text{UMF}_{\mu_l} = \frac{c_1}{2f_2(l+l_s)} \left[l_s - l \left(1 + \frac{3}{(1+f_1)f_5} \right) \right] \mu_l, \quad (\text{C1d})$$

$$\text{UMF}_\sigma = -\frac{1}{2} \frac{[3l + (l+f_1l - l_s - f_1l_s)f_5]\sqrt{t_e}\sigma}{2(1+f_1)(l+l_s)f_5 f_2 V_0}, \quad (\text{C1e})$$

$$\text{UMF}_{\rho_g} = \frac{l_s}{2(l+l_s)}, \quad (\text{C1f})$$

$$\text{UMF}_{\mu_g} = \frac{l}{2(l+l_s)}, \quad (\text{C1g})$$

$$\text{UMF}_{\lambda_g} = \frac{l}{2(1+f_1)(l+l_s)f_5}, \quad (\text{C1h})$$

$$\text{UMF}_z = -\frac{l \alpha \csc(\alpha) \sec(\alpha)}{l + l_s}, \quad (\text{C1i})$$

with auxiliary functions $f_i, i = 1, \dots, 5$ given as

$$f_1 = \frac{19.2\pi\lambda_g}{\sqrt{12}R_0 t_e^{3/2}}, \quad (\text{C2a})$$

$$f_2 = c_1 \mu_l + 3c_2^2 R_0 t_e^2 V_0 \rho_l, \quad (\text{C2b})$$

$$f_3 = \frac{c_1 V_0 \mu_l + \sqrt{t_e} \sigma}{f_2}, \quad (\text{C2c})$$

$$f_4 = c_1 V_0 \mu_l + 2\sqrt{t_e} \sigma, \quad (\text{C2d})$$

$$f_5 = \ln(f_1) - \ln(1+f_1), \quad (\text{C2e})$$

and with constants $c_1 = \sqrt{3}/2$ and $c_2^2 = 1.2$ provided by Riboux and Gordillo.^{1,19} l_l and l_s are lift forces due to lubrication and suction, respectively, which are defined in the RG model as

$$l_l = \frac{-3\sqrt{3}f_5 V_0 \cot(\alpha)^2 \mu_g}{\sqrt{t_e}}, \quad (\text{C3a})$$

$$l_s = \frac{3\sqrt{3}}{2\pi} K_u R_0 V_0^2 \rho_g \sqrt{t_e}, \quad (\text{C3b})$$

with $K_u = 0.3$ and ejection time t_e also given by Riboux and Gordillo.¹

These equations are numerically evaluated by uniformly sampling 100 000 random synthetic drop impacts³⁴ from a parameter space 20% larger than that of the experimental data. The bounds used for this sampling are given in Table II. Figure 9 presents the squared UMF distributions (rows) for all nine impact parameters (columns). The x-axis ticks are omitted for clarity but correspond to the sampling ranges given in Table II. The red histograms on the last column show frequency of respective squared UMFs across all nine parameters. As the figure shows, the uncertainties of ρ_b , μ_b , μ_g , and λ_g can be neglected since their average UMF² values are close to zero most of the time.

REFERENCES

- ¹G. Riboux and J. M. Gordillo, “Experiments of drops impacting a smooth solid surface: A model of the critical impact speed for drop splashing,” *Phys. Rev. Lett.* **113**, 024507 (2014).
- ²D. A. Burzynski and S. E. Bansmer, “Role of surrounding gas in the outcome of droplet splashing,” *Phys. Rev. Fluids* **4**, 073601 (2019).
- ³A. L. Yarin, “Drop impact dynamics: Splashing, spreading, receding, bouncing...,” *Annu. Rev. Fluid Mech.* **38**, 159–192 (2006).
- ⁴R. Gent, N. Dart, and J. Cansdale, “Aircraft icing,” *Philos. Trans. R. Soc. London, Ser. A* **358**, 2873–2911 (2000).
- ⁵S. Bansmer, *Aircraft Icing* (Cuvillier Verlag, 2020).
- ⁶D. A. Burzynski, *On the Impact of High-Speed Drops on Dry and Wetted Surfaces* (TU Braunschweig, 2021).
- ⁷D. A. Burzynski, I. V. Roisman, and S. E. Bansmer, “On the splashing of high-speed drops impacting a dry surface,” *J. Fluid Mech.* **892**, A2 (2020).
- ⁸M. Marengo, C. Antonini, I. V. Roisman, and C. Tropea, “Drop collisions with simple and complex surfaces,” *Curr. Opin. Colloid Interface Sci.* **16**, 292–302 (2011).
- ⁹C. Josserand and S. T. Thoroddsen, “Drop impact on a solid surface,” *Annu. Rev. Fluid Mech.* **48**, 365–391 (2016).
- ¹⁰A. L. Yarin, I. V. Roisman, and C. Tropea, *Collision Phenomena in Liquids and Solids* (Cambridge University Press, Cambridge, 2017).
- ¹¹C. H. R. Mundo, M. Sommerfeld, and C. Tropea, “Droplet-wall collisions: Experimental studies of the deformation and breakup process,” *Int. J. Multiphase Flow* **21**, 151–173 (1995).
- ¹²L. Xu, W. W. Zhang, and S. R. Nagel, “Drop splashing on a dry smooth surface,” *Phys. Rev. Lett.* **94**, 184505 (2005).
- ¹³H. J. Staat, T. Tran, B. Geerdink, G. Riboux, C. Sun, J. M. Gordillo, and D. Lohse, “Phase diagram for droplet impact on superheated surfaces,” *J. Fluid Mech.* **779**, R3 (2015).
- ¹⁴J. Hao and S. I. Green, “Splash threshold of a droplet impacting a moving substrate,” *Phys. Fluids* **29**, 012103 (2017).
- ¹⁵T. De Goede, N. Laan, K. D. Bruun, and D. Bonn, “Effect of wetting on drop splashing of Newtonian fluids and blood,” *Langmuir* **34**, 5163–5168 (2018).
- ¹⁶M. A. Quetzeri-Santiago, K. Yokoi, A. A. Castrejón-Pita, and J. R. Castrejón-Pita, “Role of the dynamic contact angle on splashing,” *Phys. Rev. Lett.* **122**, 228001 (2019).
- ¹⁷T. Ashida, M. Watanabe, K. Kobayashi, H. Fujii, and T. Sanada, “Hidden prompt splashing by corona splashing at drop impact on a smooth dry surface,” *Phys. Rev. Fluids* **5**, 011601 (2020).

- ¹⁸B. Zadrozny, "Learning and evaluating classifiers under sample selection bias," in *Proceedings of the Twenty-First International Conference on Machine Learning* (Association for Computing Machinery, New York, NY, 2004), p. 114.
- ¹⁹G. Riboux and J. M. Gordillo, "The diameters and velocities of the droplets ejected after splashing," *J. Fluid Mech.* **772**, 630–648 (2015).
- ²⁰J. M. Gordillo and G. Riboux, "A note on the aerodynamic splashing of droplets," *J. Fluid Mech.* **871**, R3 (2019).
- ²¹R. L. Vander Wal, G. M. Berger, and S. D. Mozes, "The splash/non-splash boundary upon a dry surface and thin fluid film," *Exp. Fluids* **40**, 53–59 (2006).
- ²²J. Palacios, J. Hernández, P. Gómez, C. Zanzi, and J. López, "Experimental study of splashing patterns and the splashing/deposition threshold in drop impacts onto dry smooth solid surfaces," *Exp. Therm. Fluid Sci.* **44**, 571–582 (2013).
- ²³C. S. Stevens, "Scaling of the splash threshold for low-viscosity fluids," *Europhys. Lett.* **106**, 24001 (2014).
- ²⁴J. Hao, "Effect of surface roughness on droplet splashing," *Phys. Fluids* **29**, 122105 (2017).
- ²⁵S. T. Thoroddsen, K. Takehara, and T. Etoh, "Micro-splashing by drop impacts," *J. Fluid Mech.* **706**, 560–570 (2012).
- ²⁶L. Xu, "Liquid drop splashing on smooth, rough, and textured surfaces," *Phys. Rev. E* **75**, 056316 (2007).
- ²⁷H. H. Ku, "Notes on the use of propagation of error formulas," *J. Res. Natl. Bur. Stand., Sect. C* **70C**, 263–273 (1966).
- ²⁸M. Padulo, M. S. Campobasso, and M. D. Guenov, "Comparative analysis of uncertainty propagation methods for robust engineering design," in *International Conference on Engineering Design (ICED-07)* (Cité des sciences et de l'industrie, Paris, France, 2007).
- ²⁹R. Semaan, "The uncertainty of the experimentally-measured momentum coefficient," *Exp. Fluids* **61**, 1–16 (2020).
- ³⁰H. W. Coleman and W. G. Steele, *Experimentation, Validation, and Uncertainty Analysis for Engineers* (John Wiley & Sons, 2018).
- ³¹R. Semaan and V. Yadav, "Scout: Signal correction and uncertainty quantification toolbox in MATLAB," *SoftwareX* **11**, 100474 (2020).
- ³²M. Assael, N. Dalaouti, and V. Vesovic, "Viscosity of natural-gas mixtures: Measurements and prediction," *Int. J. Thermophys.* **22**, 61–71 (2001).
- ³³B. E. Poling, J. M. Prausnitz, J. P. O'Connell *et al.*, *The Properties of Gases and Liquids* (McGraw-Hill, New York, 2001).
- ³⁴M. D. McKay, R. J. Beckman, and W. J. Conover, "A comparison of three methods for selecting values of input variables in the analysis of output from a computer code," *Technometrics* **42**, 55–61 (2000).
- ³⁵A. Kumar and D. K. Mandal, "Impact of emulsion drops on a solid surface: The effect of viscosity," *Phys. Fluids* **31**, 102106 (2019).
- ³⁶W. Steele, R. Taylor, R. Burrell, and H. Coleman, "Use of previous experience to estimate precision uncertainty of small sample experiments," *AIAA J.* **31**, 1891–1896 (1993).
- ³⁷C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.* **20**, 273–297 (1995).
- ³⁸C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.* **2**, 1–27 (2011).
- ³⁹T. Hastie, R. Tibshirani, and J. H. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer Series in Statistics, 2nd ed. (Springer, 2009).
- ⁴⁰C. M. Bishop, *Pattern Recognition and Machine Learning* (Springer, 2006).
- ⁴¹I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, "Gene selection for cancer classification using support vector machines," *Mach. Learn.* **46**, 389–422 (2002).
- ⁴²E. Villermoux, "Fragmentation versus cohesion," *J. Fluid Mech.* **898**, P1 (2020).
- ⁴³M. M. Driscoll, C. S. Stevens, and S. R. Nagel, "Thin film formation during splashing of viscous liquids," *Phys. Rev. E* **82**, 036302 (2010).
- ⁴⁴M. Pierzyna *et al.* (2021). "Data-driven splashing model," GitHub. https://github.com/FlowModelingControl/data_driven_splashing_model.