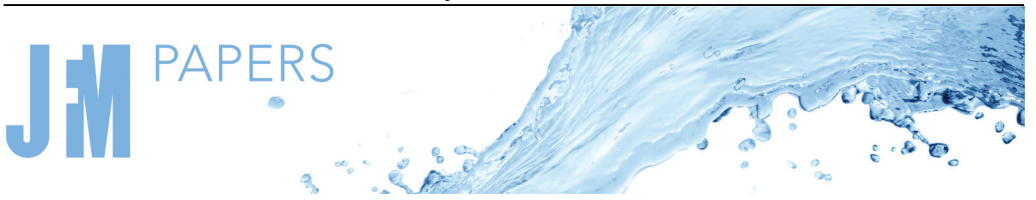
J. Fluid Mech. (2024), *vol*. 985, A7, doi:10.1017/jfm.2024.270



Characterization of partial wetting by CMAS droplets using multiphase many-body dissipative particle dynamics and data-driven discovery based on PINNs

Elham Kiyani[1,2], Mahdi Kooshkbaghi[3], Khemraj Shukla[4], Rahul Babu Koneru[5], Zhen Li[6], Luis Bravo[7], Anindya Ghoshal[7], George Em Karniadakis[4] and Mikko Karttunen[2,8,9,†]

[1]Department of Mathematics, The University of Western Ontario, 1151 Richmond Street, London, ON, Canada N6A 5B7

[2]The Centre for Advanced Materials and Biomaterials (CAMBR), The University of Western Ontario, 1151 Richmond Street, London, ON, Canada N6A 5B7

[3]Simons Center for Quantitative Biology, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA

[4]Division of Applied Mathematics, Brown University, 182 George Street, Providence, RI 02912, USA

[5]Department of Aerospace Engineering, University of Maryland, College Park, MD 20742, USA

[6]Department of Mechanical Engineering, Clemson University, Clemson, SC 29634, USA

[7]DEVCOM Army Research Laboratory, Aberdeen Proving Ground, MD 21005, USA

[8]Department of Physics and Astronomy, The University of Western Ontario, 1151 Richmond Street, London, ON, Canada N6A 3K7

[9]Department of Chemistry, The University of Western Ontario, 1151 Richmond Street, London, ON, Canada N6A 5B7

(Received 18 July 2023; revised 4 March 2024; accepted 5 March 2024)

The molten sand that is a mixture of calcia, magnesia, alumina and silicate, known as CMAS, is characterized by its high viscosity, density and surface tension. The unique properties of CMAS make it a challenging material to deal with in high-temperature applications, requiring innovative solutions and materials to prevent its buildup and damage to critical equipment. Here, we use multiphase many-body dissipative particle dynamics simulations to study the wetting dynamics of highly viscous molten CMAS droplets. The simulations are performed in three dimensions, with varying initial droplet sizes and equilibrium contact angles. We propose a parametric ordinary differential equation (ODE) that captures the spreading radius behaviour of the CMAS droplets.

† Email address for correspondence: mkarttu@uwo.ca

The ODE parameters are then identified based on the physics-informed neural network (PINN) framework. Subsequently, the closed-form dependency of parameter values found by the PINN on the initial radii and contact angles are given using symbolic regression. Finally, we employ Bayesian PINNs (B-PINNs) to assess and quantify the uncertainty associated with the discovered parameters. In brief, this study provides insight into spreading dynamics of CMAS droplets by fusing simple parametric ODE modelling and state-of-the-art machine-learning techniques.

Key words: drops, computational methods, machine learning

1. Introduction

Recent advancements in machine learning have opened the way for extracting governing equations directly from experimental (or other) data (Brunton, Proctor & Kutz 2016; Ren & Duan 2020; Delahunt & Kutz 2022). One particularly exciting use of machine learning is the extraction of partial differential equations (PDEs) that describe the evolution and emergence of patterns or features (Lee *et al.* 2020; Thiem *et al.* 2020; Meidani & Farimani 2021; Kiyani *et al.* 2022).

Spreading of liquids on solid surfaces is a classic problem (de Gennes 1985; Bonn *et al.* 2009). Although the theoretical foundations were laid by Young and Laplace already in the early 1800s (de Laplace 1805; Young 1805), there are still many open questions, and it remains a highly active research field especially in the context of microfluidics (Nishimoto & Bhushan 2013) as well as in the design of propulsion materials (Jain *et al.* 2021). As discussed in detail in the review of Popescu *et al.* (2012), there are two fundamentally different cases: non-equilibrium spreading of the droplet, and the case of thermodynamic equilibrium when spreading has ceased and the system has reached its equilibrium state. In thermodynamic equilibrium, the Laplace equation relates the respective surface tensions of the three interfaces via (de Gennes 1985; Popescu *et al.* 2012)

$$\cos\theta_{eq} = \frac{\gamma_{SG} - \gamma_{SL}}{\gamma_{LG}}, \tag{1.1}$$

where θ_{eq} is the equilibrium contact angle, and γ_{SG}, γ_{SL} and γ_{LG} are the surface tensions between solid–gas, solid–liquid and liquid–gas phases, respectively (see figure 1). Two limiting situations can be identified, namely, partial wetting and complete wetting. In the latter, the whole surface becomes covered by the fluid and $\theta_{eq} = 0^\circ$, that is, $\gamma_{SG} - \gamma_{LG} - \gamma_{SL} = 0$. When the equilibrium situation corresponds to partial wetting, $\theta_{eq} \neq 0^\circ$, it is possible to identify the cases of high-wetting ($0^\circ < \theta_{eq} < 90^\circ$), low-wetting ($90^\circ \leq \theta_{eq} < 180^\circ$) and non-wetting ($\theta_{eq} = 180^\circ$).

When a droplet spreads, it is out of equilibrium, and properties such as viscosity and the associated processes need to be addressed (de Gennes 1985; Bonn *et al.* 2009; Popescu *et al.* 2012). In experiments, the most common choice is to use high-viscosity liquids in order to eliminate inertial effects. An early classic experiment by Dussan (1979) gave a beautiful demonstration of some of the phenomena. She added tiny drops of marker dye on the surface of a spreading liquid, and observed a caterpillar-type rolling motion of the marker on the surface, giving rise to dissipation via viscous friction. Effects of viscosity and dissipation remain to be fully understood, and they have a major role in wetting phenomena (Cormier *et al.* 2012; McGraw *et al.* 2016; Edwards *et al.* 2020).

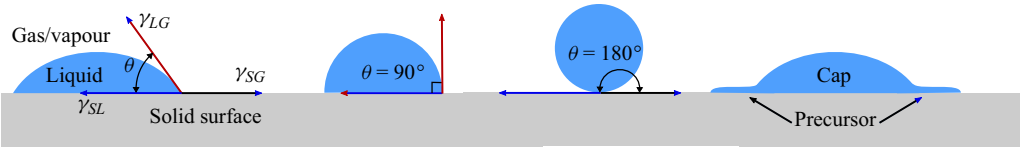


Figure 1. A schematic showing the equilibrium contact angle (i.e. $\theta \equiv \theta_{eq}$), the surface tensions (γ), the threshold between low- and high-wetting regimes ($\theta_{eq} = 90°$), and a situation of a non-wetting droplet ($\theta_{eq} = 180°$). The last image demonstrates the occurrence of a precursor that is observed in some cases. In that case, the (macroscopic) contact angle is defined using the macroscopic part of the droplet, the cap in the rightmost image. The height of the precursor is in the molecular length scales (Hardy 1919; Nieminen *et al.* 1992; Popescu *et al.* 2012).

A theoretical framework for comprehending and describing the behaviour of droplets during impact and spreading has been proposed by Gordillo, Riboux & Quintero (2019). They put forth a theoretical framework to elucidate the dynamics of spreading of Newtonian fluids. In this model, the spreading droplet is conceptualized as having a thick rim followed by a thin liquid film connected to the bulk of the drop. By employing principles of mass and momentum conservation along the rim, the liquid film and the droplet, a set of interconnected differential equations was formulated. The solution to these equations, considering appropriate boundary conditions, offers a thorough depiction of how the diameter of a Newtonian drops changes over time when impacting a smooth surface. Gorin *et al.* (2022) introduced a universal functional form to develop a model equation for estimating the time evolution of radial position, enabling predictions of transient spreading behaviour for both Newtonian and non-Newtonian fluids.

Calcium-magnesium-aluminosilicate (CMAS) is a molten mixture of several oxides, including calcia (CaO), magnesia (MgO), alumina (Al_2O_3) and silicate (SiO_2). It has a high melting point, typically around 1240 °C (Poerschke & Levi 2015) (although it can be significantly higher; see e.g. Wiesner, Vempati & Bansal (2016) and references therein), which allows it to exist in the molten state even at high temperatures encountered in modern aviation gas turbine engines (Clarke, Oechsner & Padture 2012; Ndamka, Wellman & Nicholls 2016). With high viscosity, high density and high surface tension, CMAS tends to form non-volatile droplets with $\theta_{eq} \neq 0$ rather than completely wetting the surfaces (Grant *et al.* 2007; Vidal-Setif *et al.* 2012; Nieto *et al.* 2021). When it solidifies, it forms a glass-like material that can adhere to surfaces and resist erosion. The buildup of CMAS on turbine engine components can lead to clogging of the cooling passages and degradation of the protective coatings, resulting in engine performance issues and even damage or failure (Clarke *et al.* 2012; Ndamka *et al.* 2016; Song *et al.* 2016; Wiesner *et al.* 2016).

We simplified the CMAS wetting process by studying a fully molten CMAS droplet at a constant temperature, omitting considerations of surface chemical reactions, viscosity variations and phase changes. Existing analytical models that predict the spreading dynamics of highly viscous drops often rely on assumptions such as Stokes flow and completely wetting surfaces with zero contact angles. Our contribution lies in proposing a simple and general model for predicting the spreading dynamics for a broad range of systems, including partially wetting systems with non-zero contact angles.

The spreading of a droplet over a solid surface is commonly characterized using a power law $r \sim t^{\alpha}$, which expresses the radius of the wetted area as a function of time. The relationship is called Tanner's law for macroscopic completely wetting liquids at late times, with $\alpha = 1/10$ (Tanner 1979; Bonn *et al.* 2009). Power laws have also been

demonstrated at microscopic scales (Nieminen *et al.* 1992). However, several conditions, such as surface properties, droplet shape and partial wetting, result in deviations from Tanner's law (McHale *et al.* 2004; Cormier *et al.* 2012; Winkels *et al.* 2012).

A common method, as the above suggests, for analysing the spreading dynamics is to investigate the existence of the power-law behaviour. To determine the presence of such power-law regimes in data, one can simply employ

$$\alpha(t) = \frac{\mathrm{d}\ln(r)}{\mathrm{d}\ln(t)}. \tag{1.2}$$

While this has worked remarkably well for complete wetting by viscous fluids, the situation for partial wetting is different (Winkels *et al.* 2012). In our study, we investigate the spreading behaviour of CMAS droplets using multiphase many-body dissipative particle dynamics (mDPD) simulations. We generalize (1.2) such that it includes dependence on the initial droplet radius R_0 and θ_{eq} in order to describe partial wetting, that is, $\alpha \equiv \alpha(t, R_0, \theta_{eq})$.

Our objective is to gain a comprehensive understanding of the behaviour of CMAS droplet spreading dynamics by integrating knowledge about the fundamental physics of the system into the neural network architecture. To achieve this, we employ the framework of physics-informed neural networks (PINNs) (Raissi, Perdikaris & Karniadakis 2019), an emerging machine-learning technique that incorporates the physics of a system into deep learning. The PINNs address the challenge of accurate predictions in complex systems with varying initial and boundary conditions. By directly incorporating physics-based constraints into the loss function, PINNs enable the network to learn and satisfy the governing equations of the system.

The ability of PINNs to discover equations makes them promising for applications in scientific discovery, engineering design and data-driven modelling of complex physical systems (Karniadakis *et al.* 2021). Their integration of physics-based constraints into the learning process enhances their capacity to generalize and capture the underlying physics accurately. Here, we also employ symbolic regression to generate a mathematical expression for each unknown parameter. Furthermore, we employ Bayesian physics-informed neural networks (B-PINNs) (Yang, Meng & Karniadakis 2021) to quantify the uncertainty of the predictions.

The rest of this paper is structured as follows. In § 2, we provide an overview of mDPD simulation parameters and system set-up. Simulation outcomes and the data preparation process are presented in § 3. Section 4 gives a brief introduction to the PINNs architecture, followed by a presentation of the results of PINNs and parameter discovery. The symbolic regression results and the mathematical formulas for the parameters are presented in § 5. Section 6 covers the discussion on B-PINNs as well as the quantification of uncertainty in predicting the parameters. Finally, we conclude with a summary of our work in § 7.

2. Multiphase many-body dissipative particle dynamics simulations

Three-dimensional simulations were performed using the mDPD method (Li *et al.* 2013; Xia *et al.* 2017; Rao *et al.* 2021), which is an extension of the traditional dissipative particle dynamics (DPD) model (Español & Warren 1995; Groot 2004); DPD is a mesoscale simulation technique for studies of complex fluids, particularly multiphase systems, such as emulsions, suspensions and polymer blends (Ghoufi & Malfreyt 2012; Lei *et al.* 2018; Zhao *et al.* 2021). The relation between DPD and other coarse-grained methods and

atomistic simulations has been studied and discussed by Murtola *et al.* (2009), Li *et al.* (2016), Chan, Li & Wenzel (2023) and Español & Warren (2017).

In DPD and mDPD models, the position ($\boldsymbol{r}_i$) and velocity ($\boldsymbol{v}_i$) of a particle i with mass m_i are governed by Newton's equations of motion in the form

$$\left.\begin{aligned} \frac{\mathrm{d}\boldsymbol{r}_i}{\mathrm{d}t} &= \boldsymbol{v}_i, \\ m_i \frac{\mathrm{d}\boldsymbol{v}_i}{\mathrm{d}t} &= \boldsymbol{F}_i = \sum_{j \neq i} \boldsymbol{F}_{ij}^C + \boldsymbol{F}_{ij}^D + \boldsymbol{F}_{ij}^R. \end{aligned}\right\} \tag{2.1}$$

The total force on particle i, i.e. $\boldsymbol{F}_i$, consists of three pairwise components, i.e. the conservative $\boldsymbol{F}^C$, dissipative $\boldsymbol{F}^D$, and random $\boldsymbol{F}^R$ forces. The latter two are identical in DPD and mDPD models, given by

$$\boldsymbol{F}_{ij}^D = -\gamma\, \omega_D(r_{ij})\, (\boldsymbol{v}_{ij} \cdot \boldsymbol{e}_{ij})\, \boldsymbol{e}_{ij}, \tag{2.2}$$

$$\boldsymbol{F}_{ij}^R = \zeta\, \omega_R(r_{ij})\, (\mathrm{d}t)^{-1/2} \xi_{ij}\, \boldsymbol{e}_{ij}, \tag{2.3}$$

where $\boldsymbol{e}_{ij}$ is a unit vector, ω_D and ω_R are weight functions for the dissipative and random forces, and ξ_{ij} is a pairwise conserved Gaussian random variable with zero mean and second moment $\langle \xi_{ij}(t)\, \xi_{kl}(t') \rangle = (\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})\, \delta(t - t')$, where δ_{ij} is the Kronecker delta, and $\delta(t - t')$ is the Dirac delta function. Together, the dissipative and random forces constitute a momentum-conserving Langevin-type thermostat. The weight functions and the constants γ and ζ are related via fluctuation–dissipation relations first derived by Español & Warren (1995),

$$\omega_D = (\omega_R)^2, \tag{2.4}$$

$$\zeta = \sqrt{2\gamma k_B T}, \tag{2.5}$$

in which k_B is the Boltzmann constant, and T is the temperature. This relation guarantees the canonical distribution (Español & Warren 1995) for fluid systems in thermal equilibrium. The functional form of the weight function is not specified, but the most common choice (also used here) is

$$\omega_D(r_{ij}) = \begin{cases} (1 - r_{ij}/r_d)^s & \text{for } r_{ij} \leq r_d, \\ 0 & \text{for } r_{ij} > r_d. \end{cases} \tag{2.6}$$

Here, $s = 1.0$ is used, and r_d defines a cutoff distance for the dissipative and random forces.

Although the above equations are the same for both DPD and mDPD, they differ in their conservative forces. Here, we use the form introduced by Warren (2001, 2003):

$$F_{ij}^C = A\, \omega_C(r_{ij})\, \boldsymbol{e}_{ij} + B(\rho_i + \rho_j)\omega_B\, \boldsymbol{e}_{ij}. \tag{2.7}$$

The functional form of both weight functions ω_C and ω_B is the same as ω_D in (2.6), but with different cutoff distances r_c and r_b. The first term in (2.7) is the standard expression for the conservative force in DPD, and the second is the multi-body term. The constants A and B are chosen such that $A < 0$ for attractive interactions, and $B > 0$ for repulsive

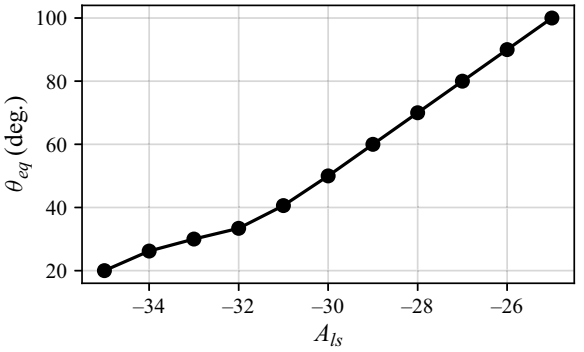


Figure 2. The equilibrium contact angles θ_{eq} for the different attraction parameters between the liquid and solid particles (A_{ls}; see (2.7)). It is worth noting that the data for this figure have been extracted from Koneru *et al.* (2022).

interactions; note that in conventional DPD, $A > 0$ and $B = 0$. The key component is the weighted local density

$$\rho_i = \sum_{j \neq i} \omega_\rho(r_{ij}). \tag{2.8}$$

There are several ways to choose the weight function (Zhao *et al.* 2021), and here, the normalized Lucy kernel (Lucy 1977) in three dimensions is used:

$$\omega_\rho(r_{ij}) = \frac{105}{16\pi r_{c\rho}^3}\left(1 + \frac{3r_{ij}}{r_{c\rho}}\right)\left(1 - \frac{r_{ij}}{r_{c\rho}}\right)^3, \tag{2.9}$$

with a cutoff distance $r_{c\rho}$ beyond which the weight function ω_ρ becomes zero.

2.1. *Simulation parameters and system set-up*

To simulate molten CMAS, the parameter mapping of Koneru *et al.* (2022) was used together with the open-source code LAMMPS (Thompson *et al.* 2022). In brief, the properties of molten CMAS at approximately 1260 °C, based on the experimental data from Naraparaju *et al.* (2019), Bansal & Choi (2014) and Wiesner *et al.* (2016), were used. In physical units, density was 2690 kg m^{-3}, surface tension 0.46 N m^{-1}, and viscosity 3.6 Pa s. Using the density and surface tension to estimate the capillary length ($\kappa = (\sigma/(\rho g))^{1/2}$) gives 4.18 mm. The droplets in the simulations (details below) had linear sizes shorter than the capillary length, hence gravity was omitted. In terms of physical units, time was 6.297×10^{-6} s, length was 17.017×10^{-6} m, and mass was 1.964×10^{-8} kg.

Using the above values, droplets of initial radii $d = 8, 9, 10, 11, 12$ in mDPD units, corresponding to $R_0 = 0.136, 0.153, 0.17, 0.187, 0.204$ mm, respectively, were used in the simulations; all of them are smaller than the capillary length. The time step was 0.002 (mDPD units) corresponding to 12.59 ns. In addition, $k_BT = 1$, $r_c = 1.0$, $r_b = r_{c\rho} = 0.75$, $r_d = 1.45$, $\gamma = 20$ and $B = 25$. The attraction parameter A in (2.7) has to be set for the interactions between the liquid particles (A_{ll}), and the liquid and solid particles (A_{ls}). The former was set to $A_{ll} = -40$, and A_{ls} was chosen based on simulations that provided the desired θ_{eq}, thus allowing for controlled variation of θ_{eq} (see figure 2).

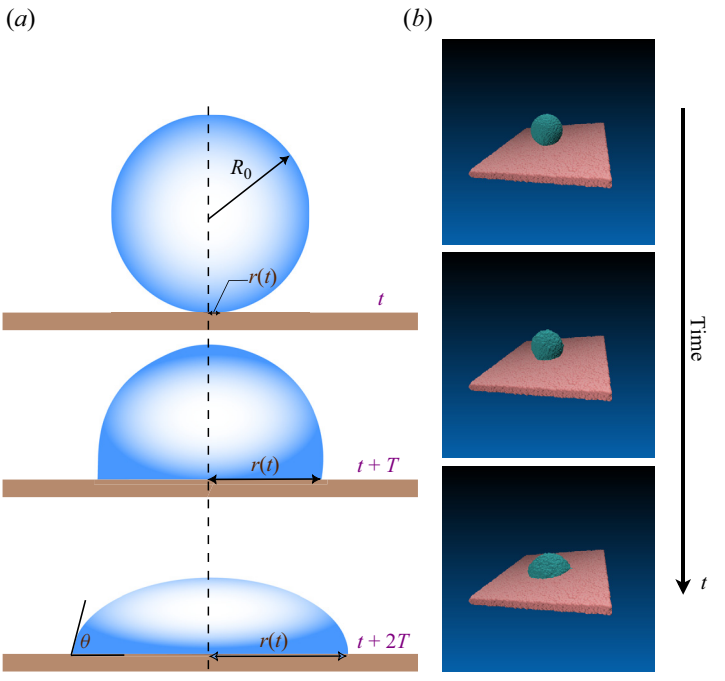


Figure 3. (*a*) Illustration of the spreading behaviour of a CMAS droplet on a hydrophilic surface at different times. The droplet with initial size R_0 spreads on the surface with radius $r(t)$ and contact angle $\theta(t)$. (*b*) A series of snapshots from a simulation of a droplet with initial size $R_0 = 0.136$ mm and equilibrium contact angle $\theta_{eq} = 93.4°$.

The initial configuration of the droplet and the solid wall was generated from a random distribution of equilibrated particles with number density $\rho = 6.74$. This amounts to approximately 60 660 particles in the wall, and depending on the initial radius, anywhere between 14 456 and 48 786 particles in the droplet. Periodic boundary conditions are imposed along the lateral directions, and a fixed, non-periodic boundary condition is imposed along the wall-normal direction. Since mDPD is a particle-based method, the spreading radius and the dynamic contact angle are approximated using surface-fitting techniques. First, the outermost surface of the droplet is identified based on the local number density, i.e. particles with $\rho \in [0.45, 0.6]$. The liquid particles closest to the wall are fitted to a circle of radius r, i.e. the spreading radius. On the other hand, a sphere with the centroid of the droplet as the centre is fitted to the surface particles to compute the contact angle. The contact angle is defined as the angle between the tangent at the triple point (liquid–solid–gas interface) and the horizontal wall. The wall in these simulations is made up of randomly distributed particles to eliminate density and temperature fluctuations at the surface. Following Li *et al.* (2018), the root mean square height (R_q) of the surface scales linearly with $1/\sqrt{N_w}$, where $N_w = (2\pi r_{cw}^3/3)\rho_w$ is the number of neighbouring particles. In this work, R_q comes out to be around 0.0708 mDPD units, or 1.2 μm.

As the CMAS droplet spreads on the substrate, it loses its initial spherical shape and begins to wet the surface as depicted in figure 3, forming a liquid film between the droplet and the substrate. Understanding how droplets behave on surfaces is important for a wide range of applications, including in industrial processes, microfluidics, propulsion

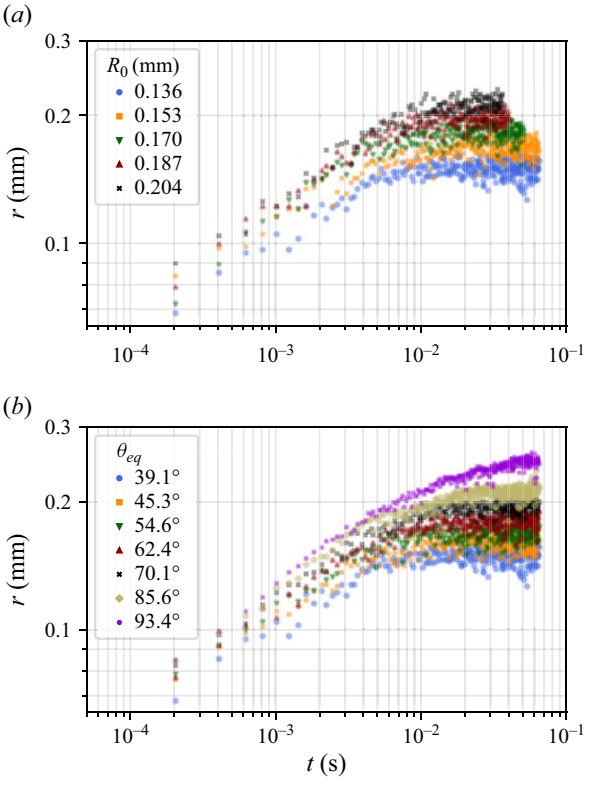


Figure 4. The impact of θ_{eq} and R_0 on the droplet radii as a function of time for various (*a*) initial drop sizes with equilibrium contact angle $\theta_{eq} = 54.6^\circ$ corresponding to $A_{ls} = 30.0$, and (*b*) equilibrium contact angles (corresponding to $A_{ls} = -25.0, -25.8, -27.0, -28.0, -29.0, -30.0, -31.4, -32.2$) and initial drop size $R_0 = 0.136$ mm.

materials, and the design of self-cleaning surfaces (Pitois & François 1999; Chen *et al.* 2016; Hassan *et al.* 2019; Jain *et al.* 2021; Nieto *et al.* 2021).

3. Simulation results

The size of a droplet changes over time. By tracking the changes, we can gain insight into the physical processes involved in spreading. The time evolution of the droplet radius ($r(t)$) is shown in figure 4. The log-log plots show the effect of the initial drop size R_0 and equilibrium contact angles θ_{eq} on the radius $r(t)$. Figure 4(*a*) displays $r(t)$ for initial drop sizes $R_0 = 0.136, 0.153, 0.17, 0.183, 0.204$ mm and equilibrium contact angle $\theta_{eq} = 54.6^\circ$. Similarly, figure 4(*b*) shows the spreading radius for different equilibrium contact angles ($\theta_{eq} = 93.4^\circ, 85.6^\circ, 77.9^\circ, 70.1^\circ, 62.4^\circ, 54.6^\circ, 45.3^\circ$ and 39.1°) with initial drop size $R_0 = 0.136$ mm.

Eddi, Winkels & Snoeijer (2013) used high-speed imaging with time resolution covering six decades to study the spreading of water–glycerine mixtures on glass surfaces. By varying the amount of glycerine, they were able to vary the viscosity over the range 0.0115–1.120 Pa s. They observed two regimes, the first one for early times with α changing continuously as a function of time from $\alpha \approx 0.8$ to $\alpha \approx 0.5$. This was followed by a sudden change to the second regime in which α settled to $0.1 < \alpha < 0.2$. As pointed out by Eddi

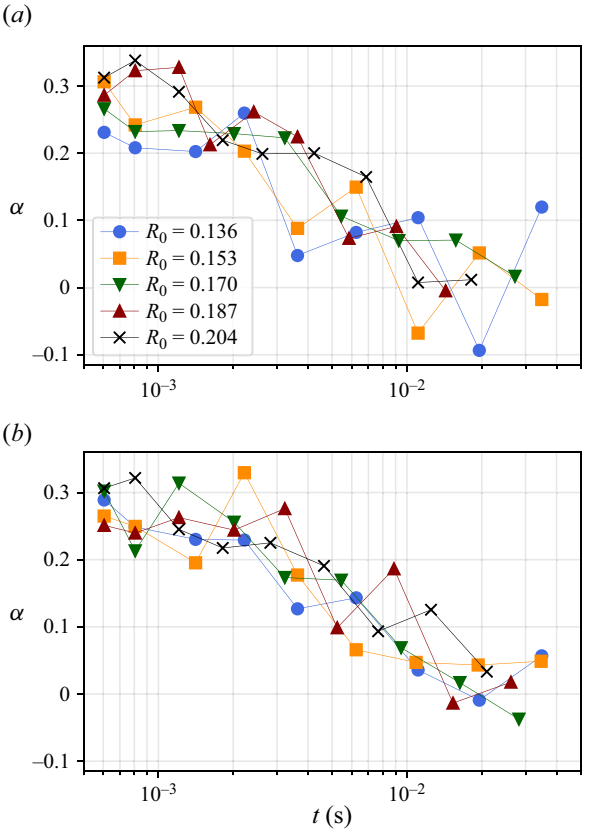


Figure 5. The value of α, calculated using (1.2), varies for different initial radii and fixed equilibrium contact angles (*a*) $\theta_{eq} = 85.6^\circ$ and (*b*) $\theta_{eq} = 62.4^\circ$. The figure illustrates that α is influenced by both the initial drop size R_0 and the equilibrium contact angle θ_{eq}.

et al. (2013), the second regime agrees with Tanner's law (Tanner 1979). All of their systems displayed complete wetting.

Based on $r(t)$ of the CMAS drops and α shown in figure 4, it can be observed that $r(t)$ (and based on the power law, α) depends on both the initial drop size and the equilibrium contact angle. The values of α for some simulation datasets are plotted over time in figure 5. The plot shows the behaviour of α for different initial drop sizes and equilibrium contact angles $\theta_{eq} = 62.4^\circ$ and 85.6°.

Inspired by the experimental results of Eddi *et al.* (2013), the simulations of Koneru *et al.* (2022) and the current simulations, we propose a simple sigmoid type dependence for α:

$$\frac{\mathrm{d}\ln(r)}{\mathrm{d}\ln(t)} = \alpha(t, R_0, \theta_{eq}) := \eta\left[\frac{1}{1+\exp(\beta(\tau - \ln(t)))} - 1\right]. \tag{3.1}$$

The two constant values of α discussed in the above references are the two extrema of the sigmoid curve, given that the transition between the two regimes occurs at $\ln(t_{transition}) = \tau$.

The parameters of (3.1) are discovered by PINNs, and their dependence on R_0 and θ_{eq} is then expressed using symbolic regression. The general steps in the discovery of the droplet

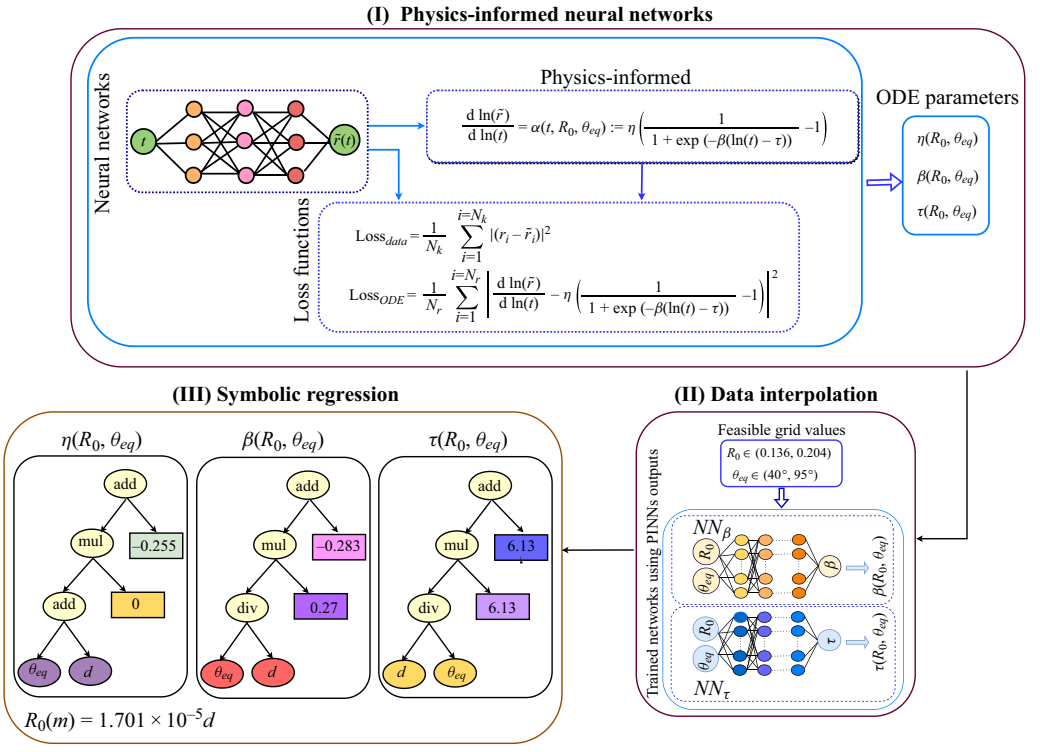


Figure 6. The process of utilizing PINNs to extract three unknown parameters of the ODE (3.1), using three-dimensional mDPD simulation data. First, a neural network is trained using simulation data, where the input is time t and the output is spreading radii $\tilde{r}(t)$. This neural network comprises four layers with three neurons, and is trained for 12 000 epochs. Subsequently, the predicted $\tilde{r}(t)$ is used to satisfy (3.1) in the physics-informed part. The loss function for this process consists of two parts: data matching and residual. By optimizing the loss function, the values of $\eta(R_0, \theta_{eq})$, $\beta(R_0, \theta_{eq})$ and $\tau(R_0, \theta_{eq})$ are determined for each set of R_0 and θ_{eq}. After predicting the unknown parameters using PINNs, two additional neural networks, denoted as NN_β and NN_τ, are trained using these parameters to generate values for the unknown parameters at points where data are not available. The outputs of these networks, together with the outputs of the PINNs, are then fed through a symbolic regression model to discover a mathematical expression for the discovered parameter.

spreading equation and the extraction of the unknown parameters $\eta(R_0, \theta_{eq})$, $\beta(R_0, \theta_{eq})$ and $\tau(R_0, \theta_{eq})$ are shown in figure 6 and can be summarized as follows.

(i) Data collection: for this study, data are collected by conducting three-dimensional simulations using the mDPD method in LAMMPS with varying initial drop sizes R_0 and equilibrium contact angles θ_{eq}.

(ii) PINNs: the input of the network is time t, and the output of the network is the spreading radii $\tilde{r}(t)$. The physics-informed part of the PINNs is encapsulated in designing the loss function. In this study, the 'goodness' of the fit is measured by (a) deviation from trained data together with (b) deviation of network predictions and those from ordinary differential equation (ODE) (3.1) solutions. This optimization process reveals the the unknown parameters $\eta(R_0, \theta_{eq})$, $\beta(R_0, \theta_{eq})$ and $\tau(R_0, \theta_{eq})$.

(iii) Data interpolation: after PINNs are trained, we used their predictions together with two additional multilayer perceptron neural networks to fill the sparse parameter space. This step helps our next goal, which is relating the ODE parameters to R_0 and θ_{eq} without performing three-dimensional simulations.

(iv) Symbolic regression: discovering a mathematical expression for each unknown parameter $\eta(R_0, \theta_{eq})$, $\beta(R_0, \theta_{eq})$ and $\tau(R_0, \theta_{eq})$ of (3.1).

(v) In order to quantify the uncertainty associated with our predictions, we utilize B-PINNs and leverage the insights gained from the PINNs prediction and symbolic regression, with a specific emphasis on the known value of η. By employing B-PINNs, we can effectively ascertain the values of two specific parameters, β and τ, which in turn enable us to quantify the uncertainty in our predictions.

4. Physics-informed neural networks

In this section, we delve into the process of uncovering the parameters of the proposed ODE using the available simulation data.

In recent years, the concept of approximating the behaviours of complex nonlinear dynamical systems with experimental or numerically simulated data has gained significant traction, driven by the exponential growth in available data. The conventional approach for model fitting involves nonlinear least squares, which relies on iterative optimization algorithms (Srinivasan & Mason 1986). These algorithms can be computationally expensive and sensitive to initial conditions, sparsity, and high dimensionality of the data. To circumvent these issues, Brunton *et al.* (2016) proposed a framework named SINDy (sparse identification of nonlinear dynamical system), which selects a class of algebraic functions along with an arbitrary order of derivatives. By performing sparse regression, which assumes that the contributions of many of these functions are negligible, SINDy can accurately discover ODEs and PDEs. On the other hand, equation learner (EQL; Martius & Lampert 2016) leverages the ability to use algebraic functions as neural network units, which brings interpretability to the model fitting process. Both SINDy and EQL share a reliance on predefined basis functions, constraining their ability to represent complex functional forms beyond those that are predefined. Therefore, applying these methods in the present study is not feasible. These coefficients not only are the coefficients of the basis functions but also cannot be utilized to define a library of functions, primarily due to the structure of the proposed ODE.

Physics-informed neural networks form another promising class of approaches that leverages the flexibility and scalability of deep neural networks to discover governing equations, incorporating physical laws (PDEs, ODEs, integro-differential equations (IDEs), etc.) and data (initial conditions plus boundary conditions) into the loss function (Raissi *et al.* 2019; Chen *et al.* 2020; Mao, Jagtap & Karniadakis 2020; Mishra & Molinaro 2020; Shukla *et al.* 2020; Karniadakis *et al.* 2021). In contrast, EQL and SINDy focus on deriving equations from data without enforcing the governing laws, lacking a guarantee of adherence to these principles. In addition, PINNs leverage automatic differentiation for computing the exact derivatives of differential equations, while SINDy, considered as an optimization method, relies on numerical differentiation for gradient approximations. The PINNs represent a more adaptable approach, capable of adjusting to complex systems by learning directly from data, even when a predefined mathematical model is completely or partially absent.

The method employed to find unknown parameters in this study exhibits potential for broader applications, such as different fluid properties, reduced-order PDEs, and scenarios in which only a part of the equation is absent (grey box learning framework; Kiyani *et al.* 2023). While, potentially, the parameters of the proposed ODE could be discovered using other aforementioned approaches for the purpose of this paper, the scalability and extensibility to PDEs or more complex ODEs, and insensitivity to initial conditions in the

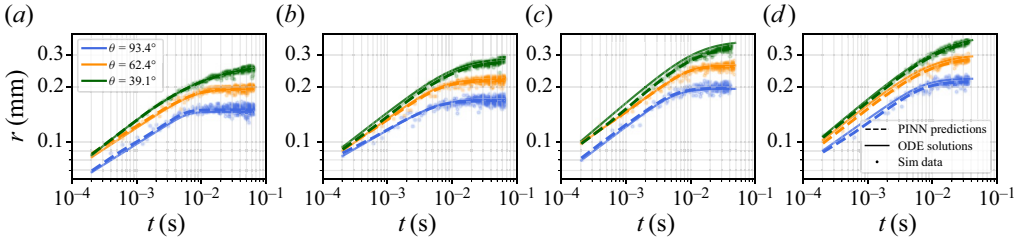


Figure 7. Comparison of the time evolution of the droplet radii: mDPD simulations (symbols), ODE model (3.1) (solid lines) and PINN predictions (dashed lines) for $\theta_{eq} = \{39.1^\circ, 62.4^\circ, 93.4^\circ\}$ and $R_0 = \{0.136, 0.153, 0.187, 0.204\}$ mm parameter sets.

fitting process compared to other methods, suggest that PINNs are the most suitable choice for this task.

4.1. *Discovering parameters of the ODE*

As discussed in § 3, our study aims to identify the values of the parameters $\eta(R_0, \theta_{eq})$, $\beta(R_0, \theta_{eq})$ and $\tau(R_0, \theta_{eq})$ in the ODE (3.1). The general steps of the framework are shown in figure 6(I). PINNs take time t as an input to predict $\tilde{r}(t)$ at each time.

For each unique combination of a contact angle and initial drop size, the data were divided randomly into training, validation and test sets at ratio 80 : 10 : 10. A neural network with four layers of three neurons each was trained on the training data using the JAX and FLAX Python libraries on the M1 Apple chip. The training process consisted of 12 000 epochs, and took approximately 5 minutes to complete for each case.

The predicted values for the time evolution of the radii $\tilde{r}(t)$ should satisfy the data and the physics-informed step, i.e. meet the requirements of the ODE (3.1). The two-component loss function, designed to meet the requirements, consists of $Loss_{data}$ and $Loss_{ODE}$, given by

$$\left.\begin{aligned} Loss_{data} &= \frac{1}{N_k}\sum_{i=1}^{i=N_k} |(r_i(t) - \tilde{r}_i(t))|^2, \\ Loss_{ODE} &= \frac{1}{N_r}\sum_{i=1}^{i=N_r} \left| \frac{\mathrm{d}\ln(\tilde{r})}{\mathrm{d}\ln(t)} - \eta\left(\frac{1}{1+\exp(-\beta(\ln(t)-\tau))} - 1\right)\right|^2, \end{aligned}\right\} \tag{4.1}$$

where $\tilde{r}(t)$ and $r(t)$ stand for the radii from the prediction and simulation, respectively. Here, N_k is the number of training points, and N_r is the number of residual points.

Figures 7, 8 and 9 illustrate the results obtained by utilizing PINNs to discover the parameters of the ODE (3.1), which describes the dynamics of the radii of the CMAS drops.

Figure 7 shows comparisons of the simulation data ($r(t)$), the prediction ($\tilde{r}(t)$) and the solution of the ODE (3.1). The figure demonstrates a remarkable degree of agreement between simulations, PINNs and our ODE model. The first three panels in figure 8 show the convergence of $\eta(R_0, \theta_{eq})$, $\beta(R_0, \theta_{eq})$, and $\tau(R_0, \theta_{eq})$ parameters during training. One can conclude that, all three parameters stabilize roughly after 12 000 epochs. In the rightmost panel of figure 8, the loss function (4.1) history is plotted against the training epochs, stabilizing around 10^{-4}. This indicates successful training of the PINNs model. The results shown in figures 7 and 8 demonstrate the capability of our proposed framework

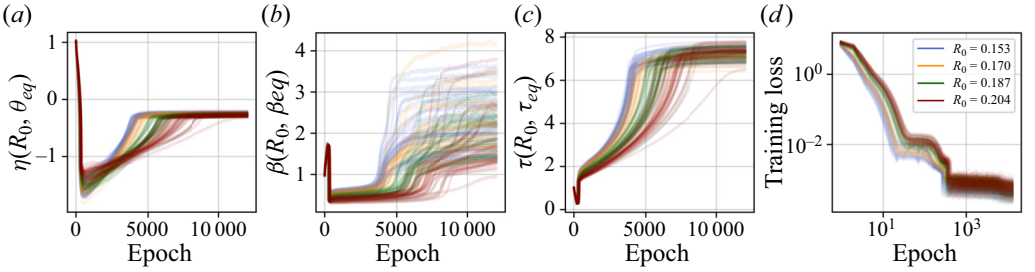


Figure 8. (*a*,*b*,*c*) The evolution of parameters $\eta(R_0, \theta_{eq})$, $\beta(R_0, \theta_{eq})$ and $\tau(R_0, \theta_{eq})$, respectively, over multiple epochs. These plots demonstrate that the parameters gradually converge to a stable state after 12 000 epochs. (*d*) The traces of the loss function for the PINNs framework. The learning curves demonstrate the decreasing trend of the loss functions, indicating that they converge to a stable point for all initial drop sizes and θ_{eq}.

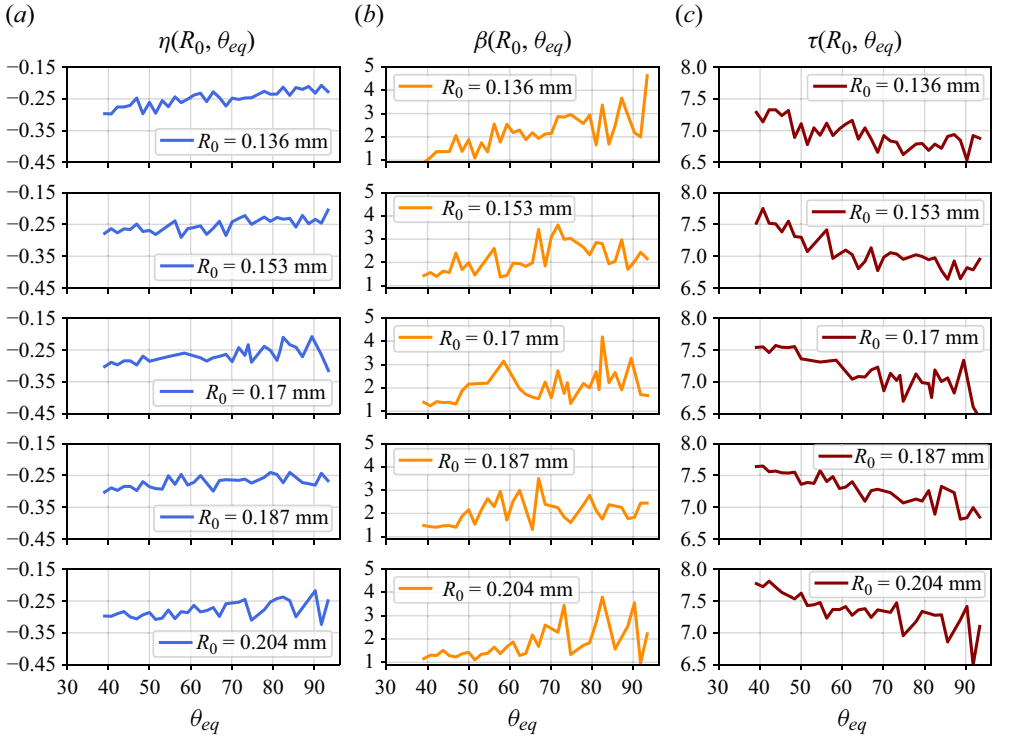


Figure 9. The values of (*a*) η, (*b*) β and (*c*) τ obtained through PINNs. These values exhibit varying behaviours depending on the initial radius R_0 and equilibrium contact angle θ_{eq}. The horizontal axes display the equilibrium contact angles θ_{eq}. The vertical axes of all plots represent the values of η, β and τ. Here, η remains nearly constant within a small range of values between -0.325 and -0.200, and β and τ change within ranges from 1.0 to 5.0, and 6.5 to 8.0, respectively.

to accurately predict the spreading radius of CMAS across the different initial radii and equilibrium contact angles.

Figures 9(*a*,*b*,*c*) show the values of $\eta(R_0, \theta_{eq})$, $\beta(R_0, \theta_{eq})$ and $\tau(R_0, \theta_{eq})$, respectively, obtained by PINNs for different initial radii R_0 ranging from 0.136 to 0.204 mm, and equilibrium contact angles θ_{eq} ranging from 39.1° to 93.4°. The results show that η changes within a small window between -0.325 and -0.200 for all R_0 and θ_{eq}.

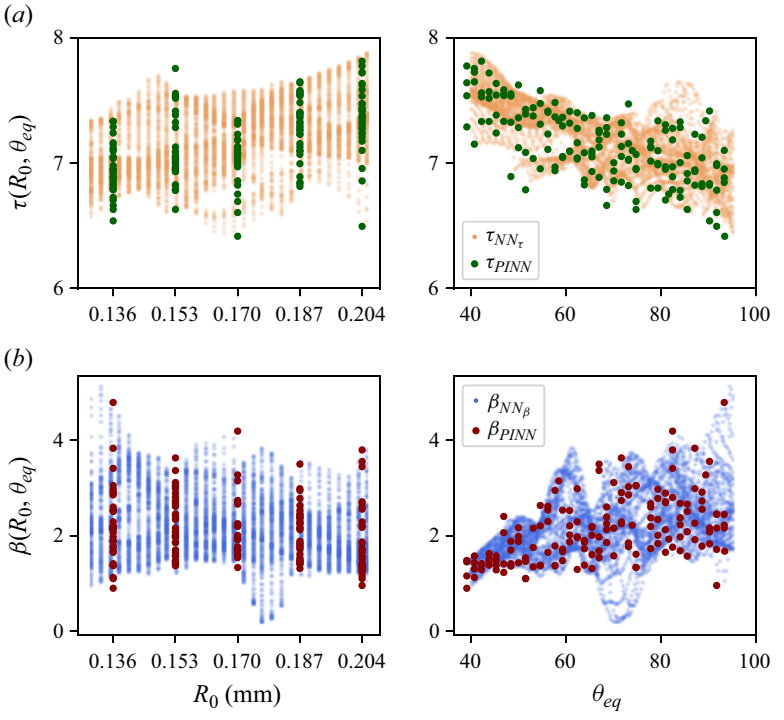


Figure 10. Predictions of the parameters (*a*) τ and (*b*) β using the trained neural networks NN_β and NN_τ. The horizontal axes show R_0 and θ_{eq}. The green and red circles correspond to the obtained values of τ and β using the PINNs that were used to train NN_β and NN_τ. Additionally, the orange and blue dots represent the predicted values for grid interpolations from $R_0 = 0.136$ mm to $R_0 = 0.204$ mm, and $\theta_{eq} = 40^\circ$ to $\theta_{eq} = 95^\circ$.

However, the changes in β (between 1 and 5) and η (between 6.0 and 8.0) are significant, indicating that these parameters depend strongly on R_0 and θ_{eq}.

4.2. *Generate more samples of feasible radii and contact angles*

As discussed earlier, the parameters in our ODE model (3.1) are functions of the initial radius and the equilibrium contact angle. Using PINNs, we were able to find the values for those parameters. To find a closed-form relation between the parameters R_0 and θ_{eq}, more data than the rather small current set are needed. Performing three-dimensional mDPD simulations is, however, computationally expensive. In this subsection, we train two additional neural networks to capture the nonlinear relation between the ODE parameters found by PINNs, and the variables R_0 and θ_{eq}. Then we will use these trained networks to fill our sparse parameter space to perform symbolic regression in the next section.

Specifically, we generate values for R_0 in the interval [0.136, 0.204] mm and θ_{eq} in the range [40°, 95°], as shown in figure 6(II).

Two fully connected networks NN_β and NN_τ consist of eight dense layers with 256/256/256/128/64/32/16/8 neurons. These networks are trained using an Adam optimizer with learning rate 10^{-2} for a total of 4000 epochs.

Discovered parameters from NN_τ and NN_β are visualized in figures 10(*a*,*b*), respectively, where the parameter values obtained from PINNs are denoted by green and red circles, indicating the training data for NN_β and NN_τ, respectively. Furthermore, the parameter values generated by NN_β and NN_τ are depicted as light orange and light blue dots. Visually,

it is evident that these dots have filled the gaps between parameters that were absent in our LAMMPS dataset.

5. Symbolic regression

In this section, we use symbolic regression to find the explicit relation between the ODE parameters, and the initial radii and equilibrium contact angles. Symbolic regression is a technique used in empirical modelling to discover mathematical expressions or symbolic formulas that best fit a given dataset (Billard & Diday 2002). The process of symbolic regression involves searching a space of mathematical expressions to find the equation that best fits the data. The search is typically guided by a fitness function that measures the goodness of fit between the equation and the data. The fitness function is optimized using various techniques, such as genetic algorithms, gradient descent, or other optimization algorithms. The equations discovered by symbolic regression are expressed in terms of familiar (i.e. more common) mathematical functions and variables, which can be understood and interpreted easily by humans.

In this study, we used the Python library gplearn for symbolic regression (Stephens 2016). As discussed in § 4.2, in order to have more accurate formulation for the ODE parameters before using symbolic regression, we trained two networks, NN_β and NN_τ, using the discovered values $\beta(R_0, \theta_{eq})$ and $\tau(R_0, \theta_{eq})$, enabling us to predict the parameter values for grid interpolations where no corresponding data points were available. The predicted parameters from both PINNs and the NN_β and NN_τ networks are fed through the symbolic regression model to discover a mathematical formulation for each parameter. For this purpose, θ_{eq} and R_0 are fed as inputs, and η, β and τ are the outputs. We set the population size to 5000, and evolve 20 generations until the error is close to 1 %. Since the equation consists of basic operations such as addition, subtraction, multiplication and division, we do not require any custom functions.

The following results of symbolic regression can be substituted in (3.1):

$$\left.\begin{aligned} \eta &= -0.255, \\ \beta &= 0.283 + 0.27\left(\frac{\theta_{eq}}{d}\right), \\ \tau &= 6.13\left(\frac{d}{\theta_{eq}} + 1\right), \end{aligned}\right\} \tag{5.1}$$

where d is the initial size of the droplet in mDPD units.

Figure 11 shows the history of $\alpha(t, R_0, \theta_{eq})$ using the the ODE (3.1) with parameters from (5.1). The conversion between d in mDPD units and R_0 in physical unit is $R_0 = d \times 1.701 \times 10^{-2}$ mm.

Figure 12(a) depicts the values of α for initial drop size $R_0 = 0.127$ mm and contact angles $\theta_{eq} = 93.4^\circ$ and 87.2°. Figure 12(b)compares the solution of (3.1), using the discovered parameters (5.1), with the simulation data. The figure demonstrates the agreement between the ODE solution and the actual simulation results for this particular, unseen dataset. It is important to note that this particular drop size lies outside the training interval for initial drop sizes, [0.136, 0.204] mm.

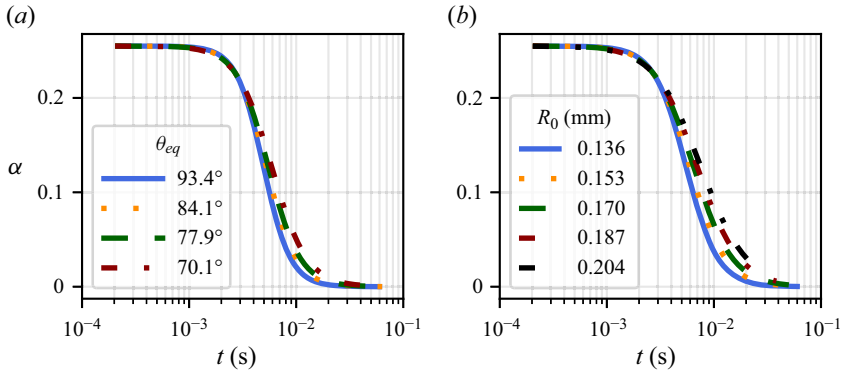


Figure 11. The behaviour of α from the right-hand side of (3.1) with parameters from (5.1): (*a*) different contact angles with fixed initial radius $R_0 = 0.136$ mm; (*b*) varying initial radii with fixed contact angle $\theta_{eq} = 77.9^\circ$.

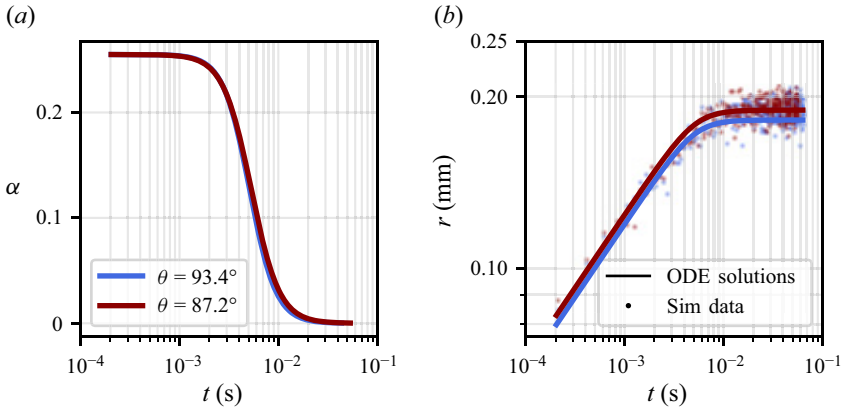


Figure 12. (*a*) The behaviour of the parameter α using (3.1) for $R_0 = 0.127$ mm, which falls outside the range of the initial drop sizes used for training the networks. (*b*) The simulation data and the solution obtained from solving the ODE (3.1) with parameters from symbolic regression (5.1).

6. Bayesian physics-informed neural network results

Bayesian physics-informed neural networks integrate the traditional PINN framework with Bayesian neural networks (BNNs; Bykov *et al.* 2021) to enable quantification of uncertainty in predictions (Yang *et al.* 2021). This framework combines the advantages of BNNs (Bishop 1997) and PINNs to address both forward and inverse nonlinear problems. By choosing a prior over the ODE and network parameters, and defining a likelihood function, one can find posterior distributions using Bayes' theorem. The B-PINNs offer a robust approach for handling problems containing uncorrelated noise, and they provide aleatoric and epistemic uncertainty quantification on the parameters of neural networks and ODEs. The BNN component of the prior adopts Bayesian principles by assigning probability distributions to the weights and biases of the neural network. To account for noise in the data, we add noise to the likelihood function. By applying Bayes' rule, we can estimate the posterior distribution of the model and the ODE parameters. This estimation process enables the propagation of uncertainty from the observed data to the predictions

made by the model. We write (3.1) as

$$\mathcal{N}_t(r;\boldsymbol{\lambda}) = f(t), \quad t \in \mathbb{R}^+, \tag{6.1a}$$

$$\mathcal{I}(r,\boldsymbol{\lambda}) = r_0, \quad t = 0, \tag{6.1b}$$

where $\boldsymbol{\lambda} = [\eta, \beta, \tau]^{\mathrm{T}}$ is a vector of the parameters of the ODE (3.1), $\mathcal{N}_t$ is a general differential operator, $f(t)$ is the forcing term, and $\mathcal{I}$ is the initial condition. This problem is an inverse problem; $\boldsymbol{\lambda}$ is inferred from the data with estimates on aleatoric and epistemic uncertainties. The likelihoods of simulation data and ODE parameters are given as

$$\left.\begin{aligned}
&P(\mathcal{D} \mid \boldsymbol{\theta},\boldsymbol{\lambda}) = P(\mathcal{D}_r \mid \boldsymbol{\theta})\, P(\mathcal{D}_f \mid \boldsymbol{\theta},\boldsymbol{\lambda})\, P(\mathcal{D}_{\mathcal{I}} \mid \boldsymbol{\theta},\boldsymbol{\lambda}), \quad \text{where} \\
&P(\mathcal{D}_r \mid \boldsymbol{\theta},\boldsymbol{\lambda}) = \prod_{i=1}^{N_r} \frac{1}{\sqrt{2\pi\sigma_r^{(i)^2}}} \exp\Big[-\frac{(r(\boldsymbol{t}_r^{(i)};\boldsymbol{\theta},\boldsymbol{\lambda}) - \bar{r}^{(i)})^2}{2\sigma_r^{(i)^2}}\Big], \\
&P(\mathcal{D}_f \mid \boldsymbol{\theta},\boldsymbol{\lambda}) = \prod_{i=1}^{N_f} \frac{1}{\sqrt{2\pi\sigma_f^{(i)^2}}} \exp\Big[-\frac{(f(\boldsymbol{t}_f^{(i)};\boldsymbol{\theta},\boldsymbol{\lambda}) - \bar{f}^{(i)})^2}{2\sigma_f^{(i)^2}}\Big], \\
&P(\mathcal{D}_{\mathcal{I}} \mid \boldsymbol{\theta},\boldsymbol{\lambda}) = \prod_{i=1}^{N_{\mathcal{I}}} \frac{1}{\sqrt{2\pi\sigma_{\mathcal{I}}^{(i)^2}}} \exp\Big[-\frac{(\mathcal{I}(\boldsymbol{t}_i^{(i)};\boldsymbol{\theta},\boldsymbol{\lambda}) - \bar{\mathcal{I}}^{(i)})^2}{2\sigma_{\mathcal{I}}^{(i)^2}}\Big],
\end{aligned}\right\} \tag{6.2}$$

where $D = D_r \cup D_f \cup D_{\mathcal{I}}$, with $\mathcal{D}_r = \{(\ln t_r^{(i)}, \ln r^{(i)})\}_{i=1}^{N_r}$, $\mathcal{D}_f = \{(t_f^{(i)}, f^{(i)})\}_{i=1}^{N_f}$, $\mathcal{D}_{\mathcal{I}} = \{(t_{\mathcal{I}}^{(i)}, \mathcal{I}^{(i)})\}_{i=1}^{N_{\mathcal{I}}}$ are scattered noisy measurements. The joint posterior of $[\boldsymbol{\theta},\boldsymbol{\lambda}]$ is given as

$$\begin{aligned}
P(\boldsymbol{\theta},\boldsymbol{\lambda} \mid \mathcal{D}) &= \frac{P(\mathcal{D} \mid \boldsymbol{\theta},\boldsymbol{\lambda})\, P(\boldsymbol{\theta},\boldsymbol{\lambda})}{P(\mathcal{D})} \\
&\approx P(\mathcal{D} \mid \boldsymbol{\theta},\boldsymbol{\lambda})\, P(\boldsymbol{\theta},\boldsymbol{\lambda}) \\
&= P(\mathcal{D} \mid \boldsymbol{\theta},\boldsymbol{\lambda})\, P(\boldsymbol{\theta})\, P(\boldsymbol{\lambda}).
\end{aligned} \tag{6.3}$$

To sample the parameters from the the posterior probability distribution defined by (6.3), we utilized the Hamiltonian Monte Carlo approach (Radivojević & Akhmatskaya 2020), which is an efficient Markov Chain Monte Carlo (MCMC) method (Brooks 1998). For a detailed description of the method, see e.g. Neal (2011, 2012) and Graves (2011). To sample the posterior probability distribution, however, variational inference (Blei, Kucukelbir & McAuliffe 2017) could also be used. In variational inference, the posterior density of the unknown parameter vector is approximated by another parametrized density function, which is restricted to a smaller family of distributions (Yang *et al.* 2021). To compute the uncertainty in the ODE parameters by using B-PINN, a noise of 5 % was added to the original dataset. The noise was sampled from a normal distribution with mean 0 and standard deviation ± 1.

Here, the neural network model architecture comprises of two hidden layers, each containing 50 neurons. The network takes time t as the input, and generates a droplet radius $r(t)$ as the output. Additionally, we include a total of 2000 burn-in samples.

The computational expense of B-PINNs compared to traditional neural networks arises primarily from the iterative nature of Bayesian inference and the need to sample from the posterior distribution. The B-PINNs involve iterative Bayesian inference,

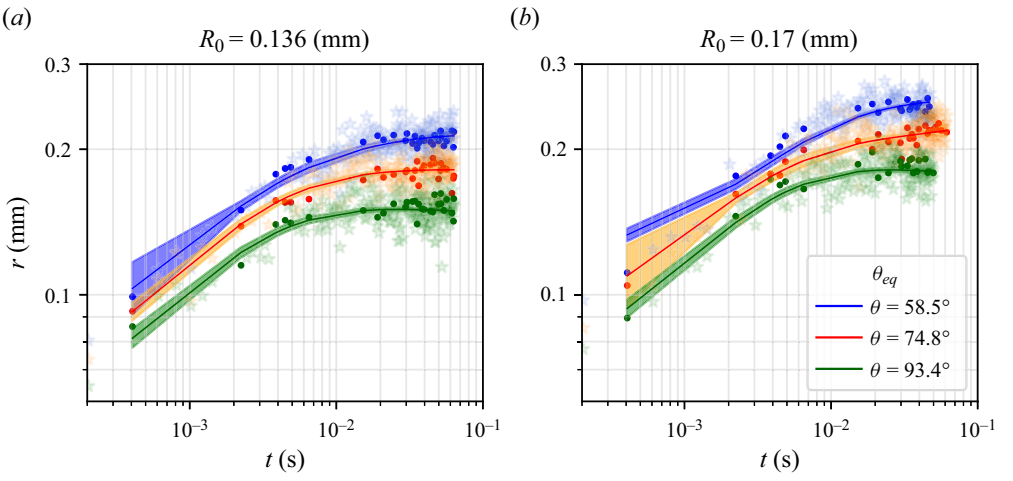


Figure 13. The mean and uncertainty (mean ±2 standard deviations) of B-PINN predictions of the spreading radii history are given as solid lines and shaded regions, respectively. The test simulation data are depicted by solid circles, and training data are indicated by stars. This analysis is carried out for two different initial drop sizes, namely $R_0 = 0.136$ and 0.170 mm, for three equilibrium contact angles.

where the posterior distribution is updated iteratively based on observed data. This iterative process requires multiple iterations to converge to a stable solution, leading to increased computational cost compared to non-iterative methods. Moreover, B-PINNs employ sampling-based algorithms such as MCMC or variational inference to estimate the posterior distribution of the model parameters. These algorithms generate multiple samples from the posterior distribution, which are used to approximate uncertainty and infer calibrated parameters.

Sampling from the posterior distribution can be computationally expensive, particularly for high-dimensional parameter spaces or complex physics models. Furthermore, B-PINNs often require running multiple forward simulations of the physics-based model for different parameter samples. Each simulation represents a potential configuration of the model parameters. Since physics-based simulations can be computationally intensive, conducting multiple simulations significantly increases the computational cost of training B-PINNs. Achieving a high acceptance rate for posterior samples, especially for high-dimensional data, demands running a large number of simulations. This further adds to the computational complexity.

Due to the computational expense associated with B-PINNs, we opt to use the method selectively for a few cases only. Utilizing the insights gained from PINNs prediction and symbolic regression, specifically the known value of $\eta = -0.255$, we can leverage the power of B-PINNs to uncover and ascertain the values of the parameters β and τ. Figure 13 showcases the comparison between the mean values of the radii $\tilde{r}(t)$ predicted by B-PINNs represented by solid lines, the corresponding standard deviations denoted by highlighted regions, and the simulation data used for training presented as stars, while the test data are indicated by coloured circles. The horizontal axes represent time, while the vertical axes depict the spreading radii $\tilde{r}(t)$. This comparative analysis is conducted for two distinct initial drop sizes, namely $R_0 = 0.136$ and 0.170 mm, considering various equilibrium contact angles.

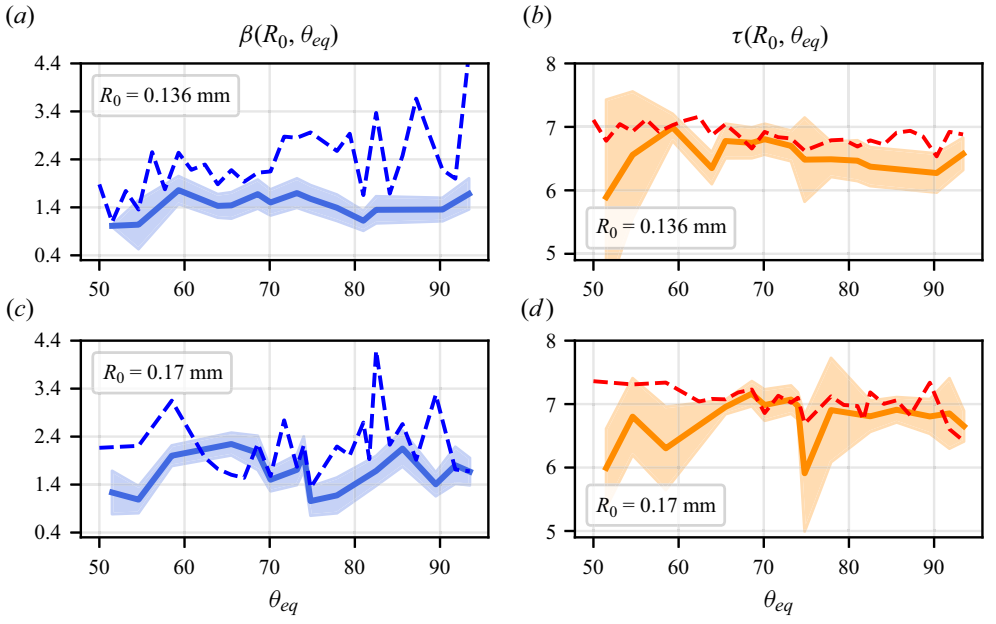


Figure 14. Comparison between B-PINNs and PINNs discovered parameters for a range of equilibrium contact angles and two initial radii. The mean values (solid lines) and the standard deviations (mean values ±2 standard deviations, shaded region) of (*a*,*c*) β and (*b*,*d*) τ. The dashed lines represent the parameters discovered by PINNs.

Figure 14 illustrates the mean values of the parameters β and τ obtained using B-PINNs along with their corresponding standard deviations. The solid lines represent the average values of the discovered parameters, while the highlighted regions indicate the standard deviations. The parameters discovered by PINNs are represented by the dashed lines. In figures 14(*a*,*c*), the vertical axes represent the values of β, while figures 14(*b*,*d*) display the values of τ. The results are presented for two initial drop sizes, $R_0 = 0.136$ and 0.17 mm. From the plots, it can be observed that the parameter β exhibits a range of values between 1.0 and 3.0. On the other hand, the parameter τ fluctuates within the range 5.0 to 7.0. These ranges provide insight into the variability and uncertainty associated with the estimated values of β and τ obtained through the B-PINN methodology.

By comparing figures 9 and 14, it becomes evident that the discovered parameters β and τ using PINNs of B-PINNs frameworks exhibit remarkable similarity. This striking similarity reinforces the efficacy and capability of our models in identifying accurately the parameters of the ODE (3.1). The close alignment between the discovered parameters in both figures demonstrates the robustness and reliability of our models. It highlights their ability to capture effectively the underlying dynamics and characteristics of the spreading behaviourof CMAS, leading to accurate parameter estimation. This consistency and agreement between the PINN and B-PINN results provide further validation of the power and effectiveness of our modelling approaches in uncovering the true values of the parameters β and τ in the ODE. Additionally, (3.1) with anticipated parameters obtained using B-PINNs is solved using Odeint. The results are presented in figure 15, which provides a comparison between the simulated spreading radii $r(t)$ (circles) and the solution of the ODE (solid lines). This comparison is conducted for initial drop sizes $R_0 = 0.136$ and 0.17 mm, considering different contact angles.

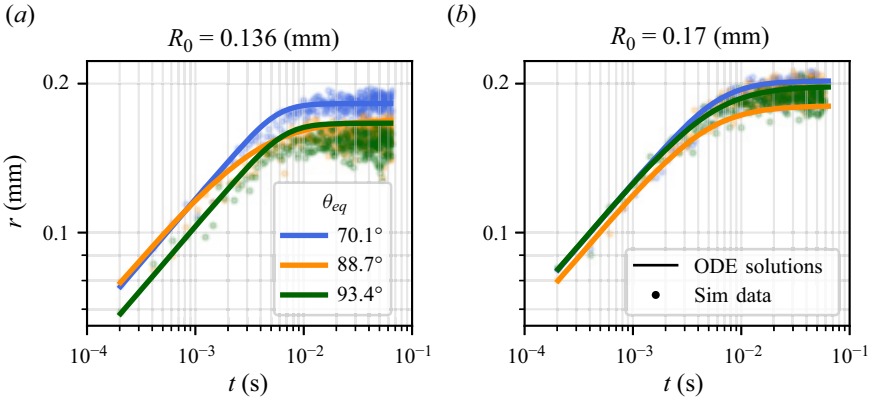


Figure 15. Comparison between the ODE solution with parameters found by B-PINNs (solid lines), and the simulation radii (circles). Two initial drop sizes, $R_0 = 0.136$ and 0.170 mm, and three equilibrium contact angles are shown.

7. Conclusions

This study introduces a new approach to model the spreading dynamics of molten CMAS droplets. In the liquid state, CMAS is characterized by high viscosity, density and surface tension. The main objective is to achieve a comprehensive understanding of the spreading dynamics by integrating the underlying physics into the neural network architecture.

The study emphasizes the potential of PINNs in analysing complex systems with intricate dynamics. To study the dynamics of CMAS droplets, we performed simulations using the mDPD method. By analysing the simulation data and observing the droplet behaviour, we proposed a parametric equation (3.1), which consists of three unknown parameters. This parametric equation aims to capture and describe the observed behaviour of the CMAS droplets based on the simulation results. Using the data from the mDPD simulations, the study employed the PINNs framework to determine the parameters of the equation. Symbolic regression was then utilized to establish the relationship between the identified parameter values, and the initial droplet radii and contact angles. As a result, a simplified ODE model was developed, accurately capturing the spreading dynamics. The model's parameters were determined explicitly based on the droplet's geometry and surface properties. Furthermore, B-PINNs were employed to assess the uncertainty associated with the model predictions, providing a comprehensive analysis of the spreading behaviour of CMAS droplets.

In reality, the CMAS attack involves a complex interplay of reaction kinetics between CMAS and the thermal barrier coating, surface morphology (roughness, porosity) occurring under highly non-uniform thermal conditions. However, our findings extend beyond the isothermal conditions used in this study and even the specific case of CMAS droplets. The relationships uncovered and methods developed in this study have broader applications in understanding the spreading dynamics of droplets in general. By leveraging the insights gained from this research, one can investigate and understand the behaviour of droplets in diverse contexts, furthering our understanding of droplet spreading phenomena. Potentially, this knowledge can be used in developing strategies for effective droplet management and optimizing processes involving droplets in a wide range of practical applications.

The key aim of this work is to create a mathematical model for the partial wetting induced by CMAS droplets with specific properties. It is clear that this model has the

capacity for expansion to encompass diverse fluid properties. By exploring different fluids with unique characteristics, we can employ the same methodology to uncover the ODE's correlation with fluid properties. For future work, we can extend the proposed equations to incorporate fluid properties by considering different fluids with different properties, as the same method can be used to find the dependency of the ODE to fluid properties. Also, it is worth pursuing a systematic comparison of our results with Gordillo's theory (Gordillo *et al.* 2019; Gorin *et al.* 2022). Additionally, in the case of a contact angle θ changing with time, a similar scaling law rationale can be applied, i.e. $\theta(t) \sim t^{\gamma}$ (see De Ruijter, De Coninck & Oshanin 1999). One can extend the current study to a system of ODEs that incorporates temporal variations in both radii and the advancing contact angle.

Funding. E.K. thanks the Mitacs Globalink Research Award Abroad and Western University's Science International Engagement Fund Award. M.Ka. thanks the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Canada Research Chairs Program. The work was partially supported by a US Department of Energy (DOE) grant (DE-SC0023389). R.B.K. would like to acknowledge the support received from the US Army Research Office Mathematical Sciences Division for this research through grant no. W911NF-17-S-0002. L.B. and A.G. were supported by the US Army Research Laboratory 6.1 basic research program in propulsion sciences. Z.L. and G.K. acknowledge support from the AIM for Composites, an Energy Frontier Research Center funded by the DOE, Office of Science, Basic Energy Sciences (BES) under Award no. DE-SC0023389. Z.L. also acknowledges support from the National Science Foundation (grants OAC-2103967 and CDS&E-2204011). Computing facilities were provided by the Digital Research Alliance of Canada (https://alliancecan.ca) and the Center for Computation and Visualization, Brown University. The authors acknowledge the resources and support provided by Department of Defense Supercomputing Resource Center (DSRC) through use of 'Narwhal' as part of the 2022 Frontier Project, Large-Scale Integrated Simulations of Transient Aerothermodynamics in Gas Turbine Engines. Work at Brown University and Clemson University was supported by AIM for Composites, an Energy Frontier Research Center funded by the DOE, Office of Science, BES under Award no. DE-SC0023389 (AI Models).

Declaration of interests. The authors report no conflict of interest.

Author ORCIDs.
Elham Kiyani https://orcid.org/0000-0002-5840-7378;
Mahdi Kooshkbaghi https://orcid.org/0000-0002-6344-7382;
Khemraj Shukla https://orcid.org/0000-0002-2262-7264;
Rahul Babu Koneru https://orcid.org/0000-0002-5212-1360;
Zhen Li https://orcid.org/0000-0002-0936-6928;
Luis Bravo https://orcid.org/0000-0003-1604-9303;
Anindya Ghoshal https://orcid.org/0000-0001-6338-1224;
George Em Karniadakis https://orcid.org/0000-0002-9713-7120;
Mikko Karttunen https://orcid.org/0000-0002-8626-3033.

References

Bansal, N.P. & Choi, S.R. 2014 Properties of desert sand and CMAS glass. *Tech. Rep.* NASA/TM-2014-218365. NASA Glenn Research Center Cleveland.

Billard, L. & Diday, E. 2002 Symbolic regression analysis. In *Classification, Clustering, and Data Analysis: Recent Advances and Applications* (ed. K. Jajuga, A. Sokołowski & H.-H. Bock), pp. 281–288. Springer.

Bishop, C.M. 1997 Bayesian neural networks. *J. Braz. Comput. Soc.* **4**, 61–68.

Blei, D.M., Kucukelbir, A. & McAuliffe, J.D. 2017 Variational inference: a review for statisticians. *J. Am. Stat. Assoc.* **112** (518), 859–877.

Bonn, D., Eggers, J., Indekeu, J., Meunier, J. & Rolley, E. 2009 Wetting and spreading. *Rev. Mod. Phys.* **81** (2), 739–805.

Brooks, S. 1998 Markov Chain Monte Carlo method and its application. *J. R. Stat. Soc.* D **47**, 69–100.

BRUNTON, S.L., PROCTOR, J.L. & KUTZ, J.N. 2016 Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proc. Natl Acad. Sci. USA* **113** (15), 3932–3937.

BYKOV, K., HÖHNE, M.M.-C., CREOSTEANU, A., MÜLLER, K.-R., KLAUSCHEN, F., NAKAJIMA, S. & KLOFT, M. 2021 Explaining Bayesian neural networks. arXiv:2108.10346.

CHAN, K.C., LI, Z. & WENZEL, W. 2023 A Mori–Zwanzig dissipative particle dynamics approach for anisotropic coarse grained molecular dynamics. *J. Chem. Theory Comput.* **19** (3), 910–923.

CHEN, L., BONACCURSO, E., DENG, P. & ZHANG, H. 2016 Droplet impact on soft viscoelastic surfaces. *Phys. Rev.* E **94**, 063117.

CHEN, Y., LU, L., KARNIADAKIS, G.E. & DAL NEGRO, L. 2020 Physics-informed neural networks for inverse problems in nano-optics and metamaterials. *Opt. Express* **28**, 11618–11633.

CLARKE, D.R., OECHSNER, M. & PADTURE, N.P. 2012 Thermal-barrier coatings for more efficient gas-turbine engines. *MRS Bull.* **37** (10), 891–898.

CORMIER, S.L., MCGRAW, J.D., SALEZ, T., RAPHAËL, E. & DALNOKI-VERESS, K. 2012 Beyond Tanner's law: crossover between spreading regimes of a viscous droplet on an identical film. *Phys. Rev. Lett.* **109** (15), 154501.

DE RUIJTER, M.J., DE CONINCK, J. & OSHANIN, G. 1999 Droplet spreading: partial wetting regime revisited. *Langmuir* **15** (6), 2209–2216.

DELAHUNT, C.B. & KUTZ, J.N. 2022 A toolkit for data-driven discovery of governing equations in high-noise regimes. *IEEE Access* **10**, 31210–31234.

DUSSAN, E.B. 1979 On the spreading of liquids on solid surfaces: static and dynamic contact lines. *Annu. Rev. Fluid Mech.* **11** (1), 371–400.

EDDI, A., WINKELS, K.G. & SNOEIJER, J.H. 2013 Short time dynamics of viscous drop spreading. *Phys. Fluids* **25** (1), 013102.

EDWARDS, A.M.J., LEDESMA-AGUILAR, R., NEWTON, M.I., BROWN, C.V. & MCHALE, G. 2020 A viscous switch for liquid–liquid dewetting. *Commun. Phys.* **3** (1), 1–6.

ESPAÑOL, P. & WARREN, P. 1995 Statistical mechanics of dissipative particle dynamics. *Europhys. Lett.* **30** (4), 191.

ESPAÑOL, P. & WARREN, P.B. 2017 Perspective: dissipative particle dynamics. *J. Chem. Phys.* **146** (15), 150901.

DE GENNES, P.G. 1985 Wetting: statics and dynamics. *Rev. Mod. Phys.* **57** (3), 827–863.

GHOUFI, A. & MALFREYT, P. 2012 Coarse grained simulations of the electrolytes at the water–air interface from many body dissipative particle dynamics. *J. Chem. Theory Comput.* **8** (3), 787–791.

GORDILLO, J.M., RIBOUX, G. & QUINTERO, E.S. 2019 A theory on the spreading of impacting droplets. *J. Fluid Mech.* **866**, 298–315.

GORIN, B., DI MAURO, G., BONN, D. & KELLAY, H. 2022 Universal aspects of droplet spreading dynamics in Newtonian and non-Newtonian fluids. *Langmuir* **38** (8), 2608–2613.

GRANT, K.M., KRÄMER, S., LÖFVANDER, J.P.A. & LEVI, C.G. 2007 CMAS degradation of environmental barrier coatings. *Surf. Coat. Technol.* **202** (4–7), 653–657.

GRAVES, A. 2011 Practical variational inference for neural networks. In *Advances in Neural Information Processing Systems* (ed. J. Shawe-Taylor, R. Zemel, P. Bartlett, F. Pereira & K.Q. Weinberger), vol. 24, pp. 2348–2356. Curran Associates.

GROOT, R.D. 2004 Applications of dissipative particle dynamics. In *Novel Methods in Soft Matter Simulations* (ed. M. Karttunen, A. Lukkarinen & I. Vattulainen), pp. 5–38. Springer.

HARDY, W.B. 1919 III. The spreading of fluids on glass. *Lond. Edinb. Dublin Philos. Mag. J. Sci.* **38** (223), 49–55.

HASSAN, G., YILBAS, B.S., AL-SHARAFI, A. & AL-QAHTANI, H. 2019 Self-cleaning of a hydrophobic surface by a rolling water droplet. *Sci. Rep.* **9**, 1–14.

JAIN, N., LEMOINE, A., CHAUSSONNET, G., FLATAU, A., BRAVO, L., GHOSHAL, A., WALOCK, M. & MURUGAN, M. 2021 A critical review of physical models in high temperature multiphase fluid dynamics: turbulent transport and particle-wall interactions. *Appl. Mech. Rev.* **73**, 040801.

KARNIADAKIS, G.E., KEVREKIDIS, I.G., LU, L., PERDIKARIS, P., WANG, S. & YANG, L. 2021 Physics-informed machine learning. *Nat. Rev. Phys.* **3**, 422–440.

KIYANI, E., SHUKLA, K., KARNIADAKIS, G.E. & KARTTUNEN, M. 2023 A framework based on symbolic regression coupled with extended physics-informed neural networks for gray-box learning of equations of motion from data. arXiv:2305.10706.

KIYANI, E., SILBER, S., KOOSHKBAGHI, M. & KARTTUNEN, M. 2022 Machine-learning-based data-driven discovery of nonlinear phase-field dynamics. *Phys. Rev.* E **106**, 065303.

KONERU, R.B., FLATAU, A., LI, Z., BRAVO, L., MURUGAN, M., GHOSHAL, A. & KARNIADAKIS, G.E. 2022 Quantifying the dynamic spreading of a molten sand droplet using multiphase mesoscopic simulations. *Phys. Rev. Fluids* **7** (10), 103602.

DE LAPLACE, P.-S. 1805 Supplément au livre X du traité de mécanique céleste. Sur l'action capillaire. In *Traité de mécanique céleste*. Gauthier-Vilars.

LEE, S., KOOSHKBAGHI, M., SPILIOTIS, K., SIETTOS, C.I. & KEVREKIDIS, I.G. 2020 Coarse-scale PDEs from fine-scale observations via machine learning. *Chaos* **30**, 013141.

LEI, L., BERTEVAS, E.L., KHOO, B.C. & PHAN-THIEN, N. 2018 Many-body dissipative particle dynamics (MDPD) simulation of a pseudoplastic yield-stress fluid with surface tension in some flow processes. *J. Non-Newtonian Fluid Mech.* **260**, 163–174.

LI, Z., BIAN, X., TANG, Y.-H. & KARNIADAKIS, G.E. 2018 A dissipative particle dynamics method for arbitrarily complex geometries. *J. Comput. Phys.* **355**, 534–547.

LI, Z., BIAN, X., YANG, X. & KARNIADAKIS, G.E. 2016 A comparative study of coarse-graining methods for polymeric fluids: Mori–Zwanzig vs iterative Boltzmann inversion vs stochastic parametric optimization. *J. Chem. Phys.* **145** (4), 044102.

LI, Z., HU, G.-H., WANG, Z.-L., MA, Y.-B. & ZHOU, Z.-W. 2013 Three dimensional flow structures in a moving droplet on substrate: a dissipative particle dynamics study. *Phys. Fluids* **25**, 072103.

LUCY, L.B. 1977 A numerical approach to the testing of the fission hypothesis. *Astron. J.* **82**, 1013.

MAO, Z., JAGTAP, A.D. & KARNIADAKIS, G.E. 2020 Physics-informed neural networks for high-speed flows. *Comput. Meth. Appl. Mech. Engng* **360**, 112789.

MARTIUS, G. & LAMPERT, C.H. 2016 Extrapolation and learning equations. arXiv:1610.02995.

MCGRAW, J.D., CHAN, T.S., MAURER, S., SALEZ, T., BENZAQUEN, M., RAPHAËL, E., BRINKMANN, M. & JACOBS, K. 2016 Slip-mediated dewetting of polymer microdroplets. *Proc. Natl Acad. Sci. USA* **113** (5), 1168–1173.

MCHALE, G., SHIRTCLIFFE, N.J., AQIL, S., PERRY, C.C. & NEWTON, M.I. 2004 Topography driven spreading. *Phys. Rev. Lett.* **93** (3), 036102.

MEIDANI, K. & FARIMANI, A.B. 2021 Data-driven identification of 2D partial differential equations using extracted physical features. *Comput. Meth. Appl. Mech. Engng* **381**, 113831.

MISHRA, S. & MOLINARO, R. 2020 Estimates on the generalization error of physics informed neural networks (PINNs) for approximating PDEs. II. A class of inverse problems. arXiv:2007.01138.

MURTOLA, T., BUNKER, A., VATTULAINEN, I., DESERNO, M. & KARTTUNEN, M. 2009 Multiscale modeling of emergent materials: biological and soft matter. *Phys. Chem. Chem. Phys.* **11** (12), 1869–1892.

NARAPARAJU, R., GOMEZ CHAVEZ, J.J., NIEMEYER, P., HESS, K.-U., SONG, W., DINGWELL, D.B., LOKACHARI, S., RAMANA, C.V. & SCHULZ, U. 2019 Estimation of CMAS infiltration depth in EB-PVD TBCs: a new constraint model supported with experimental approach. *J. Eur. Ceram. Soc.* **39** (9), 2936–2945.

NDAMKA, N.L., WELLMAN, R.G. & NICHOLLS, J.R. 2016 The degradation of thermal barrier coatings by molten deposits: introducing the concept of basicity. *Mater. High Temp.* **33** (1), 44–50.

NEAL, R.M. 2011 MCMC using Hamiltonian dynamics. In *Handbook of Markov Chain Monte Carlo* (ed. S. Brooks, A. Gelman, G. Jones & X.-L. Meng), pp. 113–162. Chapman & Hall/CRC.

NEAL, R.M. 2012 *Bayesian Learning for Neural Networks*, vol. 118. Springer Science & Business Media.

NIEMINEN, J.A., ABRAHAM, D.B., KARTTUNEN, M. & KASKI, K. 1992 Molecular dynamics of a microscopic droplet on solid surface. *Phys. Rev. Lett.* **69** (1), 124–127.

NIETO, A., AGRAWAL, R., BRAVO, L., HOFMEISTER-MOCK, C., PEPI, M. & GHOSHAL, A. 2021 Calcia–Magnesia–Alumina–Silicate (CMAS) attack mechanisms and roadmap towards sandphobic thermal and environmental barrier coatings. *Intl Mater. Rev.* **66** (7), 451–492.

NISHIMOTO, S. & BHUSHAN, B. 2013 Bioinspired self-cleaning surfaces with superhydrophobicity, superoleophobicity, and superhydrophilicity. *RSC Adv.* **3** (3), 671–690.

PITOIS, O. & FRANÇOIS, B. 1999 Crystallization of condensation droplets on a liquid surface. *Colloid Polym. Sci.* **277**, 574–578.

POERSCHKE, D.L. & LEVI, C.G. 2015 Effects of cation substitution and temperature on the interaction between thermal barrier oxides and molten CMAS. *J. Eur. Ceram. Soc.* **35** (2), 681–691.

POPESCU, M.N., OSHANIN, G., DIETRICH, S. & CAZABAT, A.-M. 2012 Precursor films in wetting phenomena. *J. Phys.: Condens. Matter* **24** (24), 243102.

RADIVOJEVIĆ, T. & AKHMATSKAYA, E. 2020 Modified Hamiltonian Monte Carlo for Bayesian inference. *Stat. Comput.* **30** (2), 377–404.

RAISSI, M., PERDIKARIS, P. & KARNIADAKIS, G.E. 2019 Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* **378**, 686–707.

RAO, Q., XIA, Y., LI, J., MCCONNELL, J., SUTHERLAND, J. & LI, Z. 2021 A modified many-body dissipative particle dynamics model for mesoscopic fluid simulation: methodology, calibration, and application for hydrocarbon and water. *Mol. Simul.* **47** (4), 363–375.

REN, J. & DUAN, J. 2020 Identifying stochastic governing equations from data of the most probable transition trajectories. arXiv:2002.10251.

SHUKLA, K., DI LEONI, P.C., BLACKSHIRE, J., SPARKMAN, D. & KARNIADAKIS, G.E. 2020 Physics-informed neural network for ultrasound nondestructive quantification of surface breaking cracks. *J. Nondestruct. Eval.* **39**, 1–20.

SONG, W., LAVALLÉE, Y., HESS, K.-U., KUEPPERS, U., CIMARELLI, C. & DINGWELL, D.B. 2016 Volcanic ash melting under conditions relevant to ash turbine interactions. *Nat. Commun.* **7**, 10795.

SRINIVASAN, V. & MASON, C.H. 1986 Nonlinear least squares estimation of new product diffusion models. *Market. Sci.* **5** (2), 169–178.

STEPHENS, T. 2016 Genetic programming in Python, with a scikit-learn inspired API: gplearn.

TANNER, L.H. 1979 The spreading of silicone oil drops on horizontal surfaces. *J. Phys.* D*: Appl. Phys.* **12** (9), 1473.

THIEM, T.N., KOOSHKBAGHI, M., BERTALAN, T., LAING, C.R. & KEVREKIDIS, I.G. 2020 Emergent spaces for coupled oscillators. *Front. Comput. Neurosci.* **14**, 36.

THOMPSON, A.P., *et al.*2022 LAMMPS – a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Comput. Phys. Comms.* **271**, 108171.

VIDAL-SETIF, M.H., CHELLAH, N., RIO, C., SANCHEZ, C. & LAVIGNE, O. 2012 Calcium–magnesium–alumino-silicate (CMAS) degradation of EB-PVD thermal barrier coatings: characterization of CMAS damage on ex-service high pressure blade TBCs. *Surf. Coat. Technol.* **208**, 39–45.

WARREN, P.B. 2001 Hydrodynamic bubble coarsening in off-critical vapor–liquid phase separation. *Phys. Rev. Lett.* **87** (22), 225702.

WARREN, P.B. 2003 Vapor–liquid coexistence in many-body dissipative particle dynamics. *Phys. Rev.* E **68** (6 Pt 2), 066702.

WIESNER, V.L., VEMPATI, U.K. & BANSAL, N.P. 2016 High temperature viscosity of calcium-magnesium-aluminosilicate glass from synthetic sand. *Scr. Mater.* **124**, 189–192.

WINKELS, K.G., WEIJS, J.H., EDDI, A. & SNOEIJER, J.H. 2012 Initial spreading of low-viscosity drops on partially wetting surfaces. *Phys. Rev.* E **85**, 055301.

XIA, Y., GORAL, J., HUANG, H., MISKOVIC, I., MEAKIN, P. & DEO, M. 2017 Many-body dissipative particle dynamics modeling of fluid flow in fine-grained nanoporous shales. *Phys. Fluids* **29** (5), 056601.

YANG, L., MENG, X. & KARNIADAKIS, G.E. 2021 B-PINNs: Bayesian physics-informed neural networks for forward and inverse PDE problems with noisy data. *J. Comput. Phys.* **425**, 109913.

YOUNG, T. 1805 An essay on the cohesion of fluids. *Phil. Trans. R. Soc. Lond.* **95**, 65–87.

ZHAO, J., CHEN, S., ZHANG, K. & LIU, Y. 2021 A review of many-body dissipative particle dynamics (MDPD): theoretical models and its applications. *Phys. Fluids* **33** (11), 112002.