Accepted to Phys. Fluids 10.1063/5.0106594

Computational Fluid Dynamics and Machine Learning Algorithms Analysis of Striking Particle Velocity Magnitude, Particle Diameter, and Impact Time Inside an Acinar Region of Human Lung

Isabella Francis[1], Suvash C. Saha[1,*]

[1]School of Mechanical and Mechatronic Engineering, Faculty of Engineering and Information Technology, University of Technology Sydney, NSW, Australia

*Corresponding author: Suvash C. Saha; Email: Suvash.Saha@uts.edu.au

Abstract. Complementing computational fluid dynamics (CFD) simulations with machine learning Algorithms (MLA) is becoming increasingly popular as the combination reduces the computational time of the CFD simulations required for classifying, predicting, or optimizing the impact of geometrical and physical variables of a specific study. The main target of drug delivery studies is indicating the optimum particle diameter for targeting particular locations in the lung to achieve a desired therapeutic effect. In addition, the main goal of molecular dynamics studies is to investigate particle-lung interaction through given particle properties. Therefore, this study combines the two by numerically determining the optimum particle diameter required to obtain an ideal striking velocity magnitude (velocity at the time of striking the alveoli, i.e. deposition by sedimentation/diffusion) and impact time (time from release until deposition) inside an acinar part of the lung. At first, the striking velocity magnitudes and time for impact (two independent properties) of three different particle diameters ($0.5\ \mu m, 1.5\ \mu m, 5\ \mu m$) are computed using CFD simulations. Then, machine learning classifiers determine the particle diameter corresponding to

these two independent properties. In this study, two cases are compared: A healthy acinus where a surfactant layer covers the inner surface of the alveoli providing low air-liquid surface tension (LST) values ($10\ mN/m$), and a diseased acinus where only a water layer covers the surface causing high surface tension (HST) values ($70\ mN/m$). In this study, the airflow velocity throughout the breathing cycle corresponds to a person with a respiratory rate of 13 breaths per minute and a volume flow rate of $6\ l/min$. Accurate machine learning (ML) results showed that all three particle diameters attain larger velocities and smaller impact times in a diseased acinus compared to a healthy one. In both cases, the 0.5 μm particles acquire the smallest velocities and longest impact times, while the 1.5 μm particles possess the largest velocities and shortest impact times.

Keywords: Computational fluid dynamics (CFD); Machine learning classification; Particle striking velocity magnitude; Time for impact; Surface tension; Surfactant

Introduction

Pulmonary surfactant is a vital biological barrier residing on the lung's inner surface. It reduces the risk of infection by protecting the lungs from toxic nanoparticles, preserving lung homeostasis, and preventing alveolar collapse at the end of exhalation by decreasing the surface tension of the air-liquid interface inside the lungs (Chroneos et al. 2010). A pulmonary acinus is a gas-exchanging lung unit located distal to a single terminal bronchus (Haefeli-Bleuer and Weibel 1988), composed of a group of respiratory bronchioles, alveolar ducts, alveolar sacs and alveoli. The alveolar surface of the lungs is composed of ultrathin lung tissue that allows maximum gas exchange but also poses a threat upon inhaling toxic particles that can easily penetrate through the tissue into the bloodstream.

Accepted to Phys. Fluids 10.1063/5.0106594

Toxic nanoparticles mainly originate from the combustion of fossil fuels, wildfires, biomass burning, and industrial emissions (Sonwani et al. 2021). An increased number and concentration of inhaled toxic nanoparticles can lead to the emergence of restrictive lung diseases such as acute respiratory distress syndrome (ARDS), where the integrity of the pulmonary surfactant is threatened by the loss of lung tissue (Veldhuizen et al. 1995). Conversely, nanoparticles can be therapeutic, and a significant determinant of the effectiveness of inhaled medications is their ability to reach suitable locations inside the lungs to maximize their therapeutic effects. Therefore, determining the fate of inhaled particles in the lung is crucial for the drug evaluation (Douafer et al. 2020, Francis et al. 2022).

Artificial Intelligence systems use MLA to predict output values based on inputs given to the system. Using a set of training data, the algorithm learns and extracts patterns and correlations between variables that enable the prediction of the output for a testing dataset. MLA can categorize and cluster data based on certain variables (classification) or find relationships between variables using regression algorithms (regression analysis). Other powerful capabilities of ML include being automatic, used in various fields such as medicine, education, and engineering, having the ability to handle varieties of data, and reducing the computational time of CFD simulations. However, the disadvantages of machine learning include the requirement of having a decent amount of data for proper training and thus accurate predictions, having the chance of some error in interpreting the data where the data may be biased and not correctly representing the entire dataset, and requiring program space to store or manage the data. As MLA decreases the computational time of CFD simulations, several studies have utilized machine learning schemes and algorithms with CFD to investigate fluid dynamics (Uçar et al. 2017, Zhu et al. 2019, Raibaudo et al. 2020, Schmidt et al. 2021) and heat transfer phenomena. For example, one study by Rüttgers et al (2022) used machine

learning algorithms to conduct respiratory flow simulations for the human nasal cavity and pharynx extracted from computer tomography data. Another study by Hanna et al. (2020) used machine learning to improve the performance of a CFD simulation with a coarse mesh grid. Mohammadpour et al. (2021) used the discrete phase model in a parametric CFD study and a hybrid support vector regression-particle swarm optimization technique using machine learning to maximize the heat transfer of a nanofluid in a microchannel heat sink. Also, a recent study by Ringstad et al. (2021) trained a Gaussian Process Regression machine learning algorithm with data extracted from ANSYS FLUENT 2019 R3 to generate databases of high-quality CFD data for a two-phase ejector under various ejector operating parameters and geometrical configurations. However, to our knowledge, no study has combined CFD and MLA to analyze particle dynamics and kinematics inside the lungs/alveoli.

Without the use of machine learning techniques, an extensive amount of computational studies have investigated micron-sized and nano-sized particle trajectories in different scenarios (Ault et al. 2016, Mollicone et al. 2019) and their deposition fractions inside the upper region of the lung (Ma and Lutchen 2009, Zhang et al. 2009, Walters and Luke 2011, Soni and Aliabadi 2013, Tian et al. 2015, Xi et al. 2020). Studies consider different airway models (Ma and Lutchen 2009), various particle sources (Deng et al. 2019), diseased lungs (Vinchurkar et al. 2012, Zhao et al. 2021), and aging effects (Rahman et al. 2021). In addition, a substantial amount of studies has examined the particle deposition inside the lower distal alveolar region of the lung (Khajeh-Hosseini-Dalasm and Longest 2015, Koullapis et al. 2020), considering different alveolar sizes (Xi et al. 2020), expanding walls (Sznitman et al. 2007, Harding Jr and Robinson 2010, April Si et al. 2021), and various breathing conditions (Balásházy et al. 2008, Ciloglu 2020).

With the ongoing threat of the SARS-CoV-2 virus, these studies are becoming more vital for toxicological, pharmaceutical, and molecular dynamics studies exploring the particle-surfactant interaction inside the lung (Muhlfeld et al. 2008, Sachan et al. 2012). This interaction and the degree of toxic or medicinal effects depend on the nanoparticles' concentration number, origin, diameter, composition, charge (Hedberg et al. 2012), and hydrophobicity (Hu et al. 2013). Nevertheless, the particle velocity striking the surface of the surfactant also plays a role in determining the level of nanoparticle interaction with the surfactant. In addition, the time needed for a particle to strike the surface of the alveoli is essential for targeting therapeutic drugs in specific lung regions. Therefore, the current study offers new insights into these two particle-intensive properties, the striking velocity magnitude and time for impact. Furthermore, this study paves the way for future studies of particle kinematics inside the lungs using machine learning techniques.

2. Methods

2.1 Acinar model:

This study utilizes the same physiologically accurate acinar geometry of a single airway path that had been developed in a previous study (arXiv:2204.01699 [physics.bio-ph]) starting from two respiratory bronchioles in generation 18 to alveolar sacs in generation 23 with a total of 150 alveoli and inner bifurcating angles of 30 degrees (Fig. 1a). The duct lengths decrease from 765 µm to 575 µm (generation 18 to generation 23). Similarly, duct diameters decrease gradually from 330 µm to 240 µm (Haefeli-Bleuer and Weibel 1988, Sznitman 2013). The computational mesh was generated using ANSYS MESHING (Ansys, Inc.), where 10 µm mesh size with 3.7 million tetrahedral volume elements and six boundary layers near the walls were implemented to obtain

accurate results (Fig. 1b). A piecewise high-order polynomial realistic inhalation and exhalation breathing velocity profile is adopted from the Russo and Khalifa (Russo and Khalifa 2011) and scaled-down according to our acinar model. The profile corresponds to a volume flow rate of 6l/min and consists of a complete breathing cycle of 4.75 *s*, with an inhalation phase of 1.75 *s*, an exhalation phase of 2 s, and a pause of 1 *s* (Fig. 1c).

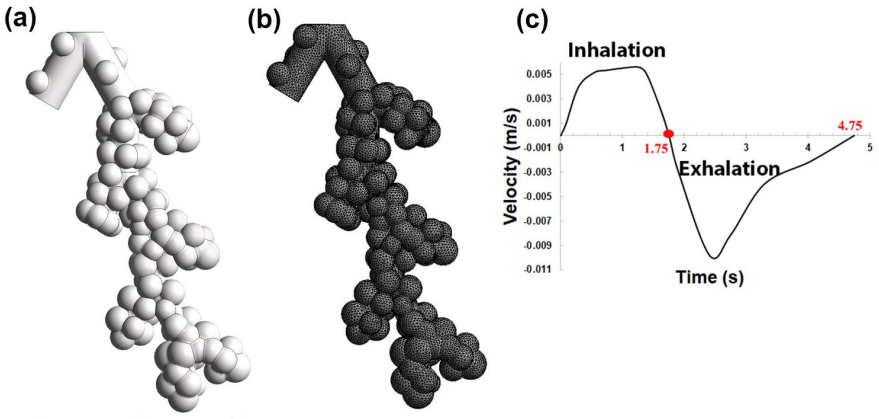

Figure 1: (a) Acinar model with hemispherical alveoli, (b) Computational meshing with 10 µm tetrahedral volume elements, (c) Realistic breathing velocity profile with positive velocity values in the inhalation phase and negative ones in the exhalation phase

2.2 Computational Fluid Dynamics Method:

Computational fluid dynamics simulation is performed considering three phases: Gas (air), liquid (surfactant layer or water layer), and solid (particles).

For the simulation of the gas phase, the flow is known to be subsonic and laminar in submillimeter acinar airspaces as the ratio of fluid to sound speed (Mach number), and the Reynolds number are

each less than unity. Therefore, the air density is relatively constant, and the flow is considered incompressible and laminar. The simulation of airflow motion is achieved by solving the conservation of mass and momentum equations for unsteady and incompressible flows using ANSYS Fluent 2020 R2 (Canonsburg, PA). For the transient flow, the Pressure-Implicit with Splitting Operators (PISO) pressure-velocity coupling scheme is chosen as the algorithm performs neighbour and skewness corrections that dramatically decrease the number of iterations required for convergence. A second-order body force weighted pressure-based spatial discretisation, a second-order momentum-based spatial discretisation, and a modified high-resolution interface capturing (HRIC) Volume of Fluid (VOF) spatial discretisation were applied. The realistic human breathing profile (Fig. 1c) is employed through a user-defined function at the inlet boundary condition.

For the liquid phase, the Multiphase Homogeneous VOF model integrates the surfactant layer on the surface of the alveoli for the first case and the water layer for the second case. The secondary phases are patched with a thickness of 1 μm on the alveolar surfaces. The VOF model is applied with an implicit formulation, implicit body force, and a sharp interface modeling with interfacial anti-diffusion to distinguish air volume from surfactant or water. The primary phase is air with isothermal properties at body temperature. The secondary phase is a surfactant with density $\rho = 25\ kg/\mathrm{m}^3$, dynamic viscosity $\mu = 0.000813\,\frac{\mathrm{kg}}{\mathrm{m}.\mathrm{s}}$, or water with $\rho = 993.3\ \mathrm{kg/m}^3$ and $\mu = 0.001003\,\frac{\mathrm{kg}}{\mathrm{m}.\mathrm{s}}$. Surface tension forces modelled the interaction between the two phases with a continuum surface force model (Brackbill et al. 1992). The surface tension is set to 10 mN/m and 70 mN/m for the surfactant and water cases, respectively.

For the solid phase, the Discrete Phase Model (DPM) is applied for Lagrangian particle tracking (Saidi et al. 2014), which considers the particle's inertia, diffusivity, and near-wall effects while tracking the individual particle trajectory before striking the alveoli surface. From the inlet of the acinar model, 132 monodispersed particles were injected at 0.05s using face normal direction with a total flow rate of $1 \times 10^{-20}\ kg/s$ scaled by face area for a duration of 1.75s. Particles were tracked using unsteady particle tracking with one-way coupling considering dilute particle suspensions as the volumetric fraction of the discrete phase compared to the continuous phase is negligible. For each surfactant and water case, three simulations were performed for three different spherically shaped micron-sized particles, $d = 0.5\ \mu m, 1.5\ \mu m, 5\ \mu m$. The walls of the acinar model trap the particles striking their surfaces. A particle's position and velocity information are stored as soon as it touches the airway wall, and its trajectory calculations are terminated.

2.2.1 VOF model equations

- The continuity equation for primary phase 1 and secondary phase 2 is the following:

$$\frac{\partial \rho}{\partial t} + \nabla . (\rho \vec{\boldsymbol{u}}) = S \qquad \textbf{(1)}$$

Where $\rho = \alpha_1 \rho_1 + \alpha_2 \rho_2$ is the total density with α and ρ being the volume fraction and the density, respectively, $\vec{\boldsymbol{u}} = \frac{1}{\rho}(\alpha_1 \rho_1 \vec{\boldsymbol{u}}_{\boldsymbol{1}} + \alpha_2 \rho_2 \vec{\boldsymbol{u}}_{\boldsymbol{2}})$ is the mixture velocity, and S is a source term.

Both phases are solved using the same momentum equation, and the resulting velocity field is shared between both phases. Based on the local value of the volume fraction, properties and variables are assigned to each control volume.

- The momentum equation of the mixture is the following:

$$\frac{\partial(\rho\vec{\boldsymbol{u}})}{\partial t} + \nabla\cdot(\rho\vec{\boldsymbol{u}}\vec{\boldsymbol{u}}) = -\nabla p + \nabla\left(\mu(\boldsymbol{\nabla}\vec{\boldsymbol{u}} + \boldsymbol{\nabla}\vec{\boldsymbol{u}}^{T})\right) + \rho\vec{\boldsymbol{g}} + \vec{\boldsymbol{T}}_{\boldsymbol{\sigma}} \quad (2)$$

$\vec{\boldsymbol{T}}_{\boldsymbol{\sigma}}$ is the surface tension force at the interface of the two phases,$\mu = \alpha_1\mu_1 + \alpha_2\mu_2$ is the total dynamic viscosity, and $\vec{g}$ is the gravitational force. According to the VOF model, the primary and secondary phases are immiscible, so in most computational cells, each phase has a volume fraction of either 0 or 1. At the interface between two phases (air and surfactant or water in this case), the volume fraction is between 0 and 1, and the interface can be tracked by solving the volume fraction equation.

- Volume fraction equation for the secondary phases:

  Interface tracking is accomplished by solving a continuity equation based on the volume fraction of the secondary phase. This equation has the following form:

$$\frac{1}{\rho_q}\left[\frac{\partial}{\partial t}(\alpha_q\rho_q) + \nabla\cdot(\alpha_q\rho_q\vec{\boldsymbol{u}}_{\boldsymbol{q}}) = S_{\alpha_q} + \sum_{p=1}^{n}(\dot{m}_{pq} - \dot{m}_{qp})\right] \quad (3)$$

where $\dot{m}_{qp}$ is the mass transfer from phase q to phase p and $\dot{m}_{qp}$ is the mass transfer from phase p to phase q. By default, the source term on the right-hand side of the equation, S_{α_q}, is zero unless specified otherwise.

- The Implicit Scheme:

Implicit schemes compute volume fraction values at the current time step, in contrast to explicit methods that require the volume fraction from the previous time step. By solving a scalar transport equation at each time step, the secondary phase volume fraction is iteratively calculated.

$$\frac{\alpha_q^{n+1}\rho_q^{n+1} - \alpha_q^n \rho_q^n}{\Delta t} V + \sum_f \left(\rho_q^{n+1} U_f^{n+1} \alpha_{q,f}^{n+1}\right) = \left[S_{\alpha_q} + \sum_{p=1}^{n} \left(\dot{m}_{pq} - \dot{m}_{qp}\right)\right] V \quad \textbf{(4)}$$

Where $n+1$ is an index for the current time step, n is an index for the previous time step, $\alpha_{q,f}^n$ is the face value of the q^{th} volume fraction, computed from the second-order upwind, modified HRIC scheme, V is the volume of a cell, Δt is the time step size, and U_f^n is the volume flux through the face based on normal velocity.

2.2.2 Particle transport

From Newton's second law, the particle's inertia equates to the summation of forces acting on the particle. Therefore, to predict the trajectory of a discrete phase particle, the following Lagrangian conservation of momentum equation is solved:

$$\frac{\partial \vec{u}_p}{\partial t} = F_D\left(\vec{u} - \vec{u}_p\right) + \frac{\vec{g}(\rho_p - \rho)}{\rho_p} \quad \textbf{(5)}$$

Where $F_D\left(\vec{u} - \vec{u}_p\right)$ is the drag force per unit particle mass with $F_D = \frac{18\mu}{d_p^2 \rho_p} \times \frac{C_D Re}{24}$,where $Re = \frac{\rho d_{p|\vec{u}_p - \vec{u}|}}{\mu}$ is the relative Reynolds number, d_p is the particle diameter, C_D is the drag coefficient, $\rho_p = 1550\, kg/m^3$ is the particle density, $\rho = 1.111\, \mathrm{kg}/m^3$ is the air density, and $\mu = 1.927 \times 10^{-5}\, \frac{\mathrm{kg}}{\mathrm{m.s}}$ is the molecular viscosity of the fluid.

In addition, Stokes' drag law is applied to spherical objects with very small Reynolds numbers in a viscous fluid. The Stokes' drag force assumes a no-slip condition which is incorrect for micron-sized particles ($< 15\ \mu m$). Therefore, the drag coefficient is divided by the Cunningham correction factor (Cunningham 1910):

$$F_D = \frac{18\mu}{d_p^2\,\rho_p C_c} \tag{6}$$

Where C_c is the Cunningham correction to Stokes' drag law, calculated through the following equation:

$$C_c = 1 + \frac{2\lambda}{d_p}\left(1.257 + 0.4e^{-\left(\frac{1.1d_p}{2\lambda}\right)}\right) \tag{7}$$

$\lambda = 0.067\ \mu m$ is the molecular mean free path of air.

2.3 Machine Learning Method

Machine learning classifiers identify and assign data to different categories based on the specific features/characteristics analyzed from a training data set. The well-known numeric computing program "MATLAB" provides different classification algorithms in the MALTAB Classification Learner Toolbox: "Decision Trees", Discriminant Analysis", "Naïve Bayes", Support Vector Machines", "K-Nearest Neighbour (KNN)", and 'Ensemble Classifiers". Two independent input variables, the striking particle velocity magnitude and time for impact, are used to classify particles into three different diameters (output). The objective is to find the optimum diameter matching a particular impaction time since its release into the acinus and corresponding to a specific particle velocity required for enhanced particle-surfactant interaction. This process is achieved for a healthy acinus with low surface tension (case 1) and a diseased acinus with high surface tension

(case 2). For both cases, all available machine learning classifiers in MATLAB R2020 analyzed and classified the training data containing the inputs (striking particle velocity magnitude and time for impact) and output (diameter) extracted from ANSYS 2020 R2. Four classifiers: "Fine KNN", "Weighted KNN", "Subspace KNN", and "Bagged Trees" predicted 100% accuracy. The Ensemble Classifier, "Ensemble Subspace KNN", and the Nearest neighbour Classifier, "Weighted KNN", were chosen to test additional input data for case 1 and case 2, respectively.

KNN is a supervised instance-based learning approach that offers a solution for problems of unknown distributions (Zhang 2016). This instance-based learning method compares each new instance with existing ones using a distance metric function and assigns a class based on the closest instance. The KNN function employs Euclidean distance through the following equation (Weinberger and Saul 2009):

$$D(p,q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \cdots + (p_n - q_n)^2}$$

Where p and q objects are to be classified according to n characteristics. Objects are classified based on a majority vote among their k nearest neighbours. For example, if k = 1 ("Fine KNN"), then the object is assigned to the class of a single nearest neighbour. If k = 10 ("Weighted KNN"), then the object is assigned to the majority class of the closest ten nearest neighbours. Choosing the appropriate value of k is essential for reducing random error and not ignoring small but significant patterns simultaneously (Zhang 2014). In addition, the "Subspace KNN" model creates an ensemble of several nearest-neighbor classifiers to improve the classification performance (Tian and Feng 2021).

3. Results and Discussion:

3.1 Example CFD simulation

Figure 2 displays the trajectories of 0.5 μm particles at different instances during the inhalation phase of a healthy acinus with LST. At 0.05 s, 132 particles were injected into the inlet of the 18th generation with the realistic inlet velocity profile shown previously in Figure 1c. Due to the very low Reynolds number in an acinus, particles act as tracer particles, following the airflow streamlines with Stokes numbers around 0.15 (Chhabra and Prasad, 2011). Figure 2a demonstrates the particle trajectories at 0.6 s, 1 s, and 1.4 s as they travel into deeper generations by rapid diffusion/sedimentation (Dailey and Ghadiali 2007), with different velocity magnitudes ranging from 0 to 7 mm/s. Particles are assumed to have an initial zero particle velocity magnitude, indicating that a particle's velocity is solely affected by the flow velocity. The particle residence time shown in Figure 2b shows the time a particle enters the acinus to its current time, where it is either trapped on the surface or still advancing to further generations. At the end of inhalation at 1.75 s, the time required for a particle to strike the alveolar walls and the velocities of the trapped particles at the time of impact are extracted from the simulation and used as input to the machine learning classifiers. The same procedure is repeated separately for the 1.5 μm and 5 μm particles in the LST case and for all three-particle sizes in the diseased case with HST.

Accepted to Phys. Fluids 10.1063/5.0106594

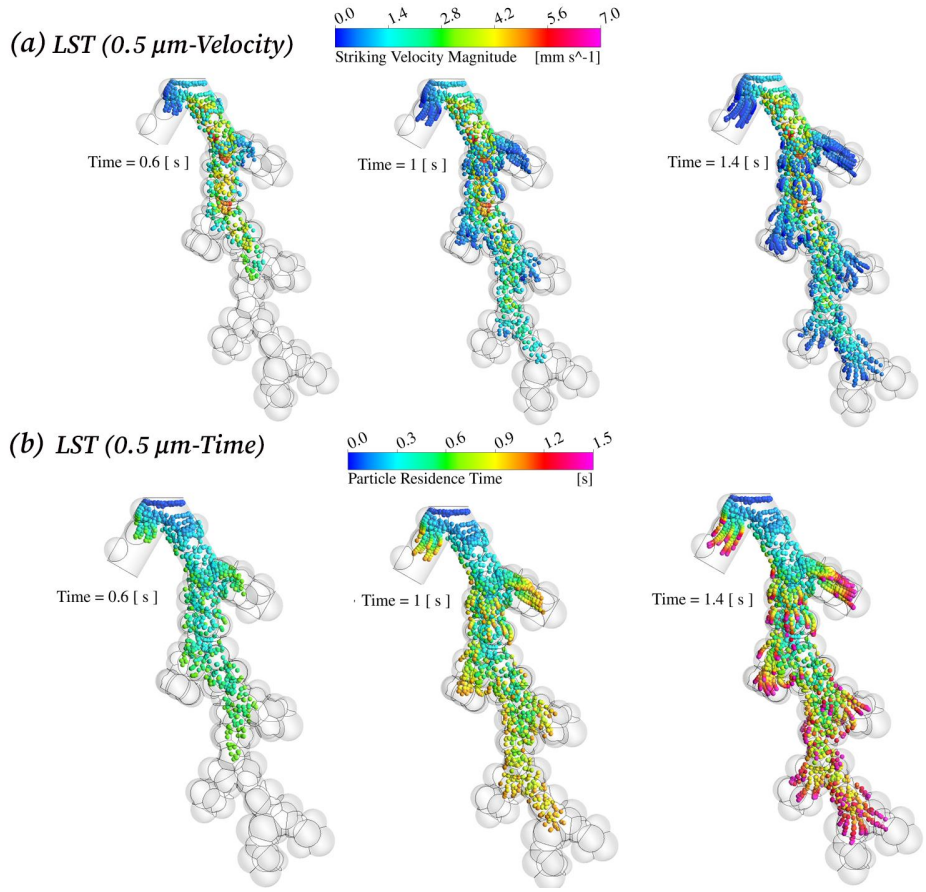

Figure 2: Three instantaneous snapshots of particle trajectories inside the acinus model in the LST case coloured according to (a) Particle Striking Velocity Magnitude (mm/s) and (b) Particle Residence Time (s)

3.2 Machine Learning Results:

3.2.1 Case 1:

In case 1, two Nearest Neighbours Classifiers, "Fine and Weighted k-nearest neighbour (KNN)", and two Ensemble Classifiers, "Subspace KNN and Bagged Trees", predicted 100% accurate

results for 274 training data. The "Ensemble Subspace KNN" model with 30 learners and one subspace dimension is selected to assess an additional group of 37 input and output data. Results showed a prediction accuracy of 94.6%,with only two misclassifications. The training data, testing data, receiver operating characteristic curve (ROC), and confusion matrix are shown in Figures 3a, 3b, 3c, and 3d, respectively.

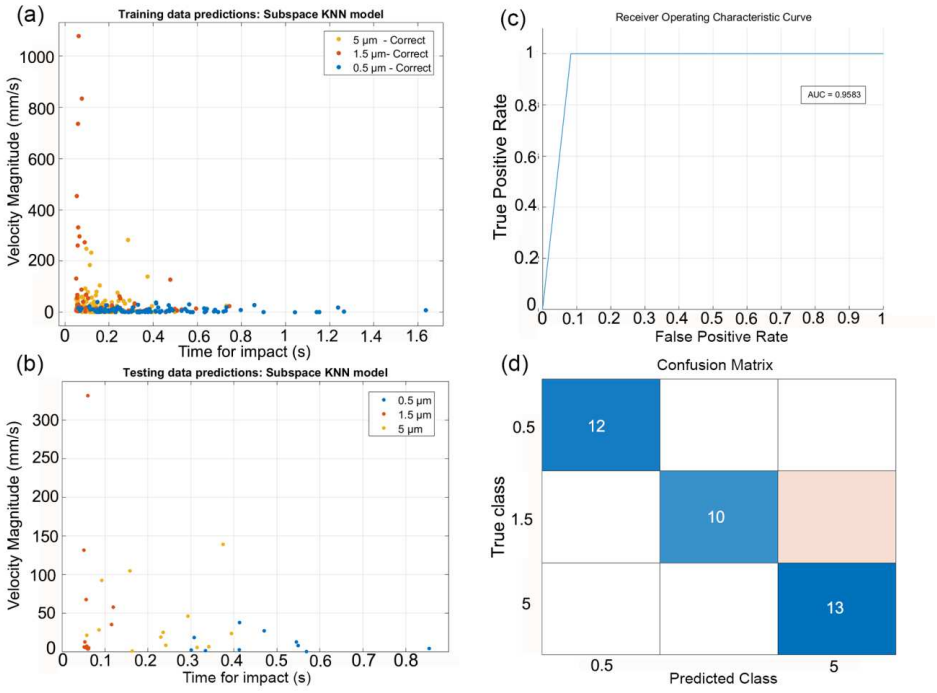

Figure 3: Case 1 healthy acinus with LST: (a) Training data predictions using Subspace KNN model for the three-particle diameters according to the velocity magnitude (mm/s) and impact time (s), (b) Testing data predictions, (c) ROC curve measuring the classification's success and perfection, (d) Confusion matrix displaying the true class versus the predicted class for each of the three-particle sizes

Particles inside the alveoli deposit either by diffusion due to a concentration gradient, advection due to the movement of air, or sedimentation due to gravitational forces (Dong et al., 2021). The combination of drag, gravitational, and diffusion forces imposed on a particle affects its striking velocity magnitude and impact time. Smaller-sized particles (0.5 μm) deposit mainly by diffusion and advection forces, larger particles deposit mainly by sedimentation (5 μm), and middle-sized particles experience a combination of both. Deposition by diffusion of small-sized particles occurs at small striking velocity magnitudes, sedimentation of large-sized particles happens with larger striking velocity magnitudes, and the combination of both with middle-sized particles produces the largest velocity magnitudes. This is shown in Figure 3a, with the striking velocity magnitude on the *y*-axis and the impact time on the *x*-axis for the three different particle sizes. An adopted "Subspace KNN" model correctly classified each particle according to these two variables. The particles of the smallest diameter, 0.5 μm, exhibit the lowest striking velocity magnitudes (average value is 4.6 mm/s) and the longest impact time from 0.05 s to 1.63s. The slightly larger particles of 1.5 μm possess much higher velocity magnitudes with an average value of 100 mm/s and much lower impact time with values ranging from 0.05s to 0.75s (average value = 0.196 mm/s). The largest 5 μm particles exhibit impact times similar to the 1.5 μm particles, with a slight decrease in the average value (0.161 mm/s), and the average of their velocity magnitudes lies between the 0.5 μm and 1.5 μm particle sizes with a value of 31.7 mm/s. The impact time of the three-particle sizes agrees with the literature findings that the smaller the particle size, the longer it takes to deposit on the surface of the lung (Heyder 2004). However, the pattern for the velocity magnitude is not consistent with the particle diameter, as the middle-sized particles (1.5 μm) experience more significant values than the largest-sized particles (5 μm). This indicates that these three-particle diameters cannot determine the type of correlation (positive or negative) between the particle

diameter and the striking velocity magnitude. Therefore, a critical/threshold diameter defining their relationship must be identified in future studies by simulating a larger number of particle diameters. The testing data in Figure 3b demonstrates similar results for the three-particle diameters. However, the "Subspace KNN model" fails to perfectly predict the diameter of every particle of the testing data and achieves an accuracy of 94.6%. Figure 3c depicts a receiver operating characteristic curve (ROC curve) utilized to evaluate the model's classification performance. The area under the ROC curve is an indicator of the model's success, where a perfect classification obtains an area of 1 with a zero false-positive rate (FPR) (x-axis) and a true positive rate (TPR) equal to 1 (y-axis). The studied case exhibits an area under the curve (AUC) of 0.9583 ($\mathrm{AUC} \approx 1$), where the optimal operating point is $\mathrm{FPR} = 0.0833$ and $\mathrm{TPR} = 1$. This slight imperfection is further illustrated in Figure 3d, where a confusion matrix displaying the true class and the predicted class for each particle size reveals a misclassification of two 1.5 μm particles by falsely predicting them as 5 μm particles.

3.2.2 Case 2:

In case 2, the same four classification models of case 1 predicted 100% accurate results for 201 training data. The "Weighted KNN" model is selected to test an additional set of 38 input and output data with the following specifications: Ten Nearest Neighbours, Euclidean distance metric, and squared inverse distance weighting for interpolating spatial sample data. The outcome is better than case 1 as the adopted model exhibits an accuracy of 100% with no misclassifications. The training data and testing data, ROC curve, and confusion matrix are shown in Figures 4a, 4b, 4c, and 4d, respectively. Figure 4a displays the same trend as case 1 for the three-particle sizes. For the striking velocity magnitude, the 5 μm particles have values between those of the other two particle sizes. For the impact time, the smaller the particle, the more time it requires for impact

($t_{0.5\,\mu m} > t_{1.5\,\mu m} > t_{5\,\mu m}$). Figure 4b shows the testing data predictions for each particle size group. A clear visual distinction between them permits 100% accurate classification by the "Weighted KNN" model where the ROC curve (Figure 3c) displays this perfect clustering with FPR = 0, TPR = 1, and AUC = 1 and the confusion matrix (Figure 4d) fits all data inside the diagonal showing no misclassifications.

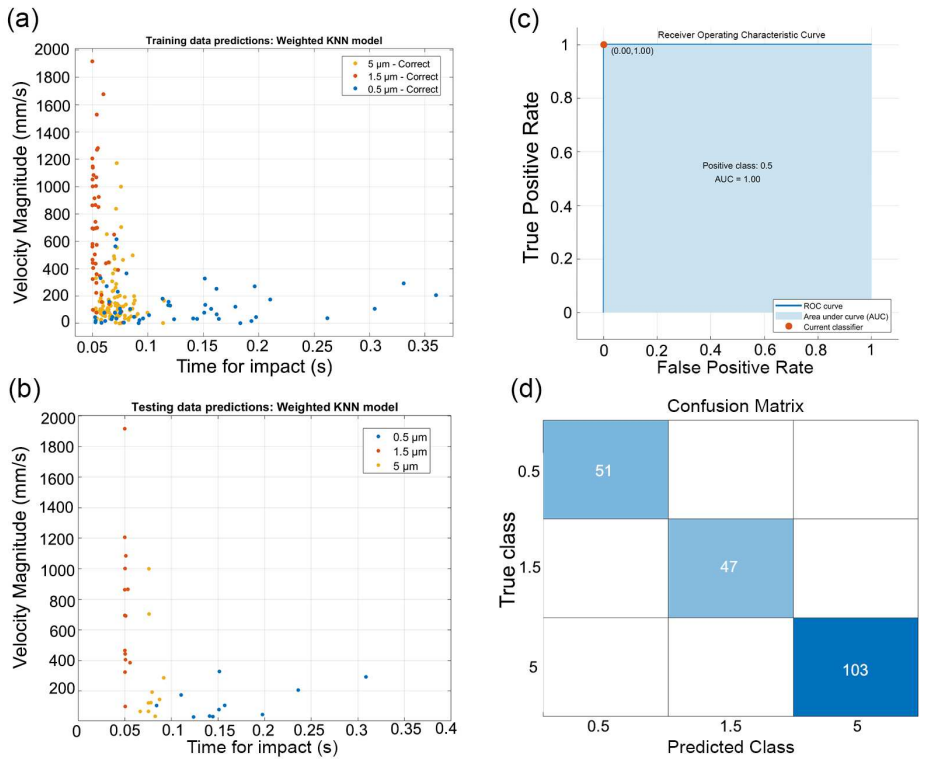

Figure 4: Case 2 diseased acinus with HST: (a) Training data predictions using Weighted KNN model (b) Testing data predictions, (c) ROC curve (d) Confusion matrix

Accepted to Phys. Fluids 10.1063/5.0106594

3.3 Comparison between the LST case and HST case:

The average striking velocity magnitude and average impact time of each particle size group ($0.5\ \mu m, 1.5\ \mu m, 5\ \mu m$) for the LST and HST cases are given in Table 1.

Table I: Comparison between case 1 (LST) and case 2 (HST) based on the average striking velocities and impact times of three-particle sizes.

	LST			**HST**		
	$0.5\ \mu m$	$1.5\ \mu m$	$5\ \mu m$	$0.5\ \mu m$	$1.5\ \mu m$	$5\ \mu m$
Average Striking Velocity Magnitude (mm/s)	4.60	100	31.7	126	692	176
Variance	0.183	0.000900	0.534	0.341	0.408	0.0042
Average Impact Time (s)	0.400	0.196	0.161	0.120	0.0750	0.0700
Variance	0.00290	0.0192	0.0000	0.0326	0.00100	0.00560

It was shown in our previous study (arXiv:2204.01699 [physics.bio-ph]) that a higher surface tension value produces more vortices (curl of the velocity), and larger airflow velocity magnitudes in the acinus. As the kinematic properties of the flow affect that of the particles, the flow's higher momentum and unsteady recirculating streamlines, due to the high surface tension, lead to larger particle velocity magnitudes and faster striking times. In other words, the drag forces acting on the particle increase, and thus the particle's inertia and striking velocity also increase. In addition, as the drag force is inversely proportional to the square of the particle's size, the smallest particle

diameter experiences the largest inertia as the surface tension increases and the largest particle diameter experiences the smallest inertia. Also, the impact time decreases with increased inertia as the particle reaches the alveolar surfaces faster with higher velocity. Table I demonstrates that the HST case has larger average particle velocity magnitudes and smaller average impact times than the LST case for each particle size group. From the LST to the HST case, the average striking velocity magnitudes increase by a factor of 27.45 for the 0.5 μm particles, by a factor of 6.92 for the 1.5 μm particles, and by a factor of 5.57 for the 5 μm particles. Meanwhile, the average impact times decrease by a factor of 3.33 for the 0.5 μm particles, by a factor of 2.61 for the 1.5 μm particles, and by a factor of 2.3 for the 5 μm particles.

The variance S^2, the deviation from the average/mean, was obtained from MATLAB R2020 for each set of data. The variance is calculated through the following formula: $S^2 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}$ where x_i is each value in the dataset, $\bar{x}$ is the mean of all values in the dataset, and n is the number of values in the dataset. For the average striking velocity magnitude and the impact time, the variance is significantly less than unity, meaning that the data points are close to the mean. Therefore, the averaged data can be used for interpreting the results.

Figures 5a and 5b present two "Parallel Coordinates Plots" for the LST and HST cases, respectively. The striking velocity magnitudes and impact times for each particle size group are arranged on vertical parallel axes to demonstrate the correlation between these two independent properties. The two figures are shown with diagonal lines that indicate an inversely proportional relationship between velocity magnitude and impact time. Figure 5a shows that 0.5 μm particles have the highest impact times (reaching 1.6 s) and the lowest velocity magnitudes and that 1.5 μm particles have the shortest impact times and the highest velocity magnitudes (around

1100 mm/s). This trend is also evident in Figure 5b, which illustrates a more apparent distinction between the particle sizes. Figures 5c and 5d display two boxplots comparing each particle size group from the HST case to that from the LST case according to their velocity magnitudes and times for impact, respectively. Similar to previous results, the two figures show that the HST values for the three-particle size groups have larger velocity magnitude values (higher medians in Figure 5c) and smaller impact times (lower medians in Figure 5d) compared to the LST case.

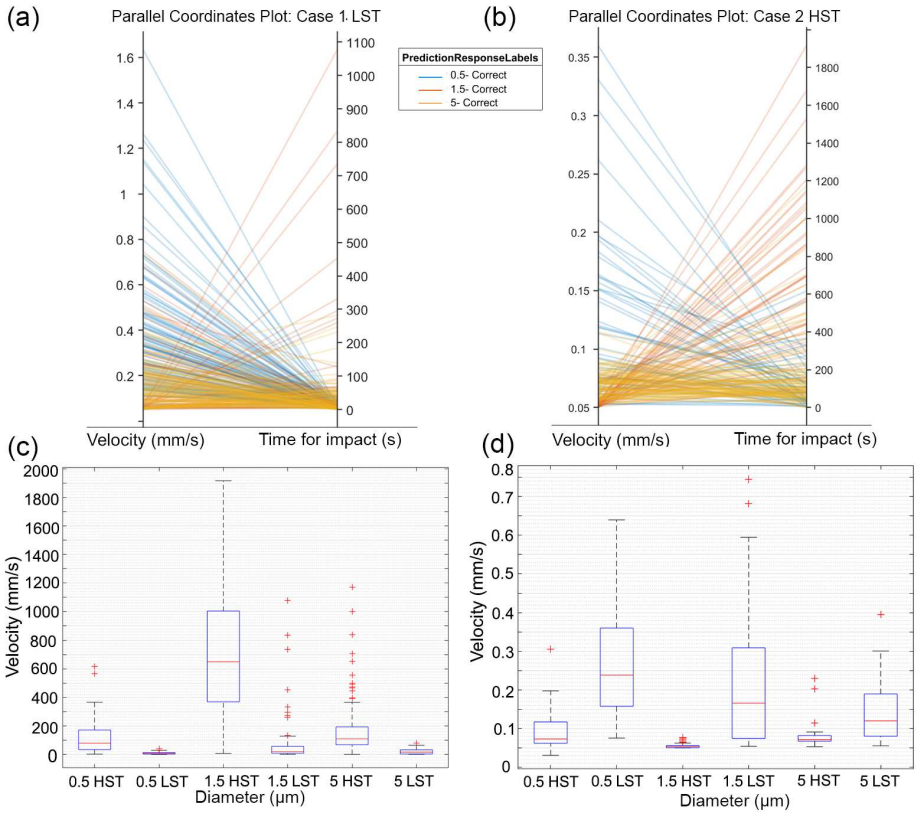

Figure 5: (a) Parallel Coordinates Plot for the LST case and (b) HST case, (c) Boxplot comparing each particle size group between the LST and HST cases according to the velocity magnitudes (mm/s) and (d) the time for impact (s)

3.4 Benefit of using ML:

The computational time used to classify particle diameters in MATLAB machine learning was 283.9s for the first LST case and 381.7s for the second HST case. On the other hand, the computational time for particle deposition in ANSYS FLUENT (using LINUX system with 2 CPU socks and 32 computing nodes) was 57.36 hours for the first case and 62.11 hours for the second case. The very high discrepancy between machine learning and ANSYS FLUENT computational times proves the benefit of combining CFD and MLA to save computational time. Therefore, extracting sample results from CFD and training the data with MLA to perform classifications/predictions/optimization is more efficient than performing simulations in CFD alone.

4. Conclusion:

A comprehensive CFD study was performed to determine the striking velocity magnitudes and impact times of three different sized particles ($0.5\ \mu m, 1.5\ \mu m,$ and $5\ \mu m$) in a healthy LST acinus and a diseased HST acinus.

Each particle size group exhibited diverse values, enabling machine learning models "Subspace KNN" and "Weighted KNN" to classify the particles according to two independent variables. The findings are summarized as follows:

- For the LST case, the "Subspace KNN" obtained a 94.6% accuracy with $0.5\ \mu m$ particles having the lowest velocities and the longest impact times. The $1.5\ \mu m$ particles experience

the highest striking velocities and the lowest impact times. The 5 μm particles have values in between.

- For the HST case, the particles show the same trend as in the LST case. However, a comparison between the two cases showed that particles have higher striking velocities and lower impact times in the HST case due to the erratic and irregular flow with vortices propelling the particles toward the alveolar walls.

The findings of this study are crucial to molecular dynamics computations studying inhaled drugs and their interaction with the surfactant layer/alveolar surfaces. The striking velocity magnitude for three micron-sized particles ($0.5\ \mu m$, $1.5\ \mu m$, and $5\ \mu m$) can be taken and inserted as an input affecting the particle/alveolar surface interaction. In pharmaceutical and medical studies, if the case (healthy or diseased lung) requires a certain striking velocity magnitude and impact time, then the values can be inserted into our testing data, and the optimal particle diameter is determined. Few limitations and assumptions include an idealised acinar model with rigid walls, zero initial particle velocity entering the 18^{th} generation of the acinus, one particle density, and two extreme surface tension cases that are not physiologically accurate. In addition, one major limitation of this study is generating particle velocity and impact time using only an acinus part of the lung. Drug delivery systems require particles traveling from the nose/mouth to the alveoli. The two mentioned particle properties would be different in this case. However, this study only indicates the applicability of such a method, and future studies will include the whole lung geometry. In particular, the initial particle velocity magnitude entering the acinus (generation 18) will be computed by conducting flow and particle simulations using a realistic lung geometry of 17 generations with a realistic inhalation breathing profile. A particle's velocity striking the outlet of the 17^{th} generation will serve as the initial velocity of the particle entering the acinar model.

Therefore, a particle's velocity will be affected by both its initial velocity and the flow's velocity. The impact time will also be realistically calculated for the whole lung. Additional studies will analyze other micro- and nanoparticle diameters with moving acinar walls (fluid-structure interaction) and physiologically-accurate surface tension values. The aim is to relate a wide range of particle diameters to specific striking velocity magnitudes and impact times and identify critical particle diameters that define the correlation between the striking velocity and particle diameter.

Declaration of competing interest

The authors declare no conflict of interest.

Acknowledgment

The authors would like to acknowledge the computing facility at the University of Technology Sydney (UTS).

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

April Si, X., Talaat, M. and Xi, J. (2021)."SARS COV-2 virus-laden droplets coughed from deep lungs: Numerical quantification in a single-path whole respiratory tract geometry." Phys. Fluids **33**(2), 023306 (2021)

Ault, J.T., Fani, A., Chen, K.K., Shin, S., Gallaire, F. and Stone, H.A. (2016)."Vortex-breakdown-induced particle capture in branching junctions." Phys. Rev. Lett. **117**(8), 084501 (2016)

Balásházy, I., Hofmann, W., Farkas, Á. and Madas, B.G. (2008)."Three-dimensional model for aerosol transport and deposition in expanding and contracting alveoli." Inhalation toxicology **20**(6), 611-621 (2008)

Brackbill, J.U., Kothe, D.B. and Zemach, C. (1992)."A continuum method for modeling surface tension." J. Comput. Phys. **100**(2), 335-354 (1992)

Chhabra, S. and Prasad, A.K. (2011)."Flow and particle dispersion in lung acini: effect of geometric and dynamic parameters during synchronous ventilation." J. Fluids Eng. **133**(7)(2011)

Chroneos, Z., Sever-Chroneos, Z. and Shepherd, V. (2010)."Pulmonary surfactant: an immunological perspective." Cellular physiology and biochemistry **25**(1), 13-26 (2010)

Ciloglu, D. (2020)."A numerical study of the aerosol behavior in intra-acinar region of a human lung." Phys. Fluids **32**(10), 103305 (2020)

Cunningham, E. (1910)."On the velocity of steady fall of spherical particles through fluid medium." Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character **83**(563), 357-365 (1910)

Dailey, H.L. and Ghadiali, S. (2007)."Fluid-structure analysis of microparticle transport in deformable pulmonary alveoli." J. Aerosol Sci **38**(3), 269-288 (2007)

Deng, Q., Deng, L., Miao, Y., Guo, X. and Li, Y. (2019)."Particle deposition in the human lung: Health implications of particulate matter from different sources." Environ. Res. **169**, 237-245 (2019)

Dong, J., Qiu, Y., Lv, H., Yang, Y. and Zhu, Y. (2021)."Investigation on Microparticle Transport and Deposition Mechanics in Rhythmically Expanding Alveolar Chip." Micromachines **12**(2), 184 (2021)

Douafer, H., Andrieu, V. and Brunel, J.M. (2020)."Scope and limitations on aerosol drug delivery for the treatment of infectious respiratory diseases." Journal of Controlled Release **325**, 276-292 (2020)

Francis, I., Shrestha, J., Paudel, K.R., Hansbro, P.M., Warkiani, M.E. and Saha, S.C. (2022)."Recent advances in lung-on-a-chip models." Drug Discovery Today(2022)

Haefeli-Bleuer, B. and Weibel, E.R. (1988)."Morphometry of the human pulmonary acinus." The Anatomical Record **220**(4), 401-414 (1988)

Hanna, B.N., Dinh, N.T., Youngblood, R.W. and Bolotnov, I.A. (2020)."Machine-learning based error prediction approach for coarse-grid Computational Fluid Dynamics (CG-CFD)." Prog. Nuclear Energy **118**, 103140 (2020)

Harding Jr, E.M. and Robinson, R.J. (2010)."Flow in a terminal alveolar sac model with expanding walls using computational fluid dynamics." Inhalation Toxicology **22**(8), 669-678 (2010)

Hedberg, J., Lundin, M., Lowe, T., Blomberg, E., Wold, S. and Wallinder, I.O. (2012)."Interactions between surfactants and silver nanoparticles of varying charge." Journal of colloid and interface science **369**(1), 193-201 (2012)

Heyder, J. (2004)."Deposition of inhaled particles in the human respiratory tract and consequences for regional targeting in respiratory drug delivery." Proc. Am. Thorac. Soc. **1**(4), 315-320 (2004)

Hu, G., Jiao, B., Shi, X., Valle, R.P., Fan, Q. and Zuo, Y.Y. (2013)."Physicochemical properties of nanoparticles regulate translocation across pulmonary surfactant monolayer and formation of lipoprotein corona." ACS nano **7**(12), 10525-10533 (2013)

Khajeh-Hosseini-Dalasm, N. and Longest, P.W. (2015)."Deposition of particles in the alveolar airways: inhalation and breath-hold with pharmaceutical aerosols." J. Aerosol Sci **79**, 15-30 (2015)

Koullapis, P., Stylianou, F., Sznitman, J., Olsson, B. and Kassinos, S. (2020)."Towards whole-lung simulations of aerosol deposition: A model of the deep lung." J. Aerosol Sci **144**, 105541 (2020)

Ma, B. and Lutchen, K.R. (2009)."CFD simulation of aerosol deposition in an anatomically based human large–medium airway model." Ann. Biomed. Eng. **37**(2), 271-285 (2009)

Mohammadpour, J., Salehi, F., Sheikholeslami, M., Masoudi, M. and Lee, A. (2021)."Optimization of nanofluid heat transfer in a microchannel heat sink with multiple synthetic jets based on CFD-DPM and MLA." International Journal of Thermal Sciences **167**, 107008 (2021)

Mollicone, J.-P., Sharifi, M., Battista, F., Gualtieri, P. and Casciola, C.M. (2019)."Particles in turbulent separated flow over a bump: Effect of the Stokes number and lift force." Phys. Fluids **31**(10), 103305 (2019)

Muhlfeld, C., Rothen-Rutishauser, B., Blank, F., Vanhecke, D., Ochs, M. and Gehr, P. (2008)."Interactions of nanoparticles with pulmonary structures and cellular responses." American Journal of Physiology-Lung Cellular and Molecular Physiology **294**(5), L817-L829 (2008)

Rahman, M.M., Zhao, M., Islam, M.S., Dong, K. and Saha, S.C. (2021)."Aging effects on airflow distribution and micron-particle transport and deposition in a human lung using CFD-DPM approach." Adv. Powder Technol. **32**(10), 3506-3516 (2021)

Raibaudo, C., Zhong, P., Noack, B.R. and Martinuzzi, R.J. (2020)."Machine learning strategies applied to the control of a fluidic pinball." Phys. Fluids **32**(1), 015108 (2020)

Ringstad, K.E., Banasiak, K., Ervik, Å. and Hafner, A. (2021)."Machine learning and CFD for mapping and optimization of CO2 ejectors." Appl. Therm. Eng. **199**, 117604 (2021)

Russo, J.S. and Khalifa, E. (2011)."Computational study of breathing methods for inhalation exposure." HVAC&R Research **17**(4), 419-431 (2011)

Rüttgers, M., Waldmann, M., Schröder, W. and Lintermann, A. (2022)."A machine-learning-based method for automatizing lattice-Boltzmann simulations of respiratory flows." Appl Intell, 1-21 (2022)

Sachan, A.K., Harishchandra, R.K., Bantz, C., Maskos, M., Reichelt, R. and Galla, H.-J. (2012)."High-resolution investigation of nanoparticle interaction with a model pulmonary surfactant monolayer." Acs Nano **6**(2), 1677-1687 (2012)

Saidi, M.S., Rismanian, M., Monjezi, M., Zendehbad, M. and Fatehiboroujeni, S. (2014)."Comparison between Lagrangian and Eulerian approaches in predicting motion of micron-sized particles in laminar flows." Atmos. Environ. **89**, 199-206 (2014)

Schmidt, D., Maulik, R. and Lyras, K. (2021)."Machine learning accelerated turbulence modeling of transient flashing jets." Phys. Fluids **33**(12), 127104 (2021)

Soni, B. and Aliabadi, S. (2013)."Large-scale CFD simulations of airflow and particle deposition in lung airway." Computers & Fluids **88**, 804-812 (2013)

Sonwani, S., Madaan, S., Arora, J., Rangra, D., Mongia, N., Vats, T. and Saxena, P. (2021)."Inhalation Exposure to Atmospheric Nanoparticles and its Associated Impacts on Human Health: A Review." Frontiers in Sustainable Cities, 87 (2021)

Sznitman, J. (2013)."Respiratory microflows in the pulmonary acinus." J. Biomech. **46**(2), 284-298 (2013)

Sznitman, J., Heimsch, F., Heimsch, T., Rusch, D. and Rösgen, T. (2007)."Three-dimensional convective alveolar flow induced by rhythmic breathing motion of the pulmonary acinus." J. Biomech. Eng. **129**(5), 658-665 (2007)

Tian, G., Hindle, M., Lee, S. and Longest, P.W. (2015)."Validating CFD predictions of pharmaceutical aerosol deposition with in vivo data." Pharm. Res. **32**(10), 3170-3187 (2015)

Tian, Y. and Feng, Y. (2021)."RaSE: Random Subspace Ensemble Classification." J. Mach. Learn. Res. **22**(45), 1-93 (2021)

Uçar, M.K., Bozkurt, M.R., Bilgin, C. and Polat, K. (2017)."Automatic detection of respiratory arrests in OSA patients using PPG and machine learning techniques." Neural Computing and Applications **28**(10), 2931-2945 (2017)

Veldhuizen, R., McCaig, L.A., Akino, T. and Lewis, J.F. (1995)."Pulmonary surfactant subfractions in patients with the acute respiratory distress syndrome." American journal of respiratory and critical care medicine **152**(6), 1867-1871 (1995)

Vinchurkar, S., De Backer, L., Vos, W., Van Holsbeke, C., De Backer, J. and De Backer, W. (2012)."A case series on lung deposition analysis of inhaled medication using functional imaging based computational fluid dynamics in asthmatic patients: effect of upper airway morphology and comparison with in vivo data." Inhalation toxicology **24**(2), 81-88 (2012)

Walters, D.K. and Luke, W.H. (2011)."Computational fluid dynamics simulations of particle deposition in large-scale, multigenerational lung models." J. Biomech. Eng. **133**(1)(2011)

Weinberger, K.Q. and Saul, L.K. (2009)."Distance metric learning for large margin nearest neighbor classification." Journal of machine learning research **10**(2)(2009)

Xi, J., Si, X.A. and Nagarajan, R. (2020)."Effects of mask-wearing on the inhalability and deposition of airborne SARS-CoV-2 aerosols in human upper airway." Phys. Fluids **32**(12), 123312 (2020)

Xi, J., Talaat, M., Si, X.A., Han, P., Dong, H. and Zheng, S. (2020)."Alveolar size effects on nanoparticle deposition in rhythmically expanding-contracting terminal alveolar models." Computers in biology and medicine **121**, 103791 (2020)

Zhang, Z. (2014)."Too much covariates in a multivariable model may cause the problem of overfitting." J. Thorac. Dis. **6**(9), E196 (2014)

Zhang, Z. (2016)."Introduction to machine learning: k-nearest neighbors." Annals of translational medicine **4**(11)(2016)

Accepted to Phys. Fluids 10.1063/5.0106594

Zhang, Z., Kleinstreuer, C. and Kim, C.S. (2009)."Comparison of analytical and CFD models with regard to micron particle deposition in a human 16-generation tracheobronchial airway model." J. Aerosol Sci **40**(1), 16-28 (2009)

Zhao, J., Feng, Y., Koshiyama, K. and Wu, H. (2021)."Prediction of airway deformation effect on pulmonary air-particle dynamics: A numerical study." Phys. Fluids **33**(10), 101906 (2021)

Zhu, L., Zhang, W., Kou, J. and Liu, Y. (2019)."Machine learning methods for turbulence modeling in subsonic flows around airfoils." Phys. Fluids **31**(1), 015105 (2019)

Accepted to Phys. Fluids 10.1063/5.0106594

Appendix

Nomenclature

CFD	computational fluid dynamics
MLA	machine learning algorithms
LST	low surface tension
HST	high surface tension
VOF	volume of fluid
DPM	discrete phase model
KNN	k-nearest neighbour
ROC	receiver operating characteristic curve
AUC	area under the curve
TPR	true positive rate
FPR	false positive rate

Symbols

ρ	density (kg/m^3)
μ	dynamic viscosity (kg/m.s)
t	time (s)
d_p	particle diameter (μm)

Accepted to Phys. Fluids 10.1063/5.0106594

$\dot{m}$	mass transfer (kg)
$\vec{T}$	surface tension force (N)
F_D	drag force (N)
D	Euclidean distance (m)
α	volume fraction
$\vec{g}$	gravitational force (N)
S	source term
V	volume of cell (m^3)
Δt	time step size (s)
Re	Reynolds number
C_D	drag coefficient
C_c	Cunningham correction

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0106594