Evaluation of multi-output machine learning models for predicting inhaled particle deposition in the human upper and central airway

Li, Xueren; Xu, Ruipeng; Fan, Jiaqi; Zhang, Liwei; Sun, Weijie; Kenjeres, Sasa; Shang, Yidan; Yang, William

DOI
10.1016/j.powtec.2025.120924

Publication date
2025

Document Version
Final published version

Published in
Powder Technology

Citation (APA)
Li, X., Xu, R., Fan, J., Zhang, L., Sun, W., Kenjeres, S., Shang, Y., & Yang, W. (2025). Evaluation of multi-output machine learning models for predicting inhaled particle deposition in the human upper and central airway. *Powder Technology*, *458*, Article 120924. https://doi.org/10.1016/j.powtec.2025.120924

Powder Technology 458 (2025) 120924

Contents lists available at ScienceDirect

Powder Technology

journal homepage: www.journals.elsevier.com/powder-technology

Evaluation of multi-output machine learning models for predicting inhaled particle deposition in the human upper and central airway

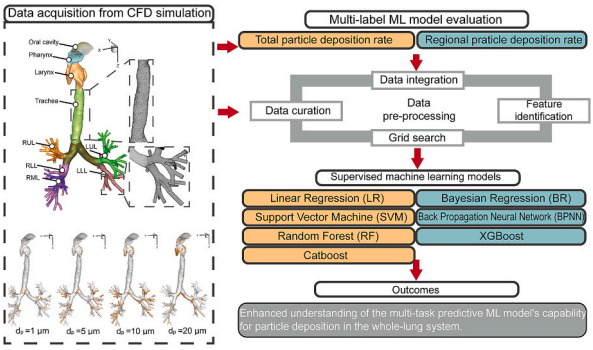
Xueren Li [a,b,1], Ruipeng Xu [c,1], Jiaqi Fan [d], Liwei Zhang [e], Weijie Sun [f], Sasa Kenjeres [c], Yidan Shang [a,g,*], William Yang [h,**]

[a] School of Mechanical and Automotive Engineering, Shanghai University of Engineering Science, 201620, China
[b] School of Engineering, RMIT University, PO Box 71, Bundoora, VIC 3083, Australia
[c] Department of Chemical Engineering, Faculty of Applied Sciences, Delft University of Technology and J.M. Burgerscentrum Research School for Fluid Mechanics, Van der Maasweg 9, 2629 HZ Delft, the Netherlands
[d] School of Safety Engineering, China University of Mining and Technology, Xuzhou 221116, China
[e] School of Low-Carbon Energy and Power Engineering, China University of Mining and Technology, Xuzhou 221116, China
[f] Department of Computing Science, University of Alberta, Edmonton, AB T6G 2R3, Canada
[g] Shanghai Xuhui Central Hospital, Fudan University, 200231, China
[h] Advanced Laser Flow Diagnostic Laboratory, CSIRO Mineral Resources, Clayton South, VIC 3169, Australia

HIGHLIGHTS

- Total and regional particle deposition rates are analyzed in a whole lung model.
- Performance of 7 multi-label ML models for predicting inhaled deposition rate is evaluated.
- LR, BR, and SVM fail to capture non-linear deposition pattern due to their limitation.
- BPNNs predict total deposition rates well but have limitations for regional patterns.
- XGBoost has superior performance in predicting both total and regional deposition rates.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Machine learning
Data-driven prediction
Oral drug delivery
Computational fluid dynamics (CFD)
Particle deposition

ABSTRACT

Targeted drug delivery to the deep lung improves therapeutic outcomes, but respiratory system variability complicates drug spray design. Numerical simulations offer insights for individualized treatments but are computationally intensive, highlighting the need for surrogate models for real-time deposition prediction. This study comprehensively explores the multi-task predictive capability of regression models, including Linear regression (LR), Bayesian regression (BR), Back Propagation Neural Network (BPNN), Support Vector Machine (SVM), Random Forest (RF), XGBoost, and CatBoost, for predicting total and regional deposition rates of inhaled particles in airway. A training dataset is obtained from well-validated CFD simulations with realistic human

* Corresponding author at: School of Mechanical and Automotive Engineering, Shanghai University of Engineering Science, 201620, China.
** Corresponding author.
E-mail addresses: yidan.shang@sues.edu.cn (Y. Shang), Wyang2001au@hotmail.com (W. Yang).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.powtec.2025.120924
Received 2 January 2025; Received in revised form 25 February 2025; Accepted 12 March 2025
Available online 21 March 2025

airway model using Euler-Lagrangian method. The results indicate that LR, BR, and SVM yield unsatisfactory predictive accuracy, with average R^2 values in range of 0.21 to 0.73. Comparatively, BPNN and decisiontree-based models show great potential in predicting total deposition rate in the upper and central airway. However, for regional deposition rate prediction, BPNN did not consistently yield high accuracy, particularly for oral deposition ($R^2 = 0.538$). Comparatively, XGBoost emerges as optimal model, achieving an R^2 approximately close to 1 on both the training and testing datasets, with predictive errors within the range of ±0.5. The overall results demonstrate that decision-tree-based models, particularly XGBoost, have superior performance in accurately predicting both total and regional deposition rates of inhaled particles within airway. Despite limitations like geometry complexity and data quantity, the workflow developed in this study is expected to pave the way for future research integrating ML models into drug delivery device design and evaluation.

1. Introduction

The growing environmental pollution driven by urbanization has become a global concern. In 2019, the World Health Organization (WHO) reported over 7 million deaths attributed to pollution [1]. The respiratory system, the only internal organ directly exposed to the external environment, is particularly vulnerable to pollutants through breathing [2]. The rise in pollution has led to an increase in respiratory diseases, including chronic obstructive pulmonary disease (COPD), asthma, and occupational lung diseases. COPD, in particular, ranks as the fourth leading cause of death in the United States, causing significant respiratory challenges due to airway rigidity, reduced elasticity, and inflammation-induced obstruction [3]. Conventional COPD treatment predominantly depends on the inhalation of therapeutic nano- and micro-particles; however, attaining the intended therapeutic results is problematic, as merely 25 % of medication particles successfully penetrate the deep lung areas [4]. Understanding the features of pulmonary air-particle movement is crucial for overcoming this barrier and enhancing drug delivery efficiency for the best treatment effects. This understanding is crucial for prompt diagnosis, efficient treatment, and the optimization of inhalation therapy.

Examining airflow and particle dynamics is crucial for optimizing inhalation drug delivery [5–7]. Nonetheless, achieving enhanced delivery efficiency continues to pose challenges. The complex structure of the lower airways, comprising bronchioles and alveoli, forms small, extensively branched passages that are challenging to access and visualize in real-time [8]. Moreover, traditional imaging modalities, such as CT and MRI, frequently lack the requisite resolution to accurately depict intricate particle dynamics within these diminutive airspaces without generating noise or artefacts [9]. The fluid dynamics of airflow, influenced by ongoing inhalation and expiration cycles, hinder the tracking of particles due to the constant alteration of flow patterns within the deeper lung. Furthermore, particles interact with lung tissue and airway walls through mechanisms such as diffusion, sedimentation, and inertial impaction. However, conducting in vivo studies is subject to ethical constraints, while in vitro and in silico models, like mock lung replicas, face inherent limitations in fully replicating the complexity of real airway environments [10]. These problems underscore the increasing demand for alternate approaches [11]. Consequently, numerical simulations have gained widespread popularity. For instance, Van de Moortele, et al. [4] numerically visualized and investigated the airway morphology and inspiratory flow features of COPD patients. Tiwari, et al. [12] used the CFD method to analyze the drug settling efficiency in different regions of the human respiratory tract for different drug particle sizes, under the condition of unsteady breathing pattern. Si, et al. [13] systematically investigated the effect of design parameters (e.g., vanes) on metered-dose inhaler drug delivery efficiency using large-eddy simulations. These studies demonstrate the robust capability of CFD methods to provide intuitive, parameter-driven insights into particle dynamics and drug delivery efficiency within the respiratory tract. A whole-lung model is essential for accurately predicting drug deposition and toxicity, however most models are limited to upper airway generations due to geometric complexity [14]. Reconstructing deeper lung generations is necessary to capture the complete airflow dynamics and particle distribution, ensuring improved drug delivery efficiency and more reliable predictions for therapeutic outcomes.

Conventional CFD models provide a comprehensive understanding of particle movement and deposition throughout the respiratory system [15–17]. Nonetheless, recent studies have highlighted that the significant variability observed in the tracheobronchial (TB) region leads to substantial differences in lung function and regional aerosol deposition, even after adjustments for age, sex, and body size [11]. This suggests that case-specific CFD simulations are necessary for each patient to obtain accurate outcomes. Individualized models reveal unique airflow patterns that significantly affect particle transport and deposition, complicating conventional predictions and limiting generalizability. Consequently, a standardized sprayer design may not consistently achieve the desired delivery efficiency across different patients. On the other hand, performing CFD simulations for each patient to optimize their patient-specific treatments requires substantial computational resources [18]. Therefore, creating rapid prediction models would significantly aid in the preliminary examination of drug deposition patterns, offering a valuable resource for inhaler design or drug delivery strategy optimisation. Given these challenges, integrating CFD with machine learning (ML) methods presents a promising solution, although ML has seen limited application in predicting particle deposition. This limited use originates from the lack of large, high-quality datasets necessary to effectively train these models [11]. The lung's complex, multi-regional structure - spanning from the upper airways to deep lung generations - requires predictive models capable of multi-task learning to accurately capture both local and systemic interactions, including regional airflow, turbulence, and deposition patterns across different lung zones. By combining machine learning with computational fluid dynamics, these integrated models may overcome current limitations and offer innovative solutions for more precise and personalized particle deposition predictions. To the best of the authors' knowledge, currently, there are still very limited studies integrating CFD with ML to predict particle deposition in the human airway.

To advance the development of a machine learning (ML)-based surrogate model for this purpose, this study first constructs a human upper and central airway model based on patient-specific data. The transport and deposition characteristics of inhaled particles via the oral pathway are analyzed using Euler-Lagrangian methods. To enable fast predictive models for local regional deposition rates, several advanced machine learning models are selected, including the Linear regression (LR), Bayesian regression (BR), Back Propagation Neural Network (BPNN), Support Vector Machine (SVM), Random Forest (RF), extreme gradient boosting (XGBoost), and categorical boosting (CatBoost), to explore their multitask prediction capabilities in particle dynamics. This study aims to guide the selection of an optimal ML algorithm for this purpose while developing a surrogate model for particle deposition rate prediction.

2. Method

2.1. Reconstruction of human mouth-lung model

The airway model used in this study is identical to that of Banko,

et al. [19], which has been adopted for analyzing the transport and deposition characteristics of inhaled particles in a upper and central airway model. The anatomical features of the model include the extra-thoracic airway (i.e., oral cavity, pharynx), upper airway (i.e., larynx), and lower airway (i.e., trachea, right and left lobes) with the lung network extending to generation 8 (G8). To quantify the regional drug deposition characteristics of the particles ejected via the inhaler or oral spray, the model is carefully segmented, as illustrated in Fig. 1. This lung model is reconstructed based on a CT scan of a 47-year-old healthy male, weighing 78 kg and measuring 174 cm in height, with no history of respiratory disease.

The reconstructed airway model is discretized using the commercial software package ANSYS Fluent Meshing 2022 R1. The polyhedral mesh with typical cell dimensions ranging from 0.05 mm to 0.1 mm is used to discretize the entire computational domain. Polyhedral meshes are less sensitive to element stretching, outperfoming the widely used tetrahedral mesh while discretizing complex multi-scale geometries [20]. Five prism layers with an aspect ratio of 5 are further employed to capture the detailed variable changes at the wall boundary, such as wall shear stress and local deposition efficiency. Mesh independence tests are conducted with 3 sets of mesh configurations prior to adding particles, i.e., 0.86 million, 1.46 million and 3.1 million.

As shown in Fig. 2, a grid independence test has been performed through the project with lines in the domain are extracted and the results of velocity filed are compared. The velocity field data was extracted from the plane near the larynx where z = 0.135 m, y = 0.205 m. By comparing the velocity variation, consider both the accuracy and the requirement of computational rescource, a 1.46-million polyhedral mesh configuration is ultimately adopted for the subsequent simulations.

2.2. Boundary conditions

For the purpose of deposition prediction via machine learning, several combinations of the independent variables are considered and parametrically investigated in this study. Four respiratory airflow rates, i.e., 15 L/min, 30 L/min, 45 L/min, and 60 L/min, from low respiratory rate to high respiratory rate are performed. The inlet airflow originates solely from the oral cavity. The inlet velocity is specified at the mouth,

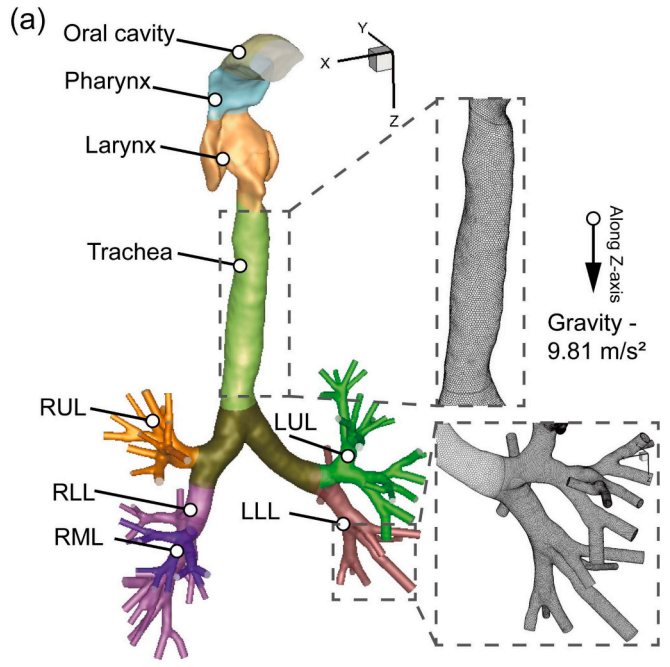


Fig. 1. Model reconstruction of the patient-specific lung with mesh details.

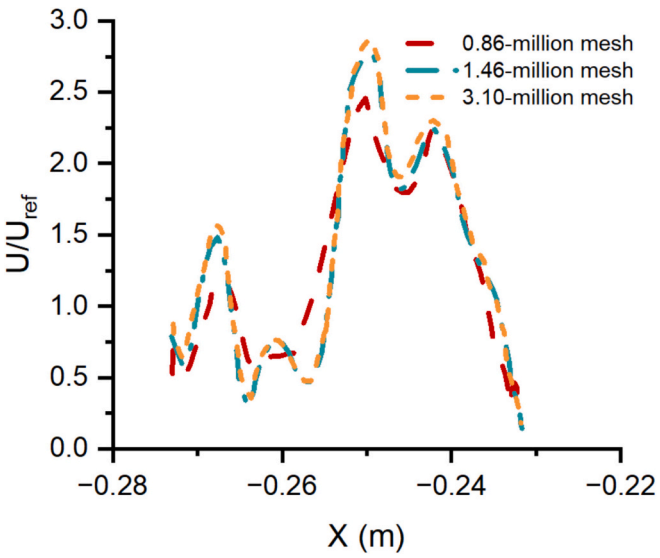


Fig. 2. Result of mesh independence test performed through the process.

while pressure outlets are set for each lobe. For general drug delivery, the particle size distribution primarily falls in the range from 0.5 μm to 20 μm [21]. Therefore, in this study, particle sizes of 0.5, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12.5, 15, and 20 μm are investigated based on the steady-state airflow field. One-way coupling is considered in this study due to relatively low volume fraction of inhaled particles. Spherical particles with a unit density of water (i.e., 998.2 kg/m^3) are adopted. For each injection with a different particle size, 30,000 particles are randomly released from the oral interface to mimic the behavior of particle inhalation patterns via oral spray based on User-defined Functions (UDFs), as shown in Fig. 3. In addition, initial particle velocities ranging from 0 % to 100 % of the initial respiratory airflow rates are also tested. The detailed range of training variables for regional deposition fraction prediction is outlined in Table 1. The selection of these three variables as training features is due to the following reasons:

1) respiratory airflow rate represents realistic breathing conditions, ranging from rest to exertion; 2) particle sizes align with the typical range for drug delivery systems; 3) initial particle velocities capture variations in spray dynamics. These variables are believed to be key factors affecting deposition patterns and ensure a diverse dataset for ML

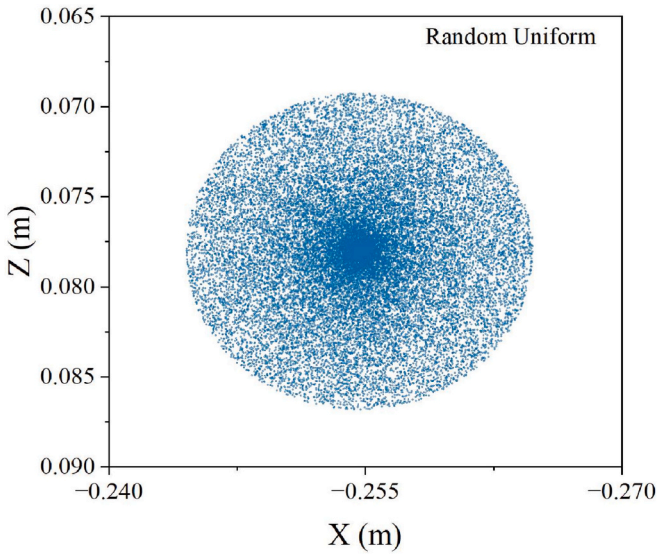


Fig. 3. Random distribution pattern of injected particle via UDF.

Table 1
Independent training variables in this study.

Respiratory rate (L/min)	Fraction of initial respiratory airflow rate (%)	Particle diameter (μm)
15, 30, 45, 60	0, 25, 50, 75, 100	0.5, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12.5, 15, 20

training.

Considering the volume fraction and size of the particle involved in current research is relatively low, one-way coupling was as applied in the simulation assuming that their effect on the airflow field can be negligible. In this study, four steady-state CFD cases based on different breathing rates were first simulated to establish the baseline airflow fields. Based on these results, the Lagrangian model was directly applied to predict particle transport and deposition under different conditions, generating the dataset for machine learning training. A total of 320 case scenarios were tested.

2.3. Mathematical models

In this study, four steady-state airflow simulations (for respiratory rates reported in Table 1) are first performed using the transitional SST turbulence model, aiming to capture the flow transition from laminar to turbulent. This turbulence model has been widely adopted, providing an acceptable balance of computational cost and accuracy compared to large-eddy simulations (LES) [22,23]. The transition SST model is based on the coupling of the SST k-omega model with two other transport equations: one for the transition onset criteria and the other for the intermittency. The production term is further multiplied by the turbulence intermittency (γ), which can be expressed as.

$$P_k \rightarrow P_k\gamma \quad 0 \le \gamma \le 1 \tag{1}$$

$P_k = 0$, indicating the laminar boundary layer, and $0 < P_k < 1$ represents the transitional boundary layer and $P_k = 1$ describes the turbulent boundary layer.

The additional transport equation for γ can be expressed as

$$\frac{\partial(\rho\gamma)}{\partial t} + \frac{\partial(\rho U_j\gamma)}{\partial x_j} = P_{\gamma 1} - E_{\gamma 1} + P_{\gamma 2} - E_{\gamma_2} + \frac{\partial}{\partial x_j}\left[\left(\mu + \frac{\mu_t}{\sigma_\gamma}\right)\frac{\partial\gamma}{\partial x_j}\right] \tag{2}$$

where $P_{\gamma 1}$ and $E_{\gamma 1}$ are the transition sources, $P_{\gamma 2}$ and E_γ are the destruction/relaminarization sources.

The trajectories of the inhaled particles were calculated using the Eulerian-Lagrangian method [24]. This method additionally calculates the force balance equation to account for the force effect on the particle surface, including the drag force, $\vec{F}D$, gravity force, $\vec{F}g$, and pressure gradient force, $\vec{F}p_{\text{grad}}$, which can be computed with the following equation

$$\vec{F}_D = \frac{18\mu(u_i^g - u_i^p)}{C_c d_p^2 \rho_p} \tag{3}$$

$$\vec{F}_g = m_p\vec{g} \tag{4}$$

$$\vec{F}_{p_{\text{grad}}} = u_p\rho_f\frac{D\vec{u}_f}{Dt} \tag{5}$$

where μ is the air viscosity, u_i^g is the airflow velocity, ρ_p is the particle density, d_p is the particle diameter, u_p is the velocity of particle, ρ_f is density of fluid, and C_c is the Cunningham correction factor.

The overall particles released from the oral interface are tracked until they either deposit on the airway walls or exit each lobe outlet. The equation for calculating the regional deposition fraction is as follows:

$$DF_i = \frac{PD_i}{TR_j} \times 100\% \tag{6}$$

where DF_i predicted the deposition fraction in anatomical region i, PD_i predicted the deposition count into anatomical region i, TR is the total number of particles, with particle size j released into the airway.

2.4. Supervised regression models

In this study, several popular supervised ML models are adopted to predict the nonlinear deposition characteristics of inhaled particles in the airway. Seven widely used ML models are selected: the Linear regression (LR), Bayesian regression (BR), Back Propagation Neural Network (BPNN), Support Vector Machine (SVM), Random Forest (RF), extreme gradient, boosting (XGBoost), and categorical boosting (CatBoost). Before the training process, data curation is first performed on the data obtained from the CFD simulations to ensure consistency and cleanliness. Relevant features are then identified for model training, validation, and testing. In this study, the input variables include human breathing rate, particle initial velocity, and particle size, while the output variables are the total and regional deposition rates (i.e., Oral, Pharynx, Larynx, Trachea, RUL, RML, RLL, LUL, and LLL).

BPNN are among the most popular machine learning algorithms, inspired by the structure and function of biological neural networks in animal brains. These models imitate the human brain structure, with neurons in an interconnected structure, i.e., layers of interconnected nodes, to effectively bridge input variables and output variables via this unique architecture [25]. The labeled information as input is fed into the input layer (X), and the predicted data is ultimately exported from the output layer (Y) via a number of in-between layers, known as hidden layers. In this study, for BPNN, training is performed using minibatch gradient descent with the Adam optimizer. The network architecture itself includes a fully connected (linear) layer with 32 hidden units, followed by a ReLU activation function to introduce non-linearity.

For limited datasets, particularly those in tabular form, decision tree-based models show superior predictive performance [26]. These algorithms model decision and predictive results with a tree-like structure that has been used for both regression and classification tasks. During training, the model splits the data into subsets based on the values of input features, forming branches that represent decision paths. Each internal node in the tree corresponds to a feature in the data, with branches representing the possible outcomes of a decision rule applied to that feature. The terminal node or leaf represents the final output or decision, which may be a class label in a classification task or a continuous value in a regression task. Based on this concept, various advanced decision tree-based models have been further proposed, including the random forest [27], XGBoost [28], and CatBoost [29], etc. The detailed schematic diagram is shown in Fig. 4, and a detailed explanation can be found in our previous study [30].

For the purpose of benchmark, several basic regression models are also tested, such as LR [31] , BR [32] and SVM [33]. LR is a very basic ML model that positing a linear relationship between input and output features, with a relationship generally in a form of a linear equation ($y = mx + b$). BR, on the other hand, incorporates probabilistic priors into the regression process, enabling it to account for uncertainty in model parameters and providing robust predictions even with limited data [32]. Support Vector Machine (SVM) is a versatile model that constructs hyperplanes in high-dimensional feature spaces to optimize regression tasks, known for its effectiveness in handling both linear and non-linear relationships through kernel functions [33].

A grid search is performed on the development set to identify the optimal hyperparameters for each model, as outlined in Table 2.

Training variables in Table 1 are used to predict the total and regional particle deposition rates based on the model segmentation. During training, cross-validation is performed, with 72 % of the random data rows used for training, 8 % for validation, and 20 % for testing. Two evaluation metrics are used Two evaluation metrics are used in this study, including the mean square errors (R2) and mean absolute error

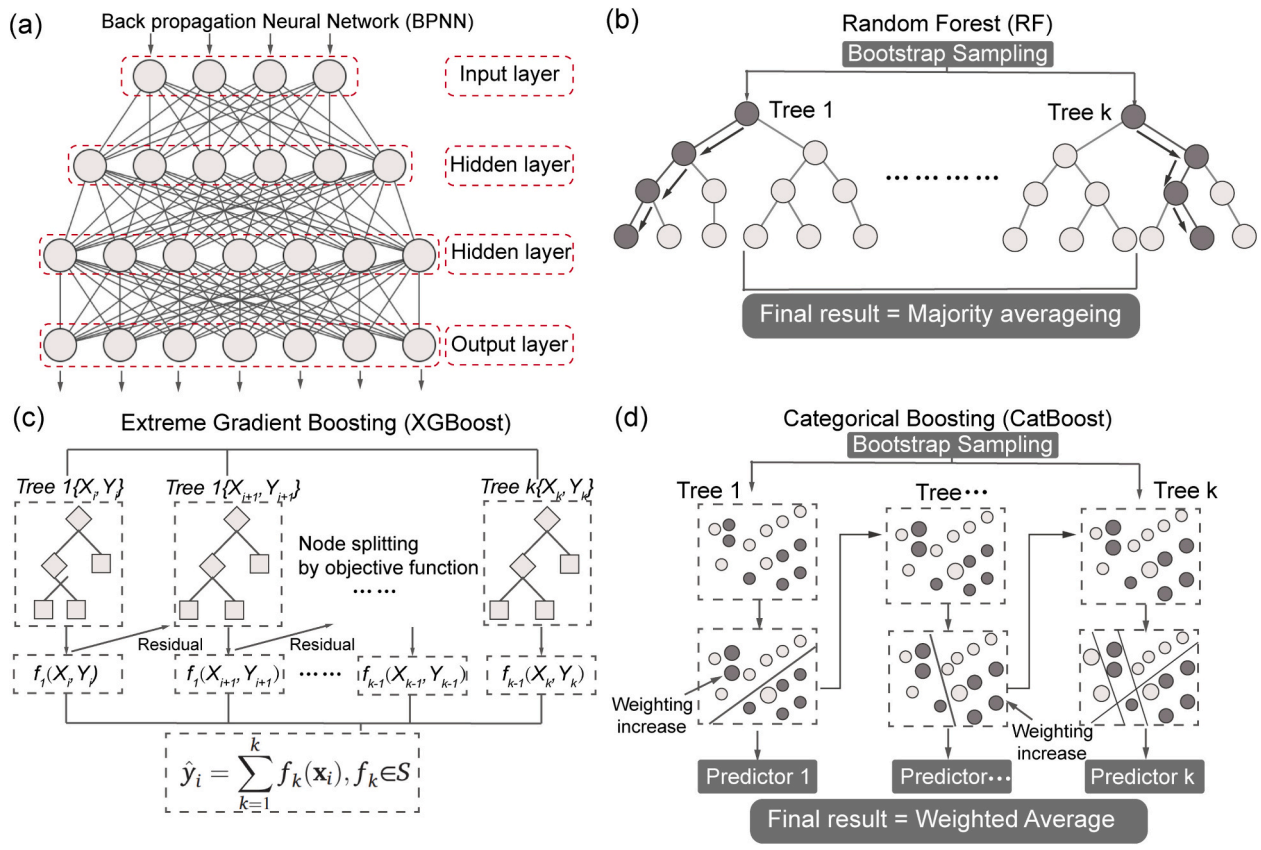


Fig. 4. Schematic diagram of (a) BPNN, (b) RF, (c) XGBoots and (d) Catboost.

Table 2
Range of hyperparameter optimization (Note: The text with bold font represents the optimal hyperparameter).

Model	Parameter	Range for tested hyperparameter
SVM	Kernel	'linear', 'poly', 'rbf', **'sigmoid'**
	C	0.1, **1**, 10, 100
	Gamma	'scale', **'auto'**, 0.01, 0.1, 1
	Epsilon	0.1, 0.2, **0.5**, 1, 128
BPNN	Learning rate	**0.01**, 0.001, 0.0001
	Hidden layer sizes	2, 4, 8, 16, **32**, 64
	Dropout rate	0.0, **0.1**, 0.2, 0.5
	Batch size	32, 64, **128**
Random Forest	N_estimators	50, **100**, 200
	Max depth	**None**, 10, 20, 30
	Min samples split	**2**, 5, 10
	Min samples leaf	**1**, 2, 4
	Max features	Auto, **sqrt**, log2
CatBoost	Learning rate	0.01, 0.05, **0.1**, 0.2
	Depth	4, 6, 8, **10**
	Iterations	5000, **10,000**, 15,000
	L2 leaf reg	1, 3, 5, **10**
	Bagging temperature	**0**, 1, 3, 5
	Grow policy	**Symmetric tree**, depthwise, lossguide
XGBoost	Max depth	6, **10**, 12
	Learning rate	**0.01**, 0.05, 0.1
	N estimators	100, **500**, 1000
	Colsample bytree	0.6, **0.8**, 1.0
	Subsample	0.8, **1.0**
	Tree method	Hist
	Multi strategy	Multi-output tree

(MAE). The equations are expressed as follows:

$$R^2 = 1 - \frac{\sum (y_i - \hat{y})^2}{\sum (y_i - \bar{y})^2} \tag{7}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{8}$$

where y_i is the actual value, $\hat{y}$ is the predicted value and $\bar{y}$ is the mean value of y, n is the overall sample size.

To prevent overfitting, several strategies were applied across models. Dropout, learning rate reduction, and early stopping were used for BPNN, while early stopping was also implemented for tree-based models like XGBoost and CatBoost. For other models, such as Linear Regression, BayesianRidge, SVM, and RandomForestRegressor, intrinsic regularization mechanisms like probabilistic frameworks and bagging were relied upon to mitigate overfitting.

3. Results and discussions

In this section, the airflow pattern is first validated against the experiment conducted by Banko, et al. [19]. The airway surface pressure and streamlines are further analyzed.

3.1. Airflow pattern and particle deposition validation

Fig. 5 compares the experimental data obtained by Banko, et al. [19] and the numerical simulation of this study at the steady flow rate of 60 L/min. Three representative planes are selected. Plane 1 is located in the pharynx and is the narrowest region of the upper airway(z = 0.118 m). Plane 2 is located in the larynx(z = 0.146 m) and plane 3 is located in the trachea downstream of the larynx(z = 0.21 m). The velocity on each plane is scaled according to the experimental data, in which U is the

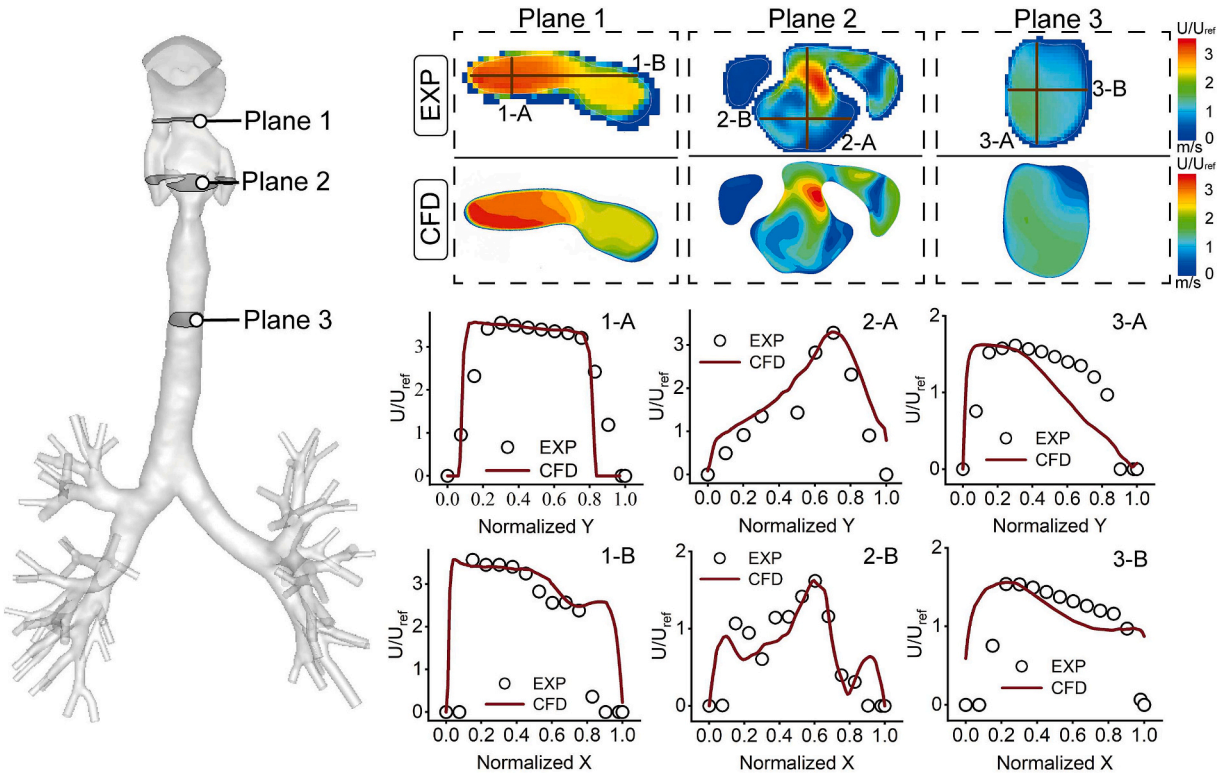


Fig. 5. Airflow pattern validation at three planes through the pharynx, larynx and trachea.

velocity magnitude and U_{ref} is the inhalation velocity at the mouth. The experiment was done with magnetic resonance velocimetry (MRV) and the original experiment material was water in the referenced work with an inlet mass flow of 3.78 L/min to achieve the relative Reynolds number and maintain the similarity of the flow pattern. Qualitatively, the velocity flow patterns obtained from numerical simulations show satisfactory agreement with experimental data, capturing most key airflow field characteristics. Velocity profiles along two lines in each plane are further extracted and quantitatively compared. The numerical results predict similar airflow behavior in Plane 2 due to high velocity but low turbulence laminar flow field. With the enhanced turbulence intensity downstream the Plane 1, the predictive discrepancy increased, particularly evident in Plane 3, Line A. However, the overall trend is accurately captured and the agreements are acceptable.

To validate the reliability of the presented airway model, the predicted profile of the total inhaled particle deposition fraction (DF) versus the Stokes number was compared with data from the literature, which all compared with the experimental outcomes. Numerical deposition predictions in the mouth-lung model were evaluated against results from [22], Kenjereš and Tjin [23], Cheng, et al. [34] and Zhang, et al. [35], as shown in Fig. 6. The comparison reveals that our simulations exhibit a similar DF pattern across various particle sizes, consistent with the literature data. However, it is worth noting that our study shows a higher overall deposition rate compared to the literature, although the general trend remains well-aligned and acceptable.

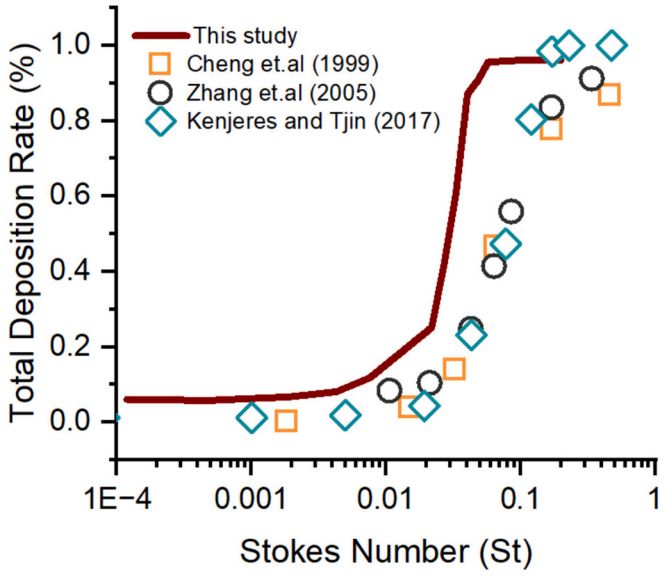


Fig. 6. Particle deposition validation compared to work by Kenjereš and Tjin [22,23], Cheng, et al. [34] and Zhang, et al. [35].

3.2. Airflow dynamics in the mouth-lung model

The airflow dynamics within the airway models are thoroughly analyzed before training the ML models. Fig. 7 presents a comparative evaluation of the surface pressure distribution across different inhalation intensities, ranging from low to high. The results indicate a notable decrease in surface pressure along the flow path, extending from the oral cavity to the distal regions of the lung. A key observation is the presence of negative pressure in regions where airway constriction occurs, particularly at the glottis, which plays a critical role in regulating airflow resistance. In the oral cavity, relatively higher pressures are observed, with a measurable increase from 5 Pa under low inhalation to nearly 90 Pa under high inhalation conditions. This pressure buildup reflects the initial resistance at the entry point, as airflow transitions from the ambient environment into the respiratory tract. Beyond the oral cavity, a pronounced pressure drop is observed downstream, particularly as the airflow passes through the pharynx, where geometric narrowing introduces additional flow resistance.

The inhalation airflow patterns are further illustrated in Fig. 8. The

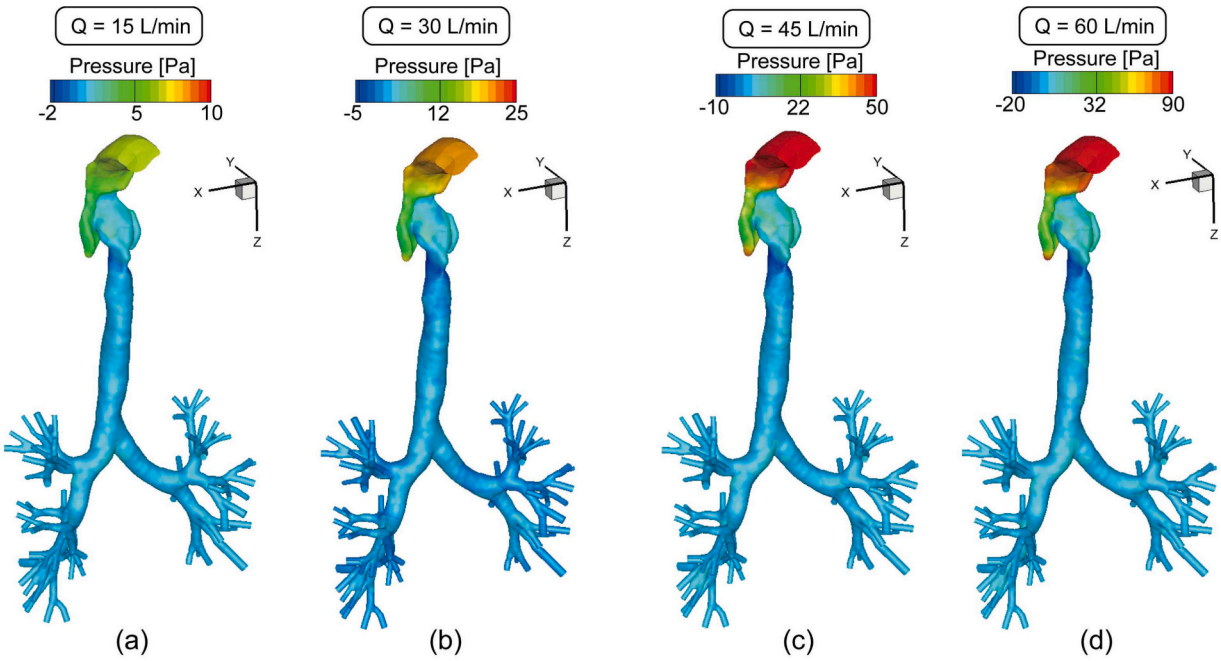


Fig. 7. Global surface pressure distribution of the airway model under an inhalation flow rate of (a) 15 L/min, (b) 30 L/min, (c) 45 L/min, and (d) 60 L/min.

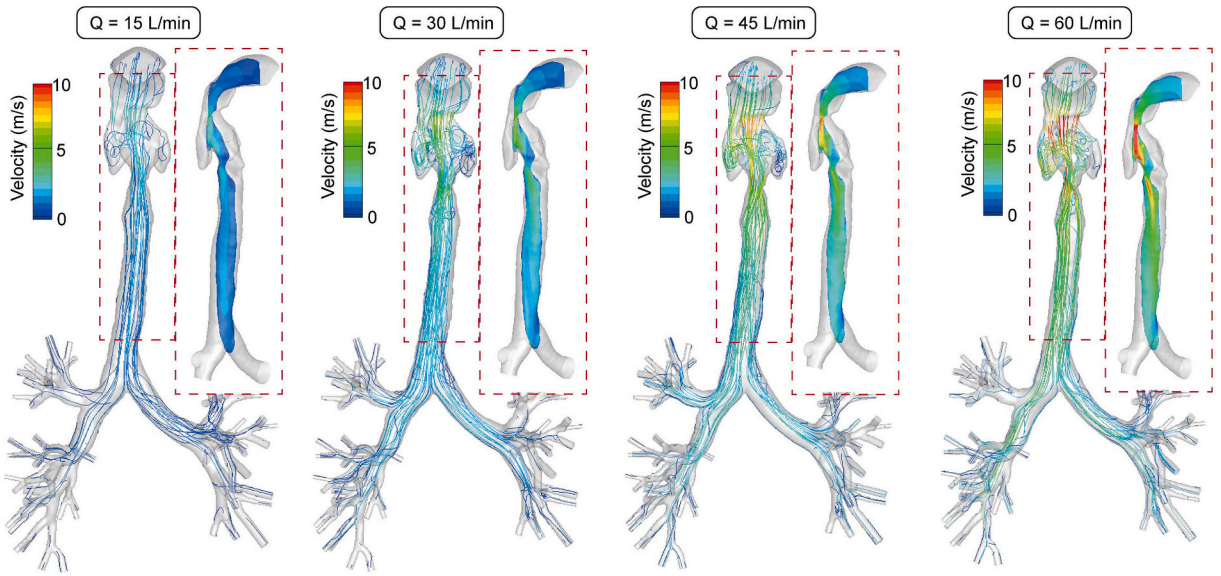


Fig. 8. Airflow streamlines and vertical contour within the airway model under low-to-high inhalation flow rate.

results reveal that airflow recirculation commonly occurs in the oropharynx, supraglottic, and subglottic regions. This phenomenon arises primarily from the narrowing of the airways, where laminar flow accelerates and develops into a laryngeal jet. The high-velocity jet impinges on the posterior airway wall, leading to flow separation and the formation of recirculation zones in the anterior regions. To facilitate an intuitive comparison of velocity variations, the legend bar in Fig. 7 remains consistent across different flow conditions. At a low inhalation flow rate of 15 L/min, the laryngeal jet reaches a velocity of approximately 3 m/s. This velocity increases to over 10 m/s when the inhalation flow rate is 60 L/min. The increase in jet velocity at higher flow rates suggests a greater likelihood of particle deposition in the airway due to enhanced inertial impaction and reduced flow adherence to the airway walls.

3.3. Particle dynamics in the mouth-lung model

Fig. 9 illustrates the spatial deposition patterns of the mouth-lung model based on selected representative inhaled particle sizes of 1, 5, 10 and 20 μm under the respiratory airflow rate of 15 L/min, 30 L/min, 45 L/min and 60 L/min, respectively. The deposition patterns demonstrate the mechanistic impact of particle size and airflow rate on deposition sites throughout the respiratory system. For small particles (1 μm), deposition predominantly occurs in the deep lung areas owing to their low inertia, enabling them to tightly adhere to the airflow streamlines and access the alveolar regions. As particle size increases to 5 μm and 10 μm, the particles demonstrate increased deposition in the central airways and upper bronchi, where alterations in airflow direction and velocity result in heightened inertial impaction. For the largest particles (20 μm), deposition primarily occurs in the upper respiratory tract,

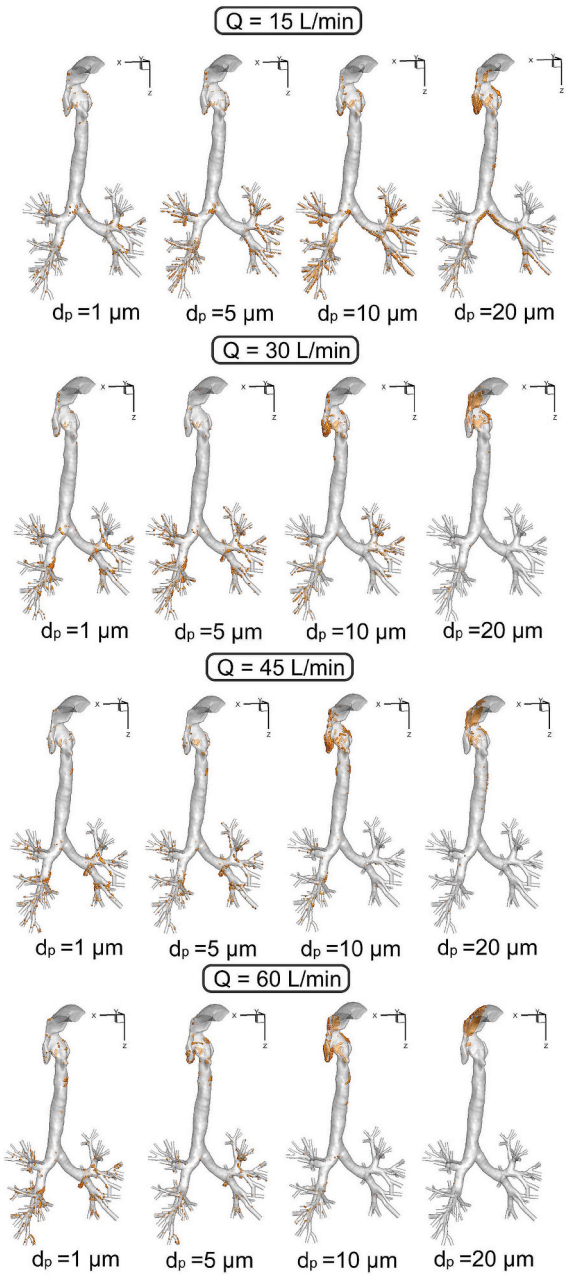


Fig. 9. Deposition patterns of the inhaled particles in mouth-lung models under inhalation flow rates of 15 L/min, 30 L/min, 45 L/min, and 60 L/min.

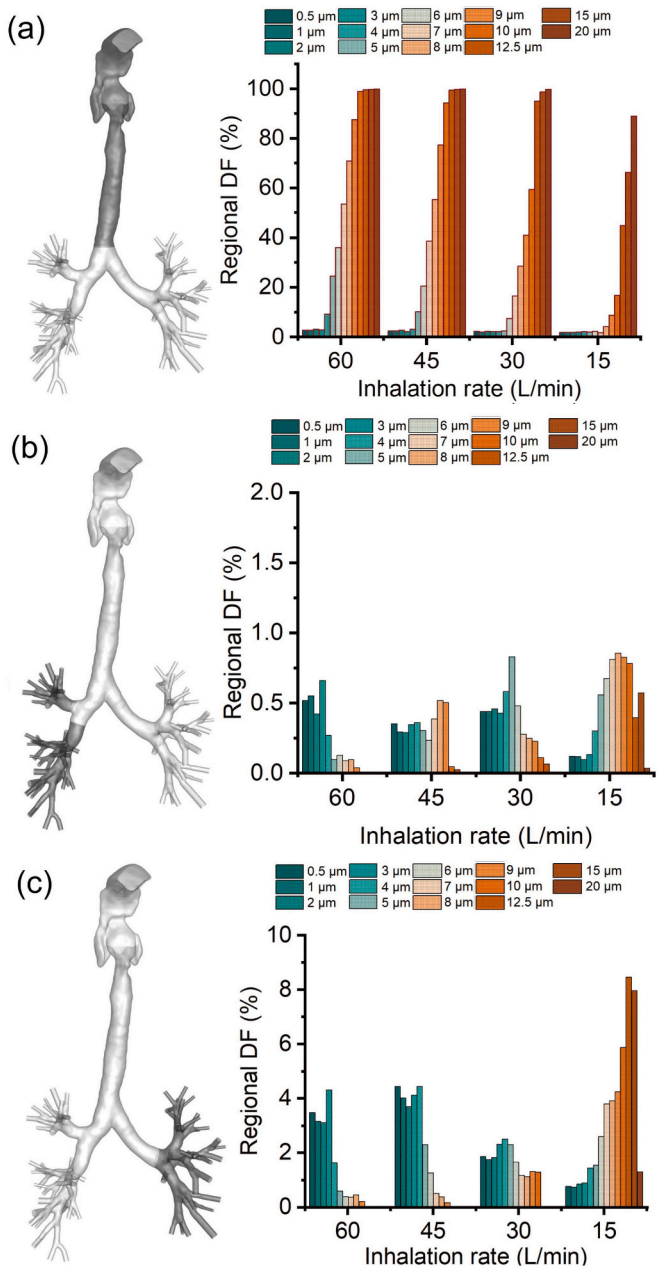


Fig. 10. Average regional deposition fraction in (a) oral-trachea, (b) lower right lobes and (c) lower left lobes.

encompassing the mouth and throat, owing to their significant inertia, which hinders their ability to adhere to airflow streamlines, resulting in early impact and deposition on airway surfaces.

With increasing airflow rate this phenomenon is further intensified as particles gain greater momentum from the respiratory jet flow. As a consequence, 20 μm-particles are more likely to settle in the oral cavity and very few particles penetrating through the larynx.

Fig. 10 illustrates the average regional deposition fraction (DF) of inhaled particles within different sections of the respiratory tract, specifically focusing on the oral-trachea (a), the lower right lobes (b), and the lower left lobes (c) at various inhalation rates of 15, 30, 45, and 60 L/min. In Fig. 10(a), it is observed that in the upper respiratory airway region, the local deposition fraction increases significantly with particle size. Larger particles (e.g., 10 μm and above) exhibit particularly high deposition fractions, even at lower inhalation rates, due to inertial impaction. Particles in the size range of 3 μm to 7 μm appear to be highly sensitive to changes in respiratory rate. The deposition fraction for these particle sizes gradually increases as inhalation rates rise, which can be attributed to a combination of increased inertial impaction from greater jet momentum and the particles' higher mass, both of which enhance deposition on airway surfaces. The deposition fraction varies differently for the lower respiratory system, as illustrated in Fig. 10(b) and (c). Smaller particles (e.g., 0.5–5 μm) are more likely to reach the deeper lung regions, and their deposition fraction is maximized at higher inhalation rates. The lower deposition efficiency of larger particles (e.g., 15–20 μm) in these distal regions is likely due to their tendency to deposit in the upper airways before reaching the lower lobes. Additionally, the distribution of particles in this region highlights the role of gravitational settling and diffusion, as smaller particles have a greater ability to follow the airflow down to the distal areas.

3.4. Predictive results via multi-output regression models

In this study, several surrogate models using ML algorithms are established based on validated CFD simulations. The overall dataset is randomly divided, with 72 % selected for training, 8 % for validation, and 20 % for testing. The training variables include respiratory rate, injected particle diameter, and initial particle injection velocity. The prediction variables include the total and regional deposition rates within the oral cavity, pharynx, larynx, trachea, right upper lobe (RUL), right middle lobe (RML), right lower lobe (RLL), left lower lobe (LLL), and left upper lobe (LUL).

Tables 3 and 4 summarize the comparison of predictive errors between the results of the ML models. As shown in Fig. 11, the R^2 values for predicting the total deposition rate are calculated as 0.8361, 0.9802, 0.8361, 0.999, 1, 1 and 0.9955 for LR, SVM, BR, RF, Catboost, XGBoost, and BPNN models, respectively, based on the training dataset. These values indicate a very strong correlation between the input and output training variables, except for the LR and BR, with their R^2 smaller than 0.9. For the other models, it can be observed that BPNN and SVM have comparatively lower R^2 value, suggesting that, for a tabular dataset with limited data, decision-tree-based models generally outperform the other ML algorithm. Additionally, CatBoost and XGBoost demonstrate the best prediction capabilities, with both models successfully correlating the training and predictive features. It is worth noting that, although the R^2 values of CatBoost and XGBoost are 1, this result is achieved by adopting an overfitting strategy. R^2 values based on the testing datasets are 0.8219, 0.9729, 0.8222, 0.9997, 0.9977, 0.9999 and 0.9932 for LR, SVM, BR, RF, Catboost, XGBoost, and BPNN models, respectively. This implies that, for predicting the particle deposition rate, these popular ML algorithms can all yield satisfactory fitting results expect for LR and BR, with XGBoost being the best option.

In Fig. 12, the mean absolute errors between the predicted and numerical data provide a detailed comparison of model performance when predicting regional deposition rates. It can be found that MAE values are comparatively higher for LR, SVM, and BR for both the training and testing datasets. For the training dataset, such MAE values for these three models are generally within 1. However, these models show the worst generalizability for the testing dataset, with MAE greatly exceeding nearly 10. It is worth mentioning that the BPNN model also shows relatively lower accuracy, especially in predicting deposition rates in the pharynx, right lower lobe (RLL), and notably the larynx, where the MAE exceeds 2. This suggests that BPNN may struggle with limited datasets while predicting the complex or non-linear relationships in the data.

On the other hand, decision-tree-based models, particularly XGBoost, demonstrate superior performance. XGBoost achieves lower MAE values across all regions, outperforming both RF and CatBoost. The MAE for each regional deposition rate prediction is consistently below 0.15 for decision-tree-based models, indicating their robustness and precision in handling intricate spatial deposition patterns within the respiratory tract. These findings underscore the suitability of decision-tree-based models, particularly XGBoost, for applications where high accuracy in predicting regional deposition is critical. The consistent performance of XGBoost across various regions suggests its potential for reliable predictions in real-world scenarios, where accurate modelling of particle deposition can significantly impact the design and optimization of inhalation therapies.

To summarize the model performance in the training and testing cases, the error histograms of each model are outlined. The error histograms for predicting the total deposition rate are first presented in Fig. 13. These histograms visualize the discrepancy between the predicted and targeted outputs, with the y-axis representing the number of samples with corresponding error sizes and the x-axis representing the error magnitude between the ground truth and the predicted outcomes. It can be observed that decision-tree-based models, particularly XGBoost and CatBoost, show significantly smaller error distributions compared to other models. Both models demonstrate a strong alignment between ground truth and predicted values, with their errors tightly centered around zero. XGBoost is the model that exhibits the smallest error ranges during the training and testing processes. Further information with regards to the regional deposition rate can be found in Appendix A. It is evident that, for regional deposition rate prediction, decision-tree-based models still show great potential for such multi-task predictive capability, particularly XGBoost.

4. Further discussion and future research prospects

This study concludes that decision-tree based models are better solutions for predicting very complicated non-linear relationships in the data of particle deposition. Numerous tabular datasets frequently exhibit objective functions that lack smoothness, revealing intricate nonlinear interactions. This attribute does not provide a difficulty for tree-based models, as they utilise a piecewise modelling technique, with each decision tree node adjusting to the local anomalies in the data distribution. It should be noticed that the underperformance of BPNN in regions like oral deposition can be attributed to its limitations in handling complex, multi-output regression tasks. BPNN models are highly dependent on large, high-quality datasets for effective training. When data is sparse or lacks diversity, the model struggles to generalize, resulting in reduced accuracy, particularly in intricate tasks where subtle input variations can significantly influence the output. Therefore, when confronted with intricate target variables featuring many extrema or irregular patterns, tree-based models such as XGBoost are more adept at precisely capturing these relationships.

In the current research, the feasibility of introducing machine learning to predict particle drug deposition within the human airway has been tested. Meanwhile, the participants of the current research have also noted that this attempt may contain limitations that need to be addressed in future studies. The limitations and recommended future directions are summarized below.

Current CFD-DPM studies have shown that airway diversity significantly influences particle deposition rates. For future ML-based prediction models, it is recommended to classify airway geometries based on population diversity, and use machine learning to learn geometric

Table 3
The summary of R^2 value of each model based on training cases.

Variables	LR	SVM	BR	RF	CatBoost	XGBoost	BPNN
Total deposition	0.8361	0.9802	0.8361	0.9999	1	1	0.9955
Oral	0.2251	0.5261	0.2246	0.9975	1	0.9968	0.9546
Pharynx	0.5346	0.9788	0.5345	0.9992	1	1	0.9916
Larynx	0.4868	0.9754	0.4867	0.9996	1	1	0.9881
Trachea	0.3536	0.6426	0.3534	0.9985	1	0.9984	0.8257
RUL	0.3359	0.8554	0.3356	0.9979	1	0.9955	0.9605
RML	0.1948	0.4215	0.1942	0.9953	1	0.994	0.741
RLL	0.0352	0.3987	0.0304	0.9979	1	0.998	0.8035
LUL	0.2923	0.8298	0.2919	0.9973	1	0.9968	0.9338
LLL	0.1158	0.7261	0.1146	0.999	1	0.999	0.9181

Table 4
The summary of R^2 value of each model based on testing cases.

Variables	LR	SVM	BR	RF	CatBoost	XGBoost	BPNN
Total deposition	0.8219	0.9729	0.8222	0.9997	0.9977	0.9999	0.9932
Oral	−0.6044	0.8407	−0.4894	0.9926	0.9688	0.9275	0.5379
Pharynx	0.447	0.9775	0.4489	0.9971	0.9943	1	0.9899
Larynx	0.5247	0.9685	0.5254	0.9987	0.9976	0.9998	0.9858
Trachea	0.4144	0.6423	0.4151	0.9913	0.9613	0.9893	0.7075
RUL	0.2736	0.8283	0.2769	0.9814	0.9818	0.9803	0.9441
RML	0.0476	0.1824	0.0574	0.9762	0.9349	0.9568	0.6328
RLL	−0.0073	0.2844	0.005	0.9956	0.9713	0.9963	0.7119
LUL	0.1838	0.7502	0.1879	0.9592	0.9581	0.9696	0.8374
LLL	0.0014	0.5411	0.0073	0.995	0.9714	0.9951	0.9115

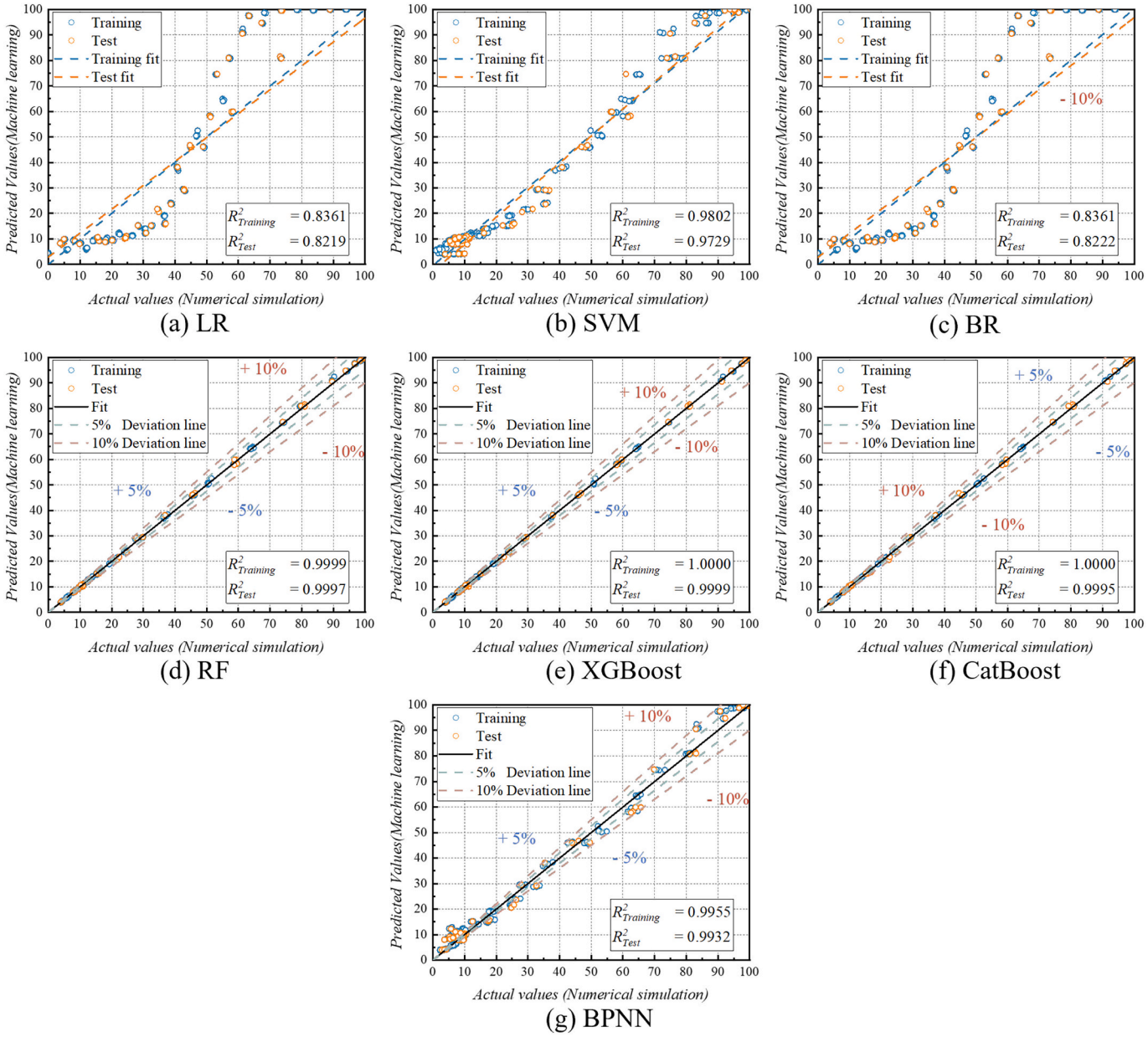


Fig. 11. The correlations of total deposition fraction between the CFD results and the predictions by the (a) LR, (b) SVM, (c) BR, (d) RF, (e) XGBoost, (f) Catboost and (g) BPNN based on the training and testing dataset.

information and flow characteristics to improve prediction accuracy. In addition, it is worth mentioning that for future studies, the effect of the proportions of the training dataset should be carefully examined, and the predictive capability using other respiratory models should be further tested to validate the ML robustness.

Considering the sensitivity of the current patient-specific research,

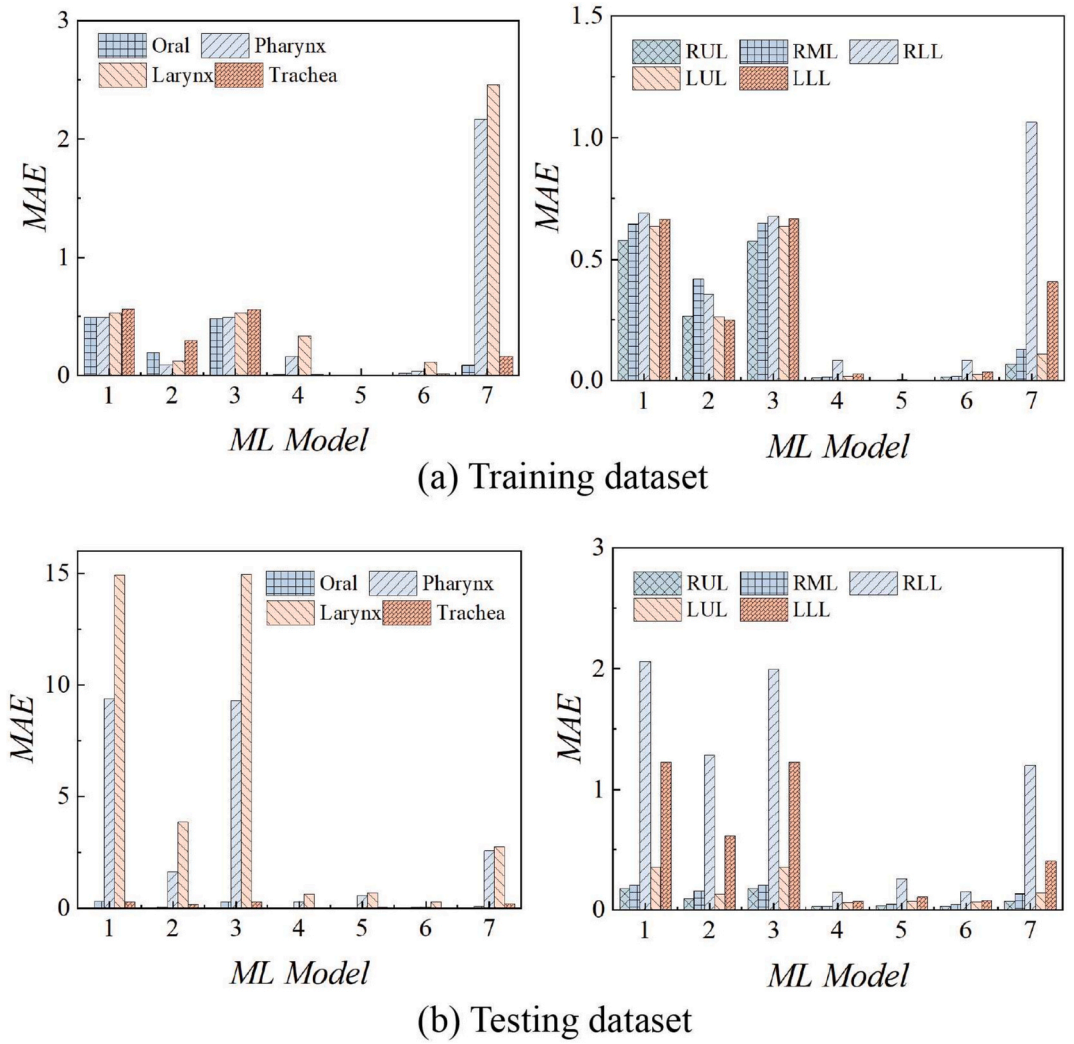


Fig. 12. The comparison of the absolute error distributions as reflected by MAE (a) based on training dataset and (b) testing dataset, with model 1) BPNN, 2) RF, 3) Catboost, 4) XGBoost, 5) SVM, 6) LR and 7) BR.

additional human airway models could be introduced to enhance model generalizability in the future. This would provide more training features and geometric information of human airways, which is beneficial for improving predictive accuracy. To improve the generalizability of ML models for inhaler design and drug delivery, additional input variables such as sprayer orifice diameter, half-angle, mass flow rate, etc., could be considered. For liquid formulations, particle density, surface tension, and viscosity are also critical factors that can be considered in future study. Furthermore, the human dynamic breathing pattern is also worth incorporating into ML studies. While the machine learning models used in this study primarily rely on static input features, transient flow variations under different breathing patterns can still be incorporated as additional input variables. However, for datasets involving strong temporal dependencies, LSTM or other recurrent neural network (RNN) architectures could be more effective in capturing sequential relationships, and future work may explore their integration to enhance predictive accuracy for dynamic inhalation conditions. While incorporating more relevant variables may enhance predictive accuracy by better capturing particle transport dynamics, it may also introduce challenges like increased model complexity, higher computational costs, and the risk of overfitting.

In real-world scenarios, inhaled aerosols typically exhibit polydispersity, which may impact the predictive accuracy of statistical models. This study utilizes monodisperse particles to establish a foundational understanding of the model's predictive capabilities across multiple output targets under controlled conditions. Additionally, employing monodisperse particles enables a faster and more precise assessment of deposition mechanisms. By initially focusing on monodisperse particles, the study aims to ensure that the statistical models accurately capture fundamental deposition trends before introducing the complexity of polydisperse aerosols.

In future research, these aforementioned factors should be taken into consideration to validate the universality of the current workflow and generate practical guidance or a well-structured tool for clinicians to apply the findings of this research in clinical treatments and the development of novel medical devices.

Extending similar machine learning approaches to predict particle deposition in nasal airways is also a valuable research direction. Given the anatomical complexity and airflow characteristics of the nasal cavity, incorporating ML models for deposition prediction could provide further insights into drug delivery optimization [36–41].

For the computational fluid dynamics (CFD) part of the research, considering the complexity of healthy lung models and computational resources, the turbulence model for the simulation has been set to the 4-equ SST model, which provides flexibility to quickly generate flow field information for the prediction of particle deposition. However, in future research and applications, more precise turbulence models, such as Detached Eddy Simulation (DES) or Large Eddy Simulation (LES), can be

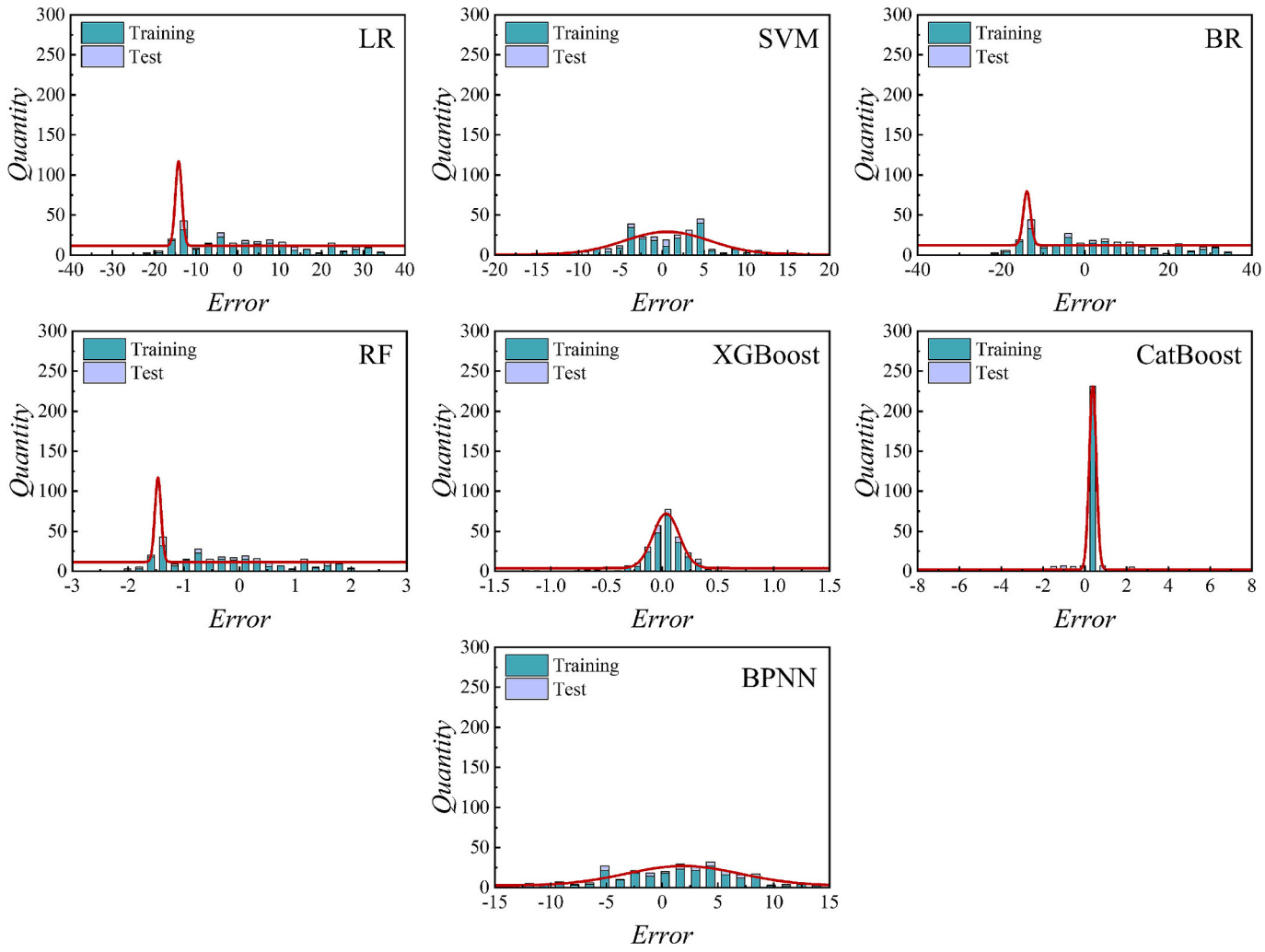


Fig. 13. The error histograms of the advanced models for predicting total deposition rate.

applied, particularly when dealing with the complexity introduced by morphological changes due to diseases such as COPD and lung cancer. The particle release strategies, such as release location and release patterns on the initial release plane, should also be considered, especially when applying relatively large particles whose movement is more influenced by inertia than by the flow.

5. Conclusion

This study investigates the multi-task predictive capabilities of several regression models, including including Linear regression (LR), Bayesian regression (BR), Back Propagation Neural Network (BPNN), Support Vector Machine (SVM), Random Forest (RF), XGBoost, and CatBoost, for predicting both the total and regional deposition rates of inhaled particles in a patient-specific mouth-lung model. The input training features include human respiratory airflow rate, particle diameters, and initial particle velocity. The database is derived from well-validated CFD simulations using the Eulerian-Lagrangian method.

The results show that LR, BR, and SVM cannot fully interpret the complex non-linear relationship of the particle deposition mechanism for both the training and testing datasets, with average R^2 values ranging from 0.21 to 0.73 and MAE reaching nearly 15. For total deposition rate prediction, all other models, including BPNN, RF, CatBoost, and XGBoost, achieve high accuracy, with R^2 values greater than 0.99. However, the BPNN model fails to achieve consistent performance when predicting the regional deposition rate, particularly for oral deposition, where $R^2 = 0.538$. From these comparisons, with such a limited dataset derived from a patient-specific model, decision tree-based models outperform the BPNN models. These decision tree-based models exhibit very small MAE values when predicting total and regional deposition rates, with XGBoost and CatBoost being the most optimal. Notably, for decision tree-based models, the MAE for regional deposition rate prediction consistently remains below 0.15, highlighting their robustness and precision in capturing intricate spatial deposition patterns within the respiratory tract. By visualizing the error histograms, it can be observed that XGBoost performs the best, as most predictive errors are within the range of ±0.5, while CatBoost and RF generally show errors within ±2.

CRediT authorship contribution statement

Xueren Li: Writing – original draft, Visualization, Validation, Formal analysis, Data curation, Conceptualization. **Ruipeng Xu:** Writing – original draft, Visualization, Validation, Formal analysis, Data curation, Conceptualization. **Jiaqi Fan:** Visualization, Validation, Software, Methodology, Formal analysis, Data curation. **Liwei Zhang:** Methodology, Investigation, Formal analysis, Data curation. **Weijie Sun:** Software, Methodology, Formal analysis, Data curation. **Sasa Kenjeres:** Writing – review & editing, Supervision. **Yidan Shang:** Writing – review & editing, Supervision, Resources, Project administration, Conceptualization. **William Yang:** Writing – review & editing, Supervision, Conceptualization.

Declaration of competing interest

The authors declare no conflict of interest.

Acknowledgments

We thank the support of the National Natural Science Foundation of China (Grant No. 82370101). The support provided by China Scholarship Council (CSC) during the study of the author Ruipeng Xu is acknowledged.

Appendix A

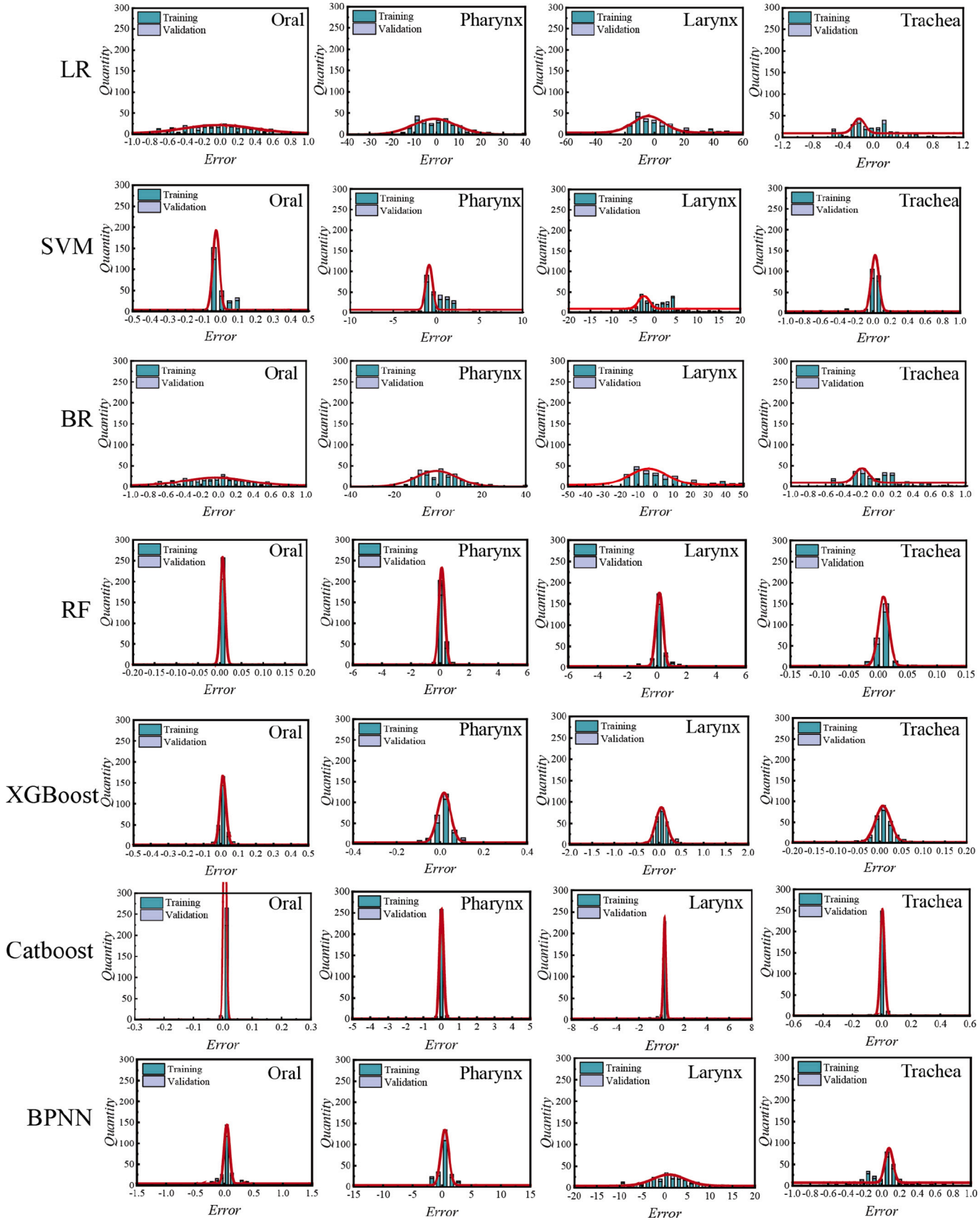


Fig. 14. The error histograms of the advanced models for predicting the upper airway deposition rate.

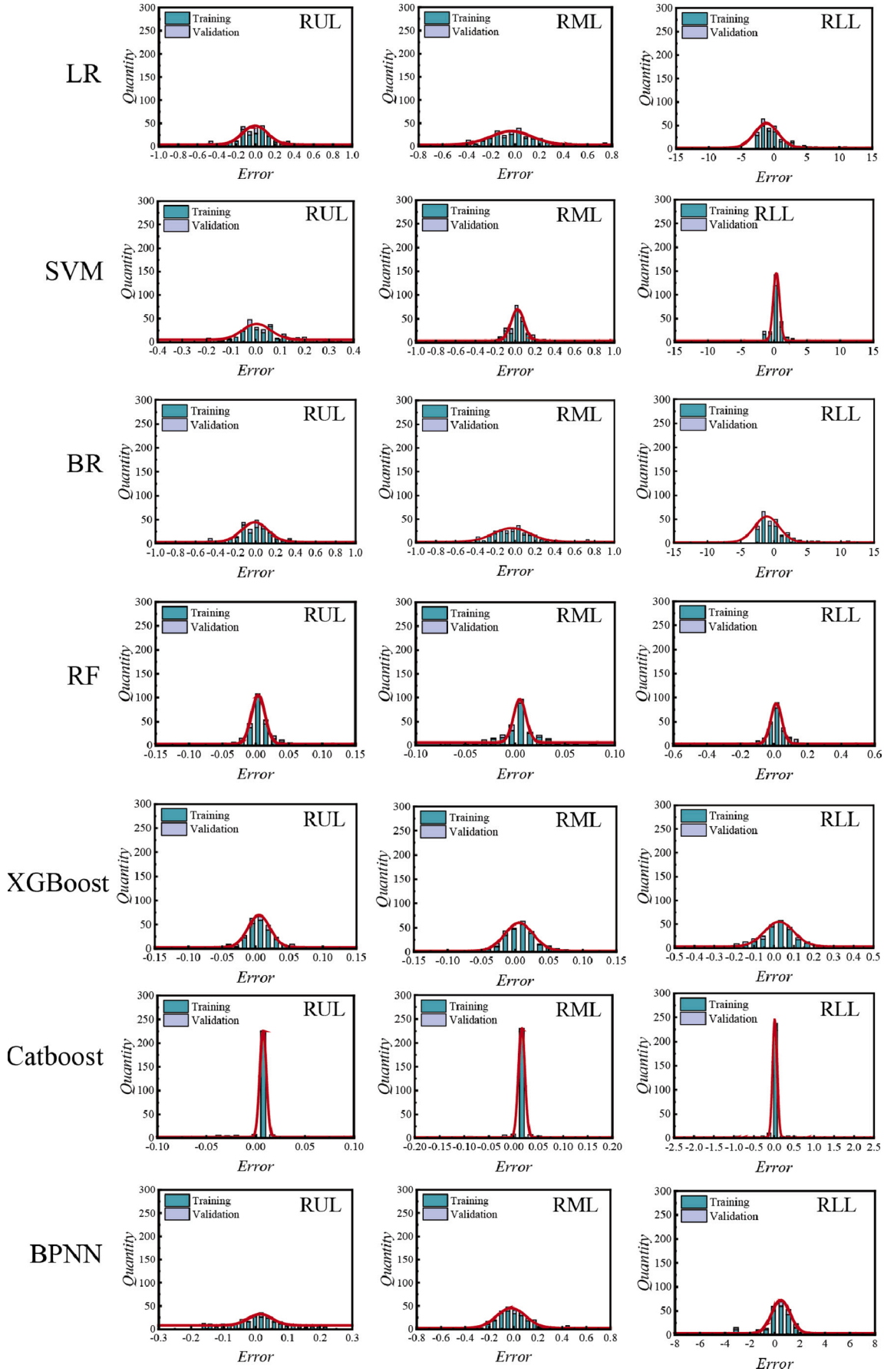


Fig. 15. The error histograms of the advanced models for predicting the right lower airway deposition rate.

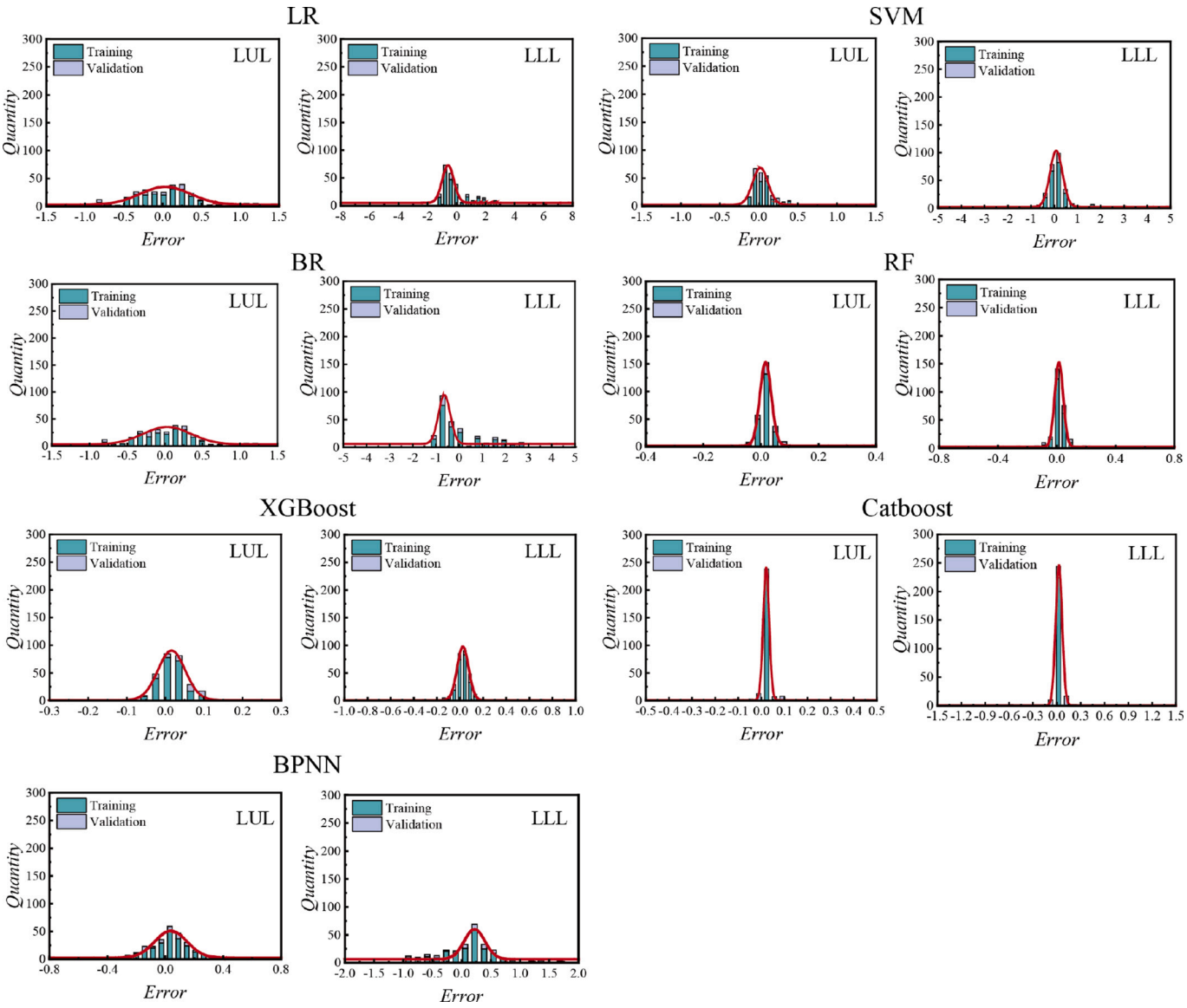


Fig. 16. The error histograms of the advanced models for predicting the left lower airway deposition rate.

Data availability

Data available on request due to privacy/ethical restrictions.

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to patient privacy restrictions.

References

[1] World Health Organization, Ambient (Outdoor) Air Pollution, URL, https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health, 2024.

[2] Yidan Shang, Rui Chen, Ru Bai, Jiyuan Tu, Lin Tian, Quantification of long-term accumulation of inhaled ultrafine particles via human olfactory-brain pathway due to environmental emissions –a pilot study, NanoImpact 22 (2021) 100322, https://doi.org/10.1016/j.impact.2021.100322.

[3] P.J. Barnes, P.G. Burney, E.K. Silverman, B.R. Celli, J. Vestbo, J.A. Wedzicha, E. F. Wouters, Chronic obstructive pulmonary disease, Nat. Rev. Dis. Primers 1 (2015) 15076, https://doi.org/10.1038/nrdp.2015.76.

[4] Tristan Van de Moortele, Ute Goerke, Chris H. Wendt, Filippo Coletti, Airway morphology and inspiratory flow features in the early stages of chronic obstructive pulmonary disease, Clin. Biomech. 66 (2019) 60–65 (ISSN 0268-0033).

[5] Yidan Shang, Jingliang Dong, Lin Tian, Kiao Inthavong, Jiyuan Tu, Detailed computational analysis of flow dynamics in an extended respiratory airway model, Clin. Biomech. (Bristol) 61 (2019) 105–111, https://doi.org/10.1016/j.clinbiomech.2018.12.006. ISSN 0268-0033.

[6] Lixing Zhang, Zhenbo Tong, Ya Zhang, Aibing Yu, Numerical study on the deposition distribution and mechanism of inhaled drug particles in various regions of the realistic inhaler-airway model, Powder Technol. 449 (2025) 120402, https://doi.org/10.1016/j.powtec.2024.120402.

[7] Gang Guo, Lixing Zhang, Hongxian Ren, Yingzhe Ding, Ya Zhang, Zhenbo Tong, Aibing Yu, Nasal drug delivery to olfactory region based on high-flow auxiliary gas method: a numerical simulation study, Powder Technol. (2024) 120567, https://doi.org/10.1016/j.powtec.2024.120567.

[8] D. García-Martínez, N. Torres-Tamayo, I. Torres-Sanchez, F. García-Río, M. Bastir, Morphological and functional implications of sexual dimorphism in the human skeletal thorax, Am. J. Phys. Anthropol. 161 (3) (2016) 467–477, https://doi.org/10.1002/ajpa.23051. ISSN 0002-9483.

[9] K. Inthavong, P. Das, N. Singh, J. Sznitman, In silico approaches to respiratory nasal flows: a review, J. Biomech. 97 (2019) 109434, https://doi.org/10.1016/j.jbiomech.2019.109434.

[10] N. Jahani, S. Choi, J. Choi, K. Iyer, E.A. Hoffman, C.L. Lin, Assessment of regional ventilation and deformation using 4d-ct imaging for healthy human lungs during tidal breathing, J. Appl. Physiol. (1985) 119 (10) (2015) 1064–1074, https://doi.org/10.1152/japplphysiol.00339.2015. ISSN 8750-7587 (Print) 0161-7567.

[11] Stavros C. Kassinos, Josué Sznitman, Multiscale modeling of respiratory transport phenomena and intersubject variability, Annu. Rev. Fluid Mech. 57 (2024)

141–165. https://www.annualreviews.org/content/journals/10.1146/annurev-fluid-031424-103721.
[12] Anurag Tiwari, Anuj Jain, Akshoy R. Paul, Suvash C. Saha, Computational evaluation of drug delivery in human respiratory tract under realistic inhalation, Phys. Fluids 33 (8) (2021) 083311.
[13] Xiuhua A. Si, Mohamed Talaat, Jinxiang Xi, Effects of guiding vanes and orifice jet flow of a metered-dose inhaler on drug dosimetry in human respiratory tract, Exp. Comput. Multiph. Flow 5 (3) (2023) 247–261 (ISSN 2661-8877).
[14] R.E. Gosman, R.M. Sicard, S.M. Cohen, D.O. Frank-Ito, A computational analysis on the impact of multilevel laryngotracheal stenosis on airflow and drug particle dynamics in the upper airway, Exp. Comput. Multiph. Flow 5 (3) (2023) 235–246, https://doi.org/10.1007/s42757-022-0151-9. ISSN 2661-8869 (Print) 2661-8869.
[15] K. Inthavong, Y. Shang, J. Tu, Surface mapping for visualization of wall stresses during inhalation in a human nasal cavity, Respir. Physiol. Neurobiol. 190 (2014) 54–61, https://doi.org/10.1016/j.resp.2013.09.004. ISSN 1569-9048.
[16] Idan Shang, Kiao Inthavong, Numerical assessment of ambient inhaled micron particle deposition in a human nasal cavity, Exp. Comput. Multiph. Flow 1 (2019), https://doi.org/10.1007/s42757-019-0015-0.
[17] Yidan Shang, Lin Tian, Yaming Fan, Jingliang Dong, Kiao Inthavong, Jiyuan Tu, Effect of morphology on nanoparticle transport and deposition in human upper tracheobronchial airways, J. Comput. Multiph. Flows 10 (2) (2018) 83–96, https://doi.org/10.1177/1757482X18756012. ISSN 1757-482X.
[18] Fen Huang, Qixuan Zhu, Xudong Zhou, Dazhao Gou, Jiaqi Yu, Renjie Li, Zhenbo Tong, Runyu Yang, Role of cfd based in silico modelling in establishing an in vitro-in vivo correlation of aerosol deposition in the respiratory tract, Adv. Drug Deliv. Rev. 170 (2021) 369–385, https://doi.org/10.1016/j.addr.2020.09.007. ISSN 0169-409X.
[19] A.J. Banko, F. Coletti, D. Schiavazzi, C.J. Elkins, J.K. Eaton, Three-dimensional inspiratory flow in the upper and central human airways, Exp. Fluids 56 (6) (2015) 117, https://doi.org/10.1007/s00348-015-1966-y. ISSN 1432-1114.
[20] Xueren Li, Yihuan Yan, Xiang Fang, Fajiang He, Jiyuan Tu, Towards understanding of inhalation exposure of pilots in the control cabin environment, Build. Environ. 242 (2023) 110572, https://doi.org/10.1016/j. buildenv.2023.110572.
[21] Xuwen Tong, Jingliang Dong, Yidan Shang, Kiao Inthavong, Jiyuan Tu, Effects of nasal drug delivery device and its orientation on sprayed particle deposition in a realistic human nasal cavity, Comput. Biol. Med. 77 (2016) 40–48, https://doi.org/10.1016/j.compbiomed.2016.08.002. ISSN 0010-4825.
[22] Saša Kenjereš, On recent progress in modelling and simulations of multi-scale transfer of mass, momentum and particles in bio-medical applications, Flow Turbul. Combust. 96 (3) (2016) 837–860, https://doi.org/10.1007/s10494-015-9669-2. ISSN 1573-1987.
[23] Saša Kenjereš, Jimmy Leroy Tjin, Numerical simulations of targeted delivery of magnetic drug aerosols in the human upper and central respiratory system: a validation study, R. Soc. Open Sci. 4 (12) (2017) 170873, https://doi.org/10.1098/rsos.170873.
[24] Yidan Shang, Yao Tao, Jingliang Dong, Fajiang He, Jiyuan Tu, Deposition features of inhaled viral droplets may lead to rapid secondary transmission of covid-19, J. Aerosol Sci. 154 (2021) 105745, https://doi.org/10.1016/j.jaerosci.2021.105745.
[25] Logan G. Wright, Tatsuhiro Onodera, Martin M. Stein, Tianyu Wang, Darren T. Schachter, Zoey Hu, Peter L. McMahon, Deep physical neural networks trained with backpropagation, Nature 601 (7894) (2022) 549–555, https://doi.org/10.1038/s41586-021-04223-6. ISSN 1476-4687.
[26] Barry de Ville, Decision trees, WIREs Comput. Stat. 5 (6) (2013) 448–455, https://doi.org/10.1002/wics.1278. ISSN 1939-5108.
[27] Leo Breiman, Random forests, Mach. Learn. 45 (2001) 5–32 (ISSN 0885-6125).
[28] Tianqi Chen, Carlos Guestrin, Xgboost: a scalable tree boosting system, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016, pp. 785–794.
[29] Liudmila Prokhorenkova, Gleb Gusev, Aleksandr Vorobev, Anna Veronika Dorogush, Andrey Gulin, Catboost: unbiased boosting with categorical features, Adv. Neural Inf. Proces. Syst. 31 (2018).
[30] Xueren Li, Weijie Sun, Chao Qin, Yihuan Yan, Liwei Zhang, Jiyuan Tu, Evaluation of supervised machine learning regression models for cfd-based surrogate modelling in indoor airflow field reconstruction, Build. Environ. 267 (2024) 112173.
[31] Xiaogang Su, Xin Yan, Chih-Ling Tsai, Linear regression, WIREs Comput. Stat. 4 (3) (2012) 275–294, https://doi.org/10.1002/wics.1198. ISSN 1939-5108.
[32] Christopher M. Bishop, Michael E. Tipping, Bayesian Regression and Classification, IOS Press, 2003, pp. 267–285.
[33] Marti A. Hearst, Susan T. Dumais, Edgar Osuna, John Platt, Bernhard Scholkopf, Support vector machines, IEEE Intell. Syst. Applic. 13 (4) (1998) 18–28. ISSN 1094-7167.
[34] Yung Cheng, Yue Zhou, Bean Chen, Particle deposition in a cast of human oral airways, Aerosol Sci. Technol. 31 (1999) 286–300, https://doi.org/10.1080/027868299304165.
[35] Z. Zhang, C. Kleinstreuer, J.F. Donohue, C.S. Kim, Comparison of micro- and nano-size particle depositions in a human upper airway model, J. Aerosol Sci. 36 (2) (2005) 211–233 (ISSN 0021-8502).
[36] Stavros C. Kassinos, Josué Sznitman, Multiscale modeling of respiratory transport phenomena and intersubject variability, Annu. Rev. Fluid Mech. 57 (2024).
[37] Zhiwei Shen, Jingliang Dong, Xinyu Cai, Hanieh Gholizadeh, Hak-Kim Chan, Ann Lee, Agisilaos Kourmatzis, Shaokoon Cheng, Computational study of the impact of nasal vestibule anatomy on nasal drug administration with nasal spray, Int. J. Pharm. 669 (2025) 125086, https://doi.org/10.1016/j.ijpharm.2024.125086.
[38] Jiayi Huang, Ya Zhang, Xiaole Chen, Yueyang Cai, Yi Zhang, Zhenning Jia, Droplet deposition dynamics in a realistic three-dimensionally printed human upper airway: an in vitro study, Phys. Fluids 36 (12) (2024) 123364, https://doi.org/10.1063/5.0238983.
[39] Yu Liu, Xiaole Chen, Jun Xie, Xiaojian Xie, Yi Zhang, Feng Tao, Numerical simulation of high-concentration droplet flow in an idealized mouth–throat airway model in the influence of environmental temperature and humidity, Phys. Fluids 36 (12) (2024) 123361, https://doi.org/10.1063/5.0240899.
[40] Hamideh Hayati, Yu Feng, Xiaole Chen, Emily Kolewe, Catherine Fromen, Prediction of transport, deposition, and resultant immune response of nasal spray vaccine droplets using a cfpd-hcd model in a 6-year-old upper airway geometry to potentially prevent covid-19, Exp. Comput. Multiph. Flow 5 (3) (2023) 272–289, https://doi.org/10.1007/s42757-022-0145-7. ISSN 2661-8877.
[41] Qinyuan Sun, Jingliang Dong, Ya Zhang, Lin Tian, Jiyuan Tu, Numerical study of the effect of nasopharynx airway obstruction on the transport and deposition of nanoparticles in nasal airways, Exp. Comput. Multiph. Flow 4 (4) (2022) 399–408, https://doi.org/10.1007/s42757-022-0143-9. ISSN 2661-8877.