Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Research Paper

Prediction of the spread of Corona-virus carrying droplets in a bus - A computational based artificial intelligence approach

Mehrdad Mesgarpour a, Javad Mohebbi Najm Abad b, Rasool Alizadeh c, Somchai Wongwises a,d, Mohammad Hossein Doranehgard e, Saeidreza Ghaderi f, Nader Karimi g,h,a

a Fluid Mechanics, Thermal Engineering and Multiphase Flow Research Lab. (FUTURE), Department of Mechanical Engineering, Faculty of Engineering, King Mongkut’s University of Technology Thonburi (KMUTT), Bangkok, Bangkok 10140, Thailand
b Department of Computer Engineering, Qucan Branch, Islamic Azad University, Quchan, Iran
c Department of Mechanical Engineering, Quchan Branch, Islamic Azad University, Quchan, Iran
d National Science and Technology Development Agency (NSTDA), Pathum Thani 12120, Thailand
e Department of Civil and Environmental Engineering, School of Mining and Petroleum Engineering, University of Alberta, Edmonton, Alberta T6G 1H9, Canada
f Department of General Surgery, Shahid Modarres Hospital, Shahid Beheshti University of Medical Sciences, Tehran, Iran
g School of Engineering and Materials Science, Queen Mary University of London, London E1 4NS, United Kingdom
h James Watt School of Engineering, University of Glasgow, Glasgow G12 8QQ, United Kingdom

1. Introduction

The spread of Corona-virus via droplets generated by human sneezing and coughing has attracted a significant attention in recent months. It has been shown that this process can enhance the virus transmission by 18 times (Chen, 2020; Meccariello and Gallo, 2020; Diwan et al., 2020; Rockett et al., 2020). Dispersion of the droplets produced by human sneezing is a complex physical process (Enserink and Kupferschmidt, 2020; Ivorra et al., 2020). Upon leaving the mouth, the saliva droplet is broken down into smaller droplets. The subsequent evaporation of droplets and reduction of diameters cause drastic changes in their deposition and suspension. This highly complicates the development of predictive models for spread of such droplets and the subsequent transmission of the virus (Chaudhuri et al., 2020; Abuhegazy et al., 2020).

A considerable amount of work has been already done on modelling of sneezing. In 2019, Hassan and Khorrarammehr (2019) studied the process of transmitting sneezing generated droplets into human airways. The results showed that for an average flow rate of 4.79 L/s, the airflow outlet velocity from the mouth reaches 5.3 and 8.4 m/s, respectively. In 2020, Singh and Tripathi (2020) numerically simulated the dispersion of sneezing particles in a room. Their investigation was complimented by an experiment and put forward a model for measuring the transfer of particles in surrounding air. This was continued by Kotb and Khalil (2020) with a numerical study of the emission of particles containing SARS-COVID 19 virus into the interior of a passenger aircraft. The result indicated that velocity of the droplets produced by the moving passengers could reach the seated passengers. However, sneezing generated droplets had more harmful impacts than those produced by coughing,

ARTICLE INFO

Editor: Dr. R Teresa

Keywords: COVID-19, Droplet distribution, Large eddy simulations, Deep neural network, Prediction

ABSTRACT

Public transport has been identified as high risk as the corona-virus carrying droplets generated by the infected passengers could be distributed to other passengers. Therefore, predicting the patterns of droplet spreading in public transport environment is of primary importance. This paper puts forward a novel computational and artificial intelligence (AI) framework for fast prediction of the spread of droplets produced by a sneezing passenger in a bus. The formation of droplets of saliva is numerically modelled using a volume of fluid methodology applied to the mouth and lips of an infected person during the sneezing process. This is followed by a large eddy simulation of the resultant two phase flow in the vicinity of the person while the effects of droplet evaporation and ventilation in the bus are considered. The results are subsequently fed to an AI tool that employs deep learning to predict the distribution of droplets in the entire volume of the bus. This combined framework is two orders of magnitude faster than the pure computational approach. It is shown that the droplets with diameters less than 250 micrometers are most responsible for the transmission of the virus, as they can travel the entire length of the bus.
while both traveled a long distance inside the cabin.

Most recently, Busco et al. (2020) conducted a comprehensive study of the spread of the virus in the environment and showed that the deposition of different particles is a function of their average diameter. They further argued that environmental effects play a major role in particle dispersion. In another study, Li et al. (2020) investigated the deposition of different particles is a function of their average diameter.

In general, numerical methods provide a convenient framework to study dispersion of the virus carrying droplets. Hence, a number of researchers tried to detect virus transmission routes via numerical simulations (Chen, 2020; Meccariello and Gallo, 2020; Diwan et al., 2020; Rockett et al., 2020; Enserink and Kupferschmidt, 2020; Ivorra et al., 2020; Chaudhuri et al., 2020; Abuhegazy et al., 2020). Table 1 presents an outline of the latest research articles on the applications of computational particle science combined with computational fluid dynamics.

Public transport contributes significantly to transmission of COVID-19 (Abuhegazy et al., 2020). As a result, there is a substantial emphasis on the mitigation of transmission risk in public transport including buses and train carriages. This, in turn, calls for prediction of the spread of the virus in indoor environments. In general, such predictions can be made by the conventional computational analyses. However, the associated computational burden forbids their use in practical risk assessments. As a remedy, the techniques from machine learning can be utilized to reduce the computational load. This will lead to the development of a data-driven model that uses the computationally generated data for a small section of the domain and extrapolates those. The effectiveness of such combined approach has been already demonstrated in the context of propulsion and process engineering (Christodoulou et al., 2020; Alizadeh et al., 2020a; Abad et al., 2020; Alizadeh et al., 2020b). It was shown that data-driven approaches could...
predict complicated spatiotemporal behaviors, while high fidelity computation was performed only for a small fraction of the domain (Christodoulou et al., 2020). Hence, the current study includes a computational part followed by artificial intelligence (AI). In the numerical study, computational fluid dynamics is used to model a three-phase flow of liquid droplets and water vapor in air during the droplet distribution process. This model is based on the changes in the particle diameters encountered in sneezing process and their transmission by air flow. Due to the large gradient in the sneezing process and the changes in the flow Reynolds number, an innovative model is developed for low cost and yet accurate analysis of high and low-speed flows. The numerical results are then fed into an AI-based soft tool, which predicts the droplet distribution inside a bus. This results in a novel combination of computational fluid dynamics and artificial

Table 1
List of the most recent investigations that employed a combination of CFD and particle science.

| No. | Model | Dimensions | Particle/Droplet size (mm) | Domain | Turbulence method | Target | Solver | Ref. |
|-----|-------|------------|-----------------------------|--------|-------------------|--------|--------|------|
| 1   | Unsteady-DPM | 2D | 0.1–0.3 | tube banks | k-ω (SST) | Deposition particle | Ansys Fluent | (Mu et al., 2020) |
| 2   | Unsteady-DPM | 2D | 0.012 | cryogenic condenser | k-ω (SST) | Deposition particle | Ansys Fluent | (Hendry et al., 2019) |
| 3   | Unsteady-Lagrangian particle tracking | 2D | 0.001–0.01 | airway | RANS | Particle flows | Ansys Fluent | (Kim et al., 2019) |
| 4   | Unsteady-DPM | 2D | 0.001–0.05 | ribbed duct air flows | RANS | Particle deposition | Ansys Fluent | (Lu and Lu, 2016) |
| 5   | Unsteady-DPM | 3D | 0.001–11.25 | 90-degree bend of natural gas pipelines | k-ω (SST) | Particle deposition | Ansys Fluent | (Seyfi et al., 2020) |
| 6   | Lagrangian Particle Track | 2D | 0.002–0.058 | cyclone separator | RANS | Particle deposition | Ansys Fluent | (Song et al., 2017) |
| 7   | Unsteady-DPM | 2D | 0.002–0.005 | Nose-to-Lung | k-ω (SST) | Aerosol Delivery | Ansys Fluent | (Sutin et al., 2020) |
| 8   | Unsteady | 3D | – | environment | LES | Breathing Simulations | Ansys Fluent | (Ivanov and Mijorski, 2019) |
| 9   | Unsteady-DPM | 2D | 0.001–0.009 | lung | LES | Aerosol deposition | Ansys Fluent | (Huang et al., 2020) |
| 10  | Unsteady-DEM | 3D | 0–0.002 | inhalation indoors | Monte-Carlo | SARS-CoV-2 transmission | In-house code | (Vuorinen et al., 2020) |
| 11  | Lagrangian method | 2D | – | combustion | RANS | Ash deposition | In-house code | (Yang et al., 2017) |

Fig. 1. The schematic view of a) the investigated bus and sneezing person, b) seat position and key dimensions.
intelligence, which forms an efficient tool for practical risk assessments.

2. Problem definition

Buses are the most common means of public transport throughout the world and have specific features that make them pertinent to transmission of the virus. These include the high surface density of passengers and the lack of strong ventilation. In this study, one of the assumed to be adiabatic. In the proceeding analysis, all fluids are shown in Fig. 1, a standing passenger (1.80 m of height) at the end of the bus is sneezing. The passenger is located on the center-plane of the bus. Steady-state airflow (V = 0.1 m/s) is injected at the end of the bus. This system is closed and stays at constant temperature and humidity (Tambient = 20°C, RH = 20%). All the walls and doors of the bus are assumed to be adiabatic. In the proceeding analysis, all fluids are assumed to be isotropic and Newtonian and, the flow is incompressible.

3. Governing equation

The multiphase flow in this study contains liquid, gas and vapor phases. To model this complex, three phase flow, the conservation of mass and momentum in the Eulerian framework for any m fluid is written as (Yeoh and Tu, 2019; Kolev and Kolev, 2007; Crowe et al., 2011):

\[
\frac{\partial (\alpha m \rho m \text{u} m)}{\partial t} + \frac{\partial (\alpha g \rho g \text{u} g)}{\partial x} + \frac{\partial (\alpha v \rho v \text{u} v)}{\partial y} + \frac{\partial (\alpha w \rho w \text{u} w)}{\partial z} = \sum_{i=1}^{N} \Gamma_{i} \quad (1)
\]

where \( \rho \) and \( \text{u} \) are the volume fraction, denoted by \( \alpha \) and the mixture velocity and enthalpy for the phase and mass-weighted, respectively, and the latter is defined as \( \rho \rho m \text{u} m \).

\[
\sum_{i=1}^{N} \sum_{j=1}^{N} \rho m \text{u} m \text{u} m \text{u} j \text{w} j = \sum_{i=1}^{N} \rho m \text{u} m \text{u} m \text{w} j + \sum_{j=1}^{N} \rho m \text{u} m \text{w} j \text{u} j
\]

In this equation, \( t \), is time and \( \chi, \gamma, \zeta \) are the spatial coordinates. Transport of momentum in \( \chi, \gamma, \zeta \) directions:

\[
\frac{\partial (\alpha m \rho m \text{u} \alpha m \text{u} m, \alpha m \text{u} m, \alpha m \text{u} m)}{\partial t} + \frac{\partial (\alpha g \rho g \text{u} \alpha g \text{u} g, \alpha g \text{u} g, \alpha g \text{u} g)}{\partial x} + \frac{\partial (\alpha v \rho v \text{u} \alpha v \text{u} v, \alpha v \text{u} v, \alpha v \text{u} v)}{\partial y} + \frac{\partial (\alpha w \rho w \text{u} \alpha w \text{u} w, \alpha w \text{u} w, \alpha w \text{u} w)}{\partial z} = \sum_{i=1}^{N} \Gamma_{i} \quad (2)
\]

\[
\frac{\partial (\alpha m \rho m \text{u} \alpha m \text{u} m, \alpha m \text{u} m, \alpha m \text{u} m)}{\partial t} + \frac{\partial (\alpha g \rho g \text{u} \alpha g \text{u} g, \alpha g \text{u} g, \alpha g \text{u} g)}{\partial x} + \frac{\partial (\alpha v \rho v \text{u} \alpha v \text{u} v, \alpha v \text{u} v, \alpha v \text{u} v)}{\partial y} + \frac{\partial (\alpha w \rho w \text{u} \alpha w \text{u} w, \alpha w \text{u} w, \alpha w \text{u} w)}{\partial z} = \sum_{i=1}^{N} \Gamma_{i} \quad (3)
\]

Furthermore, \( \mu^T \) denotes turbulent viscosity of the continuous N\textsuperscript{th} phase and \( \rho m \text{u} m, \rho g \text{u} g, \rho v \text{u} v, \rho w \text{u} w \) are the source terms. The energy equations for the multiphase model (mixture) are:

\[
\frac{\partial (\rho m \text{u} m \text{H} m)}{\partial t} + \nabla \cdot (\rho m \text{u} m \text{u} m \text{H} m) = \nabla \cdot (\rho m \text{u} m \text{u} m \text{H} m) - \nabla \cdot (\rho m \text{u} m \text{u} m \text{H} m) + \sum_{i=1}^{N} (\alpha v \rho v \text{u} v \text{H} v m)
\]

Here, \( \text{H} m, \zeta, \text{U} \text{u} m \) are enthalpy by the combination of phase-weighted, relaxation factor (functional) and drift velocity vector of the N\textsuperscript{th} phase, respectively, and the latter is defined as \( \text{U} \text{u} m = \text{U} \text{u} m - \text{U} \text{u} m \).

The combination of mixture velocity and enthalpy for the phase and mass-weighted variables leads to:

\[
\text{U} \text{u} m = \sum_{i=1}^{N} (\alpha v \rho v \text{u} v) \quad \text{H} m = \sum_{i=1}^{N} (\alpha v \rho v \text{u} v)
\]

3.1. Inter-phase mass transfer

In the vaporization process, if the kinetic energy of droplets and the work of viscosity term are neglected, the transfer of mass on the interface of the droplets and vapor can be reduced to (Ivorra et al., 2020):

\[
G m = \frac{\sum_{i=1}^{N} (\eta m \cdot \chi)}{h m} \quad (7)
\]

In Eq. (7), \( h m \) is the enthalpy of vaporization and \( \eta m \) is the heat flux for N\textsuperscript{th} fluid. This equation can be defined as the net heat transport to the interface of droplet divided by the latent heat of vaporization process.

3.2. Large eddy simulation (LES)

In LES, the large energy-containing eddies are computed directly (within the accuracy of the computational scheme) and the small-scale structures are modeled. For the bulk multiphase flow, the mathematical formulations consist of a continuous and discrete phase. The Favre-averaged transport equations of mass and momentum utilized in LES are given by (Vuorinen et al., 2020):

\[
\frac{\partial (\alpha m \rho m \text{u} m \text{u} m)}{\partial t} + \frac{\partial (\alpha g \rho g \text{u} g \text{u} g)}{\partial x} + \frac{\partial (\alpha v \rho v \text{u} v \text{u} v)}{\partial y} + \frac{\partial (\alpha w \rho w \text{u} w \text{u} w)}{\partial z} = \sum_{i=1}^{N} \Gamma_{i} \quad (8)
\]

In Eq. (9), \( S \) denotes the body forces acting on the fluid. The subgrid-scale (SGS) stress tensor \( (\tau) \) is modeled through an eddy-viscosity approach

\[
\tau_{ij} = \rho (\tau_{ij} - \tau_{ij})
\]

\[
S_{ij} = \frac{1}{2} \left( \frac{\partial (\tau_{ij})}{\partial x_{j}} \right) \quad (10)
\]

\( S_{ij} \) is the rate of strain tensor for the resolved scale. Here, the so-called 'one equation eddy-viscosity model' (OEEVM) subgrid-scale (SGS) was employed. To calculate the turbulent kinetic energy (K), the OEEVM was solved as follows.

\[
\frac{\partial (\kappa k)}{\partial t} + \nabla \cdot (\kappa k \text{u} m) = - \tau_{ij} \cdot S_{ij} + \nabla \cdot (\kappa \mu k) + \kappa
\]

\( \kappa = c_s k^{3/2} / \Delta \quad (12) \)

The kinetic energy in LES turbulence method is define as:
\[
\frac{\partial \langle u_i u_j \rangle^*}{\partial t} = u_i \frac{\partial \langle f_{ij} \rangle^*}{\partial x_j} + \frac{\partial \langle f_{ij} \rangle^*}{\partial x_i} + \frac{\partial}{\partial x_j} \left( \frac{1}{\tau_p} \right) \left( \langle U_j \rangle^* - u_j \right) \langle f_{ij} \rangle^* = -\frac{1}{\tau_p} \left( \frac{\partial \langle u_i u_j \rangle^*}{\partial u_i} \right) 
\]

(13)

where \( u_f^* = (U_f/\rho_f) - (U_f/\rho_f) \). Using Boussinesq approximation, a simple correlation for the incompressible Reynolds stress tensor is developed, which is expressed by

\[
\overline{u_i u_j} = 2 \sqrt{3\kappa_0 - \nu_f} \left( \overline{u_i}^2 + \overline{u_j}^2 \right) 
\]

(14)

### 3.3. Droplet evaporation process

The evaporation of droplets is modeled using (Chen, 2020; Meccariello and Gallo, 2020; Diwan et al., 2020; Rockett et al., 2020; Enserink and Kupferschmidt, 2020; Ivorra et al., 2020; Chaudhuri et al., 2020; Abuhagezy et al., 2020):

\[
J_v = 1 \times 10^{-4} (20.56 + 27.21V_u + 6.39V_d^2) (1 \times 10^{-3} (P_a - P_w)) \frac{b}{9} 
\]

(15)

In this regard, surface evaporation depends on \( V_u, P_a \) and \( P_w \), which depend on the air velocity, partial saturated vapor pressure, and partial water vapor pressure in air, respectively.

### 3.4. Droplet merging

The most common way to achieve the integration of droplets is to ensure that there is an electric field where the two droplets meet. Link et al. (2006) showed that there are direct relations between voltage and rate of merging. This results in oppositely charged surfaces, which will merge the particles together. Also, Chabert et al. studied the merging of individual droplets and particle pairs via electro coalescence (EC) (Chabert et al., 2005). For Newtonian and incompressible fluids at low Reynolds and Bond numbers, this process is defined as (Niu et al., 2008):

\[
\mu \nabla^2 U - \nabla p = \nabla \theta 
\]

(16)

with \( \mu \) being the viscosity of fluid, \( U \) the velocity in X direction, \( P \) the total pressure and \( \theta \) the stress field in each phase of the two-phase fluid.

### 3.5. Boundary conditions

#### 3.5.1. Sneezing profile

As shown in Fig. 2, sneezing involves a sudden change in pressure. The details of this pressure variation depends upon the person’s gender, age and other physiological features. A fluid flow under the pressure trace shown in Fig. 2 is ejected from the mouth. The inner surface of the mouth is a wall and the throat is modelled as a pressure inlet boundary condition. The fluid flow has a relative humidity of 75% and temperature and other physiological features. A fluid flow under the pressure trace shown in Fig. 2 is ejected from the mouth. The inner surface of the mouth is a wall and the throat is modelled as a pressure inlet boundary condition. The fluid flow has a relative humidity of 75% and temperature and other physiological features. A fluid flow under the pressure trace shown in Fig. 2 is ejected from the mouth. The inner surface of the mouth is a wall and the throat is modelled as a pressure inlet boundary condition. The fluid flow has a relative humidity of 75% and temperature and other physiological features. A fluid flow under the pressure trace shown in Fig. 2 is ejected from the mouth. The inner surface of the mouth is a wall and the throat is modelled as a pressure inlet boundary condition.

### 3.5.2. Droplet

For each droplet, the shape is affected by the balance of internal and external forces. According to Han et al. (2013) the maximum rate of the volume-based size distributions can be represented by Yeoh and Tu (2019):
In the second part, starting from the mouth and extending for 1 m, LES was utilized to model the spray and dispersion of droplets of saliva. Due to the relatively high velocity of the flow in this section and the presence of vortices, a full structure grid generation has been developed. Included in the model is evaporation and heat transfer within the droplet and the possibility of droplets merging. It is essential to note that unlike most existing studies on modeling of sneezing, here no model of spray or ejector was used. Instead, the entire process of two-phase flow formation and gradually turns into a two-phase flow.

4. Artificial Intelligence

To predict of the concentration and velocity, a Multi-Input Multi-Output (MIMO) model is developed. In each experiment, several inputs including time, concentration and velocity are given to the model to calculate concentration outputs. The number of input samples in each experiment is about 1240. For prediction, a Deep Neural Network (DNN) regression with a number of input, hidden and output layers was constructed (Schmidt-Hieber, 2020). One superiority of this model is achieving high accuracy on complex data. One of its main challenges is the existence of a weak structural theory (Du and Xu, 2017). In other words, there is no certain theory for the arrangement of elements in a network. Often, the model is studied experimentally and by applying

\[
F_{\text{drag}} = \frac{1}{8} \rho_f d^2 C_D (V_{fr}^2 - V_{in}^2) |V_{fr} - V_{in}|
\]

\[
F_{\text{Added}} = K \rho_f V_{fr} d (V_{fr}^2 - V_{in}^2) \frac{dt}{dt}
\]

\[
F_{\text{Body}} = \rho_f V_{fr} g
\]

where \( \rho_f \), \( dp \), \( V_{fr} \), \( C_D \), \( V_{in} \) are the fluid density, particulate diameter, volume of the second phase for a spherical particulate, partial surface pressure of the vapor and surface vapor mass concentration, respectively. Further, \( V_{fr} \) is volume of the second phase (the initial state) and \( V_{in} \) denotes the volume of fluid (the initial state).

Further, a comparison between the current numerical results and those from several well-established computational models and empirical correlations have been presented in Fig. 4. An excellent agreement is evident confirming validity of the numerical simulations.

3.7. Concentration and particle diameter measurement (PDM)

Fig. 3 illustrates the method applied to calculate the average value of the droplets in each time step. Based on this, one computational domain for each droplet around the virtual center of mass was considered. This computational domain was fitted to the surface of droplet. The fitting process was based on the adaptive grid generation. A number of vertical vectors on the surface were drawn at any point with the same distribution on the surface, the number of which increased or decreased depending on the particle diameter. The average value of droplet diameter was obtained at any given time in all ranges. The values calculated at each time step were then averaged from all three axes. The amount of virus transmitted by each droplet is directly related to the particle diameter. Therefore, the study of changes and classification of droplets based on their average diameter is an important part of calculating the distribution of droplets in the range of 1–1000 μm. This specific model examines the diameter changes at each time step, which vary by the internal and external forces and evaporation, and provides an ability to statistically evaluate the diameter of the droplets. The developed scheme further offers the ability to calculate the number of particles in a constant volume through applying an image processing technique.

3.8. Validation

Extensive tests confirmed the grid independency of the simulations. Further, a comparison between the current numerical results and those from several well-established computational models and empirical correlations have been presented in Fig. 4. An excellent agreement is evident confirming validity of the numerical simulations.
some techniques such as changing the number of layers and neurons and testing different activation functions the accuracy of the model is improved. The sequential models used in this paper have been developed by importing sequential functions from the Keras library of Python. By using the Dense function of the Keras library, 5 layers have been inserted to the network. The first layer is used as the input layer to receive attributes values. Model attributes are: time, location, pressure, temperature and density. In the proposed model, three hidden layers are used, which include 30, 20 and 15 neurons, respectively. The number of neurons and the type of activation functions were selected by several experiments.

Of course, in selecting the appropriate structure for the network, it is useful to use the proposed structures in models that have already been used for DNN regression. The ReLU activation function is used for the neurons of hidden layer, which is one of the most widely utilized and simplest nonlinear activation functions. The two advantages of this function are that it is not saturated with a large number of inputs and the errors are simply back propagate (Schmidt-Hieber, 2020). The last layer of the network is the output layer, which contains some neurons. These neurons return the result of the prediction of concentration and velocity. The defect of the ReLU activation function is that it only takes values for values greater than zero. Thus, the Sigmoid activation function is used in the output layer. Other settings made on the model are the selection of epoch value and batch size. An epoch is a count of the number of times all training data are used once to update neural weights. For batch training, all training samples go through the training algorithm simultaneously at an epoch before the weights are updated. The model weights is updated at the end of each epoch and the batch size. Adam

Fig. 4. Validation for numerical method base on a) Ranz-marshal equation (Ranz and Marshall, 1952), b) Hamey evaporation model (Hamey, 1982), c) Busco particle distribution (Busco et al., 2020), d) mass profile (Busco et al., 2020).
optimizer is also used to calculate the adaptive learning rate for each parameter. This optimizer has been utilized to update the neuron weights of hidden layer.

Finally, the steps of creating the proposed model are presented. At first, the provided dataset is loaded. Then, the data preprocessing is done, which includes data normalization. Then, the sequential model is created. Input, hidden, and output layers are created using the ReLU and Sigmoid activation functions. Afterwards, model compilation is done using Adam optimizer. The epoch value and batch size are determined for the model. Subsequently, the concentration and velocity prediction is done with the model. Finally, the model is evaluated. Tensor flow has been used as a backend to create a predictive model based on deep neural network. It is a library written for high-performance numerical calculations in Python. The Sequential Model API has also been developed and evaluated with Keras and Python3 libraries. Keras is a Python library developed for deep learning. The Sequential model API is utilized to create a model for predicting the concentration and velocity and to which the model layers are added. Also, the Keras Regressor was used to evaluate the model. Performance evaluation was performed using tenfold cross-validation. In this method, the samples are divided into 10 categories. In each step, 9 categories are used to train the model and 1 category is used to test the model. The epoch and batch size values are set to 200 and 5, respectively.

The accuracy of the model is calculated with the criteria of Mean Square Error (MSE) and Mean Absolute Error (MAE). These criteria are calculated as follow:

\[
MAE = \frac{\sum_{i=1}^{n} |y_i - t_i|}{n}
\]

\[
MSE = \frac{\sum_{i=1}^{n} (y_i - t_i)^2}{n}
\]

in which \(y_i\) and \(t_i\) are prediction and target values, respectively and \(n\) is the total number of data points. Several nonlinear models have been evaluated to select the appropriate model for concentration and velocity prediction. For this purpose, several well-known neural networks including Multi-Layer Perceptron (MLP), Support Vector Regression (SVR) and Radial Basis Function (RBF) are compared to DNN. The results of this comparisons are shown in Table 2. In this evaluation, two criteria 1 and 2 have been considered for each network. Based on the results, the DNN with a significant difference represents a better accuracy than other models. After that, RBF can be used as an appropriate choice to predict concentration and velocity.

5. Characteristics of sneezing process

In the sneezing process, the droplets with larger diameter and thus heavier weight settle under gravity and deposit on the surfaces. The rest of droplets are dispersed at longer distances from the sneezing point. This longitudinal distance depends on the ambient air temperature and the initial velocity of the droplets. Therefore, the spatiotemporal distribution of droplets generated by sneezing is dominated by the environmental parameters and the sneezing parameters. It is known that the concentration of droplets and the exposure time with the contaminated environment are important factors in determining the transmission risk of COVID-19 (Chen, 2020; Kotb and Khalil, 2020; Han et al., 2013; Burke, 2012; Pendar and Páscoa, 2020). Hence, in this section the effects of pertinent parameters influencing the dispersion of particles are examined.

Fig. 5 shows the parameters that characterize the employed sneezing model. These parameters were extracted from the latest literature on the distribution of sneezing droplets (Chen, 2020; Meccariello and Gallo, 2020; Diwan et al., 2020; Rockett et al., 2020; Enserink and Kupferschmidt, 2020; Ivorra et al., 2020). In general, the dispersion process of the sneezing droplets leads to the formation of three different regions. These are separable based on the rate of deposition and are marked in different colors in Fig. 5. In the red zone, particles with large mass fell quickly and hit the ground at a short distance from the sneezing person. The percentage of these particles is low. Yet, due to their considerable mass, they can carry a large amount of virus. Conduction of image processing on the numerical results revealed that distribution pattern of this zone is different for each person. The flow velocity in this zone varies between 50 and 100 m/s (Chen, 2020; Meccariello and Gallo, 2020; Diwan et al., 2020). The second zone, which has a lower concentration of particles compared to the first zone, is marked in orange in Fig. 5. This area, as presented in the following sections, has the largest particle diameter relative to the longitudinal distance and thus the highest risk of transmission. Zone three, marked in blue in Fig. 5, is of particular importance. Although this zone has a low density of particles, it is a very large and hence important area. The extent of this zone is because of the suspension of particles, which brings down their sedimentation rate and enables them to survive for a long time. Most importantly, these suspended particles can readily travel with the air flow and spread the virus vastly.

6. Results and discussion

6.1. Sneezing dynamics

Sneezing is a fully automated process and a natural reaction to remove any contaminants from the lungs. This process travels as a pressure wave from the lungs to the mouth (Li et al., 2020; Burke, 2012; Kitta et al., 2016). This high-pressure flow, once expelled from human mouse and nose, distributes saliva in the form of droplets in the environment. The generated flow changes the size and diameter of these droplets. Modeling of these changes during the sneezing process is one of the most important features of the current work. Previous studies often used particle distribution models based on the distribution of spherical particles (Han et al., 2015; Dan et al., 2019; Song et al., 2020; Jiang et al., 2018). This prevents examination of the changes in droplet diameter during the sneezing process. Further, since movement within the air-fluid causes deformation of the droplets, the drag force caused by the deformation is not constant and this has a significant effect on the droplet distribution.

Fig. 6 shows a typical view of the process of breaking droplets during the sneezing process in the vicinity of human mouth. This analysis was conducted by image processing applied on the LES results. It shows that large droplets fall into small particles within a short distance. It also shows that the concentration of saliva in the unit volume of the densely packed area is much higher than other areas. Fig. 6 further indicates that the break and splitting of large droplets due to the internal forces and drag forces reduce the concentration in the constant volume. This, in

| Table 2 | Mean Square Error (MSE) and Mean Absolute Error (MAE) to predict concentration for various potential models. |
|--------|----------------------------------------------------------------------------------------------------------------------------------|
| Mean Error | DNN | CFD error (%) | MLP | CFD error (%) | SVR | CFD error (%) | RBF | CFD error (%) |
| MAE | $6.72 \times 10^{-6}$ | 0.154 | $2.92 \times 10^{-5}$ | 12.36 | 1.03 $\times 10^{-5}$ | 8.45 | 1.72 $\times 10^{-5}$ | 0.369 |
| MSE | $3.62 \times 10^{-10}$ | 0.266 | $1.82 \times 10^{-8}$ | 37.85 | 4.13 $\times 10^{-9}$ | 5.44 | 9.10 $\times 10^{-10}$ | 1.254 |
turn, leads to the expansion of the contaminated space with sneezing droplets. The spread of virus in the process of sneezing depends on various parameters such as speed and direction, as well as the virus concentration in saliva. Further, it has been shown that the physical properties of saliva varies amongst different people (Ericsson and Stjernström, 1951; Briedis et al., 1980).

The statistics of droplet diameter is an essential outcome of the current analysis. It is already evident that there exists a correlation between the droplet diameter and the possibility of infection transmission (Chen, 2020; Kotb and Khalil, 2020; Busco et al., 2020; Li et al., 2020). The changes in diameter depends on the evaporation rate, flow field and the physical properties of saliva. Fig. 7 shows the statistical distribution of droplets based on the average diameter at two different instances of time. The results clearly show that the droplet diameter is strongly time dependent. Changes in the droplet diameter directly influence the droplet distribution. As such, there is a direct relation between droplet diameter and deposition rate. Also, increasing the rate of evaporation may help reducing the concentration of droplets near the sneezing person. Particles with diameters of 1–250 micrometers are more likely to become suspended. These particles are not transported by air flow to far distances from the sneezing person.

As shown in Fig. 7, it is also clear that the distribution of particles in the transversal direction within the range of 1–100 µm is the greatest. The issue of sneezing and its spread in the investigated environment (inside the bus) has some unique features. First, the infected person sneezes at the end of the bus and the droplets are spread out by natural dispersion as well as the air flow set by the ventilation system. Second, the ambient temperature influences the particle density by altering the evaporation rate. Further, it has been shown that the evaporation rate has an inverse relationship with the droplet diameter and decreases with increasing the droplet diameter (Hasan, 2020). Fig. 7 further shows the effects of temporal changes on different droplet diameters. It is important to note that during the sneezing process, the time and size of the droplets play a major role in the spread of the virus.

6.2. Droplet dispersion in the bus-CFD approach

In general, there are two important factors dominating the probability of infection: exposure time and concentration of virus (Ueba, 1978; Gerba, 1984; Jang et al., 2014). Sneezing process involves a wide range of droplet diameter with different velocity. According to the weight and diameter, each particle has a specific deposition velocity. It follows that the spatial analysis of particle diameter and concentration can help assessing the risk of infection.

Fig. 8 shows the distribution of droplets within the range of two meters from the person. The classification of regions in Fig. 8a is on the basis of the downward velocity of the droplets which drives the sedimentation process. A volume of 2 m³ was considered in front of the sneezing person and a series of image processing techniques were employed to identify the droplet velocity. This led to identification of three spatial regions. In the green region, the ensemble average of the downward velocity of droplets is less than 0.001 m/s. This quantity increases to 0.1 m/s and 0.5 m/s in the blue and brown regions. As shown in Fig. 8a, the affected area of sneezing droplets can generally be
Fig. 6. Distribution and breakup of the saliva droplets in the vicinity of the sneezing person ($T_{inf}=25^\circ\text{C}$, Velocity=90 m/s, number of droplets = 5,000,000, $t = 0.1$ s).
Fig. 7. Frequency of droplet diameter at X = 2 m and for $T_{inf} = 25^\circ$C.

The results show that the droplet diffusion is directly related to the droplet size. It is known that increasing the diameter of the droplets modifies the distribution model and alters the maximum deposition location. Fig. 9 shows that for the droplets with diameters between 1 and 250 $\mu$m, there exists a very small amount of sedimentation. That is to say that the droplets of this size practically do not sediment and remain suspended and therefore can be easily transferred along the bus. These droplets are the primary suspect of the infection parameter. As an important point, Fig. 9 further shows that the ventilation system can reduce the droplet concentration by removing the droplets. Some of these droplets, with the diameters between 250 and 500 $\mu$m, may deposit. Other droplets can be still transferred along the bus and contribute to transmission of the virus.

By increasing the droplet diameter, the deposition rate is expected to increase. According to Fig. 10, the droplets with diameters of 500–750 $\mu$m can deposit quickly within a short distance from the sneezing person. The results show that for this range of droplet diameter, there is a large concentration gradient along the bus. Overall, the simulations indicate that the droplets with diameters of 1–250 $\mu$m have the highest diffusion speed amongst all generated droplets. This is because of their small size and their ability to be suspended in air. According to Fig. 10, for the droplets with the diameter of 1–250 $\mu$m at a distance of

\[
\text{concentration (\%)} = \left( \frac{\text{number of particle (selected range)}}{\text{number of particle (selected range) at } t = 0.1 \sin C \text{ V_{close the face}}} \right)
\]

Fig. 8a and 8b, depicts the top view of droplet distribution in the bus. It is inferred from this figure that the cross-section of the particle transport may change the areas of accumulation of heavy particles. The diffusion environment varies based on the position in the three-dimensional space, the distance between the droplet and the sneezing location for different models. The results further indicate that according to the particle mass, the path of a particle can be different. Also, by comparing Figs. 8a and b, it is clear that the process of transfer and deposition of particles is strongly dependent on the time and environmental conditions. The process of distributing the sneezing particles in the bus is statistically affected by the environmental conditions.

The deposition process and distribution of droplets were studied computationally up to 2 m from the sneezing person. Fig. 9 shows the contours of droplet concentration drawn on three planes at the distances of 0.5, 1 and 2 m from the sneezing person. In order to compare the droplet concentration, Fig. 9 is based on the percentage of droplets per unit volume (number of particles or droplets in standard control volume by considering the evaporated mass of droplets).

The deposition rate of the particles and the time they remain suspended depend directly on the particle diameter. Regardless of the resultant forces, the droplet diameter has a direct effect on the deposition time. Also, due to the changes in diameter along the longitudinal direction of the bus, it can be stated that the droplet deposition time increases directly. Expectedly, heavy droplets deposit at a faster rate than those with a smaller mass.

Fig. 8b shows the distribution of droplets from the top view. It is clear that the pattern of transverse distribution varies according to the location of the individual as well as other environmental parameters. Heavy particles are scattered quickly and in front of the sneezing person due to their high mass and inertia, these particles have less deviation from the axial direction of sneezing. This figure further shows that in the areas with low density (green) particles are more prone to be affected by the suction and ventilation system of the bus.
Fig. 8. Image processing detection of the distribution of selected droplets at $t = 0.1\text{s}$, $t = 0.5\text{s}$ and $t = 2\text{s}$ a) side view b) top view ($T_{\text{inf}}=25^\circ\text{C}$).
Fig. 9. Spatial distribution of the droplet concentration in 2 m distance from the sneezing person for a) 1–250 µm, b) 250–500 µm, c) 500–750 µm, d) 750–1000 µm droplet diameter at t = 5 s after sneezing.
Fig. 10. Non-dimensional velocity for the droplets distribution within 2 m distance from the sneezing person for a) 1–250 µm, b) 250–500 µm, c) 500–750 µm, d) 750–1000 µm droplet diameter at t = 5 s after sneezing.
two meters, the velocity changes are greater than those of other droplets. Also, the results in Fig. 10 show that with increasing the diameter of the droplets, their average velocity decreases by a factor three, which clearly shows the importance of spreading the droplets with small diameters.

6.3. Droplet dispersion in the bus-AI approach

The last section demonstrated that the computational model can provide an insight into the spread of virus in the selected environment (bus). Nonetheless, the computational burden of the model poses a serious practical issue. Rapid computation for risk assessment purposes is almost impossible and therefore alternative approaches should be sought for the prediction of particle distribution. The use of AI can reduce the computational cost of the analysis and increase the processing speed up to almost 100 times. Here, the computationally generated data are used to train an AI tool (see Section 4) that predicts the temporal and spatial evolution of the particle distribution in the bus. In Fig. 11, the results obtained from the AI model and those produced through numerical simulations are presented for four different areas. A comparison between the two sets of data confirms the accuracy of the developed AI tool to predict the droplet distribution. The excellent

![Fig. 11. Comparison between the CFD results and AI predictions at different locations along the bus.](image-url)
Fig. 12. Distribution of the droplet concentration in 18 m distance from the sneezing person for a) 1–250 μm, b) 250–500 μm, c) 500–750 μm, d) 750–1000 μm droplet diameter at t = 10 s after sneezing.
agreement between the computationally and AI generated data means that artificial intelligence can predict the concentration in the entire volume of the bus. As shown in Fig. 10, these values are for the middle plane of the bus and the midline on this plane. It is important to note that with increasing the computational accuracy in region 1, fewer changes for the data obtained from artificial intelligence are encountered.

In Fig. 12, the concentration of particles has been plotted based on their diameter distribution in the whole volume of the bus over 9 different parallel planes with increasing distance from the sneezing person. It is recalled that computational modeling was performed only for the first two meters away from the sneezing person and the rest of the domain was predicted by the AI tool. In keeping with the computational result, the AI results show that smaller diameter particles can propagate widely throughout the bus. As farther distance is taken from the sneezing person, a decrease in the concentration of particles and a gradual deposition of particles in different parts are observed. The data presented in this figure are based on the Acen emission analysis model at the distance of two meters. In producing those data, the maximum particle density was plotted on each plane for 10 s. Also, the presence of ventilation in the bus and the upper part of the passengers’ heads can change the pattern of the particle distribution. It is further observed that the particles with small diameters are readily transported throughout the bus. Also, heavy particles evaporate and turn into smaller particles. The presence of flow obstacles such as passenger seats impart a major effect on the distribution and suspension of particles as well as sedimentation of the settling particles. This is because the air flow is very small in between the seats. Thus, the sedimentation model and the deposition of particles are modified in the space between the seats. The results show that artificial intelligence has the capability of predicting the temporal and spatial distribution of particles in the complex and highly varying environment of the bus. Since AI analysis is much faster (around 100 times) than the corresponding computational analysis, it can be used in practice to evaluate the spread of droplets and the associated risk of infection. An overall view of the temporal spread of droplets within the bus can be found in the companion animation file. In there, the data on the first 2 m from the sneezing passenger have been extracted from the CFD analysis and the spread of droplets in the rest of the bus was predicted by the AI tool.

7. Conclusions

Protection against COVID-19 and establishment of relevant risk assessment schemes are currently central to the health and wellbeing of the society. Understanding the patterns of virus spread and their pertinent parameters are amongst the key issues raised in the recent months. Importantly, there is a pressing need for the quick assessment of transmission risk in indoor environments with large occupancy. In particular, public transport has been identified as one of the main routes to virus transmission. This poses a major challenge on the conventional simulations based on computational fluid dynamics, as the required computational time makes them impractical for risk assessment. To address this issue, an attempt was made to provide a novel approach to the prediction of droplet distributions set by sneezing of an infected person in a bus. This was based on the high fidelity computation of the droplets formation during the sneezing process followed by the droplets dispersion in the domain up to 2 m away from the sneezing person. The processes of droplets formation and spread were modeled through applying the volume of fluid and LES methodologies to the person’s mouth and the immediate surroundings. The resultant computational data were then used to develop an AI-based tool capable of predicting the evolution of droplet distribution in the entire volume of the bus. It was shown that while the AI-based approach reduces the computational time most significantly (~100 times), it offers an excellent accuracy. Hence, the developed combined scheme on the basis of CFD and AI is deemed suitable for practical risk assessments.

In addition to this, the following physical outcomes emerged from this work.

1. The boundary conditions have substantial effects on the droplet dispersion. The ambient velocity and initial speed of droplets can widen the septic zone.
2. The droplet diameter dominates the dispersion process. The droplets with smaller diameters (less than 250 µm) are very likely to remain suspended in air and thus be transferred to other parts of the environment.
3. The results indicate that about 59% of the initial droplet will be deposited in the first 2 m away from the sneezing person. The droplets with diameters between 500 and 1000 µm are most likely to fall and hit the ground within this distance.
4. The concentration of droplets could decline to 87% in the first 3 m. Nonetheless, this process is heavily affected by the ambient temperature and airflow velocity.

CRediT authorship contribution statement

Mehrdad Mesgarpour: Conceptualization, Software, Validation, Formal analysis, Writing - original draft. Javad Mohshab Najm Abad: Software, Data curation, Investigation. Asool Alizadeh: Investigation, Project administration. Somchai Wongwises: Resources, Supervision. Mohammadhossein Dornehgard: Writing - review & editing. Saeidreza Ghaderi: Writing - original draft. Nader Karimi: Conceptualization, Supervision, Project administration, Writing - review & editing.

Declaration of Competing Interest

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

Acknowledgment

The first author (M. Mesgarpour) acknowledges Postdoctoral Fellowship from KMUTT. S. Wongwises acknowledges the support provided by the “Research Chair Grant “National Science and Technology Development Agency (NSTDA), and King Mongkut’s University of Technology Thonburi through the “KMUTT 55th Anniversary Commemorative Fluid. N. Karimi acknowledges the financial support by the Engineering and Physical Science Research Council, UK, through the grant number EP/V036777/1 Risk Evaluation Fast Intelligent Tool (RELIANT) for COVID 19.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jhazmat.2021.125358.

References

Ahbogay, M., Talaat, K., Anderoglu, O., Poroseva, S.V., 2020. Numerical investigation of aerosol transport in a classroom with relevance to COVID-19. Phys. Fluids 32 (10), 103311.
Alizadeh, R., 2020b. A machine learning approach to predicting the heat convection and thermodynamics of an external flow of hybrid nanofluid. J. Energy Resour. Technol. 1–22.
Alizadeh, R., Mohshab Najm Abad, J., Fattahi, A., Alhajri, E., Karimi, N., 2020a. Application of machine learning to investigation of heat and mass transfer over a cylinder surrounded by porous media-the radial basic function network. J. Energy Resour. Technol. 142, 1–18.
Briedis, D., Moutrie, M., Balmer, R., 1980. A study of the shear viscosity of human whole saliva. Rheol. Acta 19 (5), 365–374.
Burke, W., 2012. Why do we sneeze? Med. Hypotheses 78 (4), 502–504.
Busco, G., Yang, S.R., Seo, J., Hassam, Y.A., 2020. Sneezing and asymptomatic virus transmission. Phys. Fluids 32 (7), 073309.
Chabert, M., Dorfman, K.D., Viovy, J.L., 2005. Droplet fusion by alternating current (AC) field electrocoalescence in microchannels. Electrophoresis 26 (19), 3706–3715.
