A new fast dynamic Smagorinsky model using artificial neural network for simulation of outdoor airflow and pollutant dispersion

TingDai1, SumeiLiu1*, PunjieLiu1, Nan Piang2, Qingyan Chen3
1Tianjin Key Laboratory of Indoor Air Environmental Quality Control, School of Environmental Science and Engineering, Tianjin University, Tianjin 300072, China
2School of Mechanical Engineering, Tianjin University, Tianjin 300072, China
3Department of Building Environment and Energy Engineering, The Hong Kong Polytechnic University, Kowloon, 999077, Hong Kong SAR, China

Abstract
Fast fluid dynamics (FFD) with the dynamic Smagorinsky model could simulate outdoor airflow and pollutant dispersion accurately. However, it is time-consuming since the model coefficient is obtained by filtering the flow field twice. To improve the computing efficiency, this study proposed a new dynamic Smagorinsky model which calculated model coefficient through artificial neural network (ANN). Compared with the original dynamic Smagorinsky model, the new dynamic Smagorinsky model can save the time to calculate eddy viscosity by more than 60%, thus reducing the total computing time by more than 20%, while ensuring computing accuracy.

Highlights
- The development of a new fast dynamic Smagorinsky model which obtains coefficient by ANN model
- The ANN model is trained from a simple case and is generalizable to other complex cases
- The ANN model can predict the distribution of dynamic model coefficient accurately
- The new dynamic Smagorinsky model can ensure high computing speed and accuracy

Introduction
With the development of industrialization, a large amount of pollutants discharged into the outdoor environment has resulted in severe air pollution, causing a borad range of diseases risk (Anderson et al, 2010; Qadeer et al, 2020; Li, Lu, & Li, 2020; Slama et al, 2019; Wu et al, 2022). It is necessary to study the pollutants diffusion outdoors to prevent them from polluting the environment around buildings. Computational fluid dynamics (CFD) is an effective way for simulating airflow and pollutant dispersion around buildings, from single building to actual urban areas (Chatzidimitriou, & Yannas, 2017; Chen, Rong, & Zhang, 2021; Ghobadi, & Nasrollahi, 2021; Liu et al, 2017; Liu, Heidarinejad, Pitchurow, Zhang, & Srebric, 2018; Liu et al, 2018; Mei et al, 2019; Tominaga, & Mochida, 2016; Tominaga, & Stathopoulos, 2010). CFD is convenient and economical since the condition can be effectively controlled and adjusted, and it could provide data for the entire flow field. However, CFD needs significant computing capacity and time for the simulation of the actual urban area with large computational domain and complex building geometry.

Computing efficiency is of great importance in engineering applications. Fast fluid dynamics (FFD) is an efficient method for airflow and pollution dispersion simulation. It was proved to be 50 times faster than CFD with similar accuracy (Zuo, & Chen, 2009). The original FFD took the assumption of laminar flow, and it was adopted in some follow-up studies (Jin, Liu, & Chen, 2014; Zuo, & Chen, 2010). But the ignorance of turbulence will cause errors in the simulation. Therefore, RNG k-ε turbulence model was implemented into FFD for more accurate simulation in indoor environment (Liu, & Chen, 2018; Liu, You, Zhang, & Chen, 2017; Liu, Jin, Chen, You, & Chen, 2016) and the Smagorinsky model in the actual urban area of Montreal (Mortezaazadeh, & Wang, 2018; Mortezaazadeh, & Wang, 2020). Our previous investigation (Dai, Liu, Liu, Jiang, & Chen, 2022) compared the results of FFD with different turbulence models and found that FFD with the dynamic Smagorinsky model can provide the most accurate prediction in the outdoor environment. However, the calculation of the dynamic model coefficient was time-consuming. Because its calculation formula consisted of subgrid stress $\nu_{ij}$ and strain rate tensor $\nu_{ij}$, which were obtained by filtering the flow field at two cutoff widths (Chester, & Meneveau, 2001; Germano, Piomelli, Moin, & Cabot, 1990; Lilly, 1992; Versteeg, & Malalasekera, 2007).

Recently, an increasing number of research have used artificial neural network (ANN) to improve the computing accuracy and efficiency of subgrid-scale models. For example, it was used to develop subgrid-scale models with accurate subgrid-scale stress tensors and with accurate kinetic energy prediction (Park, & Choi, 2021; Prat, & Sautory, & Martinez, 2020; Qi, Li, Luo, & Yu, 2022; Zhou, He, Wang, & Jin, 2019). In those research, DNS data was selected as the training set to develop the ANN models, but those data were not available in complex outdoor cases. Sarghini, Felice and Santini (2003) used the ANN model to calculate the model coefficient of the hybrid Smagorinsky model, which reduced the calculation time by 20%. But the trained model was not generalizable and only suitable for simple channel flow. Therefore, those developed ANN models were not applicable for the complex outdoor cases.

In summary, FFD with the dynamic Smagorinsky model had accurate outdoor environment simulation but long computing time. The optimized Smagorinsky models were not suitable for outdoor cases and were not
integrated with FFD. Therefore, this study proposed a new dynamic Smagorinsky model in FFD which obtained the dynamic model coefficient by an ANN model. The developed model was trained using the result of the original dynamic Smagorinsky model and can be applied in outdoor environment. The new dynamic Smagorinsky model can shorten the computing time by reducing the filter times. The performance of FFD with the new dynamic Smagorinsky model was evaluated by comparing with that of other Smagorinsky models in the simulation of outdoor cases.

Method

This section introduces the research method used in this investigation, including FFD, Smagorinsky models, and ANN.

**Fast fluid dynamics**

FFD solves the Navier-Stokes equations for an incompressible Newtonian fluid:

\[
\rho \frac{\partial u_i}{\partial t} + \rho u_i \frac{\partial u_j}{\partial x_j} = - \frac{\partial p}{\partial x_i} + \mu \Delta u_i + \mu \frac{\partial^2 u_i}{\partial x_j \partial x_j} + F_i
\]

The SIPC scheme (Guermond, Minev, & Shen, 2006) is applied to split Equation (1) into two discretized equations:

\[
\frac{u_i^n - u_i^0}{\Delta t} = - \frac{\partial p^n}{\partial x_i} - \rho u_i^n \frac{\partial u_j^n}{\partial x_j} + \mu \frac{\partial^2 u_i^n}{\partial x_j \partial x_j} + F_i
\]

The pressure projection method (Guermond, Minev, & Shen, 2006) is applied in resolving the coupled pressure and velocity. Substituting Equation (4) into Equation (1) yields:

\[
\frac{\partial^2 (p^{n+1} - p^n)}{\partial x_i \partial x_i} = \rho \frac{\partial u_i^n}{\Delta t \partial x_i}
\]

where \(p^{n+1}\) is calculated by solving Equation (5). With \(u_i^n\) and \(p^n\), Equation (4) can be used to obtain \(u_i^{n+1}\).

**Smagorinsky models**

The flow is filtered into large-scale flow and small-scale flow in Smagorinsky models. The effect of small-scale flow is described by the model. In the standard Smagorinsky model (Smagorinsky, 1963), the turbulence viscosity \(\mu_t\) is calculated by:

\[
\mu_t = \rho C_d \Delta^2 \left( \frac{2S_{ij} S_{ij}}{S_{ij}} \right) = \rho C_d \Delta^2 \sqrt{S}
\]

where \(C_d\) is a dynamic model coefficient and can be solved by:

\[
C_d = \frac{(L_{ij} M_{ij})}{(M_{ij} L_{ij})}
\]

The bar and wavy line represent the variables filtered at grid scale \(\Delta\) and test scale \(\Delta\) (typically \(\Delta = 2 \Delta\)), respectively. Then Equation (12) can be expressed as:

\[
M_{ij} = 2 \Delta^2 \left( \frac{4 \Delta^6 S_{ij} - S_{ij}}{S_{ij}} \right) = 2 \Delta^2 N_{ij}
\]

Substituting Equation (13) into Equation (10) yields:

\[
C_d = \frac{(2L_{ij} N_{ij})}{(4\Delta^2 N_{ij} N_{ij})}
\]

The computing formula of added subgrid stress \(L_{ij}\) and strain rate tensor \(N_{ij}\) includes the filtered variables at two scale, so they are obtained by filtering the flow field twice. The commonly used filter method is the volume average box filter, and it is related to the gradient of variables and grid scale (Wilcox, 2006). For example, the difference between the variables \(\bar{u}_i\) and \(\bar{u}_j\) in \(L_{ij}\) is related to gradient of \(\bar{u}_i\) \(\bar{u}_j\) and the grid scale \(\Delta\). The same applies to other variables in the equations \(N_{ij}\).

**Artificial neural network**

Equation (14) is used to calculate the model coefficient in the dynamic Smagorinsky model. It takes a long time to obtain the added subgrid stress \(L_{ij}\) and strain rate tensor \(N_{ij}\) by filtering the flow field twice. This study proposes to obtain \(L_{ij}\) and \(N_{ij}\) by ANN model. We compare the performance of ANN with different inputs parameters and found that \(\text{grad} \bar{u}_i\), \(\text{grad} \bar{u}_j\), \(\text{grad} \bar{u}_i\Delta^2\) and \(\text{grad} \bar{u}_i\Delta\) are the optimal inputs of the artificial neural network of \(L_{ij}\). The comparison shows that \(\text{grad} \left[ S \right] \text{grad} S_{ij}\Delta^2\) and \(\text{grad} \left[ S \right] \text{grad} S_{ij}\Delta\) are the optimal inputs of the artificial neural network of \(N_{ij}\). Figure 1 shows the structure of ANN for \(L_{ij}\) and \(N_{ij}\).
The artificial neural network has four layers, including an input layer, two hidden layers, and an output layer. The loss function of ANN model is the mean-square-error (MSE) function. The backpropagation method (Rumelhart, Hinton, & Williams, 1986) is used to train ANN by optimizing the weight and bias coefficients to minimize the loss function. The percentage of data used for training, validation, and testing in the ANN are 70%, 15% and 15%, respectively. The activation functions of the hidden layers and output layer are the sigmoid function and linear function, respectively. The sigmoid function (Hornik, 1991) is defined as:

\[
f(x) = \frac{1}{1 + e^{-x}}
\]

(15)

The above methods were implemented into FFD in Open Field Operation and Manipulation (OpenFOAM), an open-source program (Weller, Tabor, Jasak, & Fureby, 1998).

**Result**

This section presented the result comparison of FFD with the standard Smagorinsky model, original dynamic Smagorinsky model and the new dynamic Smagorinsky model in outdoor cases.

**The single-building case**

The single-building case (Tanaka, Yoshie, & Cheng, 2006) was a benchmark case for outdoor environment simulation. It consisted of a square building and pollution source on the ground, as shown in Figure 2. It included the basic features of flow and pollution dispersion around buildings. Therefore, this case was chosen to train the ANN model. The ANN was trained with data from the instantaneous simulation results of the original dynamic Smagorinsky model in 50% grid points. According to the grid independence test, the computational domain was meshed into 0.71 million unstructured grids.

![Figure 2: Geometric configuration of the single-building case.](image)

The iteration number for convergence is in the range of 150-250 in the training process. The ANN correlation coefficient of \( L_{ij} \) and \( N_{ij} \) were 0.931 and 0.927, respectively, which meant that the ANN model predicted well. Figure 3 shows the contours of the model coefficient of FFD with the two dynamic Smagorinsky models, which are similar. The approaching flow separated on the windward side of the building, resulting in strong turbulence dissipation and hence the large model coefficient. The model coefficient was small at the upper rear of the building since the flow reverted to horizontal. A larger model coefficient was shown in the wake of the building owing to the reversal and separation there. The standard Smagorinsky model with a constant model coefficient of 0.0324 (0.18) underestimated the turbulence fluctuation in front of the building and overestimated the turbulence dissipation behind the building.

![Figure 3: Contours of model coefficient simulated by FFD with (a) the original dynamic Smagorinsky model and (b) the new dynamic Smagorinsky model in the single-building case.](image)

Figure 4 shows the normalized streamwise velocity and concentration contour of FFD with the three Smagorinsky models. In the result of standard Smagorinsky model, the reattachment length and reverse velocity behind the building were over-predicted because of the over-predicted model coefficient. Pollutants were concentrated on the leeward side of the building owing to the reverse velocity there. The new dynamic Smagorinsky model had similar simulation result to the original dynamic Smagorinsky and was greatly improved when comparing with the standard Smagorinsky model, because it can predict the dynamic model coefficient accurately. The comparison showed that the new dynamic Smagorinsky model performed well in the single-building case.

![Figure 4: The distributions of the normalized streamwise velocity \( U_j/U_H \) and concentration \( K_j \) on the centre vertical plane computed by FFD with the standard Smagorinsky model (a and b), the original dynamic Smagorinsky model (c and d), and the new dynamic Smagorinsky model (e and f). \( K_j = C_{gas}/C_{ref} \) where \( C \) was the gas concentration (ppm), \( C_{gas} \) was the released tracer-gas concentration (ppm), \( U_H \) was the reference wind](image)
velocity (m/s) measured at height H, H was the building height (0.2 m), and Q was the released gas emission (m³/s).

Table 1 summarizes the normalized root mean square error (NRMSE) for quantitative evaluation of the simulation results. We took the result of the original dynamic Smagorinsky model as the baseline to calculate the error between that result and the simulation results of the other two models. The velocity and pollutant concentration errors of the standard Smagorinsky model were 7.23% and 11.69%, respectively. This is because the model overestimated the length of the recirculation zone and the reverse velocity behind the building. Due to the correct prediction of the model coefficient, the velocity and pollutant concentration errors of the new dynamic Smagorinsky model were 1.97% and 3.15%, respectively. The new dynamic Smagorinsky model can be considered to have the same accuracy as the original dynamic Smagorinsky model since the errors were less than 5%.

### Table 1 Summary of normalized root mean square error for velocity and pollutant concentration of different turbulence models in the single-building case.

<table>
<thead>
<tr>
<th>Turbulence model</th>
<th>Velocity</th>
<th>Pollutant concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>FFD + standard Smagorinsky model</td>
<td>7.23%</td>
<td>11.69%</td>
</tr>
<tr>
<td>FFD + new dynamic Smagorinsky model</td>
<td>1.97%</td>
<td>3.15%</td>
</tr>
</tbody>
</table>

### The three-building case

The new dynamic Smagorinsky model was tested in the three-building case (Chavez et al. 2011), as shown in Figure 5. A mixture comprising nitrogen and SF₆ was released from a stack located on the roof of the B1 building. According to the grid independence test, the computational domain was meshed into 1.52 million unstructured grids.

Figure 6 shows the model coefficient contour of the two dynamic Smagorinsky models in the three-building case. The model coefficient distribution of this case had the similar characteristic as that of the single-building case. The detached flow in front of the building caused the large model coefficient. Horizontal flow at the upper rear of the upward building made the coefficient small. The model coefficient was larger above the centre building because of the vortices generated by the interference of upward building. The model coefficient distribution of the two dynamic Smagorinsky models was similar, while the coefficient of the standard Smagorinsky model was a constant of 0.0324. The invariable coefficient made the turbulence fluctuation underestimated in front of the upward building and over-predicted above the centre building in the result of standard Smagorinsky model.

Figure 7 display the streamwise velocity and normalized concentration result simulated by FFD with the three Smagorinsky models. A recirculation zone was formed at the top of the centre building under the obstruction of high upward building. In the recirculation zone, the flow direction changed from forward to reverse with the decrease of height. The pollutant was carried around the upward buildings because of the reverse flow at the chimney outlet. The underestimated model coefficient in the standard Smagorinsky model made the height of the recirculation zone over-predicted, and the pollutant spread over a larger region. Since the new dynamic Smagorinsky model had correct prediction of the model coefficient, the velocity and concentration result of it was similar to that of the original dynamic Smagorinsky model and was greatly improved over the standard Smagorinsky model.
Table 2 compares the normalized root mean square error for velocity and pollutant concentration simulated by different turbulence models. The results of the original dynamic Smagorinsky model were taken as the baseline. The velocity and pollutant concentration errors of the standard Smagorinsky model were large since it overestimated the height of the recirculation zone and the concentration diffusion. The errors of the new dynamic Smagorinsky model were less than 5%; therefore, it can be considered to be as accurate as the original dynamic Smagorinsky model.

Table 2: Summary of normalized root mean square error for velocity and pollutant concentration of different turbulence models in the three-building case.

<table>
<thead>
<tr>
<th>Turbulence model</th>
<th>Velocity</th>
<th>Pollutant concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>FFD + standard Smagorinsky model</td>
<td>8.15%</td>
<td>15.86%</td>
</tr>
<tr>
<td>FFD + new dynamic Smagorinsky model</td>
<td>2.32%</td>
<td>3.55%</td>
</tr>
</tbody>
</table>

Computing time

Computing efficiency was a major consideration in engineering applications. Therefore, this investigation compared the computing time of FFD with the three turbulence models. The total computing time and the time for calculating turbulent viscosity \( \mu_t \) of the two cases were summarized in Table 3. The main difference between the three models was the way of obtaining model coefficient. The standard Smagorinsky model had the shortest time because of the invariable coefficient. The original dynamic Smagorinsky model took the longest time since the coefficient was calculated by filtering the flow field twice. The new dynamic Smagorinsky model obtained the model coefficient through an artificial neural network. Compared with the original dynamic Smagorinsky model, the new dynamic Smagorinsky model can save the time to calculate turbulent viscosity by 62.5%-63%, so the overall computing time can reduce by 22%-22.9%.

Table 3: Summary of the computing time used by FFD with the three models

<table>
<thead>
<tr>
<th>case</th>
<th>Turbulence model</th>
<th>Time for ( \mu_t ) (s)</th>
<th>Overall time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single-building</td>
<td>FFD + standard Smagorinsky model</td>
<td>151.2</td>
<td>3754.5</td>
</tr>
<tr>
<td></td>
<td>FFD + original dynamic Smagorinsky model</td>
<td>1664.1</td>
<td>5420.8</td>
</tr>
<tr>
<td></td>
<td>FFD + new dynamic Smagorinsky model</td>
<td>615.0</td>
<td>4227.6</td>
</tr>
<tr>
<td>Three-building</td>
<td>FFD + standard Smagorinsky model</td>
<td>553.2</td>
<td>15962.1</td>
</tr>
<tr>
<td></td>
<td>FFD + original dynamic Smagorinsky model</td>
<td>6054.7</td>
<td>22787.1</td>
</tr>
</tbody>
</table>

Discussion

The training set of ANN should include the basic flow characteristics to make the model generalizable for the flow to be studied. The single building case was taken as the training set because it contained the basic flow characteristics in the outdoor environment around buildings. Therefore, the trained ANN was applicable to other complicated outdoor case. However, it was unsuitable for other type of flows, such as airflow in a building (Wang, & Chen, 2009; Yuan, Chen, Glicksman, Hu, & Yang, 1999, Liu, & Deng, 2023; Liu, Koupriyanov, Paskaruk, Fediuk, & Chen, 2022).

Conclusions

This study developed a new dynamic Smagorinsky model which obtains the model coefficient by ANN to improve the computing efficiency. The model was trained on a single building case and was tested on the three-building case. The conclusions are as followed:

1. The ANN trained on single building case has applicability to the three-building case.
2. The new dynamic Smagorinsky model can accurately predict the distribution of model coefficient. The simulation result of the new dynamic Smagorinsky model was improved when comparing with that of the standard Smagorinsky model and was similar to that of the original dynamic Smagorinsky model.
3. Compared with the original dynamic Smagorinsky model, the new dynamic Smagorinsky model can reduce the time to calculate turbulent viscosity by more than 60%, thus reducing the overall computing time by more than 20%.

Acknowledgement

This study was partially supported by the National Natural Science Foundation of China (NSFC) through grant No. 52108084, and by the Global STEM Professorship from the Hong Kong government received by Qingyan Chen.

References


