





Article

Modified Accuracy of RANS Modeling of Urban Pollutant Flow within Generic Building Clusters Using a High-Quality Full-Scale Dispersion Dataset

Mohammad Reza Kavian Nezhad , Khashayar RahnamayBahambary , Carlos F. Lange  and Brian A. Fleck 

Department of Mechanical Engineering, The University of Alberta, 5-1J Mechanical Engineering Building, Edmonton, AB T6G 2G8, Canada

* Correspondence: kavianne@ualberta.ca

Abstract: To improve the reliability of the computational fluid dynamics (CFD) models of wind-driven pollutant dispersion within urban settings, a re-calibration study is conducted to optimize the standard $k - \epsilon$ model. A modified optimization framework based on the genetic algorithm is adapted to alleviate the computational expenses and to further identify ranges for each empirical coefficient to achieve the most reliable and accurate predictions. A robust objective function is defined, incorporating both the flow parameters and pollutant concentration through several linear and logarithmic measures. The coefficients are trained using high-quality and full-scale tracer experiments in a mock urban arrangement simulating a building array. The proposed ranges are $0.14 \leq C_\mu \leq 0.15$, $1.30 \leq C_{\epsilon 1} \leq 1.46$, $1.68 \leq C_{\epsilon 2} \leq 1.80$, $1.12 \leq \sigma_\epsilon \leq 1.20$, and $0.87 \leq \sigma_k \leq 1.00$. A thorough evaluation of the predicted flow and concentration fields indicates the modified closure is effective. The fraction of predictions within the acceptable ranges from measurements has increased by 8% for pollutant concentration and 27% for turbulence kinetic energy. The generality of the calibrated model is further tested by modeling additional cases with different meteorological conditions, in which the calculated validation metrics attest to the noteworthy improvements in predictions.

Keywords: CFD; pollutant dispersion; urban setting; ABL; RANS; Closure model optimization



Citation: Kavian Nezhad, M.R.; RahnamayBahambary, K.; Lange, C.F.; Fleck, B.A. Modified Accuracy of RANS Modeling of Urban Pollutant Flow within Generic Building Clusters Using a High-Quality Full-Scale Dispersion Dataset. *Sustainability* **2023**, *15*, 14317. <https://doi.org/10.3390/su151914317>

Academic Editors: Mohammad Reza Pendar and Frederico Miguel Freire Rodrigues

Received: 23 August 2023

Revised: 23 September 2023

Accepted: 26 September 2023

Published: 28 September 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The continuous growth in the world's population in recent decades and the need for higher living standards have led to rapid urbanization worldwide [1,2]. However, despite the countless benefits of the ongoing developments, several detrimental impacts on the environment and public health are also inevitable. The emergence of industrial facilities, along with the growth in the popularity of motor vehicles, have aggravated poor urban air quality [3–5]. Additionally, the intensive construction of crucial infrastructures in response to rapid urbanization has resulted in a compact and diverse arrangement of the buildings [6]. With these growing pressures, anticipating and controlling the possible health hazards of living in compact regions has become a topic that requires attention. The constantly changing layout of urban areas affects the wind flow patterns, which, if not properly planned, can aggravate poor air quality by progressively worsening urban ventilation performances [7]. Therefore, to mitigate the possibility of pollutant accumulation, assessing the wind flow field and the dispersion patterns around the buildings prior to construction is strongly advised [8,9].

Full-scale field measurements and reduced-scale laboratory experiments (i.e., wind tunnels and water channels) have been used in the past to investigate wind flow and dispersion patterns around the buildings and to further identify design shortcomings [10–14]. Even though field measurements can account for the realistic atmospheric state, the uncontrolled meteorological conditions make it costly to independently study influencing parameters and collect data that can be used to predict dispersion patterns [15].

Reduced-scale experiments can be performed in a controlled arrangement, facilitating parametric studies. However, in addition to missing actual environmental effects, this method has some disadvantages, such as complicated scaling and similarity issues [16]. The apparent complexities, limitations, and high costs associated with the experimental methods have limited their applications to general cases mainly used for model validation studies. Therefore, the need for a more effective and practical approach has made computational fluid dynamics (CFD) very popular among environmental researchers and urban planners [17,18].

CFD can estimate the pollutant concentration and other flow characteristics on every grid point in a computational domain and is generally less costly than experiments. However, the numerical modeling of the urban dispersion flow within the atmospheric boundary layer (ABL) is quite challenging due to its turbulent nature with large-scale recirculation zones and three-dimensional strain fields [11]. Therefore, considering appropriate assumptions and modeling settings at every step of the process is essential to having a reliable and efficient CFD model. That includes efforts to evaluate the applicable approaches to treat the governing equations, examples being Reynolds-averaged Navier–Stokes (RANS) and large eddy simulations (LES) [19]. LES was found to produce relatively richer results than RANS by resolving the large and most prominent eddies; however, that comes with prohibitively greater computational demands. Given the large size of the computational domain in urban dispersion studies and the focus in the mean quantities of flow characteristics (assuming a continuous release of the pollutant from the source), RANS has remained quite popular in resolving the Navier–Stokes equations [20–23].

The time-averaged treatment of this completely chaotic and randomized phenomenon (i.e., turbulent atmospheric dispersion flow) can be achieved by applying the Boussinesq hypothesis to model the intermittent shear fluxes [24]. However, neglecting the available fluctuations during the momentum transport (using modeled turbulence viscosity) undermines the reliability of the CFD model predictions. Under-prediction of the turbulence kinetic energy (TKE) or inaccurate representation of the separation points and the reattachment lengths in flows around bluff bodies are among the few well-established flaws of RANS [25]. Several closure models, along with their modifications, have been introduced over the years to remediate these deficiencies. Notably, these models were developed in the context of conventional classical flows (e.g., fully developed turbulent channel flow, simplified wall-bounded, or free shear flow, etc.), but their applications can be extended to a wide range of engineering problems [26]. Despite the apparent disparity between the nature of the mentioned classical flows and the ABL flow, common two-equation closures, such as standard $k - \varepsilon$ [27], realizable $k - \varepsilon$ [28], renormalization group (RNG) $k - \varepsilon$ [29], and the shear stress transport (SST) $k - \omega$ [30] have been widely used in urban studies [15,31]. To ensure the reliability of an urban dispersion model, conducting a sensitivity study to select the most suitable closure for the specific case of interest is advised [32–35]. Moreover, the accuracy of the selected model in the context of ABL can be further improved by re-calibrating the empirical coefficients to better fit the critical features of this particular flow.

The majority of CFD codes employ the revised values offered by Launder and Spalding ([27]) as the default closure coefficients for the standard $k - \varepsilon$ model ($C_\mu = 0.09$, $C_{\varepsilon 1} = 1.44$, $C_{\varepsilon 2} = 1.92$, $\sigma_k = 1$, and $\sigma_\varepsilon = 1.3$). Previous researchers have made efforts to modify the closure coefficients with the purpose of improving the representation of the urban flow in various cases. As one of the initial attempts, Detering [36] modified the original definition of coefficients for modeling the neutrally stratified, one-dimensional atmospheric flow over flat and irregular terrain (e.g., hills, valleys, etc.). In this work, C_μ was assumed to be inversely proportional to the depth of the atmospheric boundary layer, which consequently led to a new set of constants ($C_\mu = 0.03$, $C_{\varepsilon 1} = 1.13$, $C_{\varepsilon 2} = 1.9$, $\sigma_k = 0.77$, and $\sigma_\varepsilon = 1.29$). Later, Bechmann proposed a hybrid RANS/LES method based on the standard $k - \varepsilon$ model and adopted a $C_\mu = 0.03$, as suggested for atmospheric flows over irregular terrain, instead of the original value of 0.09 for industrial flows [37]. The number of unknowns was reduced in this proposed adjustment by keeping $C_{\varepsilon 2}$, σ_k , and σ_ε

the same as the standard values but calculated $C_{\varepsilon 1} = 1.30$ using an empirical correlation between the model constants in the ABL. The mentioned correlation was developed by Richards and Hoaxy [38], assuming constant shear stress in the atmospheric surface layer, resulting in an equilibrium between shear production and viscous dissipation. The results of Bechmann's model of flow over Askervein hill showed an improved accuracy compared to the ones offered by Detering. However, it could not be confidently distinguished whether this progress was mainly due to the re-calibration of coefficients or the proposed hybrid model.

The modification of the standard $k - \varepsilon$ model was further extended to a more complicated scenario by Guilass et al. [39], in which the airflow within a regular street canyon was considered. The authors performed a Bayesian calibration to tune four out of the five constants by excluding σ_{ε} and calculating it directly via the correlation between the model constants within the ABL [38]. The vertical profile of the TKE at the center of the street canyon was considered as the optimization objective, and the results of the 135 CFD runs were processed to determine a set of constants that produced the preferable match with the wind tunnel measurements ($C_{\mu} = 0.12$, $C_{\varepsilon 1} = 1$, $C_{\varepsilon 2} = 2.1$, $\sigma_k = 0.46$, and $\sigma_{\varepsilon} = 0.42$). In another work by Zahid Iqbal and Chan [40], an investigation of the wind flow field around a cross-shaped building at the pedestrian level was conducted. The modification of the $k - \varepsilon$ closure model in this attempt was based on the proposed coefficients by Guilass et al. [39]. The number of unknowns was reduced to simplify the process by just varying C_{μ} at four equal intervals within the range of 0.09–0.12 while keeping the values of $\sigma_k = 0.53$ and $\sigma_{\varepsilon} = 0.5$ constant, as suggested by Edeling et al. [41]. Setting the normalized velocity field as the modification objective in their study, $C_{\mu} = 0.12$ was shown to provide the least discrepancies with wind tunnel measurements, though the superiority of the tuned model over the one proposed by Guilass et al. [39] was insignificant. This conclusion heightens the importance of calibrating the whole set of coefficients simultaneously to accomplish a worthwhile improvement, in contrast to alleviating the computational complexity by just doing a linear sensitivity study on a selected coefficient.

The demonstrated uncertainty inherent in these coefficients implies the necessity of their objective modification within the reference frame of generic case studies. Because of the non-linear relationship between coefficients, a complete closure optimization for ABL flow in large models (e.g., compact urban settings) becomes substantially more challenging. Implementing data assimilation methods, such as the Bayesian calibration, demands a large number of CFD runs, creating the need for more robust optimization approaches to obtain the best coefficient set. In this regard, Shirzadi et al. [42] used stochastic optimization combined with the Monte Carlo sampling scheme and adopted the streamwise velocity around an isolated building as the optimization objective. They used available wind tunnel measurements to introduce a set that better represents the flow characteristics compared with the standard model ($C_{\mu} = 0.146$, $C_{\varepsilon 1} = 1.489$, $C_{\varepsilon 2} = 2.801$, $\sigma_k = 1$, and $\sigma_{\varepsilon} = 0.373$). Shirzadi et al. later performed a parametric sensitivity study to evaluate the possibility of improving the numerical simulation of cross-ventilation in compact urban regions [43]. They did not use an optimization framework in this study but mitigated the associated computational costs by considering 10 uniformly distributed values for each coefficient within the recommended ranges. The resultant set only offered modifications to $C_{\varepsilon 2}$ and σ_{ε} (3.2 and 0.35, respectively), while keeping the rest of the coefficients the same as their original values. The aptness of the suggested coefficients was examined using wind tunnel measurements of the wind pressure difference over the faces of the central building (i.e., the objective function in this study). Even though an improved agreement with measurements was observed, the accuracy of the calibrated model was found to be inadequate and in need of further efforts [44].

Objectives

Following the previous endeavors, this study aims to enhance the reliability and generality of the closures in ABL dispersion studies. Based on the presented review, a

well-tested set of coefficients that accurately represents the pollutant concentration field within an urban array has yet to be developed. The current research is part of a project to improve the existing practices in dispersion modeling. In our previous paper, the effects of several critical modeling decisions (e.g., closure model, inflow boundary conditions, computational domain size, and turbulent Schmidt number) were investigated, and a relatively efficient framework was introduced [35]. Considering different venting scenarios (i.e., various source locations), the standard $k - \varepsilon$ model coupled with the locally variable turbulent Schmidt number (Sc_t) was found to be the most efficient setup with the least calculated deviations from field measurements. Therefore, the present study conducts a thorough optimization of the standard $k - \varepsilon$ closure by incorporating the recommended modeling settings.

The importance of selecting a high-quality data set for the robust calibration of the turbulence model constants is undeniable. Among all of the previous attempts reviewed, none included the pollutant concentration field in their defined focus parameters. Additionally, to the best of our knowledge, all earlier ABL calibration studies used reduced-scale wind tunnel data produced in controlled and steady test arrangements. Full-scale field measurements, however, have the effects of constantly varying the meteorological features and realistic atmospheric conditions inherent in them, making them very valuable for improving the time-averaged representation of an intrinsically unsteady phenomenon. To this aim, the unique data set of the mock urban setting tests (MUST) is employed, as it provides comprehensive measurements of both the concentration and airflow fields throughout a compact building array (74 measuring points for the concentration field and 22 measuring points for the airflow field) [45]. Using full-scale field measurements of this kind creates an unprecedented opportunity to account for realistic atmospheric features that could potentially lead to a more accurate representation of the dispersion flow by RANS simulations. Given the variation in the concentrations in a broad range inside the domain (0.001–100 ppm), an exhaustive statistical analysis is carried out that includes both the linear and logarithmic validation metrics. In addition to the optimization objective, the predicted turbulence characteristics are also assessed to evaluate uncertainties of the coefficients more extensively. Ultimately, three other test cases representing different meteorological conditions (e.g., inflow wind speed and turbulent kinetic energy, wind direction, etc.) and source locations are modeled to further examine the generality of the proposed set of constants for an idealized array simulating urban regions.

2. Mathematical Model

The pattern in which an emitted pollutant stream is dispersed in the ABL heavily depends on the wind regime and flow features. In this regard, acquiring a solid understanding of urban flows is essential prior to characterizing the physics of this phenomenon. The ABL generally refers to the lowest portion of the atmosphere, which can be divided into two main sublayers. The outer region (i.e., the Ekman layer) makes up approximately 90% of this layer and shows a balance among pressure gradient, friction, and Coriolis forces. The inner region (i.e., the surface layer) exhibits strong vertical fluxes of physical quantities with negligible variation within its depth [46]. As our work aims to model pollutant dispersion within a compact urban setting, the surface layer is the area of interest in investigating the flow characteristics. Therefore, disregarding the Coriolis effects on the direction of the urban wind flow is deemed an acceptable assumption. This claim can be further argued by assessing the non-dimensional Rossby number (R_o), which quantifies the ratio of inertial to Coriolis forces [47]. Given the typical velocity and length scales in urban studies, R_o is estimated to be in the order of 10^3 , which indicates a strong dominance of inertial forces [35].

The atmospheric flows in urban areas with a horizontal length scale of less than 10 km are typically treated at the micro-scale range [48]. The CFD simulation of pollutant dispersion within this spatial scope demands explicit modeling of the available structures. These roughness elements, such as buildings of varying shapes and dimensions, intensify

the turbulent nature of the ABL flows. This generates three-dimensional flow features, such as flow separation, recirculation, and substantial directional change of the wind around the bluff bodies. As a result, eddies with varying lengths and time scales are formed, and modeling them requires special considerations. The RANS equations would be an appropriate method to employ in cases dealing with time scales that are considerably greater than turbulent fluctuations. On pollutant dispersion modeling in full-scale urban geometries, the RANS equations provide an acceptable compromise between accuracy and computational cost and are, therefore, adopted in this study.

The physics of a steady-state, incompressible, and iso-thermal flow within a neutrally stratified atmosphere with no body forces can be described by time-averaged continuity and momentum equations:

$$\frac{\partial \bar{u}_i}{\partial x_i} = 0 \quad (1)$$

$$\bar{u}_j \frac{\partial \bar{u}_i}{\partial x_j} = -\frac{1}{\rho} \frac{\partial \bar{p}}{\partial x_i} + \nu \frac{\partial^2 \bar{u}_i}{\partial x_j^2} - \frac{\partial \overline{u'_i u'_j}}{\partial x_j} \quad (2)$$

where \bar{u}_i and u'_j are the mean and fluctuating velocity components in the Cartesian directions of x_i and x_j , respectively ($i, j = 1, 2, 3$). ρ is the air density, \bar{p} is the time-averaged pressure, and ν is the kinematic viscosity defined as $\nu = \mu/\rho$ with μ being the dynamic viscosity. The Reynolds stress tensor ($\overline{u'_i u'_j}$) further promotes an enhanced diffusive transport of momentum due to the fluctuating velocity components. However, this term introduces additional unknowns, leading to more variables than the available equations. Assuming an isotropic turbulent flow, the Boussinesq hypothesis can be applied to model these Reynolds stresses in terms of mean velocity gradients:

$$-\overline{u'_i u'_j} = 2 \left(\nu_t \bar{S}_{ij} - \frac{1}{3} k \delta_{ij} \right) \quad (3)$$

$$\bar{S}_{ij} = \frac{1}{2} \left(\frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i} \right) \quad (4)$$

$$k = \frac{1}{2} \overline{u'_i u'_i} \quad (5)$$

in which \bar{S}_{ij} represents the mean strain rate tensor, k is the turbulence kinetic energy (i.e., TKE), δ is the Kronecker delta, and ν_t denotes the eddy viscosity.

ν_t is a parameter defined as a property of the flow to reflect and control turbulence through a form of viscosity, analogous to the role of molecular viscosity in a laminar flow [49]. In order to mathematically close the governing equations, several eddy-viscosity closure models have been developed over the years to approximate ν_t and provide supplementary equations. These closures aim to model the eddy viscosity as a product of turbulent velocity and length scales. Based on the argument made in the introduction, the standard $k - \varepsilon$ model is adopted in the current study to model ν_t and estimate all the other turbulence quantities [49]. The TKE represents the turbulent velocity scale in this closure and can be calculated using a transport equation as follows:

$$\bar{u}_j \frac{\partial k}{\partial x_j} = \frac{\partial}{\partial x_j} \left[\left(\nu + \frac{\nu_t}{\sigma_k} \right) \frac{\partial k}{\partial x_j} \right] - \overline{u'_i u'_j} \frac{\partial \bar{u}_i}{\partial x_j} - \varepsilon \quad (6)$$

where σ_k is a model constant referred to as turbulence Prandtl number, and ε represents the dissipation rate of the TKE into internal energy (heat). The second term on the right-hand

side of the Equation (6) serves as a source of TKE production (P_k), which can alternatively be expressed in terms of the mean velocity shear stresses using the Boussinesq hypothesis:

$$P_k = -\overline{u'_i u'_j} \frac{\partial \bar{u}_i}{\partial x_j} \approx \nu_t \left(\frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i} \right) \frac{\partial \bar{u}_i}{\partial x_j} \quad (7)$$

In addition to the TKE, another turbulence quantity is required to estimate the associated length scales of eddies within the flow field. The appearance of the dissipation rate in the form of a sink term in Equation (6) suggests that solving a transport equation for ε is a logical choice to close the equation set. It is worth noting that the TKE equation remains the same among all variants of $k - \varepsilon$ closure, whereas the proposed equation for ε is what distinguishes these models. Using the standard version of the $k - \varepsilon$ model, the transport of ε can be resolved by:

$$\bar{u}_j \frac{\partial \varepsilon}{\partial x_j} = \frac{\partial}{\partial x_j} \left[\left(\nu + \frac{\nu_t}{\sigma_\varepsilon} \right) \frac{\partial \varepsilon}{\partial x_j} \right] + \frac{\varepsilon}{k} (C_{\varepsilon 1} P_k - C_{\varepsilon 2} \varepsilon) \quad (8)$$

σ_ε , $C_{\varepsilon 1}$, and $C_{\varepsilon 2}$ are the model's empirical constants that were derived through intensive data fitting with a number of classical flows [27]. The TKE and dissipation rates resulting from supplementary equations can be further used in Equation (9) to estimate the eddy viscosity. The factor of proportionality in this equation, C_μ , is another empirical model constant.

$$\nu_t = C_\mu \frac{k^2}{\varepsilon} \quad (9)$$

The dispersion of the pollutant within the resolved wind and turbulence fields can be modeled by solving the Eulerian advection–diffusion transport equation. Considering a neutrally buoyant and inert gas emitting from a source point without initial momentum, the Reynolds averaging method is once again employed to decompose the instantaneous quantities (i.e., velocity components and scalar concentration) into their mean and fluctuating elements:

$$\bar{u}_j \frac{\partial \bar{c}}{\partial x_j} = \frac{\partial}{\partial x_j} \left(D \frac{\partial \bar{c}}{\partial x_j} \right) - \frac{\partial \overline{u'_j c'}}{\partial x_j} + S' \quad (10)$$

where \bar{c} and c' are the mean and fluctuating scalar concentrations, respectively. S' denotes the scalar source term, and D represents the molecular diffusivity defined as the ratio of molecular viscosity to molecular Schmidt number ($D = \frac{\nu}{Sc}$). Equation (10) is coupled with the RANS equation system in one-way under the assumption of a passive scalar, which considerably reduces the computational cost and modeling complexity [50].

The convective transport of the scalar concentration due to the fluctuating velocities (i.e., unresolved flow field) is expressed by $\overline{u'_j c'}$. With an analogy similar to the Boussinesq hypothesis (i.e., the random isotropic motion), the standard gradient diffusion hypothesis (SGDH) assumes that the turbulent convective transport of a scalar occurs in the direction of the time-averaged concentration gradient [51]. From this, the turbulent scalar transport can be approximated as follows:

$$-\overline{u'_j c'} = D_t \frac{\partial \bar{c}}{\partial x_j} \quad (11)$$

Similar to eddy viscosity, eddy diffusivity ($D_t = \frac{\nu_t}{Sc_t}$) is also a property of the turbulence and not of the fluid. On another note, a normalized definition of pollutant concentration (C^*) is used in this work for making comparisons between predictions and measurements. In this equation, U_{ref} is the upstream reference velocity, C is the predicted

or measured concentration at a given location, H_{ref} is a reference length scale, and q_s is the volumetric flowrate at which the pollutant is being released from the source.

$$C^* = \frac{10^{-6} U_{ref} C H_{ref}^2}{q_s} \quad (12)$$

The turbulent Schmidt number, Sc_t , is a variable in the scalar transport equation that substantially affects the accuracy of the predicted concentration field. Having in mind that the turbulent mass flux is approximated through a closure assumption (i.e., SGDH), a universal value of Sc_t cannot be determined [52]. To this end, similar to the coefficients of the eddy-viscosity turbulence models, obtaining case-dependent Sc_t might be a practical approach. However, in addition to the time demanding sensitivity studies required by this approach, several experimental and validated numerical studies have demonstrated the strong local variability of Sc_t [53,54]. As part of our ongoing efforts to improve the reliability of the urban dispersion modeling, the proposed framework in [35] is adopted in this study. Following these recommendations, a well-tested locally variable Sc_t is incorporated into the advection–diffusion equation. The aforementioned Sc_t formulation accounts for the local state of turbulence and estimates the optimum value in every computational node [55]. The implementation of the locally variable Sc_t further strengthens the dependency of the transport equation on the flow-related parameters. Using Equation (13), Sc_t is defined as:

$$\log(Sc_t) = 0.6617 Sc - 0.8188 Re_t^{0.01} - 0.00311 S - 0.0329 \Omega \quad (13)$$

$$Re_t = \frac{k^2}{\nu \varepsilon} \quad (14)$$

$$S = \frac{k}{\varepsilon} \sqrt{\frac{1}{2} \left(\frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i} \right)^2} \quad (15)$$

$$\Omega = \frac{k}{\varepsilon} \sqrt{\frac{1}{2} \left(\frac{\partial \bar{u}_i}{\partial x_j} - \frac{\partial \bar{u}_j}{\partial x_i} \right)^2} \quad (16)$$

where Re_t is the turbulent Reynolds number, S is the strain rate invariant, and Ω is the vorticity rate invariant.

3. Closure Model Calibration

The empirical coefficients of RANS closures were primarily tuned to provide a satisfactory compromise between accuracy and applicability to a broad range of flows. However, a review of the previous studies shows that these coefficients are not universal, implying they can be further adjusted for turbulent flow in case studies that were not originally considered [56]. This paper addresses the identified research gaps by developing an adjusted set of coefficients suitable for ABL dispersion flows within a generic form of urban settings.

A vital step in this process is deciding on the appropriate output constraints upon which to base the modification of the model coefficients. To improve the reliability of the air quality assessments in urban regions, a closure re-calibration is carried out, in which the pollutant concentration field and the flow field parameters (velocity and TKE) are set as the test parameters. The training case study is taken from a comprehensive full-scale measurement campaign ([45]), which integrates realistic environmental effects. A preliminary investigation of this dispersion data set shows the wide-ranging variation in the concentrations throughout the domain, ranging from 100 ppm near the source to 0.001 ppm downstream and away from the centerline (five orders of magnitude). Thus, specific considerations are required to define and evaluate the validation metrics for scalar

concentrations by including logarithmic criteria, as using the typical linear measures alone would bias the fitting toward high-concentration zones.

We have chosen three linear and two logarithmic measures to examine the agreement of our numerical predictions with full-scale dispersion measurements. These measures were defined by Chang and Hanna [57], and their reliability and effectivity were exhaustively tested using several full-scale dispersion datasets [58]:

$$FB = \frac{(\overline{C_o} - \overline{C_p})}{0.5(\overline{C_o} + \overline{C_p})} \quad (17)$$

$$NMSE = \frac{(\overline{C_o - C_p})^2}{\overline{C_o} \overline{C_p}} \quad (18)$$

$$MG = \exp(\overline{\ln C_o} - \overline{\ln C_p}) \quad (19)$$

$$VG = \exp[(\overline{\ln C_o} - \overline{\ln C_p})^2] \quad (20)$$

$$FAC2 = \frac{1}{N} \sum_{i=1}^N n_i \quad (21)$$

$$n_i = \begin{cases} 1 & \text{if } 0.5 \leq \left| \frac{C_o}{C_p} \right| \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

C_o and C_p are observed and predicted concentrations, respectively, while an overbar denotes the average over the dataset with N sampling points. FB and MG are fractional and mean geometric biases, respectively, metrics that measure the overall under/over-prediction. NMSE is the normalized mean square error, and VG is the geometric variance, quantifying the linear and logarithmic scatter of the predictions, respectively. FAC2 is another linear metric that denotes the fraction of predictions within the factor of 2 of the measurements. The necessity of analyzing all these measures lies in the nature of the atmospheric pollutant distribution. Linear metrics could be inordinately affected by the random extreme values, while logarithmic treatments might reflect a more balanced interpretation of them [58]. The ideal value of these performance measures and their acceptable ranges for field experiments are presented in Table 1. The application of these metrics is not limited to evaluating the predicted concentration field, as they also can be effectively considered to assess all the other flow parameters. Nevertheless, FAC2 might not be an adequately robust criterion for measuring the deviations between the observed and modeled velocity and TKE fields. Preliminary investigations indicated that even models with relatively deficient accuracy exhibit FAC2 larger than 85% for velocity and TKE. Hit-Rate (HR), on the other hand, sets more strict criteria and, therefore, could be an appropriate substitution for FAC2. As it can be deduced from Equation (22), HR gives the fraction of data points in which the relative deviation of predictions is within 25% of the measured values. P_i and M_i are the predicted and measured values of a parameter at data point i , respectively. An HR of at least 66% is required for the CFD model to be considered valid [59].

$$HR = \frac{1}{N} \sum_{i=1}^N m_i \quad (22)$$

$$m_i = \begin{cases} 1 & \text{if } \left| \frac{P_i - M_i}{M_i} \right| \leq 0.25 \\ 0 & \text{otherwise} \end{cases}$$

Table 1. The ideal value and the recommended ranges for validation metrics [58].

Validation Metrics	FB	NMSE	MG	VG	FAC2	HR
Ideal value	0	0	1	1	1	1
Acceptable ranges	−0.3–0.3	0–4	0.7–1.3	1–1.6	0.5–1	0.66–1

Having the focus parameters defined as discussed, employing a systematic and so-called “economic” optimization scheme is utterly crucial to modify the input variables (i.e., closure coefficients). In our work, a recommended optimization framework ([44]) is adopted and modified to carry out a rigorous calibration of the standard $k - \epsilon$ model (adjusting all five coefficients) in the context of a large model (i.e., full-scale urban dispersion flow).

The inherent uncertainties of the empirical coefficients, combined with their highly nonlinear and synergistic effects on the output variables, make any linear sensitivity approach ineffective in obtaining an optimized set. Despite that, simple screening techniques can still be adopted to determine the input variables with the most influence on the validation metrics [42]. Here, the relative contribution of each closure coefficient to the predicted concentration field will be assessed by quantifying its elementary effect [60]. For a selected validation metric denoted by G , which is assumed to be a function of the vector of coefficients $F = (f_1, \dots, f_m)$, the first derivative term of the Taylor series can be used to give a measure of the elementary effect for each coefficient:

$$E_i(F) = \frac{[G(f_1, \dots, f_i + \Delta_i, \dots, f_m) - G(F)]}{\Delta_i} \quad (23)$$

where m is the number of coefficients to be assessed, and Δ_i are step lengths determined to account for the variation of the i th coefficient within a given range. Furthermore, the mean \bar{E}_i and the standard deviation σ_{E_i} of the elementary effects related to each input variable must be calculated. The interpretation of these quantities provides valuable insight into the possibility of prioritizing some coefficients over others, potentially simplifying the optimization exercise [61]. Therefore, any coefficient that demonstrates negligible and quasi-linear impacts on output variables can be ignored at the optimization stage by keeping it at its standard value. The modified value of this coefficient can be obtained later by performing a sensitivity analysis while the optimized values for other closure constants are considered.

The amended list of the coefficients, together with their revised ranges from the previous step, should be put through the calibration process. The eddy-viscosity closure coefficients are pragmatic in definition; hence, their optimum value for producing reliable and accurate outputs (i.e., pollutant concentration distribution and flow parameters in this study) cannot be attained through a deterministic approach. An appropriate optimization scheme can be implemented in such cases to effectively navigate the performance variation caused by the random but targeted variation of the coefficients in the considered parameter space. To mitigate the effects of the available uncertainties on output variables, two intrinsically analogous and robust objective functions are defined to be used at two levels of the optimization [62]:

$$O_{f1} = \sum_{i=1}^l (G_i - \Gamma_i)^2 \quad (24)$$

$$O_{f2} = \sum_{i=1}^l [(\bar{G}_i - \Gamma_i)^2 + \sigma_{G_i}^2] \quad (25)$$

Equation (24) is employed for assessing the performance of each analyzed set, aiding in the deliberate choice among them in every optimization trial, while Equation (25) contributes to refining the ranges following each iteration of optimization. For the validation metric i , G_i is its calculated value for each considered coefficient set with \bar{G}_i and σ_{G_i} as its mean and standard deviation over each optimization iteration, respectively. Γ_i is the

ideal value for the i th validation metric, and l is the number of metrics included in this calibration study. The ranges for all input variables should be adjusted to minimize the objective function O_{f2} . The first term in Equation (25) ensures the accuracy of the model, while the second term ensures its reliability by reducing the variation in output responses within the proposed ranges.

Considering the complex and multidimensional parameter space in this work, the genetic algorithm (GA) was adopted and carefully adjusted to carry out the calibration exercise. GA is an optimization technique inspired by the natural selection analogy that facilitates reaching an optimal solution through evolution [63]. It offers balanced exploitation (cross-over) and exploration (mutation) of the search space, which effectively produces populations of new coefficient sets with a higher potential for success. Different adjustments of GA were tested to form a framework that best serves the re-calibration study in this work. Random selection of the coefficient sets from investigated ranges is carried out by employing the Monte-Carlo sampling method [64]. Using this selected population, the probability density function (PDF) of the validation metrics is calculated by repeatedly running CFD simulations. A rank-based selection based on the calculated O_{f1} values (Equation (24)), along with a uniform cross-over approach, is found to be more efficient in exploring the parameter space. A cross-over probability of 75% is considered, which falls within the recommended limits [65]. To explore new regions in the space that are not properly examined during the cross-over step, the generated off-springs (new sets) should be also mutated. In this regard, a Gaussian operator for the mutation step is adopted to introduce controlled perturbations to each coefficient of the sets with a probability of 10%. The new sets of coefficients generated in each iteration of the optimization process replace the sets from the previous step, and the optimal ranges for each individual coefficient are refined accordingly until the defined objective function O_{f2} reaches its converged minimum values.

Further investigations are required to evaluate the performance of the modified $k - \epsilon$ closure. To this aim, both the standard and optimized $k - \epsilon$ model will be considered next for a comparative study (Table 2). In the end, the generality of the proposed set will be further tested by modeling three supplementary test cases.

Table 2. Standard and revised values of $k - \epsilon$ empirical constants with their predetermined variation ranges.

Closure Coefficients	C_μ	$C_{\epsilon 1}$	$C_{\epsilon 2}$	σ_k	σ_ϵ
Original value [27]	0.09	1.44	1.92	1.00	1.30
Ranges	0.03–0.16	Equation (32)	1.20–3.20	0.50–1.40	0.30–1.30

4. Description of Case Studies

The unique dispersion dataset of the mock urban setting test (MUST) [45] was invaluable to calculate the objective function for calibrating the $k - \epsilon$ turbulence closure. MUST was sponsored by the U.S. Defense Threat Reduction Agency (DTRA) to provide a reliable resource that includes the meteorological and dispersion data to validate and verify the accuracy of the dispersion models and CFD simulations. An idealized urban-like geometry was designed in this experiment that consisted of a 10 by 12 array of shipping containers placed in the center of the test domain over relatively flat ground with a ground roughness of $z_0 = 0.045$ m. Figure 1 illustrates the schematic of the MUST geometry, along with the dimensions of the obstacles (i.e., containers), spacing between obstacles, and site configurations. In total, 68 trials were performed in which propylene gas (C_3H_6) was released from 37 different locations that can be classified into six main types with assigned letters A to F (Figure 1a). It is noteworthy that the elevation at which the aforementioned sources are located varies among all trials and should be identified prior to finalizing the CFD models.

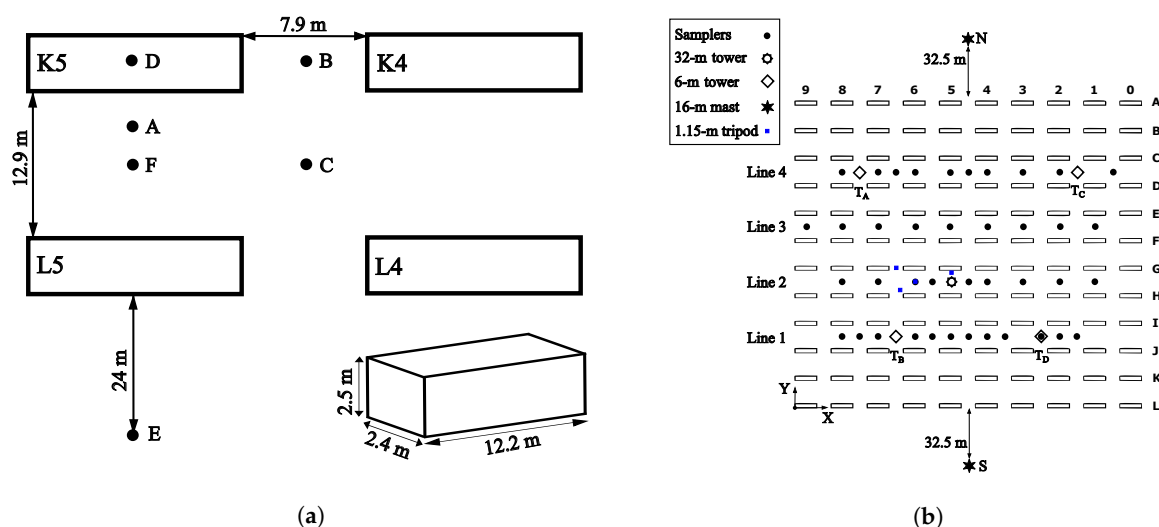


Figure 1. MUST experiment schematic; (a) Dimensions and source locations, (b) Top planar view with samplers locations.

Exhaustive measurements of the concentration field were carried out at 74 sampling points distributed throughout the array. Forty sensors were spread in four horizontal sampling lines of 1–4 (shown by black circles), and the remaining 32 sensors were installed on the central 32 m tower and four 6 m towers of T_A , T_B , T_C , and T_D , positioned in each quadrant (Figure 1b). The mean velocity and turbulence measurements were obtained using several sensors installed on all towers and masts (22 measurement points). That includes the 32 m central tower, four 6 m towers, two 16 m pneumatic masts upstream and downstream of the array, and four 1.15 m tripods. The specifications of all sampling points considered in this calibration study are given in Table 3. The MUST data were further processed for time-averaged studies by conditionally sampling all measurements to extract 200 s in each trial with the lowest recorded temporal variation in the mean upstream wind speed and direction [66].

Table 3. Specifications of all MUST sampling points considered in the calibration study.

Location	Quantity		Sensors Type		Elevation	
	Tracer	Flow	Tracer	Flow	Tracer	Flow
Lines 1, 2, 3, and 4	12, 9, 9, and 10	0	dPID ^a	—	1.6 m	—
Central tower	8	4	dPID	3D-SA ^b	1, 2, 4, 6, 8, 10, 12, and 16 m	4, 8, 16, and 32 m
T_A , T_B , T_C , and T_D	6 each	2 each	UVIC ^c	3A-UA ^d	1, 2, 3, 4, 5, and 5.9 m	2.4, and 6 m
S and N	0	3 each	—	2D-SA ^e	—	4, 8, and 16 m
Tripods	0	1 each	—	3D-SA	—	1.15 m

^a Digital photo-ionization detector; ^b 3-dimensional sonic anemometer; ^c Ultra-violet ion collector; ^d 3-axis Ultrasonic anemometer; ^e 2-dimensional sonic anemometer.

This paper considers four different trials: one as the training case to be used in the calibration study and three as test cases (TC-1 to TC-3), to evaluate the generality of the proposed coefficient set for various inflow boundary conditions and pollutant source locations. Table 4 summarizes all four cases with their main characteristics. This includes the averaged inflow velocity (S_{04}) and direction (α_{04}) at 4 m height, source type and its elevation (Z_s) with respect to ground, the tracer release rate (q_s), and the Monin–Obukhov length (L_{MO}). The positive value of α_{04} is measured counter-clockwise from the y -axis.

Table 4. Summary of selected trials of MUST experiment for RANS calibration study.

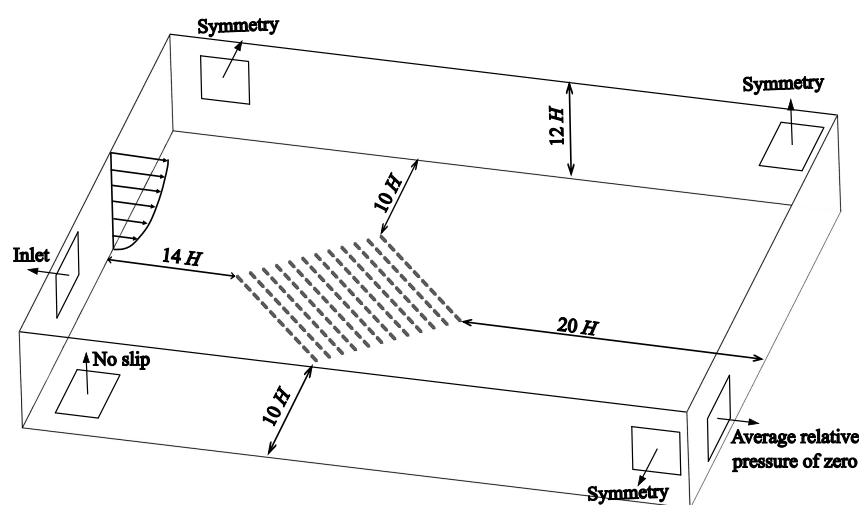
Case Study	Trial I.D.	Source Type	$q_s \left(\frac{1}{\text{min}} \right)$	$Z_s \text{ (m)}$	$S_{04} \left(\frac{\text{m}}{\text{s}} \right)$	$\alpha_{04} \text{ (deg)}$	$L_{MO} \text{ (m)}$
Training case	2682320	D	225	2.6	4.55	−39	170
TC-1	2681829	F	225	1.8	7.93	−41	28,000
TC-2	2692250	E	225	1.3	3.38	36	130
TC-3	2672150	A	200	0.15	2.30	36	150

5. CFD Simulation

Performing a reliable and accurate CFD simulation strongly depends on modeling decisions at the setup stage. To this end, the following subsections deal with the general CFD settings and considerations, mesh independence study, and the importance of remediating the artificial horizontal inhomogeneity (i.e., the undesirable streamwise gradients in the vertical profiles of flow variables).

5.1. General Settings

The pollutant dispersion within a compact urban setting is modeled on a high-resolution grid using the ANSYS CFX commercial code. The 3D steady-state RANS equations coupled with the Eulerian transport equation are discretized and solved by employing the hybrid finite element/volume method. High-resolution schemes are considered to evaluate advection terms and turbulence numerics, and the Rhie–Chow algorithm was adopted to implement pressure–velocity coupling. To ensure convergence of the solution, target values of root mean square (RMS) normalized residuals were set to 10^{-6} and 10^{-9} for flow variables and tracer concentrations, respectively. The minimum size of the computational domain was determined based on the recommendations of Tominaga et al. [67] and Franke et al. [68]. The additional domain sizes were examined to specify the proper dimensions for the MUST geometry to eliminate the possibility of any artificial flow accelerations around the buildings and destructive backflow at the boundaries. Taking the height of the obstacles as H , distances from the MUST array are revised to $14H$ from the inlet, $10H$ from lateral boundaries, $20H$ from the outlet, and $12H$ from the top boundary. This ensures that the results are domain independent. Figure 2 illustrates a schematic of the computational domain. As can be seen, the rectangular domain is oriented so that the inflow wind stream is perpendicular to the inlet and outlet faces.

**Figure 2.** Computational domain.

The next crucial step in setting up a CFD model is to define the proper constraints at the limits of the computational domain (i.e., boundary conditions). Assuming an equilibrium

atmospheric surface layer, the inflow velocity and turbulence profiles can be approximated by [38]:

$$U(z) = \frac{u_\tau}{\kappa} \ln\left(\frac{z + z_0}{z_0}\right) \quad (26)$$

$$k(z) = \frac{u_\tau^2}{\sqrt{c_\mu}} \quad (27)$$

$$\varepsilon(z) = \frac{u_\tau^3}{\kappa(z + z_0)} \quad (28)$$

where u_τ is the friction velocity, κ is the Von Kármán constant with the value of 0.4 [38], and z_0 is the aerodynamic roughness. Having a reference velocity at a reference height (Table 4), Equation (26) can be used to calculate u_τ for each case study.

An average relative pressure of zero is specified at the outlet boundary. The top and lateral limits are set as symmetry. The faces of each container are treated as smooth walls with no-slip boundary conditions, while the ground is considered to be a rough wall with a physical roughness of z_0 . The scalable wall function is employed to resolve the velocity and turbulence quantities near the walls. The given uniform ground roughness also represents the roughness of the upstream terrain that is not considered in the domain. This implies having a fully-developed atmospheric flow at the inlet, which, ideally, should not reflect any streamwise gradient as it progresses through the domain [69]. However, horizontal heterogeneity is a known issue in ABL flow simulations that needs to be carefully evaluated before claiming the reliability of the results [68].

5.2. Grid Convergence Study

Generating a high-quality grid has a pivotal impact on reducing the discretization error and enhancing the convergence of the solution. In the current study, a nested domain was defined in which the inner domain encloses the MUST array. As a consequence of taking this approach, the large empty portion of the domain was meshed using structured hexahedral elements, which improves the overall rate of convergence [70]. Unstructured tetrahedral cells are used to mesh the regions near the blocks within the inner domain, while several prismatic layers are considered near the solid surfaces (i.e., ground and faces of the blocks) to properly capture boundary layer gradients.

Following the “best-practice” guidelines in computational wind engineering, the sensitivity of the CFD model results to the grid resolution should be carefully assessed [67]. In this regard, three successive grid resolutions of coarse, medium, and fine were examined with 6.39, 9.68, and 14.62 million computational nodes, respectively. The vertical profile of normalized velocity at the central 32 m tower and the horizontal profile of normalized pollutant concentration at the second sampling line are compared to assess the uncertainty of results due to the grid size. As Figure 3a illustrates, there are trivial deviations between the resultant velocity profiles using all three grid refinements, which necessitates evaluating a more sensitive variable (i.e., concentrations). Figure 3b, however, implies that employing the grid with medium refinement would be the appropriate choice with respect to the computational cost and dependency of the predictions to mesh sizes.

Additionally, another grid sensitivity study proposed by Celik et al. [71] was conducted to calculate the grid convergence index (GCI). The predicted pollutant concentration at 74 sampling points was selected as the variable of interest. For detailed mathematical steps and calculations of this investigation, please refer to our previous paper [35]. Considering the second-order discretizations used in this work and assuming a safety factor of 1.25 as suggested by Roache [72], the average relative error and GCI for the two finer grids are 2.09% and 5.94%, respectively. Comparing these quantities with ones of the coarse-medium case (average relative error of 4.93% and GCI of 14.05%) further justifies the decision to use the grid with medium refinement. For the selected grid, the building edges were divided into 20 elements and a 5-layer inflation region was considered for all the solid surfaces with

a growth rate of 1.07 to ensure y^+ values were well within the acceptable range of 30 to 300. It is worth noting that the simulations were run on a workstation with two 18-core Intel Xeon Gold 5220 CPUs (36 cores in total) and 256GB of RAM. On average, a fully converged solution for each CFD run was achieved in approximately 4–5 h.

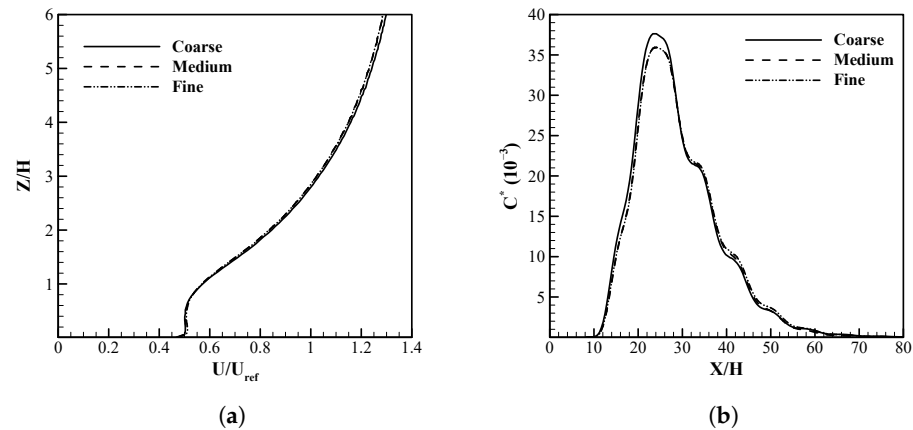


Figure 3. Mesh sensitivity of (a) normalized velocity at 32 m central tower and (b) normalized concentration at the second horizontal sampling line.

5.3. Horizontal Homogeneity

To ensure that the numerical uncertainty is kept to a minimum, the horizontal homogeneity of the flow must be first verified. A lack of homogeneity in the incident velocity profile can lead to erroneous results [70]. To account for the roughness of the wall, the majority of the commercial codes modify the wall function based on the equivalent sand grain roughness (i.e., k_s) approach [69]. As shown below, this method seems to capture the homogeneity of the flow better compared to other methods, such as replacing the no-slip with a constant shear boundary condition [73]. In order to incorporate the wall roughness into the CFX simulation, Equation (30) along with the modified wall function, Equation (31), are used [69]:

$$k_s = 29.6z_0 \quad (29)$$

$$k_s^+ = \frac{u_\tau k_s}{\nu} \quad (30)$$

$$\frac{U_p}{u_\tau} = \frac{1}{\kappa} \ln \left[\frac{u_\tau y_p}{\nu(1 + 0.3\nu k_s^+)} \right] + 5.2 \quad (31)$$

where U_p is the velocity at the center of the first cell next to the wall, and k_s^+ is the dimensionless sand grain roughness. These equations and the proper value of aerodynamic roughness were used to study the homogeneity of the incident profile in an empty domain.

Figure 4 shows the velocity, TKE, and dissipation rate at three cross-sections in an empty domain. As can be seen, the horizontal homogeneity for the velocity and dissipation rate was near completely achieved with a mean average error of 1.3% and 1.7%, respectively. However, the TKE profile seems to display a small degree of streamwise inhomogeneity in regions close to the ground, reaching a mean average error of 2.4%. The analytical solution to the horizontally homogeneous atmospheric boundary layer requires the production of turbulence to be balanced by the dissipation at a constant rate leading to a uniform value for the TKE normal to the ground. However, as indicated in the work of the previous researchers [23] and from Figure 4, TKE varies with distance to the ground. The reason for this over-generation of turbulence in the vicinity of the wall in the $k - \epsilon$ model was previously discussed by Richards [74], and the inconsistent discretization of the turbulence production term at the wall is described as the source of this excessive generation.

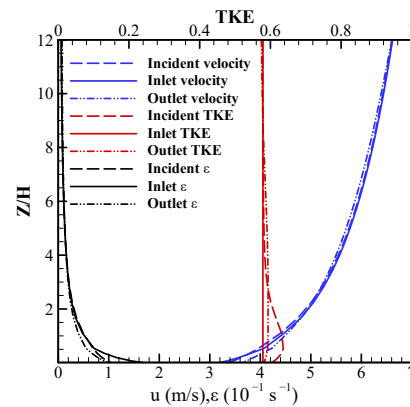


Figure 4. The velocity, TKE and turbulence dissipation rate profiles in an empty domain.

6. Results and Discussion

The suggested numerical scheme and the described methodology for closure model optimization in previous sections are employed to improve the accuracy and reliability of the RANS method in modeling the ABL dispersion flow within compact urban regions. First, the results of the optimization study are presented and discussed. Later, the modified closure is evaluated by comparing its resulting performance measures with ones of the standard model and other analogous works. Finally, the generality of the proposed model is examined by simulating three different case studies.

6.1. Closure Re-Calibration

Comprehensive optimization of the $k - \epsilon$ model demands the re-calibration of all five empirical coefficients. Given the large size of the computational domain in this work (i.e., full-scale compact urban setting), this process becomes prohibitively expensive in terms of computing time and power. In this regard, taking specific considerations is crucial to make this study more feasible and practical. A reasonable approach to achieve this goal is to reduce the number of variables that require simultaneous tuning. The necessity of maintaining a horizontally homogeneous flow in the atmospheric surface layer encouraged Richards to devise a relationship between the constants of the $k - \epsilon$ closure [38]. With this condition, the turbulence model, inflow profiles, and the resulting ground shear due to the aerodynamic roughness would be in equilibrium. Therefore, four out of the five constants (C_μ , $C_{\epsilon 1}$, $C_{\epsilon 2}$, and σ_ϵ) must take appropriate values to satisfy Equation (32). Accordingly, one of these four coefficients can be arbitrarily excluded from the re-calibration process, as it can be determined as a function of the remaining three. Following the recommendation of the previous studies, $C_{\epsilon 1}$ was omitted from the optimization process in this work [37,75].

$$C_{\epsilon 1} = C_{\epsilon 2} - \frac{\kappa^2}{\sqrt{C_\mu} \sigma_\epsilon} \quad (32)$$

A further simplification of closure optimization may be possible by investigating the elementary effects of each coefficient on the objective parameters. To this aim, each constant is varied independently in a predetermined range, while other constants, except for $C_{\epsilon 1}$, are kept at their original values. The preliminary ranges for closure coefficients in this work are specified conservatively based on the previous studies reviewed in Section 1 (Table 2). Considering 25 equally spaced values within each range, a total of 100 CFD runs was carried out and processed for this part of the study.

Prioritizing the accurate prediction of the concentration field, the elementary effects of input variables (coefficients) on this parameter are presented in Figure 5. As it can be deduced, increasing C_μ leads to more accurate predictions of the concentration field (higher FAC2), while similar arguments cannot be made for the other constants. For instance, higher levels of FAC2 are generally associated with smaller values of $C_{\epsilon 2}$. Ranging σ_ϵ

from 0.3 up to values of approximately 1.15 enhances the model accuracy, whereas its further increase reduces the resulting FAC2. The variation in σ_k within its range exhibits considerably milder but analogous impacts on FAC2 to the ones in σ_ε . As discussed in Section 3, despite the undeniable importance of FAC2 as a measure of accuracy, the quality of the predicted concentration field should not be judged solely based on this metric.

For this reason, two linear measures of FB and NMSE for the predicted concentration field, and their logarithmic counterparts, MG and VG, are also included in the objective function. Figure 5 displays a drop for both FB and MG as $C_{\varepsilon 2}$ and σ_ε increase, while opposite trends are shown for C_μ and σ_k . Given the definition of the validation metrics, smaller amounts of FB and MG indicate stronger scalar dispersion and lower predicted concentrations at sampling points. It is worth noting that a VG greater than 1 and positive values of FB reflect under-predictions of the concentration field. However, due to the unavoidable presence of compensation errors in the calculations of FB and MG, there may be instances of significantly inaccurate predictions in which these metrics take their ideal values. To avoid such a misinterpretation of the systematic errors, it is necessary to simultaneously calculate and explain NMSE and VG metrics, which also quantify the available random errors (scatter of predictions from measurements) [58]. Lower values of NMSE and VG that correspond to smaller scatters are shown in Figure 5 to be gained by larger values of C_μ and σ_k within ranges of 0.1 to 0.16 and 0.9 to 1.4, respectively. Conversely, adopting smaller values of $C_{\varepsilon 2}$ between 1.2 and 2 considerably mitigates the scatters. For σ_ε ranging from 0.3 to 0.6, the drastic fall of VG compared to the relatively moderate reduction of NMSE suggests that data points with exceedingly low concentrations may have overly influenced this logarithmic measure. Therefore, it is expected that the optimal value of σ_ε falls within the range of 0.8 to 1.2, which is in agreement with the recommended ranges acquired by analyzing the other validation metrics.

A further examination of Figure 5 implies minimal sensitivity of all validation metrics in response to the variation in σ_k compared to the rest of the coefficients. Such observation advocates the exclusion of σ_k from the optimization study to reduce computational expenses. However, the reliability of this decision must be assessed beforehand by quantifying the elementary effects of closure coefficients on output parameters using Equation (25). Heat maps are used as shown in Figure 6 to qualitatively compare the mean (\bar{E}) and standard deviation (σ_E) of the quantified contribution of each coefficient to the output parameters. By employing this color-coded scheme, we can promptly visualize the relative influence each parameter has on the validation metrics. This approach provides a concise assessment without the need for overwhelming numerical data and could serve as a complementary visualization to the other presented figures in this work. A larger value of \bar{E} for an input variable suggests its relatively greater net impact on the validation metrics, while a larger quantity of σ_E reflects its highly non-linear response or stronger interaction effects by other input variables on output parameters [60].

As Figure 6 implies, the variation in C_μ shows the highest order of influence on all validation metrics with the most significant level of interaction effects with other factors. This remark was expected as C_μ not only contributes to the modeled eddy viscosity (Equation (9)) and, consequently, scalar diffusion (Equation (11)), it also plays a part in the estimation of the inflow TKE profile (Equation (27)). The sensitivity of the validation metrics to variation in σ_ε and $C_{\varepsilon 2}$ show roughly the same order of significance, which can be categorized as the second and third most influential factors, respectively. Finally, the calculated \bar{E} and σ_E of the output responses to σ_k variation hold the lowest orders relative to ones of other coefficients. Accordingly, it is reasonable to claim that the validation metrics are considerably less sensitive to σ_k , and this coefficient can be ignored at the optimization step. This statement is moreover supported by some of the previous studies [37,56]. Given the fact that σ_k does not also correlate with other coefficients through Equation (32), its modified value can be obtained later by conducting a direct sensitivity study.

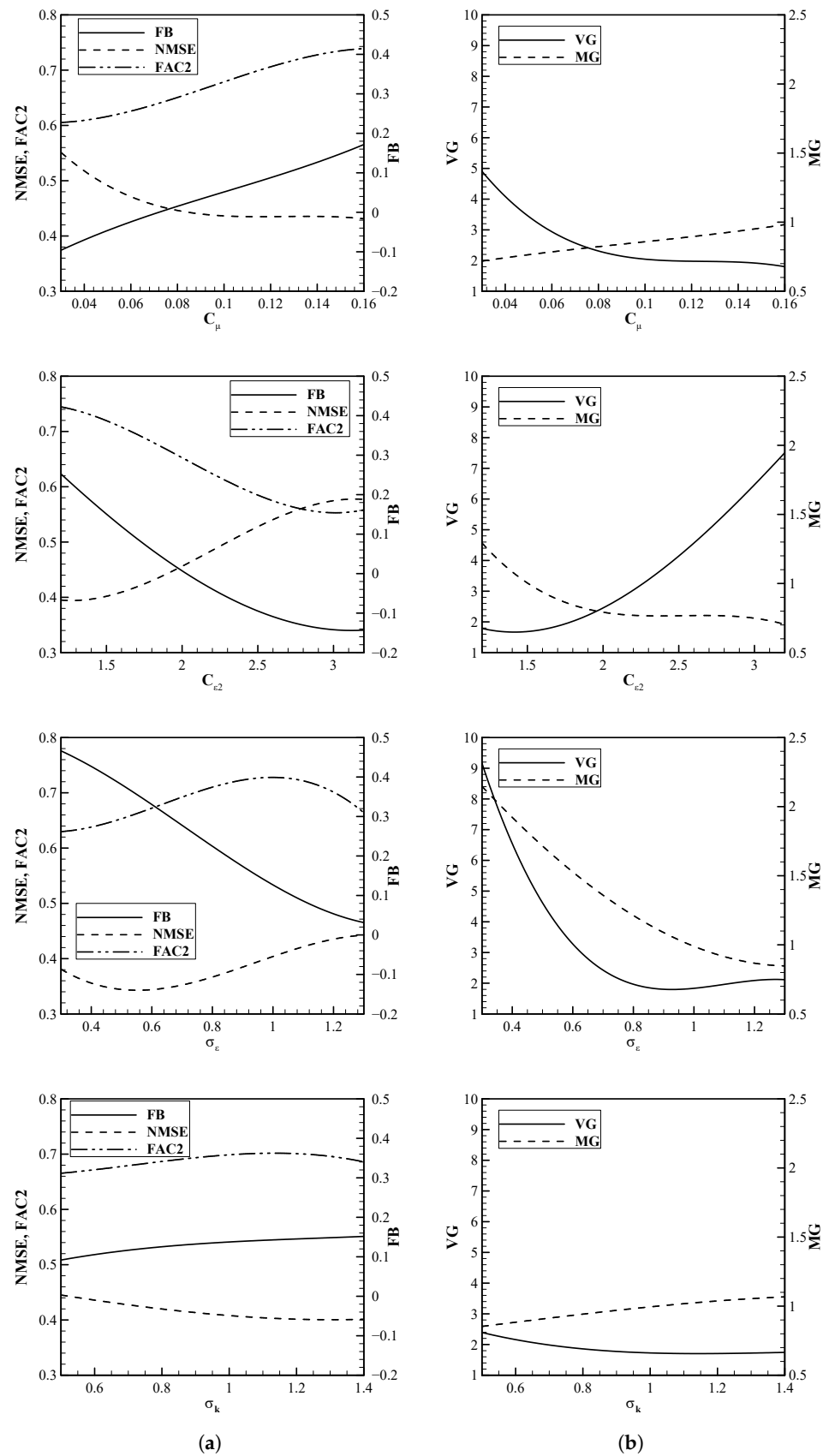


Figure 5. Elementary effects of closure coefficients on (a) linear and (b) logarithmic validation metrics.

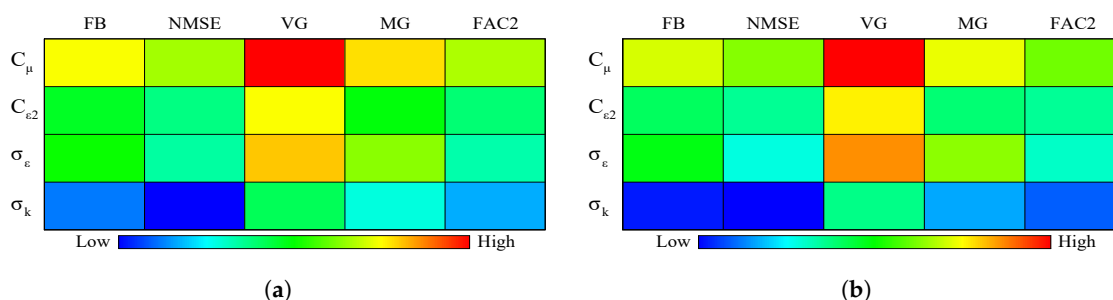


Figure 6. Comparing (a) \bar{E} and (b) σ_E of the elementary effects of closure coefficients on the validation metrics.

The distinct responses of the model's accuracy to the variation in each closure constant demonstrate the necessity of implementing a more rigorous approach to navigate towards an optimized set. To this aim, the Monte-Carlo sampling method is used to generate random selections of C_μ , $C_{\epsilon 2}$, and σ_ϵ within their associated ranges, while $C_{\epsilon 1}$ is estimated by Equation (32), and σ_k is kept at its standard value of 1. Given the definition of the proposed objective functions (Equations (24) and (25)), the GA optimization scheme aims at revising these ranges after each main iteration to detect ones in which the model coefficients collectively correspond to a minimum value of O_{f2} . Uniform distributions of probability are assumed for all coefficients in their variation ranges, and 40 sample sets are generated randomly to produce the initial population for the optimization process.

Figure 7 reveals scatter plots presenting the PDFs of FAC2 and HR for the predicted concentration and TKE fields, respectively. These plots offer valuable insights into the variation and distribution of these metrics across the last 100 investigated coefficient sets during the re-calibration process. As can be seen, both the concentration and TKE fields exhibit noteworthy sensitivity to the variations among coefficient sets, which are evidenced by the widespread and diverse clustering of data points in the provided PDFs. The observed similarity and consistency between both the output responses (i.e., FAC2 and HR), further attest to the strong dependency of the predicted concentration field on the accuracy of the predicted TKE field, which justifies the definition of the objective function as described in Section 3.

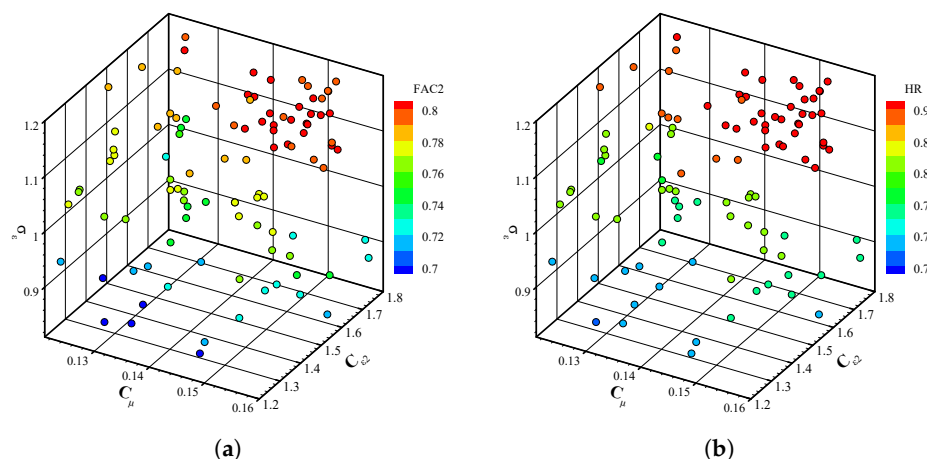


Figure 7. Variation in validation metrics during the re-calibration (a) FAC2 for pollutant concentration and (b) HR for TKE.

To better investigate the dependency of the model's output responses on closure sets, Figure 8 illustrates surface plots of FAC2 resulting from variations in the considered closure constants during the re-calibration process. As can be concluded, higher fractions of predictions within FAC2 of measurements are generally achieved for quantities of C_μ ranging from 0.13 to 0.16. Increasing C_μ from its standard value of 0.09 could result in

relatively lower eddy viscosity and inflow TKE to be estimated, which could remediate the known flaw of the standard $k - \varepsilon$ model: excessive over-prediction of the TKE [76]. The shown results in Figure 8 further suggest that decreasing σ_ε from its original value of 1.3 increases the probability of acquiring predictions with overall higher associated FAC2. Lower quantities of this constant basically result in promoting the diffusive transport of the rate of viscous dissipation; however, a general conclusion cannot be drawn due to the complex linked relationship of the TKE and ε and the heavy modeling applied to Equation (8) [77]. The optimal values of σ_ε vary between 1 and 1.2. $C_{\varepsilon 2}$ appears as a factor for the sink term available in the ε transport equation and is expected to substantially affect the predicted turbulence field and, consequently, the resulting pollutant concentration field. The most accurate representation of the pollutant concentration field for the generic case of a compact urban setting is shown to be obtained by smaller values of $C_{\varepsilon 2}$ than its standard value, differing between 1.6 and 1.8.

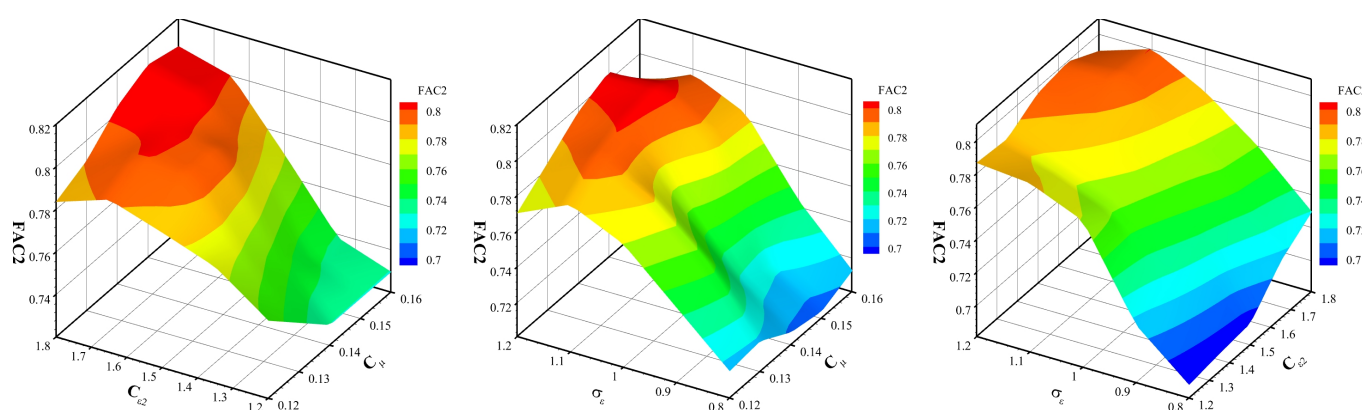


Figure 8. Surface plots of FAC2 variation for pollutant concentration field during the calibration process.

The quality of the predictions within the detected ranges for closure constants that contribute to the highest calculated FAC2 should be further examined. In this regard, the concentration-related output responses of the remaining validation metrics to different sets of closure coefficients are presented in Figure 9. Noting that the calculated NMSE remains well inside the acceptable limits and near its ideal value ($0.35 \leq \text{NMSE} \leq 0.55$), it was decided to mainly focus on the FB, VG, and MG metrics and omit the NMSE variation contours in this paper. As can be seen, both the linear and logarithmic measures of the model's systematic errors (i.e., FB and MG, respectively) exhibit roughly similar trends in response to the closure modification.

The analysis of Figure 9 suggests that the most desirable measures of FB and MG are attained as C_μ takes values between 0.12 and 0.15, while $1.68 \leq C_{\varepsilon 2} \leq 1.80$ and $1.12 \leq \sigma_\varepsilon \leq 1.20$. The variation in FB and MG within the given ranges reflects an overall under-prediction of the concentration field compared to the observations. On the other hand, evaluation of the resulting scatter from measurements (i.e., VG) identifies the optimal ranges of coefficients as $0.14 \leq C_\mu \leq 0.16$, $1.50 \leq C_{\varepsilon 2} \leq 1.80$, and $1.05 \leq \sigma_\varepsilon \leq 1.20$. These findings further emphasize the importance of collectively assessing the validation metrics, as well as the synergistic effects of closure coefficients on model outputs. The optimal ranges can be extracted from the intersections of the identified spans to ensure the greatest reliability of predictions (FAC2). The proposed ranges in this work are $0.14 \leq C_\mu \leq 0.15$, $1.68 \leq C_{\varepsilon 2} \leq 1.80$, and $1.12 \leq \sigma_\varepsilon \leq 1.20$, in which a closure set of $C_\mu = 0.147$, $C_\varepsilon = 1.344$, $C_{\varepsilon 2} = 1.693$, and $\sigma_\varepsilon = 1.196$ produces the most accurate and reliable predictions.

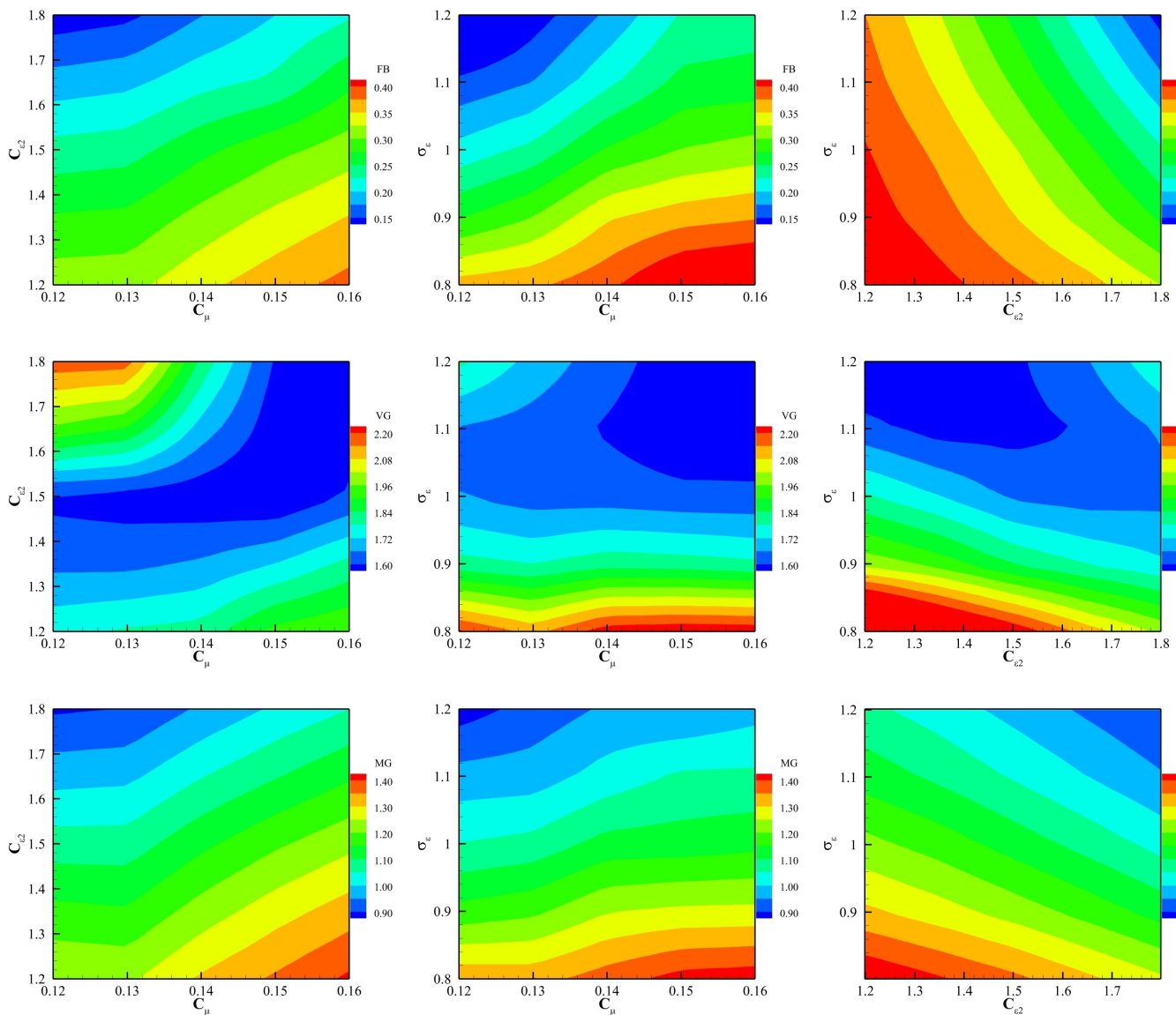


Figure 9. Variation in FB, VG, and MG for pollutant concentration field during the calibration process.

The optimal value of σ_k is successively derived by conducting a sensitivity study using the standard $k - \varepsilon$ model in which the modified coefficients are implemented. As expected, all the output parameters demonstrate relatively weak sensitivity to the variation of this coefficient, with FAC2 being the least responsive (Figure 10). Increasing σ_k weakens the diffusive transport of TKE, consequently leading to larger turbulence mass diffusivity (i.e., D_t). This behavior of D_t enhances the under-prediction of the pollutant concentrations at the plume centerline, which justifies the consistent increase in FB shown in Figure 10a. Considering all the validation metrics together while prioritizing FAC2 (i.e., the model's accuracy), the range of $0.87 \leq \sigma_k \leq 1.00$ is suggested, with $\sigma_k = 0.927$ being the optimal value.

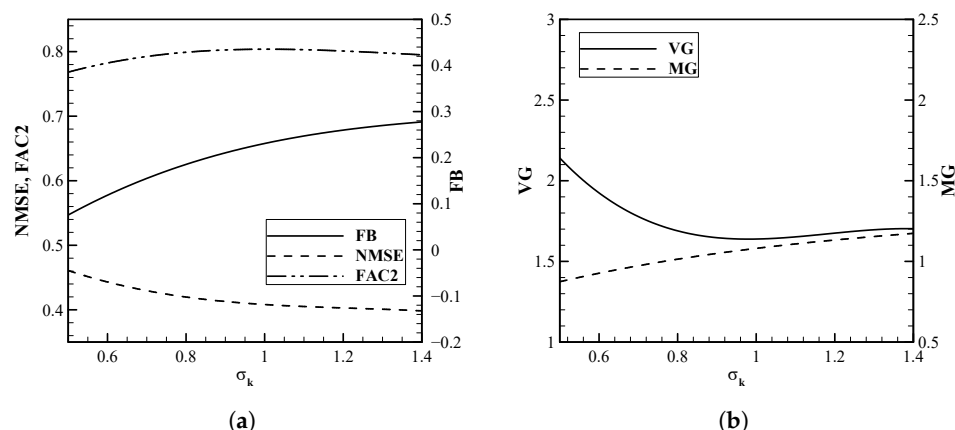


Figure 10. Sensitivity of the (a) linear and (b) logarithmic validation metrics to the variations in σ_k .

6.2. Performance and Generality Evaluation

The performance of the improved closure model with the proposed coefficient set ($C_\mu = 0.147$, $C_{\epsilon 1} = 1.344$, $C_{\epsilon 2} = 1.693$, $\sigma_\epsilon = 1.196$, and $\sigma_k = 0.927$) is evaluated in this section. Figure 11 depicts the iso-surfaces of pollutant concentration ($C^* = 10^{-2}$) flooded by TKE contours that were predicted by all the considered revisions of the $k - \epsilon$ turbulence model. Even though the overall form of the predicted plume by both the optimized and standard turbulence closures display roughly similar shapes due to the interactions between flow and structures, the predicted volumes of the selected iso-surface exhibit clear differences. As for the TKE, the proposed model in this work resulted in considerably lower values compared to the other revision (i.e., the standard version). The distinct differences observed among these models highlight the necessity of such re-calibration practices and lays the foundation for a more meticulous investigation of their performances. In this regard, rigorous statistical comparisons are subsequently required to further elucidate the strengths and weaknesses of the studied models.

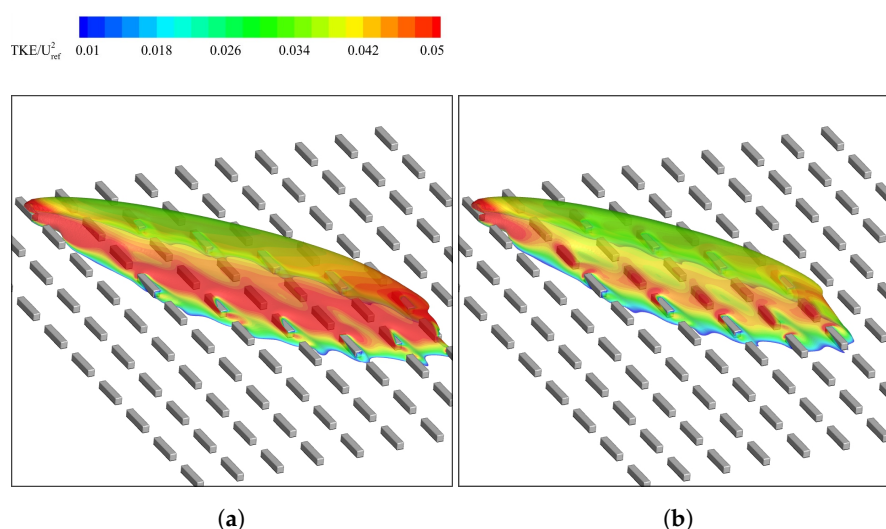


Figure 11. Normalized concentration iso-surfaces ($C^* = 10^{-2}$) flooded with TKE, predicted by (a) standard, and (b) optimized revision of $k - \epsilon$ turbulence model.

The validation metrics of the predicted concentration field by all revisions of the standard $k - \epsilon$ are calculated and presented in Table 5. The most accurate reproduction of the concentration field was obtained using the optimized closure in this work, with 80% of the predictions within FAC2 of measurements. Evaluating all validation metrics together, the decisive superiority of the optimized closure for the generic case of a compact

urban setting is apparent. The positive quantities of FB show the general under-prediction of the concentration field using all these revisions. The presented MG values further support this observation. Considering FB values of 0.13 and 0.21 by the original and optimized versions, respectively, these models demonstrate slight under-predictions of the concentration field. In terms of the recorded scatters, the optimized model in the current study strongly outperforms the other revision with $VG = 1.63$.

Table 5. Performance evaluation of the modified closure model in predicting the pollutant concentration field.

Closure Revision	FB	NMSE	MG	VG	FAC2
Standard [27]	0.13	0.41	0.98	1.82	0.72
Current study	0.21	0.40	1.05	1.63	0.80

A more rigorous interpretation of statistical measures in Table 5 is possible by analyzing the predicted concentration profiles throughout the domain. Figure 12 maps the concentration field at four horizontally distributed sampling lines. As can be seen, the proposed models by the previously revised set under-predict the concentrations at all sampling lines (more severely farther from the source), which justifies its corresponding larger values of FB and MG. On the contrary, the original and optimized models over-predict the concentrations at the first two sampling lines. As the plume progresses downstream, the intensity of over-prediction decreases gradually until concentrations are entirely under-predicted at sampling line 4. It is also worth mentioning that the variation in the field measurements along the horizontal line closely resembles the profiles generated by the present optimized model.

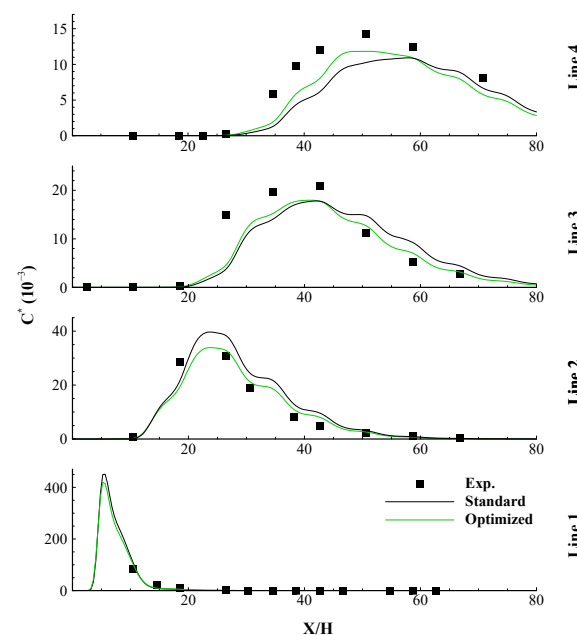


Figure 12. Comparison of the predicted concentration fields at horizontal sampling lines.

The predicted vertical profiles of pollutant concentrations are shown and compared in Figure 13. Given the direction of the upstream wind flow, two towers, T_A and T_D , did not detect any tracer quantities and were, thus, excluded from this study. As illustrated, the optimized model predicts relatively lower concentration levels along the vertical sampling lines, resulting in improved agreements with the field measurements. It should also be noted that the reproduced plume by the optimized models successfully exhibits the expected Gaussian shape as it passes through the selected vertical lines. To further evaluate the overall capability of the proposed model, a scatter diagram for all 74 sampling points is

presented in Figure 14. The improved accuracy of the predictions by the optimized model is apparent by showing smaller scatters, which further supports the lowest calculated VG of 1.63 (Table 5). In addition to sampling points with higher concentrations (near the source and plume centerline), the profound improvement in predictions for points with lower concentrations (near the plume's edges and far from the source) is also evident, which implies a more accurate reproduction of the pollutant spread throughout the domain. As a result, a greater fraction of predictions is shown between the FAC2 lines using the optimized model, with almost no data point outside the FAC5 lines.

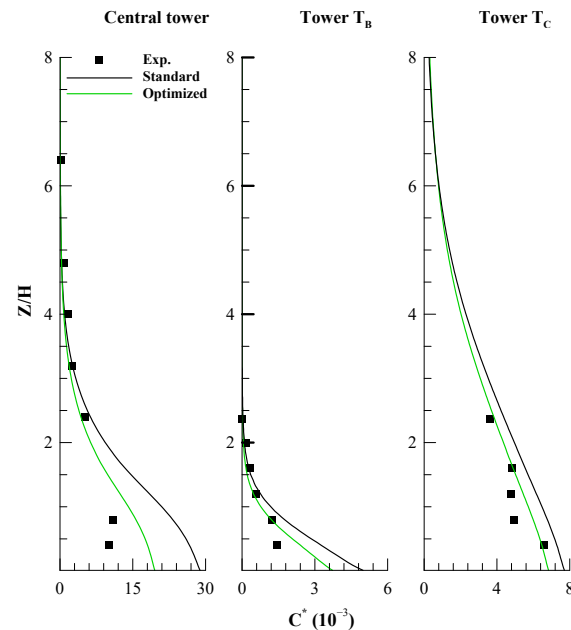


Figure 13. Comparison of the predicted concentration fields at vertical sampling lines. The solid line represents the exact match between predictions and observations.

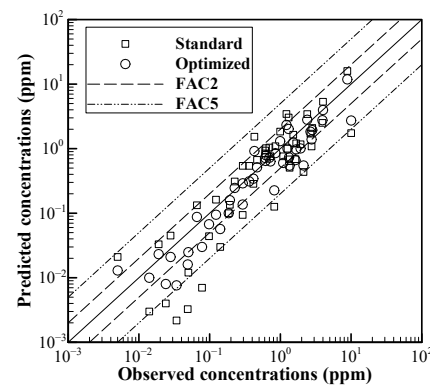


Figure 14. Scatter plot between the observed and predicted concentrations at all sampling points for the standard and optimized closure sets.

The capability of the proposed coefficient set in capturing the wind and turbulence fields must also be examined to ensure the reliability of predictions exhaustively. Table 6 compares the calculated validation measures for both the predicted velocity and TKE fields obtained by optimized and standard versions of the $k - \epsilon$ model. As previously discussed in Section 3, FAC2 is not an adequately strict metric for this part of the evaluation study and is replaced by HR. It should also be noted that the measurements of one of the sensors of tower T_A , both sensors of tower T_D , and one sensor of the upstream mast S are missing from the MUST data set and therefore are excluded from calculations. In

terms of the flow field, approximately similar performances by both models are observed, with validation metrics well within the acceptable ranges. The predicted velocity fields are slightly overestimated (negative values of FB and MG below 1), and quantities of NMSE and VG suggest low scatter. A minimum HR of 66% is required to consider predictions as valid, which is obtained by all models.

As is expected, implementing different versions of coefficient sets leads to contrasting representations of the turbulence field. As suggested by the validation metrics, the standard revision of the $k - \varepsilon$ coefficient sets substantially overestimates TKE compared to the optimized model. The cross-comparison of the considered coefficient sets and their associated impacts on the model's outputs argue that large values of $C_{\varepsilon 2}$ might be the most influencing factor in TKE over-prediction. The appearance of a greater quantity of $C_{\varepsilon 2}$ in the form of a factor to the source term in Equation (8), promotes the production of ε . Considering the coupled relationship of Equations (6) and (8), it is expected that the production rate of TKE is large where its dissipation rate is large [24], which justifies considerably lower values of TKE predicted by the optimized set in this work compared to ones obtained by the standard version.

Table 6. Performance evaluation of the modified closure model in predicting the velocity and TKE fields for the training case study.

Closure Revision	Velocity				
	FB	NMSE	MG	VG	HR
Standard [27]	−0.03	0.03	0.93	1.05	0.72
Current study	0	0.03	0.96	1.04	0.78
	TKE				
	FB	NMSE	MG	VG	HR
Standard [27]	−0.15	0.05	0.87	1.04	0.67
Current study	−0.03	0.02	0.98	1.02	0.94

The qualitative examination of the predicted turbulence field can be carried out by investigating the resulting TKE profiles at several locations within the building array. Figure 15 depicted the vertical variations of TKE at two different sample locations: at the center of the array (central tower) and on tower T_C , positioned nine rows into the array. It is worth noting that the predicted TKE fields show major dissimilarities at lower elevations, where the flow is primarily affected by the presence of objects, but they asymptotically converge as the array's influence disappears aloft. Figure 16 is provided to facilitate a point-to-point comparison and evaluation of the predicted flow field parameters using the standard and optimized $k - \varepsilon$ models at the 18 measuring points (which has measurements) spread across the MUST test domain. The statistical data for the predicted velocity field presented in Table 5 are further supported by Figure 16a, showcasing insignificant differences between the performances of these two models. Furthermore, a clear trend is observed for both the velocity and TKE, implying that the standard version of the turbulence model generally returns higher values of these parameters. The considerable difference in the predicted TKE values, despite what is noted for the velocity field, highlights the ample improvement achieved by the optimized model in the predictions. As was already mentioned, TKE plays a pivotal role in the accurate representation of the concentration field through the definitions of the eddy viscosity, Sc_t , and consequently, the eddy diffusivity. This improved agreement can be attributed to the optimized model's ability to represent turbulent mixing and dispersion more accurately, resulting in a more realistic depiction of the concentration distribution in the flow field.

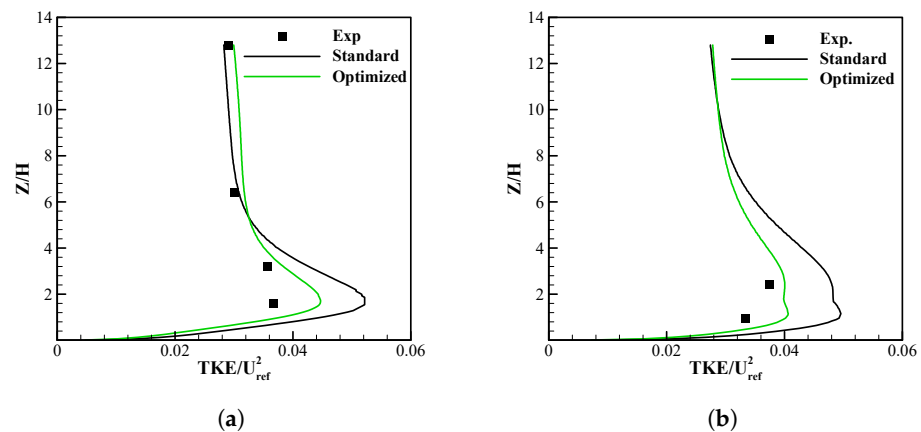


Figure 15. Vertical profiles of the predicted TKE (a) at 32 m central tower and (b) at the tower T_C .

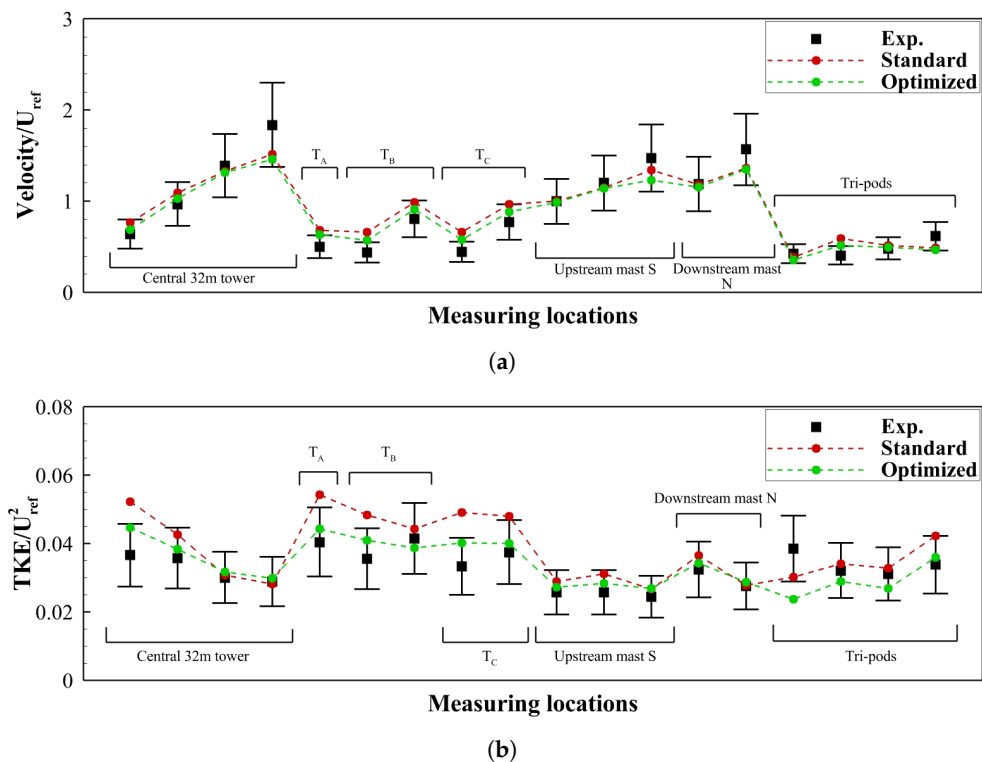


Figure 16. Point-to-point evaluation and comparison of the predicted flow field across the test domain: (a) Velocity, (b) TKE. The error bars represent a 25% deviation from the experimental measurements.

As for the final step of this optimization study, the generality of the proposed set must be tested prior to claiming its utility. To this aim, three test cases are considered (Table 4), in which the performance of the optimized model is comparatively evaluated. As can be seen, all the test cases are selected to cover a diverse range of situations in terms of the tracer source type and location, array orientation, mean wind speed, and atmospheric condition. Referring to Figure 1b, the source in the training case is roof-based (i.e., type D) and positioned on container J9. For the test cases of TC-1, TC-2, and TC-3, the sources are, respectively, located between containers K8 and L8 (type F), 24 m upstream of container L1 (type E) and immediately upstream of container J3 (type A). In order to consider the distinct array orientations, cases with different incident directions of wind flow that generate unique flow structures are chosen. The precise prediction of the scalar concentration field strongly depends on the accurate reproduction of flow and turbulence fields [15]. Therefore, by defining these test cases, opportunities to assess the extent of the applicability of the

developed coefficient set were created. To further explore the limitations and uncertainties associated with the proposed framework, two atmospheric stability classes of fully neutral (TC-1), and stable near neutral (TC-2 and TC-3) are considered. Detailed evaluation studies are carried out for all three test cases, for which the corresponding statistical measures are given in Table 7.

Table 7. Validation metrics for three selected test cases during the generalizability study.

Case	Revision	Velocity				
		FB	NMSE	MG	VG	HR
TC-1	Standard	−0.09	0.02	0.91	1.03	0.78
	Optimized	−0.07	0.01	0.94	1.03	0.89
TC-2	Standard	0.13	0.08	1.03	1.07	0.67
	Optimized	0.15	0.06	1.05	1.07	0.72
TC-3	Standard	0.25	0.30	1.07	1.16	0.78
	Optimized	0.26	0.28	1.09	1.15	0.78
TKE						
		FB	NMSE	MG	VG	HR
TC-1	Standard	0.05	0.03	1.07	1.05	0.83
	Optimized	0.09	0.01	1.09	1.01	0.94
TC-2	Standard	0.13	0.04	1.07	1.08	0.72
	Optimized	0.18	0.01	1.11	1.02	0.89
TC-3	Standard	0.08	0.03	0.97	1.06	0.83
	Optimized	0.11	0.01	1.02	1.02	1.00
Concentration						
		FB	NMSE	MG	VG	FAC2
TC-1	Standard	0.11	0.68	1.08	2.11	0.64
	Optimized	0.18	0.58	1.10	1.93	0.69
TC-2	Standard	0.12	0.47	1.10	1.52	0.70
	Optimized	0.19	0.45	1.17	1.41	0.76
TC-3	Standard	0.10	0.59	1.12	2.03	0.67
	Optimized	0.16	0.50	1.14	1.84	0.71

The analysis of the presented data demonstrates the validity of the predicted flow fields in all test cases. However, consistent with our earlier statements, the reproduced velocity fields demonstrate minimal sensitivity to the modification of the closure coefficients. The same argument does not hold for pollutant concentration and TKE fields. Concerning TKE, the modified closure achieves noticeable enhancements in the accuracy and quality of predictions. In line with the output responses of the training case, a relatively more intense under-prediction of TKE by the optimized closure is observed, which has resulted in relatively higher quantities of FB and MG. The slightly higher levels of FB and MG, however, have modified the predictions in a manner that has led to smaller overall scatters (lower VG and NMSE) and raised HR in all three test cases. Having a more accurate and reliable representation of the flow and turbulence parameters, an improvement in the predicted concentration field is consequently expected.

Any changes to solutions for Equations (6) and (8) (i.e., using a revised coefficient set) modify the estimated eddy viscosity (ν_t) and, accordingly, the turbulence mass diffusion (D_t) in Equation (10). Additionally, D_t is also affected by the definition of Sc_t in this work (i.e., Equation (13)), which shows an explicit dependency on the local characteristics of turbulence. These contributions result in strong sensitivity of the predicted concentration field to different editions of closure constants, which is established by the presented data in Table 7. The evaluation of the validation metrics together exhibits an overall boost in the quality and reliability of predictions for the three test cases. Therefore, the potential applicability of the trained closure set for time-averaged modeling of atmospheric dispersion flow in compact urban settings is attested.

7. Conclusions

In an effort to improve the accuracy of the steady atmospheric dispersion modeling in the context of a compact urban setting, a re-calibration study is carried out on the empirical constants of the standard $k - \epsilon$ model. To this aim, in addition to the flow parameters (velocity and TKE), the pollutant concentration field is also selected as the parameter of focus, implemented through the definition of five different validation metrics (three linear and two logarithmic). An optimization scheme based on the GA algorithm is adopted here with some modifications to systematically select sets of coefficients from predetermined ranges to evaluate and eventually identify variation spans with the highest associated model's accuracy. The comprehensive and high-quality dispersion dataset of full-scale field measurements in an urban-like geometry (i.e., the MUST experiment) was used to re-calibrate the model constants. In order to evaluate the performance of the optimized closure, the quality of predictions for concentration, velocity, and TKE fields are further compared with the ones of the standard model. Finally, the general applicability of the modified set to other cases is examined by modeling three distinctive test cases with different inflow velocities, source locations, building orientations, and atmospheric states. The main steps and key findings of this study can be summarized as follow:

- Considering the large size of the model, making specific arrangements is essential to alleviate the associated computational costs. In this regard, the established relationship among C_μ , $C_{\epsilon 1}$, $C_{\epsilon 2}$, and σ_ϵ within the atmospheric surface layer is utilized to omit $C_{\epsilon 1}$ from the optimization study, as it can be calculated using the other three constants.
- A screening method was used to quantify each constant's direct and interactional effects on the validation metrics. As suggested by the results, the model's outputs reflect a relatively minimal sensitivity to σ_k , which justifies the decision to exclude this coefficient from the optimization step and find its optimal value later through a simple sensitivity study.
- The rigorous analysis of all validation metrics together, while prioritizing achieving the highest quantities of FAC2, has led to recommending the optimal ranges for the generic case of a compact urban setting as follows: $0.14 \leq C_\mu \leq 0.15$, $1.68 \leq C_{\epsilon 2} \leq 1.80$, $1.12 \leq \sigma_\epsilon \leq 1.20$, and $0.87 \leq \sigma_k \leq 1.00$.
- Given the proposed ranges in this work, a closure set is found to generate predictions that agree best with the selected field measurements of this study, which consists of $C_\mu = 0.147$, $C_\epsilon = 1.344$, $C_{\epsilon 2} = 1.693$, $\sigma_\epsilon = 1.196$, and $\sigma_k = 0.927$.
- An exhausting assessment of the statistical measures resulting from the comparative study indicates that the optimized closure significantly outperforms the other revision in reproducing the concentration and TKE fields, while both editions yield roughly similar results for the velocity field. Relative to predictions by the standard model, the FAC2 for the concentrations (among 74 sampling points) and HR for the TKE field (among 18 sampling points) are increased by 8% and 27%, respectively.
- The investigation of the general applicability of the proposed modifications suggests that except for the predicted velocity field, in which only minor improvements are observed, the closure model successfully enhances the quality and reliability of predictions for concentrations and the TKE field in all three test cases.

Author Contributions: Conceptualization, M.R.K.N., C.F.L. and B.A.F.; methodology, M.R.K.N., C.F.L. and B.A.F.; software, M.R.K.N.; validation, M.R.K.N. and K.R.; formal analysis, M.R.K.N.; investigation, M.R.K.N.; resources, B.A.F.; data curation, M.R.K.N.; writing—original draft preparation, M.R.K.N.; writing—review and editing, M.R.K.N., K.R., C.F.L. and B.A.F.; visualization, M.R.K.N.; supervision, C.F.L. and B.A.F.; project administration, B.A.F.; funding acquisition, B.A.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Natural Science and Engineering Council of Canada (NSERC) with discovery grant number RGPIN-2020-03896, and the Canada First Research Excellence Fund (CFREF) with grant number CFREF-2015-00001, provided by Brian A. Fleck.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors would like to thank the Defense Threat Reduction Agency (DTRA) for providing access to the MUST dispersion dataset.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations/Nomenclature

ABL	Atmospheric boundary layer
CFD	Computational fluid dynamics
FAC2	Fraction of the predictions within a factor of 2 of the observations
FB	Fractional bias
GA	Genetic algorithm
GCI	Grid convergence index
HR	Hit-Rate
LES	Large eddy simulations
MG	Mean geometric bias
MUST	Mock urban setting tests
NMSE	Normalized mean square error
PDF	Probability density function
RANS	Reynolds-averaged Navier–Stokes
RNG	Renormalization group
SST	Shear stress transport
VG	Geometric variance
C	Scalar concentration, ppm
C_p	Predicted concentration
C_o	Observed concentration
C^*	Normalized scalar concentration
$C_{\epsilon 1}, C_{\epsilon 2}, C_{\mu}$	Constants in the $k - \epsilon$ model
D	Molecular diffusivity, m^2s^{-1}
D_t	Eddy diffusivity, m^2s^{-1}
E_i	Elementary effect of model coefficients
\bar{E}_i	Mean elementary effect
H_{ref}	Reference height, m
k	Turbulent kinetic energy (TKE), m^2s^{-2}
k_s	Sand grain roughness, m
k_s^+	Dimensionless sand grain roughness
M_i	Measured value
O_{fi}	Objective function ith
P_i	Predicted value
P_k	TKE production rate, m^2s^{-3}
q_s	Scalar volumetric flowrate, m^3s^{-1}
Re_t	Turbulent Reynolds number
Ro	Rossby number
S	Strain rate invariant
S_{ij}	Strain rate tensor, s^{-1}
Sc	Schmidt number
Sc_t	Turbulent Schmidt Number
\bar{u}_i	Component of mean velocity, ms^{-1}
u'_i	Component of fluctuating velocity, ms^{-1}
u_τ	Friction velocity, ms^{-1}
U_p	Velocity at the center of the first cell next to the wall, ms^{-1}
U_{ref}	Reference velocity, ms^{-1}
x_i	component of space coordinate, m
z_0	Aerodynamic roughness, m
δ_{ij}	Kronecker delta
ϵ	Turbulent dissipation rate, m^2s^{-3}

κ	Von Karman constant
μ	Dynamic molecular viscosity, $\text{kgm}^{-1}\text{s}^{-1}$
ν	Kinematic molecular viscosity, m^2s^{-1}
ν_t	Eddy viscosity, m^2s^{-1}
ρ	Density, kgm^{-3}
$\sigma_\varepsilon, \sigma_k$	Constants in the $k - \varepsilon$ model
σ_{Ei}	Standard deviation of the elementary effect
ω	Specific dissipation rate, s^{-1}
Ω	Vorticity rate invariant

References

1. 2018 Revision of World Urbanization Prospects; Technical Report; United Nations Department of Economics and Social Affairs (UNDESA): New York, NY, USA, 2018.
2. Yang, Y.; Guangrong, S.; Chen, Z.; Hao, S.; Zhouyiling, Z.; Shan, Y. Quantitative analysis and prediction of urban heat island intensity on urban-rural gradient: A case study of Shanghai. *Sci. Total Environ.* **2022**, *829*, 154264. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Li, Z.; Ming, T.; Liu, S.; Peng, C.; de Richter, R.; Li, W.; Zhang, H.; Wen, C.Y. Review on pollutant dispersion in urban areas-part A: Effects of mechanical factors and urban morphology. *Build. Environ.* **2021**, *190*, 107534. [\[CrossRef\]](#)
4. Lu, C.; Cao, L.; Norbäck, D.; Li, Y.; Chen, J.; Deng, Q. Combined effects of traffic air pollution and home environmental factors on preterm birth in China. *Ecotoxicol. Environ. Saf.* **2019**, *184*, 109639. [\[CrossRef\]](#)
5. Jin, X.; Yang, L.; Du, X.; Yang, Y. Sensitivity analyses of ultrafine particle dispersion inside an isolated street canyon. *Powder Technol.* **2016**, *304*, 143–156. [\[CrossRef\]](#)
6. Xia, Y.; Guan, D.; Jiang, X.; Peng, L.; Schroeder, H.; Zhang, Q. Assessment of socioeconomic costs to China's air pollution. *Atmos. Environ.* **2016**, *139*, 147–156. [\[CrossRef\]](#)
7. Li, Z.; Ming, T.; Shi, T.; Zhang, H.; Wen, C.Y.; Lu, X.; Dong, X.; Wu, Y.; de Richter, R.; Li, W.; et al. Review on pollutant dispersion in urban areas-part B: Local mitigation strategies, optimization framework, and evaluation theory. *Build. Environ.* **2021**, *198*, 107890. [\[CrossRef\]](#)
8. Hu, Y.; Xu, F.; Gao, Z. A Comparative Study of the Simulation Accuracy and Efficiency for the Urban Wind Environment Based on CFD Plug-Ins Integrated into Architectural Design Platforms. *Buildings* **2022**, *12*, 1487. [\[CrossRef\]](#)
9. Blocken, B.; Stathopoulos, T. CFD simulation of pedestrian-level wind conditions around buildings: Past achievements and prospects. *J. Wind. Eng. Ind. Aerodyn.* **2013**, *121*, 138–145. [\[CrossRef\]](#)
10. Gough, H.L.; Luo, Z.; Halios, C.H.; King, M.F.; Noakes, C.J.; Grimmond, C.S.; Barlow, J.F.; Hoxey, R.; Quinn, A.D. Field measurement of natural ventilation rate in an idealised full-scale building located in a staggered urban array: Comparison between tracer gas and pressure-based methods. *Build. Environ.* **2018**, *137*, 246–256. [\[CrossRef\]](#)
11. Mattar, S.J.; Nezhad, M.R.K.; Versteeg, M.; Lange, C.F.; Fleck, B.A. Validation Process for Rooftop Wind Regime CFD Model in Complex Urban Environment Using an Experimental Measurement Campaign. *Energies* **2021**, *14*, 2497. [\[CrossRef\]](#)
12. Li, B.; Li, C.; Yang, Q.; Tian, Y.; Zhang, X. Full-scale wind speed spectra of 5Year time series in urban boundary layer observed on a 325m meteorological tower. *J. Wind. Eng. Ind. Aerodyn.* **2021**, *218*, 104791. [\[CrossRef\]](#)
13. He, Y.; Liu, Z.; Ng, E. Parametrization of irregularity of urban morphologies for designing better pedestrian wind environment in high-density cities—A wind tunnel study. *Build. Environ.* **2022**, *226*, 109692. [\[CrossRef\]](#)
14. Hang, J.; Chen, G. Experimental study of urban microclimate on scaled street canyons with various aspect ratios. *Urban Clim.* **2022**, *46*, 101299. [\[CrossRef\]](#)
15. Lateb, M.; Meroney, R.N.; Yataghene, M.; Fellouah, H.; Saleh, F.; Boufadel, M.C. On the use of numerical modelling for near-field pollutant dispersion in urban environments: A review. *Environ. Pollut.* **2016**, *208*, 271–283. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Hajra, B.; Stathopoulos, T. A wind tunnel study of the effect of downstream buildings on near-field pollutant dispersion. *Build. Environ.* **2012**, *52*, 19–31. [\[CrossRef\]](#)
17. Blocken, B. 50 years of Computational Wind Engineering: Past, present and future. *J. Wind. Eng. Ind. Aerodyn.* **2014**, *129*, 69–102. [\[CrossRef\]](#)
18. Hassan, A.M.; ELMokadem, A.A.; Megahed, N.A.; Abo Eleinen, O.M. Urban morphology as a passive strategy in promoting outdoor air quality. *J. Build. Eng.* **2020**, *29*, 101204. [\[CrossRef\]](#)
19. Blocken, B. LES over RANS in building simulation for outdoor and indoor applications: A foregone conclusion? *Build. Simul.* **2018**, *11*, 821–870. [\[CrossRef\]](#)
20. Guo, D.; Zhao, P.; Wang, R.; Yao, R.; Hu, J. Numerical simulations of the flow field and pollutant dispersion in an idealized urban area under different atmospheric stability conditions. *Process Saf. Environ. Prot.* **2020**, *136*, 310–323. [\[CrossRef\]](#)
21. Reiminger, N.; Vazquez, J.; Blond, N.; Dufresne, M.; Wertel, J. CFD evaluation of mean pollutant concentration variations in step-down street canyons. *J. Wind. Eng. Ind. Aerodyn.* **2020**, *196*, 104032. [\[CrossRef\]](#)
22. Nezhad, M.R.K.; Lange, C.F.; Fleck, B.A. Evaluating the Validity of Computational Fluid Dynamics Model of Dispersion in a Complex Urban Geometry Using Two Sets of Experimental Measurements. *Int. J. Mech. Mechatronics Eng.* **2023**, *17*, 78–86.
23. Keshavarzian, E.; Jin, R.; Dong, K.; Kwok, K.C. Effect of building cross-section shape on air pollutant dispersion around buildings. *Build. Environ.* **2021**, *197*, 107861. [\[CrossRef\]](#)

24. Versteeg, H.K.; Malalasekera, W. *An Introduction to Computational Fluid Dynamics*, 2nd ed.; Pearson Education: Harlow, UK, 2007.
25. Tominaga, Y. Flow around a high-rise building using steady and unsteady RANS CFD: Effect of large-scale fluctuations on the velocity statistics. *J. Wind. Eng. Ind. Aerodyn.* **2015**, *142*, 93–103. [\[CrossRef\]](#)
26. Glover, N.; Guillas, S.; Malki-Epshtein, L. Statistical calibration of CFD modelling for street canyon flows. In Proceedings of Building Simulation 2011: 12th Conference of International Building Performance Simulation Association, Sydney, Australia, 14–16 November 2011; pp. 1513–1520.
27. Launder, B.E.; Spalding, D.B. The numerical computation of turbulent flows. *Comput. Methods Appl. Mech. Eng.* **1974**, *3*, 269–289. [\[CrossRef\]](#)
28. Shih, T.H.; Liou, W.W.; Shabbir, A.; Yang, Z.; Zhu, J. A new k- ϵ eddy viscosity model for high reynolds number turbulent flows. *Comput. Fluids* **1995**, *24*, 227–238. [\[CrossRef\]](#)
29. Yakhot, V.; Orszag, S.A. Renormalization-group analysis of turbulence. *Phys. Rev. Lett.* **1986**, *57*, 1722–1724. [\[CrossRef\]](#) [\[PubMed\]](#)
30. Menter, F.R. Two-equation eddy-viscosity turbulence models for engineering applications. *AIAA J.* **1994**, *32*, 1598–1605. [\[CrossRef\]](#)
31. Pantusheva, M.; Mitkov, R.; Hristov, P.O.; Petrova-Antonova, D. Air Pollution Dispersion Modelling in Urban Environment Using CFD: A Systematic Review. *Atmosphere* **2022**, *13*, 1640. [\[CrossRef\]](#)
32. Lateb, M.; Masson, C.; Stathopoulos, T.; Bédard, C. Comparison of various types of k- ϵ models for pollutant emissions around a two-building configuration. *J. Wind. Eng. Ind. Aerodyn.* **2013**, *115*, 9–21. [\[CrossRef\]](#)
33. Hosseinzadeh, A.; Keshmiri, A. Computational Simulation of Wind Microclimate in Complex Urban Models and Mitigation Using Trees. *Buildings* **2021**, *11*, 112. [\[CrossRef\]](#)
34. Narjisse, A.; Abdellatif, K. Assessment of RANS turbulence closure models for predicting airflow in neutral ABL over hilly terrain. *Int. Rev. Appl. Sci. Eng.* **2021**, *12*, 238–256. [\[CrossRef\]](#)
35. Kavian Nezhad, M.R.; Lange, C.F.; Fleck, B.A. Performance Evaluation of the RANS Models in Predicting the Pollutant Concentration Field within a Compact Urban Setting: Effects of the Source Location and Turbulent Schmidt Number. *Atmosphere* **2022**, *13*, 1013. [\[CrossRef\]](#)
36. Detering, H.W.; Etling, D. Application of the E- ϵ turbulence model to the atmospheric boundary layer. *Bound. Layer Meteorol.* **1985**, *33*, 113–133. [\[CrossRef\]](#)
37. Bechmann, A.; Sørensen, N.N.; Bechmann, C.A.; Dtu, R. Hybrid RANS/LES method for wind flow over complex terrain. *Wind. Energy* **2010**, *13*, 36–50. [\[CrossRef\]](#)
38. Richards, P.J.; Hoxey, R.P. Appropriate boundary conditions for computational wind engineering models using the k- ϵ turbulence model. *J. Wind. Eng. Ind. Aerodyn.* **1993**, *46–47*, 145–153. [\[CrossRef\]](#)
39. Guillas, S.; Glover, N.; Malki-Epshtein, L. Bayesian calibration of the constants of the k- ϵ turbulence model for a CFD model of street canyon flow. *Comput. Methods Appl. Mech. Eng.* **2014**, *279*, 536–553. [\[CrossRef\]](#)
40. Zahid Iqbal, Q.M.; Chan, A. Pedestrian level wind environment assessment around group of high-rise cross-shaped buildings: Effect of building shape, separation and orientation. *Build. Environ.* **2016**, *101*, 45–63. [\[CrossRef\]](#)
41. Edeling, W.N.; Cinnella, P.; Dwight, R.P.; Bijl, H. Bayesian estimates of parameter variability in the k- ϵ turbulence model. *J. Comput. Phys.* **2014**, *258*, 73–94. [\[CrossRef\]](#)
42. Shirzadi, M.; Mirzaei, P.A.; Naghashzadegan, M. Improvement of k-epsilon turbulence model for CFD simulation of atmospheric boundary layer around a high-rise building using stochastic optimization and Monte Carlo Sampling technique. *J. Wind. Eng. Ind. Aerodyn.* **2017**, *171*, 366–379. [\[CrossRef\]](#)
43. Shirzadi, M.; Naghashzadegan, M.; Mirzaei, P.A. Improving the CFD modelling of cross-ventilation in highly-packed urban areas. *Sustain. Cities Soc.* **2018**, *37*, 451–465. [\[CrossRef\]](#)
44. Shirzadi, M.; Mirzaei, P.A.; Tominaga, Y. RANS model calibration using stochastic optimization for accuracy improvement of urban airflow CFD modeling. *J. Build. Eng.* **2020**, *32*, 101756. [\[CrossRef\]](#)
45. Biltoft, C. *Customer Report for Mock Urban Setting Test*; U.S. Defense Threat Reduction Agency: Fort Belvoir, VA, USA, 2001.
46. Emeis, S. *Wind Energy Meteorology*, 2nd ed.; Green Energy and Technology; Springer: Zug, Switzerland, 2018; pp. 31–56. [\[CrossRef\]](#)
47. Speranza, A.; Lucarini, V. Environmental Science, Physical Principles and Applications. In *Encyclopedia of Condensed Matter Physics*; Bassani, F., Liedl, G.L., Wyder, P., Eds.; Elsevier: Oxford, UK, 2005; pp. 146–156. [\[CrossRef\]](#)
48. Blocken, B.; Tominaga, Y.; Stathopoulos, T. CFD simulation of micro-scale pollutant dispersion in the built environment. *Build. Environ.* **2013**, *64*, 225–230. [\[CrossRef\]](#)
49. Wilcox, D.C. *Turbulence Modeling for CFD*, 3rd ed.; DCW Industries: La Canada, CA, USA, 2006.
50. Warhaft, Z. Passive Scalars in Turbulent Flows. *Annu. Rev.* **2003**, *32*, 203–240. [\[CrossRef\]](#)
51. Rossi, R.; Iaccarino, G. Numerical simulation of scalar dispersion downstream of a square obstacle using gradient-transport type models. *Atmos. Environ.* **2009**, *43*, 2518–2531. [\[CrossRef\]](#)
52. Tominaga, Y.; Stathopoulos, T. Turbulent Schmidt numbers for CFD analysis with various types of flowfield. *Atmos. Environ.* **2007**, *41*, 8091–8099. [\[CrossRef\]](#)
53. Koeltzsch, K. The height dependence of the turbulent Schmidt number within the boundary layer. *Atmos. Environ.* **2000**, *34*, 1147–1151. [\[CrossRef\]](#)

54. Longo, R.; Fürst, M.; Bellemans, A.; Ferrarotti, M.; Derudi, M.; Parente, A. CFD dispersion study based on a variable Schmidt formulation for flows around different configurations of ground-mounted buildings. *Build. Environ.* **2019**, *154*, 336–347. [\[CrossRef\]](#)
55. Longo, R.; Bellemans, A.; Derudi, M.; Parente, A. A multi-fidelity framework for the estimation of the turbulent Schmidt number in the simulation of atmospheric dispersion. *Build. Environ.* **2020**, *185*, 107066. [\[CrossRef\]](#)
56. Shirzadi, M.; Mirzaei, P.A.; Naghashzadegan, M.; Tominaga, Y. Modelling enhancement of cross-ventilation in sheltered buildings using stochastic optimization. *Int. J. Heat Mass Transf.* **2018**, *118*, 758–772. [\[CrossRef\]](#)
57. Chang, J.C.; Hanna, S.R. Air quality model performance evaluation. *MEteorology Atmos. Phys.* **2004**, *87*, 167–196. [\[CrossRef\]](#)
58. Hanna, S.R.; Hansen, O.R.; Dharmavaram, S. FLACS CFD air quality model performance evaluation with Kit Fox, MUST, Prairie Grass, and EMU observations. *Atmos. Environ.* **2004**, *38*, 4675–4687. [\[CrossRef\]](#)
59. Wang, Q.; Wang, J.; Hou, Y.; Yuan, R.; Luo, K.; Fan, J. Micrositing of roof mounting wind turbine in urban environment: CFD simulations and lidar measurements. *Renew. Energy* **2018**, *115*, 1118–1133. [\[CrossRef\]](#)
60. Campolongo, F.; Braddock, R. Sensitivity analysis of the IMAGE greenhouse model. *Environ. Model. Softw.* **1999**, *14*, 275–282. [\[CrossRef\]](#)
61. Campolongo, F.; Braddock, R. The use of graph theory in the sensitivity analysis of the model output: A second order screening method. *Reliab. Eng. Syst. Saf.* **1999**, *64*, 1–12. [\[CrossRef\]](#)
62. Koch, P.N.; Yang, R.J.; Gu, L. Design for six sigma through robust optimization. *Struct. Multidiscip. Optim.* **2004**, *26*, 235–248. [\[CrossRef\]](#)
63. Goldberg, D.E. *Genetic Algorithms in Search Optimization and Machine Learning*; Addison-Wesley: New York, NY, USA, 1988.
64. Hammersley, J.; Handscomb, D. *Monte Carlo Methods*; Chapman and Hall: London, UK, 1964.
65. Gimenez, J.M.; Bre, F. Optimization of RANS turbulence models using genetic algorithms to improve the prediction of wind pressure coefficients on low-rise buildings. *J. Wind. Eng. Ind. Aerodyn.* **2019**, *193*, 103978. [\[CrossRef\]](#)
66. Yee, E.; Bilitoft, C.A. Concentration Fluctuation Measurements in a Plume Dispersing Through a Regular Array of Obstacles. *Bound.-Layer Meteorol.* **2004**, *111*, 363–415. [\[CrossRef\]](#)
67. Tominaga, Y.; Mochida, A.; Yoshie, R.; Kataoka, H.; Nozu, T.; Yoshikawa, M.; Shirasawa, T. AIJ guidelines for practical applications of CFD to pedestrian wind environment around buildings. *J. Wind. Eng. Ind. Aerodyn.* **2008**, *96*, 1749–1761. [\[CrossRef\]](#)
68. Franke, J.; Hellsten, A.; Schlünzen, K.; Carissimo, B. Best practice guideline for the CFD simulation of flows in the urban environment—a summary. In Proceedings of the 11th Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes, Cambridge, UK, 2–5 July 2007.
69. Blocken, B.; Stathopoulos, T.; Carmeliet, J. CFD simulation of the atmospheric boundary layer: Wall function problems. *Atmos. Environ.* **2007**, *41*, 238–252. [\[CrossRef\]](#)
70. Blocken, B. Computational Fluid Dynamics for urban physics: Importance, scales, possibilities, limitations and ten tips and tricks towards accurate and reliable simulations. *Build. Environ.* **2015**, *91*, 219–245. [\[CrossRef\]](#)
71. Celik, I.B.; Ghia, U.; Roache, P.J.; Freitas, C.J.; Coleman, H.; Raad, P.E. Procedure for estimation and reporting of uncertainty due to discretization in CFD applications. *J. Fluids Eng. Trans. ASME* **2008**, *130*, 0780011–0780014. [\[CrossRef\]](#)
72. Roache, P.J. Quantification of uncertainty in computational fluid dynamics. *Annu. Rev. Fluid Mech.* **1997**, *29*, 123–160. [\[CrossRef\]](#)
73. RahnamayBahambary, K.; Fleck, B.A. Effects of Inflow Parameters and Disk Thickness on an Actuator Disk inside the Neutral Atmospheric Boundary Layer. *Wind* **2022**, *2*, 733–746. [\[CrossRef\]](#)
74. Richards, P.J.; Norris, S.E. Appropriate boundary conditions for computational wind engineering models revisited. *J. Wind. Eng. Ind. Aerodyn.* **2011**, *99*, 257–266. [\[CrossRef\]](#)
75. Toja-Silva, F.; Peralta, C.; Lopez-Garcia, O.; Navarro, J.; Cruz, I. Roof region dependent wind potential assessment with different RANS turbulence models. *J. Wind. Eng. Ind. Aerodyn.* **2015**, *142*, 258–271. [\[CrossRef\]](#)
76. Tominaga, Y.; Stathopoulos, T. CFD simulation of near-field pollutant dispersion in the urban environment: A review of current modeling techniques. *Atmos. Environ.* **2013**, *79*, 716–730. [\[CrossRef\]](#)
77. Ferziger, J.H.; Perić, M.; Street, R.L. *Computational Methods for Fluid Flow*, 4th ed.; Springer: Zug, Switzerland, 2020. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.