SAND2019-10539C

An Urban Dispersion Inspired Scenario for CFD Model Validation

Alexander L. Brown[a*], Michael D. Clemenson[b], Michael Benson[c], Christopher Elkins[d], and Samuel T. Jones[e]

[a]Sandia National Labs, PO Box 5800, Albuquerque, NM, USA, albrown@sandia.gov

[b]Sandia National Labs, PO Box 5800, Albuquerque, NM, USA, mdcleme@sandia.gov

[c]United States Military Academy, 752 Thayer Rd., West Point, NY, USA, michael.benson@westpont.edu

[d]Stanford University, 488 Escondido Mall, Stanford, CA, USA, celkins@stanford.edu

[e]American International College of Arts and Sciences, Jabberwock Road, Coolidge, Antigua, W.I, sjones@aicasa.net

*Corresponding author

Abstract:

Momentum, advection, diffusion, and turbulence are component physics relating to fire simulation tools like computational fluid dynamics (CFD). Magnetic Resonance Velocimetry and Magnetic Resonance Concentration MRV/MRC techniques can produce heretofore unrivaled detailed measurements of three-component velocity and concentration fields in turbulent flows. This study exhibits 3D flow comparisons between velocity and concentration fields obtained using MRC/MRV and SIERRA/Fuego for an urban geometry based on a section of downtown Oklahoma City. A 1:2500 scale water flow scenario provides 0.8 mm resolution data. Various techniques are employed to quantify the accuracy of the simulation results. The techniques all generally suggest a good comparison between the model and experiments throughout the compared volume. The selected metrics provide benchmark accuracy measures that can be used to indicate quantitative accuracy of the simulations, as well as for targets for future simulation improvements.

Keywords: CFD; fluid dynamics; validation

1. Introduction

Validation is a credibility process that helps quantify the accuracy of simulation tools for use in predicting physical systems [1]. It is normally a persistent process that continually needs revisiting to update with code modifications. For complex simulation tools like computational fluid dynamics (CFD) it is a continual process that benefits from additional high-relevancy data as new experimental test regimes are devised. Fire CFD simulation tools are complex, and validation is a regularly evaluated challenge. Radiation, soot, chemical reactions, flow, heat transport, diffusion, and turbulence are examples of many complicating features that are frequently modeled with these tools. While validation can involve full-physics comparisons between data and models, it is often more instructive (and certainly easier) to identify potential

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

issues by performing component validation studies. These isolate selected code features in simpler datasets that can be used to build confidence in the accuracy of simulation tools. The fire community is taking this approach, endeavoring to improve the credibility of model validation by engaging in a community effort known as Measurement and Computation of Fire Phenomena MaCFP [2]. The MaCFP community has been using a 2D buoyant mixing helium plume dataset for evaluating turbulent mixing and buoyant behavior at scales relevant to fire conditions [3].

Magnetic Resonance Velocimetry (MRV) and Magnetic Resonance Concentration (MRC) are methods that employ existing magnetic resonance imaging (MRI) systems outside their normal healthcare applications to interrogate engineering problems of interest [4-9]. Specific scan sequences can obtain the three velocity components or concentrations and are independently performed over measurements that occur during a period of several hours. The data are averaged and for the MRC technique, stitched together during processing to minimize experimental uncertainty. Since the systems are commonly tuned to measure resonance patterns of hydrogen protons, the best data are achieved for water flows. From a fire and modeling perspective, water is not of particular interest. It is more dense and viscous than air and is incompressible. These issues notwithstanding, the technique has applicability to model validation as applied to relevant scenarios. The simulation tools solve the same transport equations regardless of the fluid type, and self-similarity is a common and proven fluid mechanics concept that permits qualified application of fluid data between fluids of differing types by matching key non-dimensional parameters, such as Reynolds numbers. Liquid water experiments have been used previously for validating fire simulation tools (e.g. [10]).

This team has collaborated previously on two datasets with similar conditions. Notional scale geometries were constructed to simulate urban contaminant transport [11-13]. This work differs from many of the prior comparisons by focusing on the full 3D dataset simultaneously rather than focusing the point, line, or planar data extractions. This work also focuses on a scenario motivated by a plume release dataset that was part of a test campaign to study the release of contaminants in urban settings [14]. A report of the MRC/MRV dataset is presently in review for publication [15]. A portion of downtown Oklahoma City is modeled with a contaminant injection just south of the intersection of W. Main St. and Broadway Ave. Flows are fully turbulent, with calculated Reynolds numbers based on the channel cross-section greater than 10^4. While in a turbulent regime, the Reynolds number is smaller than would be desired for a characteristic self-similar approximation to a more probable 1-2 m/s wind condition. Nonetheless, the scenario provides a rich platform for performing a uniquely detailed assessment of CFD accuracy in the turbulent regime.

This work seeks to apply the data from the MRC/MRV experiments towards a validation study for a fire simulation tool. A prime objective is to exhibit and quantify the accuracy of the simulation tool. An equally important objective is to develop and assess comparison methods suited to the novel validation problem relating to the unusually detailed data and model.

2. Material and Methods

2.1 Geometry

A CAD representation of downtown Oklahoma City consistent with the datasets [14] was obtained and scaled to 1:2500. Additive manufacturing using Accura 60 resin was used to generate a model for insertion in the flow channel illustrated in Fig. 1. The flow section of the channel was 196 mm wide and 110 mm high. Additional details of the channel and the flow development section may be found in the data report [15].

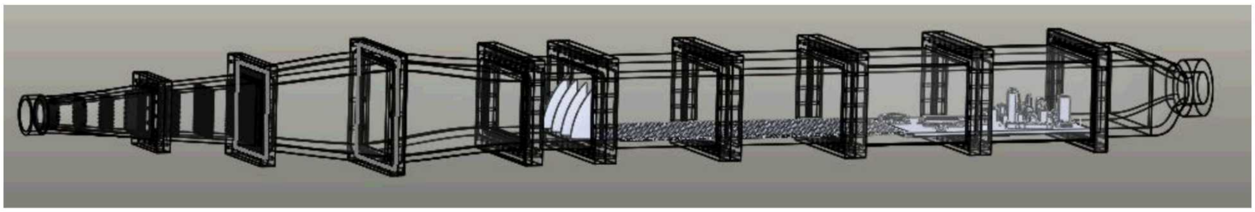

Fig. 1. A graphical rendering of the flow apparatus.

Just south of the intersection of W. Main and N. Broadway, an injector (Fig. 2) with a 15-degree angle relative to the street was placed to provide a constant rate injection of a $CuSO_4$ doped water stream that was nearly neutrally buoyant and provided increased signal to noise ratio for concentration measurements. The up-wind boundary condition was measured by MRV and applied by interpolation to the modeling. Remaining simulation boundary conditions were assumed to be smooth walls or open boundaries (at the exit of the domain).

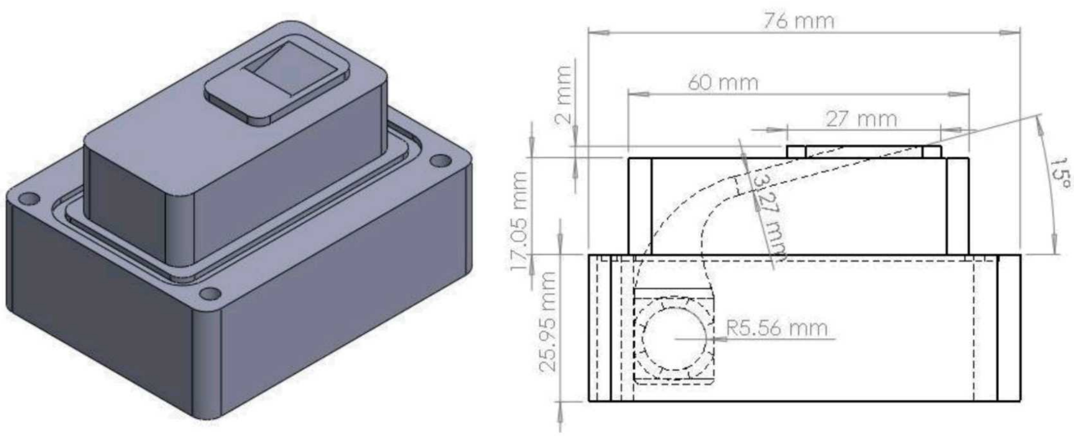

Fig. 2. The injector used for the tests.

A 3.0T GE Discovery MR750 research MRI system at the Richard M. Lucas Center at Stanford University was used for testing. An undergraduate student team from the U.S. Military Academy helped construct the geometries and prepare the 3D printed test objects. The same students traveled to California to perform the tests and conducted initial analyses. Datasets were post-processed for scale accuracy and stitched together for a comprehensive dataset for the scenario. This resulted in the 3-components of velocity and the concentration data in a data matrix of approximately 244x127x309, or approximately 9.5 million measurement locations (voxels).

2.2 Alignment and Filtering

Prior work [16] has identified the mutual alignment of the simulation and experimental datasets as a parameter of uncertainty. Achieving alignment to within an isotropic voxel length of about 0.8 mm is difficult. In this case, geometric features of the skyline structures provide good guidelines for producing an accurate comparison. Further, as suggested by prior work [16], the gradients of the measured variables (vorticity and gradient of concentration) provide a higher sensitivity measure for assessing and aiding in achieving good alignment. They diverge much quicker than the concentration scalar or velocity vector. A manual overlay technique suggested the default dataset be moved 1 mm in the z-direction and scaled 0.99 in the x-direction to attain the alignment of the data as presented in this report. This provided high confidence in the x- and z-direction alignment. Y-direction alignment was more difficult to verify and may have larger uncertainty in position than the other coordinates.

Filtering of the data is necessary for several reasons. First, it enables analysis with existing common tools. Performing comparative calculations across a dataset with millions of points starts to get into the application space of parallel algorithms, and this work has not progressed to that level of maturity. Second, there are reasons to clip the data for the sake of accuracy. The data for concentration are only reliable to the detection threshold, and interpolated velocities are more uncertain at surfaces such as channel walls and buildings where partial fluid-volume effects are increased. While the techniques employ strategies to detect and remove these potential sources of error, this work also employs the use of several data clips that improves the validity of the field data comparisons as defined by:

- Experimental concentration <= 0.02
- Experimental velocity magnitude < 1.0 mm/s (2.5% of mean)
- Boundary cells omitted (using the y+ variable from the model)
- y < 0.0 (omits the injector where the experimental data were questionable)
- One building had a courtyard or otherwise similar cavity where the test team identified a build-up of concentration. This was omitted using a spatial clipping algorithm (0.00315>x>-0.00785, y<0.00719, -0.0534<z<-0.06044).

Gradient variables are calculated using Kitware/paraview before clipping occurs. The boundary clips are expected to reduce or eliminate artifacts of the surface from the gradient data, while retaining relevant domain data.

2.3 Computational

The SIERRA suite of computational tools is composed of engineering science analysis software designed to support DOE mission applications. SIERRA is a common architecture designed to take advantage of constantly improving computational platforms for high-fidelity simulation. One of the fluid mechanics modules available is the low-Mach number reacting flow code known as Fuego [17]. A suite of simulations was performed, but here we focus on the nominal intermediate mesh results and the KSGS [18] LES turbulence model scenario. The fluid is simulated as a constant density fluid with a mixture fraction for the concentration variable. Simulations were run for 5 seconds, after which the concentration and velocity were averaged for 15 seconds. Prior work on similarly scaled scenarios suggests the average over that time is an adequate representation of the mean flow. Fig. 3 illustrates the coarse mesh and illustrates some light biasing applied to the top and down-stream regions of the flow. Table 1 gives some information on the meshes for the study, with the baseline ‘medium’ results being the focus of this paper.

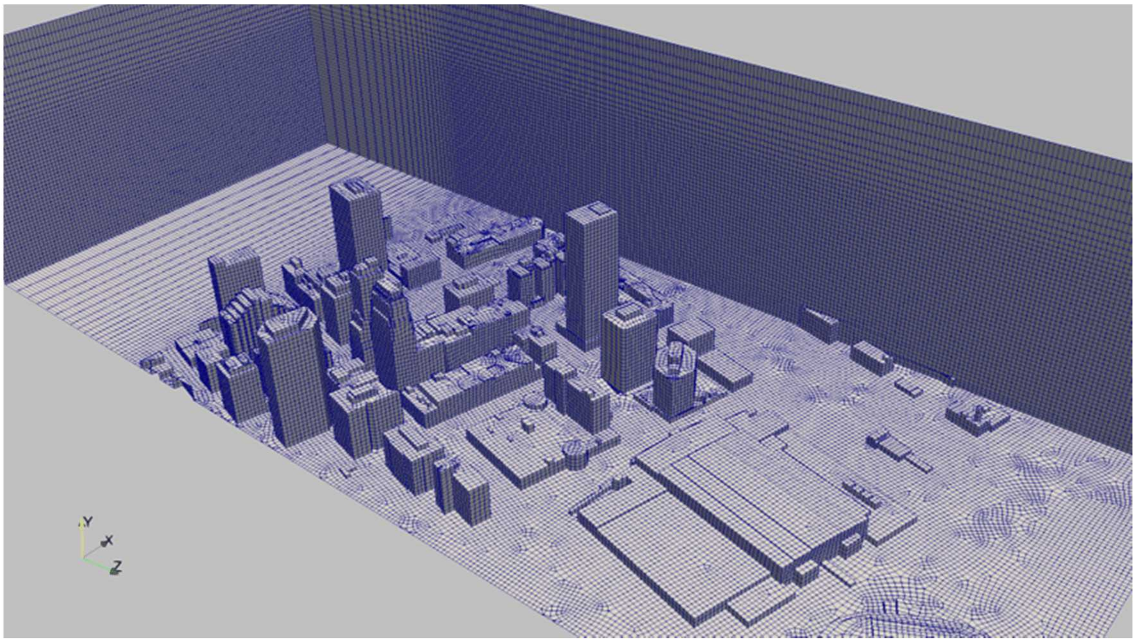

Fig. 3. An illustration of the coarse mesh

Table 1. Meshes used for the broader simulation study

Mesh	Smallest	Most Likely	Largest
Coarse	0.35 mm	1.7 mm	4.4 mm
Medium	0.175 mm	0.85 mm	2.2 mm
fine	0.12 mm	0.57 mm	1.47 mm

Grid convergence and a parameter study was conducted at the medium mesh resolution to propagate model uncertainties and provide a better understanding of the accuracy of the model closely following the Brown and Benavidez methods [11-12]. This presentation of the data omits more than a mention of this study, but this additional quantitative information is expected to become available in other forthcoming documentation.

2.4 Comparisons

We interpolate experimental results from a 0.8 mm data mesh onto the computational medium mesh with comparable resolution. We have found previously that results are generally insensitive to the selection of source and destination mesh. This comparison work was initially guided by the approach Hanna and Chang (2012) [19] proposed for comparing campaign-level test results to simulations (heretofore referred to as HC-2012). They proposed the following metrics for assessing performance of computational models:

- The Fractional Mean Bias (FB): $2\left(\overline{C_o} - \overline{C_p}\right)/\left(\overline{C_o} + \overline{C_p}\right)$

- The Normalized Mean-Square Error (NMSE): $\overline{\left((C_o - C_p)^2\right)}/(\overline{C_o} \times \overline{C_p})$
- The Geometric Mean (MG): $exp(\overline{lnC_o}) - exp(\overline{lnC_p})$
- The Geometric Variance (VG): $exp\left(\overline{(lnC_o - lnC_p)^2}\right)$
- The Fraction of Predictions with in a Factor of 2 of the Observations (FAC2): $0.5<(C_p/C_o)<2$
- The Normalized Absolute Difference (NAD): $|\overline{C_o - C_p}|/(\overline{C_o} + \overline{C_p})$

They propose an urban acceptance criteria of |FB| < 0.67; NMSE < 6; FAC2 > 0.3; and NAD < 0.5 for concentration data (c_p being the predicted variable and c_o being the observed). Under the assumption that these statistical parameter comparisons are agnostic to the type of variable, we apply these to the velocity as well.

We introduce in this context three additional measures for quantifying model and experimental differences. Two of these are identified as:

The local normalized difference (LND):

$$LND = \frac{|C_p - C_o|}{max(C_p, C_o)} \tag{1}$$

The local logarithmic ratios (LLR):

$$LLR = ln\left(\frac{max(C_p, C_o)}{min(C_p, C_o)}\right) \tag{2}$$

These two additional measures provide alternate expressions of accuracy. Both quantities approach zero for ideal comparisons. LND cannot exceed 1.0, which represents the worst-case comparison between model and experiment. LLR is not similarly bounded, however larger numbers represent poorer comparisons. These have advantages over many of the HC-2012 metrics in that they do not have terms in the denominator that depend on the entire dataset. The third additional measure is the square of the correlation coefficient. This is assessed on the raw velocity and concentration data, as well as on the relevant accuracy measures listed above and variable gradients including the vorticity and q-criterion. The R^2 term is the square of the correlation coefficient, also known as the coefficient of determination.

Additionally, we explore the mapping of the lowest performing spatial locations on selected error measures back onto the geometry model. This provides information on the locations of the lowest quality predictions, which can be helpful in discovering the reasons for the error and in addressing shortcomings.

3. Results and Discussion

Two data selection technique results are illustrated herein. Because concentration measurements are only accurate to a mass fraction of 0.02, velocity and concentrations are compared to just the data with viable concentration and velocity data. Velocities are clipped to ones greater than 1 mm/s to eliminate boundary node effects, and concentration to 0.02. Concentration clips omit the building courtyard data, while the box clip (a more comprehensive comparison for velocity only comparisons) retains these data.

3.1 Concentration Clipped Comparisons

The concentration clipping technique left more than 275,000 points at which velocity and concentrations could be compared in the domain. Fig. 4 shows a comparison of experimental and simulation velocity magnitude and concentration. The data are filtered showing only 1 in 5 data points. The velocity magnitudes are clearly similar, with some limited outlier points moderately to significantly outside the perfect fit trend. Concentration comparisons appear more scattered, with good linear trends along the 45° best fit in these plots. Concentration data are more sparse approaching the peak concentration of 1.0, with the bulk of the data at much smaller concentrations. The slope of the velocity magnitude fit line constrained to an intercept of 0.0 was 0.983 with an R^2 of 0.824. The simulations were very well scaled to the data (poor scaling would be evident with a slope significantly varying from 1.0). The slope of the concentration fit line constrained to an intercept of 0.0 was 0.924 with an R^2 of 0.549. This fit is not as good, but the slope suggests the concentrations are scaled similarly.

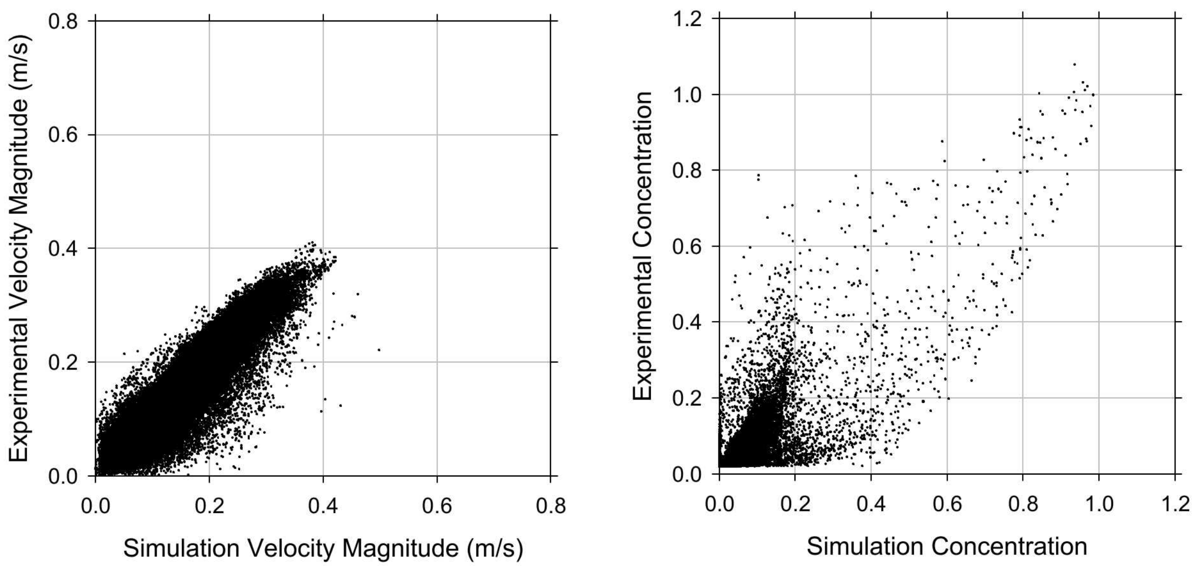

Fig. 4. Scatter plot showing the velocity magnitude and concentration comparisons for the 0.02 concentration clipped data.

Table 2 shows the quantitative performance to the presented accuracy metrics. The LND of 0.197 is the lowest observed yet for comparisons of this nature. The HC-2012 metrics are all passed (accepted) and mostly passed easily with metrics well away from threshold values. The concentration metrics are generally poorer than those found in prior comparison activities with more simplified geometries [16]. Observations of cut-plane results comparing concentrations along various streets suggested that the concentration results were poorest nearest the injector. The reason for this is not obvious but could be related to the truncated flow development in the simulation for the injection.

Table 2. Quantitative metrics for the accuracy of the model versus prediction

Parameter	Concentration	Pass/Fail	Velocity Magnitude	Pass/Fail
FB	-0.1251	Pass	-0.2177	Pass
NMSE	0.5630	Pass	0.09576	Pass
MG	-0.01039		-0.00626	
VG	2.377		6.963	
FAC2	0.9128	Pass	0.9222	Pass
NAD	0.1803	Pass	0.08879	Pass
LND	0.2619		0.1971	
LLR	0.4253		0.2550	

3.2 Box Domain Comparison

While concentration comparisons are better suited for a more limited suite of points which have data within the experimental accuracy window, velocity comparisons are credible throughout a greater section of the simulated domain. A domain was selected that is limited to a region 10.3 x 8.3 x 14.5 cm that encompasses the bulk of the high-concentration prediction area horizontally. The vertical selection was made to sample data into the free-stream flow above the urban canopy. A similar analysis is performed on these data as was performed with the concentration limited selection, with this selection providing over 910,000 individual points of comparison focused on the velocity measurements.

Fig. 5 shows velocity magnitude and the three velocity component comparisons for the full domain. There were nearly 2 million comparison points in the raw comparison. This was reduced by half for data analysis because the analysis software could not handle the size of the full dataset for the full analysis. For plotting, the points shown are reduced to 1 in every 20 of the reduced dataset (for a net reduction of 1:40 from the raw dataset). One can observe in the scatter plots that the velocity magnitude and component experimental and simulation results compare very well, with the bulk trends suggesting a strong linearity and correlation between the two. The raw dataset was inverted compared to the simulations in the x- and z-directions, which had north being negative-z and east being positive-x. The comparisons here are all aligned in the coordinate frame of the model. The slope of the velocity magnitude fit line constrained to an intercept of 0.0 was 1.01 with an R^2 of 0.951. The simulations were very well scaled to the data, and the fit is very good.

Comparison metrics are shown in Table 3. These measures suggest excellent similarity between model and data. These comparisons were made on the half-sized dataset of 910,000 data points. Prior LND and LLR magnitudes were around 0.2-0.3 [16]. This LND and LLR measure is much better, with LND below 0.1 and LLR just above 0.1. These metrics suggest a high similarity between data and model.

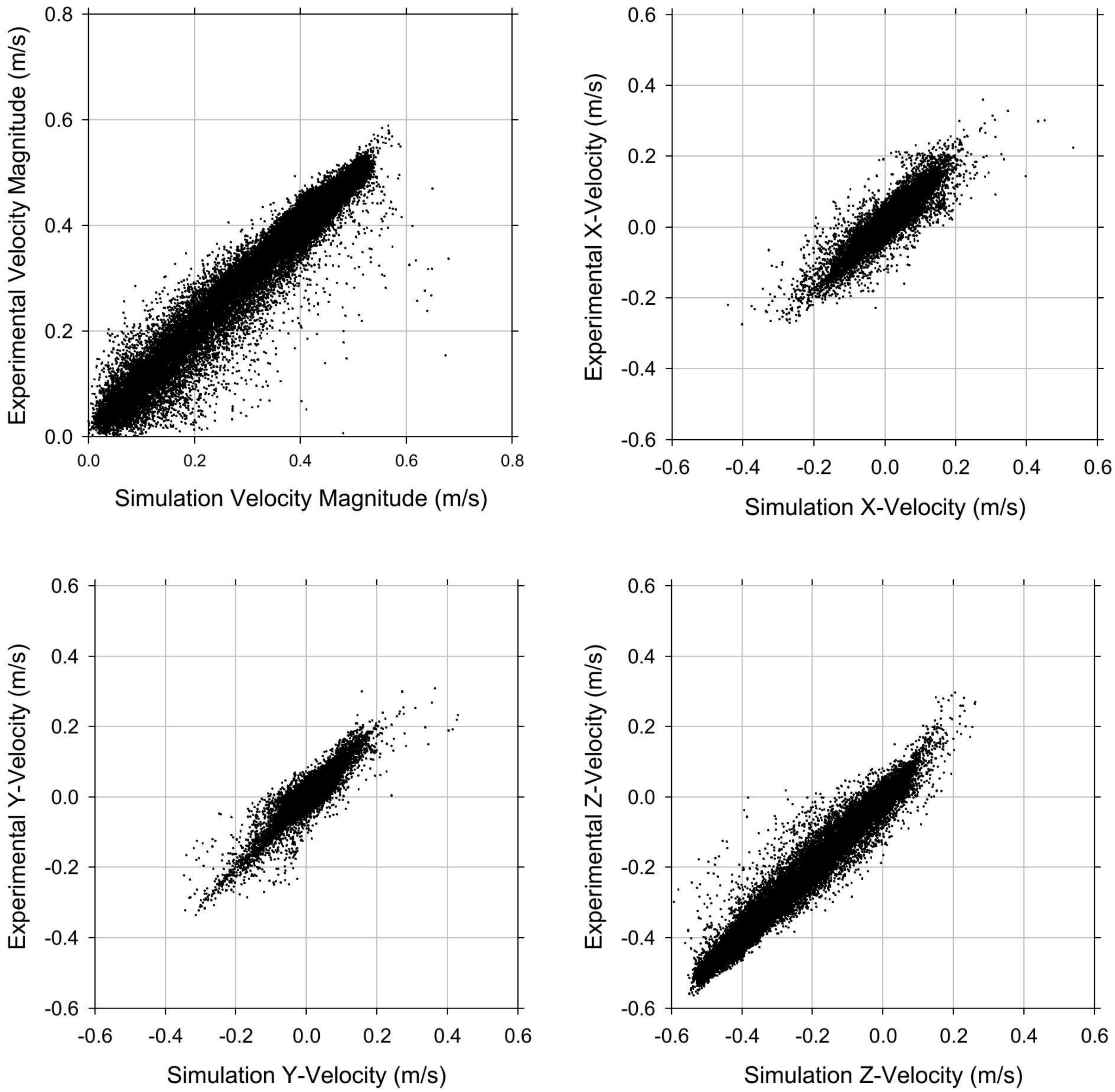

Fig. 5. Velocity magnitude and component comparisons for the full dataset

Table 3. Quantitative metrics for the accuracy of the model versus prediction for the rectangular region

Parameter	Velocity Magnitude	Pass/Fail
FB	-0.0111	Pass
NMSE	0.0076	Pass
MG	-0.0053	
VG	1.0560	
FAC2	0.9754	Pass
NAD	0.0296	Pass
LND	0.0862	
LLR	0.1059	

3.3 Correlations

Gradients of the measured data are helpful for alignment, but also for assessing general accuracy. Using a linear correlation analysis and illustrating the coefficient of determination (R^2) between model and simulation results, the data in Table 4 give an extended indication of the accuracy of the comparisons using the two methods for clipping the data.

Table 4. Correlation analysis results between data and model for the concentration and box clip data

Variable	Concentration Clip R^2	Box Clip R^2
Concentration Magnitude	0.603	0.691
Velocity Magnitude	0.841	0.952
X Velocity	0.778	0.825
Y Velocity	0.687	0.848
Z Velocity	0.928	0.968
Concentration Gradient Magnitude	0.542	
Concentration X Gradient	0.295	
Concentration Y Gradient	0.299	
Concentration Z Gradient	0.124	
Vorticity Magnitude	0.433	0.630
X Vorticity	0.423	0.502
Y Vorticity	0.514	0.579
Z Vorticity	0.268	0.286
Q-criterion	0.069	0.051

Prior work on a different dataset did not suggest good comparison for Q-criterion between data and model [16]. This continues to hold true, with poor coefficient of determination (R^2) below 0.1. The box clip generally results in better correlations between the data and model. This is consistent with prior findings [16] and is believed to be because the concentration release point at the base of the structures is in the urban canyon where it also coincides with higher fluctuations and lower average velocities and velocity accuracy. This point merits future exploration. The gradient variables did not exhibit as strong of a correlation as some prior work [16], which suggests the potential for improved comparisons with better gradient alignment. Concentration gradients did not exceed an R^2 of 0.55, and vorticity R^2 values were as high as 0.63. These are indicators of good alignment with the possibility of improvement with some subtle changes.

3.4 Error Locations

A particularly powerful analysis method is possible with a full 3D dataset. It makes possible identifying not only the cumulative accuracy of the full dataset over a selected domain, but also indicating where the comparisons were of the lowest accuracy. This might in the future enable identification and targeting of regions of complex flow for model improvement. Here we select the LND comparison metric as the selection criterion. For both datasets, the maximum LND values are sorted from the dataset. These are illustrated and mapped back onto the 3D geometry to give a graphical spatial picture of where the data and model comparisons are the least similar.

For the concentration clipped dataset, the LND parameter for velocity and concentration were multiplied with each other to identify the locations where both comparisons were poor with high magnitudes of the product variable. 100 of the highest error points were selected for illustration. Figure 6 shows the locations of these points. Generally, the simulation concentrations were very low, with the experimental concentrations also being low, but much higher than the simulated values such that the error was high. The velocity magnitudes were generally low (suggesting wake regions), with the experimental and modeling parameter varying as to which is much higher than the other. Figure 7 graphically identifies the locations plotted back on the original geometry. A cluster of points are found on W Main Avenue in the wake of the up-wind building. Most of the rest are found on the complex parking structure northeast of the release. The surface mesh in this and like figures is colored by elevation, with the darker surfaces representing higher surfaces on taller structures. Points are colored by LND.

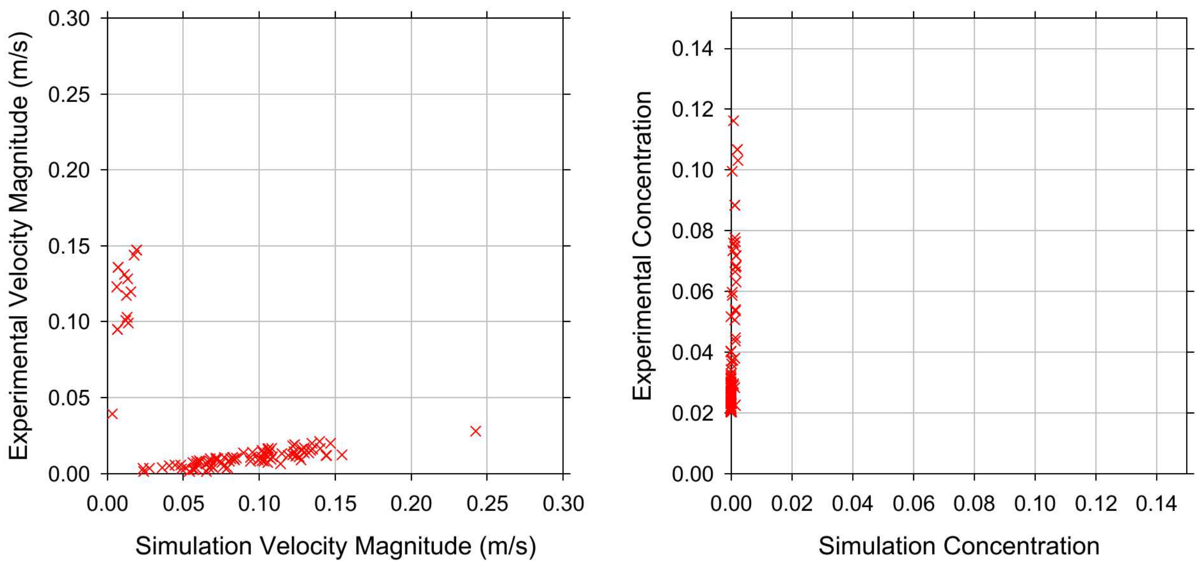

Fig. 6 scatter plots highlighting selected points with the highest error for the concentration clipped data

The box clipped data suggest somewhat different error locations. Here displayed are the 200 points with the highest LND error. With few exceptions, the highest errors based on the LND parameter for velocity magnitude were on or very near the surfaces. The scatter plot in Fig. 8 suggests errors were due to a combination of much lower velocities from the experiment and the model. Were these only low experimental values, we might conjecture that we were unsuccessful at clipping away all of the erroneous surface data that are distorted by the boundaries. The majority appear to fit that categorization, but some do not. Fig. 9 helps locate the points of highest error by providing two views of the high error locations. With the exception of a cyan point close to the release plane in the wake of an intermediate sized building above W. Main St., most other points are near surfaces. This is perhaps indication of the errors associated with CFD surface modeling techniques, or errors relating to variable surface roughness in the printed experimental model. A more thorough evaluation of the error data is left to subsequent work, but these comparisons suggest this effort has utility.

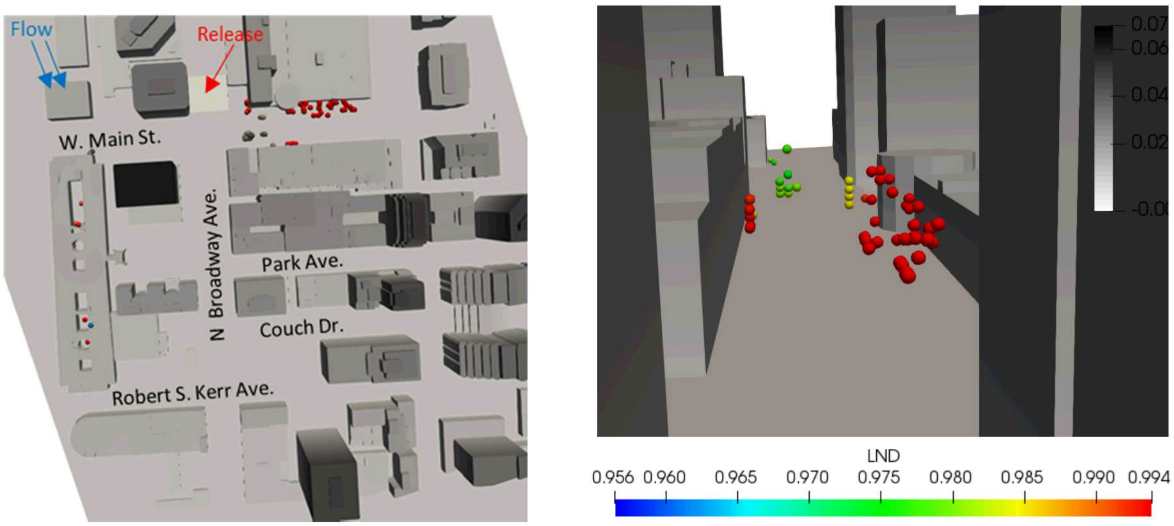

Fig. 7. Highest error points for the concentration clipped data mapped back onto the computational surface mesh. Top view (left) and side view down W. Main (right).

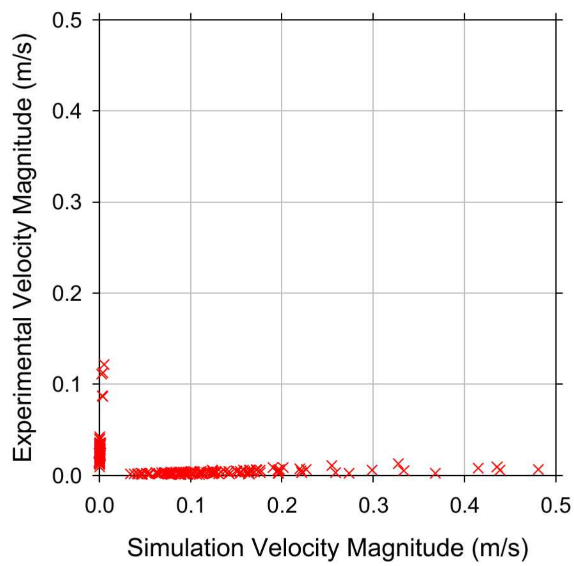

Fig. 8. Scatter plot highlighting selected points with the highest error for the box clipped data

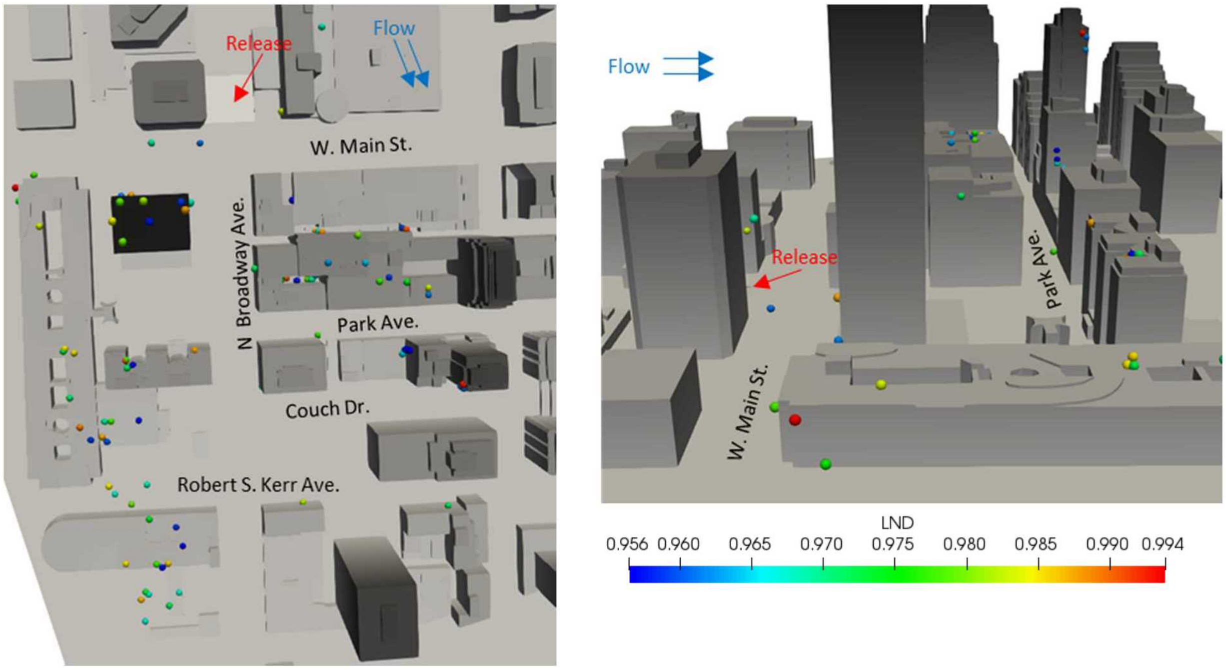

Fig. 9. Highest error points for the box clipped data mapped back onto the computational surface mesh. Top view (left) and side view (right).

3.5 General Discussion

This work exhibits the application of some analysis techniques that demonstrate utility to quantify and identify error in CFD simulations in relation to a very rich experimental dataset. The 0.8 mm experimental resolution provided comparable resolution to that of the CFD, which is unusually detailed for flow comparisons. Such a detailed comparison represents a significant improvement over traditional point, line, and planar data comparisons that risk giving a false sense of adequacy when local comparisons could be good, but global comparisons could differ substantially. The fact that concentration and velocity were simultaneously measured also adds value to the comparison.

This case exhibited generally good velocity data comparisons but was lacking in the concentration predictions. This is suggestive of problems with the concentration simulations. The general advective nature of the velocity appears to be quite accurate based on the quantitative comparison, although the lower performance metrics for the concentration clipped dataset suggest the velocities are off preferentially where the concentrations are higher. This suggests a potential issue with the injection model, a possibility that was noted in the initial planar comparisons between the data and model (observed, but not exhibited in this paper). The boundary conditions and/or flow development in and around the injector might be causal. Other error sources are also worth consideration, like the diffusion methods used in the models and the slight density gradients in the tests caused by introduction of the contaminant that enables detection. The identification of the highest error datapoints in the wake of structures on W. Main St. suggest potential difficulties with comparisons in the wakes of structure.

The poor Q-criterion comparisons are surprising given the high accuracy exhibited in the velocity comparisons. Vorticity comparisons suggested a correlation, but not as strong as the velocity. One potential interpretation of this could be that the code is struggling with vorticity that is sourced from vortical fluid motion, the contribution intended to be better isolated by the Q-citerion. The vorticity equations also account viscous shear in flows as 'vorticity', and this could be the component of the computed vorticity that lends to the improved correlation. The Q-criterion is also in different units than vorticity, which may play a role in the accuracy. As previously noted [16], gradient parameters have been shown to be leading indicators of inadequate spatial alignment, and correlations for the primary variables can be fairly good even in mis-alignment situations. It merits consideration to perform a more detailed follow-up study that evaluates the gradient parameters and the mesh alignment. Even though the velocity vector comparisons suggest very good alignment, optimal alignment cannot be certified without a more rigorous study of the dataset comparisons. This alignment challenge was not anticipated in pre-comparison work and seems to be a consideration for additional effort in subsequent studies. The general issues with gradient comparisons merits further attention.

This data comparison exposes another unique challenge associated with the existence of such detailed experimental and simulation data. The computational methods for analysis restricted what could have been an even more detailed comparison. The consequences of thinning the comparisons as was done in this case is probably not severe, but it would be advisable to verify this. Such a task is relegated to follow-on work. Another potentially productive follow-on activity might also be to develop or identify the techniques that enable more detailed complete comparisons of the datasets through large-scale data techniques and automated analysis.

The methods for dealing with full 3D dataset comparisons such as this one deserve further consideration. We are unaware of similar prior work on CFD comparisons with comparably rich datasets. The HC-2012 suggested comparison statistical methods along with the additional ones we have introduced to this problem appear to have value in several ways. First, the quantitative measures give a way to express the accuracy of the comparisons in a way that stands alone quantitatively, but also serves as a relative magnitude of potential accuracy. We see value in these methods as well as the comparison plots produced here for future studies of this nature. As this type of comparison becomes more common, this paper hopefully provides guidance that was not initially obvious relating to how to perform a valuable comparison for these types of rich datasets. The HC-2012 thresholds for acceptance were passed, but these do not appear to be as relevant to this type of comparison as it was for the author's application. Future work may focus on defining a more discriminating success measure for CFD applications. It is anticipated that there is utility in studying ways to improve comparison and assessment techniques to improve the quality of the validation effort.

A main objective of validation is to be able to quantify simulation accuracy and to identify potential improvements or shortcomings. This work that includes the novel experimental techniques coupled with detailed and comprehensive analysis methodologies represents a step towards generally improved modeling capabilities that will enable modeling to impact a greater range of problems in the future.

4. Conclusions

The MRC/MRV techniques were used to perform a validation comparison with Sandia's SIERRA/Fuego fire simulation CFD software. A model for Oklahoma City was used as a geometric case for a highly complex flow geometry. Remarkable accuracy was found in the velocity comparisons, with many of the quantitative benchmark measures exhibiting very low error. Concentration predictions were less accurate, but still suggest good representation of the bulk trends. This novel application of technical capabilities provides confidence and quantifiable accuracy metrics for calculations including the momentum, species, and turbulence models in a CFD code. Results of high error locations are mapped back onto the original geometry, suggesting regions where models are least accurate compared to the data. This work illustrates a novel and promising methods approach to validation that helps quantify model accuracy for CFD applications in 3 dimensions.

5. Acknowledgements

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525.

6. References

[1] Oberkampf, W.L. and Roy, C.J., 2010. Verification and validation in scientific computing. Cambridge University Press.

[2] Brown, A., Bruns, M., Gollner, M., Hewson, J., Maragkos, G., Marshall, A., McDermott, R., Merci, B., Rogaume, T., Stoliarov, S. and Torero, J., 2018. Proceedings of the first workshop organized by the IAFSS Working Group on Measurement and Computation of Fire Phenomena (MaCFP). *Fire safety journal*, 101, pp.1-17.

[3] O'hern, T.J., Weckman, E.J., Gerhart, A.L., Tieszen, S.R. and Schefer, R.W., 2005. Experimental study of a turbulent buoyant helium plume. Journal of Fluid Mechanics, 544, pp.143-171.

[4] Benson, M.J., Elkins, C.J. and Eaton, J.K., 2011. Measurements of 3D velocity and scalar field for a film-cooled airfoil trailing edge. Experiments in fluids, 51(2), pp.443-455.

[5] Chang, S., Elkins, C., Alley, M., Eaton, J. and Monismitha, S., 2009. Flow inside a coral colony measured using magnetic resonance velocimetry. Limnology and Oceanography, 54(5), pp.1819-1827.

[6] Elkins, C.J. and Alley, M.T., 2007. Magnetic resonance velocimetry: applications of magnetic resonance imaging in the measurement of fluid motion. Experiments in Fluids, 43(6), pp.823-858.

[7] Elkins, C.J., Alley, M.T., Saetran, L. and Eaton, J.K., 2009. Three-dimensional magnetic resonance velocimetry measurements of turbulence quantities in complex flow. Experiments in Fluids, 46(2), pp.285-296.

[8] Elkins, C.J., Markl, M., Iyengar, A., Wicker, R. and Eaton, J.K., 2004. Full-field velocity and temperature measurements using magnetic resonance imaging in turbulent complex internal flows. International journal of heat and fluid flow, 25(5), pp.702-710.

[9] Elkins, C.J., Markl, M., Pelc, N. and Eaton, J.K., 2003. 4D Magnetic resonance velocimetry for mean velocity measurements in complex turbulent flows. Experiments in Fluids, 34(4), pp.494-503.

[10] Laskowski, G.M., Kearney, S.P., Evans, G. and Greif, R., 2007. Mixed convection heat transfer to and from a horizontal cylinder in cross-flow with heating from below. International journal of heat and fluid flow, 28(3), pp.454-468

[11] Brown, A.L., Benavidez, E. "Dispersion Validation for Flow Involving a Large Structure," SANLLR018-9380, August 2018

[12] Brown, A.L., E. Benavidez, M.D. Clemenson, M.J. Benson, C.J. Elkins, "Contaminant Dispersion Validation Simulations for an Urban Inspired Scenario," 4th Thermal and Fluids Engineering Conference (TFEC) April 14–17, 2019 Las Vegas, NV, USA, TFEC-2019-27457.

[13] Shim, G., Prasad, D., Elkins, C.J., Eaton, J.K. and Benson, M.J., 2019. 3D MRI measurements of the effects of wind direction on flow characteristics and contaminant dispersion in a model urban canopy. *Environmental Fluid Mechanics*, 19(4), pp.851-878.

[14] Allwine, K.J., M.J. Leach, L.W. Stockham, J.S. Shinn, R. P. Hosker, J.F. Bowers and J.C.Pace, 2004: Overview of joint urban 2003 - an atmospheric dispersion study in Oklahoma city. Symp. Plan. Nowcast. Forecast. Urban Zone, AMS, Seattle, WA.

[15] Benson, M., Wilde, N., Brown, A.L., Elkins, C., "Detailed Measurements of a Contaminant Dispersed in an Oklahoma City Model," submitted to Atmospheric Environment, 2019.

[16] Brown, A.L., Jones, S.T., Clemenson, M.D., Benson, M.J., Elkins, C.J., "3D Analysis of Concentration and Flow using the MRC/MRV 90 Degree Flow Data and CFD Predictions from SIERRA/Fuego," accepted to the Second Pacific Rim Thermal Engineering Conference December 13-17, 2019, Maui, Hawaii, USA, PRTEC-24360.

[17] Aro, C., A. Black, A. Brown, S. Burns, B. Cochran, S. Domino, G. Evans, D. Glaze, L. Gritzo, H. Hewson, B. Houf, M. Martinez, C. Moen, E. Newren, V. Nicolette, J. Sutherland, W. Tauber, J. Templeton, S. Tieszen and G. Wagner. (2018). Sierra Fuego Theory Manual – Version 4.50. Sandia National Laboratories, Albuquerque, New Mexico. SAND 2018-12012.

[18] Kim, W.W., Menon, S., Kim, W.W. and Menon, S., 1997, January. Application of the localized dynamic subgrid-scale model to turbulent wall-bounded flows. In 35th aerospace sciences meeting and exhibit (p. 210).

[19] Hanna, S. and Chang, J. "Acceptance criteria for urban dispersion model evaluation." Meteorology and Atmospheric Physics, 116(3-4), pp.133-146, 2012.