

Reduced-Order Urban Wind Interference

Samuel Wilkinson*, Gwyneth Bradbury, Sean Hanna

Institute for Environmental Design and Engineering, University College London
*ucftsmw@ucl.ac.uk

Abstract

A novel approach is demonstrated to approximate the effects of complex urban interference on the wind-induced surface pressure of tall buildings. This is achieved by decomposition of the domain into two components: the obstruction model (OM) of the static large-scale urban context, for which a single computational fluid dynamics (CFD) simulation is run; and the principal model (PM) of the isolated tall building under design, for which repeatable reduced-order model (ROM) predictions can be made. The ROM is generated with an artificial neural network (ANN), using a set of feature vectors comprising an input of local shape descriptors and a range of wind speeds from a training geometry, and an output response of pressure. For testing, the OM CFD simulation provides the flow boundary condition wind speeds to the PM ROM prediction. The result is vertex-resolution surface pressure data for the PM mesh, intended for use within generative design exploration and optimisation. It is found that the mean absolute prediction error is around 5.0% (σ :7.8%) with an on-line process time of 390s, 27-times faster than conventional CFD simulation; considering full process time, only 3.2 design iterations are required for the ROM time to match CFD. Existing work in the literature focuses solely on creating generalised rules relating global configuration parameters and a global interference factor (IF). The work presented here is therefore a significantly alternative approach, with the advantages of increased geometric flexibility, output resolution, speed, and accuracy.

1 Introduction

In the context of wind simulation for tall buildings, meaningful results can only be achieved through an appropriate application of boundary conditions. One example is the effect of the surrounding environment on the wind field around a chosen building. In practice, the difference between flow behaviour with and without such context, termed ‘wind interference,’ can have a significant effect on predictions.

Computational fluid dynamics (CFD) analysis in architectural practice typically involves response times that are obstructive to the fast iterations of contemporary generative design. In this parametric paradigm, architects can easily generate immense numbers of alternative scenarios but are then faced with the time-consuming task of evaluation and selection. The assessment of isolated tall buildings in itself is an intensive task, which is exacerbated when extending the scope to include context. A coarser resolution is required due to the larger domain size and computational restrictions, therefore slower simulations means fewer options can be evaluated and optimisation is infeasible.

Previous work [1] demonstrated the speed and accuracy of a reduced-order model (ROM) based on the use of a geometric feature vector ($P\{z, \mathbf{n}, \mathbf{n}\sigma^r, \mathbf{u}\}$). The objective, as here, is to match the rapid generation of design alternatives with accurate analysis of equal speed. The method was applied to the prediction of wind-induced surface pressure on isolated tall buildings, and aimed at parametric CAD tools such as *GenerativeComponents*. The ROM was generated with an artificial neural network (ANN) trained on a set of procedural tall building models which are evaluated with steady, time-averaged RANS (Reynolds-Averaged Navier-Stokes) CFD. A limitation of this previous work however was the exclusion of surrounding context; that is, the predictions were in isolation with unrealistic boundary conditions.

In our work, their problem definition and feature vector is extended to include local fluid properties ($P\{v, \mathbf{n}, \mathbf{n}\sigma^r, \mathbf{u}\}$) in order to support complex urban scenarios. For the training set, v_S is derived as the vertex’s upstream wind speed from a set of principal models (PMs, the isolated

design geometry) evaluated under a range of wind speeds; whereas for the test set, v_T is derived from the vertex’s position in the field of the obstruction model (OM, the context geometry without the design). This equates to a superimposition or combination of a large-scale, one-off contextual CFD simulation and a small-scale, repeatable design ROM prediction.

Many attempts have been made to approximate or generalise wind interference, i.e. the effect of multiple buildings in the domain (Tables 1 and 2). However, all have relied on a top-down problem definition in relating the position of identical surrounding building cuboids with a global Interference Factor (IF) for the design building. Significant improvement can be made here by: (i) increasing the geometric complexity of both the context and design models; and (ii) increasing output resolution from a single global factor to the vertex-level.

In this paper, a background review is initially presented to identify existing limitations, followed by the proposed methodology. This is then demonstrated for a realistically complex design case along with parametric sensitivity analyses on the training set size, number of required training simulations, and location of the test wind speed; leading to a discussion on the speed, accuracy, and limitations of the method.

1.1 Contribution

The primary aim of this work is to test the scalability of the ROM to cases with complex urban interference by extending the shape-based feature vector to include local wind speed. This work is therefore a development on the methodology and results of [1, 2]. Predictions of the isolated tall buildings are inadequate when considering the significant effects that dense urban environments can have on the wind-induced surface pressure. Following this, existing methods found in the literature to include interference are both limiting and not amenable to the basic ROM methodology (§3.2). The difference in approaches is fundamental: existing work focuses on explicitly describing the OM, PM, and their relation to one another with global parameters; our approach describes the PM through local shape features and the effect of the OM implicitly through local wind speeds.

2 Literature Review

Our work investigates the design of tall buildings, particularly with respect to the use of generative CAD tools. In conjunction with these tools, computational analysis can be used for guidance, exploration and optimisation of an increasingly broad selection of potential designs. The review covers analysis for generative design and the time-accuracy tradeoffs inherently involved, approaches to resolving this problem through solver (CFD) and solution (model reduction) approximation, and finally sensitivity analyses and generalisation of wind interference.

2.1 Performance-Based Generative Design

In current generative design practice, enabled by the ubiquity of computation and advances in computer aided design, integrating performance behaviours into generative models has entered the foreground [3]. Examples can be seen in the use of structural, energy and thermal, materiality, fabrication, and air movement (either internally for comfort and indoor air quality; or externally for structural or facade aerodynamics, pedestrian comfort, or pollution dispersal).

Air movement, predicted through CFD, suffers the most from restrictive response times, predominantly because of the historical focus on accuracy rather than speed (due to low-tolerance high-risk scenarios in aerospace engineering). Arguably, the margins for acceptable error are more tolerant in building design, meaning that the simulation accuracy requirements can be relaxed or traded off for speed improvements (particularly at early design stages).

In these early stages of light-weight (fast and less-accurate) performance feedback, there can be more allowance for design exploration and optimisation. This is supported by the idea of speed-accuracy trade-offs (SATs) [4], which suggests that for low-risk problems, it is often better to make faster, less accurate decisions. In other words, in the scope of the larger problem of building design, it is better to have a broader perspective on the performance variability rather than an extremely accurate but narrow perspective on fewer cases.

2.2 Time-Accuracy Tradeoffs

CFD simulation can be one of the most intensive and time-consuming stages in the assessment of building performance. Difficulty therefore typically arises in guidance for early project stage decisions due to the slow feedback from conventional CFD approaches. Paradoxically, although the simulations are most valuable at early stages, the slow-and-accurate CFD simulation is better suited to later stages where the design scope is more constrained. This leads to the fundamentally opposing requirements of a fast-yet-accurate tool.

It is therefore prudent to consider compromises in the speed-accuracy trade-off during these early stages; that is, by sacrificing accuracy for speed, so that more design options can be explored. This ongoing problem was recognised over two decades ago by [5], as the need for both speed and simulation accuracy to meet the demands of early design stages. They generate a range of reduced-order models of a combustion engine simulation, with varying accuracy and speed that can be used throughout the design process. Their solution is posed as a Pareto front of non-dominated solutions, rather than a simpler trade-off curve based on biological decision making as suggested by [4].

Approximation of some description is key to this relationship between time and accuracy. An ideal approximation of reality can replicate it instantly with no error, at least for the variables of interest. In our work, the variable of interest is wind-induced surface pressure. This in itself is a product of a number of field variables which can effectively be discarded, so long as the pressure is approximated well. This is the fundamental premise of modelling, of which approximation is a like-term, and of which two distinct approaches will now be introduced.

2.3 Solver Approximation

Most approaches towards CFD approximation focus on simplification of the solver itself. For instance: simplified meshes (spatial discretisation); the use of lower-order equations; or the treatment of turbulence through modelling. These methods can be classed as type-one, *solver approximation* (Figure 1). For instance, RANS (Reynolds-Averaged Navier-Stokes), LES (Large Eddy Simulation), and DNS (Direct Numerical Simulation) all treat turbulence with different numerical approaches, i.e. temporally, spatially, and directly.

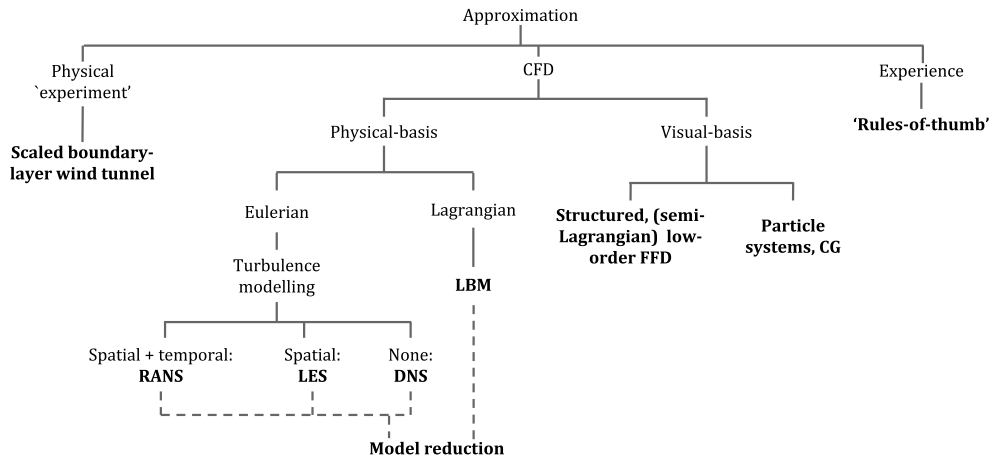


Figure 1: CFD solver approximation taxonomy.

Another example is the ‘Stable Fluids’ fast fluid dynamics (FFD) solver developed by [6] for the computer graphics and games industries, which has subsequently been developed and tested for architectural applications [7, 8, 9, 10]. Development and application for architectural design was motivated by three factors: a limited, low Reynolds number validation which suggested it as suitable for purposes beyond the scope of the validation [11, 12]; the qualitative appearance of accuracy for turbulent flows; and its remarkable speed compared to traditional CFD methods like RANS (Reynolds-Averaged Navier-Stokes). [11] implemented the FFD with a zero-equation turbulence model but found that it performed worse since it was not designed or suited to the FFD approach. It should be noted, however, that with a lack of turbulence model, the solver

relies on continuous interaction (such as game character movement) to compensate for numerical dissipation.

The general benefit of solver over solution approximation is the availability of full fluid field data, since this is discarded with the latter. Although currently with the FFD, production of surface data is also difficult due to the structured mesh approximation (voxelisation) which gives a poor geometric representation. Improving the FFD mesh is possible, and would significantly increase accuracy, but at the cost of its attractive real-time speed.

2.4 Solution Approximation

Another possible approach to this problem, type-two, is *solution approximation*. CFD originated in aeronautics and astronautics, as such there is a large quantity of work directed towards modelling and optimisation of aerofoils, fuselages, and turbine blades. An optimisation routine will often generate large data sets of simulation data, from which knowledge of the problem can be extracted. The following fall into the greater category of supervised machine learning approaches where the relationship between an input feature vector extracted from some geometry and the ground truth data output from full CFD simulation is learnt.

In one such case, a large model set of turbine blades is used with a decision tree to analyse the relationship between point deformation of models and their change in surface pressure [13, 14]. Areas of high sensitivity can then be mapped onto a pre-defined base geometry and used to focus subsequent analysis. [15] extend this work further to incorporate an evolutionary optimisation process, so as to use the information extracted from previous cases to create non-random initial populations of solutions and to guide the evolution.

Analyses that are potentially obstructive to the design process may involve partial differential equations (PDEs), such as the Navier-Stokes equations of fluid flow and the Maxwell equations in electromagnetism [16]. Whilst these methods give high-accuracy results, they are computationally expensive and cannot be computed in real-time. As a result, a design process using high-accuracy techniques has inherently slow response times and loses any desired interactivity.

Significant efforts have been made to reduce the complexity of these systems in order to make them interactive; this is generally referred to as *model reduction*. Reduced-order models (ROMs) approximate representations of system behaviours, namely for computational simulations with slow response times; with the aim to create a lower-dimensional system model whilst retaining predictive fidelity [17, 18].

The distinction between solver and solution approximation is shown in Figure 2. Solver approximation (Figure 2a), i.e. the various approaches to CFD simulation, have boundary conditions (X) and geometry as input, and direct field (Z) and derived surface (Y) simulation outputs. Typical reduced-order models (Figure 2b) act as approximations between boundary conditions (X) and a key output of interest (Y). In our case however (Figure 2c), we alter this to create a model between geometry input and Y.

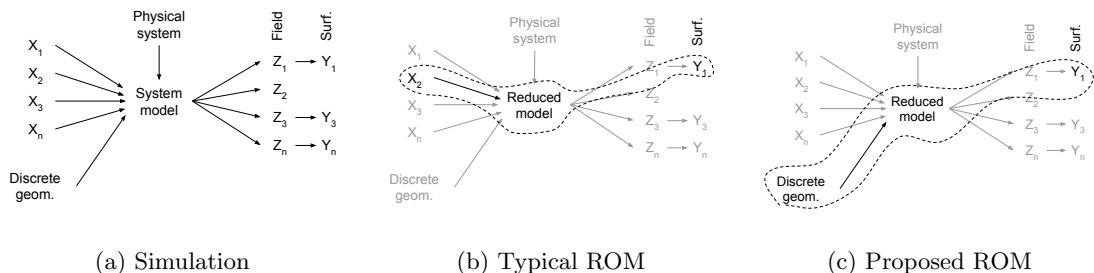


Figure 2: Reduced-order model schematic.

[19] use spatial and behavioural parameters as input feature vectors to a radial basis function (RBF). The RBF is used to interpolate and merge CFD and wind-tunnel data on pressure coefficient values (lift and drag) for aerofoil analysis. They use an input feature vector $C_p\{x, y, z, a, M, Re\}$: where x, y, z is the spatial position; a the angle of attack; M the Mach number; and Re the Reynolds number.

Whilst this method proved successful for linking behavioural characteristics (a, M , and Re) to data sources (CFD and wind-tunnel), it is limited to a single geometry, thus the use of

explicit spatial positions (x , y , and z). For cases of differing geometry between training and testing, spatial positions become non-unique and can therefore not be used within the feature vector. This necessitates the use of either explicit global design parameters or implicit local shape description.

Using spatial positions (or mesh node numbers) for a feature vector is also proposed by [20]. In this case, an ANN is used to predict post-processed CFD data for rapid visualisation and interpolation of boundary conditions with an augmented reality user display system. The input feature vector, \mathbf{X} , and output response, \mathbf{Y} , are defined as: $\mathbf{X}\{n, S, P\}$, $\mathbf{Y}\{T, V\}$, where T is the air temperature and V is the air speed at a node, n is the mesh node number (1224 nodes in the cubic room), S is the supply temperature, and P is the supply pressure. Again the proposal is strongly limited to not only a single geometry but to a single mesh by the use of spatial positions or node numbers corresponding consistently with fixed locations. The limitation again is that for differing training/test geometry the positions are non-unique.

[1] developed the ROM feature vector using $P\{z, \mathbf{n}, \mathbf{n}\sigma^r, \mathbf{u}\}$ for tall building surface pressure prediction. Predictions are made through training an ANN on these local shape features extracted from a set of procedural tall building models evaluated with RANS. The separate test set of 10 models is a selection of existing tall building models (Figure 3), showing results of a mean absolute error of between 1.994 and 4.440% (σ :2.096 to 5.088%), and an on-line response time of between 14.720 and 809.983s.

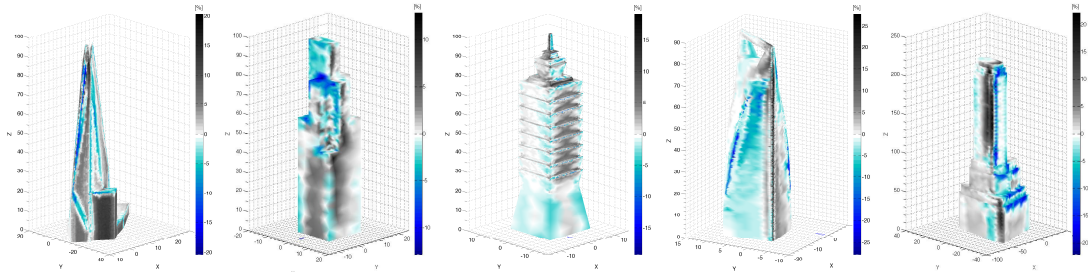


Figure 3: ROM test model prediction errors, $Y'-Y$ [%]: (a) The Shard; (b) Willis Tower; (c) Taipei 101; (d) Shanghai World Financial Centre; (e) Exchange Place. [1]

Existing work on extracting features from the fluid field is primarily focused on knowledge extraction or data mining from large sets. [21] review a number of applications such as flow topology classification, vortex identification and tracking, shock wave detection, and separation/attachment detection. Identification of flow characteristics (vortex cores, separation/attachment, and shock waves) can be ‘mined’ during a CFD simulation [22]. In extremely high resolution problems (up to 300 million grid-point transonic turbofan simulation with RANS) it can often take weeks or months for a single run to converge on a stable result. Therefore employing feature detection during the simulation can give insight into the development and stability of the actual fluid structure rather than physical properties of the simulation (e.g. mass and momentum residuals). This results in considerable time-savings if features of interest can be observed and tracked directly, potentially allowing the user to cut short the simulation when they are confident that the flow is stable enough for the accuracy requirements of the problem.

2.5 Wind Interference

Interference refers to the positive or negative effect that nearby buildings may have upon the wind behaviour of one another. Within an urban situation this is very common, and since the effects can be significant it is necessary to consider the context within the simulation; that is, independently designed buildings can not be treated in isolation.

A common misconception is that interference always reduces wind loads in comparison to the isolated case. Whilst this may be true for a uniformed array of similar buildings in close proximity, wind loads can be increased in the more complex, realistic case. The key factors in determining the effects of interference are the size, shape, and configuration of the buildings with respect to the direction of flow. The effects have been shown to be as great as up to 46% under-prediction and 525% over-prediction from regulatory loads on simple prismatic buildings [23]. An over-prediction of wind pressure is less dangerous than an under-prediction, since the

latter may cause discomfort or safety issues. [24] present a thorough review of the full past and present state of research in interference. A summary of typical studies can be found in Table 1.

Within the paradigm of such global parametric analysis, simplifications of both the problem and the solution are necessary. For all the cases in Table 1, simple cuboids are used with typical variables such as aspect ratio and position configuration; in other words, translating the objects over the two-dimensional horizontal plane. These studies analysed the effects of a small number of adjacent structures, leading to the development of the Interference Factor (IF). This is a ratio between the wind loads with and without the interference from adjacent structures. No such studies, however, have been undertaken which consider realistically complex shapes or contexts since they are typically esoteric and difficult to generalise.

Table 1: Summary of existing interference global parameter sensitivity studies.

No.	Evaluation method	O.	SD.	AR.	C.	α	Source
2	WT		• _{X,Y}	•		0.14	[25]
2	WT		• _{X,Y}	•		-	[26]
2	WT		• _{X,Y}	•		0.14	[27]
2	WT		• _{X,Y}	•		0.14	[28]
2	WT		• _{X,Y}	•		0.14	[29]
2	WT		• _{X,Y}	•		0.14	[30]
2	WT	•	• _{X,Y}	• _Z		0.19	[31]
2	WT + CFD (RNG k- ϵ)	•				0.16	[32]
2	WT		• _{X,Y}	• _{X,Y}		0.14	[30]
2 & 3	WT		• _{X,Y}	• _{X,Z}		0.16	[33], [34]
5	WT	•	•		•	•	[35]
5	WT	•		• _{X,Y}		0.15	[36]
Multi.	CFD (RNG k- ϵ)	•			•	0.22	[37]

• varied in study; - no data; WT wind-tunnel; **No.** Number of Study Buildings; **O.** Orientation; **SD.** Separation Distance; **AR.** Aspect Ratio; **C.** Configuration; α Wind profile exponent; *X* is direction perpendicular to flow; *Y* stream-wise; and *Z* vertical.

2.6 Interference Approximation

In a few cases generalisation, or regression, has been attempted (Table 2) with the IF used as output response and basic scenario parameters as input features.

Table 2: Summary of existing interference global parameter generalisation studies.

No.	Evaluation / Regression method	O.	SD.	AR.	C.	α	Source
2	WT / Polynomial		• _{X,Y}	•		0.14	[38]
2	WT / RBF		• _X	•		•	[39]
2	WT / RBF		• _{X,Y}				[40]
2	WT / RBF	•	• _{X,Y}	•		•	[38]
2	WT / RBF		•	•		•	[41]

No. Number of Study Buildings; **O.** Orientation; **SD.** Separation Distance; **AR.** Aspect Ratio; **C.** Configuration; α Wind profile exponent; *X* is direction perpendicular to flow; *Y* stream-wise; and *Z* vertical.

In the first case by [38], regression curves are fitted with either second-, third-, or fourth-order polynomials; whilst in the remaining four studies a radial basis function (RBF) ANN was used. In all five cases, the IF was collected from new or existing wind-tunnel data. The common limitations of all these studies are the simplistic geometries (cuboids of a single height), basic configurations (typically two or three buildings), and lack of output data (only a single performance metric: the IF, rather than the varied surface pressure distribution). It should be noted that in every case the studies were constrained to a limited number of cuboids, a significant simplification in attempting to create generalised interference rules.

3 Methodology

The approach here is to combine: i) a large-scale CFD simulation of an urban context, the *obstruction model (OM)*, which remains static throughout the design process and can therefore

be simulated only once; and ii) a small-scale ROM prediction of an isolated tall building, the *principal model* (PM), which can be iteratively changed through generative design.

Considerable time can be saved if it can be demonstrated that a CFD simulation can effectively be used for boundary conditions to a ROM. The clear advantage is that the full OM does not need to be re-run with every change of PM.

A realistic OM of the dense City district in London is used (Figure 4a), along with a realistic design PM (Figure 4b). These are put together for the full validation model, OM+PM (Figure 4c). Note that the geometry in Figure 4a also shows the level of detail typically found from source, and Figure 4c shows the same geometry after simplification and meshing. The geometry intends to replicate a scenario that would be found in practice. The design model (Figure 4b) is arbitrary, but is based on prior models generated at competition, massing, or form-finding project stages within practice. The design model has a height of $310m$ (Z -axis), a cross-wind (X -axis) width of $55.4m$, and an along-wind (Y -axis) depth of $41.8m$; the aspect ratio (width:height) is therefore roughly 1:6. In comparison, the upstream Swiss Re is $180m$ and the downstream Tower 42 is $183m$.

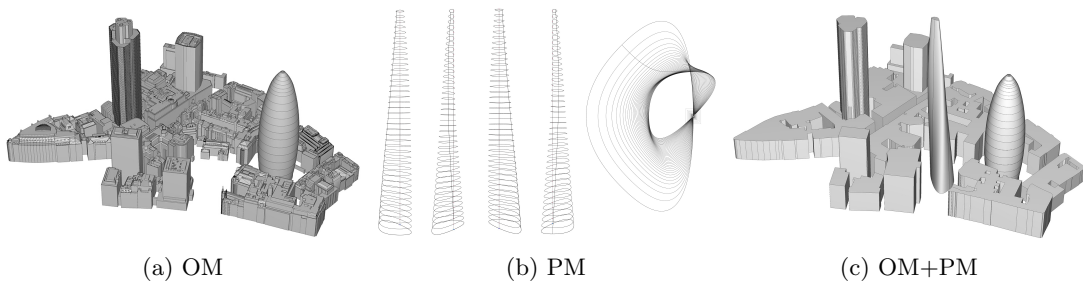


Figure 4: Components of the principal model (PM) and obstruction model (OM).

3.1 Simulation Methodology

CFX 13.0 [42] is used for the steady-state Reynolds-Averaged Navier-Stokes (RANS) simulations. Due to the wind speed, and the complexity and scale of the geometry, flow separation occurs, as well as the Reynolds number being relatively high, therefore the flow is turbulent; as such, a $k-\epsilon$ turbulence model was used. Typically the models are meshed with roughly an equal number of cells (up to the maximum available computational resources), of around four million elements. The PM simulations, for the training set, therefore have a finer resolution than the OM. The geometry is created in *GenerativeComponents* [43]. Process times are based on a 2.66GHz *i7* quadcore, 4GB RAM.

The CFD simulation domains and significant dimensions are given for the OM and PM in Figure 5. For the ground, a no slip smooth wall is assigned (i.e. fluid velocity at wall boundary is zero); for the sides and top parallel to the flow, a free slip wall (i.e. zero shear stress from wall friction); and for the outlet, a zero relative pressure opening. For the inlet, the wind profile is applied as described below, with a medium intensity turbulence and eddy viscosity ratio [44].

Basic simulation parameters are: high-resolution advection and turbulence numerics; isothermal fluid at $25^{\circ}C$; a scalable wall function; and a convergence residual target of $1.0e^{-6}$ RMS. The following meshing parameters are used: an unstructured tetrahedral domain mesh, with patch independence; a boundary surface element size of $5m$; a model surface minimum size of $0.20m$ and maximum face size $0.25m$; for prismatic expansion, a growth rate of 1.2, a transition ratio of 0.77, and a maximum of 3 layers. The wind direction is shown in Figure 6, where the flow streamlines are visualised for both OM (left) and OM+PM (right). Note that, unlike for an isolated building where the wind direction can be easily changed by rotating the model, now with the contextual OM the two are independent of one another.

In the test case (the OM simulation), the wind speed is applied at an upstream inlet with a reference speed (v_r) of $10m \cdot s^{-1}$ at a reference height (z_r) of $10m$. The most commonly used distribution of wind speed with height is the ‘power-law’ expression:

$$v_x = v_r \cdot (z_x/z_r)^{\alpha} \quad (1)$$

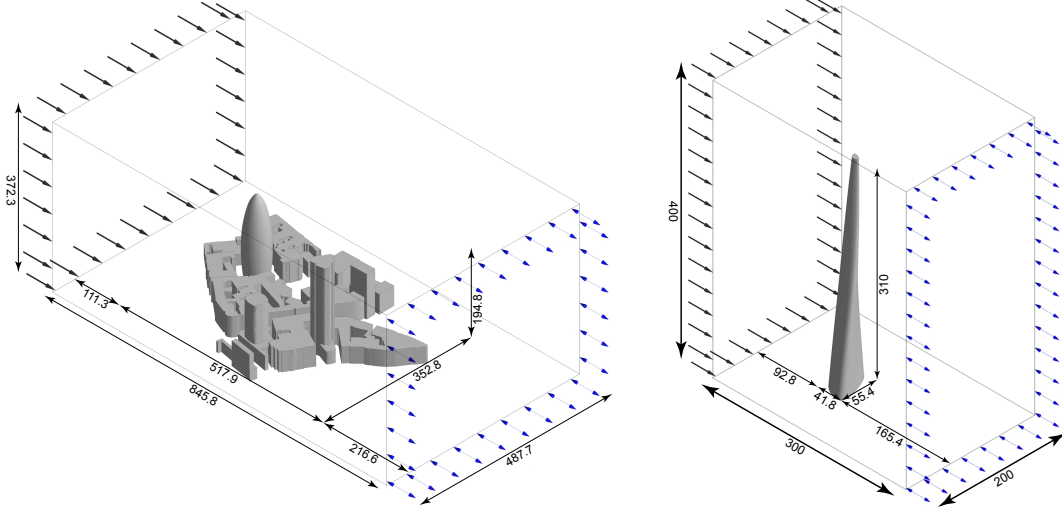


Figure 5: Simulation domain sizes: (left) OM; (right) PM.

The exponent α is an empirically derived coefficient that is dependent on the stability of the atmosphere. For neutral stability conditions it is approximately 0.143, and is appropriate for open-surroundings such as open water or landscape [45]. In the training models a constant vertical wind profile is used, albeit with varying speeds, so as to generate a range of upstream wind speeds across the simulated training set for every vertex, i.e. $v_x = v_r$.

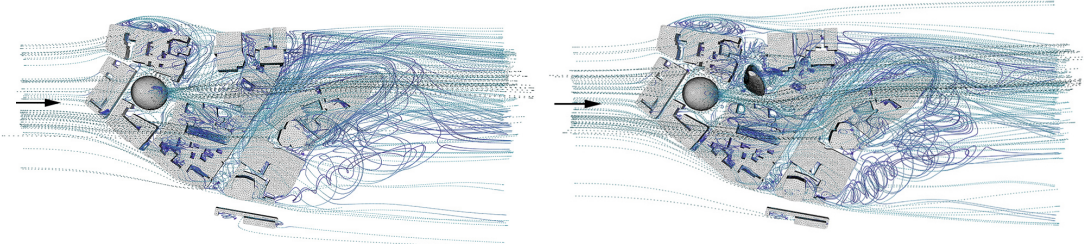


Figure 6: CFD flow field of (a) OM for testing and (b) OM+PM for validation.

A transient large eddy simulation (LES) could alternatively be used instead of RANS to achieve more accurate and time-dependent peak pressures. However, due to time and resource limitations it was not possible to include a comparison in this study.

3.2 Reduced-Order Model Generation

For a training set, \mathbf{S} , consisting of vertex feature vectors and simulated pressure extracted from the CFD, the ANN approximates the function $f^{ANN} : \mathbf{X} \rightarrow P$ where \mathbf{X} is the vertex feature vector and P is the vertex pressure. \mathbf{X} is defined as:

$$\mathbf{X}\{v, \mathbf{n}, \mathbf{n}\sigma^{1-5}, \mathbf{u}\} \quad (2)$$

where $\mathbf{n}_{x,y,z}$ are the vertex normal components (Figure 7a); $\mathbf{n}\sigma^{1-5}_{x,y,z}$ are the vertex-ring (one through five) neighbourhood curvature (non-absolute) standard deviation components (Figure 7b); and $\mathbf{u}_{x,y,z}$ are the normalised relative vertex position within the model limits (Figure 7c). For training, $v=v_S$, simply the inlet wind speed of the training simulation which is constant with height, i.e. no profile. For testing however, this is replaced with $v=v_T$, the wind speed at the vertex's transformed position in the OM fluid field.

This is essentially the same as in previous work [1], except that previously the vertex height z was used instead of the wind speed v . This was acceptable for interference-free predictions since the wind profile was the same in both the training and test models, which is now not the case.

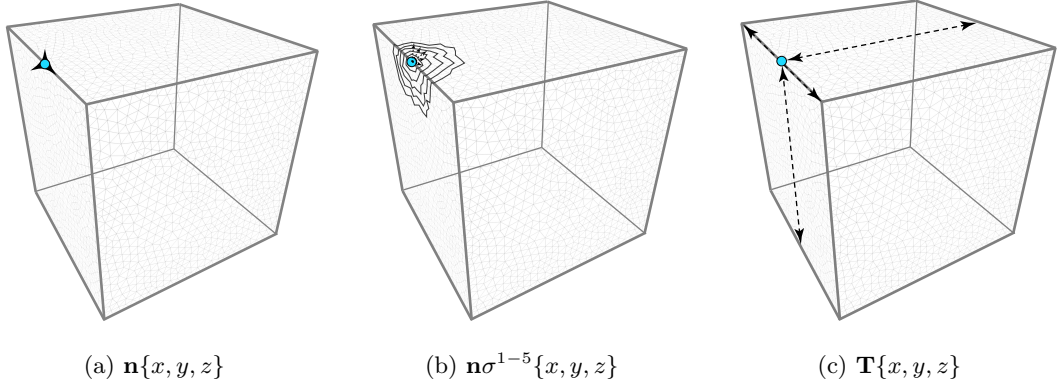


Figure 7: Locality of feature definitions.

The transformation of the vertex is either a normal offset (Figure 8a) or a projection upstream (Figure 8b) from the original location. This results in either $v_{T.offset}$ or $v_{T.upstream}$, both of which are tested for different distances d in the following section. For both transformations, d is increased from 0 at increments of $1m$ until an obstruction is encountered; $12m$ for the normal offset, and $9m$ for the upstream projection.

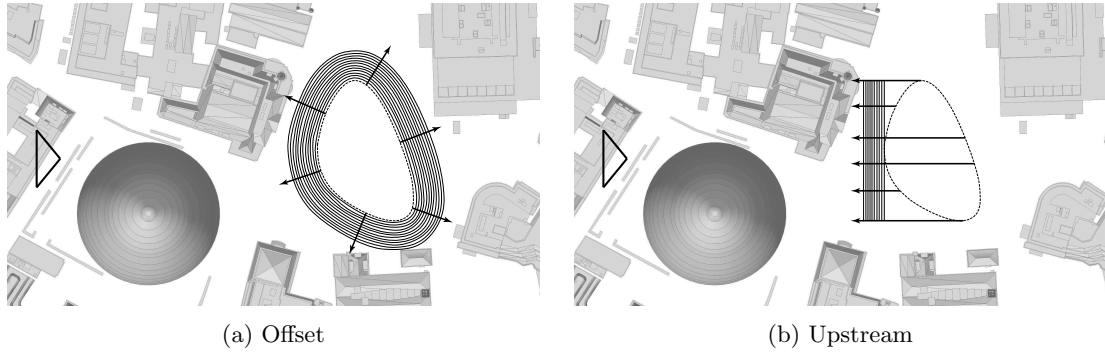


Figure 8: Test feature wind speed location from OM.

From the training feature set, the reduced-order model is generated by a back-propagation artificial neural network (ANN), with a hyperbolic tangent sigmoid transfer function [46]:

$$tansig(x) = 2/(1 + \exp(-2 \cdot x)) - 1 \quad (3)$$

The ANN structure $X:H:Y$ is 22:20:1, i.e. 22 input neurons, 20 hidden layer neurons, and 1 output. The sensitivity analysis on the number of neurons in the hidden layer, and the number of layers, is not included here; although 20 in a single layer has been seen to be sufficient. There is generally no rule-of-thumb or guidance to define either, necessitating sensitivity analysis for each problem.

4 Results

The distinction is drawn between the simulation output response Y from CFD, and the prediction output response Y' from the reduced-order model. For a single vertex sample i , the difference between the Y and Y' is used to calculate the sample prediction error, δ_i :

$$\delta_i = (Y'_i - Y_i)/(Y_{max.} - Y_{min.}) \quad (4)$$

The descriptive statistics used for reporting the errors throughout are:

There are two types of test used here: sample-based (Figure 9a) is the training error, and model-based (Figure 9b) is the validation error. In the sample-based assessment, test data set \mathbf{T} of size m and training data set \mathbf{S} of size n are drawn from the same set of available data D ,

$\delta_{min.}$	real-valued minimum of the error range [%]
$\delta_{max.}$	real-valued maximum of the error range [%]
$ \bar{\delta} $	mean of the absolute error range [%] ($ \bar{\delta} \neq \bar{\delta} $)
$\sigma_{ \delta }$	standard deviation of the absolute error range [%] ($\sigma_{ \delta } \neq \sigma_{\delta}$)

meaning that $m = D - n$. Both \mathbf{T} and \mathbf{S} are randomised in this case, and are used to monitor error convergence during the ANN training. For model-based tests, a completely different test set is used so that \mathbf{T} and \mathbf{S} are independently generated. In this case, the procedural geometry is simply the single PM model, simulated under a range of inlet wind speeds; and the test geometry is the OM.

However, the study is focused on the difference between the wind speed component in the training and test features. The sample-based test here, more precisely, is purely composed of data from the training simulation; whereas the model-based test derives its data from the OM simulation.

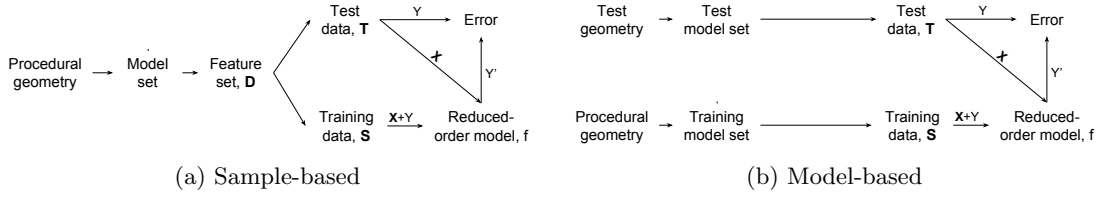
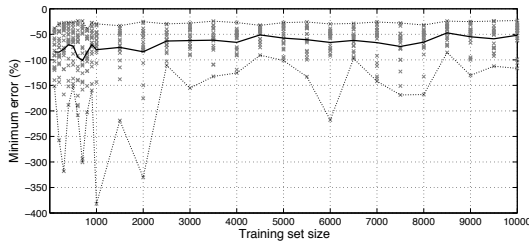


Figure 9: Testing of sample- and model-based data sets.

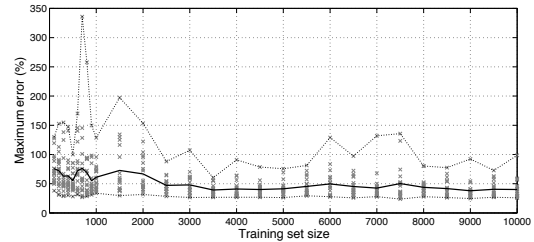
Whilst the descriptive statistics for the sample- and model-based accuracy are a good indicator of the ROM's performance, a qualitative visual comparison of simulated and predicted surface pressure is also included at the end.

4.1 Training Set Size (n)

Sample-based errors are given at the converged training set size (Figure 10) of $n=10000$. The test set size m , being the full data set D minus the training set n , is therefore $210000 - 10000 = 200000$. These are randomly selected for each training run, which is repeated 20 times. The individual runs are shown as grey crosses, with the black lines showing the limits and the blue line the mean over the 20 re-runs. The training set size is increased incrementally, by an increment of 100 from 100 to 1000, and an increment of 500 from 1000 to 10000. The converged sample-based errors are: $\delta_{min.} = -51.906\%$, $\delta_{max.} = 40.115\%$, $|\bar{\delta}| = 1.217\%$, and $\sigma_{|\delta|} = 1.756\%$.



(a) Minimum error, $\delta_{min.}$



(b) Maximum error, $\delta_{max.}$

4.2 Training Set Wind Speed (v_S)

Varying the increments of inlet wind speed in the training set simulations has an impact on the time required for initially generating the ROM. The range of wind speeds is kept constant, between 1 and $15m \cdot s^{-1}$, and the increments varied between 1 , 2 , 7 , and $14m \cdot s^{-1}$. The difference in error between an increment of 1 and $2m \cdot s^{-1}$ is minimal, yet the time saving is substantial with nearly half the number of simulations required. In fact, even with an increment of $7m \cdot s^{-1}$ the difference in error is still minimal, but with a fifth of the time required for generating training data.

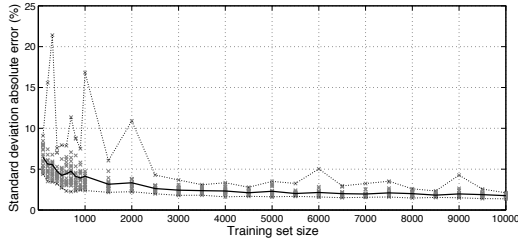
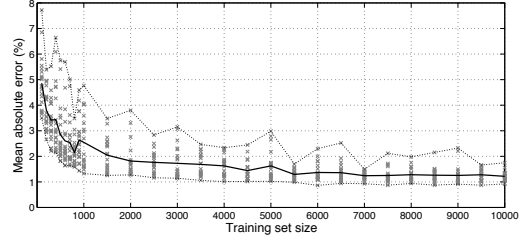
(c) Standard deviation absolute error, $\sigma_{|\delta|}$ (d) Mean absolute error, $|\bar{\delta}|$

Figure 10: Sample-based training error convergence.

Table 3: Error results summary for v_S sensitivity analysis: model-based.

Inc.	$v_S [m \cdot s^{-1}]$	No. Models	D	$\delta_{min.}$	$\delta_{max.}$	$ \bar{\delta} $	$\sigma_{ \delta }$
1	{1,2,3,...,15}	15	210000	-57.435	30.801	5.082	7.793
2	{1,3,5,...,15}	8	112000	-66.017	29.099	5.055	7.699
7	{1,8,15}	3	42000	-55.179	28.301	5.673	8.122
14	{1,15}	2	28000	-65.854	25.884	9.995	10.162

Figure 11 shows the probability density distribution of the wind speeds in the test data set from the OM with an offset of $0m$, i.e. no transformation. The probability density distribution uses a smoothing kernel with a normal distribution and a width of $0.02m \cdot s^{-1}$.

A peak (about 46%) at $11.2m \cdot s^{-1}$ is clear due to the inlet wind profile reference speed of $10m \cdot s^{-1}$ at a reference height of $10m$. Further work should establish a methodology for robustly sampling wind speeds: that is, a training set that fits the above distribution would be optimal for this case, but not for another case.

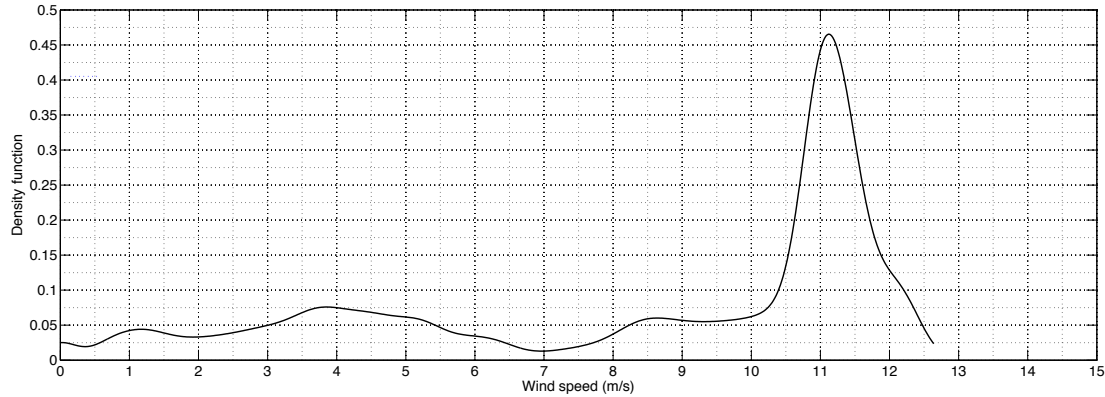


Figure 11: Test set wind speed probability distribution at PM.

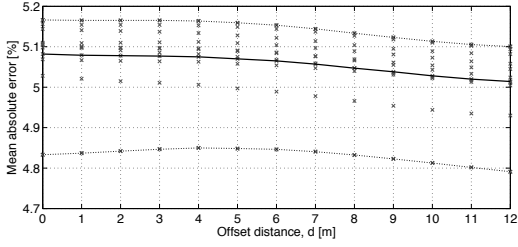
4.3 Test Set Wind Speed Location ($v_{T.offset}$ vs. $v_{T.upstream}$)

Two transformation methods, normal offset and upstream projection, are tested for varying distances d (Figure 8). The geometric transformation is applied to the PM mesh, positioned in the OM field; from which the wind speeds for the test feature vector, $v_{T.offset}$ or $v_{T.upstream}$, are calculated.

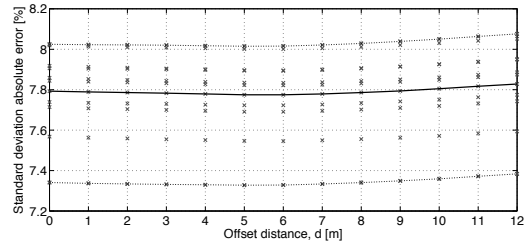
In fact, the prediction errors in Table 4 for both transformation methods with a varying d , suggest they are relatively invariant to the test feature wind speed location. For the offset there is a standard deviation over the range of absolute mean errors of only 0.024%, and only 0.014% for the upstream. Figures 12 and 13 plot the mean absolute (left) and standard deviation absolute (right) errors against distance for both the offset and upstream transformations. The individual runs are shown as grey crosses, with the black lines showing the limits and the blue line the mean over the 10 re-runs.

Table 4: Error results summary for test location sensitivity analysis: model-based.

$v_{T.offset}$					$v_{T.upstream}$				
d [m]	$\delta_{min.}$	$\delta_{max.}$	$ \bar{\delta} $	$\sigma_{ \delta }$	d [m]	$\delta_{min.}$	$\delta_{max.}$	$ \bar{\delta} $	$\sigma_{ \delta }$
0	-57.435	30.801	5.082	7.793	0	-57.711	30.793	5.240	7.934
1	-57.404	30.807	5.079	7.789	1	-57.796	30.793	5.232	7.942
2	-57.386	30.811	5.078	7.786	2	-57.744	30.794	5.221	7.946
3	-57.363	30.814	5.077	7.783	3	-57.688	30.794	5.209	7.952
4	-57.342	30.817	5.075	7.779	4	-57.653	30.795	5.203	7.965
5	-57.323	30.819	5.070	7.775	5	-57.669	30.795	5.201	7.981
6	-57.304	30.819	5.065	7.775	6	-57.686	30.795	5.201	7.997
7	-57.286	30.818	5.057	7.779	7	-57.705	30.796	5.204	8.012
8	-57.303	30.813	5.047	7.786	8	-57.747	30.796	5.209	8.027
9	-57.335	30.806	5.038	7.794	9	-57.828	30.796	5.216	8.045
10	-57.386	30.794	5.028	7.805					
11	-57.427	30.773	5.020	7.817					
12	-57.425	30.772	5.014	7.828					

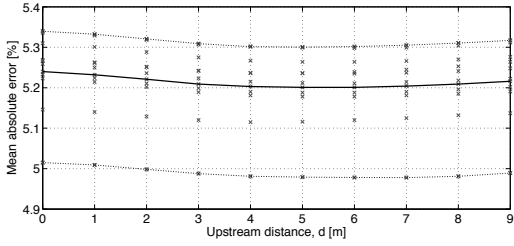


(a) Mean absolute error, $|\bar{\delta}|$

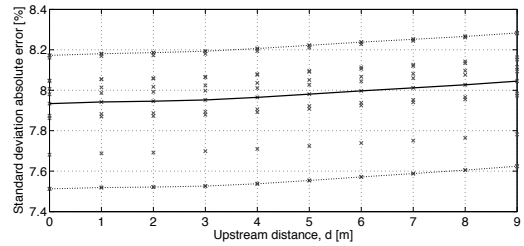


(b) Standard deviation error, $\sigma_{|\delta|}$

Figure 12: Test wind speed location: $v_{T.offset}$ distance vs. error.



(a) Mean absolute error, $|\bar{\delta}|$



(b) Standard deviation error, $\sigma_{|\delta|}$

Figure 13: Test wind speed location: $v_{T.upstream}$ distance vs. error.

With the normal offset transformation, the mean absolute error decreases as the offset distance increases so the minimum is at the greatest distance, $d=12m$. For the upstream projection transformation, the mean absolute error decreases with distance until $d=5$ or $6m$, at which point it starts to increase again.

As mentioned, the variation of error with distance or transformation method is small so the optimal location at which to measure the wind speed for the test feature is still unclear. However, it is likely to be esoteric for each problem or OM and should be studied further.

4.4 Model-Based Pressure Distribution

The following model-based results use $n=10000$, a training wind speed increment of $1m \cdot s^{-1}$, and no transformation on the test feature wind speed location. The model-based errors are: $\delta_{min.} = -57.435\%$, $\delta_{max.} = 30.801\%$, $|\bar{\delta}| = 5.082\%$, and $\sigma_{|\delta|} = 7.793\%$. Figures 14 and 15 visualise the surface pressures and errors on the PM.

In Figure 14, the top three figures are from an upstream perspective; the bottom three from downstream. Within each triplet: the left figure is the CFD simulation (Y [Pa]), centre the ANN prediction (Y' [Pa]), and right the difference (δ [%]) between predicted and simulated pressures. The inlet wind direction is indicated by an arrow on each figure.

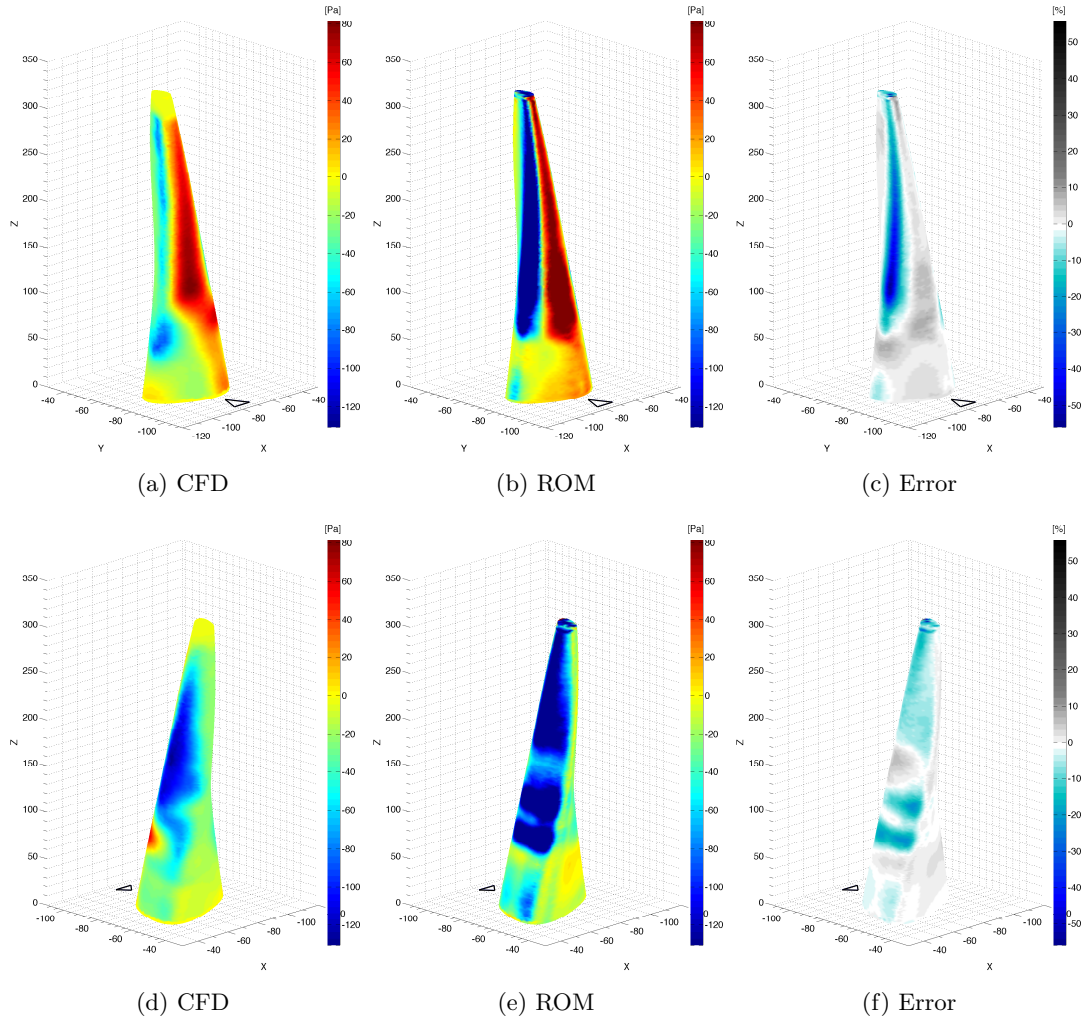


Figure 14: Model-based test: (a-c) upstream; (d-f) downstream side.

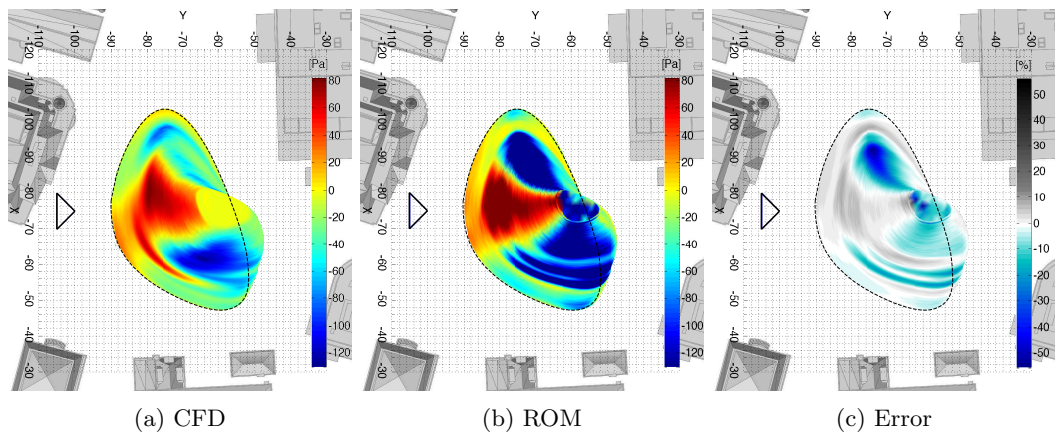


Figure 15: Model-based test: plan view.

There is generally good agreement in the spatial distribution of positive and negative pressure between the simulated and predicted models. Although the localised over-prediction (i.e. $|Y'| > Y$) of positive and negative pressure values, relative to the simulated values, is apparent in an exaggeration of the visual results. The probability distribution, or density function, of non-absolute errors [%] for 10 re-runs (grey) and their mean (black) are shown in Figure 16. A training set of 10000, an increment of $1m \cdot s^{-1}$ in the training set wind speeds, and no test

feature wind speed location transformation are applied.

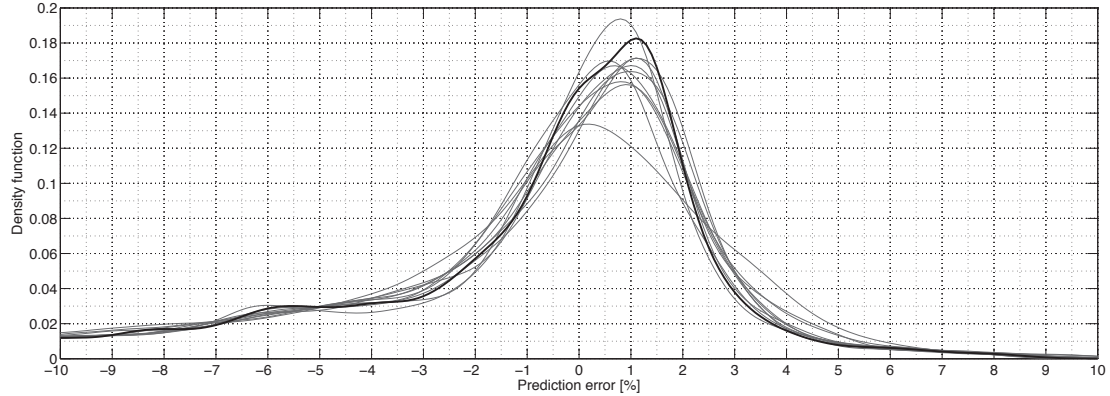


Figure 16: Probability distribution of prediction errors.

A smoothing kernel with a normal distribution and a width of 0.01% is used. The errors fit a normal distribution, with a peak probability of about 18% that the error will be 1.1%. Towards the minimum and maximum of the error range (-10 and 10%), the probability is only between 1.5 and 0% respectively.

5 Discussion

These developments represent an alternative approach that is fundamentally different to previous attempts at interference generalisation found in the literature. The use of local features rather than global parameters allows for arbitrary complexity in the obstruction model and for vertex surface pressure visualisation rather than the global interference factor.

Compared to solver approximation techniques, such as the FFD solver, solution approximation has the benefit of being based on a conventional, higher accuracy CFD solver. As such, the validity of the basis data can, to a larger extent, be trusted or verified. The comparative disadvantage is that the FFD can produce field rather than surface data which is useful for identifying flow patterns, assessing pedestrian comfort, and to gauge the secondary downstream effects that a new building will have on others.

Three sensitivity analyses were run on the training set size, the training set wind increment, and the test set wind speed location. The first found that a sample size of 10000 was adequate to reach error convergence during the ANN training. Secondly, the number of training simulations can be reduced safely from an increment of $1m^{-1}$ to 2 without much effect on the result, or even to 7 with still a reasonable effect considering the time saving. And thirdly, the results were found to be mostly invariant to the test set wind speed location, i.e. the transformation of the PM in the OM field.

Following these, the final model-based test was visualised to check the predicted pressure distribution qualitatively against the simulation. Generally, over-prediction can be seen, but general patterning or distribution of both positive and negative pressure remains intact.

5.1 Process Time Analysis

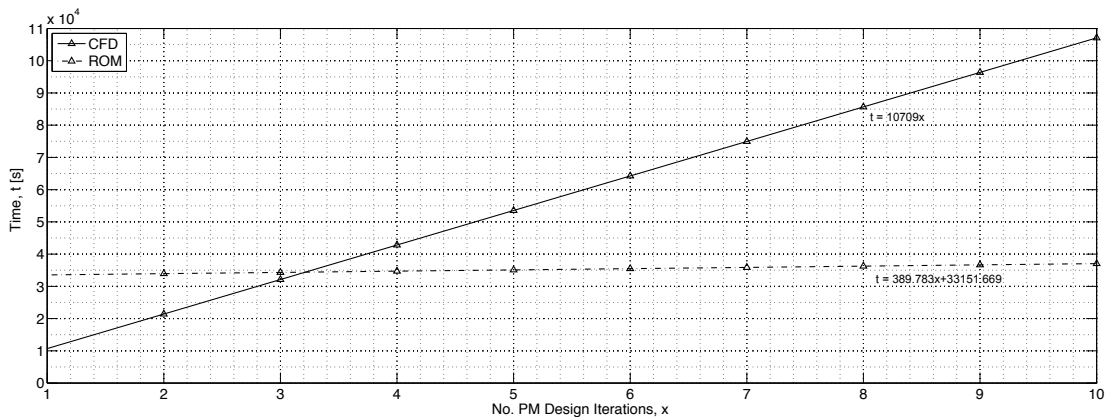
The feature extraction times are based on a calculation speed of $0.02784s/sample$, for $n=10000$ off-line ROM generation and $m=14000$ on-line prediction. And the number of PM models in the training set is 15, although it has been shown that this can be reduced to 8 or even 3 without significant reduction in accuracy. The PM simulation time is $1517s$ (28minutes) per model. The ANN training time, for $n=10000$, is averaged over 20 runs; the mean time is $38.269s$ ($\sigma:17.143s$).

The model-based prediction times show that, in comparing only on-line processes, the ROM is 27.47-times faster than the conventional CFD method. However, this does not take into account the full process. By comparing the off-line plus on-line processes for repetition, where x is the number of design iterations, the CFD time= $10709x$ and the ROM time= $389.783x+33151.669$

Table 5: CFD and ROM process times.

Conventional CFD Process		Test time [s]
PM + OM simulation	Total	10709
Off-line ROM Process		Test time [s]
PM simulations (15no.)	22755	
OM simulation	10080	
Feature extraction	278.4	
ANN training	38.269	
	Total	33151.669
On-line ROM Process		Test time [s]
PM feature extraction	389.76	
PM prediction	0.023	
	Total	389.783

(Figure 17). In solving for x , the minimum number of iterations before the full ROM process time equals the CFD is $x=3.21$.

Figure 17: Process time t against number of design iterations x for CFD and ROM.

The point at which the design team should switch from this methodology to conventional analysis depends upon the exact application and a specific consideration of the balance between accuracy and speed. Typically this might occur once the number of design alternatives has been reduced to, or any optimisation procedure suggested, a small enough number to feasibly manage with CFD or wind-tunnel. It is perhaps naive to suggest an absolute saving in time, since whatever project time is available will be filled one way or another, but rather that this fixed amount of time can be used to explore more options to the benefit of the final building.

5.2 Limitations

It is key to stress the primary limitation of this work is the similarity between the training and test geometry, i.e. the training model is the same as the PM. The paper is therefore constrained in the conclusions it can draw without subsequent testing. Further improvements and generalisation could be made by generating training shape features from a procedural model. This remains to be tested due to the combinatorial problem of sampling a procedural model under a large range of wind speeds. For instance, [1] shows that a set of 400 procedural models evaluated with RANS could be used to make predictions on isolated tall building models. Following that method, the training set would be up to 400 multiplied by the number of training wind speed iterations, i.e. 3, 8, or 15 simulations per procedural model instance. Although this would require substantial investment, it is an off-line process which would only be required once.

Further testing on alternative geometries and optimisation of the feature vector and its calculation will likely improve robustness, accuracy, and speed. The benefits of the ROM also increase with the cost of the basis simulation, for instance if a more costly RANS or LES is used.

6 Conclusion

In summary, the methodology and results presented here demonstrate an alternative approach to urban wind interference approximation for tall building design. The results indicate that significant improvements in response time (27-times faster when comparing on-line prediction times with conventional CFD) can be made with a reasonable trade-off in accuracy (mean absolute errors of around 5.0% σ :7.8%). It should be highlighted that this method is intended for use in conjunction with existing CFD and wind-tunnel analysis which, at later stages, retain their role in safe-guarding the final structural safety and performance of a building. For generative design exploration and optimisation, the current alternatives to including complex interference in wind simulations are either fast-inaccurate CFD simulations or the application of rules-of-thumb. The novel method presented in this paper has been demonstrated as a viable alternative which moves closer to the notional fast-accurate ideal.

Sensitivity analyses suggested that: i) a training set size of $n=10000$ samples was adequate; ii) wind speed increments of the training simulations v_S can be increased to 2 or even $7m \cdot s^{-1}$ without great effect, although it may potentially affect the generalisability to other scenarios; and iii) the accuracy with the test set wind speed location v_T is relative invariant to both offset or upstream transformation and distance d . Although there remain limitations to our approach before use in practice, predominantly the applicability to varying PM geometry, we believe that in conjunction with previous work it represents a significant step towards interactive building design via reduced-order wind interference modelling.

Acknowledgements

This research was sponsored by the UK Engineering and Physical Sciences Research Council, Bentley Systems, and PLP Architecture.

References

- [1] Samuel Wilkinson, Sean Hanna, Lars Hesselgren, and Volker Mueller. Inductive Aerodynamics. In *Proceedings of eCAADe 2013: Computation and Performance*, Delft, NL., 2013.
- [2] Samuel Wilkinson, Gwyneth Bradbury, and Sean Hanna. Approximating Urban Wind Interference. In *Proceedings of SimAUD SCS SpringSim'14*, number Malkawi, Tampa, FL., 2014.
- [3] Ali M. Malkawi. Developments in Environmental Performance Simulation. *Automation in Construction*, 13(4):437–445, 2004.
- [4] Lars Chittka, Peter Skorupski, and Nigel E. Raine. Speed-Accuracy Tradeoffs in Animal Decision Making. *Trends in Ecology & Evolution*, 24(7):400–407, 2009.
- [5] S. C.-Y. Lu, D. K. Tchong, and S. Yerramareddy. Integration of Simulation, Learning and Optimization to Support Engineering Design. *Annals of the CIRP*, 40(1):143–146, 1991.
- [6] Jos Stam. Stable Fluids. Technical report, Alias-Wavefront, Seattle, 1999.
- [7] Angelos Chronis, Martha Tsigkari, Adam Davis, and Francis Aish. Design Systems, Ecology and Time. In *Proceedings of ACADIA12: Synthetic Digital Ecologies*, San Francisco, CA, 2012.
- [8] Angelos Chronis, Alasdair Turner, and Martha Tsigkari. Generative Fluid Dynamics: Integration of Fast Fluid Dynamics and Genetic Algorithms for Wind Loading Optimization of a Free Form Surface. In *Proceedings of SimAUD SCS SpringSim'11*, pages 79–86, 2011.
- [9] Chrysanthi Sandy Karagkouni, Ava Fatah, Martha Tsigkari, and Angelos Chronis. Façade Apertures Optimization : Integrating Cross-Ventilation Performance Analysis in Fluid Dynamics Simulation. In *Proceedings of SimAUD SCS SpringSim'13*, number Malkawi 2004, 2013.

- [10] Panagiota Athanailidi, Ava Fatah gen Schieck, Vlad Tenu, and Angelos Chronis. Tensegrity Systems Acting as Windbreak: Form Finding and Fast Fluid Dynamics Analysis to Address Wind Funnel Effect. In *Proceedings of SimAUD SCS SpringSim'14*, pages 127–134, Tampa, FL, 2014.
- [11] Wangda Zuo and Qingyan Chen. Real-Time or Faster-than-Real-Time Simulation of Airflow in Buildings. *Indoor Air*, 19(1):33–44, 2009.
- [12] Wangda Zuo and Qingyan Chen. Fast and Informative Flow Simulations in a Building by using Fast Fluid Dynamics Model on Graphics Processing Unit. *Building and Environment*, 45(3):747–757, 2010.
- [13] Lars Graening and Bernhard Sendhoff. Shape Mining: A Holistic Data Mining Approach for Engineering Design. *Advanced Engineering Informatics*, 2014.
- [14] Lars Graening, Stefan Menzel, Martina Hasenjäger, Thomas Bihrer, Markus Olhofer, and Bernhard Sendhoff. Knowledge Extraction from Aerodynamic Design Data and its Application to 3D Turbine Blade Geometries. *Journal of Mathematical Modelling and Algorithms*, 7(4):329–350, 2008.
- [15] Soundappan Ramanathan and Lars Graening. *Knowledge Incorporation into Evolutionary Algorithms to speed up Aerodynamic Design Optimizations Soundappan Ramanathan*. PhD thesis, Universitat Stuttgart, 2009.
- [16] Joris Degroote, Jan Vierendeels, and Karen Willcox. Interpolation Among Reduced-Order Matrices to Obtain Parameterized Models for Design, Optimization and Probabilistic Analysis. *International Journal for Numerical Methods in Fluids*, 2010(63):207–230, 2010.
- [17] T. Bui-Thanh, Karen Willcox, and O. Ghattas. Parametric Reduced-Order Models for Probabilistic Analysis of Unsteady Aerodynamic Applications. *AIAA*, 46(10), 2008.
- [18] Wil Schilders. Introduction to Model Order Reduction. In *Model Order Reduction: Theory, Research Aspects and Applications*. Springer, 13 edition, 2008.
- [19] T. C. S. Rendall and C. B. Allen. Multi-Dimensional Aircraft Surface Pressure Interpolation using Radial Basis Functions. *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, 222(4):483–495, 2008.
- [20] Ravi S. Srinivasan and Ali M. Malkawi. The Use of Learning Algorithms for Real-Time Immersive Data Visualisation in Buildings. In *SIGRADI*, pages 329–332, 2004.
- [21] Frits H. Post, Benjamin Vrolijk, Helwig Hauser, Robert S. Laramee, and Helmut Doleisch. Feature Extraction and Visualisation of Flow Fields. In *Eurographics*, 2002.
- [22] Michael R. Gosnell, Robert S. Woodley, and Steven E. Gorrell. Results of Mining Data Features During Computational Fluid Dynamics Simulations. Technical report, Brigham Young University, Provo, UT, 2012.
- [23] Theodore Stathopoulos. Adverse Wind Loads on Low Buildings Due to Buffeting. *Journal of Structural Engineering*, 110(10):2374–2392, 1984.
- [24] A. C. Khanduri, Theodore Stathopoulos, and C. Bedard. Wind-Induced Interference Effects on Buildings - A Review of the State-of-the-Art. *Engineering Structures*, 20(7):617–630, 1998.
- [25] B. E. Lee and G. R. Fowler. The Mean Wind Forces Acting on a Pair of Square Prisms. *Building Science*, 10:107–110, 1975.
- [26] H. Sakamoto and H. Haniu. Aerodynamic Forces Acting on Two Square Prisms placed Vertically in a Turbulent Boundary Layer. *Journal of Wind Engineering and Industrial Aerodynamics*, 31:41–66, 1988.
- [27] Yoshihito Taniike. Interference Mechanism for Enhanced Wind Forces on Neighboring Tall Buildings. *Journal of Wind*, 42:1073–1083, 1992.

- [28] J. W. Saunders and W. H. Melbourne. Buffeting Effects of Upwind Buildings. In *Proceedings of the 5th International Conference on Wind Engineering*, pages 593–605, Fort Collins, CO, 1979.
- [29] P. A. Bailey and K. C. S. Kwok. Interference Excitation of Twin Tall Buildings. *Journal of Wind Engineering and Industrial Aerodynamics*, 21:323–338, 1985.
- [30] Yoshihito Taniike and Hideki Inaoka. Aeroelastic Behavior of Tall Buildings in Wakes. *Journal of Wind Engineering and Industrial Aerodynamics*, 28(1-3):317–327, 1988.
- [31] Nikhil Agrawal, Achal Kr. Mittal, and V. K. Gupta. Along-Wind Interference Effects on Tall Buildings. In *National Conference on Wind Engineering, India*, pages 193–204, 2012.
- [32] Aishe Zhang and Ming Gu. Wind-Tunnel Tests and Numerical Simulations of Wind Pressures on Buildings in Staggered Arrangement. *Journal of Wind Engineering and Industrial Aerodynamics*, 96(10-11):2067–2079, 2008.
- [33] Z.-N. Xie and Ming Gu. Mean Interference Effects among Tall Buildings. *Engineering Structures*, 26(9):1173–1183, 2004.
- [34] Ming Gu and Z.-N. Xie. Interference Effects of Two and Three Super-Tall Buildings under Wind Action. *Acta Mechanica Sinica*, 27(5):687–696, 2011.
- [35] K. M. Lam, J. G. Zhao, and M. Y. H. Leung. Wind-Induced Loading and Dynamic Responses of a Row of Tall Buildings under Strong Interference. *Journal of Wind Engineering and Industrial Aerodynamics*, 99(5):573–583, 2011.
- [36] Zhao Jianguang. *Interference Effects on Wind Loading of a Group of Tall Buildings in Close Proximity*. PhD thesis, The University of Hong Kong, 2008.
- [37] Aishe Zhang, Cuilan Gao, and Ling Zhang. Numerical Simulation of the Wind Field around Different Building Arrangements. *Journal of Wind Engineering and Industrial Aerodynamics*, 93(12):891–904, 2005.
- [38] A. C. Khanduri. *Wind-Induced Interference Effects on Buildings - Integrating Experimental and Computerized Approaches*. PhD thesis, Concordia University, Montreal, CA, 1997.
- [39] E. C. English and F. R. Fricke. The Interference Index and its Prediction using a Neural Network Analysis of Wind-Tunnel Data. *Journal of Wind Engineering and Industrial Aerodynamics*, 83:567–575, 1999.
- [40] A. C. Khanduri, C. Bedard, and Theodore Stathopoulos. Modelling Wind-Induced Interference Effects using Backpropagation Neural Networks. *Journal of Wind Engineering and Industrial Aerodynamics*, 72:71–79, 1997.
- [41] Aishe Zhang and Ling Zhang. RBF Neural Networks for the Prediction of Building Interference Effects. *Computers & Structures*, 82(27):2333–2339, 2004.
- [42] ANSYS. CFX v13.0, 2014.
- [43] Bentley Systems. GenerativeComponents v08.11.08.296, 2013.
- [44] ANSYS. CFX-Solver Theory Guide. Technical report, ANSYS Inc., Canonsburg, PA, 2009.
- [45] S. A. Hsu, Eric A. Meindl, and David B. Gilhousen. Determining the Power-law Wind-Profile Exponent under Near-Neutral Stability Conditions at Sea. *Journal of Applied Meteorology*, 33(1994):757–772, 1994.
- [46] T. P. Vogl, J. K. Mangis, A. K. Rigler, W. T. Zink, and D. L. Alkon. Accelerating the Convergence of the Back-propagation Method. *Biological Cybernetics*, 59(4-5):257–263, 1988.