CWE-8
June 7-11, 2026
London, ON, Canada

A Hierarchical Deep Learning Framework for High-Fidelity Prediction of Pedestrian-Level Urban Winds

Reda Snaiki [a], Jiachen Lu [b,c], Shaopeng Li [d], Negin Nazarian [b,c,e]

[a]Department of Construction Engineering, École de Technologie Supérieure, Université du Québec, Montréal, Québec, Canada, reda.snaiki@etsmtl.ca
[b]School of Built Environment, UNSW Sydney, Sydney, NSW 2052, Australia, jiachen.lu@unsw.edu.au
[c]ARC Centre of Excellence for Climate Extremes, University of New South Wales, Sydney, Australia
[d] Department of Civil Engineering and Environmental Engineering, University of Louisiana at Lafayette, LA, USA, shaopeng.li@louisiana.edu
[e] ARC Centre of Excellence for the 21st Century Weather, UNSW Sydney, Sydney, Australia, n.nazarian@unsw.edu.au

SUMMARY

Deep learning-based surrogate models provide a computationally efficient alternative to high-fidelity Computational Fluid Dynamics (CFD) for urban wind prediction but often yield low-frequency approximations that miss critical details like peak wind speeds and sharp gradients. To address this limitation, a hierarchical ‘predictor-refiner’ framework is proposed. The first stage utilizes a U-Net architecture to generate a baseline prediction from urban geometry, while the second stage employs a conditional Generative Adversarial Network (cGAN) to refine this baseline and restore high-frequency flow content. The cGAN incorporates a multi-scale generator with stepwise kernel sizes to facilitate the simultaneous learning of global flow structures and fine-grained local features. Trained and validated on the UrbanTALES dataset, the framework significantly outperforms standard baseline predictors. Quantitative analysis reveals a reduction in validation Root Mean Square Error (RMSE) by 60%, alongside marked qualitative improvements in resolving complex turbulent wakes and high-speed jets. This methodology offers a robust solution for urban planning and wind safety analysis and is integrated into the interactive web platform, Feilian Version 2.

Keywords: *Urban wind flow; Urban climate modeling; Image-to-image translation; Hierarchical model; Deep learning.*

1. INTRODUCTION

The accurate characterization of wind flow within the urban canopy is a critical challenge in environmental engineering and city planning, with direct implications for pedestrian comfort, pollutant dispersion, and structural safety (Blocken, 2015; Mittal et al., 2018). While high-fidelity Computational Fluid Dynamics (CFD) methods, particularly Large-eddy Simulation (LES), serve as the state-of-the-art for generating comprehensive wind datasets (Buccolieri et al., 2021), their immense computational cost renders them impractical for rapid design iteration or large-scale analysis. To overcome this bottleneck, deep learning-based surrogate models have emerged as efficient alternatives, utilizing architectures such as Convolutional Neural Networks (CNNs) and Fourier Neural Operators (FNOs) to emulate physics-based solvers (Xie et al., 2020; Wu and Snaiki, 2022; Peng et al., 2024; Vargiemezis and Gorlé, 2025). However, a persistent limitation remains: standard pixel-wise regression models suffer from spectral bias, producing overly smooth predictions that compress wind speed distributions. These approaches effectively filter out high-frequency details, missing critical flow features such as sharp velocity gradients and peak wind speeds.

To address this gap, a novel two-stage, hierarchical predictor-refiner framework is introduced to significantly enhance the fidelity of surrogate model predictions. The first stage employs a U-Net-based predictor to generate a stable, low-frequency baseline from urban geometry. The second stage utilizes a conditional Generative Adversarial Network (cGAN) as a refiner, specifically designed to restore the high-frequency details absent in the baseline. The cGAN's generator integrates a multi-scale architecture with stepwise kernel sizes, enabling it to capture both large-scale flow structures and small-scale turbulent features. The proposed model is trained and validated on the high-resolution UrbanTALES dataset, which encompasses both idealized building arrays and realistic urban neighborhoods, to demonstrate its ability to resolve complex wind extremes that conventional models fail to capture.

2. METHODOLOGY

The study utilizes the UrbanTALES dataset, which comprises 512 high-resolution Large-eddy Simulations (LES) of pedestrian-level wind flow across diverse urban configurations, ranging from idealized arrays to realistic neighborhoods. The prediction problem is framed as an image-to-image translation task: the input is a 2D map of building heights, and the output is the corresponding time-averaged wind speed field at a height of 1.5 meters. To address the spectral bias common in single-step regression, a two-stage predictor-refiner framework is employed. The first stage utilizes a pre-trained U-Net to generate a stable, low-frequency baseline prediction that captures dominant wind paths. The second stage employs a conditional Generative Adversarial Network (cGAN) as a refiner, taking the baseline as input to generate a high-fidelity output. Crucially, the cGAN generator incorporates a modified U-Net architecture with stepwise kernel sizes (ranging from 32 down to 4) in the encoder, allowing it to simultaneously capture broad spatial dependencies and fine-grained turbulent features.

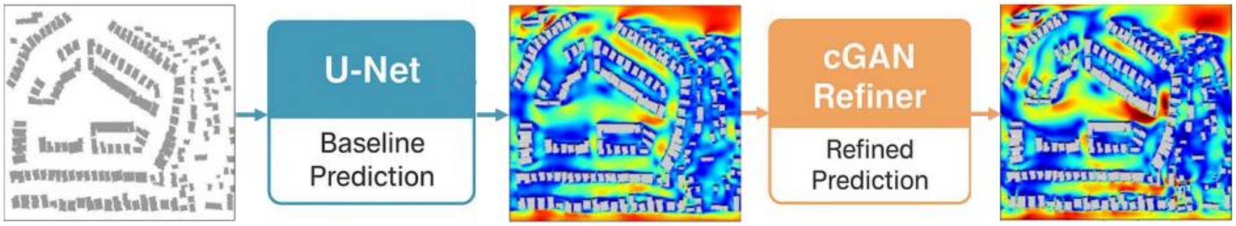

Figure 1: Overview of the two-stage predictor-refiner framework.

The training process follows a paired image-translation paradigm based on the Pix2Pix architecture. A PatchGAN discriminator is used to classify local image patches as real or fake, encouraging the generation of sharp, realistic textures. The objective function combines a standard adversarial loss (Binary Cross-Entropy) with a pixel-wise L1 reconstruction loss to ensure quantitative accuracy, with the L1 term weighted by $\lambda = 100$. Data is pre-processed by replacing building pixels with zeros and scaling all values to the range [-1, 1]. The model is trained using the Adam optimizer with a batch size of 4, utilizing an early stopping protocol based on validation L1 loss to select the optimal checkpoint and prevent overfitting.

3. RESULTS

The proposed hierarchical framework demonstrates significant quantitative improvements over the baseline U-Net predictor. Evaluated on the UrbanTALES validation split, the median Root Mean Square Error (RMSE) decreased by 59.7% (from 0.139 m/s to 0.056 m/s), while the median Coefficient of Determination (R^2) increased from 0.573 to 0.945. The model exhibits robust generalization across diverse urban morphologies. For the specific real-world configurations of Paris (FR-PA-V2_d00) and Barcelona (ES-Bar-V1_d00), the hierarchical model achieved RMSE reductions of 66.0% and 66.3%, respectively.

Figure 1 illustrates the qualitative performance for these two cities. As shown in the left panels, the baseline U-Net correctly identifies broad wind corridors but produces overly smoothed fields that fail to capture localized extremes. In contrast, the hierarchical model (middle panels) successfully restores high-frequency details, sharpening velocity gradients and accurately resolving complex turbulent wakes. The refined predictions exhibit a high degree of fidelity to the Ground Truth LES data (right panels), confirming the refiner's ability to correct spectral bias without introducing artifacts.

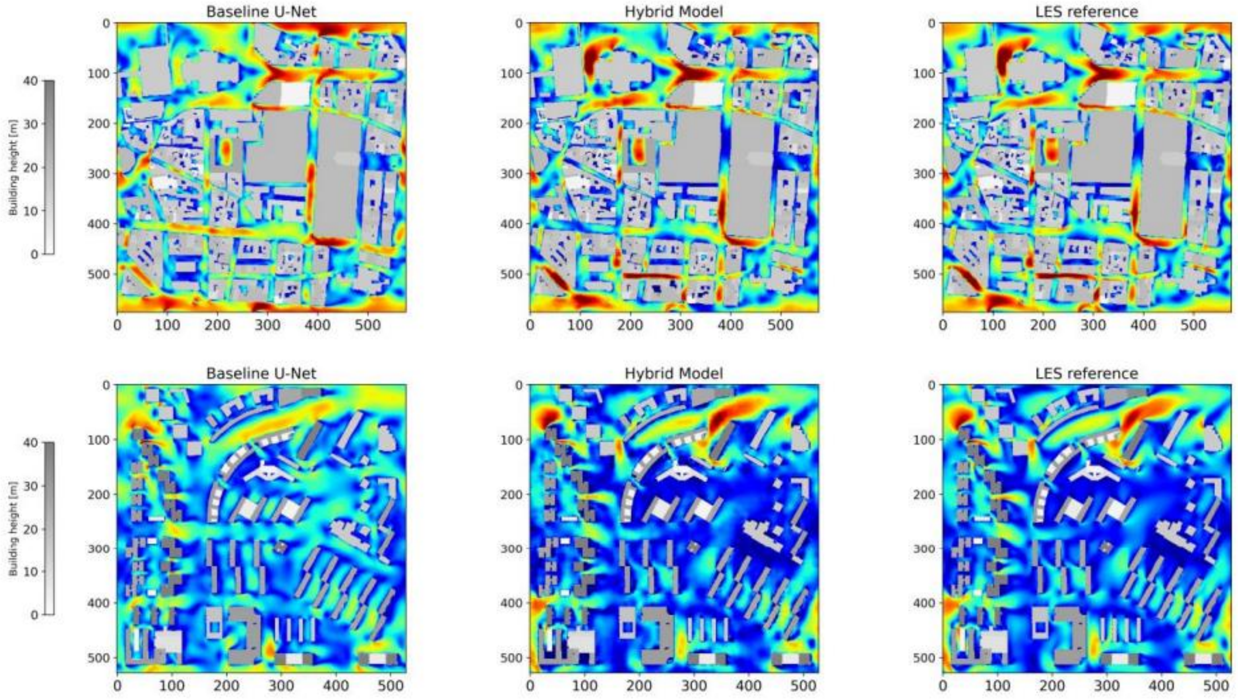

Figure 2: Qualitative comparison of wind speed predictions for (Top) Paris (FR-PA-V2_d00) and (Bottom) Barcelona (ES-Bar-V1_d00). The columns display: (Left) the smooth Baseline U-Net prediction; (Middle) the Proposed Hierarchical Model prediction, which restores sharp gradients and wake structures; and (Right) the high-fidelity LES Ground Truth reference.

4. CONCLUSIONS

This study successfully demonstrated a two-stage deep learning framework for urban wind prediction, coupling a U-Net predictor with a Pix2Pix cGAN-based refiner. This design leverages the stability of the predictor to capture large-scale flow structures while utilizing the adversarial refiner to restore critical high-frequency content missing from the baseline. Training proved robust, identifying a stable checkpoint that exhibited a minimal generalization gap and productive adversarial dynamics. Quantitative evaluations confirm consistent accuracy gains across diverse testing regimes. For realistic urban morphologies, the framework reduced RMSE by approximately 50–80% compared to the baseline. Robustness to wind direction was also established, with error reductions of 73–83% across inflow angles from 0° to 90°, maintaining tight absolute error bands between 0.044 and 0.085 m/s5. Qualitatively, the model sharpens gradients and accurately recovers complex wake and jet features, achieving high agreement with LES ground truth ($R^2 \geq 0.9$). Given that the inference cost is negligible compared to high-fidelity simulations, this approach offers a highly effective tool for early-stage urban design, pedestrian comfort assessment, and rapid hazard screening.

ACKNOWLEDGEMENTS

This work was partly supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) [grant number CRSNG RGPIN 2022-03492].

REFERENCES

Blocken, B., 2015. Computational Fluid Dynamics for urban physics: Importance, scales, possibilities, limitations and ten tips and tricks towards accurate and reliable simulations. Building and environment, 91, pp.219-245.

Buccolieri, R., Santiago, J.L. and Martilli, A., 2021, June. CFD modelling: The most useful tool for developing mesoscale urban canopy parameterizations. In Building Simulation (Vol. 14, No. 3, pp. 407-419). Beijing: Tsinghua University Press.

Mittal, H., Sharma, A. and Gairola, A., 2018. A review on the study of urban wind at the pedestrian level around buildings. Journal of Building Engineering, 18, pp.154-163.

Peng, W., Qin, S., Yang, S., Wang, J., Liu, X. and Wang, L.L., 2024. Fourier neural operator for real-time simulation of 3D dynamic urban microclimate. Building and Environment, 248, p.111063.

Vargiemezis, T. and Gorlé, C., 2025. From large-eddy simulations to deep learning: A U-net model for fast urban canopy flow predictions. arXiv preprint arXiv:2507.06533.

Wu, T. and Snaiki, R., 2022. Applications of machine learning to wind engineering. Frontiers in Built Environment, 8, p.811460.

Xie, P., Li, T., Liu, J., Du, S., Yang, X. and Zhang, J., 2020. Urban flow prediction from spatiotemporal data using machine learning: A survey. Information Fusion, 59, pp.1-12.