Evaluation of uWRF Performance and Modeling Guidance Based on WUDAPT and NUDAPT UCP Datasets for Hong Kong

Introduction

The urban Weather Research and Forecasting Model (uWRF) (WRF BEP/BEM, Salamanca et al., 2010) has been widely used to study the urban boundary layer physics of several major cities. However, its performance in modeling Hong Kong and the Pearl River Delta (PRD) region has received less attention. One reason is the lack of a complete regional dataset of urban canopy parameters (UCPs), i.e. the National Urban Data and Access Portal Tool (NUDAPT) (Ching et al., 2009). The World Urban Database and Access Portal Tools (WUDAPT) (Ching et al., 2014) approach provides an alternative estimation of building morphology dimensions based on satellite-retrieved local climate zones (LCZs) (Stewart et al., 2012). Such an approach provides a simple open source means of generating the input data required for uWRF modeling. The implementation of WUDAPT in uWRF simulation involves uncertainties that arise from various sources, as compared with the NUDAPT approach in which more accurate building data are used.

These uncertainties include the following.

(1) The supervised classification of different LCZs in the WUDAPT approach by satellite images introduces uncertainty in identifying the correct LCZs.

(2) Given that the current WUDAPT approach based on LCZs requires selecting UCPs from a range of values associated with each LCZ, the choices may not accurately represent local conditions, thus introducing some degree of uncertainty into the modeling results.

(3) As the WUDAPT approach divides urban grids into classes (LCZs), compared with the continuous UCPs of the NUDAPT approach, this discretization introduces another source of uncertainty.

(4) The subsampling of WUDAPT and NUDAPT datasets into the uWRF's domain also generates uncertainty.

The quantification of these sources of uncertainty can improve the understanding of the efficacy of the WUDAPT model compared with the results from using NUDAPT-based UCP values. Some uncertainties could be reduced. Identifying the source of relatively large but reducible uncertainty could provide a framework and guidance for other regions, thus contributing to progress in the development of next-generation WUDAPT levels 1 and 2 datasets, which are more accurate than the current WUDAPT level 0.

In this study, different methods of WUDAPT level 0 preprocessing methods are carried out to isolate the source of uncertainties (1) to (4) above. The corresponding WRF result is compared with NUDAPT, which acts as the baseline "reference"; the extent of deviation of the WUDAPT cases vis à vis the NUDAPT case provides the means to quantify different sources of uncertainty.

After quantifying those different sources of uncertainty, we offer guidance for implementing WUDAPT in uWRF to possibly minimize uncertainty. Then we evaluate the improvement in the performance of uWRF with NUDAPT, along with the corresponding and suggested WUDAPT approach as input, for comparison with the traditional Noah bulk scheme and local and non-local planetary boundary layer (PBL) schemes, by considering surface observation station data.

Study area and experiment setup

Four typical calm wind days (December 18 to December 22, 2010) with a one-day spin-up (December 18) were chosen as the study period, when the urban effect was relatively dominant. In the simulations, a total of five domains with horizontal resolution of 27 km (domain 1), 9 km (domain 2), 3 km (domain 3), 1 km (domain 4), and 500 m (domain 5) were configured respectively. Domain 5, covering Hong Kong, was the area of interest in this study. To evaluate the performance of WUDAPT in uWRF applications compared with NUDAPT, and to quantify the different sources of uncertainty, we performed a set of experiments for domain 5. The setups of each experiment and their purposes are listed below. All of the cases share the same initial and boundary conditions and domain size, as inherited from domain 4. The only differences between them are the urban scheme, the PBL scheme, and the UCPs dataset used.

Case 1: NoahACM2.

In this case, for the PRD region the Noah bulk urban scheme was selected with the non-local ACM2 PBL scheme as recommended by Xie et al. (2012).

Purpose: To determine whether a more appropriate PBL scheme or an urban scheme is more important in simulation over an urban area.

Case 2: NoahBoulac.

In this case, the Noah bulk urban scheme was selected with a local Boulac PBL scheme that was a consistent PBL scheme for the WRF BEP/BEM model for a fairer comparison, because ACM2 is currently incapable of coupling with the BEP/BEM schemes in uWRF.

Purpose: As a control to demonstrate the improvement of the WRF BEP/BEM multi-layer urban scheme compared with the bulk scheme.

Case 3: BEPreal.

In this case, only the BEP was turned on with the Boulac PBL scheme with NUDAPT-type data in Hong Kong.

Purpose: To determine whether WRF BEP (without anthropogenic heating) performs better than the Noah bulk scheme. This case can also act as a reference to quantify the uncertainties from different WUDAPT Cases 5 to 10.

Case 4: BEPcat.

In this case, the continuous building morphology data (NUDAPT) were discretized into six categories based on the LCZs 1–6 criteria.

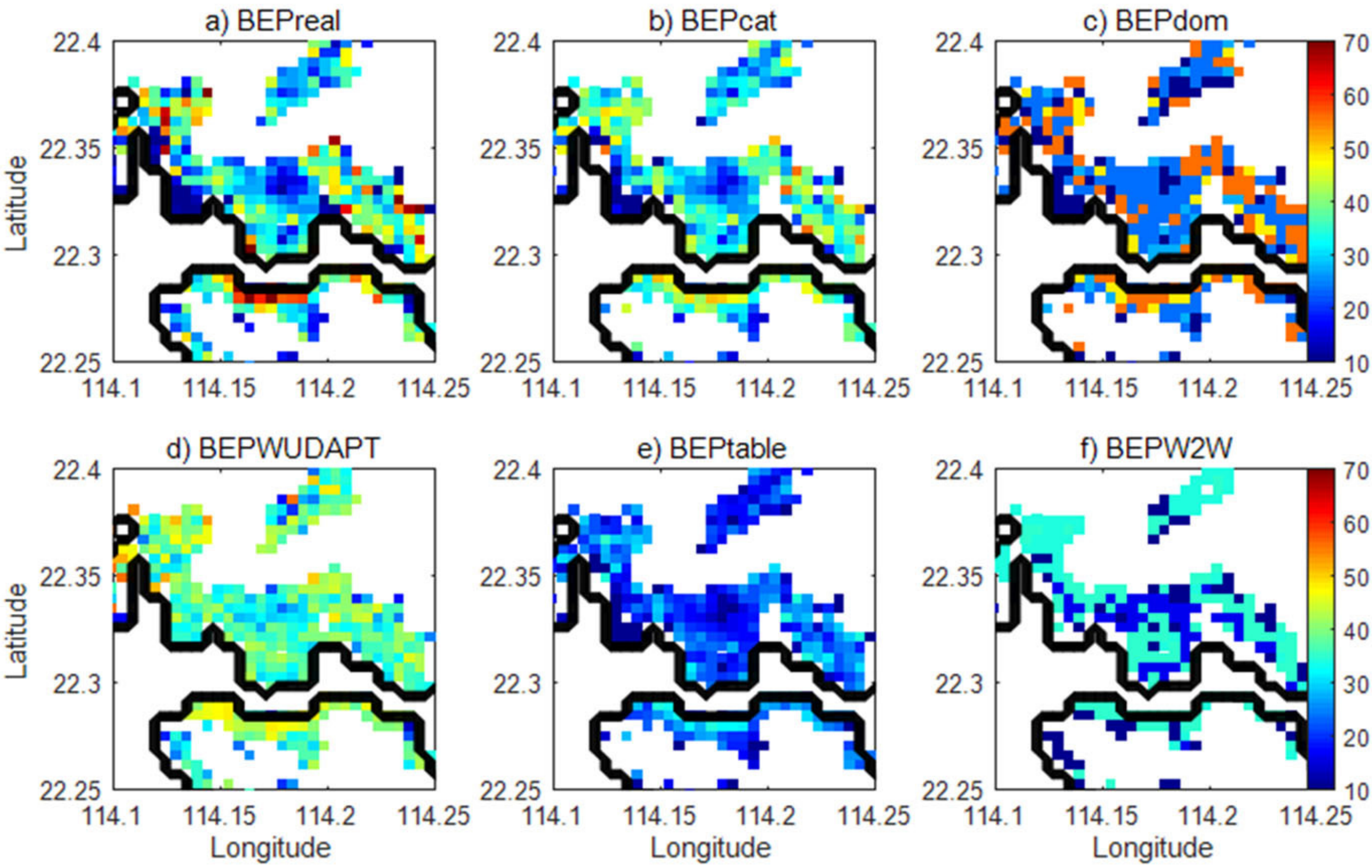


Figure 1. The zoomed-in spatial distribution of average building height [m] in Hong Kong for different cases. a) corresponds to case 3 and b) to f) correspond to cases 4-8.

Purpose: To quantify the uncertainty (by comparison with the reference case BEPreal) generated by discretizing continuous building morphology data. Given a distribution of real building morphology data (NUDAPT UCPs), how much uncertainty is generated by the simplification of these distributions by classifying them into LCZs?

Case 5: BEPdom

In this case, the data in BEPcat were subsampled on a dominant category basis rather than by subgrid averaging (Hammerberg et al., 2018, as indicated in Figure 5 and mentioned in section 3.3). Except for this case and Case 8, all of the other cases were performed based on a subgrid-averaging method.

Purpose: To quantify the uncertainty (by comparison with the reference case BEPreal) generated by the subsampling method (dominant vs. subgrid-averaging approach).

Case 6: BEPWUDAPT.

In this case, the WUDAPT dataset was fused with the NUDAPT categories in Case 4 (BEPcat) as the input data. The WUDAPT level 0 LCZs generated from the supervised classification method in Hong Kong was used with the UCP values (look-up table for Hong Kong) generated in Case 4 (BEPcat). In this way, the uncertainty was a sum of supervised classification method (machine-learning algorithm) and categorization (as in Case 5) because the look-up table values from Stewart and Oke (2012) were not used.

Purpose: To quantify the uncertainty (by comparison with the ground-truth) generated by the supervised classification method.

Case 7: BEPtable.

In this case, the UCPs in Case 4 (BEPcat) were replaced by look-up table values (Stewart & Oke, 2012). The uncertainty involved in this case was generated by the look-up table values, as the supervised classification (WUDAPT level 0 data) was not involved. The LCZs were derived from the actual building database in Case 3 (BEPreal).

Purpose: To quantify the uncertainty (by comparison with the ground-truth) generated by the lack of local information.

Case 8: BEPW2W.

In this case, the W2W protocol was used with WUDAPT level 0 LCZ data, and all of the uncertainties were involved.

Purpose: To examine the difference between the WRF output from the NUDAPT-type dataset and the WUDAPT dataset when all of the uncertainties are added up.

Case 9: BEPBEMreal.

In this case, both the BEP and BEM were turned on with the Boulac PBL scheme with NUDAPT-type data and the default setting for thermal properties in Hong Kong.

Purpose: To determine whether WRF BEPBEM (with default thermal settings) performs better than the WRF BEP alone.

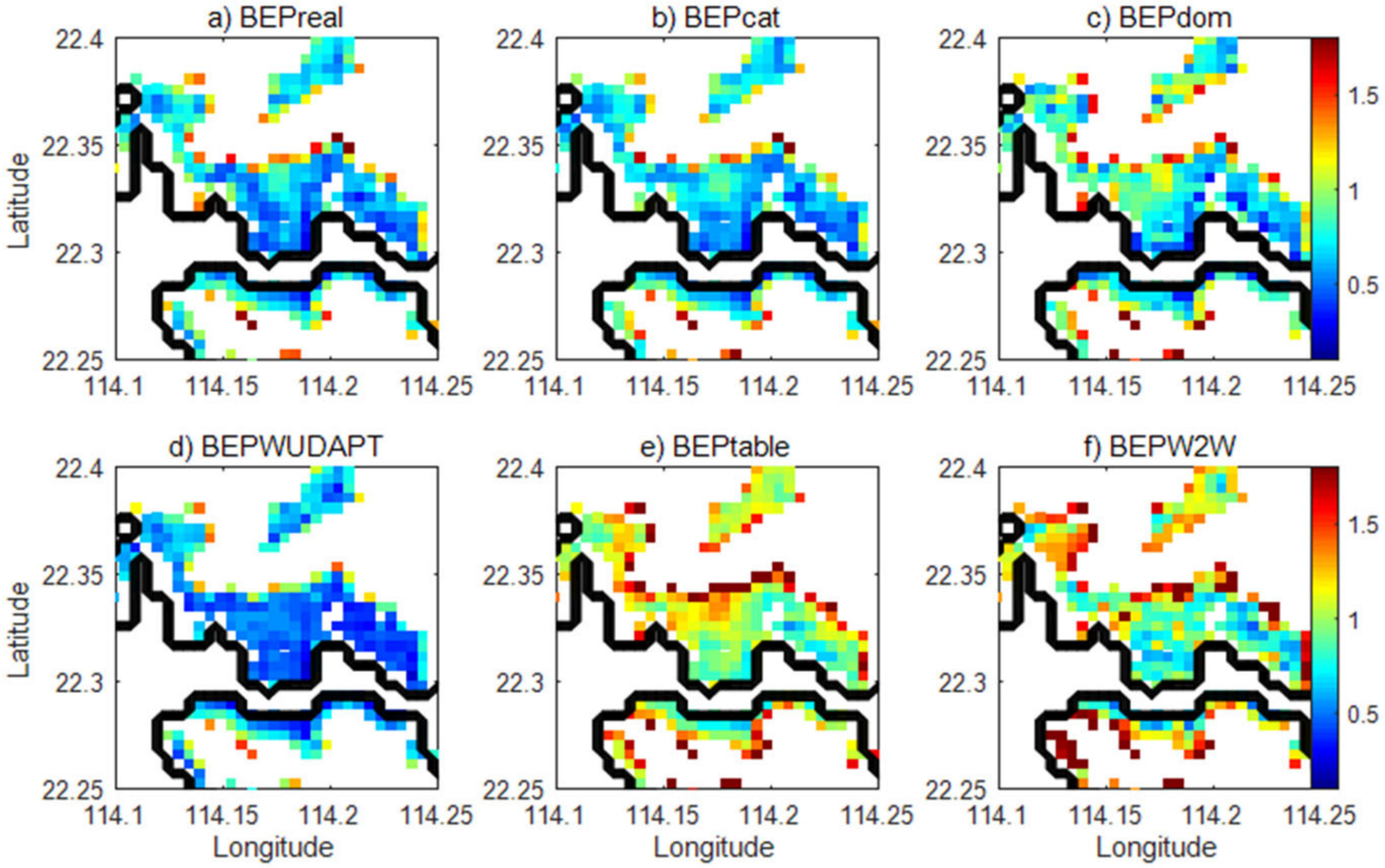


Figure 2. Corresponding averaged urban surface wind speed [m/s] of the NUDAPT case and different WUDAPT cases.

Simulation results

A. Comparison of the NUDAPT case and WUDAPT cases (Uncertainty from Different Cases).

Figure 1 and 2 show the zoomed-in (Kowloon Peninsula and Hong Kong Island) spatial distribution of input UCPs (average building height, A_h) and the corresponding output meteorological field (wind speed) for the NUDAPT case and the various WUDAPT cases. Spatially, the categorization (BEPcat) shows the closest A_h to the real case; the other cases all show greater differences. Furthermore, by using the dominant subsampling method (BEPdom) or the default W2W (Case 8, BEPW2W), the UCP input is more discrete because the dominant LCZs are considered for each modeling grid. In contrast, by using the look-up table by Brousse et al. (2016), the building height is severely underestimated because the buildings in European countries are generally shorter than those in Hong Kong. When the WUDAPT level 0 dataset is used along with the Hong Kong look-up table derived from BEPcat (BEPWUDAPT), the average building height is generally overestimated compared with the real case (mean deviation of 11m for the whole of Hong Kong), which suggests that the supervised classification algorithm tends to recognize more LCZ 1 and LCZ 4 (high-rise) categories in Hong Kong, as has been reported in other studies (Cai et al., 2016).

In Figure 2, as expected, it is visually evident that the deviation of BEPcat from BEPreal is smallest (0 m/s on average over the entire domain) regarding the magnitude and spatial pattern. This is the case because the only uncertainty involved, when compared with the real case, is the discretization of the UCPs from the NUDAPT type (continuous field) into the WUDAPT type (categories). This small difference can be related to Figures 1a and 1b in which the input UCPs for BEPcat is also very similar to the one in BEPreal (similar to other UCPs that are not shown here). The deviation of surface urban wind speed from the real case starts increasing when other sources of uncertainty become involved. The BEPdom case deviates on average 0.17 m/s from BEPreal over the whole urban area and can be up to 0.8 m/s or more in the zoomed-in region as shown in Figure 2. For example, in the highly dense Kowloon peninsula area, BEPreal estimated a surface wind speed of about 0.2 m/s, whereas BEPdom estimated around 1 m/s on average for the 3-day simulation period. This overestimation of surface wind speed suggests that the dominant LCZs approach generally leads to underestimates of the building height for the case of Hong Kong. This uncertainty arises from the considerable heterogeneity of the UCPs distribution in Hong Kong; subgrid tall building clusters can be ignored if the dominant type of buildings is relatively short. Subgrid averaging is preferable for minimizing this uncertainty because modelers can more or less control it via subsampling algorithms, compared with other sources of uncertainties.

Table 1: Averaged RMSE and BIAS for wind speed and temperature in different cases.						
Cases		BEPreal	BEPBEMreal	BEPWUDAPT	NoahACM2	NoahBoulac
WIND [m/s]	RMSE	0.94	0.97	1.09	1.54	1.6
	BIAS	-0.44	-0.25	-0.61	0.74	0.57
TEMPERATURE [°C]	RMSE	1.36	1.43	1.41	1.5	1.52
	BIAS	-0.12	-0.03	0.03	-0.31	-0.41

For the BEPWUDAPT case, Figure 2d shows that spatially there is a general underestimation of urban surface wind speed because the supervised classification causes a misrecognition of LCZs (LCZ 1 and LCZ 4), which overestimate the building height and plan area ratio and thus increase urban drag. This suggests that an improvement of the WUDAPT level 0 data might be required for high-density urban areas such as Hong Kong, which deserves attention for level 1 and level 2 data with a better algorithm or the complement of different data sources (e.g., crowdsourcing, higher resolution satellite images). Furthermore, if the default look-up table is used (BEPtable in Figure 2e), the uncertainty is greatest among all the cases because local information (a localized look-up table) is lacking. Finally, for the default BEPW2W, spatially the uncertainty is similar to that in the BEPtable case and greater than in all of the other cases. A similar conclusion is drawn from the statistics for the whole of Hong Kong (domain 5), which are not shown here.

B. Improvement of WRF-BEP/BEM Compared with the Noah Bulk Scheme

Table 1 shows different averaged statistics of the models for a comparison of the observations in the BEP cases and The Noah bulk scheme cases. The following points can be made regarding improvements in surface wind speed simulation:

1. BEP's wind speed performs better than the Noah bulk scheme in terms of RMSE (from about 1.6 m/s for NoahBoulac to 0.94 m/s for BEPreal), regardless of whether WUDAPT or NUDAPT datasets are used, and it is similar to the BEP/BEM module (RMSE of 0.97 m/s). The magnitude of bias for BEPreal (-0.44 m/s) and BEPBEMreal (-0.25 m/s) is also smaller than that of NoahACM2 (0.74 m/s) and NoahBoulac (0.57 m/s). BEPWUDAPT has a magnitude of bias (-0.61 m/s) comparable with NoahBoulac because BEPWUDAPT overestimates the plan area ratio and building height of the urban area due to the uncertainty generated by the supervised classification of WUDAPT level 0 dataset (misrecognition of LCZs).

2. ACM2 performs better than the Boulac scheme in terms of RMSE, as demonstrated by Xie et al. (2012).

3. Over urban areas, the urban scheme overrides the impact of the PBL scheme because BEPreal, BEPBEMreal, and BEPWUDAPT perform better than ACM2-Noah in terms of RMSE.

4. BEM is not necessarily better than BEP alone, probably because of the lack of urban fraction data, thermal properties, and detailed anthropogenic heating data, all of which deserve further research on the WUDAPT level 1 and level 2 dataset that is currently under active development by the WUDAPT team.

With the non-local PBL scheme ACM2, the wind speed is higher in the daytime because of a better simulation of the non-local effect. The underestimation of wind speed for uWRF runs (WUDAPT/NUDAPT cases) in the daytime suggests a future need to couple the BEP/BEM urban scheme with a non-local scheme to better simulate the convective atmosphere.

Conclusions

This study evaluates the performance of the WRF BEP/BEM model in Hong Kong, which has unique hilly topographic conditions and landscape features with a highly inhomogeneous building morphology. The results show that combined with the multi-layer WRF BEP/BEM urban scheme (uWRF), the model gives a better simulation of wind speed and a slight improvement in the simulation of temperature over the urban area, in the context of urban sites. This indicates that the choice of a multi-layer urban scheme's benefit dominates the non-local PBL scheme's benefit in determining the quality of a surface simulation over an urban area, which is crucial for urban climate studies. Furthermore, different methods of implementing WUDAPT datasets in this study quantified the different sources of uncertainty. The ascending order for the uncertainty is as follows: 1) data discretization; 2) the dominant-type subsampling method, comparable with supervised classification in WUDAPT; and 3) the lack of local information (look-up tables). A compensation error may occur when all of the uncertainties accumulate in the W2W protocol, but the error in the region with maximum uncertainties is still the largest compared with other cases. However, WUDAPT is still a suitable alternative in regions where NUDAPT-type datasets are not available, provided that building morphology for different LCZs is estimated based on local expertise (assisted by the help of 3D maps) with a subgrid-averaging approach. BEM is not necessarily better than BEP alone, probably because of the lack of urban fraction data, thermal properties, and detailed anthropogenic data. This suggests the need of level 1 and level 2 WUDAPT datasets that give more detailed input for driving the BEP model.

Further work

Regarding directions for future study, the development of level 1 and level 2 WUDAPT datasets is crucial to further minimize sources of uncertainty. This may be accomplished by incorporating more and higher resolution satellite images in the supervision classification of LCZs or by improving the machine-learning algorithm, which is an ongoing effort for the WUDAPT team, as in the work of Xu et al. (2017). Besides building morphology data, more work is also needed to arrive at better estimations of buildings' thermal properties and the urban green fraction, based on building uses. Higher resolution satellite images may also provide more adequate information for repeating the current experiments of the BEM scheme. Coupling the uWRF schemes with a non-local PBL scheme is also expected to produce a better simulation in the context of convective conditions.

Acknowledgement

The study is supported by the RGC Grants 16300715 and 14611015 respectively. It is also partially supported by The Vice-Chancellor's Discretionary Fund of The Chinese University of Hong Kong. The authors would like to thank Miss Xinwei Li and Dr. Yong Xu for their help on training sample collection and the LCZ map development. We also appreciate the assistance of the Hong Kong Observatory (HKO), which provided the meteorological data.

References

Brousse, O., Martilli, A., Foley, M., Mills, G., & Bechtel, B. (2016). WUDAPT, an efficient land use producing data tool for mesoscale models? Integration of urban LCZ in WRF over Madrid. *Urban Climate*, 17, 116-134.

Cai, M., Ren, C., Xu, Y., Dai, W., & Wang, X. M. (2016). Local Climate Zone Study for Sustainable Megacities Development by Using Improved WUDAPT Methodology–A Case Study in Guangzhou. *Procedia Environmental Sciences*, 36, 82-89.

Ching, J., Brown, M., Burian, S., Chen, F., Cionco, R., Hanna, A., ... & Williams, D. (2009). National Urban Database and Access Portal Tool (NUDAPT): Facilitating a new generation of advanced urban meteorology and climate modeling with community-based urban database system. *Bulletin of the American Meteorological Society*.

Ching, J., See, L., Mills, G., Alexander, P., Bechtel, B., Feddema, J., Oleson, K. L., Stewart, I., Neophytou, M., Chen, F., Wang, X., & Hanna., A. (2014). WUDAPT: Facilitating advanced urban canopy modeling for weather, climate and air quality applications.

Hammerberg, K., Brousse, O., Martilli, A., & Mahdavi, A. (2018). Implications of employing detailed urban canopy parameters for mesoscale climate modelling: a comparison between WUDAPT and GIS databases over Vienna, Austria. *International Journal of Climatology*, 38, e1241-e1257.

Salamanca, F., Krpo, A., Martilli, A., & Clappier, A. (2010). A new building energy model coupled with an urban canopy parameterization for urban climate simulations—part I. formulation, verification, and sensitivity analysis of the model. *Theoretical and applied climatology*, 99(3-4), 331.

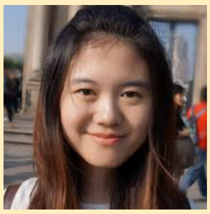
Stewart, I. D., & Oke, T. R. (2012). Local climate zones for urban temperature studies. *Bulletin of the American Meteorological Society*, 93(12), 1879-1900.

Xie, B., Fung, J. C., Chan, A., & Lau, A. (2012). Evaluation of nonlocal and local planetary boundary layer schemes in the WRF model. *Journal of Geophysical Research: Atmospheres*, 117(D12).










Michael Mau Fung Wong[1], Jimmy Chi Hung Fung[1,2], Jason Ching[3], Peter Pak Shing Yeung[1], Jason Wai Po Tse[4], Ran Wang[5], Meng Cai[5], Chao Ren[5,6]

[1] Division of Environmental and Sustainability, The Hong Kong University of Science and Technology, Hong Kong (michaelwong22mu@gmail.com); [2] Department of Mathematics (HKUST); [3] Institute for the Environment, University of North Carolina, Chapel Hill NC, USA; [4] Environmental Science Program (HKUST); [5] Institute of Future Cities, The Chinese University of Hong Kong; [6] Hong Kong Faculty of Architecture, The University of Hong Kong