Geophysical Research Letters

RESEARCH LETTER
10.1029/2021GL094737

Deep Learning-Based Super-Resolution Climate Simulator-Emulator Framework for Urban Heat Studies

Yuankai Wu[1], Bernardo Teufel[1], Laxmi Sushama[1], Stephane Belair[2], and Lijun Sun[1]

[1]Department of Civil Engineering, Trottier Institute for Sustainability in Engineering and Design, McGill University, Montreal, QC, Canada, [2]Meteorological Research Division, Science and Technology Branch, Environment and Climate Change Canada, Dorval, QC, Canada

Key Points:

- Innovative simulator-emulator framework proposed to generate super-resolution climate information
- New framework applied at city-scale demonstrates considerable skill for urban heat-related variables
- The generic nature of the framework enables adaptability to other climate variables and regions

Correspondence to:
Y. Wu,
yuankai.wu@mcgill.ca

Citation:
Wu, Y., Teufel, B., Sushama, L., Belair, S., & Sun, L. (2021). Deep learning-based super-resolution climate simulator-emulator framework for urban heat studies. *Geophysical Research Letters*, *48*, e2021GL094737. https://doi.org/10.1029/2021GL094737

Received 21 JUN 2021
Accepted 13 SEP 2021

Abstract This proof-of-concept study couples machine learning and physical modeling paradigms to develop a computationally efficient simulator-emulator framework for generating super-resolution (<250 m) urban climate information, that is required by many sectors. To this end, a regional climate model/simulator is applied over the city of Montreal, for the summers of 2019 and 2020, at 2.5 km (LR) and 250 m (HR) resolutions, which are used to train and validate the proposed super-resolution deep learning (DL) model/emulator. The DL model uses an efficient sub-pixel convolution layer to generate HR information from LR data, with adversarial training applied to improve physical consistency. The DL model reduces temperature errors significantly over urbanized areas present in the LR simulation, while also demonstrating considerable skill in capturing the magnitude and location of heat stress indicators. These results portray the value of the innovative simulator-emulator framework, that can be extended to other seasons/periods, variables and regions.

Plain Language Summary One of the major barriers in undertaking super-resolution (<250 m) urban climate simulations to generate climate and climate change information at high spatial and temporal resolutions, as required by many sectors, is their high computational cost. New approaches are therefore required to overcome this barrier. This paper makes use of the unique opportunity to couple machine learning and physical modeling to develop a computationally efficient simulator-emulator framework to generate super-resolution climate information. The trained deep neural network model generates high-resolution urban climate data from low-resolution (LR) inputs, considering also the urban morphology fields and inter-variable relationships to improve output realism. The developed framework, applied to urban heat-related variables, demonstrates the high potential of this approach as it captures well the magnitude and location of heat stress indicators. The generic nature of the developed framework makes it even more promising as it can be applied to other climate variables, periods and regions.

1. Introduction

Climate change is projected to increase the frequency and intensity of extreme weather events, including heat waves (IPCC, 2013). These projected changes in heat waves coupled with the urban heat island (UHI) effect, as manifested by elevated near-surface air temperatures in urban areas compared to their non-urban surroundings, exposes urban dwellers to additional heat stress. The higher urban temperatures are largely related to thermal and radiative properties of built surfaces, substantially different from its surrounding natural environment, and to a lesser vegetation coverage with limited evaporative cooling. UHI is further enhanced by heat emitted from transportation, heating and air conditioning systems (Oke, 1982). Adaptation and mitigation strategies considered to reduce UHI, which could reduce the impacts during heat waves, include increasing the reflectivity of urban regions and increasing vegetation coverage (Alexandri & Jones, 2008; Costanzo et al., 2016; Touchaei & Akbari, 2013). Development of effective adaptation and mitigation strategies will require information of mean and extreme temperature changes at super-resolution (<250 m). Climate change information at such super-resolution are not available for cities, primarily due to the inadequate representation of urban regions in climate models due to their coarse resolution. Climate processes are complex–more so in urban regions (Bai et al., 2018). Super-resolution climate modeling that includes good representation of urban regions is required to capture the urban-climate feedbacks. With significant developments in high performance computing, it has now become possible to execute regional climate simulations at 4 to 1 km resolutions, but for shorter periods. High computational cost continues to be

a major barrier in undertaking super-resolution simulations. A simulation at 250 m requires more than 100 times the computing resources required at 2.5 km over the same domain. Furthermore, larger ensembles of climate simulations are generally required to quantify uncertainty in climate projections, making it even more expensive. Recent advances in the area of machine learning provide an opportunity for developing hybrid approaches to overcome this obstacle and advance studies on climate-urban infrastructure interactions (Brenowitz & Bretherton, 2018; Scher, 2018).

Traditionally, physical modeling and machine learning have been treated as two different fields with very different scientific paradigms, i.e., theory-driven versus data-driven. The synergy between the two approaches has been gaining much attention recently (Reichstein et al., 2019). In recent years, deep-learning based image super-resolution models have been actively developed (Dong et al., 2015; Wang et al., 2015; Zhang et al., 2018), and they have been used to produce high-quality physical fields in various domains such as fluid dynamics (Xie et al., 2018), medical imaging (Trinh et al., 2014), and astronomical observations (Li et al., 2009). In general, the family of image super-resolution (SR) algorithms uses convolutional neural networks (CNNs) as their basic building blocks (Lai et al., 2017), and purely data-driven principle like generative adversarial nets (Ledig et al., 2017). Due to its data-driven nature, the CNN-based SR approach does not require solving complex equations. Therefore they can dramatically decrease the computational cost of generating high-resolution data once the parameters of CNNs are properly trained. Given the high computational cost for simulating high-resolution climate data, there has been a few attempts to directly apply models from image super-resolution domain to climate model outputs (Stengel et al., 2020; Vandal et al., 2017). However, the scales considered in those studies are too coarse to capture features that are critical for urban-climate feedbacks. Furthermore, multivariate dependencies and surface features have large impacts on climate, which were ignored in those studies.

This study focuses on the development, validation and application of a framework combining the physically based regional climate model GEM (Global Environmental Multi-scale) with deep learning techniques to generate super-resolution climate information efficiently. Given the importance of heat stress in urban regions, the focus is on emulating the two variables that have the most influence in determining heat stress: near-surface air temperature and dew point temperature for the summer season. A deep learning super-resolution model that can make full use of inter-variable relationships and dependencies on auxiliary fields is developed. The goal is to keep the development as generic as possible so that the framework can be extended to other seasons, variables, periods and regions.

The remainder of this manuscript is organized as follows: Section 2 describes the physical climate and machine learning models, and the experiments performed. Section 3 presents the super-resolution climate information obtained from the machine learning models and compares them to the climate model. Finally, Section 4 provides discussion and conclusions.

2. Methodology

2.1. Climate Model and Simulations

The physically based climate model considered in this study is the limited area version of the GEM model (Côté et al., 1998; Girard et al., 2014), used for numerical weather prediction at Environment and Climate Change Canada (ECCC). It employs semi-Lagrangian transport and a (quasi) fully implicit time stepping scheme. In its fully elastic nonhydrostatic formulation (Yeh et al., 2002), it uses a vertical coordinate based on hydrostatic pressure (Laprise, 1992). Condensation processes are computed by a double-moment microphysics scheme (Milbrandt & Yau, 2005). More details on the parameterizations used can be found in Diro and Sushama (2019). The land part of the model is represented using the Canadian Land Surface Scheme–CLASS (Verseghy, 1991, 2011), while the urban regions are represented by the Town Energy Balance (TEB; Masson, 2000) model. TEB is a physically based scheme for urban energy budget based on a generalization of local canyon geometry. It is a single-layer urban canopy model. Buildings are assumed to be located along identical roads leading to the canyon structure. The lengths of the roads are considered to be far greater than their widths; any road orientation is possible, and all exist with the same probability. The buildings bordering the street canyon are assumed to have the same height and width within a model grid cell. In TEB, the urban morphology for each grid cell is defined in terms of the local building and road fractions,

building height, building and canyon aspect ratio and roughness length. The radiative characteristics such as albedos and emissivity for roofs, roads and walls and thermal properties such as thermal conductivity and heat capacity of each roof, road and wall layer are prescribed.

GEM simulations for the period from April 2019 to August 2020 are performed at 2.5 km and 250 m resolutions; the experimental domains for the two resolutions are shown in Figure 1a, with the smaller 250 m resolution domain covering the city of Montreal. The predominant land cover for Montreal and surroundings is shown on Figure 1b. The super-resolution urban climate simulation at 250 m is driven at the lateral boundaries by the 2.5 km GEM simulation, which is in turn driven by ERA5 reanalysis data (Hersbach et al., 2020) from the European Centre for Medium-Range Weather Forecasts. GEM outputs for the summer months are used as the input data for the deep learning framework.

2.2. Deep Learning Framework

The task of the deep learning super resolution (SR) model is to emulate climate variables at 250 m resolution from the corresponding climate variables at 2.5 km resolution. In this work, the input is partitioned into two parts: the 2.5 km low resolution (LR) climate data I^{LR}, and the high-resolution (HR) auxiliary fields A^{HR} (land cover and urban morphology fields). The goal is to train a deep neural network $\mathcal{G}$ with parameters θ that can produce realistic HR climate variables Y^{HR}:

$$Y^{HR} = \mathcal{G}_{\theta}\left(I^{LR}, A^{HR}\right). \tag{1}$$

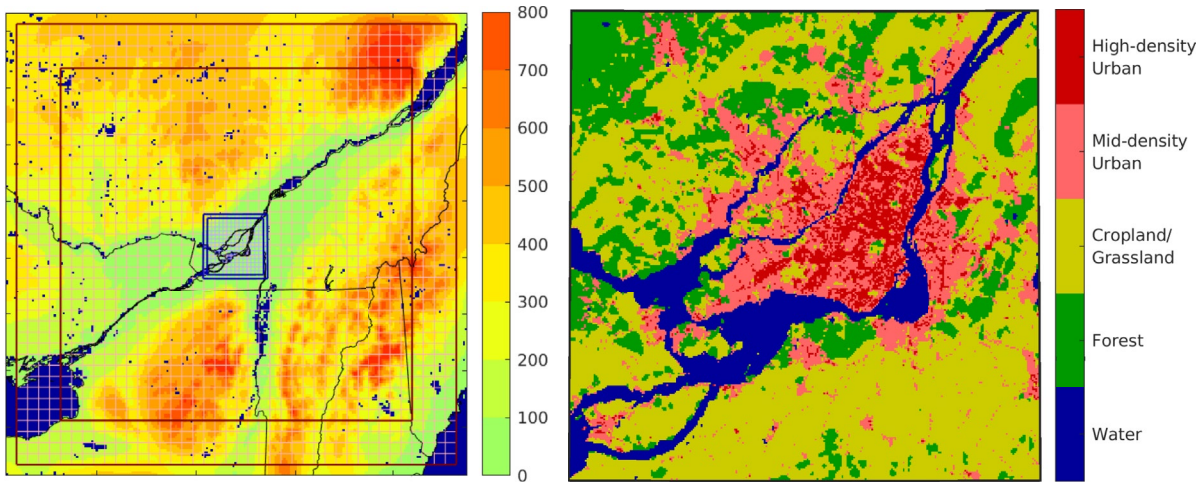
The variables considered in this study (for I^{LR} and Y^{HR}) are 2-m air temperature and dew point temperature, which are essential for estimating heat-stress indices. Twenty-one auxiliary variables, covering building characteristics, surface types, and land types of the study area constitute A^{HR}. The corresponding HR (250 m) and LR (2.5 km) domains consist of 280×280 and 28×28 pixels (grid cells), respectively. SR is considered a data generation problem and a deep learning framework based on generative adversarial network (GAN), named fused sub-pixel convolutional generative adversarial network (FSPCGAN), is proposed.

2.2.1. Neural Network Structure

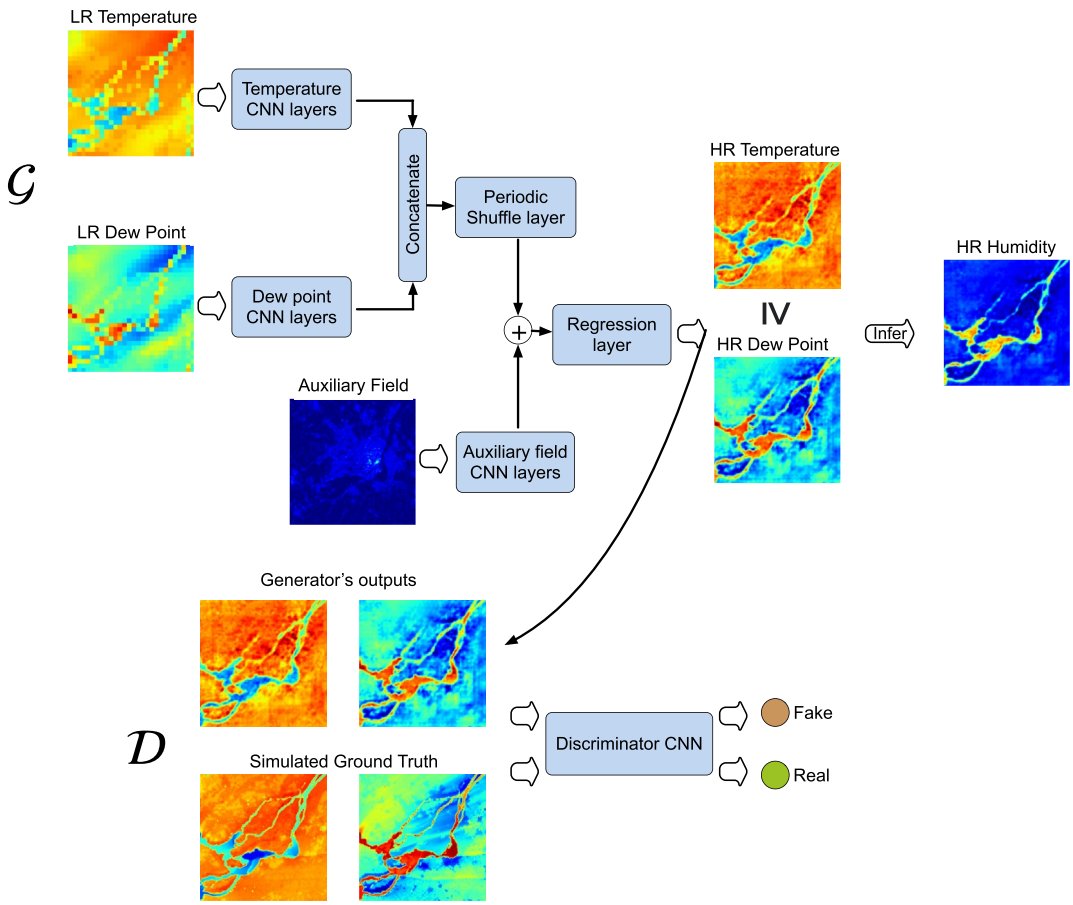
To design a neural network for climate variable at SR, a straightforward way is to use pixel-wise loss functions (errors for every grid) such as the mean squared error (MSE) of the reconstruction as an objective function. However, it has been found in general image SR studies that the pixel-wise loss function struggles to handle the high-frequency details in images—they tend to encourage the model to find an unrealistic solution with small MSE (Estrach et al., 2016; Johnson et al., 2016). Moreover, pixel-wise loss functions cannot ensure the physical consistency of climate fields (Stengel et al., 2020). Recent studies (Ledig et al., 2017; Yu & Porikli, 2016) in image SR applications have applied GANs to tackle this problem. The GAN procedure encourages reconstructions to move toward regions of the search space with high probability of containing physically realistic samples (Goodfellow et al., 2014).

Figure 1c depicts the architecture of the proposed FSPCGAN, consisting of two neural networks—a generator $\mathcal{G}$ and a discriminator $\mathcal{D}$. GANs provide a powerful scheme to map LR input fields to realistic HR field space. The generator $\mathcal{G}$ (as in Equation 1) attempts to generate synthetic HR climate variables, and the discriminator $\mathcal{D}$ attempts to determine whether the generated HR fields are real (i.e., coming from the HR simulation) or not. As both input and output are spatial matrices, CNN are used as the key component to build both $\mathcal{G}$ and $\mathcal{D}$.

For the input I^{LR}, two separate CNNs are applied to process LR temperature and dew point fields, respectively, instead of building a single CNN model to reconstruct multiple climate variables like RGB channels in image processing. Different CNN layers are used to establish multichannel flow of information and produce LR temperature and dew point features, respectively. A concatenation layer is then applied to build a joint feature containing information from both temperature and dew point fields. In order to upscale the joint LR features to HR (i.e., 10 times larger) feature space, the concatenation is fed to a periodic shuffle layer (Shi et al., 2016), which requires far less computational power than directly using the upsampling and/or deconvolution operators. CNNs are also used for the auxiliary A^{HR} fields to obtain features. The two HR features generated from I^{LR} and A^{HR} are added together, and then fed into a regression layer to get final HR output Y^{HR}, which consists of both temperature and dew point fields.



(a) GEM computational domains (b) Predominant land cover



(c) Architecture of the proposed deep learning models

Figure 1. (a) Global Environmental Multi-scale computational domain at 2.5 km (red; every fifth grid cell shown) and 250 m (blue; every twentieth grid cell shown). The outer thick lines represent the model domain, while the inner thick lines represent the model free domain. Water bodies are shown in dark blue. (b) Predominant land cover in the 250 m domain. High-density urban is defined as urban fraction exceeding two thirds coverage, while mid-density urban is defined as urban fraction between one third and two thirds coverage. (c) The architecture of the proposed deep learning model. The super-resolution (SR) model consists of a generator (upper part) and a discriminator (bottom part). The generator is used to generate high-resolution (HR) temperature and dew point from low-resolution inputs, the discriminator is used to distinguish between real HR simulations and SR model outputs.

There are several points to note in the design of generator $\mathcal{G}$. First, the periodic shuffle layer used for upscaling often suffers from the checkerboard artifact (Aitken et al., 2017), which can cause unrealistically high gradients in the north-south and east-west directions for HR outputs of climate variables. To alleviate the checkerboard artifact, the nearest neighbor strategy (Aitken et al., 2017) is adopted to initialize the periodic shuffle layer. In addition, a total variation term L_{TV} is added to the loss function to ensure spatial smoothness in adjacent pixels:

$$L_{TV} = \frac{1}{2 \times 280^2} \sum_{c,i,j} \sqrt{\left| Y^{HR}\left[c,i+1,j\right] - Y^{HR}\left[c,i,j\right] \right|^2 + \left| Y^{HR}\left[c,i,j+1\right] - Y^{HR}\left[c,i,j\right] \right|^2}. \quad (2)$$

Second, to ensure physical consistency, constraints are imposed on the generator during training to ensure that the generated dew point does not exceed 2-m temperature.

2.2.2. Training Strategy

In this work, the Wasserstein GAN (WGAN) framework (Arjovsky et al., 2017) is used, which is an improved version of original GAN. In WGAN, the discriminator network $\mathcal{D}$ is a K-Lipschitz continuous function to compute the Wasserstein distance between the real probability and the generated probability. To keep the discriminator K-Lipschitz continuous, its parameters are confined to a range $(-c, c)$. To produce more realistic temperature fields, the discriminator $\mathcal{D}$ and generator $\mathcal{G}$ are trained against each other iteratively over time. In WGAN, the training loss for discriminator is the Wasserstein distance L_W between the generated climate variables and those from HR simulation. During training, parameters of the discriminator are confined by clipping values, i.e., $[-3, 3]$. More details on the Wasserstein distance can be found in (Arjovsky et al., 2017).

For the generator $\mathcal{G}$, the final training loss is defined as

$$L_{\mathcal{G}_\theta} = L_{MSE} + \gamma L_{TV} - \lambda L_W, \quad (3)$$

where λ and γ are loss weights. L_{MSE} quantifies the total reconstruction error for both temperature and dew point fields and L_{TV} is the aforementioned total variation loss to alleviate the checkerboard artifact.

2.3. Deep Learning Experiments and Baseline

To evaluate the benefits of using auxiliary fields, adversarial training and physical constraints, four deep learning variants are evaluated in this work: (a) **SPCNN** (sub-pixel convolutional neural network), (b) **FSPCNN** (fused sub-pixel convolutional neural network), (c) **FSPCGAN-S** (S refers to separate models) and (d) **FSPCGAN** (see Figure 1c), with the first three being special cases of the proposed **FSPCGAN** model. **SPCNN** and **FSPCNN** differ from **FSPCGAN** in that they do not have the discriminator component. Furthermore, **SPCNN** does not consider auxilliary variables, while **FSPCNN** takes them into account. **FSPCGAN-S** differs from **FSPCGAN** in that two separate models are developed for temperature and dew point fields in **FSPCGAN-S**, without joint training.

The output frequency for the LR and HR GEM simulations (cf. Section 2.1) is 1 hr. Of the 3,600 outputs from April 1st, 2019 to August 29th, 2019, 3,240 samples are randomly picked as the training data set, and the remaining 360 samples from the same period are used as the validation data set. The 2928 samples from May 1st, 2020 to August 31st, 2020 are used as the test data set. To illustrate the superiority of deep learning over traditional methods, a simple linear Cokriging model is also considered. The mean absolute error (MAE) and root mean squared error (RMSE) are used to measure the reconstruction performance of different deep learning variants and the Cokriging baseline:

$$\begin{aligned} MAE &= \frac{1}{N}\sum_{n=1}^{N}\left|o_n^{HR} - y_n^{HR}\right| \\ RMSE &= \sqrt{\frac{1}{N}\sum_{n=1}^{N}\left(o_n^{HR} - y_n^{HR}\right)^2}, \end{aligned} \quad (4)$$

where o_n^{HR} denotes the nth GEM-simulated ground truth, y_n^{HR} denotes the nth SR model outputs and N is the sample size.

For the deep learning models, 3-layer CNNs are used to capture LR temperature/dew point features. The convolutional kernel of the first layer is 5×5, and the kernels of the following two layers are both 3×3.

Table 1
Average Errors and Their Standard Deviations for Temperature and Dew Point for the Last 50 Learning Steps; and for Inferred Relative Humidity for the Last Step

	Temperature (°C)		Dew point (°C)		Relative humidity (%)	
Models	MAE	RMSE	MAE	RMSE	MAE	RMSE
Cokriging	1.2818	1.8468	1.1203	1.6405	7.0178	9.6141
SPCNN	0.7900 ± 0.0206	1.0959 ± 0.0216	0.9703 ± 0.0236	1.4362 ± 0.0265	5.3736	7.3140
FSPCNN	**0.7892 ± 0.0350**	**1.0851 ± 0.0369**	**0.9674 ± 0.0187**	1.4326 ± 0.0294	5.1908	7.1213
FSPCGAN-S	0.7920 ± 0.0340	1.0901 ± 0.0364	0.9698 ± 0.0228	**1.4158 ± 0.0274**	5.6680	7.5672
FSPCGAN	0.8828 ± 0.0601	1.1996 ± 0.0625	0.9940 ± 0.0310	1.4308 ± 0.0345	**5.1496**	**7.0700**

Note. The bold values used to indicate the best performance. MAE, mean absolute error; RMSE, root mean squared error.

The number of feature maps for the three CNN layers are 128, 128, and 64, respectively. Two-layer CNNs are used to process HR auxiliary fields; the kernels of the first and second layers are 5×5 and 3×3, respectively. The number of feature maps for auxiliary CNNs are 32 and 64 for the two layers. The discriminators for FSPCGAN-S and FSPCGAN are built upon 4-layer CNNs. For the discriminators, the convolutional kernel of the first layer is 5×5, and all the three following layers use a 3×3 kernel. The number of feature maps are 128, 128, 64, and 64 for the four layers. The rectified linear unit function is used as the activation functions for all convolutional layers. The RMSprop optimization strategy is used to train SPCNN, FSPCNN and the generators of FSPCGAN-S and FSPCGAN, and the learning rate is set to 0.0001. RMSprop is also used for discriminators of FSPCGAN-S and FSPCGAN, with a learning rate of 0.00005 for training. The loss weights λ and γ in Equation 3 are set to 0.001 and 1, respectively. All the deep learning models use the same validation data set, and training is stopped if the MAEs with respect to the validation set do not improve in 100 training episodes.

3. Results

Table 1 compares the MAE and RMSE of temperature and dew point across the test set for each SR method. All deep learning methods significantly outperform the Cokriging model with respect to these two metrics. FSPCNN gives lower errors than SPCNN, which indicates that deep SR models can take advantage of the information from auxiliary fields such as building fractions and land cover to improve their SR performance. However, the FSPCGAN-S with an adversarial training scheme results in higher errors than FSPCNN and SPCNN. This is because the adversarial training strategy favors more physically consistent data and encourages the generator to insert more small-scale features into SR results (Stengel et al., 2020). Moreover, FSPCGAN gives higher errors than the other three models. It jointly produces HR temperature and dew point, and its discriminator forces the relationship between SR temperature and dew point to be consistent with training data. This constraint causes FSPCGAN to give higher pixel-wise errors since they are hard to infer from training data. The auxiliary fields information and adversarial training strategy also cause higher variance during training (the raw SPCNN model gives the lowest variances in Table. 1), which indicates that they make deep learning SR methods harder to converge. Lastly, whether deep learning models can learn the relationship between two physically related climate variables is investigated by looking at the performance with respect to relative humidity rh, which can be inferred from temperature t and dew point d as

$$rh = 100e^{\frac{cb(d-t)}{(c+t)(c+d)}} \tag{5}$$

with $b = 17.625$ and $c = 243.04$. Of the four deep learning models, only FSPCGAN jointly uses both dew point and temperature as inputs and outputs, other models process them separately. The inferred relative humidity is compared to that obtained from the GEM simulation. The resulting MAEs and RMSEs for the rh data are given in the last column of Table 1. All deep learning models still significantly outperform Cokriging on humidity data. Interestingly the pixel-wise errors of FSPCGAN for rh are lower than other models that separately process dew point and temperature. This implies that FSPCGAN can learn the physical relationship between temperature and dew point. Another finding is that FSPCGAN-S gives higher errors

than SPCNN and FSPCNN. This could be caused by the inserted small features from adversarial training. Those features potentially violate the physical relationship between dew point and temperature. Moreover, FSPCGAN avoids unrealistic (above 100%) humidity values due to imposed constraints between dew point and temperature, while other models occasionally produce unrealistic values. These results demonstrate that jointly training correlated climate variables is useful and necessary.

To qualitatively validate the results of different SR models, the temperature, dew point and inferred rh of different models are shown in Figure 2. It should be noted that the HR ground truth only follows the LR input data at the boundaries, and the HR simulation has a significant degree of freedom inside the boundaries, thus it can be significantly different from LR simulation. This makes the SR tasks more challenging than those in existing works (Stengel et al., 2020). In Figure 2, HR truths contain many small-scale features that are not reproduced by the LR simulation. The simple Cokriging method entirely smooths out those features, as expected. FSPCNN with auxiliary field information can additionally introduce more small-scale features to the outputs compared with SPCNN. Finally, the SR results from FSPCGAN-S and FSPCGAN are associated with more refined small-scale features. Results of FSPCGAN qualitatively appear to be more comparable with the HR simulation, although it sometimes gives higher pixel-wise errors than other deep learning models. The above is also confirmed by semivariogram analysis (not shown).

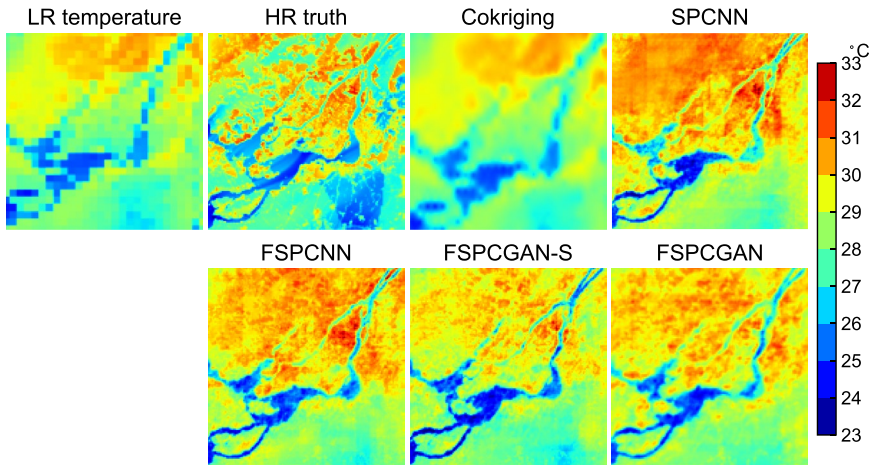
The behavior of temperature and dew point is expected to be strongly dependent on the underlying land cover (e.g., water, vegetation, urban), and higher heat stress is expected over urbanized areas. Thus, the performance of deep learning methods needs to be assessed for each of the HR grid cells (pixels). To achieve this, the RMSE with respect to the HR truth is calculated for FSPCGAN and compared to the LR data (Figure 3a). It can be seen that the LR data performs quite well (RMSE <2°C) in regions with relatively homogeneous land cover (e.g., the southeast region), while exhibiting higher errors (especially for temperature) over highly heterogeneous regions such as the Montreal metropolis and adjacent river areas. It is for these heterogeneous regions that deep learning methods such as FSPCGAN are able to produce super-resolution data much closer to the HR truth, often reducing RMSE by over 50% for temperature and around 20% for dew point. The relatively larger improvements for temperature can be explained by the fact that temperature has a stronger dependence on land cover (which is among the auxiliary fields provided to FSPCGAN) and also because of the already strong performance of the LR simulation for dew point, which leaves only little room for further improvement.

In addition to time-averaged metrics such as RMSE, assessing the performance of deep learning methods in reproducing extreme events is of utmost importance, given the large impacts that such events can entail. As this study focuses on summertime temperatures, an extreme of interest is the number of hours during which heat stress is experienced. As in previous studies (Teufel et al., 2021), a hot hour can be defined as an hour where the temperature exceeds 30°C. A potentially better approximation to heat stress can be obtained using the humidex, which combines temperature and humidity information (Davis et al., 2016; Ho et al., 2016). Figure 3b shows the performance of both the LR simulation and FSPCGAN in modeling the number of hot hours over the study region when compared to the HR truth. While some of the finest details in the HR truth are not reproduced by FSPCGAN, it clearly outperforms the LR simulation and demonstrates considerable skill in capturing both the magnitude and location of extreme temperatures.

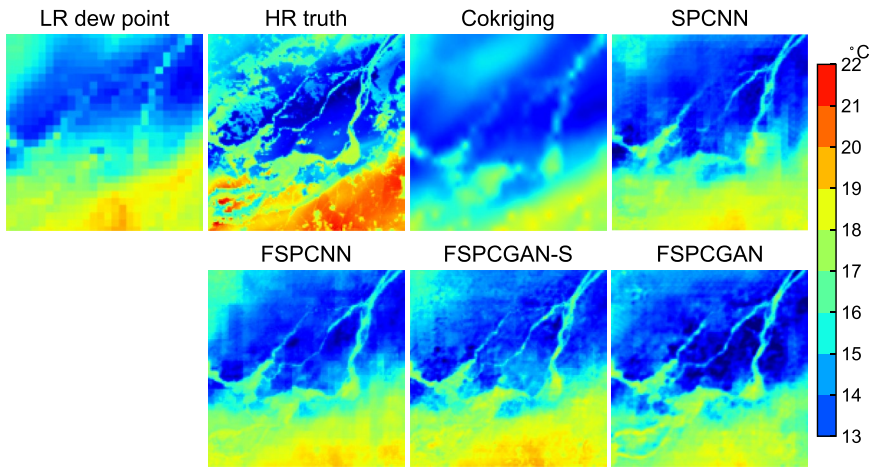
4. Discussion and Conclusions

This study demonstrates that deep learning can be a highly effective way to support numerical climate simulators in the generation of super-resolution information. A deep learning framework that increases the resolution of temperature and dew point from 2.5 km to 250 m (i.e., by a factor of 10) is presented. The deep learning framework fully considers the relationship between variables and dependencies on urban morphology fields. The model is trained on 2019 summer simulation data, and its performance is tested with respect to unseen data for the 2020 summer period. Results show that the deep learning model is capable of capturing both the magnitude and location of extreme temperatures.

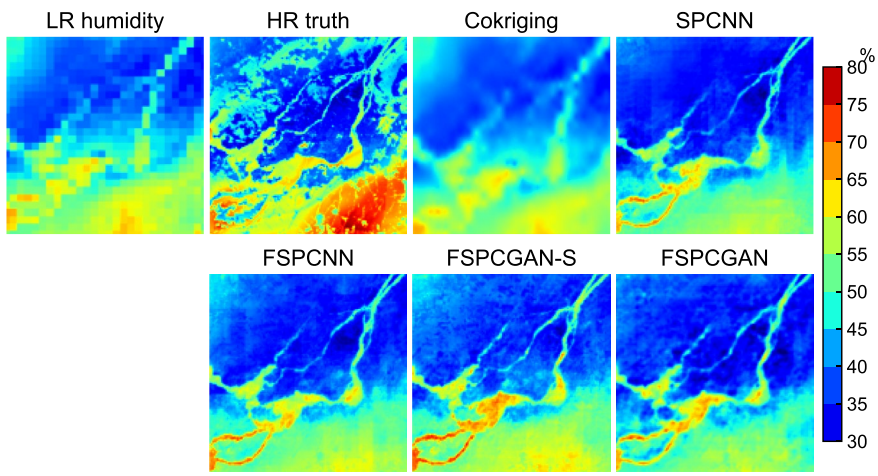
This study is to be seen as a proof of concept, in which it is shown that in principle it is possible to let a deep neural network generate high-resolution climate fields for multiple interrelated variables. One could train a deep learning model on a relatively short time period with high resolution data (250 m), following which



(a) Temperature



(b) Dew point



(c) Relative humidity

Figure 2. GEM simulated fields at low-resolution and high-resolution and super-resolution results obtained with co-kriging and deep learning models for (a) temperature, (b) dew point and (c) relative humidity at one time step.

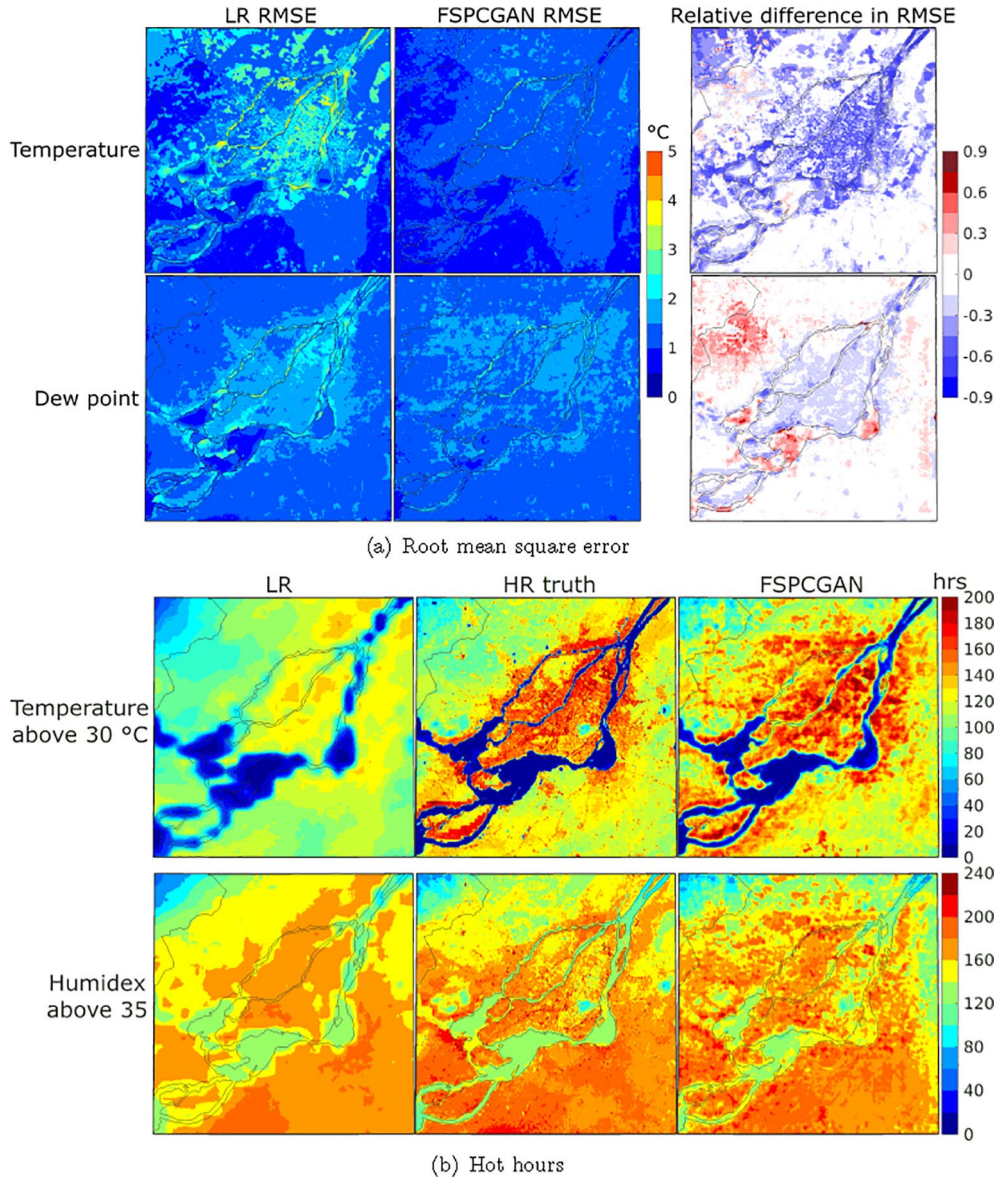


Figure 3. (a) Root mean square error (RMSE) in temperature (top row) and dew point (bottom row), with respect to the high-resolution (HR) truth for low resolution data (first column) and fused sub-pixel convolutional generative adversarial network (FSPCGAN) (second column). The relative difference in RMSE is also shown (third column). Red colors denote lower RMSE in the low resolution data, while blue colors denote lower RMSE in FSPCGAN. (b) Number of hours with temperature above 30°C (top row) and humidex above 35 (bottom row) for the May–August 2020 period, from low resolution data (first column), HR truth (second column) and FSPCGAN (third column).

the deep learning model can be fed with low resolution climate data to generate the high resolution data. Given the high flexibility of deep learning in capitalizing physical relationships across multiple variables and dependencies on urban morphology fields, the next natural step for future studies would be to apply the framework to other complex variables, such as precipitation, which involves more complex dynamics, and also for other regions and periods. Still, the principle opens up the possibility of a new type of super resolution climate simulation framework, which could also be used to refine operational weather forecasts over cities at short and medium ranges.

Data Availability Statement

All data used for training the deep learning models can be accessed at: https://zenodo.org/record/5008611#.YNDww_KSlPY.

Acknowledgments
This research was funded by the Canadian Space Agency (under grant 21SUESDFIM), Institute for Data Valorization, Natural Sciences and Engineering Research Council of Canada, Trottier Institute for Sustainability in Engineering and Design and McGill Sustainability Systems Initiative. The GEM simulations considered in this study were performed on the supercomputer managed by Compute Canada and Calcul Québec.

References

Aitken, A., Ledig, C., Theis, L., Caballero, J., Wang, Z., & Shi, W. (2017). *Checkerboard artifact free sub-pixel convolution: A note on sub-pixel convolution, resize convolution and convolution resize*. arXiv preprint arXiv:1707.02937.

Alexandri, E., & Jones, P. (2008). Temperature decreases in an urban canyon due to green walls and green roofs in diverse climates. *Building and Environment*, *43*(4), 480–493. https://doi.org/10.1016/j.buildenv.2006.10.055

Arjovsky, M., Chintala, S., & Bottou, L. (2017). Wasserstein generative adversarial networks. In *Proceedings of the 34th international conference on machine learning-volume*. 70 (pp. 214–223).

Bai, X., Dawson, R. J., Ürge-Vorsatz, D., Delgado, G. C., Barau, A. S., Dhakal, S., et al. (2018). *Six research priorities for cities and climate change*. Nature Publishing Group.

Brenowitz, N. D., & Bretherton, C. S. (2018). Prognostic validation of a neural network unified physics parameterization. *Geophysical Research Letters*, *45*, 6289–6298. https://doi.org/10.1029/2018gl078510

Costanzo, V., Evola, G., & Marletta, L. (2016). Energy savings in buildings or uhi mitigation? comparison between green roofs and cool roofs. *Energy and Buildings*, *114*, 247–255. https://doi.org/10.1016/j.enbuild.2015.04.053

Côté, J., Gravel, S., Méthot, A., Patoine, A., Roch, M., & Staniforth, A. (1998). The operational cmc–mrb global environmental multiscale (gem) model. Part I: Design considerations and formulation. *Monthly Weather Review*, *126*(6), 1373–1395. https://doi.org/10.1175/1520-0493(1998)126<1373:TOCMGE>2.0.CO;2

Davis, R. E., McGregor, G. R., & Enfield, K. B. (2016). Humidity: A review and primer on atmospheric moisture and human health. *Environmental Research*, *144*, 106–116. https://doi.org/10.1016/j.envres.2015.10.014

Diro, G. T., & Sushama, L. (2019). Simulating Canadian arctic climate at convection-permitting resolution. *Atmosphere*, *10*(8). https://doi.org/10.3390/atmos10080430

Dong, C., Loy, C. C., He, K., & Tang, X. (2015). Image super-resolution using deep convolutional networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *38*(2), 295–307.

Estrach, J. B., Sprechmann, P., & LeCun, Y. (2016). Super-resolution with deep convolutional sufficient statistics. In *4th international conference on learning representations iclr 2016*.

Girard, C., Plante, A., Desgagné, M., McTaggart-Cowan, R., Côté, J., Charron, M., et al. (2014). Staggered vertical discretization of the Canadian environmental multiscale (gem) model using a coordinate of the log-hydrostatic-pressure type. *Monthly Weather Review*, *142*, 1183–1196. https://doi.org/10.1175/mwr-d-13-00255.1

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., & Bengio, Y. (2014). Generative adversarial nets. In *Advances in neural information processing systems* (pp. 2672–2680).

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The era5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. https://doi.org/10.1002/qj.3803

Ho, H. C., Knudby, A., Xu, Y., Hodul, M., & Aminipouri, M. (2016). A comparison of urban heat islands mapped using skin temperature, air temperature, and apparent temperature (humidex), for the greater Vancouver area. *Science of the Total Environment*, *544*, 929–938. https://doi.org/10.1016/j.scitotenv.2015.12.021

IPCC. (2013). Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. In Stocker, T. F., et al. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535. https://doi.org/10.1017/CBO9781107415324

Johnson, J., Alahi, A., & Fei-Fei, L. (2016). Perceptual losses for real-time style transfer and super-resolution. In *European conference on computer vision* (pp. 694–711). https://doi.org/10.1007/978-3-319-46475-6_43

Lai, W.-S., Huang, J.-B., Ahuja, N., & Yang, M.-H. (2017). Deep laplacian pyramid networks for fast and accurate super-resolution In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 624–632). https://doi.org/10.1109/cvpr.2017.618

Laprise, R. (1992). The Euler equations of motion with hydrostatic pressure as an independent variable. *Monthly Weather Review*, *120*(1), 197–207. https://doi.org/10.1175/1520-0493(1992)120<0197:TEEOMW>2.0.CO;2

Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., & Acosta, A. (2017). Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the ieee conference on computer vision and pattern recognition* (pp. 4681–4690).

Li, L., Yu, Q., Yuan, Y., Shang, Y., Lu, H., & Sun, X. (2009). Super-resolution reconstruction and higher-degree function deformation model based matching for chang'e-1 lunar images. *Science in China - Series E: Technological Sciences*, *52*(12), 3468–3476. https://doi.org/10.1007/s11431-009-0334-7

Masson, V. (2000). A physically-based scheme for the urban energy budget in atmospheric models. *Boundary-Layer Meteorology*, *94*(3), 357–397. https://doi.org/10.1023/a:1002463829265

Milbrandt, J., & Yau, M. (2005). A multimoment bulk microphysics parameterization. part ii: A proposed three-moment closure and scheme description. *Journal of the Atmospheric Sciences*, *62*(9), 3065–3081. https://doi.org/10.1175/jas3535.1

Oke, T. R. (1982). The energetic basis of the urban heat island. *Quarterly Journal of the Royal Meteorological Society*, *108*(455), 1–24. https://doi.org/10.1002/qj.49710845502

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., et al. (2019). Deep learning and process understanding for data-driven earth system science. *Nature*, *566*(7743), 195–204. https://doi.org/10.1038/s41586-019-0912-1

Scher, S. (2018). Toward data-driven weather and climate forecasting: Approximating a simple general circulation model with deep learning. *Geophysical Research Letters*, *45*, 12–616. https://doi.org/10.1029/2018gl080704

Shi, W., Caballero, J., Huszár, F., Totz, J., Aitken, A. P., Bishop, R., & Wang, Z. (2016). Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1874–1883). https://doi.org/10.1109/cvpr.2016.207

Stengel, K., Glaws, A., Hettinger, D., & King, R. N. (2020). Adversarial super-resolution of climatological wind and solar data. *Proceedings of the National Academy of Sciences*, *117*(29), 16805–16815. https://doi.org/10.1073/pnas.1918964117

Teufel, B., Sushama, L., Poitras, V., Dukhan, T., Bélair, S., Miranda-Moreno, L., & Bitsuamlak, G. (2021). Impact of covid-19-related traffic slowdown on urban heat characteristics. *Atmosphere*, *12*(2). https://doi.org/10.3390/atmos12020243

Touchaei, A. G., & Akbari, H. (2013). The climate effects of increasing the albedo of roofs in a cold region. *Advances in Building Energy Research*, *7*(2), 186–191. https://doi.org/10.1080/17512549.2013.865558

Trinh, D.-H., Luong, M., Dibos, F., Rocchisani, J.-M., Pham, C.-D., & Nguyen, T. Q. (2014). Novel example-based method for super-resolution and denoising of medical images. *IEEE Transactions on Image Processing*, *23*(4), 1882–1895. https://doi.org/10.1109/tip.2014.2308422

Vandal, T., Kodra, E., Ganguly, S., Michaelis, A., Nemani, R., & Ganguly, A. R. (2017). Deepsd: Generating high resolution climate change projections through single image super-resolution. In *Proceedings of the 23rd acm sigkdd international conference on knowledge discovery and data mining* (pp. 1663–1672).

Verseghy, D. L. (1991). Class—A Canadian land surface scheme for gcms. I. soil model. *International Journal of Climatology*, *11*(2), 111–133.

Verseghy, D. L. (2011). *Class–the Canadian land surface scheme version 3.5 technical documentation (version 1)*. Environment Canada.

Wang, Z., Liu, D., Yang, J., Han, W., & Huang, T. (2015). Deep networks for image super-resolution with sparse prior. In *Proceedings of the IEEE international conference on computer vision* (pp. 370–378). https://doi.org/10.1109/iccv.2015.50

Xie, Y., Franz, E., Chu, M., & Thuerey, N. (2018). Tempogan: A temporally coherent, volumetric gan for super-resolution fluid flow. *ACM Transactions on Graphics*, *37*(4), 1–15. https://doi.org/10.1145/3197517.3201304

Yeh, K.-S., Côté, J., Gravel, S., Méthot, A., Patoine, A., Roch, M., & Staniforth, A. (2002). The cmc–mrb global environmental multiscale (gem) model. Part iii: Nonhydrostatic formulation. *Monthly Weather Review*, *130*(2), 339–356. https://doi.org/10.1175/1520-0493(2002)130<0339:TCMGEM>2.0.CO;2

Yu, X., & Porikli, F. (2016). Ultra-resolving face images by discriminative generative networks. In *European conference on computer vision* (pp. 318–333). https://doi.org/10.1007/978-3-319-46454-1_20

Zhang, Y., Tian, Y., Kong, Y., Zhong, B., & Fu, Y. (2018). Residual dense network for image super-resolution. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2472–2481). https://doi.org/10.1109/cvpr.2018.00262