DeepUrbanDownscale: A physics informed deep learning framework for high-resolution urban surface temperature estimation via 3D point clouds

Linwei Chen [a], Bowen Fang [c], Lei Zhao [c], Yu Zang [a,*], Weiquan Liu [a], Yiping Chen [a], Cheng Wang [a], Jonathan Li [a,b]

[a] *School of Information Science and Technology, Xiamen University, Xiamen 361005, China*
[b] *Department of Geography and Environmental Management, University of Waterloo, Waterloo, ON N2L 3G1, Canada*
[c] *Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, USA*

ARTICLE INFO

ABSTRACT

Accurate high-resolution downscaling of surface climate variables (such as surface temperature) over urban areas has long been a critical yet unresolved research problem in the field of urban climate and environmental sciences. In this paper, we propose a novel physics informed neural network (PINN) based framework: Deep-UrbanDownscale (DUD) for high-resolution urban surface temperature estimation. Anchored in process-based modeling and satellite remote sensing, the DUD network leverages the high-precision 3D point clouds to achieve accurate urban land surface temperature (LST) estimation at an ultra-high spatial resolution. This network, ingesting the high-precision land surface geometry information derived from 3D point clouds and guided by the atmospheric physics related to surface temperature, constructs a physics informed data-driven framework to fit high-resolution temperature distribution, which is otherwise difficult to be obtained by physical (numerical) simulations or traditional machine learning. Specifically, the proposed DUD network contains two branches: The Global Feature Perception (GPFP) branch and Local Urban Surface Perception (LUSP) branch. The former considers the broader-scale urban physical parameters, constraining the estimation results in accordance with the relevant physical laws. The latter, by employing a proposed local spatial coefficient index (LSCI), which is based on 3D point clouds, the estimation performance is further improved at a very high resolution. Results from designed experiments demonstrate that the proposed DUD network predicts the urban LST on a 30-by-30 m grid with the estimated error less than 0.2 Kelvin compared to the satellite measurement, which is well below the errors of other traditional methods.

1. Introduction

Cities, as the hotspots of concentrated population and infrastructure, are where major climate-driven impacts occur (Grimm et al., 2008; Mora et al., 2017). Effective urban planning and infrastructure-based growth strategies rely on high-resolution and high-precision urban climate predictions (Matei Georgescu et al., 2014; Krayenhoff et al., 2018). The urban land surface temperature (LST) is one of the most critical climate variables that is of great public interest, as it ties closely to the concerns of urban climate change, public health, city energy planning, infrastructure security, and system resilience to the climate extreme events. For example, excessive high temperature in cities would cause a substantial increase in human mortality/morbidity (Anderson and Bell, 2011; Huang et al., 2011; Patz et al., 2005), energy demand and power

grid failure (Isaac and van Vuuren, 2009; Perera et al., 2020), and a large reduction in workplace productivity (Dunne et al., 2013). Urban LST is also a critical factor that determines the aerodynamic, biophysical, and surface energy balance processes occurring at the urban land – atmosphere interface (Zhao et al., 2014), which affects virtually all the engineered systems and the residents' everyday lives in cites.

To achieve the goal of predicting high-resolution urban LST, downscaling is one of most widely used techniques which describes the family of methods or algorithms to generate high-resolution estimates based on coarse-resolution input data and some other auxiliary data. Such efforts, however, have been extremely challenging, especially for the urban surface temperature estimation, because of the large spatiotemporal variability in urban LST caused by the complex urban surface heterogeneity and the ever-changing atmospheric state (Gaffin et al., 2008).

Existing downscaling methods for urban LST can be categorized into two major groups: dynamic downscaling and statistical downscaling, each of them facing its own challenges in urban settings.

Dynamic downscaling leverages process-based dynamic models that aim to resolve the physical processes within the urban atmospheric boundary layer to solve for the temperature field. These models are highly computational expensive and can hardly operate at a high spatial resolution, limited by the physics represented in the model and the availability of urban surface characteristics dataset. For example, the dynamic downscaling using the widely-used mesoscale model – the Weather Research and Forecast (WRF) model (Kusaka et al., 2001) – are usually conducted at 1–2 km resolution at the finest (Moustaoui Georgescu et al., 2013; Krayenhoff et al., 2018); whereas using the computational fluid dynamics (CFD) based models can only be limited to very small domains such as a single urban block, street canyon, or neighborhood (Gromke et al., 2015; Middel et al., 2015). Their prediction accuracy is further subjected to the accuracy of the parameterization and representation of the physical processes.

Statistical downscaling, on the other hand, seeks to establish empirical relationships between LST and the auxiliary data such as land cover, vegetation indices, and/or other observational data (Bonafoni and Stefania, 2016; Keramitsoglou et al., 2013; Kustas et al., 2003; Zaksek and Ostir, 2012). Empirical at their core, the traditional statistical downscaling methods have been limited by: (i) the complexity of statistical methods used, (ii) the availability and reliability of the observed records, (iii) the relatively arbitrary choices of the features (predictors), and (iv) the lack of physics represented in the statistical models (Fowler et al., 2007; Pu, 2021; Spak et al., 2007; Tang et al., 2016; Wei et al., 2021). These barriers have significantly limited the traditional downscaling from generalization both spatially (i.e. up-scaled to a larger region or applied to other study locations) and temporally (i.e. future forecast).

Because of the challenges described above, high-resolution, high-precision, and computational-efficient prediction of urban LST remains a critical yet largely unresolved research gap. Recent efforts have started to explore the applications of both deep neural network and physics informed machine learning to tackle Earth and environmental science problems. These applications pointed to some potentially promising avenues to address the aforementioned downscaling challenges. Specifically, (Karpatne et al., 2017) modeled the lake temperatures across depth and over time by combining physics-based models and deep learning methods. (Ham et al., 2019) leveraged the advantages of convolutional neural network (CNN) in processing of multi-channel images (He et al., 2016; Krizhevsky et al., 2017; Ronneberger et al., 2015), applying CNN to sea surface temperatures map and the oceanic heat content map. Importantly, (Zhao et al., 2020) recently developed an urban climate emulator to predict the citywide average temperatures on the global scale using the physics informed machine learning paradigm, further demonstrating the large potential of incorporating physical mechanisms into machine learning for the urban temperature prediction. These studies have shed light on a promising direction of using physics informed neural network (PINN) to achieve the intra-city urban LST prediction at a higher spatial resolution.

In this study, we propose a novel PINN-based framework – DeepUrbanDownscale (DUD) – to downscale the urban LST. Anchored in process-based modeling and satellite remote sensing, the DUD network leverages the high-precision 3D point clouds to achieve accurate urban land surface temperature (LST) predication at an ultra-high spatial resolution. Fig. 2 shows the overview of the DUD network. By ingesting the atmospheric forcing fields and the high precision land surface geometry
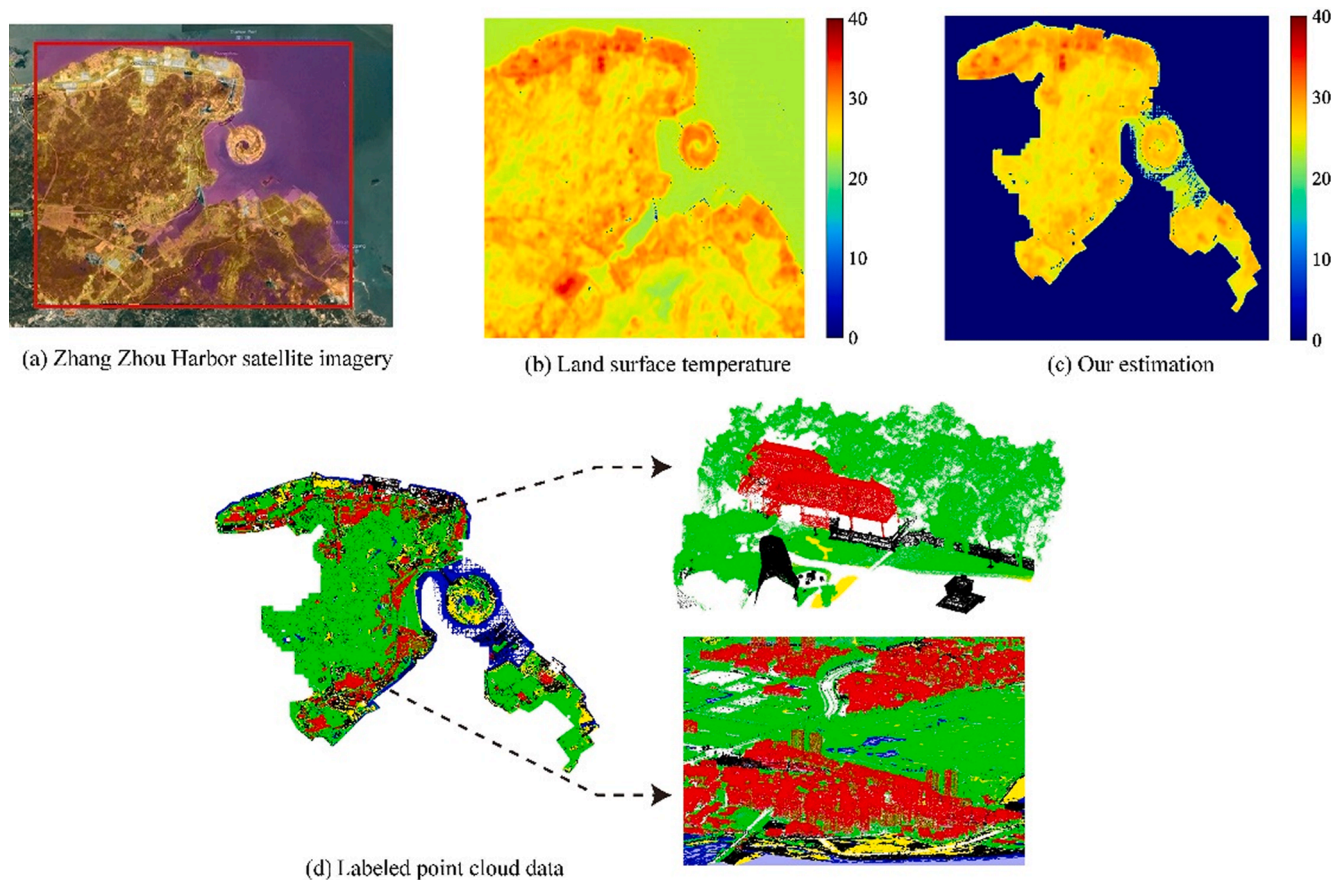

Fig. 1. Overview of the datasets and the results. (a) Testing area of Zhang Zhou Harbor. (b) Visualized map of the land surface temperature (LST) captured by Landsat satellite. (c) Our estimation results. (d) Visualized map of the labeled 3D point cloud.
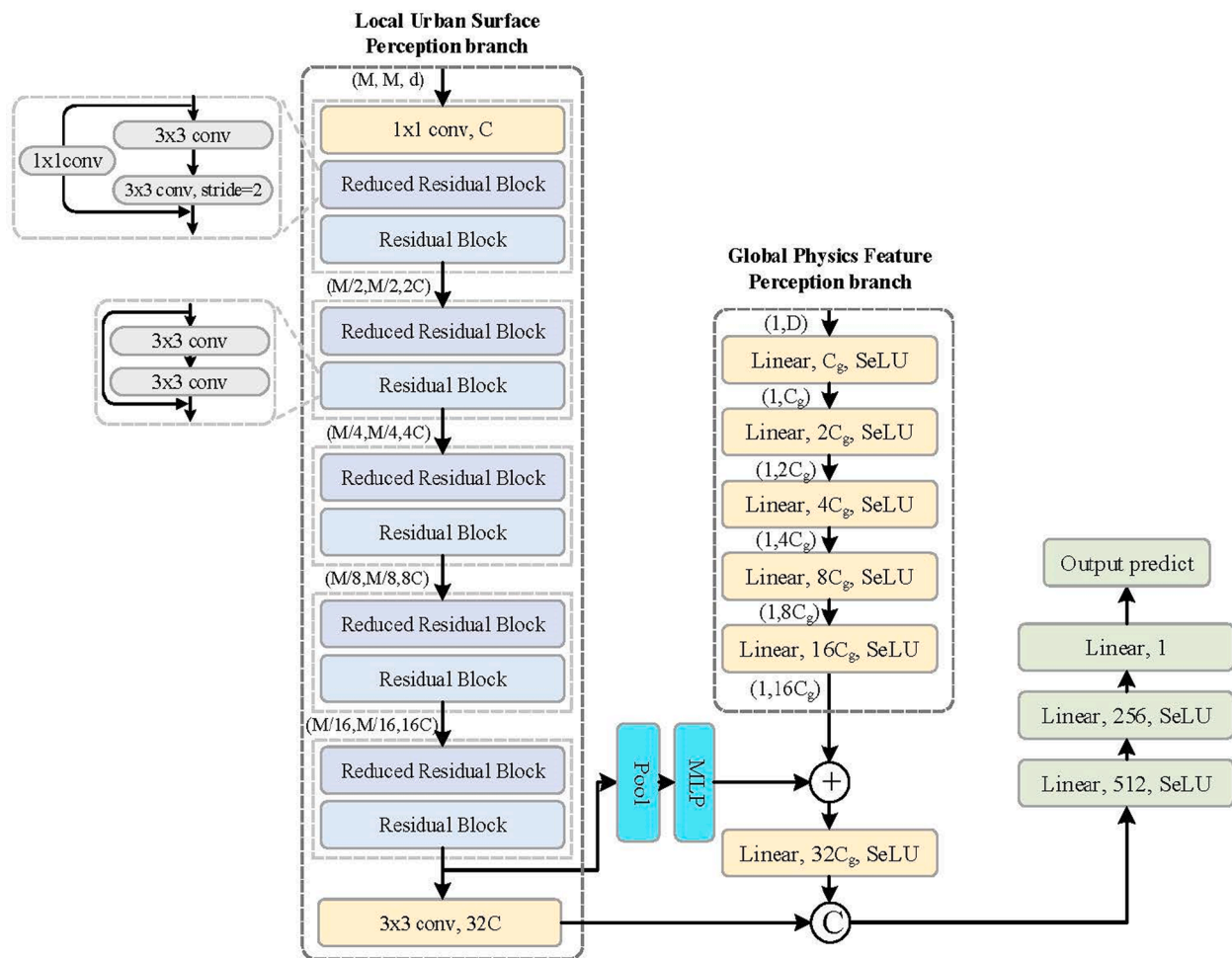
Fig. 2. Architecture of the proposed DeepUrbanDownscale system.

information, the proposed DUD network preserves not only the relevant physics represented in a typical dynamic downscaling model, but also the very high-resolution urban surface properties, which make the whole network both accurate at very fine resolutions and generalizable at large spatial scales. For evaluation, we have compared our DUD network with several traditional downscaling methods, we achieved higher spatial resolution (from 1000 m to 30 m) and lower estimation error (less than 0.2 Kelvin), and the results demonstrate that the proposed DUD network outperforms the state-of-the-art downscaling algorithms.

2. Rationale and method

2.1. Rationale

An important rationale of the proposed algorithm can be ultimately traced back to the physical processes to determine the urban surface temperature which are also the key processes. At the urban land – atmosphere interface, the urban surface temperature (i.e. LST) is governed by the surface energy balance in which several key radiative and heat transfer processes are involved. These processes include incoming and reflected shortwave (solar) radiation, longwave radiation feedback between urban land surface and the lower atmosphere, heat convection between land and the lower atmosphere, evapotranspiration from the previous surfaces in cities (such as vegetation, soil and parks), and the stored heat in buildings and other artificial materials. These processes are further determined by the atmospheric forcing state (variables) at a forcing height (usually the bottom atmospheric level in a dynamic

model) and the surface characteristics. In the process-based dynamic models, the land component takes the atmospheric forcing variables and the urban surface dataset (specifying the urban land surface properties) as input to solve the physical equations for urban LST. However, this usually can only be done for a relatively coarse spatial resolution, because at very high resolutions these physical processes are too complex to resolve due to highly complex urban surfaces. Even if a model of high complexity were available to resolve the high-detailed physics, it is almost infeasible to apply to a large-scale domain due to the unattainable computational burden.

Based on the considerations described above, the rationale of this study can be summarized as:

- For the underlying surface such as urban area, the dynamic physical process can be very complex, thus make it almost impossible to estimate the accurate LST by resolving the detailed physics.
- This work, for the first time, attempts to explore the underlying relationship between the LST and all key factors in the dynamic physical processes, by utilizing a constructed physics informed deep neural network – DeepUrbanDownscale (DUD).

These rationale further implies the necessary datasets (see Section 3.1).

2.2. DUD network

2.2.1. Overview

The urban LST is controlled by the biophysical processes associated

with the urban surface energy balance including surface radiation balance (both shortwave and longwave), turbulent transport of sensible heat and surface evapotranspiration (Zhao et al., 2014). These factors largely depend on (i) the atmospheric forcing fields such as incoming solar (shortwave) and atmospheric (longwave) radiation, air temperature at the forcing height, wind, air pressure and humidity, and (ii) the urban local surface properties such as the greenness fraction, aerodynamic roughness of the surface, building heights, and the configuration of the buildings and streets. The broader-scale urban surface climate (for example, citywide average temperature) is primarily controlled by the atmospheric forcing (i.e. meteorology) (Zhao et al., 2020), whereas the local-scale spatial variation (i.e. high-resolution details) is mostly determined by the local urban surface features such as the perviousness, greenness (vegetation), roughness and the 3D man-made structures (Gromke et al., 2015). Guided by these physical principles, we design the overall architecture of the DUD network as two component branches: global physics feature perception (GPFP) branch and local urban surface perception (LUSP) branch, as shown in Fig. 2.

The GPFP branch is mainly based on a Multi-Layer perceptron (MLP) to encode the global physical forcing factors which determine the broader-scale temperature distribution. This branch is designed to constrain the estimation results in accordance with the forcing meteorology and climatology (for example, synoptic states, changing climate or seasonality), but does not support accurate temperature estimation at a higher resolution. The LUSP branch, therefore, acts as a localization branch to fit the high-resolution variabilities. One key innovation in the LUSP branch is that it utilizes the high-precision urban 3D point cloud. Previous LST downscaling methods only use the 2D information of the surface properties such as the satellite-measured normalized difference vegetation index (NDVI) (Kustas et al., 2003). However, it has been shown that the 3D geometry of the surface (such as the surface roughness) plays a critical role in determining the urban surface temperature through affecting the efficiency of the turbulent heat transport (Zhao et al., 2014). This verticality has never been incorporated in the traditional downscaling methods. This shortcoming is overcome in our network in which a local spatial coefficient index (LSCI) derived from the urban 3D point cloud model is introduced. Details about the two branches are described below.

2.2.2. Global physics feature perception branch (GPFP)

The GPFP branch aims to embed primary atmospheric forcing variables that are used in the process-based urban climate models (Kusaka et al., 2001; Oleson et al., 2008) into the deep neural network. Specifically, a list of these features is shown in Table 1, which contain the main components of the atmospheric forcing data. The atmospheric forcing data include much physical dimension information, such as longwave radiation, shortwave land radiation and surface pressure etc. Their will bring global physics guide to our method. This branch takes all the required forcing variables in the process-based models, in a way to mimic the dynamic simulations. Therefore, this branch can be considered as a deep learning "solver" of the physical equations in those process-based models. Here a multi-layer perceptron (MLP) is designed

to utilize these forcing variables which are provided by the MERRA-2 reanalysis dataset (Gelaro et al., 2017) in this study. For the MLP, we use five layers to encode the feature vector into multiple dimensions, the first layer maps atmospheric forcing date from $1 \times D$ to $1 \times C_g$, D represent the number of atmospheric forcing data property, shown in Table 1. The remainder of layers are used to encode the output of first layer into multiple dimensions $[C_g - 2C_g - 4C_g - 8C_g - 16C_g]$. Then, we concatenate the intermediate output from LUSP branch, applying the last layer to produce a final latent vector, with size $1 \times 32C_g$. After each full-connected layer, the SeLU (Klambauer et al., 2017) is applied to avoid exploding and vanishing gradients. This GPFP branch is designed to confine the predicted coarse-resolution temperatures in accordance with the driving atmosphere. The high-resolution variability will be refined by the local surface perception branch described in detail below.

2.2.3. Local urban surface perception branch (LUSP)

Local Spatial Coefficient Index. To introduce the local geometric information of the 3D man-made structures into the neural network, one straightforward way is to directly import the entire 3D point cloud of a local region to the network by a point-based network, such as (Meng et al., 2021; Qi et al., 2017; Thomas et al., 2019; Xu et al., 2020; Zhang et al., 2020). However, we find that this approach can hardly capture the general relationship between the regional average temperature and the local geometric information because of the high complexity of the urban surface features that lead to poor generalizability of the whole system. Therefore, we instead design a descriptor to aggregate the potential factors that influence the local-scale urban surface temperature. This descriptor is denoted as the local spatial coefficient index (LSCI) thereafter.

The urban LST data is obtained from NASA's Landsat satellite measurement, which is 30-by-30 m in spatial resolution shown in Fig. 1**(b)**. We further divide each 30-by-30 m grid into m $\times$ m square cells (that is 10×10 in this study). Each cell is described by a high dimensional index which implies the potential surface characteristics that affect the local temperature. The index consists of two major components: the surface property index and the local geometry index. Each component is calculated via the local 3D point cloud within each cell as described below.

The urban surface property index is defined by the proportion of different semantic structures of a region. In our approach, five major temperature related structure categories (include water, building, vegetation, soil and road/pavement) are labeled by previous semantic labeling works (Hackel et al., 2017), as shown in Fig. 3**(a)**, various colors in the point cloud denote different urban surface properties. The statistics of each structure category is aggregated to form the urban surface property index.

The local geometry index is defined by the average height of a certain cell multiplied by the urban building index to abstract the spatial configuration of the buildings, the surface roughness, and the verticality which affect the local atmospheric turbulence over the urban surfaces.

Specifically, the whole LSCI is defined by the following equation:

$$I_{des}(l) = \frac{C(S_l)}{C(S)}, S \in T.\rho \tag{1}$$

$$H_{des}(l_b) = I_{des}(l_b) \cdot avgz(S_{lb}) \tag{2}$$

where *S* represents point cloud set centered at a certain cell, $C(\cdot)$ denotes the number of points in *S*, S_l denotes the number of points in *S* with certain category l, l_b denote the label of building, and $avgz(\cdot)$ calculates the average height of all points in *S*. $T.\rho$ represents all point cloud set (see Section 3.1)

Fig. 3 (b) and (c) show the example of constructed LSCI. For each 30-by-30 m grid, a m $\times$ m $\times$ d standard matrix (d denotes the dimension of the feature vector at a certain cell) is constructed to define the local spatial coefficient index (LSCI).

Table 1
Atmospheric and location data.

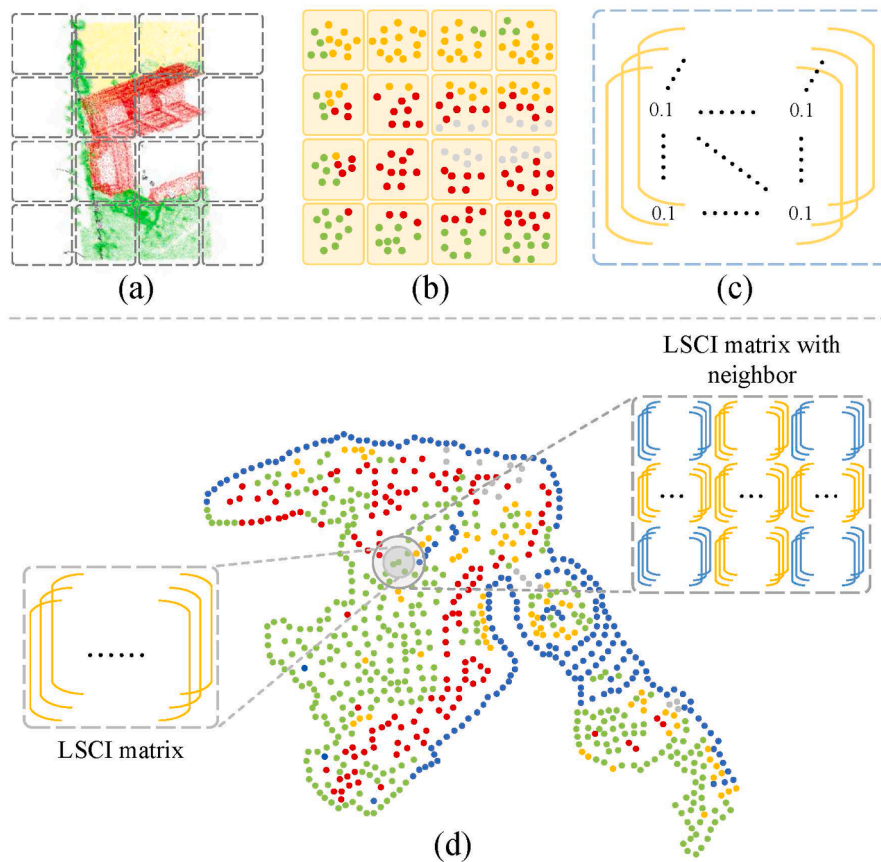|	Feature information	Unit
1	Surface absorbed longwave radiation	$W \cdot m^{-2}$
2	Incident shortwave land	$W \cdot m^{-2}$
3	Total precipitation over land	$kg \cdot m^{-2} \cdot s^{-1}$
4	Surface pressure	Pa
5	Air temperature at the reference height	K
6	Eastward wind	m^{-1}
7	Northward wind	m^{-1}
8	Specific humidity	$kg \cdot kg^{-1}$
9	Longitude	km
10	Latitude	km

Fig. 3. The illustration how to generate the LSCI. (a) An example of the labeled point cloud for a 30-by-30 m grid. (b) Visualized result of different categories in a certain cell. (c) Example of the LSCI matrix with size m × m × d. (d) Illustration of the aggregation for a grid.

The LSCI was designed with the following considerations: i) LSCI could introduce a richer physical information, give more physical orientation to the network and help network better discover hidden physical laws. In this case some environment-related physical variables (such as NDVI) was to achieve this purpose we borrowed the design of. ii) LSCI could make full use of the relative location information within the region (the distribution of surface properties between cities has a greater impact on the overall surface temperature-distribution (Alexander, 2020, 2021)), we divided the point cloud of each large region into smaller cells, and then parameterize each cell, which could have a better use of the CNN's ability to process the information and extract the corresponding features. iii) This design can greatly reduce the computational effort, making it possible to process the neighborhood information (since the point cloud has a huge amount of data, processing a region and its surrounding areas simultaneously will lead to an order of magnitude increase in computation).

Local urban surface perception branch. The LUSP branch aims to extract the local surface features via the proposed LSCI that embeds the local 3D geometric information of the urban surface in the descriptor. This branch enables the neural network to capture the high-resolution variability of the urban surface temperature (as demonstrated in Section 3.2). The LUSP branch is a 5-stage deep residual network (He et al., 2016), where each stage contains two residual blocks. The first block is composed sequentially by two 3 × 3 convolution layers. The stride is set as 2 and 1 respectively, and a skip connection is used to align output shape. The second block consists of two convolution layers where the stride is set as 1. The last stage includes a 3 × 3 convolution operator, which replaces the pooling operation to adjust the output size. Batch normalization and ReLU layers are applied after each convolution layer. Note that the matrix size is reduced in the LUSP branch only by adjusting the stride of convolution so that the local urban surface features are able

to be better preserved.

Furthermore, we aggregate the areas around the center grid within k meters, to build a larger LSCI matrix with size M × M × d. More details about such design is shown in Fig. 3(d). Where, different colors represent different land surface properties; the left bottom of LSCI matrix shows an example to depict a 30-by-30 m grid; and the right top of the matrix shows LSCI matrix with neighborhood information. Our experiments demonstrate that such scheme brings about 0.1 K improvements of the estimation error, because of the perception of the context information. Processed by the LUSP branch, the local land surface features and the 3D geometric structure information are encoded to a latent vector with size 1 × 32C (C is a hyper-parameter with value 32).

At last, the outputs from LUSP and GPFP branches will be concatenated to form a latent feature vector with size 32C + 16Cg (Cg is the hyper-parameter that defines the number of neurons for each layer in the MLP, in our approach, it is set as 32). This vector contains the information of broader-scale atmospheric forcing factors, high-resolution local urban surface features, along with the 3D geometric structure information. This is then fed into a regression branch which is comprised of three fully-connected layers.

2.3. Loss functions

Following the previous work (Karpatne et al., 2017), we employ the mean squared error and the L2 normalization of the network weights to measure the loss. The overall loss can be written as:

$$\underset{\mathbf{W},\text{b}}{\operatorname{argmin}}\ L(Y,\widehat{Y}) + \lambda R(\mathbf{W})$$

$$L(Y,\widehat{Y}) = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \widehat{y_i}\right) \tag{3}$$

$$\text{R}(\mathbf{W}) = ||\mathbf{W}||_2$$

Here, Y and $\widehat{Y}$ represent the ground truth set and predicted results set, W and b are the combined coefficient of weights and bias terms, $y_i \in Y, \widehat{y_i} \in \widehat{Y}, i = 1, \ldots, N$, N is the data size of Y and $\widehat{Y}$, λ is the weight of regularization term.

3. Experiments

3.1. Dataset and Implementation detail

Dataset. In this work, Zhang Zhou Harbor[1], a small city in China has been selected as the testing area. As shown in Fig. 1**(a)**, such region covers 150.56 square kilometers and contains various terrains such as urban, mountain, water, etc. (as highlighted in the red box). We collected a suite of datasets from various sources in the study region to implement our DUD model, described below.

- **LST**. The first dataset is the land surface temperature calculated by NASA's Landsat satellite imagery. Landsat provides the longest continuous space-based record of Earth's land since 1973. The imagery includes different types of band (Blue, Red, Green, Near Infrared, etc.), special bands can be used to calculate LST via classic algorithms (Weng et al., 2004; Yu et al., 2014). The dataset is publicly accessible from the website of American USGS[2]. In this study, we collected the data from Jul. 2013 to Jul. 2020, with a return period of 16 days. The spatial resolution of such data is 30-by-30 m. A visualization map for illustrative purpose is shown in Fig. 1**(b)**, where each pixel represents the LST of a 30-by-30 m area.
- **NDVI**. The second dataset is Normalized Difference Vegetation Index data, which is also measured by NASA's Landsat satellite. The value range of NDVI is from -1 to 1, the larger index implies the denser plant. We filtered the data with the same temporal interval and location of the LST data set. The resolution of such data is 30-by-30 m. The download site is American USGS website or NASA's MODIS[3].
- **Atmospheric forcing**. The third dataset is the atmospheric forcing data, which is provided by the NASA MERRA-2 reanalysis data system (Gelaro et al., 2017). The data describes the comprehensive characteristics of specific area, because each region has Atmospheric features, respectively. Table 1 list the main components of the atmospheric forcing data. This data is publicly accessible from the NASA MERRA-2 website[4]. The resolution of such data is 0.5° latitude × 0.625° longitude. This atmospheric forcing data also increases the generalizability of our DUD model.
- **Land surface 3D structure**. The last dataset is the 3D point cloud data of the entire Zhang Zhou Harbor region, which describe the precise 3D structure of the area. Such data is constructed by the UAV 3D model (DaJiang Inspire-1 UAV) along with a RIEGL VMX-450 mobile laser scanning system (with two full-view RIEGL VQ-450 laser scanners, and can produce 1.1 million range measurements per second, capable to acquire nearly 100 GB point clouds data in 1 h). This data cover different scenes such as urban, town, village, etc. Such point cloud data is then manually labeled to eight main categories: water (noted as blue), building (red), vegetation (green), soil (yellow), road (gray), pavement (white), vehicle (purple), and other

(black). We remove some objects which are extremely incomplete or subtle influence for land surface temperature, such as the streetlights, pedestrians and hard scape like garden walls, fountains, etc.

Finally, we align all data, let each cell (a 30-by-30 m area) data can be described by a set of attributes denoted by 4-tuple:

$$\text{T} = \{\text{t}|\text{t} = (\tau,\eta,\alpha,\rho)\}$$

Where τ, η, α, ρ are the value of LST, NDVI, atmosphere features and the set of point cloud. We have open-sourced all of the datasets, they can be downloaded from FTP server[5]. Due to the point cloud data contains some of the sensitive geography information, we release the constructed LSCI for Zhang Zhou Harbor instead.

Fig. 1**(b)** shows the examples of different datasets, in which the Fig. 1 **(a)** is the visualized result of the label with the different urban surface properties denoted by different colors; The Fig. 1**(c)** is an example of the labeled point cloud data of a 30-by-30 m region, with the same resolution of the NDVI and LST data.

Implementation Detail. In this study, we implement the proposed DUD network based on the PyTorch (Paszke et al., 2019) framework. In the training phase, we adopt Adam solver (Kingma & Ba, 2014) with an initial learning rate 0.0002, which decayed by $0.5^{1/1000}$ for each epoch. Then the network is trained for 3000 epochs on a RTX2080Ti GPU. The hyper-parameter C, C_g and k are set at 32, 64 and 60, respectively.

3.2. Evaluations of the proposed DUD system

3.2.1. Performance

To evaluate the performance of our approach, we select ten pieces of data that cover different seasons of a year. In accordance with previous work (Daw et al., 2020; Karpatne et al., 2017), the root mean squared error (RMSE) is applied to measure the performance of our proposed approach. The estimation error is measured by the unit of Kelvin, and the ratio of training data is 70%, corresponding results are shown in Table 2. Results show that the average error of the ten groups of data is around 0.11 K (last column), demonstrating that the proposed approach is able to provide consistent high-accuracy results independent of seasonal changes. The corresponding visualized results are shown in Fig. 4. The first line shows the ground truth (the color bar is from deep blue to deep red, represents 0 Celsius to 40 Celsius correspondingly), and the third line shows the estimation error map that depicts the level of error (from 0 K to 1.8 K) by different colors.

3.2.2. Ablation studies

In this part, a series of experiments are designed to evaluate how the local structures in urban land surface affect the performance of the DUD network. We select 10 pieces of data between 01 January 2017 and 31 December 2017. These data are selected to cover different seasons in a year. To guarantee the reliability of evaluation results, all selected data are cloud-free to avoid the incorrect temperature sampling by the satellite. Corresponding results are shown in Table 3. We list the average RMSE of the ten selected pieces of data with and without the LUSP branch. Besides, we also adjust the range of neighborhood k in the LUSP branch to further ameliorate the generalization of our approach, here

Table 2
The results of our approach in different seasons.

Quarter-1	Quarter-2	Quarter-3	Quarter-4	Avg. Error(K)
0.102	0.122	0.107	0.109	0.116

---

[1] Located in Xiamen, Fujian, China
[2] https://landsat.gsfc.nasa.gov/
[3] https://modis.gsfc.nasa.gov/
[4] https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/
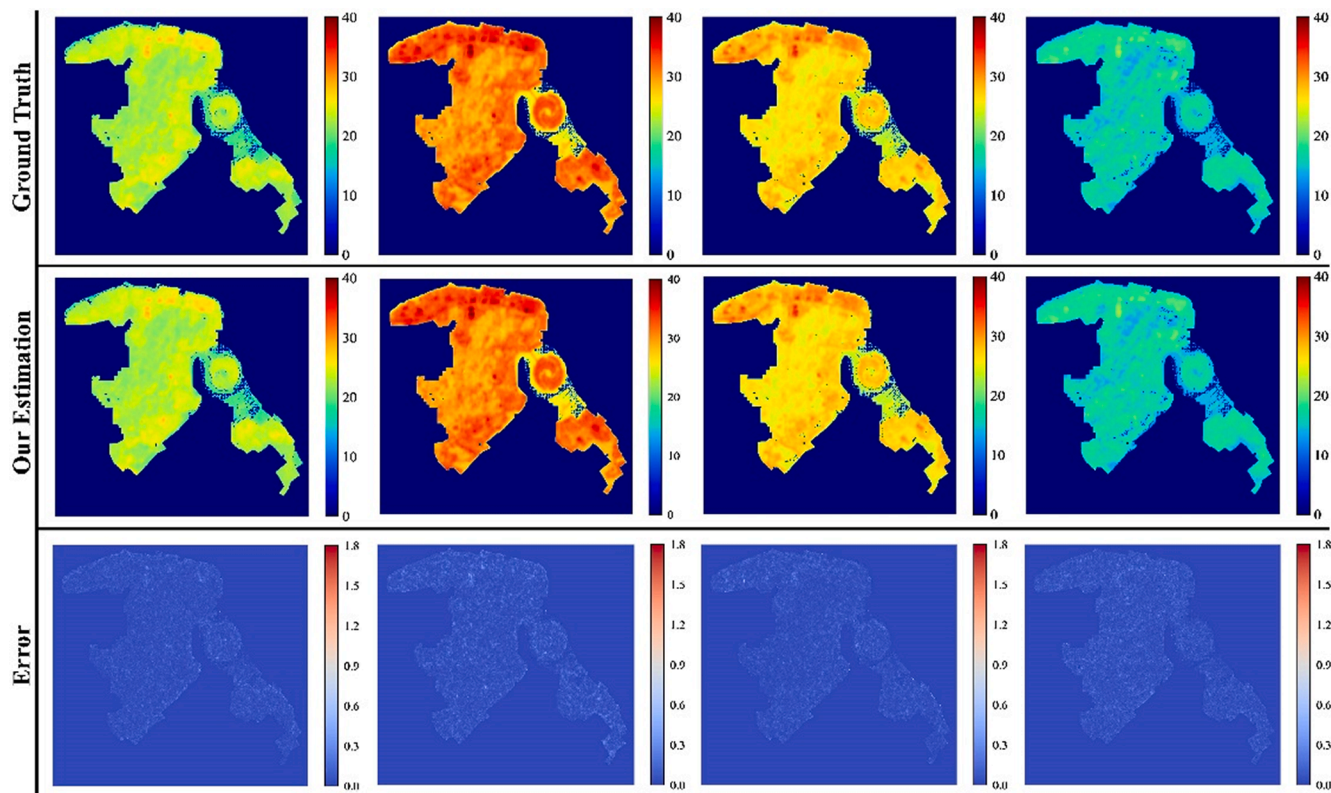
[5] ftp://182.61.174.17/DUD_dataset/

Fig. 4. The visualized results of the ground truth, our approach and the error map for quarter 1 to 4.

Table 3

Ablation studies. k is the range of neighborhood. LSCI is removed for $DUD^{0,2}$.

Method	k(meter)	LSCI	Avg. Error(K)
DUD^0	30	No	1.064
DUD^1	30	Yes	0.223
DUD^2	60	No	1.053
DUD^3	60	Yes	0.131

two classic settings k = 30 and k = 60 are applied. The results

demonstrate that the proposed LUSP branch enables the network to capture the urban local geometric structures, thus leading to a significant performance improvement of LST prediction, as shown in the results of the second (DUD^0) and four rows (DUD^2) of the Table 3. Increasing the reception field further reduces the estimation error, as shown in the third (DUD^1) and fifth (DUD^3) rows of the Table 3.

3.2.3. Effectiveness of the LSCI

In this part, to validate the effectiveness of the proposed LSCI, a point cloud oriented network PointNet (Qi et al., 2017) is applied for
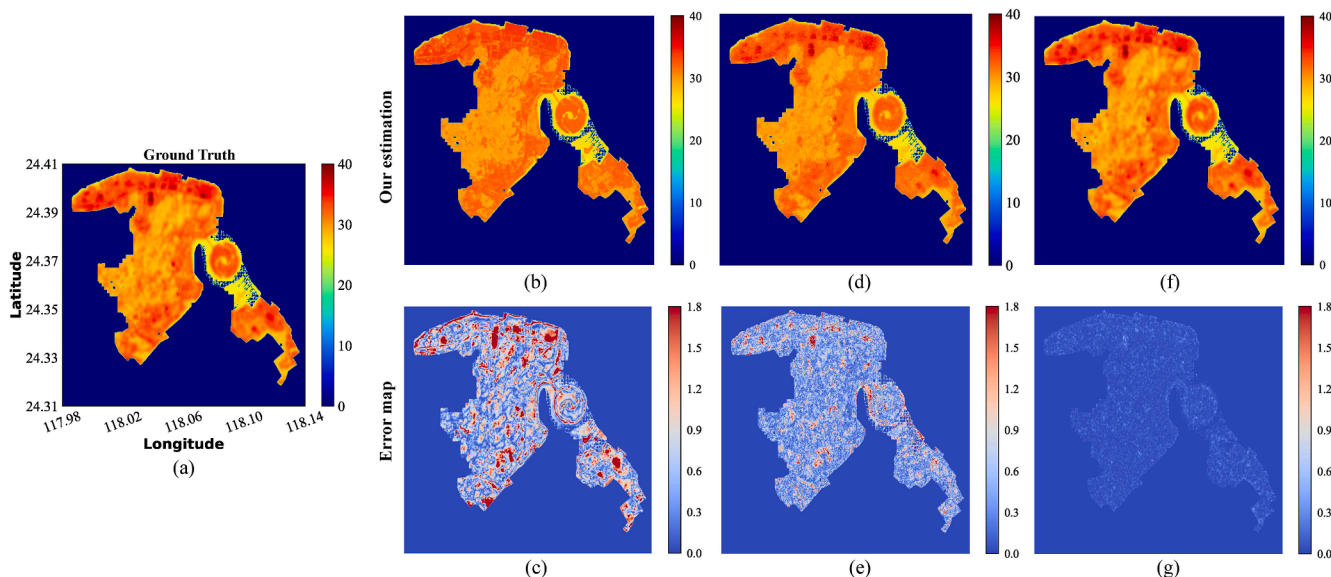

Fig. 5. The visualization of results. (a) Is LST map. (b) and (c) are our estimation and visualized error map without LUSP branch. (d) and (e) are corresponding results that directly import the point cloud to the network. (f) and (g) show our results.

comparison. We adjust the LUSP branch by the framework of PosPool (Liu, Hu, Cao, Zhang, & Tong, 2020) that integrates point could into the deep residual network, such framework captures the point cloud feature via point-wise transformation and local aggregation operator to generate latent vector as a global feature. The results are shown in Fig. 5, where (b) and (c) are our estimation and visualized error map without LUSP branch, (d) and (e) are corresponding results that directly import the point cloud to the LUSP branch of DUD network, (f) and (g) show our DUD network results. The estimation error is shown to be reduced by introducing point cloud, although such improvement is limited. Comparing with the version of without LUSP branch, LUSP reduce the average of predict error, Fig. 5(e) and Fig. 5(g).

The corresponding statistics are listed in Table 4. Directly importing point cloud into the network improves the result of 0.166 K, but such a scheme leads to a significant over-fitting. On the contrary, the network performance is substantially improved when the raw 3D point cloud data is replaced by the proposed LSCI.

3.2.4. Performance under various ratios of training samples

In this part, the estimation performances under different ratios of training samples are applied to evaluate the DUD network under limited sample conditions. Specifically, the same ten groups of data in year 2017 are collected, corresponding statistics are shown in Table 5. The first row denotes various ratios of the training data (from 22% to 78%), and the second row shows the corresponding RMSE. These results demonstrate that, as the decreases of training sample ratio, the RMSE increases from 0.118 to 0.881. It is noteworthy that such a downtrend is not linear. When the ratio of training data drops down to 50%, the performance of the network suffers a sharp drop (from 0.131 to 0.681). On the other hand, even under the scenario of only 22% training samples, the average RMSE is still lower than 1 K, a typical error magnitude of the dynamic downscaling models (Krayenhoff et al., 2018). The corresponding visualized error map of a typical case (data of 3 Jan. 2017) under different ratios of testing samples are shown in Fig. 6, where (b)-(e) represents five stages error map respectively.

3.2.5. Performance under point cloud with different densities

In practice, one might argue that the high precision point cloud may be too strict for the system, therefore, we conduct a group of experiments based on the urban point cloud models with different densities, to explore the potential feasibility for the replacement of the high precision point cloud.

Specifically, the original high precision point cloud is downsampled to simulate the lower-resolution surface model. Here, we apply the farthest point sampling (FPS) to each cell's $t_i.\rho$(where $t_i \in T, 1 \leq i \leq n$), and six levels of point cloud models (i.e. with the number of points 100, 200, 500, 1000, 2000 and 5000 respectively) are built. Then, the models with different levels are applied to test the performance of the system at various times, under 50% training samples. The average error of estimations is shown in Table 6.

Furthermore, we visualize the result at a typical time (Oct. 2017) in Fig. 7, where (a) shows the ground truth of LST; (b) shows the line chart of the estimation errors based on different levels of models, the horizontal axis represents the number of points in each cell. Overall, the result is consistent with the tendency of Table 6. Estimation results under rougher models slightly suffer reductions of accuracy. Where (c) and (d) are corresponding estimation and error maps under the model with the number of points 500 and 2000. Such a result demonstrates the

Table 5
The evaluations of different ratios training samples.

Train Data Ratio	78%	67%	55%	33%	22%
Average RMSE (K)	0.118	0.116	0.131	0.681	0.881

potential strong upscaling capability of this new network, i.e. replacing the high precision point cloud with the digital elevation model (DEM) or even multi-view remote sensing images as inputs to the DUD without signifcantly compromising the prediction performance. This lays the ground of easily generalizing the network to a much broader scale such as regional or national domain. More results can be viewed in Fig. 8 and Table 8.

3.3. Comparison with traditional method

As is described above, the traditional LST downscaling methods either employ the process-based physical models at great computational cost or fit a traditional statistical or machine learning model such as linear regression or random forests. The dynamic downscaling models, such as the Weather Research and Forecast (WRF) model (Moustaoui Georgescu et al., 2013) and the computational fluid dynamics (CFD) based model (Gromke et al., 2015), usually with an accuracy about 1–2 km and the average error about 2 Kevin, can hardly achieve such a high spatial resolution (30-by-30 m) over the whole city-scale coverage. Because traditional dynamic model's calculation time is beyond exception when high resolution land surface model be applied. Therefore, we focused on the comparisons with previous statistical downscaling methods, such as linear regression (Montgomery et al., 2012), KNN regression (Cover and Hart, 1967), and random forest regression (Liaw et al., 2002). All these methods were implemented based on Scikit-learn (Pedregosa et al., 2011).

Here, the same 10 pieces of data in different seasons in the year 2017 are selected to evaluate different methods. The average RMSE are listed in Table 7. To ensure the fairness of comparison, we also integrate the proposed LSCI into all the methods to embed the local geometry information. Specifically, each dimension of the m × m × d matrix is averaged to a specific value, and the original matrix is reshaped to a 1 × d vector. This vector is then imported into various regression methods.

For the linear regression model, the average error is more than 1 K. Because there is no apparent linear relationships between the variables. The corresponding statistics of four selected cases along with the average value of the ten pieces of data are listed in the second row of the Table 7.

For the KNN regression and random forest regression, the hyperparameters are manually adjusted to best fit the ten pieces of data. Specifically, the number of neighbors in the KNN regression is set as 4. The maximum depth of the tree and the number of trees in the random forest regression are set as 30 and 150, respectively. The corresponding results are shown in the third and fourth row of Table 7. Overall, the results of the traditional machine learning methods perform much better than the linear regression model. The errors of lower than 1 K for both methods are observed. However, these methods might suffer from the generalizability of the hyper-parameter settings when the algorithms are applied to multiple cities on a large scale.

For the proposed DUD network, under the 50% training samples, the average estimation error is about 0.13 K, which is well below the error magnitude of the traditional statistical downscaling methods tested above. And comparing with raw method, such as PointNet (Qi et al., 2017), RandLA (Hu et al., 2020) point-based network, that extract 3D point cloud feature to regress LST, our method demonstrates fine performance. In the future, we plan to apply more comprehensive testing under various cities.
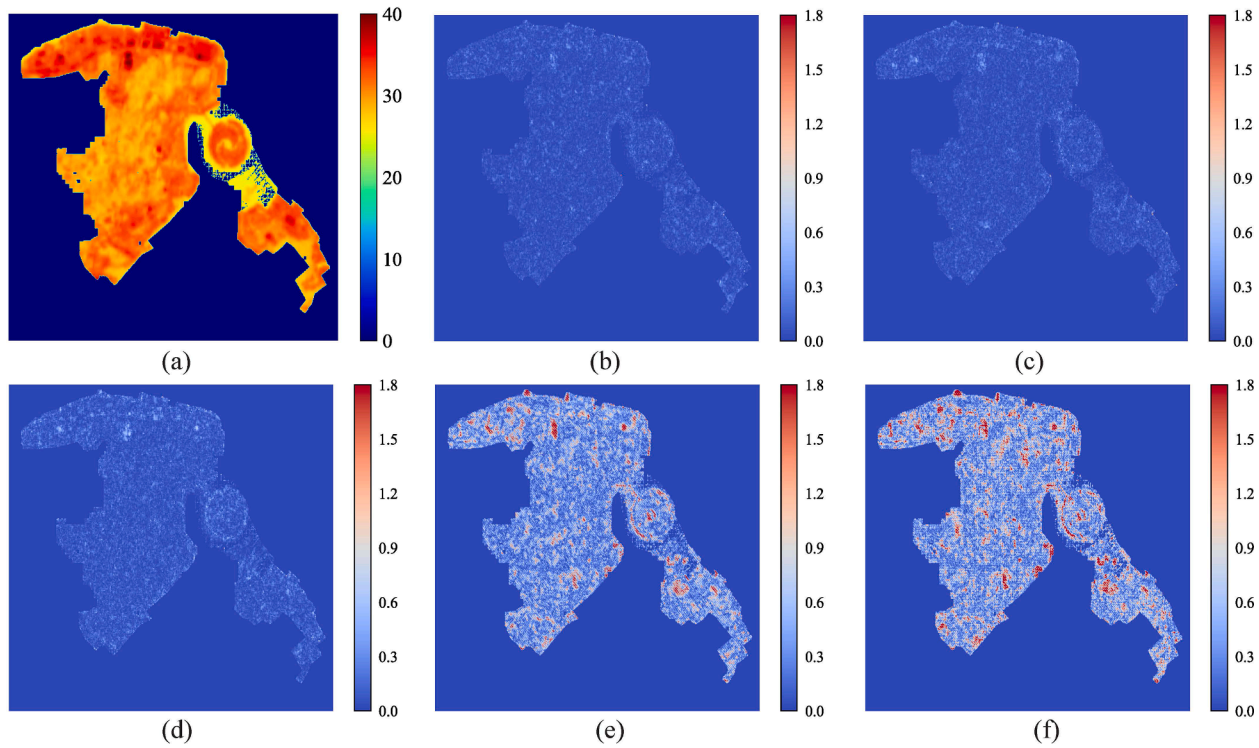
Table 4
The evaluation of how the LSCI affect the results.

	Without LUSP branch	LUSP branch based on point cloud	LUSP branch based on the LSCI
Train/Test Error(K)	1.103/1.168	0.701/1.002	0.112/0.122

Fig. 6. Visualized the error map. (a) Shows the satellite captured land surface temperature, (b)-(f) are error map under 78%, 67%, 55%, 33%, 22% ratio of training data, respectively.

Table 6
The results in different densities of point cloud.

Points	100	200	500	1000	2000	5000
Avg. Error(K)	0.139	0.143	0.141	0.136	0.139	0.135

4. Discussion

- **Contribution**. From the Experiment 3 section, we find that the proposed DUD system performs considerably well for the testing area – Zhang Zhou Harbor. Benefiting from the incorporation of the 3D

point cloud, average errors about 0.15 K is observed with 50% of data as training samples. Major contributions of this work include:

i) it provides a first-of-its-kind solution of surface temperature downscaling over highly-complex urban areas by implementing a PINN-based architecture to incorporate both process-based insights and data-driven information.

ii) Future extension of this work to larger-scale domains (such as regional, national and global scales) and to include more predicted physical quantities (such as surface solar radiation, turbulence, surface wind speed, etc.) world bring new inspirations to the global climate chance, energy flow, and other fields.
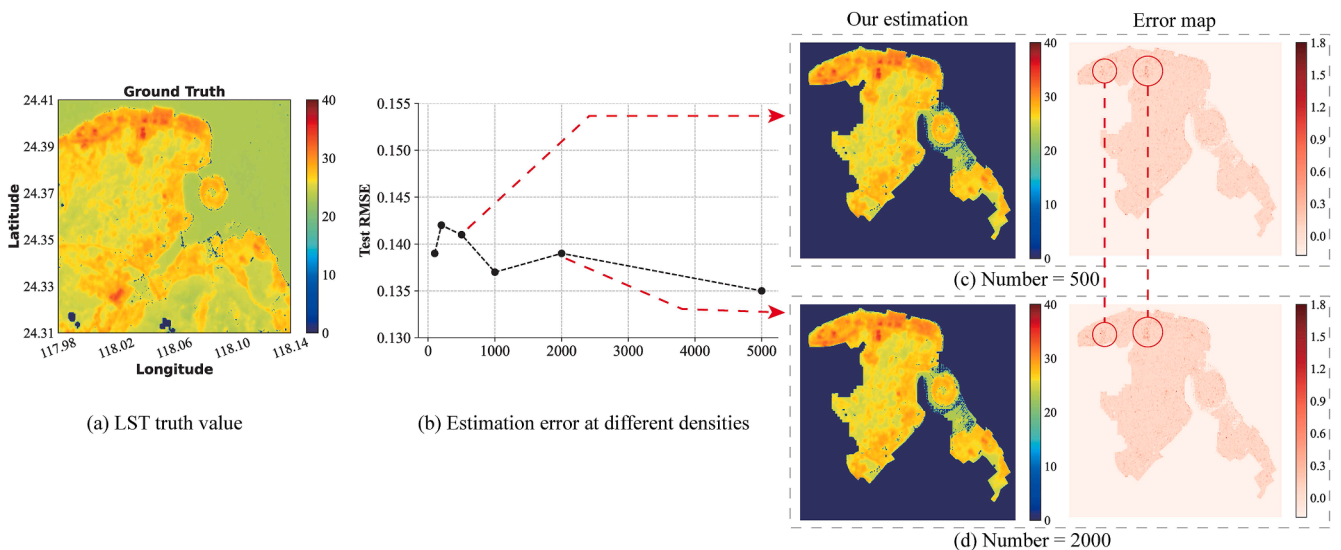

Fig. 7. The visualization of results under different densities of point cloud.
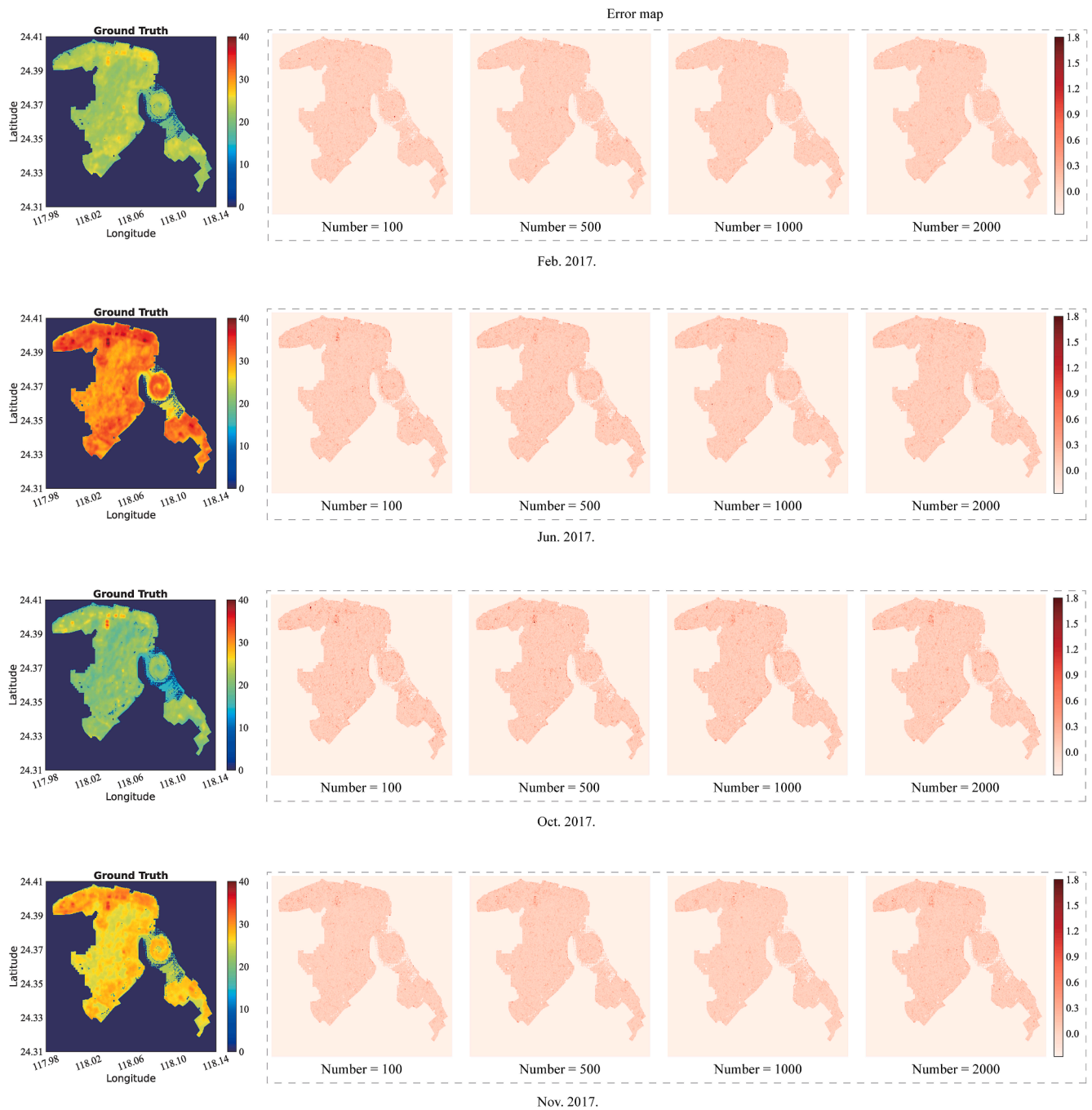
Fig. 8. Additional visualized results under point cloud with various densities at different times.

Table 7
Average error for 10 pieces of data for different seasons in year 2017.

Method	Feb.	Jun.	Oct.	Dec.	Avg. Error(K)
Linear Reg	1.634	1.350	1.586	1.533	1.304
KNN Reg	1.099	0.828	1.041	1.031	0.861
Random forest Reg	0.541	0.452	0.539	0.460	0.426
PointNet	1.001	0.808	0.926	0.902	0.855
RandLA	0.980	0.901	0.913	0.823	0.837
DUD	**0.156**	**0.151**	**0.144**	**0.149**	**0.131**

Table 8
Average error results under point cloud with various densities at different seasons in year 2017.

	Number = 100	Number = 500	Number = 1000	Number = 2000
Feb. 2017	0.155	0.154	0.151	0.149
Jun. 2017	0.124	0.141	0.137	0.137
Oct. 2017	0.139	0.141	0.136	0.139
Nov. 2017	0.147	0.139	0.137	0.132

- **Limitation**. One limitation of the proposed work lies in the range of the testing area. Due to the data availability (labeling high precision 3D point cloud data is labor intensive and thus rather limited for larger-scale experiments.), the experiment of the DUD system at the

current stage focuses on a single city in China to validate the algorithm.

- **Future Work**. Firstly, our work of expanding the DUD system to cover more cities over a much larger domain is underway. To achieve this, the most challenging part is to replace the labeled high precision point cloud data with other surrogates. Actually, some efforts had already been made in the current work 3.2.5, we design a group of experiments to explore the relationship between the density of the point and the performance. Such experiments may provide benefit inspirations for the potential replacement solutions of the high precision point cloud, such as the digital surface model (DSM) or even multi-view remote sensing data. Secondly, the state-of-art semantic segmentation methods are being tested to develop an automatic data labeling scheme to be incorporated in the DUD system. Lastly, more key urban surface variables of public interest that can be prognostically predicted by the dynamic urban models will be included in the DUD system of future versions.

5. Conclusions

In this paper, we propose PINN-based framework DUD network for high-resolution high-precision urban surface temperature downscaling. The DUD network leverages the global feature perception (GPFP) branch to capture broader-scale influences by the atmospheric forcing. Furthermore, the local urban surface perception (LUSP) branch extracts the high-precision land surface geometry information by employing a proposed local spatial coefficient index (LSCI). With both modules, the DUD network achieves high-accuracy temperature prediction with the estimated error of less than 0.2 K. The DUD network combines process-based modeling and deep learning approach to provide ultrahigh resolution urban LST predictions in a computationally efficient manner. This network can be adopted in other urban surface climate prediction applications that otherwise would require either computationally-expensive (and maybe unattainable) dynamic downscaling or less-accurate traditional statistical methods. For future work, we will up-scale the test region and estimate other surface climate variables over multiple metropolitan areas.

CRediT authorship contribution statement

Linwei Chen: Conceptualization, Methodology, Software, Writing – original draft. **Bowen Fang:** Validation, Data curation, Visualization. **Lei Zhao:** Formal analysis, Data curation, Writing – review & editing. **Yu Zang:** Conceptualization, Resources, Supervision, Funding acquisition. **Weiquan Liu:** Writing – review & editing, Formal analysis. **Yiping Chen:** Software, Formal analysis. **Cheng Wang:** Data curation, Writing – review & editing. **Jonathan Li:** Resources, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

References

Alexander, C., 2020. Normalised difference spectral indices and urban land cover as indicators of land surface temperature (LST). Int. J. Appl. Earth Obs. Geoinf. 86, 102013. https://doi.org/10.1016/j.jag.2019.102013. Retrieved from.

Alexander, C., 2021. Influence of the proportion, height and proximity of vegetation and buildings on urban land surface temperature. Int. J. Appl. Earth Obs. Geoinf. 95, 102265. https://doi.org/10.1016/j.jag.2020.102265. Retrieved from.

Anderson, G.B., Bell, M.L., 2011. Heat waves in the United States: mortality risk during heat waves and effect modification by heat wave characteristics in 43 US communities. Environ. Health Perspect. 119 (2), 210–218. https://doi.org/10.1289/ehp.1002313. Retrieved from.

Bonafoni, S., 2016. Downscaling of Landsat and MODIS Land Surface Temperature Over the Heterogeneous Urban Area of Milan. IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens. 9 (5), 2019–2027. https://doi.org/10.1109/JSTARS.2016.2514367. Retrieved from.

Cover, T., Hart, P., 1967. Nearest neighbor pattern classification. IEEE Trans. Inf. Theory 13 (1), 21–27.

Daw, A., Thomas, R.Q., Carey, C.C., Read, J.S., Appling, A.P., Karpatne, A., 2020. Physics-Guided Architecture (PGA) of Neural Networks for Quantifying Uncertainty in Lake Temperature Modeling. In: Demeniconi, C., Chawla, N. (Eds.), Proceedings of the 2020 SIAM International Conference on Data Mining. Society for Industrial and Applied Mathematics, Philadelphia, PA, pp. 532–540. https://doi.org/10.1137/1.9781611976236.60.

Dunne, J.P., Stouffer, R.J., John, J.G., 2013. Reductions in labour capacity from heat stress under climate warming. Nat. Clim. Change 3 (6), 563–566. https://doi.org/10.1038/nclimate1827. Retrieved from.

Fowler, H.J., Blenkinsop, S., Tebaldi, C., 2007. Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. Int. J. Climatol. 27 (12), 1547–1578. Retrieved from.

Gaffin, S.R., Rosenzweig, C., Khanbilvardi, R., Parshall, L., Mahani, S., Glickman, H., Goldberg, R., Blake, R., Slosberg, R.B., Hillel, D., 2008. Variations in New York city's urban heat island strength over time and space. Theor. Appl. Climatol. 94 (1-2), 1–11. https://doi.org/10.1007/s00704-007-0368-3. Retrieved from.

Gelaro, R., McCarty, W., Suárez, M.J., Todling, R., Molod, A., Takacs, L., Randles, C.A., Darmenov, A., Bosilovich, M.G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A.M., Gu, W., Kim, G.-K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J.E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S.D., Sienkiewicz, M., Zhao, B., 2017. The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). J. Clim. 30 (14), 5419–5454. Retrieved from.

Georgescu, M., Morefield, P.E., Bierwagen, B.G., Weaver, C.P., 2014. Urban adaptation can roll back warming of emerging megapolitan regions. Proc. Natl. Acad. Sci. 111 (8), 2909–2914. https://doi.org/10.1073/pnas.1322280111. Retrieved from.

Georgescu, M., Moustaoui, M., Mahalov, A., Dudhia, J., 2013. Summer-time climate impacts of projected megapolitan expansion in Arizona. Nat. Clim. Change 3 (1), 37–41. https://doi.org/10.1038/NCLIMATE1656. Retrieved from.

Grimm, N.B., Faeth, S.H., Golubiewski, N.E., Redman, C.L., Wu, J., Bai, X., Briggs, J.M., 2008. Global change and the ecology of cities. Science 319 (5864), 756–760. https://doi.org/10.1126/science.1150195. Retrieved from.

Gromke, C., Blocken, B., Janssen, W., Merema, B., van Hooff, T., Timmermans, H., 2015. CFD analysis of transpirational cooling by vegetation: Case study for specific meteorological conditions during a heat wave in Arnhem, Netherlands. Build. Environ. 83, 11–26. https://doi.org/10.1016/j.buildenv.2014.04.022I. Retrieved from.

Hackel, T., Savinov, N., Ladicky, L., Wegner, J.D., Schindler, K., Pollefeys, M., 2017. SEMANTIC3D.NET: A new large-scale point cloud classification benchmark. In: ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. IV-1-W1, pp. 91–98.

Ham, Y.-G., Kim, J.-H., Luo, J.-J., 2019. Deep learning for multi-year ENSO forecasts. Nature 573 (7775), 568–572. Retrieved from.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep Residual Learning for Image Recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

Hu, Q., Yang, B., Xie, L., Rosa, S., Guo, Y., Wang, Z., Markham, A., 2020. Randla-net: Efficient semantic segmentation of large-scale point clouds. Paper presented at the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition.

Huang, C., Barnett, A.G., Wang, X., Vaneckova, P., FitzGerald, G., Tong, S., 2011. Projecting future heat-related mortality under climate change scenarios: a systematic review. Environ. Health Perspect. 119 (12), 1681–1690. https://doi.org/10.1289/ehp.1103456. Retrieved from.

Isaac, M., van Vuuren, D.P., 2009. Modeling global residential sector energy demand for heating and air conditioning in the context of climate change. Energy Policy 37 (2), 507–521. Retrieved from.

Karpatne, A., Watkins, W., Read, J., Kumar, V., 2017. Physics-guided neural networks (pgnn): An application in lake temperature modeling. arXiv preprint arXiv: 1710.11431.

Keramitsoglou, I., Kiranoudis, C.T., Weng, Q., 2013. Downscaling Geostationary Land Surface Temperature Imagery for Urban Analysis. IEEE Geosci. Remote Sens. Lett. 10 (5), 1253–1257. Retrieved from.

Kingma, D.P., Ba, J., 2014. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

Klambauer, G., Unterthiner, T., Mayr, A., Hochreiter, S., 2017. Self-normalizing neural networks. In: Advances in neural information processing systems, pp. 971–980.

Krayenhoff, E.S., Moustaoui, M., Broadbent, A.M., Gupta, V., Georgescu, M., 2018. Diurnal interaction between urban expansion, climate change and adaptation in US cities. Nat. Clim. Change 8 (12), 1097–1103. https://doi.org/10.1038/s41558-018-0320-9. Retrieved from.

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2017. ImageNet classification with deep convolutional neural networks. Commun. ACM 60 (6), 84–90. https://doi.org/10.1145/3065386. Retrieved from.

Kusaka, H., Kondo, H., Kikegawa, Y., Kimura, F., 2001. A simple single-layer urban canopy model for atmospheric models: Comparison with multi-layer and slab models. Bound.-Layer Meteorol. 101 (3), 329–358.

Kustas, W.P., Norman, J.M., Anderson, M.C., French, A.N., 2003. Estimating subpixel surface temperatures and energy fluxes from the vegetation index-radiometric temperature relationship. Remote Sens. Environ. 85 (4), 429–440. https://doi.org/10.1016/S0034-4257(03)00036–1l. Retrieved from.

Liaw, A., Wiener, M., et al., 2002. Classification and regression by randomForest. R news 2 (3), 18–22.

Liu, Z., Hu, H., Cao, Y., Zhang, Z., Tong, X., 2020. A closer look at local aggregation operators in point cloud analysis. In: European Conference on Computer Vision, pp. 326–342.

Meng, Q., Wang, W., Zhou, T., Shen, J., Jia, Y., Van Gool, L., 2021. Towards a weakly supervised framework for 3d point cloud object detection and annotation. IEEE Trans. Pattern Anal. Mach. Intell.

Middel, A., Chhetri, N., Quay, R., 2015. Urban forestry and cool roofs: Assessment of heat mitigation strategies in Phoenix residential neighborhoods. Urban For. Urban Greening 14 (1), 178–186. https://doi.org/10.1016/j.ufug.2014.09.010.

Montgomery, D.C., Peck, E.A., Vining, G.G., 2012. Introduction to linear regression analysis, Vol. 821. John Wiley & Sons.

Mora, C., Dousset, B., Caldwell, I.R., Powell, F.E., Geronimo, R.C., Bielecki, C., Counsell, C.W.W., Dietrich, B.S., Johnston, E.T., Louis, L.V., Lucas, M.P., McKenzie, M.M., Shea, A.G., Tseng, H., Giambelluca, T.W., Leon, L.R., Hawkins, E. d., Trauernicht, C., 2017. Global risk of deadly heat. Nat. Clim. Change 7 (7), 501–506. Retrieved from.

Oleson, K.W., Bonan, G.B., Feddema, J., Vertenstein, M., Grimmond, C., 2008. An urban parameterization for a global climate model. Part I: Formulation and evaluation for two cities. J. Appl. Meteorol. Climatol. 47 (4), 1038–1060. https://doi.org/10.1175/2007JAMC1597.1. Retrieved from.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., . . . others. 2019. Pytorch: An imperative style, high-performance deep learning library. In: Advances in neural information processing systems, pp. 8026–8037.

Patz, J.A., Campbell-Lendrum, D., Holloway, T., Foley, J.A., 2005. Impact of regional climate change on human health. Nature 438 (7066), 310–317.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Duchesnay, E., 2011. Scikit-learn: Machine Learning in Python. J. Mach. Learn. Res. 12, 2825–2830.

Perera, A., Nik, V.M., Chen, D., Scartezzini, J.-L., Hong, T., 2020. Quantifying the impacts of climate change and extreme climate events on energy systems. Nat. Energy 5 (2), 150–159. https://doi.org/10.1038/s41560-020-0558-0. Retrieved from.

Pu, R., 2021. Assessing scaling effect in downscaling land surface temperature in a heterogenous urban environment. Int. J. Appl. Earth Obs. Geoinf. 96, 102256. https://doi.org/10.1016/j.jag.2020.102256. Retrieved from.

Qi, C.R., Su, H., Mo, K., Guibas, L.J., 2017. Pointnet: Deep learning on point sets for 3d classification and segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 652–660.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical image computing and computer-assisted intervention, pp. 234–241.

Spak, S., Holloway, T., Lynn, B., Goldberg, R., 2007. A comparison of statistical and dynamical downscaling for surface temperature in North America. J. Geophys. Res.: Atmospheres 112 (D8). https://doi.org/10.1029/2005JD006712. Retrieved from.

Tang, J., Niu, X., Wang, S., Gao, H., Wang, X., Wu, J., 2016. Statistical downscaling and dynamical downscaling of regional climate in China: Present climate evaluations and future climate projections. J. Geophys. Res.: Atmospheres 121 (5), 2110–2129. https://doi.org/10.1002/2015JD023977. Retrieved from.

Thomas, H., Qi, C.R., Deschaud, J.-E., Marcotegui, B., Goulette, F., Guibas, L.J., 2019. KPConv: Flexible and Deformable Convolution for Point Clouds. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV).

Wei, B., Bao, Y., Yu, S., Yin, S., Zhang, Y., 2021. Analysis of land surface temperature variation based on MODIS data a case study of the agricultural pastural ecotone of northern China. Int. J. Appl. Earth Obs. Geoinf. 100, 102342. https://doi.org/10.1016/j.jag.2021.102342. Retrieved from.

Weng, Q., Lu, D., Schubring, J., 2004. Estimation of land surface temperature-vegetation abundance relationship for urban heat island studies. Remote Sens. Environ. 89 (4), 467–483. https://doi.org/10.1016/j.rse.2003.11.005.

Xu, Q., Sun, X., Wu, C.-Y., Wang, P., Neumann, U., 2020. Grid-GCN for Fast and Scalable Point Cloud Learning. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).

Yu, X., Guo, X., Wu, Z., 2014. Land surface temperature retrieval from Landsat 8 TIRS—Comparison between radiative transfer equation-based method, split window algorithm and single channel method. Remote Sens. 6 (10), 9829–9852. https://doi.org/10.3390/rs6109829. Retrieved from.

Zakšek, K., Oštir, K., 2012. Downscaling land surface temperature for urban heat island diurnal cycle analysis. Remote Sens. Environ. 117, 114–124. Retrieved from.

Zhang, D., He, F., Tu, Z., Zou, L.u., Chen, Y., 2019. Pointwise geometric and semantic learning network on 3D point clouds. Integr. Comput.-Aided Eng. 27 (1), 57–75.

Zhao, L., Lee, X., Smith, R.B., Oleson, K., 2014. Strong contributions of local background climate to urban heat islands. Nature 511 (7508), 216–219. https://doi.org/10.1038/nature13462. Retrieved from.

Zhao, L., Oleson, K.W., Bou-Zeid, E., Krayenhoff, E.S., Bray, P., Zhu, Q., . . . Oppenheimer, M., 2020. Global Multi-Model Projections of Local Urban Climates. Nat. Clim. Change (in press). Retrieved from 10.1038/s41558-020-00958-8.