

# UC Davis

## UC Davis Previously Published Works

### Title

Mapping urban form into local climate zones for the continental US from 1986-2020.

### Permalink

<https://escholarship.org/uc/item/37g468vt>

### Journal

Scientific Data, 11(1)

### Authors

Qi, Meng

Xu, Chunxue

Zhang, Wenwen

et al.

### Publication Date

2024-02-13

### DOI

10.1038/s41597-024-03042-4

Peer reviewed



OPEN

DATA DESCRIPTOR

# Mapping urban form into local climate zones for the continental US from 1986–2020

Meng Qi<sup>1</sup>, Chunxue Xu<sup>2</sup>, Wenwen Zhang<sup>3</sup>, Matthias Demuzere<sup>4</sup>, Perry Hystad<sup>5</sup>, Tianjun Lu<sup>6</sup>, Peter James<sup>7,8</sup>, Benjamin Bechtel<sup>4</sup> & Steve Hankey<sup>1</sup>✉

Urbanization has altered land surface properties driving changes in micro-climates. Urban form influences people's activities, environmental exposures, and health. Developing detailed and unified longitudinal measures of urban form is essential to quantify these relationships. Local Climate Zones [LCZ] are a culturally-neutral urban form classification scheme. To date, longitudinal LCZ maps at large scales (i.e., national, continental, or global) are not available. We developed an approach to map LCZs for the continental US from 1986 to 2020 at 100 m spatial resolution. We developed lightweight contextual random forest models using a hybrid model development pipeline that leveraged crowdsourced and expert labeling and cloud-enabled modeling – an approach that could be generalized to other countries and continents. Our model achieved good performance: 0.76 overall accuracy (0.55–0.96 class-wise F1 scores). To our knowledge, this is the first high-resolution, longitudinal LCZ map for the continental US. Our work may be useful for a variety of fields including earth system science, urban planning, and public health.

## Background & Summary

The global population is increasingly shifting to urban areas (i.e., 56% of the population currently lives in urban areas and is projected to increase to 68% by 2050)<sup>1</sup>. While much of the future urbanization is expected to occur in Asia and Africa, studying historical urban development patterns for areas that have already experienced high degrees of urbanization can provide valuable data to inform global urban planning and growth management efforts. Cities will be a key component of solutions that aim to achieve sustainable development goals since they are home to large populations that influence a large share of total energy consumption and anthropogenic emissions<sup>2,3</sup>. Targeted urban planning and growth management will be crucial for sustainable development in every dimension, for example, economic growth, societal and political well-being, public health, and environmental sustainability<sup>4–8</sup>.

A fundamental requirement for studying the historic impact of urban development and expansion is a reliable longitudinal dataset that measures urban form and urban functions. Urban form is defined as the physical characteristics of built-up areas (at various scales), including building density, building heights, construction materials, street layouts, and fraction of green space<sup>9,10</sup>. Urban functions are made possible by various infrastructure systems from diverse sectors (e.g., energy, transport, waste management, etc.). As a holistic system, the interplay of urban form and urban functions shapes the built and natural environment, which in turn influences people's daily activities, exposures to environmental pollutants (e.g., air pollution, heat flux, hazardous materials), and impacts to human health<sup>11,12</sup>. Urbanization also drives local and regional changes in climate, leading to phenomena such as the urban heat island [UHI] and urban dry island [UDI] effects<sup>3,13,14</sup>. The UHI effect (i.e., the

<sup>1</sup>School of Public and International Affairs, Virginia Polytechnic Institute and State University, Blacksburg, VA, 24060, USA. <sup>2</sup>College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis, OR, 97331, USA. <sup>3</sup>Edward J. Bloustein School of Planning and Public Policy, Rutgers University, New Brunswick, New Jersey, 08901, USA. <sup>4</sup>Urban Climatology Group, Department of Geography, Ruhr-University Bochum, Bochum, 44801, Germany. <sup>5</sup>College of Health, Oregon State University, Corvallis, OR, 97331, USA. <sup>6</sup>Department of Epidemiology and Environmental Health, University of Kentucky, Lexington, KY, 40536, USA. <sup>7</sup>Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, MA, 02215, USA. <sup>8</sup>Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, MA, 02215, USA. ✉e-mail: [hankey@vt.edu](mailto:hankey@vt.edu)

phenomenon of relatively higher temperatures in urban areas compared to surrounding non-urban areas) is an emerging risk factor for urban residents and has been extensively studied<sup>15–21</sup>. Along with global climate change, urban residents are expected to face more frequent extreme heat events and exacerbated heat stress, which poses significant risks to human health and well-being<sup>22–24</sup>.

In urban climate science, urban datasets that provide key parameters of the urban environment are crucial as they are a required component for urban climate and weather modeling. Datasets that only report simplified urban versus nonurban or a low level of thematic classification limit urban and climate analyses. For example, in-depth UHI analysis by different urban and natural forms necessitates a more detailed urban form classification. In addition, some empirical urban parameterizations derived for specific cities or regions may not transfer to other cities, regions or generalize to larger scales (e.g., continental or global). A detailed and harmonized urban classification scheme tailored to climate impacts is a critical need for urban studies. Driven by this need, Stewart and Oke proposed a generic, culturally-neutral urban form classification scheme called Local Climate Zones [LCZ] which is meant to be local in scale, climatic in nature, and zonal in representation<sup>25</sup>. The LCZ scheme consists of 10 built classes and 7 land cover classes and was originally designed to provide field metadata for UHI observational studies<sup>25,26</sup>. Stewart and Oke estimated and provided representative physical characteristics of each LCZ category including geometric and surface cover properties (e.g., sky view factor, impervious/pervious surface fraction, roughness), thermal, radiative and metabolic properties (e.g., surface albedo, anthropogenic heat output) – all of which can play a key role in urban climate modeling<sup>25</sup>. The scheme is generic and thus considered to be an urban classification solution that is applicable worldwide. A number of studies have shown improved model performance for urban climate modeling after using urban parameters derived from LCZ datasets<sup>27–30</sup>.

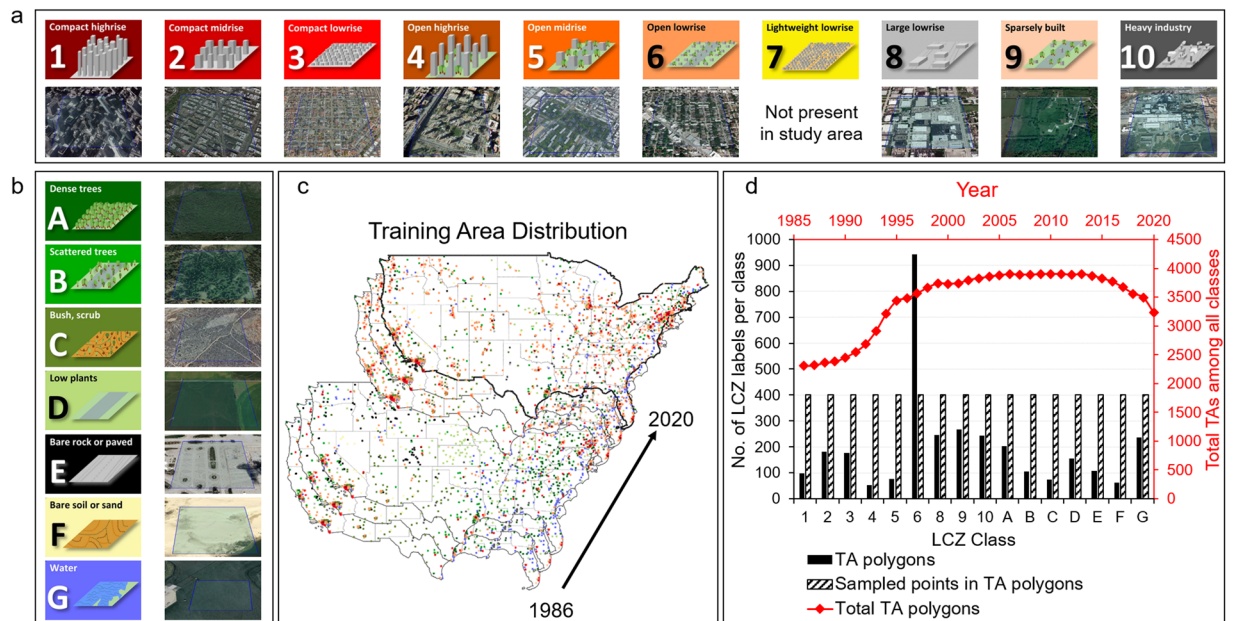
Due to these advantages, applications of the LCZ concept have become increasingly common<sup>31–34</sup>. In 2015, the World Urban Database and Access Portal Tools [WUDAPT] project initiated a community-based effort to collect global LCZ training data and proposed a straightforward LCZ mapping workflow that relies on freely available remote sensing data, crowdsourced training areas, and open software for supervised LCZ classification<sup>35,36</sup>. The WUDAPT project facilitated a remarkable growth in remote sensing based LCZ mapping studies – of which most are developed at the city and regional scale<sup>34,37–42</sup>. Also, the LCZ Generator<sup>43</sup> recently contributed to a significant increase in city-level LCZ maps. Time-series LCZ maps are available at the city or regional scale for a small number of locations (e.g., Bangkok in Thailand, Kunming and Pearl River Delta in China) for a few specific years<sup>44–46</sup>. In contrast, only a limited number of LCZ maps have been generated at the national (i.e., USA<sup>47</sup>), continental (i.e., Europe<sup>48</sup>), and global scale<sup>9,49</sup>. Many of these efforts have been aided by Google Earth Engine [GEE], a cloud-based platform that provides remote sensing datasets at the petabyte-scale and powerful online cloud-computing capabilities<sup>34,50</sup>. A main limitation of these existing large-scale maps is that they are for a single year. To date, no longitudinal LCZ maps are available at national, continental or global scales. The lack of large-scale time-series LCZ maps represents a significant gap in the current LCZ literature. Addressing this gap holds importance to the scientific community interested in exploring the impacts of the growth and change in urban environments overtime.

In this study, we introduce longitudinal LCZ maps with fine spatial resolution (100 meters) for the continental US [CONUS] from 1986 to 2020. Our modeling pipeline generally followed WUDAPT's protocol but was modified to allow large scale mapping with temporal consistency across years. We developed contextual random forest [RF] classifiers with hybrid model development, including (1) crowdsourced and expert training area labeling, (2) local model fine-tuning and cloud model reproduction, and (3) large-scale LCZ surface mapping and post-classification processing. We incorporated publicly available Landsat imagery as well as existing land use and land cover [LULC] datasets and Census information to further improve our model accuracy. To our knowledge, this is the first LCZ dataset that spans over 35 years for the continental US. We anticipate that this dataset will aid future studies with a range of potential applications, e.g., UHI studies, urban climate and weather modeling, urban expansion, etc.

## Methods

The geographic focus of our study is CONUS from 1986 to 2020. Training Areas [TAs] were collected for each year distributed across CONUS for 16 LCZ classes (Fig. 1). Following previous work<sup>47</sup>, we did not include LCZ 7 (lightweight lowrise) in our model since it rarely occurred in our study area. Our hybrid modeling pipeline included (1) hybrid TA labeling, (2) cloud-enabled RF classifier development, and (3) model prediction and post-classification processing. After model development and prediction, we used a number of approaches to evaluate model performance, including 5-fold spatial cross validation for accuracy evaluation, benchmarking against external land use and LCZ maps, and an urban sprawl analysis based on our LCZ product.

**Hybrid TA labeling.** Our approach to collect TAs included two steps. First, we leveraged the labeling power of a crowdsourcing platform – Amazon Mechanical Turk [MTurk] (<https://www.mturk.com/>) following the work of Demuzere *et al.*<sup>47</sup> Briefly, we sampled 500 m × 500 m TAs across CONUS. For each area, we showed 10 unique MTurk workers the corresponding satellite and aerial image from Google Earth (<https://www.google.com/earth/>) and street-level images from Google Street View (<https://www.google.com/streetview/>). The MTurk workers voted for the LCZ category that each TA best represented. Only TAs with at least 70% agreement among the 10 MTurk workers were considered as valid crowdsourced labels. Details on this labeling approach can be found in Demuzere *et al.*<sup>47</sup> The base year for the crowdsourced training labels is 2017 since our TAs expanded on the effort from Demuzere *et al.*<sup>47</sup> We further collected TAs across CONUS by randomly sampling from an existing LCZ map<sup>47</sup> for each LCZ category to ensure balance across LCZ types in our training data. For example, to sample more TAs for LCZ 1, we randomly sampled from all pixels that were classified as LCZ 1 in the LCZ map representative for year 2017<sup>47</sup>, created corresponding 500 m × 500 m polygons and sent them to MTurk to follow the

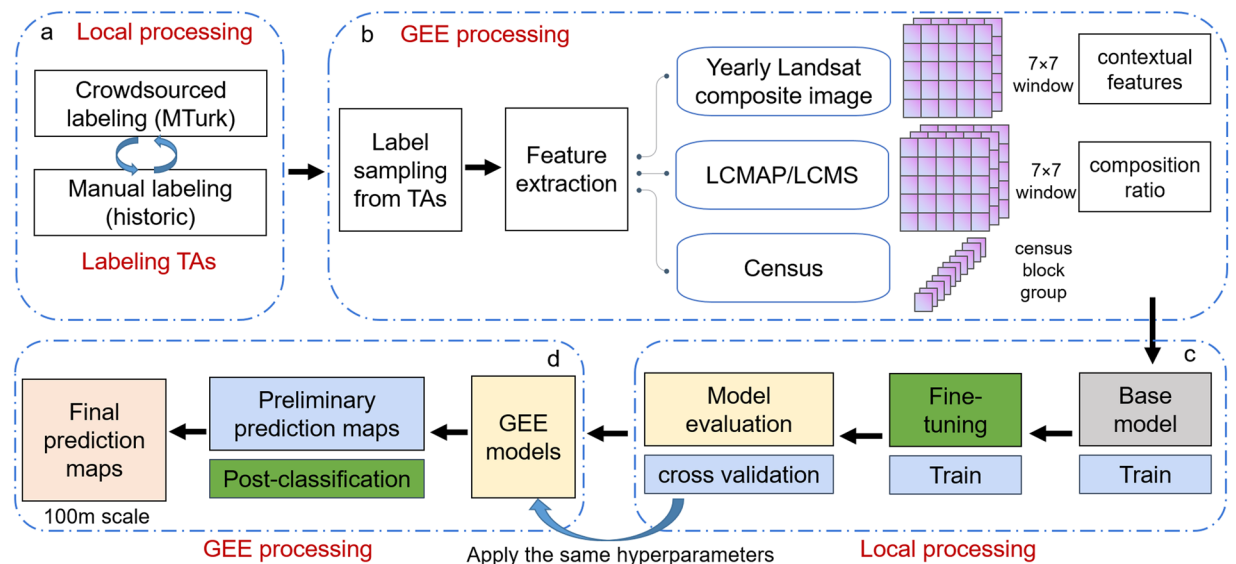


**Fig. 1** LCZ classification scheme including (a) 10 built classes (i.e., LCZs 1–10) and (b) 7 land cover classes (i.e., LCZs A–G). Panels (a) and (b) are adapted from Fig. 1 in Demuzere *et al.*<sup>47</sup>; representative aerial maps (© Google Earth 2020) for each LCZ type are shown. (c) The spatial-temporal distribution of LCZ TAs across CONUS. The colors of TAs are consistent with the color palette presented in panels (a) and (b). (d) Counts of (1) total TAs by year (red line), (2) TAs by LCZ class in 2020 (black solid bars), and (3) sampled points within TA polygons by class for each year (black stripe bars) for model development.

approach above. To improve the balance among all LCZ categories, after requiring 70% agreement among MTurk workers, we manually checked satellite imagery on Google Earth to add additional TAs for under-sampled LCZs.

To expand the temporal coverage of TAs from 1986 to 2020, we developed a manual labeling procedure to add temporal labels to the single year crowdsourced labels. Specifically, we manually examined the valid MTurk labels and added labels for other years by comparing historical satellite imagery using the Google Earth time slider and assessing when the land cover or urban form changed (e.g., due to development). The imagery in Google Earth was collected through various data providers and platforms. Hence, the temporal and spatial resolution of historical imagery available in Google Earth varies largely depending on the source data and TA location. Typically, less historical imagery is available in early years due to the limited availability of satellite imagery sources. We used three steps to ensure the quality of the TAs. First, when the expert's opinion differed from the consensus of MTurk votes (i.e., over 70% of agreement), the experts would provide their classification as well as their confidence level (high/medium/low) in that classification. Among 4,422 valid MTurk TAs, the expert agreed with 83% of classifications for the base year and marked them as high confidence. Second, when examining historical satellite imagery for each TA, if the land cover or urban form changed significantly and the expert believed that the LCZ category should be changed to another LCZ type, they either stopped historical labeling for that TA to ensure accuracy of the TA or labelled it to another LCZ class with a confidence level. Third, we only considered TAs that had high confidence for model development. TA labels from 1986–2020 were compiled after the hybrid labeling process. Due to the availability of historical aerial maps in early years (e.g., gaps during some early years from Google Earth) and our approach to stop historical labeling when the TA substantially changed, fewer TAs were available in early years (e.g., 1986–1995) for model development (Fig. 1d). This effect is also apparent (although less so) for TAs during 2017–2020 mainly due to changes (e.g., development) that happened after the base year.

**Cloud-enabled model development.** The WUDAPT protocol uses a pixel-based random forest modeling approach<sup>51</sup>. Other work showed that a contextual RF classifier, which incorporates neighborhood information, can significantly increase model accuracy (e.g., as high as 13% increase in overall accuracy)<sup>52</sup>. Some recent studies developed and applied more advanced algorithms such as convolutional neural networks [CNN] and recurrent residual network [Re-ResNet] for LCZ classification<sup>49,53–56</sup>. Algorithm comparisons have shown that CNN can significantly outperform RF<sup>55,56</sup>. For example, Rosentreter *et al.* compared CNN, pixel-based RF, and contextual RF for LCZ classification in multiple cities. In general, CNN achieved the highest model performance, with an average increase in overall accuracy of 16.5% and 4.8% compared to pixel-based RF and contextual RF<sup>55</sup>. Yoo *et al.* found that the improvement of model performance by incorporating the neighboring area of a focus pixel is significant, not only among RF models but also among CNN models<sup>56</sup>. However, deep learning algorithms require more intense computation resources than lightweight RF models. Our preliminary analyses showed that applying a deep learning model is still a difficult task for up-scaling to our target scale (i.e., national scale over 35 years) even with the aid of the cloud-computing GEE platform. For example, CNN requires input images with



**Fig. 2** Model development workflow. (a) TA polygon labeling including crowdsourced labeling using the MTurk platform for the base year (2017) and manual labeling by experts for other years (1986–2020). (b) Balanced label sampling from TA polygons and model feature extraction including features collected from yearly Landsat composite layers, LULC layers (LCMAP and LCMS), and Census layers. (c) Model training, fine-tuning, and evaluation on a local machine using Python scikit-learn library. (d) Model transfer from the local machine to the GEE platform, post-classification processing, and final LCZ prediction surfaces at 100 m  $\times$  100 m resolution from 1986–2020 across CONUS.

high enough resolution to achieve deep layers due to the convolution, pooling or drop out techniques. Using 10 m satellite images (a higher resolution would be even better) for instance, our preliminary analyses showed that the image file size for a single year would exceed two Terabytes, requiring huge computation resources for image processing and model application. In our study, we aim to develop lightweight contextual RF models for large-scale time-series LCZ mapping based on a hybrid model development pipeline (Fig. 2). The choice of the lightweight contextual RF algorithm as well as the local-cloud hybrid modeling pipeline were mainly driven by the feasibility for large scale LCZ mapping.

Since an imbalanced training set generally hinders the performance of a machine learning model, we adjusted our label sampling strategy accordingly. As shown in Fig. 1d, the number of TAs collected during labeling varied among LCZ types. For example, we obtained far more LCZ 6 (open lowrise) as compared to other categories and few TAs for LCZ 4 (open highrise) and LCZ F (bare soil or sand). To assemble a balanced training dataset for model development, we undersampled from LCZ 6 TAs and oversampled from other LCZ TAs so that each year and each LCZ category had the same amount of training samples (i.e., 400 sample points per year per class). In addition, to improve model performance, we only sampled from a 200 m  $\times$  200 m area near the centroid of the TA polygons instead of the entire 500 m  $\times$  500 m polygon based on our sensitivity analysis (see Supplementary Information: **Sensitivity analysis for sampling regions** and Table S1).

After sampling from TA polygons, we extracted model predictors through the GEE platform (Fig. 2b). Due to the lengthy study period (i.e., 1986–2020), only a limited number of datasets are available for extracting predictor features over the entire period. Our models used three types of input feature layers: (1) yearly Landsat composite imagery, (2) existing yearly LULC layers including LCMAP (Land Change Monitoring, Assessment, and Projection)<sup>57</sup> and LCMS (Landscape Change Monitoring System)<sup>58</sup>, and (3) Census data (i.e., total population and population density at the Census Block Group level)<sup>59</sup>. The Landsat composite imagery was derived from Landsat imagery datasets retrieved via GEE platform<sup>50</sup>. We also included year as a dummy variable. For Landsat variables, we extracted features from 12 bands, including 7 Landsat bands and 5 derived spectral indices, such as BCI (Biophysical Composition Index) and NDBAI (Normalized Difference BAREness Index)<sup>47</sup>. We also included 16 land cover bands from LCMAP and 22 land cover/land use bands from LCMS. More details can be found in Supplementary Information Table S2. For our contextual RF models, we collected contextual features for both Landsat and LULC features. Specifically, we collected features within a 7  $\times$  7 window for each sampling point, which was equivalent to a 210 m  $\times$  210 m neighboring area patch since both Landsat and LULC layers have a 30 m spatial resolution. For the Landsat variable, we collected 6 statistics of the input features within the sampling window: mean, max, min, median, the 25<sup>th</sup> percentile and the 75<sup>th</sup> percentile. For LCMAP and LCMS, we calculated the composition ratio for each land cover/land use type within the sampling window. For the Census layer, we directly collected the census values at the locations of samples given that the census block group size varies across regions. We also tested the introduction of another type of census data (i.e., segmented employment data) into the model. However, employment data was only available after 2002 thus was not considered for our final pipeline. Despite this, we report the details in Supplementary Information (**Incorporating segmented**

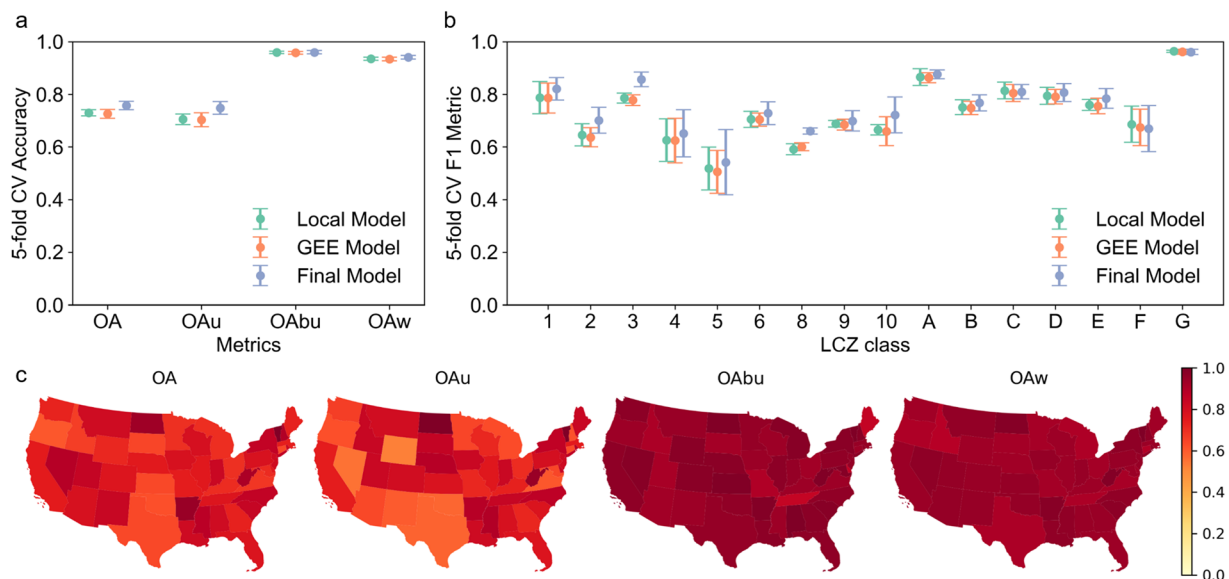
**employment data into the LCZ model** and Fig. S1) as employment data helped in distinguishing between certain urban classes and may be useful for future efforts.

After assembling the training set, we conducted two modeling steps: (1) fine-tuning a RF model on a local machine using the Python scikit-learn library<sup>60</sup> and (2) reproducing the RF model on the GEE platform<sup>50</sup> by applying the same hyperparameters for generating predictions. This workflow was designed to take advantage of both the efficiency of fine-tuning hyperparameters on the local machine (hyperparameter fine-tuning on GEE is relatively trickier) and the power of cloud computation for model prediction, i.e., large scale LCZ mapping. During the local model fine-tuning, we used cross validation to evaluate model performance (see more details in **Methods for model evaluation**). Once we achieved the best local model performance, we applied the same hyperparameters to build a GEE model and generated the preliminary prediction surfaces across CONUS at 100 m spatial resolution. It's worth noting that the RF parameters by the scikit-learn and GEE package have different names, but some serve the same function (e.g., `n_estimators` by scikit-learn vs. `numberOfTrees` by GEE). Important hyperparameters for our final RF model included 50 decision trees, a minimum of 4 samples to create a leaf node, etc. Furthermore, to reduce the “salt-and-pepper” effect (e.g., misclassification of individual pixels into disparate classes despite their similarity to neighboring pixels), we applied post-classification processing to further improve the model results. We adopted the spatial Gaussian filter developed by Demuzere *et al.*<sup>47</sup> Briefly, this approach applies Gaussian kernels with LCZ class-dependent standard deviation ( $\sigma_i$ ) and kernel size ( $>2\sigma_i$ ) to reduce noise in the predictions. Specifically, we chose  $\sigma_i$  values of 100 m for LCZ 1, 250 m for LCZs 8 and 10, 150 m for the rest of urban classes; we chose 150 m for LCZ E, 25 m for LCZ G and 100 m for other natural classes<sup>47</sup>. As mentioned in Demuzere *et al.*<sup>47</sup>, the choice of  $\sigma_i$  was derived by the priori knowledge from experts. For example, LCZ 1 zones across global cities are typically confined to rather small areas in a city. LCZs 8 and 10 are typically rather large zones. For other urban classes, 150 m is chosen as a scale that is slightly larger than the pixel resolution to ensure that single LCZ pixels do not constitute an LCZ class. For natural classes, the scales are smaller to make sure it is still possible to retain (often small scale) natural features within an urban environment, such as a river/canal, or a small urban park or grass field. Such prior knowledge deserves further investigation and adjustment in future studies and optimal  $\sigma_i$  are assumed to differ among regions<sup>35,47</sup>. After the Gaussian filter, we applied a temporal filter to improve temporal consistency among predictions. Specifically, for each pixel in our LCZ maps, we used a 5-year window (target year  $\pm 2$  year) to smooth any abrupt temporal change. The mode of the 5-year window was chosen as the result of the temporal majority vote. When multiple mode values were found, the original classification was kept for these edge cases.

**Methods for model evaluation.** We conducted 5-fold spatial cross validation for model evaluation in an effort to mitigate the impact of spatial autocorrelation. Spatial autocorrelation can inflate model test accuracy yet can lead to failure in model generalization and final LCZ mapping<sup>61</sup>. Previous studies have offered approaches to address this issue and reduce the impact of spatial autocorrelation in model performance evaluation by setting a minimum threshold for distance between two training samples, using polygon hold-out or city hold-out strategies to split training and testing sets<sup>55,56,62</sup>. In our study, we created 1000 m  $\times$  1000 m grids across CONUS with a unique geohash code for each spatial grid (detailed results of sensitivity analysis by grid size can be found in Supplementary Information: **Sensitivity analysis for spatial hold-out strategy** and Table S3). The LCZ training samples were then linked to the corresponding grid cells. We randomly divided the LCZ training samples into 5 groups based on the spatial geohash code and conducted 5-fold cross validation. This spatial hold-out method ensured that no samples collected from the same 1000 m spatial grid and TA polygon can be both in the train and test sets. Since LCZ TAs are 500 m  $\times$  500 m and we only sampled from the inner 200 m  $\times$  200 m areas of TAs, our spatial hold-out train-test splitting strategy is stricter than or at least as strict as the polygon hold-out method.

We reported several metrics for model comparison based on the 5-fold spatial cross validation, including the overall accuracy (OA), overall accuracy for urban classes (OA<sub>u</sub>), overall accuracy for built versus natural classes (OA<sub>bu</sub>), weighted accuracy (OA<sub>w</sub>), and class-wise F1 score. The F1 score is the harmonic mean of precision (i.e., the proportion of positive predictions that are actually correct) and recall (i.e., the proportion of actual positives that are correctly identified). The weighted accuracy was designed to account for the (dis)similarity between LCZ classes, such that misclassification between dissimilar LCZ classes (e.g., an urban class and a natural class) received more penalty than misclassification among similar classes (e.g., a midrise class and a lowrise class)<sup>63,64</sup>. To check temporal consistency, we reported normalized average transition rate for each LCZ category. Specifically, we calculated the number of pixels for each LCZ class in each year and created a confusion matrix for every two consecutive years across the modelled 35 years. We then averaged the confusion matrices to obtain yearly transition rates. For straightforward comparison, we normalize the yearly confusion matrix by the LCZ category.

In addition to traditional accuracy assessment, we also conducted a thematic benchmark against an external dataset, i.e., the National Land Cover Database [NLCD]<sup>65</sup>. The NLCD 2019 dataset provides both land cover classification and impervious surface information for 8 years during 2001 and 2019. Considering the various criteria for different classification schemes, we chose to conduct the thematic benchmark based on the impervious fraction. Stewart and Oke provided the range of impervious surface fraction for each LCZ category<sup>25</sup>. According to their values, all natural LCZ classes except LCZ E (bare rock or paved) have impervious surface fraction less than 10%, while all urban LCZ classes except LCZ 9 (sparsely built) are above 10%. In addition, the range of impervious surface fraction for many urban LCZ classes overlap with each other, making it difficult to directly compare impervious surface fraction. Hence, we converted both the NLCD and LCZ datasets into binary maps with a threshold of 10% impervious percent, considering pixels with more than 10% impervious as urban and the other as natural. For our LCZ dataset, since it is possible for LCZ 9 (sparsely built) to be either above or below the threshold, we removed the LCZ 9 pixels during thematic benchmarking. Assuming that the NLCD dataset reports the ground truth, we compared the two binary datasets for 8 years and report the overall



**Fig. 3** Model performance based on the 5-fold spatial cross validation. **(a)** Overall model performance by the local model, GEE model, and final model (i.e., GEE model with spatial-temporal post-classification processing). Reported metrics include the overall accuracy (OA), overall accuracy for urban classes (OA<sub>u</sub>), overall accuracy for built versus natural classes (OA<sub>bu</sub>), and weighted accuracy (OA<sub>w</sub>). **(b)** Class-wise model performance by the local model, GEE model, and final model. Reported metrics include OA, OA<sub>u</sub>, OA<sub>bu</sub>, and OA<sub>w</sub>.

accuracy (OA) and the F1 score for each class. We conducted the comparison not only for CONUS, but also for the urbanized areas across CONUS (abbreviated as CONUS-UA) based on the Census 2020 Urban Areas boundaries, since the urban areas are more of interest for the urban studies. Additionally, we compared against an existing CONUS-wide LCZ product that is available for 2017 by Demuzere *et al.*<sup>47</sup> Since this product doesn't include LCZ 7 and LCZ 9, we excluded pixels that are classified as LCZ 9 in our LCZ map from the corresponding year. Metrics including OA, OA<sub>u</sub>, OA<sub>bu</sub>, and OA<sub>w</sub> were calculated and reported.

### Data Records

Our CONUS time-series LCZ maps provide LCZ classification across the continental US spanning 35 consecutive years (1986–2020) with high spatial resolution (100 m). The projection of the LCZ product is USA Contiguous Albers Equal Area Conic (EPSG = 5070). The annual LCZ maps are delivered in the Geo TIFF file format and presented individually for each year. In addition to the final LCZ product, this study also provides LCZ maps of raw results without post-classification to allow data users the flexibility to employ different post-classification techniques as needed. The Training Areas are provided in the form of shapefiles. All LCZ maps and TAs are available via the figshare platform<sup>66</sup>. All the model input feature layers are publicly accessible. The Landsat and LCMS input layers are freely available through Google Earth Engine (<https://earthengine.google.com/>). The LCMAP layers are freely available through U.S. Geological Survey's Earth Resources Observation and Science (EROS) Center (<https://www.usgs.gov/special-topics/lcmap>). The Census data are freely available through the IPUMS National Historical Geographic Information System (NHGIS at <https://data2.nhgis.org/main>) and the Longitudinal Employer-Household Dynamics (LEHD at <https://lehd.ces.census.gov/data/#lodes>).

### Technical Validation

**Model accuracy assessment.** We developed different LCZ classifiers as part of our hybrid modeling pipeline: a local model, a GEE model, and our final model (i.e., the GEE model combined with spatial-temporal post-classification processing). Figure 3a shows the overall model performance at different stages of our modeling pipeline. As expected, the local and GEE models had very similar performance. The application of the spatial-temporal post-classification processing further increased the model performance by 0.03 for OA and 0.05 for OA<sub>u</sub>. In summary, our final LCZ model achieved 0.76 for the overall accuracy (OA), 0.75 for the overall accuracy of the urban classes (OA<sub>u</sub>), 0.96 for the overall accuracy of the built versus natural classes (OA<sub>bu</sub>), and 0.94 for the weighted accuracy (OA<sub>w</sub>).

To illustrate the spatial heterogeneity in model performance, we compared OA, OA<sub>u</sub>, OA<sub>bu</sub>, and OA<sub>w</sub> by state using our final LCZ model (Fig. 3c). The state-level OA varies from 0.61 to 0.99, with the lower OA generally occurring in central CONUS. State-level OA<sub>u</sub> shows the largest variation among the four overall accuracy metrics, ranging from 0.53 to 1.00. For OA<sub>bu</sub> and OA<sub>w</sub>, all states show high model performance (i.e., 0.85–1.00 for OA<sub>bu</sub>, 0.89–1.00 for OA<sub>w</sub>). To investigate the impact of the spatial distribution of TAs on model performance, we summarized the counts of TAs for each state (Fig. S2). The distribution of TAs is highly right skewed. States

		Predicted Label															Recall	F1		
		1	2	3	4	5	6	8	9	10	A	B	C	D	E	F			G	
Ground Truth	1	11900	960	195	299	8	162	201		27	143							9	0.86	0.82
	2	1379	10530	752	254	162	41	628		144					11			4	0.76	0.70
	3	58	1068	12158	30	152	256	50		36	1	79			2	2			0.88	0.86
	4	1159	713	222	8151	2051	998	369	83	51								17	0.59	0.66
	5	53	1082	616	1944	6490	3304	273	116			8				2	39		0.47	0.55
	6		53	341	94	332	12052	185	626	61	31	15	11	1	1	15	9		0.87	0.73
	8	117	1252	187	131	180	886	9172	491	1021		1	1	48	286	53			0.66	0.66
	9				1	9	1006	11	10478	7	772	447	274	520	8	279	67		0.75	0.70
	10	301	314	21	115	10	308	1454	1033	9592	47	1	23	37	243	238	133		0.69	0.72
	A	54						4	546		12515	565	40	15		5	98		0.90	0.88
	B				8	23	68	13	1112		892	10631	341	751		100	3		0.76	0.77
	C						1	10	143	23	96	1077	11464	271	196	580			0.83	0.81
	D			1			1		971		121	654	255	11671		113	85		0.84	0.81
	E	41	174	19	39	73	185	1303	104	1187	56	108	217	1	9898	249	54		0.72	0.78
F	3	15	15		5	101	258	344	569	14	111	1805	1748	878	8001	92		0.57	0.68	
G	3			6	7	21	4	150	50	49	72	1	20	10	77	13367		0.97	0.96	
Precision		0.79	0.65	0.84	0.74	0.68	0.62	0.66	0.65	0.75	0.85	0.77	0.79	0.77	0.86	0.82	0.96			

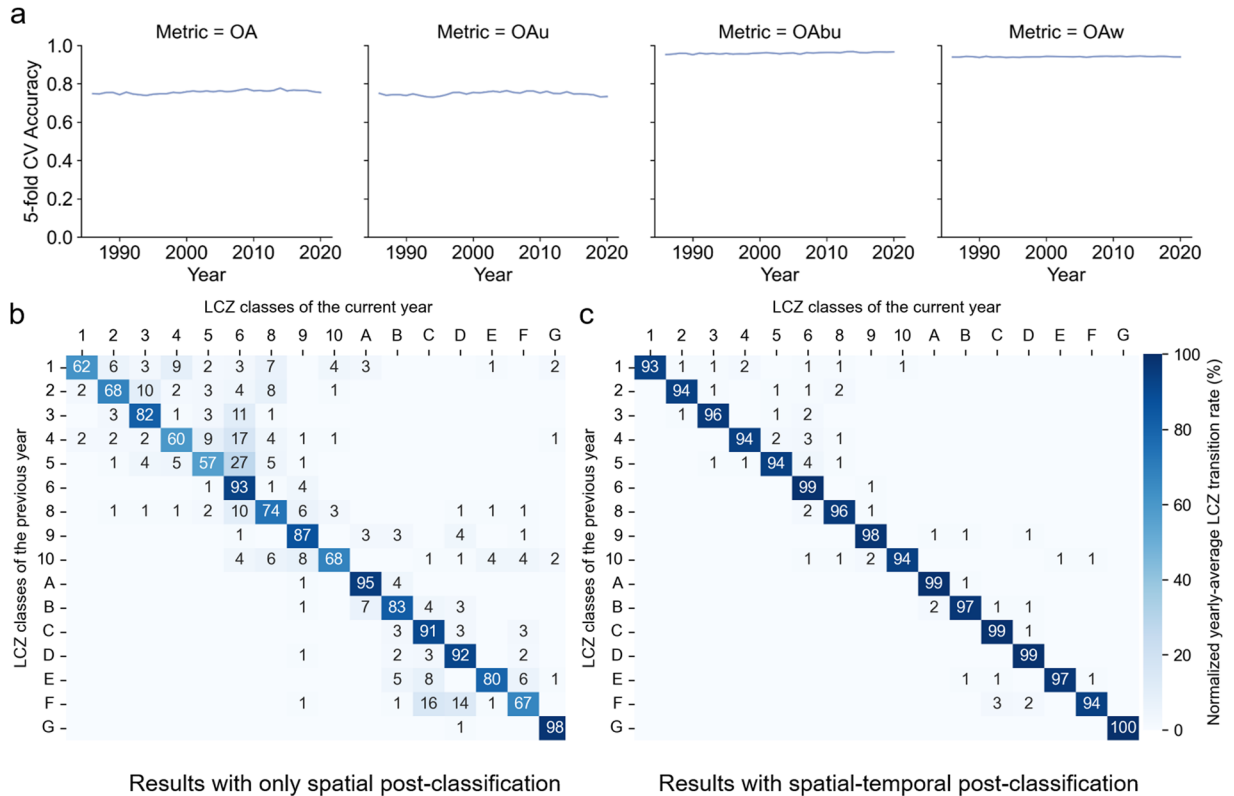
**Fig. 4** Confusion matrix for the final LCZ model. The value in each cell is the number of test samples. Precision is shown to the bottom of the confusion matrix. Recall and F1 scores are shown to the right of the confusion matrix.

such as California, Texas and Illinois have more TAs than other states. However, the number of TAs doesn't show significant relationship with  $OA_u$ ,  $OA_{bu}$ , or  $OA_w$  (Fig. S2c,d), indicating the robustness of our model.

Our final LCZ classifier showed a range of performance by LCZ class (Fig. 3b). Natural LCZ classes (5-fold CV F1 metric: 0.68–0.96) generally had better performance than urban classes (F1: 0.55–0.86). The confusion matrix for all LCZ classes is shown in Fig. 4. LCZ 3 (compact lowrise; F1: 0.86) and LCZ 1 (compact highrise; F1: 0.82) achieved the best performance among the urban classes. LCZ 5 (open midrise; F1: 0.55) and LCZ 4 (open highrise; F1: 0.66) performed worst, which is an issue also identified in other studies<sup>47,48,67</sup>. A potential reason is that many urban classes (e.g., LCZs 4 and 5) share similar surface cover fractions with their main difference being building height. Our model inputs lack features that can distinguish among building heights which may result in misclassification to similar LCZ types. For example, LCZ 5 (open midrise) was misclassified mostly as LCZ 6 (open lowrise) and LCZ 4 (open highrise); LCZ 4 (open highrise) was misclassified mostly as LCZ 5 (open midrise) (Fig. 4). This misclassification may or may not be significant depending on the application, however in general, misclassifying between two similar classes (e.g., compact highrise and compact midrise) is considered less problematic than misclassifying between dissimilar categories (e.g., compact highrise and heavy industry). For the natural classes, LCZ G (water) and LCZ A (dense trees) performed best with F1 scores of 0.96 and 0.88, respectively. LCZ F (bare soil or sand) showed relatively bad performance and was misclassified mainly as LCZ C (bush, scrub) and LCZ D (low plants). One explanation for this result may be the phenological change where certain areas may have varying LCZ classes by season (e.g., dry and wet seasons). Our model misclassification is primarily among similar class types (e.g., LCZ 5 misclassified as LCZs 6 and 4; LCZ 4 misclassified as LCZ 5; LCZ F misclassified as LCZs C and D). These misclassifications generally have high weighted accuracy<sup>63,64</sup> (i.e., larger than 0.8) thus are less problematic.

We also examined model performance by year (Fig. 5). Overall accuracy metrics were consistent over time. The WUDAPT LCZ classifier includes a spatial post-classification step. Since our objective is to create time-series LCZ maps, we introduced a temporal filter to reduce abrupt LCZ changes in consecutive years for any single pixel. Compared to the LCZ model that only applied a Gaussian spatial filter, adding a temporal filter further increased  $OA$  and  $OA_u$  by 0.01. More importantly, the temporal consistency of our time-series LCZ maps





Results with only spatial post-classification

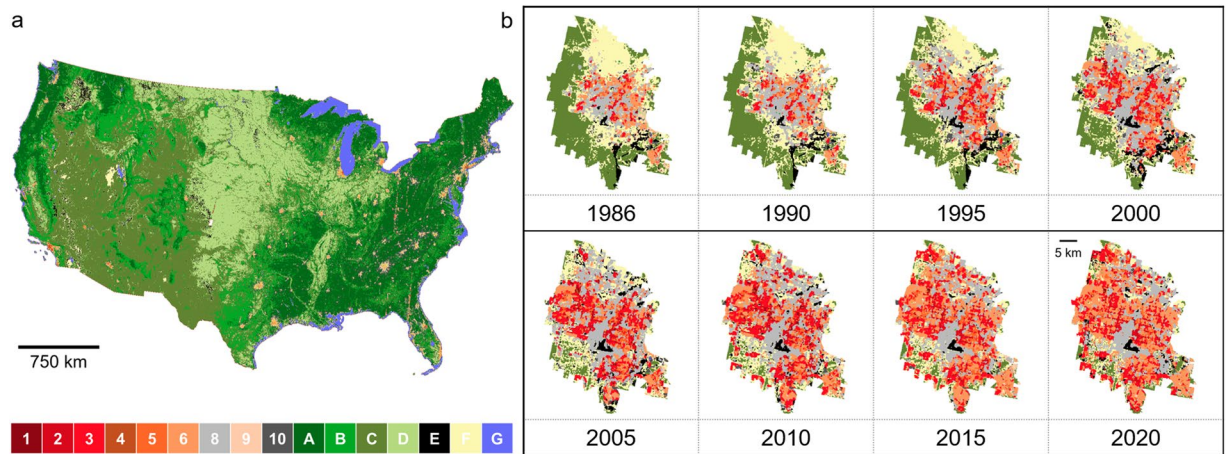
Results with spatial-temporal post-classification

**Fig. 5** Temporal consistency for the final LCZ model. **(a)** Overall model performance from 1986 to 2020, including the overall accuracy (OA), overall accuracy for urban classes ( $OA_u$ ), overall accuracy for built versus natural classes ( $OA_{bu}$ ), and weighted accuracy ( $OA_w$ ). **(b,c)** show the average yearly transition (unit: %) among LCZ categories from 1986 to 2020 for CONUS. **(b)** shows the transition for the GEE model with only spatial filtering. **(c)** shows the transition of our final LCZ model, i.e., the GEE model with both spatial and temporal filtering. The transition direction is from rows to columns. For example, in **(c)** row 1, 93% of LCZ 1 pixels in CONUS remained LCZ 1 from a previous year to the current year on average.

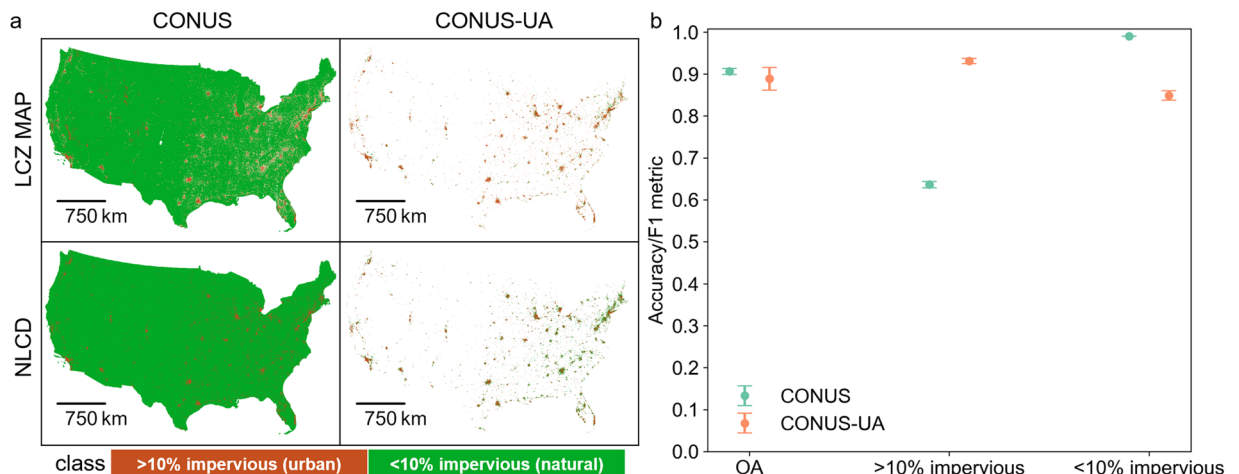
significantly improved. As shown in Fig. 5b,c, the average yearly LCZ transition became more consistent after temporal smoothing. This consistency is expected since land use types change slowly. Some model misclassification persists, e.g., Fig. 5c shows 1% of LCZ 1 pixels across the CONUS converted to LCZ 2 from a previous year to the current year on average, which is unlikely and may be from model misclassification.

In addition to model accuracy assessment, we compared the feature importance scores among different model input features. As shown in Fig. S3, the Landsat features contributed most to our LCZ model. The sum of the feature importance scores for Landsat, LCMS, LCMAP, and Census predictors are 0.54, 0.18, 0.17, and 0.09, respectively. In terms of the individual contribution from a model predictor, population density, BCI, and Landsat shortwave infrared 1 surface reflectance were the three most important features. Compared to solely relying on Landsat imagery, our modeling approach that incorporates LULC time-series features is able to improve the model temporal consistency since the LULC layers are usually derived from several change-detection algorithms<sup>58</sup>. Constrained by the data availability throughout our entire study period, very limited auxiliary data (i.e., LCMAP, LCMS and Census data in this work) is available for model development and improvement. The lack of suitable features that include information on building heights and other urban structure characteristics is one of the reasons for confusion between some urban categories. The inclusion of total population partially mitigated this issue since the distribution of population is closely related to density of buildings. Our attempt to include segmented employment data into the model also showed further improvement in accuracies over urban classes. However, it's worth noting that the total population is not available for each year and interpolation was used for non-decennial years, which can be one factor of model uncertainty.

**Longitudinal LCZ maps for CONUS.** We applied our final LCZ model to make prediction surfaces with high spatial resolution (i.e., 100 m × 100 m) for CONUS from 1986 to 2020. Figure 6a shows the LCZ map for 2020 for illustration. The spatial pattern of our LCZ maps is generally consistent with a previously published CONUS LCZ map representative for the year 2017 with an overall agreement of 0.73<sup>47</sup>. Specifically, we achieved 0.73 for OA, 0.87 for  $OA_u$ , 0.99 for  $OA_{bu}$ , and 0.95 for  $OA_w$ , when compared to this single-year LCZ map (see **Methods for model evaluation**). To our knowledge, this study is the first that develops a LCZ dataset which spans 35 years across CONUS. Our dataset may enable various longitudinal analyses (i.e., from 1986–2020) including urban sprawl analyses, urban heat analyses, urban climate and weather modeling, etc. For example,



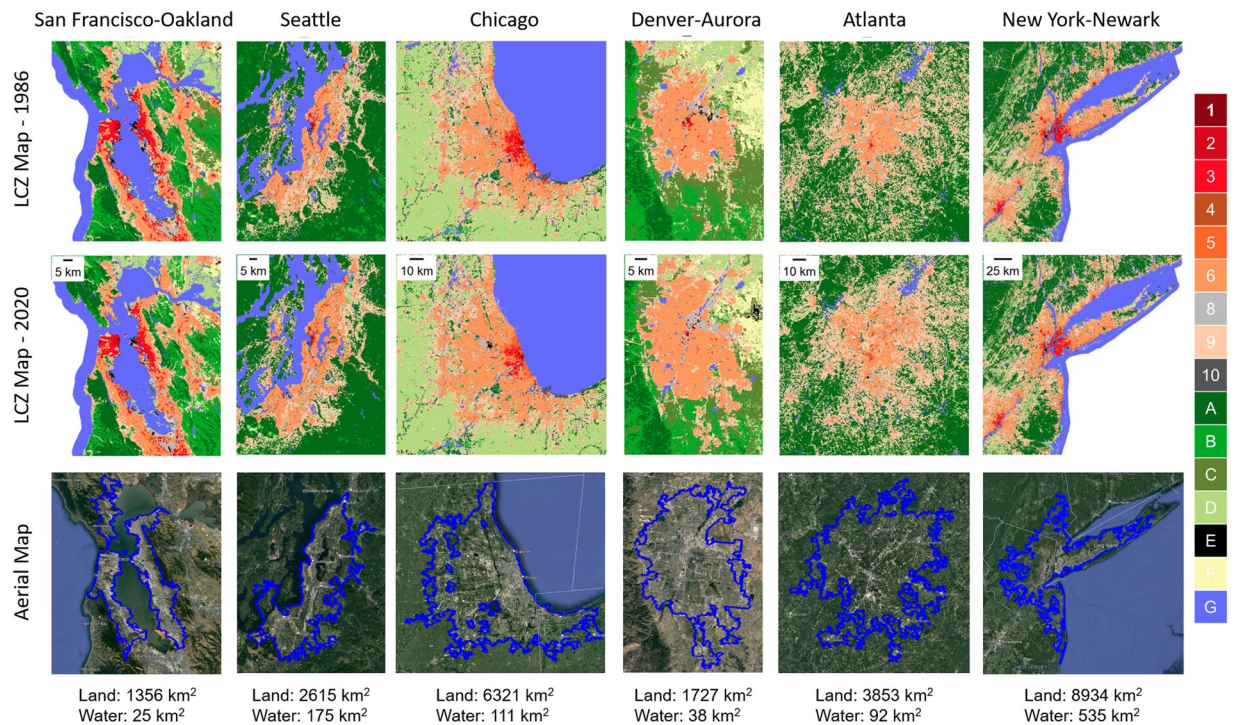
**Fig. 6** LCZ mapping. **(a)** CONUS prediction surface at  $100\text{ m} \times 100\text{ m}$  resolution using year 2020 for illustration. **(b)** LCZ mapping for an example city (Las Vegas, Nevada) at  $100\text{ m} \times 100\text{ m}$  resolution by year. For illustration, change in urban form is shown based on a  $\sim 5$ -year interval and the Census 2020 Urban Areas boundary. Full results for all years (1986 to 2020) are available in our LCZ dataset.



**Fig. 7** Thematic benchmarking. **(a)** The binary (urban vs. natural) LCZ and NLCD maps using 2019 for illustration. The threshold for defining urban vs. natural was 10% impervious surface fraction. LCZ 9 (sparsely built) was not considered in the thematic benchmarking because the impervious fraction of LCZ 9 pixels may fall into either category. The comparison was conducted for both CONUS, and CONUS-UA which used Census 2020 Urban Areas boundary. **(b)** Comparison results for overall accuracy and the urban and natural pixels.

Las Vegas is well-known as one of the fastest growing cities in the US. The LCZ maps extracted from our dataset clearly show the expected urban expansion of Las Vegas (Fig. 6b). For illustration, we used the 2020 Las Vegas urbanized area boundary to calculate the proportion of urban pixels within the boundary for each year. In 1986, only 29% of the pixels were classified as urban LCZ classes; in 2020 the proportion increased to 81%. The proportion of compact classes (i.e., LCZs 1–3) increased from 3% in 1986 to 21% in 2020 and from 11% to 36% for open urban classes (i.e., LCZs 4–6).

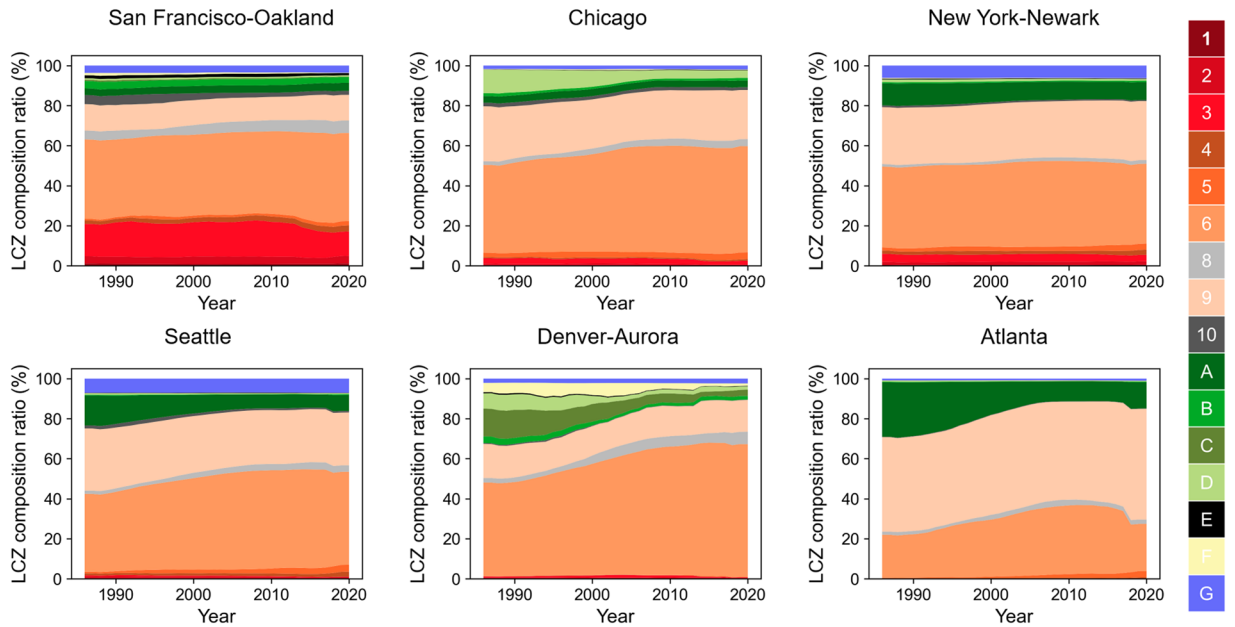
We conducted thematic benchmarking using the National Land Cover Database for 8 available years based on the impervious fraction. We created binary variables for both the LCZ and the NLCD maps into two classes (urban vs. natural) using a threshold of 10% impervious surface fraction. Since there are far more natural pixels than urban pixels across the entire CONUS, we compared the two datasets not only for CONUS but also for only the urban areas within CONUS [CONUS-UA]. As shown in Fig. 7b, the overall consistency between the two datasets was satisfactory, with average OAs of 0.91 and 0.89 for CONUS and CONUS-UA, respectively. The class-wise F1 scores across CONUS were 0.63 for the urban pixels (i.e., > 10% impervious fraction) and 0.99 for the natural pixels (i.e., < 10% impervious fraction). The relatively lower performance of the urban class was attributed to the discrepancies between the NLCD and LCZ datasets in identifying some urban pixels, especially in regions outside the CONUS-UA. However, within CONUS-UA, both urban and natural classes performed well, with F1 scores of 0.93 and 0.85, respectively. These results indicate that our LCZ maps achieved satisfactory and robust consistency to an independent, high-quality national land cover dataset.



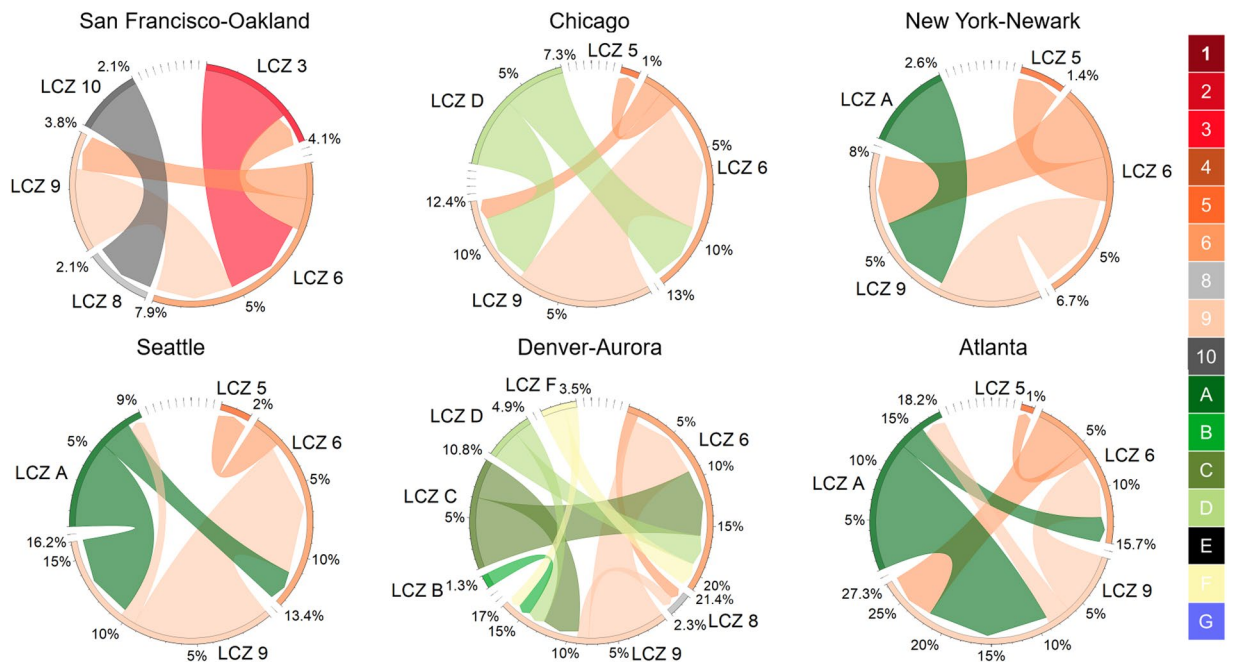
**Fig. 8** LCZ maps for 6 example metropolitan areas (two cities each for the western, central, and eastern US) in 1986 and 2020 with corresponding aerial maps from Google Earth (© Google Earth 2020). Blue lines in the aerial maps are the Census 2020 Urban Areas boundaries. Land and water area within the boundaries are listed for reference.

**LCZ distribution and change mapping.** To further validate our product, we investigated additional metropolitan areas in the US and mapped their LCZ distributions and land use change over time. Specifically, we used 6 metropolitan areas in the US (i.e., San Francisco-Oakland, Seattle, Chicago, Denver-Aurora, Atlanta, and New York-Newark) to summarize the LCZ distribution and change over 35 years. We calculated the proportion of pixels for each LCZ category within the boundary of each metropolitan area. The transition among LCZ categories were tracked between 1986 and 2020. We calculated the absolute number of pixels that were LCZ X in 1986 and turned to LCZ Y in 2020 (e.g., pixels that were LCZ 9 in 1986 and converted to LCZ 6 in 2020). We then divided by the total number of LCZ pixels in 1986 for normalization. For consistency, all calculations were based on pixels within the Census 2020 Urban Areas boundaries. Figure 8 shows the LCZ maps of the example cities using 1986 and 2020 for illustration. The aerial maps show the Census 2020 Urban Areas boundaries. The land and water area are listed alongside for context when comparing among cities. By comparing the LCZ distribution pattern between 1986 and 2020, Denver-Aurora and Atlanta show the most obvious urban sprawl patterns.

As shown in Fig. 9, 10, our longitudinal LCZ dataset is capable of measuring change over time of LCZ composition and tracking where major transitions occurred. Among the 6 example cities, the fastest urban expansion (i.e., large increases in urban LCZ ratios) occurred in Denver-Aurora. In 1986, natural LCZ classes covered 32% of the Denver-Aurora area – mostly LCZ C (bush, scrub), LCZ D (low plants), and LCZ F (bare soil or sand). During urban expansion across 35 years, the proportion of natural classes continuously shrunk, and the composition ratio of the natural classes decreased to 11% in 2020 (Fig. 9). The LCZ type that grew the most was LCZ 6 (open lowrise) with a 20% increase in composition ratio – 8% was converted from pixels of LCZ 9 (sparsely built), 7% from LCZ C (bush, scrub), 3% from LCZ D (low plants) and 2% from LCZ F (bare soil or sand) (Fig. 10). Similarly, Atlanta, Chicago, and Seattle also showed signs of urban sprawl. Like Denver-Aurora, both Chicago and Seattle had the most growth in LCZ 6 (open lowrise) with an increase of 9% and 8% in area percentage, respectively. The major transition was both from LCZ 9 (sparsely built) to LCZ 6 (open lowrise). In Atlanta, the major source of urban sprawl came from deforestation and forest fragmentation<sup>68</sup>. In 1986, LCZ A (dense trees) were 27% of the city area compared to only 13% in 2020. Most of the dense trees lost were converted to LCZ 9 (sparsely built). Another main transition was that 6% of pixels were converted from LCZ 9 (sparsely built) to LCZ 6 (open lowrise) demonstrating the gradual urban development and expansion in Atlanta. We also observed an abrupt drop in the composition ratio for LCZ 6 (open lowrise) in 2018. This might be a result of misclassification from LCZ 6 (open lowrise) to LCZ 9 (sparsely built). The other two example cities (San Francisco-Oakland and New York-Newark) showed relatively slower urban expansion during the study period. This could be partially explained by major urban expansion in these two urban agglomerations before 1986. Still, we found 3% pixels converted from LCZ 9 (sparsely built) to LCZ 6 (open lowrise) and 2% pixels converted from LCZ 10 (heavy industry) to LCZ 8 (large lowrise) in San Francisco-Oakland from 1986 to 2020.



**Fig. 9** Trends of LCZ composition ratio from 1986 to 2020 for 6 US metropolitan areas. All composition ratios were calculated based on the Census 2020 Urban Areas boundary shown in Fig. 8.



**Fig. 10** The chord diagram to show the major transition of LCZ classes between year 1986 and year 2020 for 6 US metropolitan areas. The directions of the transitions are denoted by the arrows. The colors of the chord are consistent with the colors of the transition sources. The ticks on the arcs represent the transition rate. For example, as shown by the LCZ C arc in Denver-Aurora, 7% of pixels transit from LCZ C to LCZ 6 (chord color consistent with LCZ C); 4% from LCZ C to LCZ F (chord color consistent with LCZ C); 0% from other LCZ types to LCZ C. To illustrate major transitions, any LCZ transition that occurred in less than 1% of city pixels were removed from the chord diagram; pixels remaining the same type were also removed. Full transition details can be found in Supplementary Information Fig. S4. All transition rates were calculated based on the Census 2020 Urban Areas boundary shown in Fig. 8.

In New York-Newark, 3% pixels converted from LCZ 9 (sparsely built) to LCZ 6 (open lowrise) and 3% from LCZ A (dense trees) to LCZ 9 (sparsely built).

Figure S5 shows more localized LCZ mapping using the New York County and San Francisco County for illustration. Satellite imagery for sample TAs in these two regions are also shown to demonstrate the prediction results versus the real labels. Overall, our crowdsourced-expert hybrid labeling was scalable and efficient. The local-cloud hybrid modeling approach leveraged the speed of fine-tuning model hyperparameters on a local machine while leveraging the power of model transfer and reproduction on a cloud platform (GEE) for making predictions at a large scale. Our lightweight LCZ classifier achieved good model performance (e.g., 0.76 for overall accuracy and 0.94 for weighted accuracy) and showed good consistency with NLCD (an external high-quality land cover dataset) for measuring the footprint of built-up areas. In summary, our modeling framework could be applied to other areas for large-scale longitudinal LCZ mapping. Our LCZ dataset has potential to support a wide range of research fields, such as urban weather and climate modeling, urban expansion and development, risk analysis of various urban hazards, etc. Moreover, our longitudinal LCZ dataset can serve as a valuable resource for urban planners, policy makers, and researchers to assess local, regional, and national policies related to urbanization, supporting informed decision-making for sustainable urban development.

### Code availability

Python scripts for training data sampling, earth observation and census input feature collection, random forest model fine tuning and mode prediction on Google Earth Engine are available at [https://github.com/QiMengEnv/CONUS\\_Longitudinal\\_LCZ](https://github.com/QiMengEnv/CONUS_Longitudinal_LCZ). All data processing and visualizations are done in Python 3.9. The post-classification processing is done in JavaScript on Google Earth Engine Code Editor and is also available with the same URL.

Received: 29 August 2023; Accepted: 3 February 2024;

Published online: 13 February 2024

### References

1. UN DESA. *World Urbanization Prospects: the 2018 Revision, Methodology*. (New York, 2018).
2. International Energy Agency. *World Energy Outlook 2021*. (IEA, 2021).
3. Qian, Y. *et al.* Urbanization impact on regional climate and extreme weather: Current understanding, uncertainties, and future research directions. *Adv. Atmos. Sci.*, 1–42 (2022).
4. Brondizio, E. S. *et al.* Re-conceptualizing the Anthropocene: A call for collaboration. *Global Environ. Change* **39**, 318–327 (2016).
5. Giles-Corti, B. *et al.* City planning and population health: a global challenge. *The Lancet* **388**, 2912–2924 (2016).
6. Henderson, V. The urbanization process and economic growth: The so-what question. *J. Econ. Growth* **8**, 47–71 (2003).
7. Li, W. & Yi, P. Assessment of city sustainability—Coupling coordinated development among economy, society and environment. *J. Clean. Prod.* **256**, 120453 (2020).
8. Liang, W. & Yang, M. Urbanization, economic growth and environmental pollution: Evidence from China. *Sustain. Comput. Inform. Syst.* **21**, 1–9 (2019).
9. Demuzere, M. *et al.* A global map of local climate zones to support earth system modelling and urban-scale environmental science. *Earth Syst. Sci. Data* **14**, 3835–3873 (2022).
10. Williams, K. *Urban Form and Infrastructure: A Morphological Review*. (Foresight, 2014).
11. Fathi, S. *et al.* The role of urban morphology design on enhancing physical activity and public health. *Int. J. Environ. Res. Public Health* **17**, 2359 (2020).
12. Ye, C. *et al.* Toward healthy and liveable cities: a new framework linking public health to urbanization. *Environ. Res. Lett.* (2022).
13. Krayenhoff, E. S., Moustauoui, M., Broadbent, A. M., Gupta, V. & Georgescu, M. Diurnal interaction between urban expansion, climate change and adaptation in US cities. *Nat. Clim. Change* **8**, 1097–1103 (2018).
14. Luo, M. & Lau, N. C. Urban expansion and drying climate in an urban agglomeration of East China. *Geophys. Res. Lett.* **46**, 6868–6877 (2019).
15. Akbari, H. & Kolokotsa, D. Three decades of urban heat islands and mitigation technologies research. *Energy Build.* **133**, 834–842 (2016).
16. Deilami, K., Kamruzzaman, M. & Liu, Y. Urban heat island effect: A systematic review of spatio-temporal factors, data, methods, and mitigation measures. *Int. J. Appl. Earth Obs. Geoinf.* **67**, 30–42 (2018).
17. Howard, L. *The Climate of London: Deduced From Meteorological Observations Made In The Metropolis And At Various Places Around It*. Vol. 3 (Harvey and Darton, J. and A. Arch, Longman, Hatchard, S. Highley [and] R. Hunter, 1833).
18. Kim, S. W. & Brown, R. D. Urban heat island (UHI) intensity and magnitude estimations: A systematic literature review. *Sci. Total Environ.* **779**, 146389 (2021).
19. Oke, T. R. The energetic basis of the urban heat island. *Q. J. R. Meteorol. Soc.* **108**, 1–24 (1982).
20. Peng, S. *et al.* Surface urban heat island across 419 global big cities. *Environ. Sci. Technol.* **46**, 696–703 (2012).
21. Zhao, L., Lee, X., Smith, R. B. & Oleson, K. Strong contributions of local background climate to urban heat islands. *Nature* **511**, 216–219 (2014).
22. Chapman, S., Watson, J. E., Salazar, A., Thatcher, M. & McAlpine, C. A. The impact of urbanization and climate change on urban temperatures: a systematic review. *Landsc. Ecol.* **32**, 1921–1935 (2017).
23. Li, D. & Bou-Zeid, E. Synergistic interactions between urban heat islands and heat waves: The impact in cities is larger than the sum of its parts. *J. Appl. Meteorol. Climatol.* **52**, 2051–2064 (2013).
24. Tuholske, C. *et al.* Global urban population exposure to extreme heat. *Proc. Natl. Acad. Sci. USA* **118**, e2024792118 (2021).
25. Stewart, I. D. & Oke, T. R. Local climate zones for urban temperature studies. *Bull. Am. Meteorol. Soc.* **93**, 1879–1900, <https://doi.org/10.1175/bams-d-11-00019.1> (2012).
26. Stewart, I. D. A systematic review and scientific critique of methodology in modern urban heat island literature. *Int. J. Climatol.* **31**, 200–217 (2011).
27. Brousse, O., Martilli, A., Foley, M., Mills, G. & Bechtel, B. WUDAPT, an efficient land use producing data tool for mesoscale models? Integration of urban LCZ in WRF over Madrid. *Urban Clim.* **17**, 116–134 (2016).
28. Hammerberg, K., Brousse, O., Martilli, A. & Mahdavi, A. Implications of employing detailed urban canopy parameters for mesoscale climate modelling: a comparison between WUDAPT and GIS databases over Vienna, Austria. *Int. J. Climatol.* **38**, e1241–e1257 (2018).
29. Patel, P. *et al.* Modeling large-scale heatwave by incorporating enhanced urban representation. *J. Geophys. Res.: Atmos.* **127**, e2021JD035316 (2022).
30. Varentsov, M., Samsonov, T. & Demuzere, M. Impact of urban canopy parameters on a megacity’s modelled thermal environment. *Atmosphere* **11**, 1349 (2020).
31. Jiang, S. *et al.* Mapping local climate zones: A bibliometric meta-analysis and systematic review. *OSF preprints*. p.1–106 (2021).
32. Quan, S. J. & Bansal, P. A systematic review of GIS-based local climate zone mapping studies. *Build. Environ.* **196**, 107791 (2021).

33. Xue, J., You, R., Liu, W., Chen, C. & Lai, D. Applications of local climate zone classification scheme to improve urban sustainability: A bibliometric review. *Sustainability* **12**, 8083 (2020).
34. Huang, F. *et al.* Mapping local climate zones for cities: A large review. *Remote Sens. Environ.* **292**, 113573 (2023).
35. Bechtel, B. *et al.* Mapping local climate zones for a worldwide database of the form and function of cities. *ISPRS Int. J. Geo-Inf.* **4**, 199–219 (2015).
36. Mills, G., Ching, J., See, L., Bechtel, B. & Foley, M. An introduction to the WUDAPT project. *Proc. 9th Int. Conf. Urban Clim.* 20–24 (2015).
37. Bechtel, B. *et al.* SUHI analysis using Local Climate Zones—A comparison of 50 cities. *Urban Clim.* **28**, 100451 (2019).
38. Brousse, O. *et al.* Using local climate zones in Sub-Saharan Africa to tackle urban health issues. *Urban Clim.* **27**, 227–242 (2019).
39. Danylo, O., See, L., Bechtel, B., Schepaschenko, D. & Fritz, S. Contributing to WUDAPT: A local climate zone classification of two cities in Ukraine. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **9**, 1841–1853 (2016).
40. Ren, C. *et al.* Assessment of local climate zone classification maps of cities in China and feasible refinements. *Sci. Rep.* **9**, 1–11 (2019).
41. Wang, R., Ren, C., Xu, Y., Lau, K. K.-L. & Shi, Y. Mapping the local climate zones of urban areas by GIS-based and WUDAPT methods: A case study of Hong Kong. *Urban Clim.* **24**, 567–576 (2018).
42. Xu, Y., Ren, C., Cai, M., Edward, N. Y. Y. & Wu, T. Classification of local climate zones using ASTER and Landsat data for high-density cities. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **10**, 3397–3405 (2017).
43. Demuzere, M., Kittner, J. & Bechtel, B. LCZ Generator: a web application to create Local Climate Zone maps. *Front. Environ. Sci.* **9**, 637455 (2021).
44. Khamchiangta, D. & Dhakal, S. Future urban expansion and local climate zone changes in relation to land surface temperature: Case of Bangkok Metropolitan Administration, Thailand. *Urban Clim.* **37**, 100835 (2021).
45. Vandamme, S., Demuzere, M., Verdonck, M.-L., Zhang, Z. & Van Coillie, F. Revealing kunming's (china) historical urban planning policies through local climate zones. *Remote Sens.* **11**, 1731 (2019).
46. Wang, R. *et al.* Detecting multi-temporal land cover change and land surface temperature in Pearl River Delta by adopting local climate zone. *Urban Clim.* **28**, 100455 (2019).
47. Demuzere, M. *et al.* Combining expert and crowd-sourced training data to map urban form and functions for the continental US. *Sci. Data* **7**, 1–13 (2020).
48. Demuzere, M., Bechtel, B., Middel, A. & Mills, G. Mapping Europe into local climate zones. *PLoS One* **14**, e0214474 (2019).
49. Zhu, X. X. *et al.* The urban morphology on our planet—Global perspectives from space. *Remote Sens. Environ.* **269**, 112794 (2022).
50. Gorelick, N. *et al.* Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **202**, 18–27 (2017).
51. Bechtel, B. *et al.* Generating WUDAPT Level 0 data—Current status of production and evaluation. *Urban Clim.* **27**, 24–45 (2019).
52. Verdonck, M.-L. *et al.* Influence of neighbourhood information on 'Local Climate Zone' mapping in heterogeneous cities. *Int. J. Appl. Earth Obs. Geoinf.* **62**, 102–113 (2017).
53. Liu, S. & Shi, Q. Local climate zone mapping as remote sensing scene classification using deep learning: A case study of metropolitan China. *ISPRS J. Photogramm. Remote Sens.* **164**, 229–242 (2020).
54. Qiu, C., Mou, L., Schmitt, M. & Zhu, X. X. Local climate zone-based urban land cover classification from multi-seasonal Sentinel-2 images with a recurrent residual network. *ISPRS J. Photogramm. Remote Sens.* **154**, 151–162 (2019).
55. Rosentreter, J., Hagensieker, R. & Waske, B. Towards large-scale mapping of local climate zones using multitemporal Sentinel 2 data and convolutional neural networks. *Remote Sens. Environ.* **237**, 111472 (2020).
56. Yoo, C., Han, D., Im, J. & Bechtel, B. Comparison between convolutional neural networks and random forest for local climate zone classification in mega urban areas using Landsat images. *ISPRS J. Photogramm. Remote Sens.* **157**, 155–170 (2019).
57. Rover, J. *et al.* *Land Change Monitoring, Assessment, and Projection*. Report No. 2327-6932, (US Geological Survey, 2020).
58. USDA Forest Service. USFS Landscape Change Monitoring System Conterminous United States version 2021-7 (Conterminous United States and Southeastern Alaska). (Salt Lake City, Utah, 2022).
59. Manson, S. M. IPUMS National Historical Geographic Information System: Version 15.0. (2020).
60. Pedregosa, F. *et al.* Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
61. Xu, C. *et al.* Application of training data affects success in broad-scale local climate zone mapping. *Int. J. Appl. Earth Obs. Geoinf.* **103**, 102482 (2021).
62. Qiu, C., Tong, X., Schmitt, M., Bechtel, B. & Zhu, X. X. Multilevel feature fusion-based CNN for local climate zone classification from sentinel-2 images: Benchmark results on the So2Sat LCZ42 dataset. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **13**, 2793–2806 (2020).
63. Bechtel, B. *et al.* Quality of crowdsourced data on urban morphology—the human influence experiment (HUMINEX). *Urban Sci.* **1**, 15 (2017).
64. Bechtel, B., Demuzere, M. & Stewart, I. D. A weighted accuracy measure for land cover mapping: comment on Johnson *et al.* local climate zone (LCZ) map accuracy assessments should account for land cover physical characteristics that affect the local thermal environment. *Remote Sens.* 2019, 11, 2420. *Remote Sens.* **12**, 1769 (2020).
65. Dewitz, J. & U. S. Geological Survey. National Land Cover Database (NLCD) 2019 products (ver. 2.0, June 2021). *U.S. Geological Survey* <https://doi.org/10.5066/P9KZCZM54> (2021).
66. Qi, M. *et al.* CONUS longitudinal local climate zone maps from 1986 to 2020, *Figshare*, <https://doi.org/10.6084/m9.figshare.c.6806736.v1> (2024).
67. Yokoya, N. *et al.* Open data for global multimodal land use classification: Outcome of the 2017 IEEE GRSS data fusion contest. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **11**, 1363–1377 (2018).
68. Miller, M. D. The impacts of Atlanta's urban sprawl on forest cover and fragmentation. *Appl. Geogr.* **34**, 171–179 (2012).

## Acknowledgements

This material is based upon work supported by the National Institutes of Health under grant No. R01 HL150119. Matthias Demuzere and Benjamin Bechtel were supported by the ENLIGHT project, funded by the German Research Foundation (DFG) under grant No. 437467569. The authors thank Dr. Andrew Larkin for his valuable suggestion on large-scale mapping and thank Griffin Kearns for his contribution in TAs collection.

## Author contributions

Meng Qi designed the research, performed the modeling, analysis, visualization, and wrote the original draft, with review and feedback from all the other authors. Chunxue Xu contributed to the conceptualization and modeling design. Wenwen Zhang and Matthias Demuzere contributed to the modeling design and analysis. Perry Hystad contributed to the conceptualization. Tianjun Lu contributed to the data collection. Peter James contributed to the conceptualization and analysis. Benjamin Bechtel contributed to the methodology. Steve Hankey contributed to the conceptualization, methodology design, analysis, visualization and supervision.

### Competing interests

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.

### Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41597-024-03042-4>.

**Correspondence** and requests for materials should be addressed to S.H.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024