

Urban Forests for Carbon Sequestration and Heat Island Mitigation

LEVENTE J KLEIN*, IBM TJ Watson Research Center, USA

CONRAD M ALBRECHT*, German Aerospace Center, Germany

Urban forests serve both as a carbon sequestration pool and heat island mitigation tool. Climate change will increase the frequency and severity of urban heat islands. Thus, new urban planning strategies demand our attention. Based on multimodal, remotely sensed data, we map the tree density, its carbon sequestered, and its impact on urban heat islands for Long Island, NY and Dallas, TX. Using local climate zones we investigate concepts of urban planning through optimized tree planting and adjusting building designs to mitigate urban heat islands.

Additional Key Words and Phrases: heat island mitigation, carbon sequestrations, urban planning, local climate zones, big geospatial data analytics, LiDAR, Landsat 8 surface temperature

ACM Reference Format:

Levente J Klein and Conrad M Albrecht. 2022. Urban Forests for Carbon Sequestration and Heat Island Mitigation. 1, 1 (June 2022), 8 pages. <https://doi.org/XXXXXXX.XXXXXXX>

1 INTRODUCTION & MOTIVATION

By 2050, more than 2/3 of the world's population will live in urban areas. It implies the relevance of adaptation to climate hazards, and to develop strategies for sustainable resource management [1]. Increased temperatures due to climate change require redesign of current city layouts, and it calls for rethinking the blueprints of future urban spaces [2]. Most urban areas historically grew over decades or even centuries—with buildings, roads, and green spaces developed based on social, economical, or safety considerations. Adaptation to climate change may require refurbishing some areas in cities, e.g. through improved urban forests layout and enhanced buildings designs. In this work, we demonstrate that exiting remote sensing analytics has capacity to implement identification and mitigation of urban heat islands to turn cities into more attractive living spaces.

The impact of heat islands in cities has been recognized in the past. Metrics were developed to identify the main driving factors [3]. A popular metric is *Local Climate Zones* (LCZ) that are most commonly based on visual inspections and interpretation of photographs [4]. The risk of errors in visual assessment was quickly realized, and the methodology did get automated to consistently generalize the assessments [5]. Although automated, supervised learning approaches still require significant effort to manually label images [6, 7]. To overcome manual efforts in image labeling, rapid label generation may get achieved by noisy and imperfect data to train machine learning models [8].

*Both authors contributed equally to this research.

Authors' addresses: Levente J Klein, kleinl@us.ibm.com, IBM TJ Watson Research Center, 1101 Kitchawan Rd, Yorktown Heights, NY, 10598, USA; Conrad M Albrecht, conrad.albrecht@dlr.de, German Aerospace Center, Muenchener Str. 20, 82234, Wessling, Germany.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2022 Association for Computing Machinery.

XXXX-XXXX/2022/6-ART \$15.00

<https://doi.org/XXXXXXX.XXXXXXX>

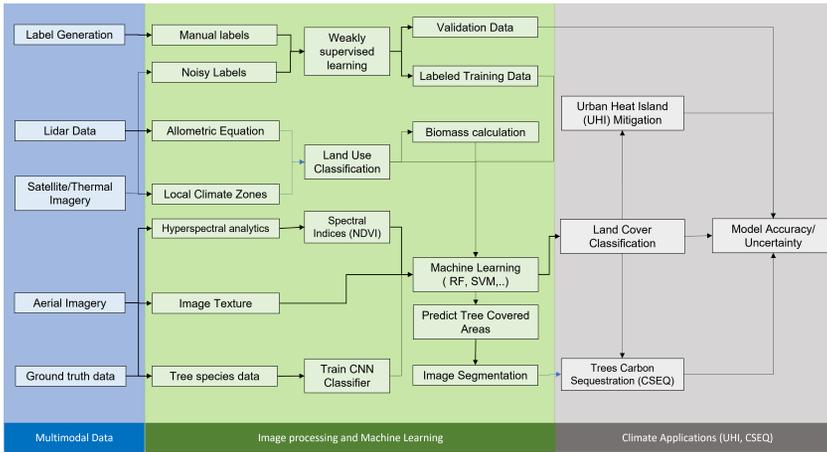


Fig. 1. Flow diagram with data processing and geospatial analytics to calculate carbon sequestered (CSEQ) in trees, and to determine the impact of land cover classes on Urban Heat Islands (UHI).

Terrain rich in vegetation is naturally cooler when compared to an area that is sealed by concrete and asphalt. Cooling the ambient temperature by vegetation is due to evapo-transpiration and tree shades. To identify the height and distribution of vegetation, Light Detection and Ranging (LiDAR) is a commonly used remote sensing approach. Access to high quality LiDAR data may not be always possible due to significant financial resources required to conduct surveys. Exploiting open source, non-classified point cloud data may be an alternative in combination with orthoimagery or satellite data. High quality LiDAR may always serve as a source to calibrate satellite images across small geospatial patches. After, the use of satellite imagery in other areas serves to detect and calibrate tree growth models [9]. To scale such models globally, multimodal data processing (LiDAR, satellite, orthoimagery, radar, and crowdsourcing data) is required to generate consistent maps across different geographies, and to identify the best available data for a given geo-location and analytics task [10].

In addition to heat island mitigation, trees serve as natural sink for atmospheric carbon dioxide when entrapped into the tree's biomass. The number of trees in urban areas constitute a significant storage of carbon with direct impact on a city's carbon footprint assessments. A major challenge in nature-based carbon sequestration is effective quantification [11]. There is ongoing interest to exploit satellite imagery to model carbon stored in trees and soils. Modeling is commonly based on dimensional scaling models where tree height is related to canopy diameter, and tree biomass gets correlated with carbon storage. These scaling models termed *allometric equations* are popular within the remote sensing community [12, 13]. Previous studies did demonstrate the utility of allometric equations [12] and the methodology got adopted to large geographies and eco-regions [14]. Uncertainty quantification in addition to estimates of carbon sequestered must be part of carbon sequestration modeling.

As the world's population continues to grow clustered in future mega-cities while, at the same time, the climate crisis unleashes its threats for life on Earth, we need to take immediate action in order to redesign our living spaces resilient to natural hazards. Our contribution paves the way to rethink urban planning in order to reduce our carbon footprint, and to mitigate the increased frequency of heat waves as projected by reports of the IPCC [15].

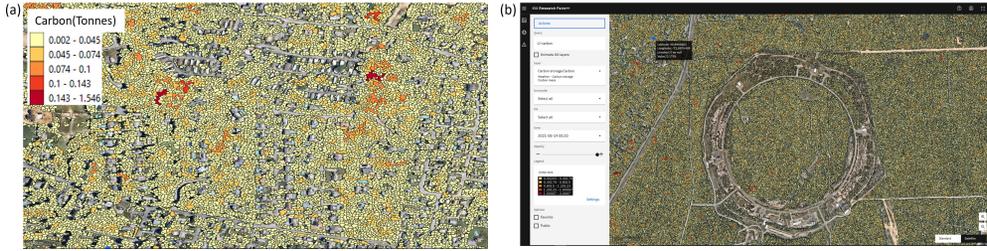


Fig. 2. (a) Carbon stored in trees for a residential area. Our analytics identified a total of about 36 million trees on all of Long Island, NY including Brooklyn and Queens, boroughs of New York City (NYC). We estimate the total carbon sequestered in trees to be $5.1 \cdot 10^6$ tonne. (b) A snapshot of the PAIRS graphical web user interface visualizing carbon sequestered in trees mapped out around Brookhaven National Laboratory's Relativistic Heavy Ion Collider.

2 METHODS & DATASETS

2.1 Multimodal Geospatial Platform

The data volume generated by satellites in combination with climate and weather modeling easily does exceed hundred of petabytes [16]. Many sets in the corpus of satellite and climate data is hosted by dedicated databases and cloud platforms. A well-known challenge of geospatial data processing poses the variability in spatial resolution, a plurality of geographic projections, and various geodata formats. Curation of training data for machine learning models provided a given geospatial area requires significant compute, and detailed domain and data knowledge drawn from corresponding metadata analysis. The PAIRS platform [10] was developed to overcome such challenges through a harmonized, nested spatial grid system enabling geo-data science at scale. At the same time, the geospatial platform minimizes the need of users to aggregate domain knowledge in order to operate [17, 18]. PAIRS has been demonstrated to process and to analyze all multimodal data presented in this work. It is tightly coupled to an AI engine for generation of training data required for tree delineation, tree species detection, and local climate zone classification. In Figure 1 we highlight the flow of data for processing by machine learning models.

2.2 Spatial Areas of Study

For tree distribution we analyze a combination of orthoimagery, LiDAR three-dimensional point clouds, and a Landsat 8 thermal band for Long Island, NY—including New York City, (NYC)—and for Dallas, TX. Our main objective is dedicated to identification of individual trees, calculate corresponding carbon sequestration, and to quantify the impact of tree density on urban heat islands.

2.3 Geospatial Data

High resolution aerial orthoimagery. The U.S. Department of Agriculture (USDA) provides 0.5m-resolution orthoimagery through the National Agriculture Imagery Program (NAIP) [19] every other year. Images get acquired during growing season—ideally suited to detect agricultural land and forestry. The imagery ships in 4 spectral bands: standard Red-Green-Blue (RGB) channels, and a Near InfraRed (NIR) band. All images are atmospherically corrected and do get adjusted for illumination.

Thermal imagery. The Landsat 8 satellite currently captures global thermal images at 100m spatial resolution. Thermal images measure surface reflectance expressed in units of Kelvin. Such may get

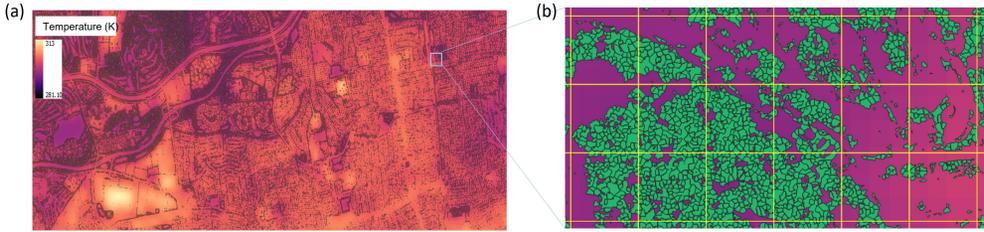


Fig. 3. Variation in surface temperature (a), and overlay of delineated trees (b) for an area on Long Island, NY. 100m-by-100m grid cells serve to assess tree impact on surface temperatures.

utilized to estimate local surface temperature. As common to all optical satellites, Landsat 8 data suffer disturbance by cloud cover obscuring direct observation of land surface.

Point cloud data. LiDAR data assembled in 2017 with a point density of about 10 points per m^2 enable high resolution, three-dimensional mapping. A simple, yet efficient point cloud data rasterization technique is implemented. It derives statistical features of the three-dimensional LiDAR point cloud that receives further processing through a rule-based classification algorithm. The reclassification serves as generator of *noisy labels*. Those proof appropriate to separate urban trees and buildings from other land cover classes [20]. Further, we utilize the LiDAR data as calibration reference for tree height estimates from orthoimagery.

Tree species survey. Tree species data was collected with the aid of crowdsourcing where tree location, tree species, tree health, and tree dimension values were measured or estimated based on visual inspections. This survey features more than 600,000 trees made available to the public domain including identification of more than 250 tree species [21]. New York City hosts more than 3 million trees in parks, as well as on public and private lands—but tree census captured a fraction of those, only. In geo-locations where there exists no tree species information, the trees delineated by a model fed with NAIP imagery get reclassified into dominant tree species [9].

2.4 Remote Sensing Analytics

To identify land cover, the Normalized Difference Vegetation Index (NDVI) is calculated from the Near Infrared (NIR) and Red bands of the NAIP imagery, $NDVI = (NIR - Red) / (NIR + Red)$. In order to separate trees from other lands, the source for training data either renders manual annotations or it stems from automated, noisy labels based on airborne LiDAR surveys. A Random Forest (RF) classifier is trained on NAIP data in accordance with image texture information. In consequence, the trained classifier will separate area covered by trees vs. other land surface types. Once trained, the RF has capacity to infer tree-covered areas from NAIP imagery. The NAIP-based NDVI value serves to segment the images into individual tree crowns using a Watershed Segmentation Algorithm [22]. Random seeding and area-growing computer vision methods applied to the NDVI-masked area delineate tree crowns. Image pixel grouping is based on pixel similarity. For tree species delineation, training data is created around the location of trees by cropping a 32-by-32 pixels field [18]. Training a ResNet43-classifier, the model accuracy reaches 80% for the 4 tree species considered [9]. All trees delineated from the NAIP imagery are classified in one of the 4 species data and carbon sequestered is calculated based on tree density and form factor of the tree's geometric dimensions.

3 RESULTS & DISCUSSIONS

3.1 Carbon Sequestration in Urban Forests

Carbon sequestered in trees can be estimated based on allometric equations [13, 23]. Tree height, canopy diameter and tree species information get aggregated to determine above/below ground biomass (A/BGB) [9]. Biomass can be considered the sum of AGB and BGB. The AGB is the biomass stored in tree trunks, leaves, tree branches, etc. In contrast BGB is the biomass stored in roots and decaying leaves buried in soil. The AGB and BGB depend on the physical dimensions of the tree, tree mass density, and tree's shape form factor. The tree shape factor characterizes the tree crown relative to other tree dimensions [24]. Many studies were carried out to model and to measure tree shape factors.

A standard tree biomass allometric equation reads:

$$AGB = F \cdot \rho \cdot H \cdot V = F \cdot \rho \cdot H \cdot \frac{\pi D^2}{4} \quad (1)$$

where the tree height H is extracted from LiDAR data and the canopy diameter D is estimated by orthoimagery segmentation models. F denotes the tree form factor and ρ encodes density with both specific to tree species. BGB is based on a fractional contribution of AGB, namely $BGB = f * AGB$ where the fraction, f , depends on upper soil depth, i.e. on how deep and how far roots extend in the horizontal direction. f also depends on tree species, the tree age, and tree health. Commonly used fraction f values are tabulated in the literature [25]. The total biomass is the simple sum $TB = AGB + BGB$. Sequestered carbon is calculated as a fraction of the total biomass, with a standard discount factor of 40% off the total biomass [26].

Given the strategy outlined in section 2.4, we mapped more than 36 million trees for Long Island, NY based on NAIP imagery. All trees are classified into any of the four species: Oak, Honey Locust, Callary Pear, and Planetree [9]—those denote the most dominant trees for New York City [21]. It is assumed that the NYC tree census data is representative for Long Island, too. As part of the image segmentation process, tree canopy diameter D and tree height H get extracted for individual trees. Subsequently, the carbon sequestered is determined based on eq. (1), and $TB = .4AGB \cdot (1 + f)$. Summing the carbon sequestered in individual trees does result in $5.1 \cdot 10^6$ tonne of carbon stored in Long Island's forests. Figure 2 (right) provides a snapshot of the trees as detected around Brookhaven National Laboratory's Relativistic Heavy Ion Collider. Iterating the procedure for Dallas, TX, the overall carbon sequestered in vegetation equates to 780,000 tonne—a value at the same scale as reported by others, e.g. [27].

3.2 Cooling Potential of Tree-Populated Spaces

Trees enable ambient cooling by changing the surface temperature. As sample reference, Figure 3 provides a snapshot from Long Island, NY with surface temperature distribution as recorded by the Landsat 8 satellite (left) with corresponding tree distribution (right). Areas of larger tree coverage reveal lower surface temperature values when compared to areas less populated by trees. To quantify the benefit of urban forests on local temperature, a (virtual) 100m-by-100m grid defines cells to aggregate individual trees into densities, cf. Figure 3 (b). At the same time, a Landsat 8 data-derived mean temperature gets associated with each grid cell.

A snapshot of surface temperature distributions for New York City, Dallas, TX and Long Island, NY presents Figure 4 (top row) alongside with temperature response to tree density within 100m-by-100m squares (bottom row). All three geo-locations indicate that there exist urban heat islands mixed with lower temperature regions. Cooling potential is most prominent for Long Island, NY where the eastern part of the island is much cooler than the west side where NYC boroughs is

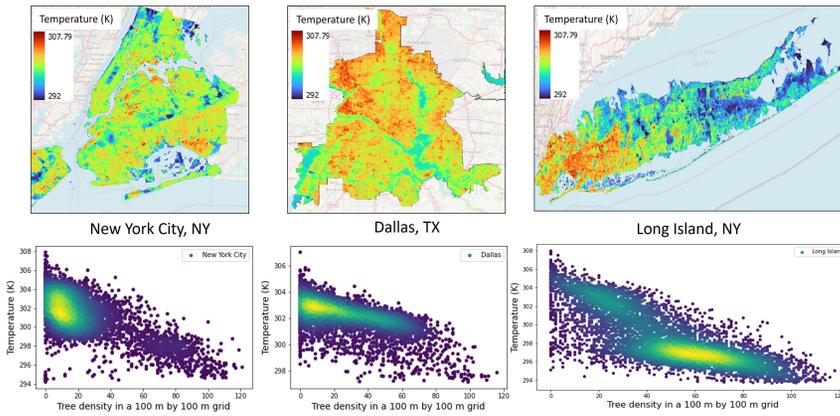


Fig. 4. Variation in surface temperature for New York City, Dallas, TX, and Long Island, NY (top row). Corresponding average mean temperature vs. tree density value plots provides the bottom row.

located. For all three sample areas, we identified a noticeable cooling of up to 2 Kelvin for areas densely covered in trees.

While urban spaces like New York City, NY and Dallas, TX do exhibit a single linear scaling of temperature vs. tree density, for Long Island, NY we observe a double-linear dependence. It traces back to the hybrid nature of Long Island's land cover where densely populated areas are mixed with agricultural and wetlands. The wetland tends to stay cooler due to surrounding waters such that the impact of trees is less pronounced. Applying linear regression to the change in surface temperature as function of tree density, the slope of temperature change for tree count is 0.03 K/tree for New York City, 0.072 K/tree for Dallas, and 0.125 K/tree and/or 0.04 K/tree for Long Island.

3.3 Urban Planning for Heat Island Mitigation

Based on airborne LiDAR surveys it has been demonstrated that rule-based processing is able to extract (noisy) semantic segmentation maps for buildings and vegetation [20]. Those may get utilized in order to train deep neural networks for change detection models as rapid response to identification of damage to vegetation after an event of natural hazard such as flooding triggered by hurricanes [8]. Here, we exploit the rule-based identification in order to correlate the fraction of buildings in patches of about 200m-by-200m with the average surface temperature as recorded by the Landsat 8 satellite mission. In addition, we fuse in survey information characterizing elements of urban planning in terms of Local Climate Zones (LCZ), [28]. In particular, we classify blocks of buildings in 200m-by-200m sample patches from the NYC boroughs Manhattan, Queens, and Brooklyn into ensembles of compactly (LCZ 1 to 3) and openly (LCZ 4 to 6) arranged blocks; details in Table 2 of [29], p. 1885.

All data fusion is managed by the Big Geospatial Data analytics platform PAIRS that allows to auto-generate co-registered raster statistics on ingestion of raw LiDAR point clouds. Those 0.5m resolution rasters serve as basis to derive the rule-based semantic segmentation maps as referenced above. Figure 5 summarizes our findings: While Figure 4 demonstrates the cooling capacity of urban forests, we notice that cooling trends critically depend on the density and composition of buildings. As expected, urban planning with open design (LCZ 4-6; gray \circ in Figure 5) allows for systematically lower local average temperatures when compared to compact architectures (LCZ 1-3; red +, green \times , and blue \bullet in Figure 5).

Further drilling into the sub-classes of compact building arrangements reveals a striking insight, i.e. the data uncover orthogonal trends for increasing urbanization when the fraction of buildings

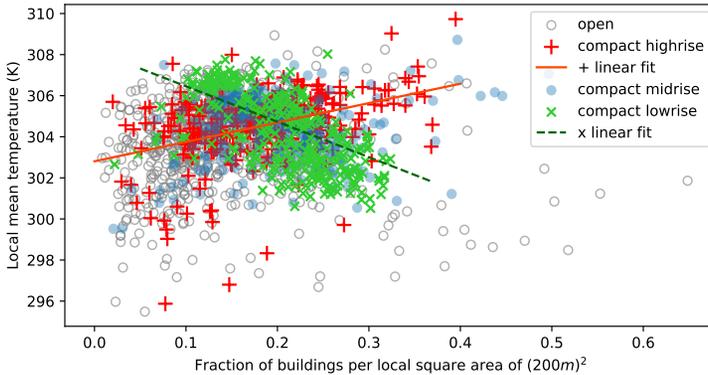


Fig. 5. Local temperature dependency on building density in New York City for selected local climate zones.

inflates: In a range where 10% to 30% of the local area is developed by low-rise architectures (×) such as e.g. characteristic for suburbs where tree crowns typically exceed building heights, cooling effects may persist (green, dashed linear trend down). However, in high-rising, compact settings a clear increase in local temperatures (red, solid linear trend up) on the ground results from intensified land use by building construction (+). Indeed, while the locally averaged ratio of tree-to-building heights equates to ~ 1.2 for *compact lowrise* (LCZ 3), it approximately reads 1 for *compact highrise* (LCZ 1). Moreover, LCZ 1 scenes for the NYC patches analyzed yield a ratio of tree cover vs. building population of about 1/2, while LCZ 3 equates to a ratio of 6/9 such that the natural cooling effect of trees remains prominent.

4 CONCLUSIONS & PERSPECTIVES

In this work, by means of large-scale remote sensing analytics, we presented the relevance of trees for urban planning from the perspective of climate change, namely: a/ trees provide a local sink to sequester carbon, b/ urban forests serve to reduce local land surface temperature, and c/ proper utilization of tree arrangements in accordance with building infrastructure in densely populated areas allows to preserve the cooling capacity of vegetation in order to mitigate heat island. Urban forest management provides a straightforward strategy in order to simultaneously mitigate urban heat islands and increase carbon sequestration in cities.

Technically, our work bases on multimodal data analytics where satellite imagery, airborne-collected 3D point clouds, and city surveys got fused to draw the conclusions summarized above. Besides our results on the relevance of urban forests, we did allude on the correlation of urban-planning defined *Local Climate Zones* and climate proxies such as mean surface temperature. The right combination of trees and building in urban areas has potential to overcome urban heat island formation.

5 ACKNOWLEDGMENTS

We would like to thanks Fernando Marianno, Wang Zhou, Siyuan Lu, Johannes Schmude, and Hendrik Hamann for insightful discussions and continuous PAIRS platform development. Conrad and his team *Large-Scale Data Mining in Earth Observation* including Chenying Liu and Yi Wang gratefully acknowledge support by the HelmholtzAI initiative (<https://www.helmholtz.ai/themenmenu/our-research/research-groups/albrecht-group>) and Zhu lab (<https://ai4eo.de>) headed by Xiao Xiang Zhu.

REFERENCES

- [1] N. B. Grimm et al. Global change and the ecology of cities. *Science*, 319(5864):756–760, 2008.
- [2] B.J. He. Towards the next generation of green building for urban heat island mitigation: Zero uhi impact building. *Sustainable Cities and Society*, 50:101647, 2019.
- [3] Y.H. Ryu and J.J. Baik. Quantitative analysis of factors contributing to urban heat island intensity. *Journal of Applied Meteorology and Climatology*, 51(5):842–854, 2012.
- [4] I. D. Stewart, T.R. Oke, and E.S. Krayenhoff. Evaluation of the ‘local climate zone’ scheme using temperature observations and model simulations. *International journal of climatology*, 34(4):1062–1080, 2014.
- [5] M. L. Verdonck and A. Okujeni et al. Influence of neighbourhood information on ‘local climate zone’ mapping in heterogeneous cities. *International Journal of Applied Earth Observation and Geoinformation*, 17:102–113, 2017.
- [6] C. Qiu, L. Mou, M. Schmitt, and X.X. Zhu. Local climate zone-based urban land cover classification from multi-seasonal sentinel-2 images with a recurrent residual network. *ISPRS Journal of Photogrammetry and Remote Sensing*, 154:151–162, 2019.
- [7] M. Demuzere, J.Kittner, and B. Bechtel. Lcz generator: A web application to create local climate zone maps. *Frontier in Environmental Sciences*, 2021.
- [8] C. M. Albrecht, C. Liu, Y. Wang, L. J. Klein, and X. X. Zhu. Monitoring urban forests from auto-generated segmentation maps. *2022 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, in press, 2022.
- [9] L.J. Klein, W. Zhou, and C. M. Albrecht. Quantification of carbon sequestration in urban forests. *ICML Climate Change Workshop*, 2021.
- [10] L.J. Klein, F.J. Marianno, and C.M. Albrecht et al. Pairs: A scalable geo-spatial data analytics platform. *2015 IEEE Conference on Big Data*, pages 1290–1298, 2015.
- [11] S. Brown. Measuring carbon in forests: current status and future challenges. *Environmental Pollution*, 116(13):363–372, 2002.
- [12] T. Jucker and J. Caspersen et al. Allometric equations for integrating remote sensing imagery into forest monitoring programmes. *Global change biology*, 23(1):177–190, 2017.
- [13] Q. Song, C. M. Albrecht, Z. Xiong, and X.X. Zhu. Towards global forest biomass estimators from tree height data. In *International Geoscience and Remote Sensing Symposium (IGARSS) 2022*, pages 1–4. IEEE, 2022.
- [14] J.M. Omernik and G.E. Griffith. Ecoregions of the conterminous united states: evolution of a hierarchical spatial framework. *Environmental Management*, 54(6):1249–1266, 2014.
- [15] P. Shukla, J. Skea, R. Slade, A. Al Khourdajie, R. van Diemen, D. McCollum, M. Pathak, S. Some, P. Vyas, R. Fradera, et al. Climate change 2022: Mitigation of climate change. *Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*.
- [16] J.T. Overpeck, G.A. Meehl, S. Bony, and D.R. Easterling. Climate data challenges in the 21st century. *Science*, 331(6018):700–702, 2011.
- [17] C.M. Albrecht et al. Pairs (re)loaded: System design benchmarking for scalable geospatial applications. *2020 IEEE Latin American GRSS ISPRS Remote Sensing Conference (LAGIRS)*, pages 488–493, 2016.
- [18] W. Zhou, L.J. Klein, and S. Lu. Pairs autgeo: an automated machine learning framework for massive geospatial data. *2020 IEEE International Conference on Big Data (Big Data)*, pages 1755–1763, 2020.
- [19] USDA. National agriculture imagery program (naip), 2021.
- [20] C. M. Albrecht, F. Marianno, and L. J. Klein. Autogeolabel: Automated label generation for geospatial machine learning. *2021 IEEE International Conference on Big Data (Big Data)*, pages 1779–1786, 2021.
- [21] NYC. New york city street tree map, 2015.
- [22] L. Wang, P. Gong, and G.S. Biging. Individual tree-crown delineation and treetop detection in high-spatial-resolution aerial imagery. *Photogrammetric Engineering Remote Sensing*, 70(3):351–357, 2004.
- [23] J.H. Lee, Y. Ko, and E. G. McPherson. The feasibility of remotely sensed data to estimate urban tree dimensions and biomass. *Urban Forestry Urban Greening*, 16:208–220, 2016.
- [24] J. Chave et al. Improved allometric models to estimate the aboveground biomass of tropical trees. *Global change biology*, 20(10):3177–3190, 2014.
- [25] D.L. Cheng and K.J. Niklas. Above-and below-ground biomass relationships across 1534 forested communities. *Annals of Botany*, 99(1):95–102, 2007.
- [26] R. A. Birdsey. Carbon storage and accumulation in united states forest ecosystems. *US Department of Agriculture, Forest Service*, 59, 1992.
- [27] Texas Trees Foundation. State of the dallas urban forest, 2015.
- [28] M. Demuzere, S. Hankey, G. Mills, W. Zhang, T. Lu, and B. Bechtel. Combining expert and crowd-sourced training data to map urban form and functions for the continental us. *Scientific data*, 7(1):1–13, 2020.
- [29] Ian D Stewart and Tim R Oke. Local climate zones for urban temperature studies. *Bulletin of the American Meteorological Society*, 93(12):1879–1900, 2012.