

High-Resolution Urban Heat Modeling: CNN-Based Predictions in San Francisco

Joshua Hikida

Received October 23, 2024

Accepted May 09, 2024

Electronic access May 31, 2024

This study explores the predictive capabilities of Convolutional Neural Networks (CNNs) in modeling urban heat distributions, focusing on the Urban Heat Island (UHI) effect in San Francisco. Leveraging high-resolution rasterized temperature data from the Urban Heat Watch Program and urban morphological data, the research employs a CNN model to predict microscale temperature variations across 212 subgrids and 3 times of day in the city. The model incorporates features such as building height and baseline rural temperature data to predict the temperature of each 10m x 10m cell. The model achieved a Mean Squared Error of 0.0216 in the normalized scale, far above other baseline machine learning models, demonstrating a promising proficiency in capturing the relationships between urban form and microclimatic variations. This research not only substantiates the potential of deep learning in predicting UHI effects in a lightweight and straightforward package for urban planners but also paves the way for future studies to explore and enhance the model's applicability and accuracy across diverse urban contexts.

Introduction

As the world's population continues to increase, urbanization is increasingly prevalent¹, with over 50% of the world's population residing in cities² and some estimates projecting growth in urban area by over one million kilometers by 2030³. Consequences of urban development, such as anthropogenic heat, impervious surfaces, and lack of vegetative cover^{4,5}, fuel what is known as the Urban Heat Island (UHI) effect, where higher temperatures in urban settings are present compared to surrounding rural and suburban areas, thus creating "islands" of heat⁶. With this expansion, the effects of UHIs have become an increasingly prevalent factor in urban planning and academic research⁷.

Especially important as global warming continues to drive temperatures up across the world, and further exacerbated by heat islands, the base impact of increased temperatures is itself damaging. Extreme heat results in fatalities through factors such as heatstroke⁸ and cardiovascular disease⁹, and whether in Athens¹⁰, Shanghai¹¹, or Ho Chi Minh City¹², studies consistently link deteriorating health and deaths across the world to the heat island effect¹³. Furthermore, subgroups such as Black people, women, those of lower socioeconomic status, the elderly, and young children are disproportionately affected by heat^{9,14}, underscoring the urgency of working to resolve the issue.

The pervasive effect of UHIs also dramatically impacts local meteorology, human lives, and animals. One particularly harmful impact is an increased concentration of airborne pollutants¹⁵, leading to effects ranging from economic losses to death, with over 6.4 million deaths worldwide in 2015 attributed to air pollution¹⁶. UHIs also alter the process of natural selection¹⁷,

which is extremely damaging to local ecosystems in these cities. Ironically, urban heat is also self-perpetuating; through greater adoption of air conditioning units, refrigerants with 2,000 times the global warming potential as carbon dioxide¹⁸ are emitted, resulting in a cyclic worsening of this effect.

To potentially aid solutions around policy and urban planning to mitigate this effect, this study aims to develop a machine-learning platform to predict temperatures at a fine granularity in cities due to easily gatherable features such as building footprints, zoning, water features, and roads. Compared to traditional simulation programs such as UMEP-SOLWEIGH¹⁹ and ENVI-MET²⁰, such a data-driven approach requires less data and reduces the need for expert knowledge that can lead to human error²¹.

Machine learning techniques such as regression²², artificial neural networks²³, dense neural networks²¹, and random forests²⁴ have been used to predict varying related metrics, such as UHI intensity, but these models often inherently do not take into account spatiality and thus cannot accurately predict the exact distribution of temperatures throughout the target area. For example, a study determining the most influential factors in the UHI effect by Yoo⁴ uses a random forest algorithm built on decision trees and thus cannot consider spatial factors (such as the proximity of buildings).

Thus, it is evident that a gap exists in considering these factors, which this paper aims to help fill.

Methods

Data Acquisition and Preprocessing

Temperature data across various zones in San Francisco was meticulously gathered by the Urban Heat Watch (UHW) Program on September 2nd, 2022²⁵. The program modeled gridded heatmaps with a fine granularity of 10 meters by 10 meters, capturing traverse measurements during distinct time intervals: morning (6-7 am), afternoon (3-4 pm), and evening (7-8 pm), resulting in three unique heatmaps. We divide each heatmap into 212 "zones", each of size 640 meters by 640 meters, resulting in a total of 636 subgrids when considering all time intervals. To facilitate model training and interpretation of predictions, temperature values were normalized between 0 and 1 with 0.5 representing the average temperature.

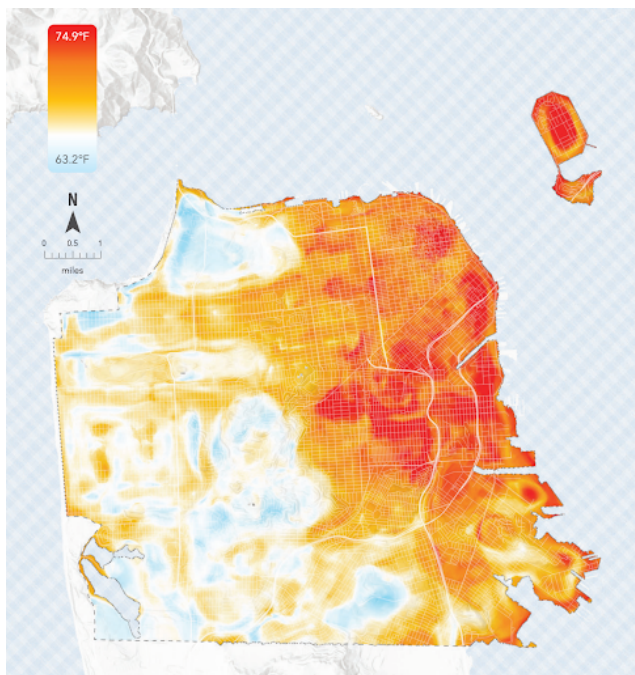


Fig. 1 Initial Temperature Data²⁵

Building footprint data, including building heights from Microsoft²⁶, and OpenStreetMap data²⁷, which encapsulates categories such as land use, bodies of water, natural features, and roads, were extracted for each subgrid. Each 10m x 10m cell within a subgrid was annotated for land use type (e.g., residential, commercial), presence of water (e.g., lakes, ponds), natural features (e.g., scrub, rock), and road types (e.g., highway, side road). Additionally, each cell was assigned a height corresponding to the tallest building within its bounds, or zero if no building was present.

A crucial aspect of our model's input involves integrating a reference temperature²⁸ from a rural location, specifically

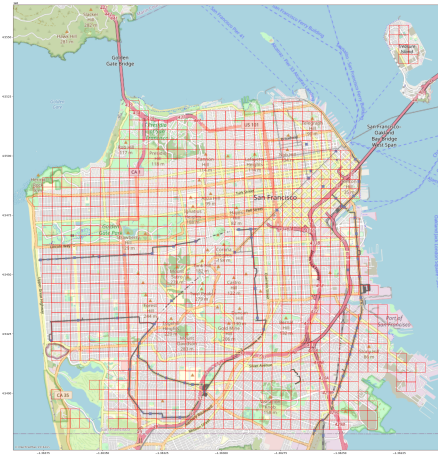


Fig. 2 Division of Temperature Data into Zones

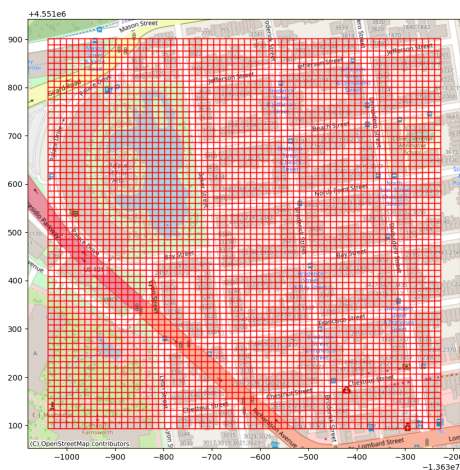


Fig. 3 Further Subdivision of Zones



Fig. 4 Building Footprints

Moss Beach, CA. This inclusion is pivotal for establishing a baseline temperature, representing what the temperature in San Francisco might have been without urbanization. Moss Beach was selected as the reference station adhering to the World Meteorological Organization's guidelines, which advocate for a location embedded in relatively flat terrain and exhibiting rural characteristics²¹. Furthermore, its coastal nature, proximity to open spaces, and location approximately 20 miles south of San Francisco mirror some of the geographical and meteorological aspects of San Francisco, thereby making it a pertinent baseline for our study. The reference temperature from Moss Beach is incorporated into the model, aiding it in discerning the urban heat island effect by providing a comparative rural temperature metric against the urban temperatures observed in San Francisco.

The extensive data available in San Francisco minimized the impact of missing and erroneous data. Nevertheless, we encountered limited instances of missing data, particularly in the UHW dataset. In these cases, we applied mean/median imputation, standard for handling such gaps in a data-driven manner. Furthermore, both the UHW and Microsoft Building Footprints datasets included mechanisms to flag questionable entries, which we subsequently excluded from our analysis. Additionally, we conducted a thorough manual verification of the OpenStreetMap data, particularly focusing on green spaces and tree coverage, and found these data to be accurate and reliable. It is important to note that our modeling approach, including data augmentation and regularization as described below, heavily mitigates the impact of minor data inaccuracies.

Data Augmentation

We also employed data augmentation techniques to enhance the robustness of our model and ensure it generalizes well to unseen data. Each subgrid, along with its associated feature maps (e.g., heights, bodies of water), was spatially rotated by 90, 180, and 270 degrees, and these rotated grids were added to the dataset. This approach not only quadrupled our dataset size but also enabled the model to learn invariant features regardless of their orientation in the input space.

Augmented data was randomly selected for manual inspection and found to be as expected. The rotated feature maps still represent plausible urban scenarios, as similar configurations of buildings, green spaces, and roads could occur in different configurations in real-world settings. Furthermore, the model was evaluated on an independent test set not exposed to the augmentation process, ensuring that the model's predictive accuracy and generalizability were not compromised by the inclusion of augmented data. The consistent performance of the model on this untouched test set shows that the spatial rotation of data, as part of our augmentation strategy, did not noticeably artificially inflate the model's efficacy or introduce systematic biases.

Model Architecture

A Convolutional Neural Network (CNN) was employed for its unparalleled ability to discern and learn spatial patterns within the data, an essential feature for analyzing urban environments. CNNs excel in processing two-dimensional data and capturing local spatial relationships and hierarchies, which are crucial in understanding the complex dynamics of urban heat distributions. Their effectiveness in this domain is well-documented in recent literature, with successful applications in classifying functional zoning at a fine scale²⁹, identifying urban features such as buildings and roads³⁰, and analyzing vegetation cover and properties³¹. Compared to other architectures like Dense Neural Networks or Random Forests, CNNs offer a more direct and efficient method for processing spatial data without extensive feature engineering. While preliminary exploration of alternative architectures was considered, the decision to use a CNN was ultimately guided by its proven track record in similar studies and its inherent suitability for the spatial nature of our data.

In the context of our CNN, each subgrid is treated as an input feature map, where the various urban characteristics serve as channels of information that the network leverages to comprehend and predict temperature distributions. The CNN scans through each cell and its neighbors, identifying localized patterns and relationships between urban features and temperature.

The architecture of the CNN was designed with multiple input layers to accommodate the diverse feature maps and additional variables. Each spatial input was passed through a dedicated CNN block, consisting of convolutional and max-pooling layers, to extract pertinent spatial features. These extracted features were then flattened and concatenated with non-spatial inputs, such as the time of day and a reference temperature from Moss Beach, CA. The concatenated features were passed through dense layers, reshaped into a grid structure, and finally, through an output layer to predict the final 64x64 temperature grid.

Model Training and Evaluation

The model was compiled using the Adam optimizer and Mean Squared Error (MSE) as the loss function, given the regression nature of the task. It was trained over 100 epochs with a batch size of 32, using 80% of the data for training and validating on 20%. The model was then evaluated on a separate test set, ensuring an unbiased assessment of its predictive capabilities.

To address overfitting, we implemented regularization techniques and employed data augmentation strategies, such as spatial rotations of input grids. Underfitting was mitigated by iteratively increasing the model complexity through additional layers and neurons, ensuring the model was sufficiently expressive to capture the nuanced relationships inherent in the urban heat data. By examining the low training and testing error, it is clear that

these strategies prevailed, and no significant issues with model fit were present.

Results

CNN Results, Metrics

This study focuses on two key metrics: normalized Mean Squared Error (MSE), and unscaled Root Mean Squared Error (RMSE).

The MSE metric allows us to assess how well the model captures the relative distribution of temperatures. This is particularly important for urban planners interested in understanding spatial temperature patterns rather than absolute temperature values. For example, when considering the impact of adding green spaces or altering zoning regulations, the relative differences in temperature distribution are more informative than the exact temperature values. As our temperature data is scaled between 0 and 1, the normalized MSE of 0.00216 indicates that the relative difference between temperature predictions was only off by 0.22%.

The unscaled RMSE metric provides a more direct figure; the temperature data is returned to its pre-normalized state (degrees Fahrenheit), and so RMSE represents how far the model's predictions were from the true temperature values. The unscaled RMSE indicates that the model's predictions were, on average, just 0.214 degrees Fahrenheit units away from the actual values, a figure that is easily understandable by policymakers and other stakeholders.

The low MSE and RMSE values suggest that the model can accurately predict temperature distributions based on the provided urban features and reference temperature, supporting the hypothesis that these variables can be utilized to predict urban heat distributions.

Baseline Comparison Outline

No existing studies precisely replicate our task of predicting micro-scale temperatures influenced by the urban heat island effect using spatial features. This unique challenge necessitates a comparative evaluation of our Convolutional Neural Network (CNN) approach against a diverse range of potential machine learning architectures to underscore its effectiveness. We have selected six widely used models as baselines: Multi-layer Perceptron (MLP), Support Vector Machine (SVM), Random Forest (RF), Linear Regression (LR), k-Nearest Neighbors (k-NN), and Decision Trees (DT). The choice of MLP allows us to demonstrate the advantages of CNN over more basic neural network structures, particularly in handling spatially structured data. LR provides a contrast from the perspective of powerful linear and non-linear approaches that require the transformation of spatial data. Random Forest and Decision Trees, as representatives

of ensemble and tree-based methods, offer insights into how traditional machine learning models perform on this complex spatial prediction task. k-NN is included to explore the efficacy of instance-based learning in capturing local spatial patterns. Through these comparisons, we aim to highlight the specific strengths of CNNs in capturing and interpreting the intricate spatial relationships crucial for accurate micro-scale temperature prediction in urban environments.

Baseline Comparisons

All baseline models were trained with a flattened transformation of the original dataset, as they are unable to take into account spatial factors like the CNN. As expected, the CNN outperformed all other models, confirming its superior ability to capture spatial relationships embedded in the dataset.

Following closely behind, RF performed well despite its lack of explicit spatial awareness. Its performance suggests that ensemble methods can serve as alternatives for this spatial analysis task, but also reinforces the precision of a well-tuned CNN.

DT and MLP models reported significantly higher errors, likely due to difficulty in managing the dataset's complexity without CNN-like spatial processing capability. PCR and KNN performed modestly well, indicating that these methods, while they can reduce dimensionality or capture local similarities, they are likely inadequate alone for dealing with the high complexity of spatial data involved in temperature predictions.

Linear regression, unsurprisingly, as the most naive model selection, performed extremely poorly, with an RMSE indicating that predictions were billions of degrees off from the true temperature values. The data contained non-linear, complex spatial relationships which were a significant mismatch with the model's assumptions (linearity).

Discussion

The results align with the initial hypothesis that urban features, when integrated with machine learning models, can effectively predict spatial temperature distributions. This alignment is particularly evident in the model's ability to discern and learn from the complex, multi-dimensional data, thereby providing predictions that closely mirror the actual temperature distributions recorded by the Urban Heat Watch Program. Comparatively, our findings echo the sentiments of previous studies that have identified a tangible link between urban structures and localized temperature anomalies, further solidifying the foundational understanding that urban planning and design are pivotal in managing urban thermal environments.

A critical theme emanating from the findings is the potential for utilizing such predictive models in urban planning and policy-making to mitigate UHI effects. The model's ability to accurately predict temperatures based on urban features and a

Table 1 Performance Comparison of Different Regression Models

Architecture	MSE (Normalized)	RMSE (Unscaled)
Convolutional Neural Network (CNN)	2.16×10^{-3}	2.14×10^{-1}
Random Forest Regressor (RF)	2.45×10^{-3}	2.29×10^{-1}
Decision Tree Regressor (DT)	4.55×10^{-3}	3.12×10^{-1}
Principal Component Regression (PCR)	5.75×10^{-2}	1.11
Multi-Layer Perceptron (MLP)	5.82×10^{-2}	1.11
K-Nearest-Neighbors Regression (k-NN)	6.77×10^{-2}	1.20
Linear Regression (LR)	2.23×10^{-2}	2.18×10^{11}

reference temperature suggests that urban planners and policy-makers can leverage such technologies to simulate and evaluate the potential impact of various urban design strategies on localized temperatures. This could guide the development of urban spaces that are functionally efficient and thermally comfortable, especially in the context of a warming climate.

While our model demonstrates a robust predictive capability, it is imperative to juxtapose these findings against prior studies. Some research in the domain has leaned heavily into the physical aspects of urban planning, such as green spaces and building materials, to understand and mitigate UHI effects. Our model, which integrates a myriad of urban features into a predictive framework, adds a computational lens to this narrative, suggesting that digital simulations can serve as a viable tool in understanding and potentially mitigating UHI effects.

The practical implementation of our model in urban planning and policy-making contexts does pose certain challenges that need to be acknowledged and addressed. The foremost challenge lies in the realm of data quality: the model's predictive accuracy is contingent upon the precision and completeness of urban data inputs, such as building footprints and land use details, and usage by real-world planners could result in failure to reduce urban heat if data is not properly vetted or acquired. While the model incorporates a degree of robustness to mitigate the impact of suboptimal data quality, the integrity of input data remains a critical factor, and manual verification of the data will be necessary. Moreover, the process of assembling and processing these extensive datasets can be resource-intensive, demanding above-average time and computational power, which may exceed the typical resources of urban planning departments. It is important to note, however, that the technical setup of the model itself is relatively straightforward; for the hundreds of cities covered by the Urban Heat Watch Program, only the city parameter must be changed for the model to adapt to any location. Thus, the model can run as a "black box" that requires little to no expert intervention to generate results. With appropriate support and resources, including leveraging widely available cloud computation, these hurdles can be effectively and readily overcome. Ultimately, it is crucial to verify the input and output of the model and avoid solely relying on its projections in a

vacuum.

Though the authors face time and computational limitations in incorporating all possible features, future research could explore the integration of dynamic meteorological variables, like transient weather patterns or long-term climatic changes, into the model, providing a more holistic and overall more accurate predictive tool by considering additional features. Additionally, exploring the model's applicability and accuracy in different urban contexts and climates could provide valuable insights into its generalizability and potential for global applicability. Even without these additional features, this study still achieves its purpose of demonstrating a promising base for other studies to build upon in forecasting urban heat.

Conclusion

The exploration into the predictive capabilities of Convolutional Neural Networks (CNNs) in forecasting urban heat distributions has unveiled a potent tool that intertwines urban morphology and microclimatic variations. Our model, substantiated by its low error metrics, not only affirms the hypothesis that urban features can be pivotal predictors of localized temperature patterns but also underscores the potential of computational models in advancing our understanding and management of Urban Heat Island (UHI) effects. The findings provide a digital framework that can simulate, analyze, and potentially mitigate the thermal implications of various urban design strategies. One especially relevant use of this model is in urban planning, where departments on a local to national level could leverage this tool to advise them in creating thermally amiable layouts in terms of zoning, building heights, infrastructure, and many more factors.

Moving forward, augmenting the model by integrating dynamic meteorological variables is imperative, thereby enhancing its predictive accuracy and temporal adaptability, as well as allowing for more factors to be considered by planners. Future studies should also explore the model's applicability across diverse urban and climatic contexts, ensuring its robustness and generalizability in a global scenario. Moreover, synthesizing such predictive models with practical urban planning and policy-making will be pivotal, guiding the development

of both thermally comfortable and environmentally sustainable urban spaces. Ultimately, this study shows that a combination of computational modeling and pragmatic urban development by policymakers can serve to mitigate UHI effects and foster urban futures safer for humans and animals alike.

Acknowledgments

Map data copyrighted OpenStreetMap contributors and available from <https://www.openstreetmap.org>

References

- 1 M. Beckline, *American Journal of Agriculture and Forestry*, **5**, 33.
- 2 Jansson, *Ecological Economics*, **86**, 285–291.
- 3 J. Foley, R. DeFries, G. Asner, C. Barford, G. Bonan, S. Carpenter, F. Chapin, M. Coe, G. Daily, H. Gibbs, J. Helkowski, T. Holloway, E. Howard, C. Kucharik, C. Monfreda, J. Patz, I. Prentice, N. Ramankutty and P. Snyder, *Science*, **309**, 570–574.
- 4 S. Yoo, *Journal of Big Data*, **5**, year.
- 5 L. Myrup, *Journal of Applied Meteorology*, **8**, 908–918.
- 6 M. Santamouris, *Science of the Total Environment*, **512-513**, 582–598.
- 7 L. Kleerekoper, M. Esch and T. Salcedo, *Resources, Conservation and Recycling*, **64**, 30–38.
- 8 B. Doyon, D. Bélanger and P. Gosselin, *International Journal of Health Geographics*, **7**, year.
- 9 R. Basu, *Environmental Health*, **8**, year.
- 10 J. Paravantis, M. Santamouris, C. Cartalis, C. Efthymiou and N. Kontoulis, *Sustainability*, **9**, 606.
- 11 J. Tan, Y. Zheng, X. Tang, C. Guo, L. Li, G. Song, X. Zhen, D. Yuan, A. Kalkstein, F. Li and H. Chen, *International Journal of Biometeorology*, **54**, 75–84.
- 12 T. Dang, D. Van, H. Kusaka, X. Seposo and Y. Honda, *American Journal of Public Health*, **108** 2, 137–143.
- 13 K. Wong, A. Paddon and A. Jimenez, *Journal of Energy Resources Technology*, **135**, year.
- 14 A. Hsu, G. Sheriff, T. Chakraborty and D. Manya, *Nature Communications*, **12**, year.
- 15 U. Energy, *U. Energy*, Assessment of International Urban Heat Island Research (Navigant Consulting, Author).
- 16 P. Landrigan, *The Lancet Public Health*, **2**, 4–5.
- 17 E. Shochat, P. Warren, S. Faeth, N. McIntyre and D. Hope, *Trends in Ecology Evolution*, **21**, 186–191.
- 18 *18 National Renewable Energy Laboratory*, <https://www.nrel.gov/news/press/2022/nrel-shows-impact-of-controlling-humidity-on-greenhouse-gas-emissions.html>.
- 19 F. Lindberg, C. Grimmond, A. Gabey, B. Huang, C. Kent, T. Sun, N. Theeuwes, L. Järvi, H. Ward, I. Capel-Timms, Y. Chang, P. Jonsson, N. Krave, D. Liu, D. Meyer, K. Olofson, J. Tan, D. Wästberg, L. Xue and Z. Zhang, *Environmental Modelling Software*, **99**, 70–87.
- 20 M. Taleghani, D. Sailor and G. Ban-Weiss, *Environmental Research Letters*, **11**, 024003.
- 21 J. Oh, J. Ngarambe, P. Duhirwe, G. Yun and M. Santamouris, *Scientific Reports*, **10**, year.
- 22 Y.-F. Su, G. Foody and K.-S. Cheng, *Landscape and Urban Planning*, **107**, 172–180.
- 23 S. Liu, J. Zhang, J. Li, Y. Li, J. Zhang and X. Wu, *Building and Environment*, **202**, 108051.
- 24 H. Hou, Q. Longyang, H. Su, R. Zeng, T. Xu and Z.-H. Wang, *International Journal of Applied Earth Observation and Geoinformation*, **122**, 103411.
- 25 V. Shandas, J. Voelkel, J. Williams and J. Hoffman, *Climate*, **7**, 5.
- 26 Microsoft, *USBuildingFootprints*, <https://github.com/microsoft/USBuildingFootprints>.
- 27 G. Boeing, *Computers, Environment and Urban Systems*, **65**, 126–139.
- 28 *28 Weather Visual Crossing Corporation*, <https://www.visualcrossing.com/>, Retrieved from.
- 29 W. Zhou, D. Ming, X. Lv, K. Zhou, H. Bao and Z. Hong, *Remote Sensing of Environment*, **236**, 111458.
- 30 X. He, A. Wang, P. Ghamisi, G. Li and Y. Chen, *IEEE Geoscience and Remote Sensing Letters*, **16**, 125–129.
- 31 T. Kattenborn, J. Leitloff, F. Schiefer and S. Hinz, *ISPRS Journal of Photogrammetry and Remote Sensing*, **173**, 24–49.