Senseable Amsterdam Laboratory

Assessing the effect of street-level urban morphology on Land Surface Temperature

USING STREET-VIEW IMAGES TO MAP URBAN MORPHOLOGY TO INVESTIGATE THE EFFECTS ON URBAN TEMPERATURES IN AMSTERDAM

THESIS MSC MADE

Author:

Michiel VAN SELM

Studentnumber:

1164376

Date:

December 13, 2023

Supervisor:

Prof. Dr. Fabio DUARTE

Supervisor:

Prof. Dr. Ir. Eveline VAN LEEUWEN

Daily Supervisor:

MSc Titus VENVERLOO

WAGENINGEN UNIVERSITY & RESEARCH

TU Delft

Abstract

According to the latest IPCC reports heat waves will increase in both intensity and frequency. Due to the physical nature of cities, they are particularly vulnerable to this increase in heat. Urban heat island (UHI) effect analysis often relies on satellite imagery, which gives a planar representation of often three-dimensional features. In this study, we propose to integrate street view data to create a large-scale approximation of local street-level micro-climates in urban environments, using Amsterdam as a case study. We present a method that incorporates street view images with a semantic segmentation model to capture finer urban elements from the panoramic images, such as sky, buildings, trees, and pervious and impervious surfaces. Furthermore, our approach also involves the calculation of view factors derived from panoramic street view images, employing a hemispherical azimuthal projection technique to accurately capture the 3D element of the urban environment. This allows us to assess the impact of various environmental features on LST, considering elements such as tree view factor (TVF), sky view factor (SVF) and building view factor (BVF). We then use the extracted features to model the relationship between these features and LST using machine learning algorithms such as Support Vector Regression, Gradient Boosting Tree, and a Random Forest. The Random Forest model turned out to be the best-performing model. To address research questions, these features are analyzed for their correlation with Land Surface Temperature (LST). The study reveals strong correlations between LST and buildings/trees, while the sky % correlation is surprisingly weak. The SVF-LST relationship is intricate, with larger SVF leading to increased heat absorption yet potentially diminishing the UHI effect through enhanced airflow. Application of results in the Amsterdam context demonstrates practical insights for climate adaptation. Areas with low tree or pervious surface coverage exhibit higher LST values, emphasizing the importance of urban greening for heat mitigation. Additionally, the study uncovers unexpected hotspots, challenging existing climate strategies. However, limitations exist, such as LST not directly equating to thermal comfort, potential exaggeration of cooling effects, and the focus on summertime daytime temperatures.

Keywords – Computer vision, Urban Heat Island, Energy surface balance, Semantic segmentation

CONTENTS

1. Introduction

1.1. Growing importance

Heat waves are known as the silent killer among extreme weather events. According to Zhao et al. (2021) excess heat is responsible for around 100.000 deaths in South Asia between 2000 and 2019. It is also forecasted that as a result of climate change the deaths that can be attributed to excess heat will increase (Huang et al., 2011). According to the latest report by the (IPCC, 2021) the effects of anthropogenic climate change on urban areas will be significant. The report specifically mentions the effects of increased frequency and intensity of heat waves. As a result of their physical characteristics, Urban areas are particularly vulnerable to heat waves due to the urban heat island effect. The absorption and retention of heat, and the low albedo of urban surfaces such as roads, buildings, and pavements can lead to significantly higher temperatures than the surrounding areas (IPCC, 2021; Kalnay and Cai, 2003). In this IPCC report the consequences of 1.5 degrees of global warming are set out. However, the growing consensus among the same authors who wrote the IPCC report, is that 1.5 degrees may no longer be realistic (Mooney, 2023). This means that the adverse outcomes as described will most likely be even more frequent and at a higher intensity. Combined with the fact that half of the population of the world currently lives in an urban environment and that the number of urban inhabitants is predicted to rise by an extra 2.5 billion by 2050 (Massaro et al., 2023). This combination of rapidly increasing heat and a growing urban population means that the number of people who are exposed to extreme heat is rising rapidly (Tuholske et al., 2021).

1.2. Societal impact

Exposure to extreme heat has also been linked to several adverse health outcomes. These include heat-related illnesses, cardiovascular diseases, and respiratory diseases (Heaviside et al., 2017). At one hospital in Australia, the effect of heat waves on hospital admissions was measured. It was found that temperatures above 30°°Ccan increase the number of hospital admissions from 2.25 per 10.000 people to 4.92 per 10.000 (Watson et al., 2020). The elderly, children, and people with preexisting conditions are especially vulnerable to the health risks related to the urban heat island. Factors that contribute to this risk include housing characteristics and socioeconomic status. Housing with poor insulation or lack of air conditioning can result in unhealthy indoor temperatures increasing the risk of the aforementioned health risks (Loughnan et al., 2012). The growing population that is exposed to heat and the health effects that accompany it, therefore, experiences an increasing risk of fatality, as a result of the environment they live in(Massaro et al., 2023). This shows that environmental factors influence both the physical and psycholog-

ical health of the people living in that said environment (Raymond et al., 2020). Studies have also shown that urban thermal comfort can influence human outdoor activity (Li and Ratti, 2018). Huang et al. (2015) found in their study that during the hot season, there were fewer people in outdoor spaces, and the people present preferred more static activities. At the same time between 70% and 90% of the world economy is located in those urban areas (Wei et al., 2023). The combination of these factors means that the additional warming of the Urban Heat Island (UHI) could bring 2.6 times the economic costs to cities when compared to lower-density areas (Li and Ratti, 2018). The combination of lower labor productivity due to heat waves and the amplifying effect of UHI on heat waves result in this higher economic loss (Estrada et al., 2017). Given that cities are both the economic and population centers of the world the amplifying effect can have a large impact on the economy(Wei et al., 2023; Massaro et al., 2023). For both the health of the population and the economic activity in cities, it is important to keep them livable.

1.3. URBAN MORPHOLOGY

In order to investigate the UHI, a temperature dataset is needed. High-resolution air temperature datasets are difficult to obtain. For this reason, Land Surface Temperature (LST) is commonly used to investigate UHI. LST is a measurement of the earth's surface temperature. In many studies, this variable is used as a representation of urban thermal properties and is commonly used to investigate the relationship between heat stress and the urban environment (Wei et al., 2023). Currently the most common way the quantify the relationship between urban morphology and land surface temperature is the use of remote sensing to map the land cover composition using remote sensing (Khalil et al., 2021; Song et al., 2014). Commonly associated land cover classes with cooling include water bodies and vegetation, while impervious surfaces like roads and buildings are associated with warming effect (Khalil et al., 2021; Wei et al., 2023; Song et al., 2014). However, satellite images present a planar representation of 3-dimensional reality. In recent studies the importance of the vertical dimension when researching urban LST has been highlighted (Wei et al., 2023; Zeng et al., 2018). Urban morphology has a significant influence on urban microclimates. Different types of 3D data are often used to model street-level urban morphology. For example, LiDAR data, 3D building data, and Street view images. One common element used as an indicator of urban morphology and constructed using 3D data is the Sky View Factor (SVF). Several studies have looked at the relationship between SVF and urban heat stress (Zeng et al., 2018; Li et al., 2017). There are several methods to calculate the SVF. For example, 3D city model-based methods, GPS methods, or using fisheye photos (Zeng et al., 2018; Li et al., 2017). In the analysis of this vertical space, the vegetation factor is rarely considered. However, using LiDAR data research has found that the height of a tree and the height difference between trees and buildings is very important in LST prediction (Chen et al., 2022). This means that not only the

SVF but also the Building View Factor (BVF) and the Tree View Factor (TVF) are important 3D aspects of the urban canyon. One widely available data source that captures the vertical dimension accurately and provides essential information about street canyons is panoramic street-view images (Wei et al., 2023). Unlike remote sensing, which only offers a planar view of the environment (Choi et al., 2022). In previous studies using fisheye imagery, the focus has been on the SVF. However, variations in the BVF and the TVF could also influence the temperature in the urban canyon (Gong et al., 2018). The effect of trees and buildings has been considered in both microclimate modeling on a small scale and remote sensing on a city-wide scale (Xu et al., 2021; Zhou et al., 2017). Including these factors combines the 3D elements of microclimate modeling with the city-wide scale of remote sensing. This could also be beneficial for policymakers in their assessment of the effectiveness of the green infrastructure on a large scale and inform new developments in a more efficient way than complicated microclimate modeling. This resulted in the following research questions.

1.4. Research question and methodological approach

Question: What is the effect of street-level urban morphology on Land Surface Temperature?

Sub-Questions:

- Can street-view be used to create an accurate representation of the street level microclimate?
- To what extent can these features, combined with street-level urban morphology features, be used to predict street-level LST?

In this paper, a workflow to extract segmentation classes from fisheye images, as well as from the original panoramic street view images is proposed. The segmentation can then be used to investigate the effects of the different view factor compositions and street-level urban features on LST. After the images are collected, a semantic segmentation model will classify the panoramic images using the Ade20k dataset. In Python a geometric model was then developed using the numpy package, to reproject the segmentation mask from the panoramic images to an azimuthal hemispherical projection. The extracted features, from both the panoramic and fisheye images, can then be used in machine or deep learning models to explore the relationships between the different parameters and LST. This method is in line with current literature on the use and detection of the SVF, BVF, and TVF, while also using novel methods such as semantic segmentation for the sky detection, to increase the efficiency of the workflow. The contributions of this study can be summarized as follows:

- Workflow to extract view factor features at a large scale from panoramic street view images

- Investigating the effects of the different compositions of sky, tree, and building view factor features on LST

This methodology can give urban planners and designers insight into the effects of street design on the temperature. This can help the implementation of mitigation strategies with scientific grounding.

2. Theoretic Framework

2.1. Urban Heat Island formation factors

In this research, The Urban Heat Island is characterized by the spatially averaged temperature difference, influenced by a spectrum of factors associated with urbanizations, with temperature variations scaling in relation to the intricate combination of these factors (Stewart, 2011; U.S. Environmental Protection Agency, 2023). Factors such as low city albedo, high urban fabric heat capacity, heat-trapping through thermal radiation, and low surface evapotranspiration. Urban areas typically have more built-up surfaces, such as concrete and asphalt, which have a lower albedo than natural surfaces like vegetation. As a result, urban areas absorb more solar radiation, leading to higher temperatures. The high heat capacity of these materials also means that the heat that is absorbed during the day is stored, and still radiated at night leading to higher nighttime temperatures. The low surface evapotranspiration is the result of the typically lower level of vegetation and natural surfaces compared to outlying areas (Oke, 1982; van Hove et al., 2015). As a result of low wind speeds and atmospheric conditions, this heat is also trapped in the urban area. The UHI is more likely to occur under weather conditions with low wind speeds and high solar insolation (Oke, 1982; van Hove et al., 2015). Previous studies conducted through crowdsourcing in the Netherlands revealed that the mean daily maximum UHI for urban areas was 2.3°C, with the 95th percentile being 5.3°C (Steeneveld et al., 2011). In Amsterdam during the summer of 2015, van Hove et al. (2015) found an average evening UHI of about 1°C, with a maximum of 4.5°C. However, it is notable that this study refers to an annual average, and includes some very low-density outlying areas that are within the city boundaries of Amsterdam.

2.2. Energy surface balance

The Urban Heat Island is characterized by the spatially averaged temperature difference, influenced by a spectrum of factors associated with urbanizations, with temperature variations scaling in relation to the intricate combination of these factors (Rizwan et al., 2008; U.S. Environmental Protection Agency, 2023). This means that there is a difference in the energy balance between the surface and the atmosphere in areas with different degrees of urbanization (Oke, 1982). Generally, higher-density urban areas have a higher sensible heat, whereas more outlying areas have higher latent heat. Latent heat refers to the heat that is necessary to conduct a phase change of substance (Wang et al., 2015). While sensible heat refers to heat that can be felt as it radiates from a surface. A simplified version of the surface energy balance can

be defined as:

$$Q = H + LE + G \tag{2.1}$$

Where Q is the total energy flux. H is the sensible heat flux, the heat that is transferred from the surface materials to the air, LE is the latent heat flux, or heat that is used for evapotranspiration (the phase change of liquid water to water vapor) by vegetation, and G is the ground heat flux, which is the heat radiated into the subsurface(Wang et al., 2015). The height-to-width ratio of the urban canyon, wind speed, roughness of the surface materials, and canopy all influence the sensible heat flux (Voogt and Grimmond, 2000; Nazarian and Kleissl, 2015). However, a localized representation of the sensible heat flux can be defined as the following:

$$H = (T_0 - T_a) * C_a * C_H * U \tag{2.2}$$

Here T_0 - T_a represents the temperature difference between the surface and the air, C_a is the heat capacity of the air, C_H is the heat transfer coefficient, and U is the wind speed. The urban features that influence the latent heat flux are mainly the presence (or lack thereof) of impervious surfaces and vegetation (Oke, 1982; Wang et al., 2015). From these formulas, it can be concluded that the difference in energy balance between outlying and urban areas is a result of the difference in the partitioning of heat between sensible and latent heat (Oke, 1982). This difference in heat stems from the morphological differences between areas with different degrees of urbanization and the factors associated with that as mentioned in 2.1. Three key elements of urban geometry that give a characterization of the partition between sensible and latent heat are the view factors, sky view, tree view, and building view (Yıldız et al., 2023). Therefore, these view factors are often used in urban climate studies, it is a critical geometric factor in the large-scale estimation of urban microclimates and is used to assess urban thermal comfort and urban heat island effect (Yıldız et al., 2023; Oke, 1981). The SVF influences the thermal environment of urban areas by shifting the radiant effect, meaning that variations in the SVF in different locations change the amount of irradiance that is received by each surface, changing the energy surface balance in that location (Yıldız et al., 2023). Apart from the SVF, the composition of the urban canopy layer is also an important factor in controlling the radiation in the street canyon (Yıldız et al., 2023; Gong et al., 2018). The composition of the urban canopy layer is a determining factor in the way it influences the surface temperature. Trees, as depicted by the view factor(TVF), provide shade and cooling through the latent heat used for evapotranspiration. Buildings, as depicted by the building view factor (BVF), usually store heat and emit heat (sensible heat) as a result of high heat capacity materials and thermal acceptance of the surface of buildings (Yıldız et al., 2023; Gong et al., 2018). Examining these properties and the composition of these three factors on a city-wide scale

could provide valuable information on the interaction between street canyons and urban thermal comfort.

2.3. Data Science in Earth Science

2.3.1. Convolutional neural network architecture

Computer vision can help to extract these features on a city-wide scale from images. It enables us to process a large amount of images in a relatively short time using semantic segmentation. Semantic segmentation is built on an algorithm called a Convolutional neural network. This is an application of the neural network architecture specifically tailored to images. The artificial neural network is a computational model inspired by the human brain. Neural networks are essential to many modern technologies and have recently been adopted into earth science disciplines (Maskey et al., 2018). A basic version of the neural network architecture can be described as a set of interconnected nodes known as perceptions 2.1. These nodes take one or more real-valued inputs and output a single real value which they can pass on to other nodes (Maskey et al., 2018; Albawi et al., 2017). Its output is usually calculated by applying an activation function to introduce non-linearity $f(x)$ (Albawi et al., 2017; Schmidhuber, 2015). A simple representation of one node can be found in 2.1. These nodes are part of one or more hidden layers where the input of each node is s an n-dimensional input vector $\vec{x}$ and n-dimensional weight vector $\vec{w}$. The value of that node is the weighted sum of the input $\vec{x}$ with the weights $\vec{w}$.

$$x_j = \sum_i o_i w_{ji} + b_i \tag{2.3}$$

The output of such a network is calculated by propagating the input through the network. The value of each node x in layers after the input layer follows from 2.3. Each o_i is the output value from a node in the previous layer and each w_{ji} is the weight corresponding to that output. Typically each layer will have a bias node b_i connected to all the nodes in the following layer with a constant value of 1. The weights of connections to each bias node are optimized during the training process.

The combination of multiple perceptrons in a network structure can be used to solve complex problems. The specific type of neural network that is relevant for this research is the convolutional neural network (CNN). In this chapter, the architecture of this network will be explained (Albawi et al., 2017; Schmidhuber, 2015). The convolutional neural network combines three different layer types, a fully connected layer, a convolutional layer, and a pooling layer. This type of network is commonly used for image-related tasks, such as object identification, classification, and segmentation. The first part of the CNN is the convolutional layer. This layer was introduced, as can be derived from the name, to keep the number of parameters in the neural network relatively limited compared to a fully connected layer (Albawi et al.,

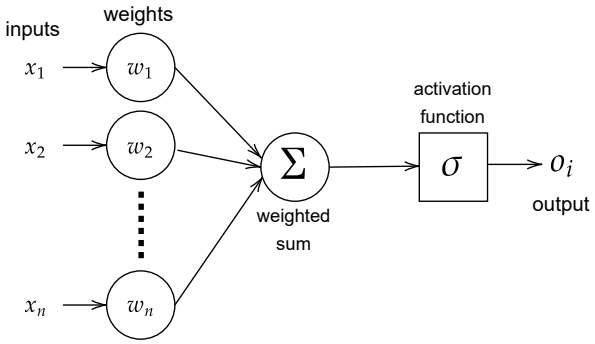

Fig. 2.1: The perceptron

2017; Li et al., 2022). Especially in image classification tasks, this is important. As illustrated by the following example of a modern image with a width and height of 1920x1080. If every pixel is considered an input parameter, this means that even for a hidden layer with only one neuron there are $1920 * 1080 * 3$ weight connections that need to be calculated for an RGB image. One neuron is also not considered useful for visual applications. Considering that even the most basic neural network has several layers with tens of neurons the number of calculations with spiral out of control quickly for modern computers. This problem of dimensionality resulted in the development of a more efficient approach (Li et al., 2022; O'Shea and Nash, 2015). Instead of connecting all the layers completely, the next layer is only connected to a subset of the corresponding image, called a filter. This filter slides over the image to create what is called a feature map 2.2. It is corresponding to a single value for each of the locations of the filter. After the creation of the feature map in the convolutional layer.

The dimensionality of the feature map is further reduced in the pooling layer. There are two common types of pooling layers, max, and mean pooling. For max pooling the maximum value of a region in the feature map is extracted, these max values combine to create a new smaller feature map. For mean pooling the mean value of this region is used to create the second feature map (Li et al., 2022; O'Shea and Nash, 2015). Another way to reduce dimensionality in neural networks is shared weights. The regions from the pooling layers are also used in this method. Neurons for the same region in the image also use the same weight. So if both of these are combined the number of weights that need to be calculated is drastically reduced by reducing both the number of connections and by reducing the variability of the weights within each layer (Li et al., 2022; Albawi et al., 2017; O'Shea and Nash, 2015).

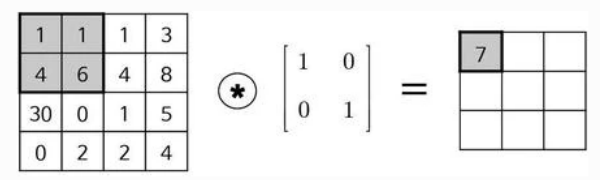

Fig. 2.2: Represention of sliding a filter over a simple grayscale image to create a feature map from (Kim and Kim, 2017)

The last layer is the fully connected layer. In this layer, the neuron connections are similar to a traditional neural network 2.3. Each neuron is fully connected. Meaning that each neuron is fully connected to both the previous and the next layer. The fully connected layer takes the produced feature maps from the convolutional layer and the pooling layer as inputs to make the final prediction for the image classification tasks (Li et al., 2022; O'Shea and Nash, 2015).

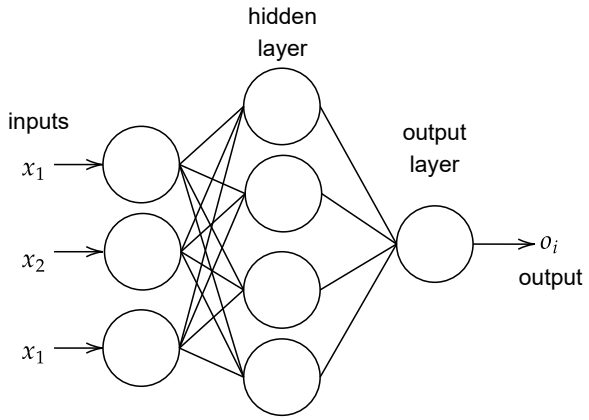

Fig. 2.3: A fully connected neural network

The activation function $\phi(x) = o$ used for calculating the output value of each node in this DPT model is the Gaussian Error Linear Unit (GeLu) function (Kenton and Toutanova, 2019; Hendrycks and Gimpel, 2016; Ranftl et al., 2021).

$$gelu(x) = \frac{1}{2}x(1 + \tanh(\sqrt{2/\pi}(x + 0.044715x^3))) \tag{2.4}$$

This activation function is similar to the Rectified Linear Unit (ReLU) activation function.

$$ReLU(x) = \max(0, x) \tag{2.5}$$

However, it has added smoothing compared to the ReLU function. This smoothness makes it well-suited for image-related tasks because of its benefits related to convergence. The nonlinearity of this function also enables it to learn more complex relationships which helps it to better understand the long-

range dependencies in images. It is also more efficient, giving it an edge for large datasets(Kenton and Toutanova, 2019; Hendrycks and Gimpel, 2016)

3. METHODOLOGY

3.1. DATA SOURCES

3.1.1. LANDSAT 8 AND 9

Landsat 8 is a satellite mission operated by the United States Geological Survey (USGS) that provides multispectral data. The data is collected by the Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS) aboard the satellite. The OLI has nine spectral bands, ranging from the visible to the shortwave infrared. The TIRS collects data in two thermal bands with a spatial resolution of 100 meters, which is converted to a spatial resolution of 30 meters by using cubic convolution. Landsat 8 has a revisit time of 16 days, meaning that it captures images of the same location on the Earth's surface every 16 days. Landsat 8 and 9 data are widely used in various applications one of which is urban heat island studies. For this study cloudless images from one of the three hottest months in the Netherlands June, July, and August in 2018 until 2022 were selected. The values from the selected day were then averaged to get a representative value for normal summer daytime temperatures in Amsterdam. The retrieved images are preprocessed with the atmospheric corrections by the USGS and transfer constants are provided to transform the cell values into degrees Celsius using the following equation.

$$T(°C) = \frac{K2}{ln(\frac{K1}{Tr} + 1)} - 273.15 \tag{3.1}$$

K1 and K2 are the constants provided by the USGS and Tr is the original raster band value.

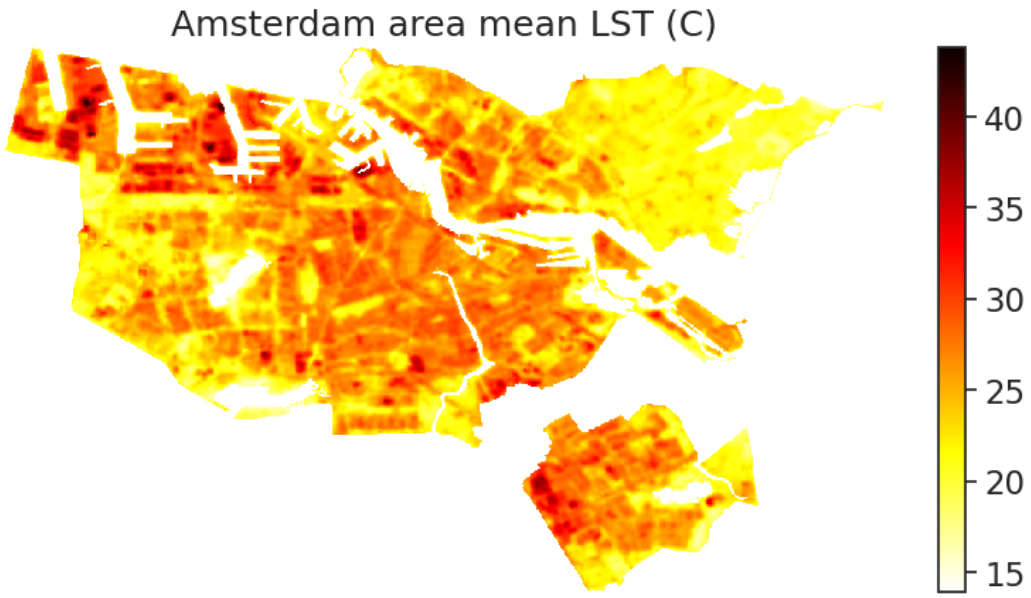

Fig. 3.1: Image of Land Surface Temperature in Amsterdam area prost-temperature transformation

As can be seen in figure 3.1, the large waterbodies have cut out of the temperature data. That was done because of the street-based nature of the research and the second data source (Street view images). In this Streetview dataset, the large waterbodies are usually not visible and therefore cannot be detected by the segmentation. The smaller canals are detectable as they are closely intertwined with the road network in the inner city. Therefore the canals were incorporated into the LST dataset.

3.1.2. Streetview images

The dataset consists of Streetview images 3.2 from Amsterdam that have been classified using a semantic segmentation model. The dataset consists of a 360° degree panoramic street view image every 50 meters on the entire road network of Amsterdam. In total, this is over 60.000 thousand images. Using semantic segmentation several categories were extracted from these images such as Sky, Building, Road, Tree, pervious surface, and impervious surface. It is important to note that street view is limited to roads and streets. That means that areas that are not directly adjacent to the road are not considered in this analysis.

For the semantic segmentation, a state-of-the-art Dense Prediction Transformer was used (Ranftl et al., 2021). It has a global perception field and combines traditional CNN architecture with transformers. It follows an encoder-decoder structure. The encoder is the transformer that learns the long-range dependencies in the image using the attention mechanism and the decoder is the convolutional layer that will learn the local features. The model used in this paper was pre-trained on the Ade20k dataset (Zhou et al., 2019). This setup was used to classify the full image dataset and the categories mentioned above were extracted.

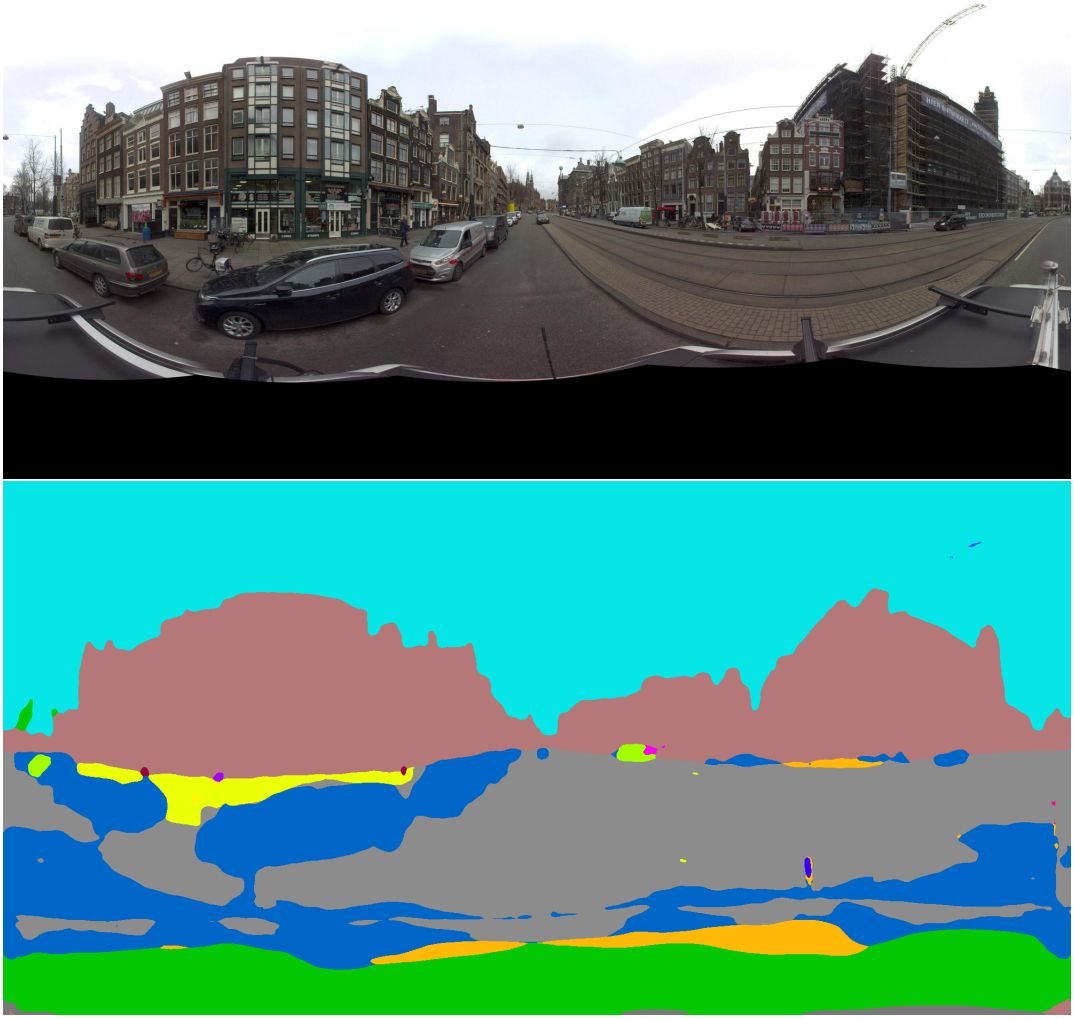

Fig. 3.2: Segmentation example

3.2. Analysis methods

3.2.1. Data collection

First, the dataset needs to be created in order to run the analysis. In this case that means adding Land Surface Temperature to the Streetview image dataset classification. In order to do this the LST for Amsterdam must first be calculated based on the thermal bands of the Landsat 8 satellite. Next, the classification data is aggregated to the 30m cell size of the LST dataset. Aggregating the data to larger cell sizes was not considered because it would increase the chances of incorporating areas that cannot be captured by Streetview imagery.

With this dataset, an analysis will be done of the impact of different factor compositions on the LST. This will be done by plotting a single view factor against the temperature while keeping the other two

factors relatively constant. This way we can isolate the effect of the single-view viewfactor. A second analysis will combine the view factors and the street view classifications to see if a predictive model can be created. For this part, several machine learning and deep learning models will be deployed, as well as feature selection to get insight into the weights the different variables have in determining the outcome.

3.2.2. Segmentation classes calculation

The panoramic images provide the 3D element and therefore a real-world approximation that 2D satellite imagery can not provide. By applying the earlier-mentioned DPT semantic segmentation model pretrained on the ade20k dataset we can calculate the proportions of the images in the classes mentioned in section 3.1.2. After the classification, the number of pixels of the color that corresponds to the classes is divided by the total amount of pixels in the panorama.

$$Tree\% = \frac{Pixels_{Tree}}{Pixels_{total}} \tag{3.2}$$

This same calculation is done for the different classes to create a table with the class % for each class and for all images. Table 3.2.2 is a small sample of the total table that was created this way.

Features segmentation results table						
Image ID	Building	Trees	Sky	Water	Impervious	Pervious
51	0.15	0.07	0.35	0.0	0.35	0.0
3	0.21	0.07	0.26	0.0	0.24	0.01
53	0.35	0.04	0.24	0.0	0.28	0.0
44	0.1	0.05	0.40	0.0	0.24	0.0
35	0.19	0.1	0.32	0.0	0.22	0.0

3.2.3. View factor calculation

In order to calculate the View factors from the panoramic street view images. The images must be projected into the hemispherical azimuthal projection with the lower hemisphere of the street view image cut off. In order to do this each pixel in the upper hemisphere of the panoramic image must be projected into the hemispherical projection based on the geometric model from (Li and Ratti, 2018). First, we need to cut off the lower hemisphere of the panoramic image by simply doing $H/2$ where H is the height of the image. Then we calculate the center coordinates of the the image and the radius of the resulting fisheye image.

$$C_x = \frac{W}{2\pi} \qquad C_y = \frac{W}{2\pi} \qquad radius = \frac{W}{2\pi} \tag{3.3}$$

The next step is to calculate the location of the pixels in the new fisheye image by calculating the angle and the distance from the center of the fisheye image. Where θ is the angle and ρ is the distance.

$$if\, x < C_x, \theta = \frac{3}{\pi} - \arctan(\frac{y - C_y}{x - C_x}) \qquad else, \theta = \frac{\pi}{2} - \arctan(\frac{y - C_y}{x - C_x}) \tag{3.4}$$

$$\rho = \sqrt{(x - C_x)^2 + (y - C_y)^2} \tag{3.5}$$

Then we calculate the corresponding pixels from the input image in regards to the resulting fisheye image with the following formulas.

$$x_p = \frac{W}{2\pi} * \theta \qquad y_p = \frac{\rho}{radius} * H \tag{3.6}$$

The application of these formulas results in a file of fisheye images as shown in figure 3.3.

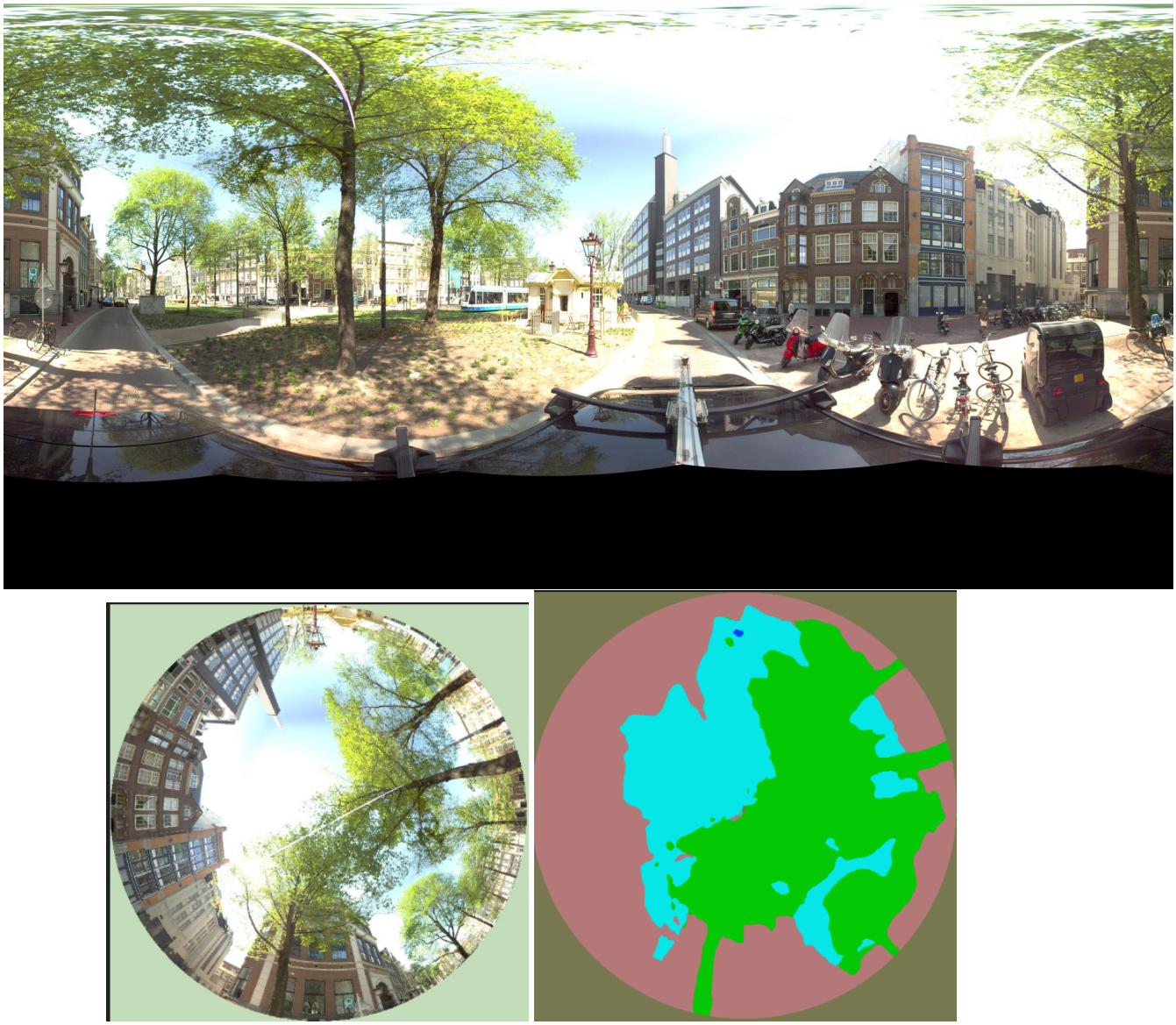

Fig. 3.3: Example of fisheye transformation

To calculate the view factors the same set of calculations is performed on the generated segmentation mask. The resulting fisheye segmentation mask will then be processed in the same way as the original segmentation masks using the same formula as given in 3.2. The result is a table with the image ID and the different view factors.

Fisheye results table			
Image ID	Building view	Tree view	Sky view
10080	0.24	0.27	0.49
57046	0.16	0.37	0.47
10740	0.32	0.36	0.31
21614	0.0	0.42	0.57
40436	0.02	0.43	0.54

In figure 3.4 a schematic representation is given of the full data pipeline from the raw data input to the dataset that is used for the prediction models and correlation analysis.

3.2.4. Predictive modelling

Before the data is used for predictive modeling a feature selection algorithm is used. This is done to improve the interoperability of the predictive models and to provide insight into the relationship between the different features and the LST. In this research, we use an L1 regression or a Lasso regression model.

$$\sum_{i=1}^{n}(Y_i - \sum_{j=1}^{p} X_{ij}\beta_j)^2 + \lambda \sum_{j=1}^{p} |\beta_j|$$

Cost function

Fig. 3.5: L1 regression

This regression model is similar to an Ordinary Least Squares (OLS) regression with an added penalty function highlighted in 3.5. Where:

- Y_i is the predicted output (in this case, the predicted LST).
- n is the number of data points
- X_{ij} is the input feature at index j.
- β_j is the coefficient that corresponds to the feature
- λ is the added penalty for feature β_j

If λ is zero there is no added penalty and the L1 regression will be the same as the OLS regression. If λ is high then the cost function will penalize the coefficients of the features with the least predictive power.

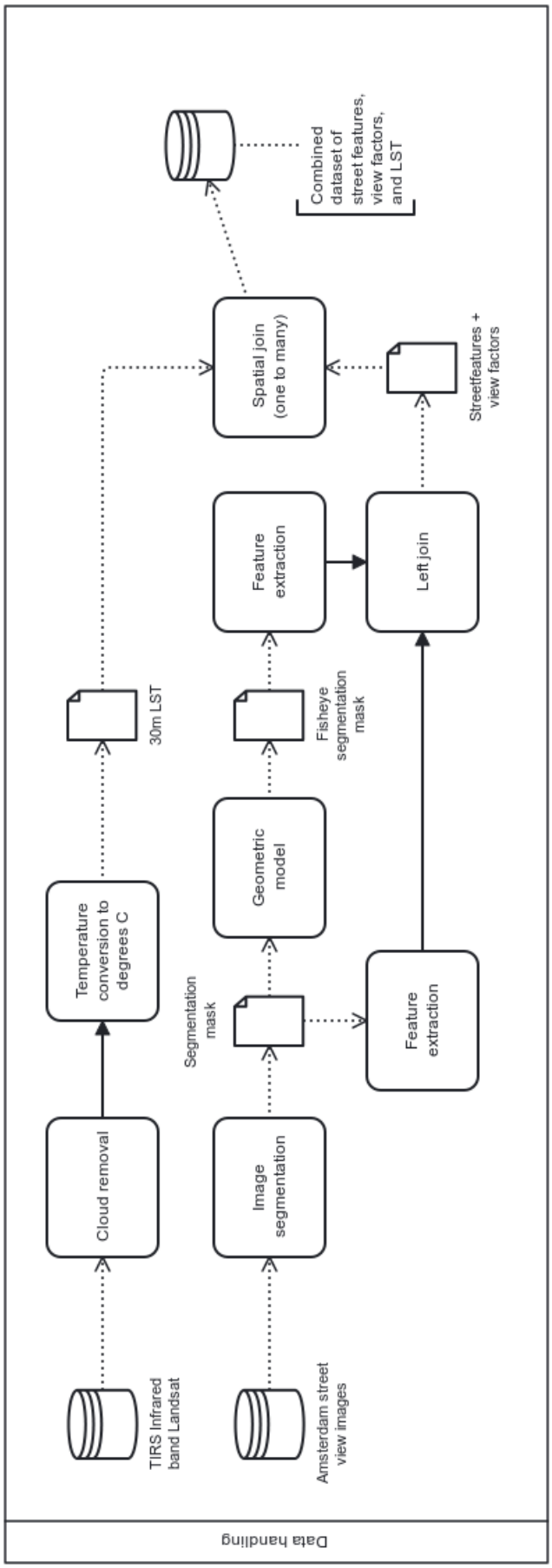

Fig. 3.4: Schematic representation of data handling

This way the model shows which features are actually important and have a strong relationship with the dependent variable, in this case, the LST. Typically the λ value is chosen through hyperparameter tuning using cross-validation (Fonti and Belitser, 2017).

The features that show the strongest correlation in the feature selection will then be used to train several machine learning models to test the predictive performance of environmental features for LST. In this study the following models were trained and tested: Support vector regression (SVR), Gradient boosting tree (GBT), and a Random Forest model (RF). On top of that, a deep-learning model was also trained. In this case a multilayer perception regression (MLP regression). An MLP regression model is a type of neural network that is often used to perform predictive tasks. The high complexity of the model allows it to find highly non-linear relationships. The drawback is that it needs a lot of data and can be computationally intensive.

The dataset for these models was randomly split into a train (80%) and test (20%) set. All models were fitted using the training data. In the training process, a 5-fold cross-validation was performed to ensure the adaptability of the model on unseen data. During the training process, hyperparameter tuning was also performed to prevent overfitting. After the models were trained they were used on the test dataset. To assess the performance of the trained models the R^2 value and the Mean Squared Error (MSE) were calculated. Equation 3.7 shows an example of a multilayer perceptron.

$$\hat{y} = f_2\left(\sum_{i=1}^{n_2} w_{2i}^{(2)} f_1\left(\sum_{j=1}^{n_1} w_{ji}^{(1)} x_j + b_i^{(1)}\right) + b_2^{(2)}\right) \tag{3.7}$$

Where:

- $\hat{y}$ is the predicted output (in this case, the predicted LST).
- f_1 and f_2 are activation functions (e.g. ReLU, sigmoid) applied to the weighted sums of inputs in the first and second layers, respectively. The best option determined by hyperparameter tuning. Although typically the ReLU function is used.
- x_j is the input feature at index j (e.g. the percentage of road in the cell).
- $w_{ji}^{(1)}$ and $w_{2i}^{(2)}$ are the weights of the connections between the input and first layer, and between the first and second layer, respectively.
- $b_i^{(1)}$ and $b_2^{(2)}$ are the biases added to the weighted sums of the first and second layer, respectively.
- n_1 and n_2 are the number of nodes in the first and second layers, respectively.

These models were all chosen because they are commonly used, and perform well, for regression analysis on larger datasets, it has also been used in similar studies regarding Urban Heat Island (Choe and Yom, 2017; Zhou et al., 2010).

4. Results

4.1. Exploratory data analysis

In this section, the result of the feature selection model will be presented along with some exploratory analysis. The first step is to examine the correlation between the features that were extracted from the street view images and the LST that was extracted from satellite imagery. As a first step, the lowest and highest LST data points and their corresponding panoramic images were looked up. In 4.1 and 4.2 the results of this are shown.

Fig. 4.1: Example of green area on the outskirts of Amsterdam with a low Land Surface Temperature

Fig. 4.2: Example of industrial area in Amsterdam with a high Land Surface Temperature

In this image, the trend seems clear. In image 4.1 with the lowest temperature, there is no building in sight. The image is filled with greenery and a small piece of road. In image 4.2 with the highest LST value in the entire dataset the visuals are reversed. There is no greenery visible at all in the image, which is completely filled with buildings and pavement. This anecdotal evidence shows a clear trend, in line with

the literature, where increased green lowers the LST and the presence of buildings increases the LST. The cold image 4.1 is located in the far northeast of Amsterdam. In 3.1 is also visible that the LST in this area is generally low. This area has a low degree of urbanization and consists mainly of grassland, trees, and water. Whereas the second image 4.2 is located in the northwest harbor area of Amsterdam. This is a highly industrialized area with some heavy industry. It is also very built up, and as can be seen in the image quite open. This means that the surface receives a lot of direct radiation, contributing to the high LST in this area. The following analysis of 60,000 street view images will show if this trend holds up city-wide for Amsterdam.

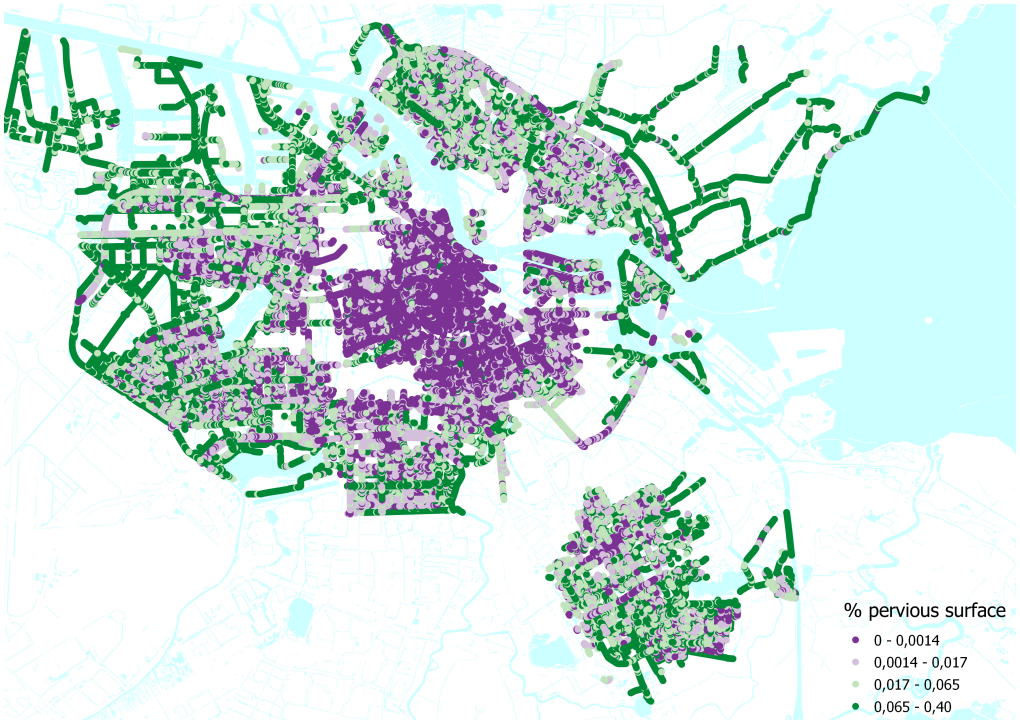

Fig. 4.3: Spatial distribution of building view factor

For the correlation analysis of the different features with LST, the Spearman correlation coefficient was used, because of the non-parametric nature of this analysis. In other words, it does not assume a normal distribution of the data. This is a result of the evaluation of the values based on their rank relative to the other values in the dataset instead the of absolute value. As most of the features do not follow a Gaussian distribution the selection of a non-parametric method was necessary. In this analysis, with the result shown in 4.4, the trees, pervious, and water features were all negatively correlated with LST. While impervious and building were positively correlated with LST. While sky had a very small negative correlation with LST. For all of these features, the P-value was very close to zero, meaning that these results all indicate a significant statistical relationship. The biggest positive correlation with LST is the buildings. The biggest negative correlation with LST is the presence of pervious surfaces. In **??** this is also

visible, in general, the LST cold spots in 3.1 in the west and far northeast correspond to the high upper end of the pervious surface value distribution. There are two notable exceptions, which are the earlier mentioned western harbor area and a small area in the southeast. This exception is probably the cause of the high density of industrial activities in both of these areas.

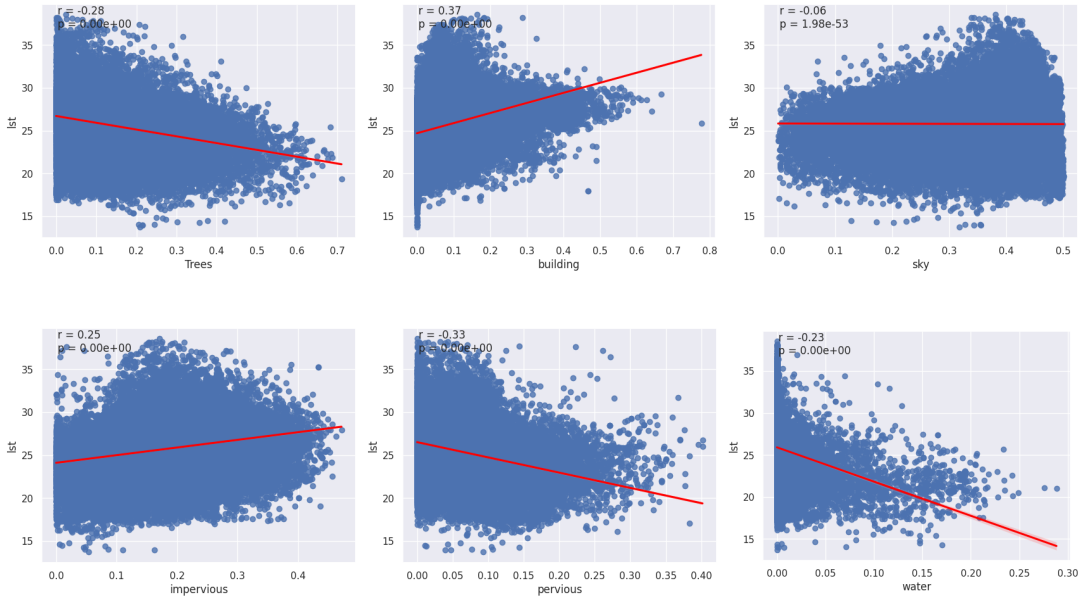

Fig. 4.4: Scatter plots of the different features including the Pearsons R statistical analysis to determine the degree of correlation and the P value for statistical significance

The results of this correlation analysis are also consistent with the result from our feature selection using the L1 regression, as can be seen in 4.1. Water is the exception here. While it was also negatively correlated in the Spearman analysis, it has a particularly large coefficient in this L1 regression. Even though the large waterbodies that also have the lowest surface temperature were removed. The IJ River, the Sloterplas, and the IJ Lake have all been removed as can be seen in 3.1. Then there is also the fact that water is sparsely visible from the street. With barely any data points with more than 0.15 water cover.

L1 regression results	
Feature	Coefficient
Trees	-4.40
Building	8.28
Sky	0.00
Pervious	-5.32
Impervious	3.59
Water	-20.15

4.2. Isolate effect of street-features

In this section variables that were extracted from the panoramic street view images before were analyzed. In the next section, the calculated view factors will be analyzed and compared to the original panoramic values. we are interested in the effect of the street canyon composition and specifically in the ratio of buildings versus trees in the street canyon. This approach in combination with the results of the first analysis, showed that the sky view factor was only slightly correlated with temperature. This is interesting because the spatial distribution of sky view factor 4.7 looks similar to that of the pervious surface 4.3. As a result of the low observed low impact of the SVF it was decided to plot the BVF and TVF against the temperature while keeping the SVF relatively constant. This was done by splitting the plots into sky view quantiles.

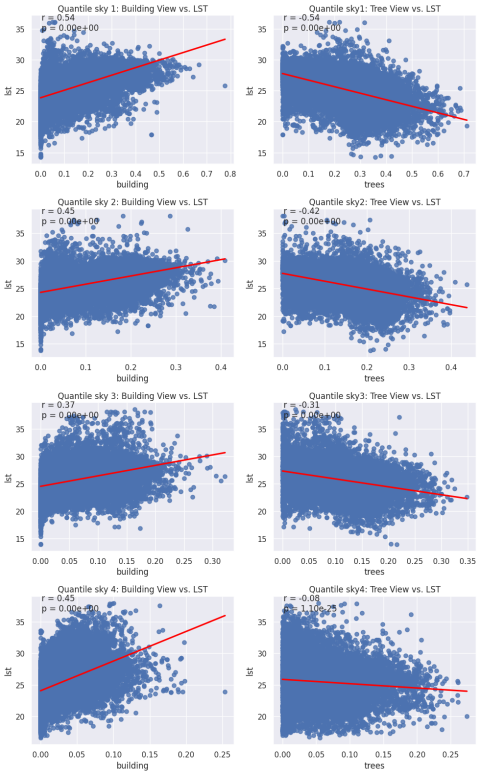

Fig. 4.5: Plotting tree and building from the panoramic projection against LST for each sky% quantile. Including the Spearman Rho correlation coefficient

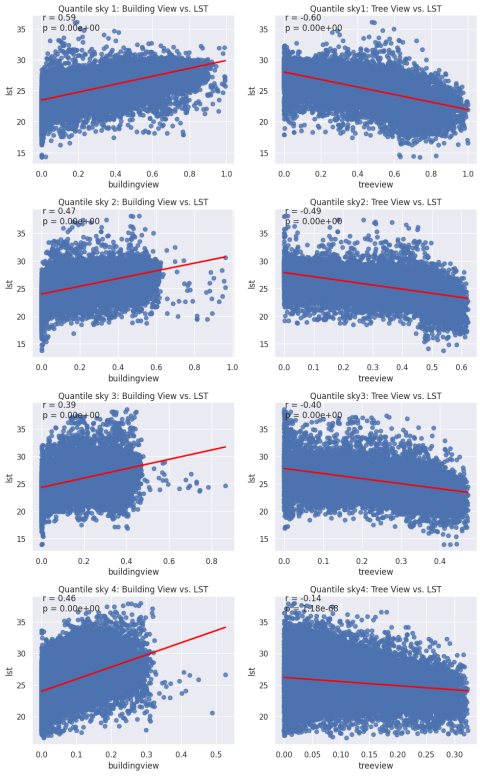

Fig. 4.6: Plotting TVF and BVF from the fisheye projection against LST for each SVF quantile. Including the spearman Rho correlation coefficient

In figure 4.5, the features are directly extracted from the panoramic street view images. In this 4.6 plot, the trend for both trees and buildings is more uniform in the lower quantiles, as represented by the higher correlation coefficients in the lower quantiles. This trend is especially clear for the trees. While the correlation with temperature gets weaker as the sky% increases, this decline is less steep in the case of

the buildings when compared to the features from the original panoramas. This is also represented in a lower score for the Spearman correlation coefficient. Where the score in the original plot was $r = -0.26$ for trees and $r = 0.46$ for buildings. In the case of tree %, it is almost double in the lowest quantile, while it is less than half in the highest quantile of sky %. This trend follows a similar pattern in the building % although less pronounced and with the exception of the highest quantile. Where the presence of buildings is still strongly positively correlated with higher temperatures. This plot indicates that especially in highly dense urban environments where the sky is usually more obstructed the amount of tree cover has a larger cooling effect on surface temperatures.

In figure 4.5 the % of each panoramic image covered by the tree and building class was examined. However, in the current literature regarding large-scale micro-climate, the features are usually processed into view factors. Because of the different orientations of the view factors, they are believed to provide a relatively accurate representation of the, receiving, and trapping of radiation and the airflow in the urban canyon. For this reason the same analysis as in figure 4.5 has also been done on the TVF and the BVF in figure 4.6. This plot shows a slightly more pronounced relationship with an increase in Spearman's correlation coefficient in every quantile. The increase is slightly higher for the TVF compared to the tree %, compared to the increase from building % to the BVF. The stronger statistical relationship between the Building and Tree View Factor and LST indicates that the features extracted from fisheye imagery provide a better estimation of the local microclimate compared to features extracted directly from the panoramic images.

From the spatial distribution of the SVF, TVF, and BVF in 4.7 4.9 4.8 it seems that the SVF and TVF both decrease as we come closer to the city center, while the BVF increases. This seems consistent with the LST in 3.1 if we consider the highly industrialized areas in the southeast and northwest as outliers.

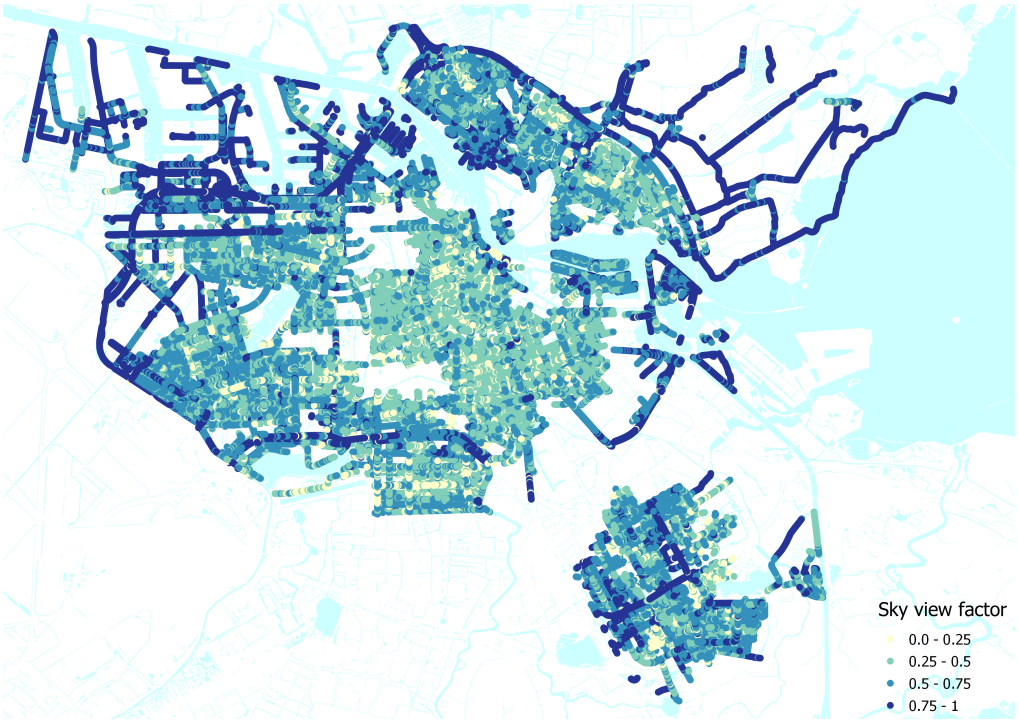

Fig. 4.7: Spatial distribution of sky view factor

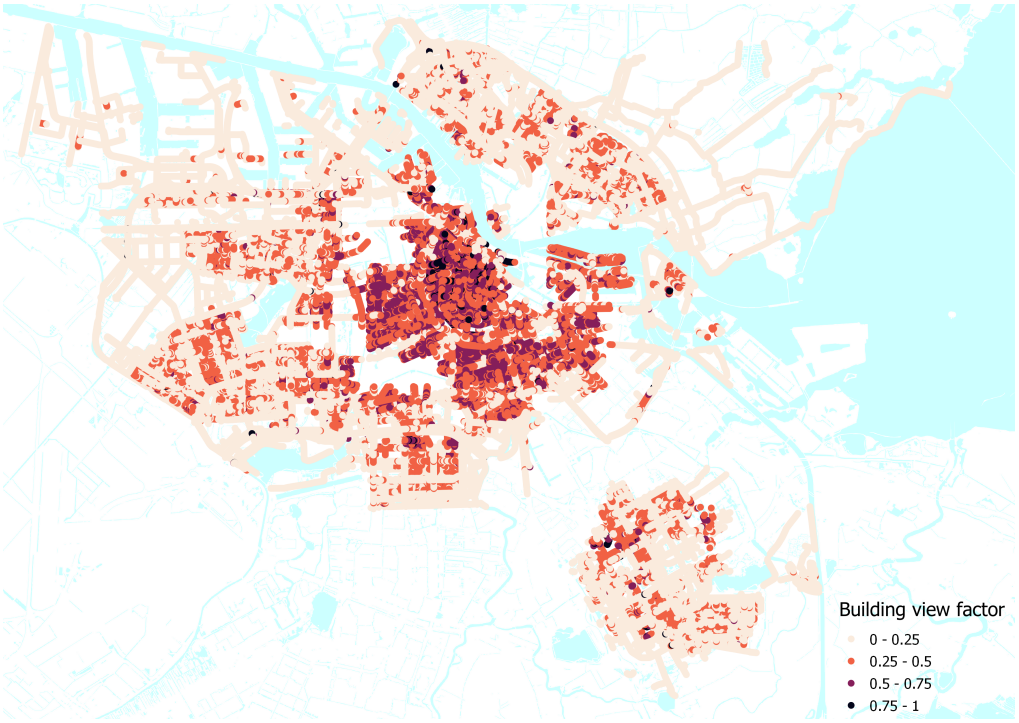

Fig. 4.8: Spatial distribution of building view factor

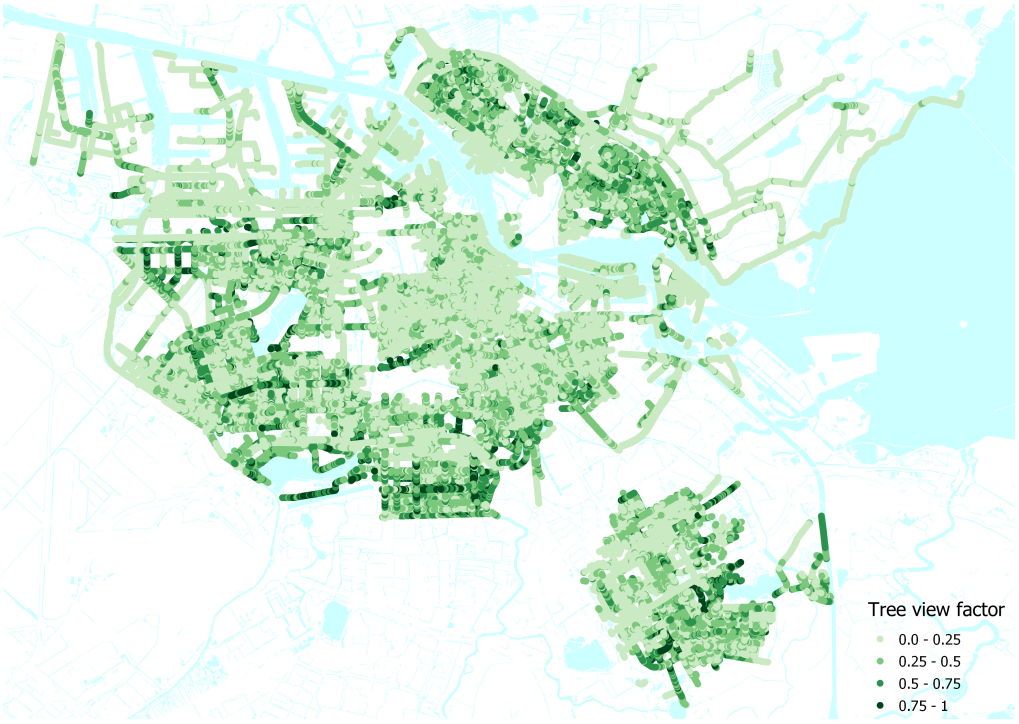

Fig. 4.9: Spatial distribution of tree view factor

4.3. Predictive peformance

In this section, a combination of the features from the original panoramic images and the converted fish eye images is used to train several machine-learning models. This will do two things, first, it will help us understand the relationship between these features and LST even better, and secondly, it will help to test the predictive power of street view features for LST estimation. The models that were trained are A l1 or Lasso regression, an RF regression, an SVR, a GBT, and the deep learning method MLP regression. The models vary significantly in performance. The L1 egression has an R^2 of 0.24 and an MSE of 6.03. The RF performance is significantly better with a R^2 of 0.52 and an MSE of 3.66. The GBT and SVR are in between the two with an R^2 of 0.39 and 0.27 respectively. The MLP regressor was only evaluated on MSE as is common with more complex deep learning methods. The complexity of deep learning makes evaluation by R^2 unreliable. After trying different configurations of hidden layers, activation functions, and nodes per layer. The lowest MSE achieved was 6.02. This MSE is almost twice as high as the MSE from the RF model. The full results can be found in table 4.3.

Model Evaluation		
Model	R^2	MSE
L1 regression	0.24	6.03
SVR	0.27	5.76
GBT	0.40	4.72
RF	0.52	3.66
MLP regression	–	6.02

From these model results it seems that features extracted from street view at 50m density do not provide enough information to accurately predict LST. While the random forest performance explains about half of the variation in the dataset, this is not enough to consistently make accurate predictions. The other models also perform considerably worse than the RF. With only an R^2 of 0.24, 0.27, and 0.4 respectively. It is noticeable that a relatively complex model such as an MLP regressor has a similar performance to the least complex model, the L1 regression, in this selection. Both scored an MSE of 6.0, considering the main purpose of the L1 is also feature selection and not prediction makes it more notable. Previous research indicates that increasing the number of street view images per LST cell could increase the model performance significantly (Wei et al., 2023). However, this comes as the cost of considerably higher computational costs.

5. Discussion

The causes of the urban heat island are complex, but it is widely accepted that street-level urban morphology has a considerable effect on the shaping of urban thermal properties. In the current literature, the research's main focus is, on the 2D representation of the urban environment, with this method large areas can be covered in the analysis. This large scale is beneficial in getting a wider overview of the situation. Another common method in current research is microclimate modeling, this method does take into account 3D data and gives an extremely detailed representation of the interplay between the different features at street-level. However, these studies are often focused on a small area and use a controlled environment. To answer the main research question it is important to take into account the third dimension, but also focus on the real urban environment at a spatial scale.

In this study, therefore, an estimation of the microclimate at a large spatial scale was conducted using 3D data from street view images. Important 3D elements include the Height/Width ratio of the urban canyon, the building height, and the tree height. These aspects are represented by the SVF, BVF, and TVF. With street view images and semantic segmentation, it is possible to extract these features at a large

scale from these street view images. This method combines both the large-scale associated with remote sensing approaches and captures detailed information about the street canyon similar to microclimate modeling.

To answer the research questions these features were then used to investigate the relationship between the 3D features of the urban canyon and LST. We looked at both the statistical relationship of individual features and the differences between features using the statistical correlation test Spearman ρ, as well as predicting the land surface temperature based on a combination of the extracted features using the following machine learning models: RF, GBT, SVR, and MLP.

5.1. Results in the scientific context

From our feature analysis, it can be concluded that urban features have a strong correlation with LST. Especially the buildings and trees appear to be strongly correlated with LST. Surprisingly the correlation between LST and the sky% was very weak. Based on the literature a higher negative correlation with LST was expected (Scarano and Sobrino, 2015; Zhang et al., 2019). The relationship between SVF and LST is complicated. A larger SVF means that the surfaces receive more direct sunlight, allowing these surfaces to absorb more heat (Huang and Wang, 2019). On the other hand a large SVF also typically corresponds to increased airflow through the urban canyon, which can diminish the UHI effect (He et al., 2020). There are also studies that indicate that the cooling effect of the SVF is larger during nighttime (Scarano and Mancini, 2017; Zhang et al., 2019). This is a result of the balance between increased airflow and higher radiation acceptance shifting. As the sun goes down this balance shifts to increased airflow, increasing the cooling effect of the SVF (Scarano and Sobrino, 2015; Zhang et al., 2019). Both of these studies however still show a small cooling effect during daytime. This small negative correlation for daytime LST was not observed in this study. It is possible that the reduction of the SVF and its corresponding temperature effect is also dependent on the property that is responsible for said reduction. For example, SVF reduction from street trees can have a cooling effect (Li et al., 2017). While SVF reduction as a result of buildings can have a heating effect (Guo et al., 2016). This analysis is also supported by the statistical analysis conducted in this research. Examining the view factors extracted from the reprojected fish eye images in 4.6 showed the same correlation between trees and buildings and LST as in these papers. In the analysis, the SVF was kept at a relatively constant value by dividing the dataset into sky view quantiles. This was done to further isolate the effect of the tree view factor and building view factor, and to analyze how the relationship changes under different sky openness scenarios. This showed that the LST correlation is strongest according to the Spearman ρ test, for both the trees and buildings when the sky is more obstructed. It seems therefore that at similar view factors, sky obstruction by trees has a cooling effect while sky obstruction by

buildings has a warming effect and that this effect increases in the lower quantiles of SVF. One possible explanation for the lack of observed cooling from the SVF is the difference in composition of the obstructed part between this case study and the areas observed in other studies.

The result of the regression models was also quite positive. The random forest achieves a MSE of 3.81 while being trained on 3D features from street view exclusively. This was the highest performing model compared to an MSE of 5.76, 4,72, and 6.02 for the SVR, GBT, and MLP respectively. One interesting observation here is that the most complex model out of the set, the MPL, is also the worst-performing model for this task. While some different model architectures were tested, due to the computational requirements for this model it was not within the scope of this study to do extensive hyperparameter tuning. Therefore it remains possible that with more computational resources to do more extensive hyperparameter tuning the model performance of this MLP could still improve drastically.

In the current literature, some of the most accurate models are around R^2 0.7 (Wei et al., 2023; Li et al., 2016). These studies also use models similar to the RF model and do not use complex deep learning models such as an MLP. In this context our RF model achieving a R^2 of 0.52 while both of these models in these papers include significantly more data sources is promising. In Wei et al. (2023), both 2D remote sensing data and street view data were used. The data was also aggregated here to a 210m cell size. Whereas, in this study, the data was kept at the highest LST resolution available of 30m. In Li et al. (2016), a combination of 2D remote sensing data and socio-economic indicators was used to predict LST. Model performance in this study was also limited because of a low number of street view images per cell used due to computational constraints. Increasing the number of street view images per cell could potentially increase the model performance significantly. This could be achieved by utilizing GPU acceleration for the reprojection of the panoramic images and taking away the computational bottleneck.

5.2. RESULT IN THE CASE STUDY CONTEXT

We can also formulate practical applications of these results by applying them to the Amsterdam context. Amsterdam has set out a climate adaptation strategy for their future city vision (Amsterdam, 2022). In this vision, they highlight the benefits of urban greening, such as biodiversity, climate adaptation, and health. The negative health impacts of heat are also highlighted as one of the major problems. This study indicates that areas with low tree or pervious surface coverage generally relate to higher LST values. This is also visible in some interesting cases. In the LST map 3.1 for example, it is visible that the new boulevard behind the central station next to the IJ River is a hot spot with regards to LST. Looking at the maps of TVF and pervious surfaces 4.9 4.3, it is visible that both show very low values for this area. At the same time, the SVF for this area is significantly larger than for the adjacent area more into the city center.

This hotspot is therefore likely a result of a high degree of openness, meaning that the surface receives a high amount of radiation, while the surface itself has a high thermal capacity and emissivity. Considering that this development is quite recent, and the strategy laid out by the Municipality itself in Amsterdam (2022) this is surprising. Other hotspots around the IJ River have similar problems.

Apart from the industrial areas, it is also visible that from comparing 3.1 and 4.8, that as the density of the build-up area increases the LST also rises. This is also consistent with the analysis of the relationship between the BVF and LST in 4.6. In Amsterdam, this means that the richer areas in the central and south areas of the city actually have higher LST values than less affluent neighborhoods in the southeast, north, and west of the city. This is interesting as previous research in the US indicates that less affluent areas often experience higher temperatures (Hsu et al., 2021).

5.3. Limitations

5.3.1. Physiological aspects of LST

This study has some limitations. First of all, LST is not directly equivalent to thermal comfort. In general, LST is accepted as an indicator of urban heat, but an air temperature or radiant temperature is often more desirable. Mean Radiant Temperature (MRT), which is used as a proxy for the heat exchange between an individual and their surroundings is often referred to as the best indicator of urban thermal comfort. While air temperature and LST have a significant positive correlation, research has shown that the hot and cold spots identified by both temperature variables do not always correspond (Cao et al., 2021; Naserikia et al., 2023). LST and MRT are both measures of radiation, there are however some key differences. MRT is a ground-based measure of radiation from a multitude of directions and at a significantly higher resolution, the resolution can differ but in this paper, 1m was used. Even though MRT provides a much more detailed picture on a hyper-local level, the general pattern of heat city-wide shows a similar trend compared to LST (Li et al., 2023). LST is captured by satellites and represents the average heat radiated back into space from each cell. At a 30m resolution, that means that hyperlocal differences in surface temperature are not accounted for. On top of that the satellite view takes only one direction of heat radiation into account. Because it registers heat radiated back into space the only direction that it captures is up.

There are also studies that indicate that the impact on thermal comfort of reducing LST is often overstated. For example, reflective pavements increase the albedo and reduce the LST by lowering the solar radiation absorption, but this solar radiation is then reflected back onto the street. This solar radiation can then still affect thermal comfort (Schneider et al., 2023). The impact of trees can also be overstated by LST. Trees absorb radiation and use this energy for evapotranspiration, this cools down the surface layer but at the same time increases the relative humidity. Higher humidity can have a negative effect on

thermal comfort. This tradeoff between higher humidity and lower surface temperature is not visible in surface-based temperature measurements(Chakraborty et al., 2022). Another limitation of the 30m resolution is, that 30m is usually about the size of a street segment, but it can occur that there are parts of a courtyard or backyard that are within the cell. In this instance, the street view would not be able to detect the contents of said courtyard or backyard leading to small inaccuracies in the dataset.

In this study, the focus was also on summertime daytime temperatures. However, research on the health effects of the UHI indicates that high nighttime temperatures are also an important factor in higher mortality rates as a result of heat (Murage et al., 2017). The focus on the summertime also means that the potential benefits of the UHI in wintertime are also not accounted for (Macintyre et al., 2021).

5.3.2. Proximity amplification

A second limitation is inherent in the segmentation of street-view images. The vantage point of the street-view image often dictates the size of the different features in the segmentation. If for example a street segment is captured with only one tree present close to the location where the image was captured, this can lead to an overrepresentation of the tree class in this image compared to the ground truth. The other way around if a feature is far away from the location of the picture it can also lead to underrepresentation of that class. To compensate for this the model needs additional spatial information such as depth estimation or the object location (Seiferling et al., 2017).

5.3.3. Data gaps in street-view and segmentation

The third limitation has to do with the segmentation as well. Most of the pre-trained semantic segmentation models such as the Ade20k dataset used in this research, but also the cityscapes dataset, are trained on Chinese or North American cities (Zhou et al., 2019). This means that the results for a complex urban environment such as Amsterdam are less accurate than they would be for case studies in China, the United States, or Canada.

On top of that semantic segmentation models can detect general street-level features but can not make distinctions within a certain feature class. Some of those within class differences however can have a significant impact on the urban temperatures. Building materials for example can differ both in heat capacity and radiation emissivity. Both of these material characteristics influence the temperature in their surroundings through their radiation output, heat capacity by the higher temporal duration of the output, and emissivity by the higher intensity of the output (Morais et al., 2019). The variations in this radiation output for both vertical features (building facades) and horizontal features (roads, sidewalks) are significant. For vertical features, there can be a 70% difference in performance, and for horizontal

features as much as 30% (Alchapar et al., 2014). The higher difference in vertical features is most likely due to the higher variation in used materials for facades. The difference in thermal performance, between, for example, glass and bricks is higher than the difference between asphalt and concrete (Tukiran et al., 2016; Vijayan et al., 2021). Adding this information to the prediction models could help to explain more of the variation in the target temperature dataset. However, currently, there are no semantic segmentation models, or they are not accessible, that can accurately classify material differences in an urban setting.

Another data gap in street view segmentation has to do with the characteristics of street view. This dataset is inherently a representation of the street network of the city. This means that the parts of a city that are not adjacent to a street are not taken into account in the analysis. This means that large green areas such as parks, courtyards, and even backyards are generally not included in the results. Even though these large green areas can have a significant cooling effect even in areas that are not immediately next to them (Peng et al., 2021).

6. Conclusion

In conclusion, this study delved into the intricate relationship between urban morphology and the urban heat island (UHI) phenomenon. While existing literature predominantly focused on 2D representations and microclimate modeling within controlled environments, this research aimed to bridge the gap by conducting a large-scale estimation of microclimate using 3D data extracted from street view images. The study emphasized the importance of considering the third dimension and maintaining a spatial scale relevant to real urban environments.

The feature analysis revealed compelling insights, highlighting the strong correlation between urban features, particularly buildings, and trees, with land surface temperature (LST). Contrary to expectations, the study found a weak correlation between LST and sky percentage, challenging previous assumptions about sky obstruction and its relationship with LST. The nuanced relationship between sky view factors (SVF) and LST, influenced by factors such as airflow and radiation, demonstrated the complexity of the UHI effect, especially during daytime.

Machine learning models, including Random Forest (RF), Support Vector Regression (SVR), Gradient Boosting Trees (GBT), and Multi-layer Perceptron (MLP), were employed to predict LST based on 3D features. The RF model emerged as the highest-performing model, showcasing promise with an impressive Mean Squared Error (MSE) of 3.66. While the MLP model displayed the lowest performance, limitations in hyperparameter tuning underscored the potential for further improvement with additional computational resources.

In the context of the case study focused on Amsterdam, the practical applications of the research findings were explored. The study identified areas with low tree or pervious surface coverage as potential hotspots, aligning with the city's climate adaptation strategy. Notably, affluent areas exhibited higher LST values, challenging conventional expectations and emphasizing the need for context-specific urban planning.

However, the study acknowledged certain limitations, including the reliance on Land Surface Temperature (LST) as an indicator of thermal comfort, potential overestimation of certain features in street-view segmentation, and data gaps related to the representativeness of pre-trained segmentation models for complex urban environments like Amsterdam.

In conclusion, this research contributes valuable insights to the understanding of UHI dynamics and provides a foundation for informed urban planning strategies. The findings underscore the necessity of considering 3D features at a large spatial scale and emphasize the potential of machine learning models in predicting and mitigating the impacts of urban heat islands.

7. References

References

Albawi, S., Mohammed, T. A., and Al-Zawi, S. (2017). Understanding of a convolutional neural network. In *2017 international conference on engineering and technology (ICET)*, pages 1–6. Ieee.

Alchapar, N. L., Correa, E. N., and Cantón, M. A. (2014). Classification of building materials used in the urban envelopes according to their capacity for mitigation of the urban heat island in semiarid zones. *Energy and Buildings*, 69:22–32.

Amsterdam, M. o. (2022). Our city of tomorrow.

Cao, J., Zhou, W., Zheng, Z., Ren, T., and Wang, W. (2021). Within-city spatial and temporal heterogeneity of air temperature and its relationship with land surface temperature. *Landscape and Urban Planning*, 206:103979.

Chakraborty, T., Venter, Z., Qian, Y., and Lee, X. (2022). Lower urban humidity moderates outdoor heat stress. *Agu Advances*, 3(5):e2022AV000729.

Chen, C., Bagan, H., Yoshida, T., Borjigin, H., and Gao, J. (2022). Quantitative analysis of the building-level relationship between building form and land surface temperature using airborne lidar and thermal infrared data. *Urban Climate*, 45:101248.

Choe, Y.-J. and Yom, J.-H. (2017). Downscaling of modis land surface temperature to landsat scale using multi-layer perceptron. *Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography*, 35(4):313–318.

Choi, K., Lim, W., Chang, B., Jeong, J., Kim, I., Park, C.-R., and Ko, D. W. (2022). An automatic approach for tree species detection and profile estimation of urban street trees using deep learning and google street view images. *Isprs Journal of Photogrammetry and Remote Sensing*, 190:165–180.

Estrada, F., Botzen, W. W., and Tol, R. S. (2017). A global economic assessment of city policies to reduce climate change impacts. *Nature Climate Change*, 7(6):403–406.

Fonti, V. and Belitser, E. (2017). Feature selection using lasso. *VU Amsterdam research paper in business analytics*, 30:1–25.

Gong, F.-Y., Zeng, Z.-C., Zhang, F., Li, X., Ng, E., and Norford, L. K. (2018). Mapping sky, tree, and building view factors of street canyons in a high-density urban environment. *Building and Environment*, 134:155–167.

Guo, G., Zhou, X., Wu, Z., Xiao, R., and Chen, Y. (2016). Characterizing the impact of urban morphology heterogeneity on land surface temperature in guangzhou, china. *Environmental Modelling & Software*, 84:427–439.

He, B.-J., Ding, L., and Prasad, D. (2020). Urban ventilation and its potential for local warming mitigation: A field experiment in an open low-rise gridiron precinct. *Sustainable Cities and Society*, 55:102028.

Heaviside, C., Macintyre, H., and Vardoulakis, S. (2017). The urban heat island: implications for health in a changing environment. *Current environmental health reports*, 4:296–305.

Hendrycks, D. and Gimpel, K. (2016). Gaussian error linear units (gelus). *arXiv preprint arXiv:1606.08415*.

Hsu, A., Sheriff, G., Chakraborty, T., and Manya, D. (2021). Disproportionate exposure to urban heat island intensity across major us cities. *Nature communications*, 12(1):2721.

Huang, C., Barnett, A. G., Wang, X., Vaneckova, P., FitzGerald, G., and Tong, S. (2011). Projecting future heat-related mortality under climate change scenarios: a systematic review. *Environmental health perspectives*, 119(12):1681–1690.

Huang, K.-T., Lin, T.-P., Lien, H.-C., et al. (2015). Investigating thermal comfort and user behaviors in outdoor spaces: A seasonal and spatial perspective. *Advances in Meteorology*, 2015.

Huang, X. and Wang, Y. (2019). Investigating the effects of 3d urban morphology on the surface urban heat island effect in urban functional zones by using high-resolution remote sensing data: A case study of wuhan, central china. *ISPRS Journal of Photogrammetry and Remote Sensing*, 152:119–131.

IPCC (2021). Climate change 2021: The physical science basis. contribution of working group i to the sixth assessment report of the intergovernmental panel on climate change [masson-delmotte, v., et al. (eds.)]. *Intergovernmental Panel for Climate Change*. https://www.ipcc.ch/report/ar6/wg1/.

Kalnay, E. and Cai, M. (2003). Impact of urbanization and land-use change on climate. *Nature*, 423(6939):528–531.

Kenton, J. D. M.-W. C. and Toutanova, L. K. (2019). Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of naacL-HLT*, volume 1, page 2.

Khalil, U., Aslam, B., Azam, U., and Khalid, H. M. D. (2021). Time series analysis of land surface temperature and drivers of urban heat island effect based on remotely sensed data to develop a prediction model. *Applied Artificial Intelligence*, 35(15):1803–1828.

Kim, P. and Kim, P. (2017). Convolutional neural network. *MATLAB deep learning: with machine learning, neural networks and artificial intelligence*, pages 121–147.

Li, X., Chakraborty, T. C., and Wang, G. (2023). Comparing land surface temperature and mean radiant temperature for urban heat mapping in philadelphia. *Urban Climate*, 51:101615.

Li, X., Li, W., Middel, A., Harlan, S. L., Brazel, A. J., and Turner Ii, B. (2016). Remote sensing of the surface urban heat island and land architecture in phoenix, arizona: Combined effects of land composition and configuration and cadastral–demographic–economic factors. *Remote Sensing of Environment*, 174:233–243.

Li, X. and Ratti, C. (2018). Mapping the spatial distribution of shade provision of street trees in boston using google street view panoramas. *Urban Forestry Urban Greening*, 31:109–119.

Li, X., Ratti, C., and Seiferling, I. (2017). Mapping urban landscapes along streets using google street view. In *Advances in Cartography and GIScience: Selections from the International Cartographic Conference 2017 28*, pages 341–356. Springer.

Li, Z., Liu, F., Yang, W., Peng, S., and Zhou, J. (2022). A survey of convolutional neural networks: Analysis, applications, and prospects. *IEEE Transactions on Neural Networks and Learning Systems*, 33(12):6999–7019.

Loughnan, M., Nicholls, N., and Tapper, N. J. (2012). Mapping heat health risks in urban areas. *International Journal of Population Research*, 2012.

Macintyre, H. L., Heaviside, C., Cai, X., and Phalkey, R. (2021). The winter urban heat island: Impacts on cold-related mortality in a highly urbanized european region for present and future climate. *Environment International*, 154:106530.

Maskey, M., Ramachandran, R., Miller, J. J., Zhang, J., and Gurung, I. (2018). Earth science deep learning: Applications and lessons learned. In *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*, pages 1760–1763. IEEE.

Massaro, E., Schifanella, R., Piccardo, M., Caporaso, L., Taubenböck, H., Cescatti, A., and Duveiller, G. (2023). Spatially-optimized urban greening for reduction of population exposure to land surface temperature extremes. *Nature Communications*, 14(1):2903.

Mooney, A. (2023). Is 1.5c still realistic? the crumbling consensus over key climate targets. *Financial Times*.

Morais, M. V. B. d., Urbina Guerrero, V. V., Freitas, E. D. d., Marciotto, E. R., Valdés, H., Correa, C., Agredano, R., and Vera-Puerto, I. (2019). Sensitivity of radiative and thermal properties of building material in the urban atmosphere. *Sustainability*, 11(23):6865.

Murage, P., Hajat, S., and Kovats, R. S. (2017). Effect of night-time temperatures on cause and age-specific mortality in london. *Environmental Epidemiology (Philadelphia, Pa.)*, 1(2):e005.

Naserikia, M., Hart, M. A., Nazarian, N., Bechtel, B., Lipson, M., and Nice, K. A. (2023). Land surface and air temperature dynamics: The role of urban form and seasonality. *Science of The Total Environment*, 905:167306.

Nazarian, N. and Kleissl, J. (2015). Cfd simulation of an idealized urban environment: Thermal effects of geometrical characteristics and surface materials. *Urban Climate*, 12:141–159.

Oke, T. R. (1981). Canyon geometry and the nocturnal urban heat island: comparison of scale model and field observations. *Journal of climatology*, 1(3):237–254.

Oke, T. R. (1982). The energetic basis of the urban heat island. *Quarterly journal of the royal meteorological society*, 108(455):1–24.

O'Shea, K. and Nash, R. (2015). An introduction to convolutional neural networks. *arXiv preprint arXiv:1511.08458*.

Peng, J., Dan, Y., Qiao, R., Liu, Y., Dong, J., and Wu, J. (2021). How to quantify the cooling effect of urban parks? linking maximum and accumulation perspectives. *Remote Sensing of Environment*, 252:112135.

Ranftl, R., Bochkovskiy, A., and Koltun, V. (2021). Vision transformers for dense prediction. *ArXiv preprint*.

Raymond, C., Matthews, T., and Horton, R. M. (2020). The emergence of heat and humidity too severe for human tolerance. *Science Advances*, 6(19):eaaw1838.

Rizwan, A. M., Dennis, L. Y., and Chunho, L. (2008). A review on the generation, determination and mitigation of urban heat island. *Journal of environmental sciences*, 20(1):120–128.

Scarano, M. and Mancini, F. (2017). Assessing the relationship between sky view factor and land surface temperature to the spatial resolution. *International Journal of Remote Sensing*, 38(23):6910–6929.

Scarano, M. and Sobrino, J. (2015). On the relationship between the sky view factor and the land surface temperature derived by landsat-8 images in bari, italy. *International Journal of Remote Sensing*, 36(19-20):4820–4835.

Schmidhuber, J. (2015). Deep learning. *Scholarpedia*, 10(11):32832.

Schneider, F. A., Ortiz, J. C., Vanos, J. K., Sailor, D. J., and Middel, A. (2023). Evidence-based guidance on reflective pavement for urban heat mitigation in arizona. *Nature communications*, 14(1):1467.

Seiferling, I., Naik, N., Ratti, C., and Proulx, R. (2017). Green streets- quantifying and mapping urban trees with street-level imagery and computer vision. *Landscape and Urban Planning*, 165:93–101.

Song, J., Du, S., Feng, X., and Guo, L. (2014). The relationships between landscape compositions and land surface temperature: Quantifying their resolution sensitivity with spatial regression models. *Landscape and urban planning*, 123:145–157.

Steeneveld, G., Koopmans, S., Heusinkveld, B., van Hove, B., and Holtslag, A. (2011). Quantifying urban heat island effects and human comfort for cities of variable size and urban morphology in the netherlands. *Journal of Geophysical Research: Atmospheres*, 116.

Stewart, I. D. (2011). A systematic review and scientific critique of methodology in modern urban heat island literature. *International Journal of Climatology*, 31(2):200–217.

Tuholske, C., Caylor, K., Funk, C., Verdin, A., Sweeney, S., Grace, K., Peterson, P., and Evans, T. (2021). Global urban population exposure to extreme heat. *Proceedings of the National Academy of Sciences*, 118(41):e2024792118.

Tukiran, J. M., Ariffin, J., and Ghani, A. N. A. (2016). Comparison on colored coating for asphalt and concrete pavement based on thermal performance and cooling effect. *J. Teknol*, 78(5):63–70.

U.S. Environmental Protection Agency (2023). Learn About Heat Islands. Website. Accessed: October 2023.

van Hove, L., Jacobs, C., Heusinkveld, B., Elbers, J., van Driel, B., and Holtslag, A. (2015). Temporal and spatial variability of urban heat island and thermal comfort within the rotterdam agglomeration. *Building and Environment*, 83:91–103. Special Issue: Climate adaptation in cities.

Vijayan, D., Mohan, A., Revathy, J., Parthiban, D., and Varatharajan, R. (2021). Evaluation of the impact of thermal performance on various building bricks and blocks: A review. *Environmental Technology & Innovation*, 23:101577.

Voogt, J. A. and Grimmond, C. (2000). Modeling surface sensible heat flux using surface radiative temperatures in a simple urban area. *Journal of Applied Meteorology*, 39(10):1679–1699.

Wang, L., Gao, Z., Miao, S., Guo, X., Sun, T., Liu, M., and Li, D. (2015). Contrasting characteristics of the surface energy balance between the urban and rural areas of beijing. *Advances in Atmospheric Sciences*, 32:505–514.

Watson, K. E., Gardiner, K. M., and Singleton, J. A. (2020). The impact of extreme heat events on hospital admissions to the royal hobart hospital. *Journal of Public Health*, 42(2):333–339.

Wei, X., Guan, F., Zhang, X., Van de Weghe, N., and Huang, H. (2023). Integrating planar and vertical environmental features for modelling land surface temperature based on street view images and land cover data. *Building and Environment*, 235:110231.

Xu, H., Wang, C. C., Shen, X., and Zlatanova, S. (2021). 3d tree reconstruction in support of urban microclimate simulation: a comprehensive literature review. *Buildings*, 11(9):417.

Yıldız, N. D., Erdem, F., Acet, S. B., and Avdan, U. (2023). Analyzing the effect of view factors on surface heat flux, surface temperature, and vegetation cover. *Environmental Science and Pollution Research*, 30(15):43843–43859.

Zeng, L., Lu, J., Li, W., and Li, Y. (2018). A fast approach for large-scale sky view factor estimation using street view images. *Building and Environment*, 135:74–84.

Zhang, Y., Middel, A., and Turner, B. (2019). Evaluating the effect of 3d urban form on neighborhood land surface temperature using google street view and geographically weighted regression. *Landscape Ecology*, 34:681–697.

Zhao, Q., Guo, Y., Ye, T., Gasparrini, A., Tong, S., Overcenco, A., Urban, A., Schneider, A., Entezari, A., Vicedo-Cabrera, A. M., et al. (2021). Global, regional, and national burden of mortality associated with non-optimal ambient temperatures from 2000 to 2019: a three-stage modelling study. *The Lancet Planetary Health*, 5(7):e415–e425.

Zhou, B., Zhao, H., Puig, X., Xiao, T., Fidler, S., Barriuso, A., and Torralba, A. (2019). Semantic understanding of scenes through the ade20k dataset. *International Journal of Computer Vision*, 127(3):302–321.

Zhou, J., Chen, Y., Wang, J., and Zhan, W. (2010). Maximum nighttime urban heat island (uhi) intensity simulation by integrating remotely sensed data and meteorological observations. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 4(1):138–146.

Zhou, W., Wang, J., and Cadenasso, M. L. (2017). Effects of the spatial configuration of trees on urban heat mitigation: A comparative study. *Remote Sensing of Environment*, 195:1–12.