DISSERTATIONES
GEOGRAPHICAE
UNIVERSITATIS
TARTUENSIS
88

ISAAC NEWTON KWASI BUO

Multi-scale Thermal Remote Sensing, Machine Learning and Radiative Flux Modeling to Assess Urban Overheating

DISSERTATIONES GEOGRAPHICAE UNIVERSITATIS TARTUENSIS
88

DISSERTATIONES GEOGRAPHICAE UNIVERSITATIS TARTUENSIS
88

ISAAC NEWTON KWASI BUO

Multi-scale Thermal Remote Sensing, Machine Learning and Radiative Flux Modeling to Assess Urban Overheating

UNIVERSITY OF TARTU
Press
1632

Department of Geography, Institute of Ecology and Earth Sciences, Faculty of Science and Technology, University of Tartu, Estonia.

This dissertation has been accepted for the commencement of the degree of Doctor of Philosophy in geoinformatics at the University of Tartu on 5th of June 2023 by the Scientific Council of the Institute of Ecology and Earth Sciences, University of Tartu.

Supervisors: Dr. Valentina Sagris
Institute of Ecology and Earth Sciences
University of Tartu
Estonia

Prof. Jaak Jaagus
Institute of Ecology and Earth Sciences
University of Tartu
Estonia

Opponent: Prof. Krzysztof Fortuniak
Institute of Meteorology and Hydrology
University of Lodz
Poland

Commencement: Senate Hall, University Main Building, Ülikooli 18, Tartu, on 29th of August 2023, at 14:15.

Publication of this dissertation is granted by the Institute of Ecology and Earth Sciences, University of Tartu.

ISSN 1406-1295 (print)
ISBN 978-9916-27-250-3 (print)
ISSN 2806-2302 (pdf)
ISBN 978-9916-27-251-0 (pdf)

University of Tartu Press
www.tyk.ee

TABLE OF CONTENTS

LIST OF ORIGINAL PUBLICATIONS 7

LIST OF ABBREVIATIONS 8

1. INTRODUCTION 10
   1.1 Urban heat islands, heatwaves, and urban overheating 10
   1.2 Assessing urban overheating at different urban climate scales 10
   1.3 Assessing UHI and overheating with remote sensing at the local scale 11
   1.4 Improving remote sensing datasets to assess SUHI and overheating at the local scale 12
   1.5 Microscale assessment of urban overheating for heat action planning 13
   1.6 Research questions 14

2. MATERIALS AND METHODS 16
   2.1 Enhancing remote sensing products to assess heatwaves in Estonia at local scales 16
      2.1.1 Overview of the random forest algorithm 16
      2.1.2 Seamless LST products for countrywide evaluation of heatwaves 17
      2.1.3 Synthetic NDVI and LST for city-wide evaluation of heatwaves 20
   2.2 Microscale thermal data and shade distribution for heat action planning 24
      2.2.1 Study sites 25
      2.2.2 Mean radiant temperature modeling with SOLWEIG 25
      2.2.3 Validation of the SOLWEIG model 26
      2.2.4 Sidewalk shade assessment 27

3. RESULTS 29
   3.1 Enhancing remote sensing products to assess heatwaves in Estonia at local scales 29
      3.1.1 Seamless LST products for countrywide evaluation of heatwaves 29
      3.1.2 Synthetic NDVI and LST for city-wide evaluation of heatwaves 31
   3.2 Microscale thermal data and shade distribution for heat action planning 39
      3.2.1 Validation of the SOLWEIG model 39
      3.2.2 Hourly mean radiant temperature and shade maps 40
      3.2.3 Sidewalk shade assessment 41

4. DISCUSSION 44
4.1 Enhancing remote sensing products to assess heatwaves in Estonia at local scales 44
4.1.1 Gap-filling satellite LST 44
4.1.2 Generating synthetic NDVI and LST products 45
4.1.3 Evaluation of the models and resulting products 47
4.1.4 Limitations and prospects 47
4.2 Microscale thermal data and shade distribution for heat action planning 48
4.2.1 Validation of the SOLWEIG model 48
4.2.2 Sidewalk shade coverage 49
4.2.3 Limitations and prospects 50
5. CONCLUSIONS 51
REFERENCES 52
SUMMARY IN ESTONIAN 64
ACKNOWLEDGEMENTS 67
PUBLICATIONS 69
CURRICULUM VITAE 108
ELULOOKIRJELDUS 110

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following publications, which are referred to in the text by Roman numerals:

Published papers are reproduced in print with the permission of the publisher.

I **Buo, I.**, Sagris, V., Jaagus, J., 2022. Gap-Filling Satellite Land Surface Temperature Over Heatwave Periods with Machine Learning. *IEEE Geoscience and Remote Sensing Letters*,19. http://dx.doi.org/10.1109/LGRS.2021.3068069

II **Buo, I.**, Sagris, V., Jaagus, J., 2022. Generating synthetic daily remote sensing products suitable for surface heat island and heatwaves assessments at urban scale. *Int. J. Environ. Sci. Technol*. https://doi.org/10.1007/s13762-022-04510-3

III **Buo, I.**, Sagris, V., Jaagus, J., Middel, A. 2023. High-resolution thermal exposure and shade maps for cool corridor planning. *Sustainable Cities and Society* 93. https://doi.org/10.1016/j.scs.2023.104499

Author's contribution to the articles denotes: '*' a minor contribution, '**' a moderate contribution, '***' a major contribution.

	Articles		
	I	II	III
Original idea	***	***	***
Study design	***	***	***
Data processing and analysis	***	***	***
Interpretation of the results	***	***	***
Writing the manuscript	***	***	***

LIST OF ABBREVIATIONS

ASTER	Advanced spaceborne thermal emission and reflection radiometer
AZMET	Arizona meteorological network
BSM	Building surface model
BUHI	Boundary layer urban heat island
CDSM	Canopy digital surface model
CUD	Central urban district
CUHI	Canopy layer urban heat island
d	Willmott's index of agreement
DEM	Digital elevation model
DisTrad	Disaggregation procedure for radiometric surface temperature
DOY	Day of the year
ETAK	Estonian topography database
EWS	Estonian weather service
FDSM	Full digital surface model
GEE	Google Earth engine
ISO	International organization for standardization
LCZ	Local climate zone
LST	Land surface temperature
LULC	Land use and land cover
MAE	Mean absolute error
MBE	Mean bias error
ML	Machine learning
MODIS	Moderate resolution imaging spectroradiometer
NDVI	Normalized difference vegetation index
PET	Physiologically equivalent temperature
OLS	Ordinary least squares
PDP	Partial dependence plot
R	Coefficient of correlation
R^2	Coefficient of determination
RF	Random forest
RMSE	Root mean square error
RMSEz	Systematic root mean square error
RMSEu	Unsystematic root mean square error
SAT	Surface-air temperature
SE	Statistics Estonia
SOLWEIG	Solar longwave environmental irradiance geometry
SUHI	Surface urban heat island

T_{MRT}	Mean radiant temperature
TsHARP	Thermal sharpening
TUA	Tallinn urban agglomeration
UTCI	Universal thermal comfort index
UHI	Urban heat island
WUDAPT	World urban database and access portal tools

1. INTRODUCTION

1.1 Urban heat islands, heatwaves, and urban overheating

Presently, over half of the world's population lives in cities. Estimates indicate that by 2050, 2.5 billion people will live in urban areas (Nazarian et al., 2022; Wahba Tadros et al., 2021). While urbanization benefits economic growth (Wahba Tadros et al., 2021), it also creates a never-ending demand for housing and other amenities to ensure comfortable living. Pre-urban land cover is replaced with artificial materials that cause urban areas to be warmer than their surrounding rural areas (Giridharan & Emmanuel, 2018; Haddad et al., 2020; Oke, 1973; Rajulapati et al., 2022), an effect known as the Urban Heat Island (UHI).

Heatwaves are death-dealing natural disasters characterized by sustained high temperatures that threaten human livability and impact regional economies (Meehl & Tebaldi, 2004; Rajulapati et al., 2022). For example, over 30,000 heat-related deaths were recorded across Europe during the 2003 summer heatwave (Basara et al., 2010; Cheval et al., 2009; Li & Bou-Zeid, 2013). Heatwaves have increased in frequency in the 21st century due to climate change (Banholzer et al., 2014; Perkins-Kirkpatrick & Lewis, 2020). Climate change projections reveal that the frequency, duration, and intensity of heatwaves will likely increase further (Basara et al., 2010; García-León et al., 2021; Meehl & Tebaldi, 2004; Seneviratne et al., 2012).

While heatwaves elevate the ambient temperatures of rural and urban areas, their associated risks are exacerbated in urban areas because of the physical characteristics and anthropogenic activities responsible for the UHI effect. In the event of heatwaves, urban areas become overheated, i.e., temperatures exceed locally set thermal thresholds, resulting in an adverse impact on people's health, comfort, and productivity, as well as related urban systems (García-León et al., 2021; Nazarian et al., 2022; Perkins-Kirkpatrick & Lewis, 2020). The thresholds can be absolute or percentile-based (Perkins & Alexander, 2013).

The severe impacts of urban overheating require local mitigation strategies and early warning systems. Evaluating heat hazards at different spatial scales, i.e., city-wide or on the street, building (indoor/outdoor), and human (indoor/outdoor) scales, will inform mitigation strategies and increase urban resilience to heat hazards (Imhoff et al., 2010; Middel & Krayenhoff, 2019; Nazarian & Lee, 2021).

1.2 Assessing urban overheating at different urban climate scales

The arrangement of urban surfaces, relevant thermal metrics, and sensor capabilities determine the scale at which urban overheating is assessed (Nazarian et al., 2022; Oke et al., 2017). A basic unit of the urban surface is the urban (building) canyon, which constitutes a street and the buildings adjacent to it

(Middel et al., 2019; Oke et al., 2017). A group of adjacent urban canyons with identical structures forming regular shapes (squares, rectangles, or triangles) or irregular shapes make up a block (Oke et al., 2017). Up the hierarchy after the block is the local (also neighborhood) scale, which includes land use classes frequently used in urban planning (Stewart & Oke, 2012). The final scale of the hierarchy is the city, which comprises different neighborhoods.

Overheating in urban canyons, quantified by metrics such as air temperature and Mean Radiant Temperature (T_{MRT}), can be observed using weather stations, net radiometers, and globe thermometers (Crank et al., 2020; Middel et al., 2019; Nazarian et al., 2022). At the block, local, and city scales, Surface-air Temperature (SAT) measurements made at 2 m above ground, Land Surface Temperature (LST), and intra-urban temperature variability are some of the relevant metrics (Nazarian et al., 2022). Sensing instruments used at the block, local, and city scales include meteorological networks, remote sensors aboard Earth-orbiting satellites, and terrestrial mobile sensors (Oke et al., 2017; Voogt & Oke, 2003). Advances in computational capabilities have allowed for evaluating overheating at different scales using climate models (Bruse & Fleer, 1998; Chen et al., 2011; Järvi et al., 2011; Lindberg et al., 2008; Matzarakis et al., 2010; Middel et al., 2019; Steuri et al., 2020).

1.3 Assessing UHI and overheating with remote sensing at the local scale

UHI describes the difference between urban and rural thermal environments (Fortuniak et al., 2006). The UHI is qualified as either Surface UHI (SUHI), Canopy layer UHI (CUHI), or Boundary layer UHI (BUHI) depending on the thermal metric and height of measurement (Oke et al., 2017; Zhou et al., 2019). The radiative temperature of urban and rural surfaces describes SUHI (Deilami et al., 2018; Voogt & Oke, 2003). Air temperatures between the surface and the roof level (urban canopy layer) characterize the CUHI, while those between the roof and urban boundary layer characterize the BUHI (Cheval et al., 2022; Zhou et al., 2019).

The application of remote sensing to detect UHIs and assess overheating at local scales has continued to increase since Rao (1972), according to Voogt & Oke (2003). The remote sensing approach allows assessing the UHI at larger scales, which is challenging to achieve using weather stations (Zhou et al., 2019). Thermal remote sensing is practical for detecting SUHIs because the thermal sensors record the upwelling thermal radiance from the Earth's surface (Voogt & Oke, 2003). The Land Surface Temperature (LST) parameter, calculated from the thermal radiance, helps identify thermal hotspots and determine SUHI intensities (Kapwata et al., 2022; Keellings & Moradkhani, 2020; Schwarz et al., 2011). In addition to the thermal radiance, remote sensing imagery provides other measurements useful in classifying Land Use and Land Cover (LULC) and calculating metrics such as the Normalized Difference Vegetation Index (NDVI) (Fichera

et al., 2012; Tucker, 1979). When combined with LST, these additional products give insight into how different surface types contribute to the urban surface energy balance and their relationships with LST (Osborne & Alvares-Sanches, 2019; Voogt & Oke, 2003; Weng, 2009; Zhou et al., 2019).

Despite all the advantages remote sensing provides, there are limitations. Usually, a trade-off between spatial and temporal resolutions must be made when using remote sensing data to understand urban overheating and UHI (Nazarian et al., 2022). Properties such as satellite altitude, orbit, and sensor viewing angle influence the remotely sensed images' temporal and spatial resolutions. For example, the Moderate Resolution Image Spectroradiometer (MODIS) LST has a high temporal resolution (daily) and a coarse spatial resolution of 1 km. On the other hand, Landsat satellites provide LST at a spatial resolution of 30 m (re-sampled from 90 m) every 16 days. Furthermore, there is no guarantee that a sensor will provide seamless images because clouds often obstruct a sensor's view, resulting in gaps (Sarafanov et al., 2020). Effective anisotropy caused by a sensor's biased view of the Earth's surface introduces errors in LST estimates (Voogt & Oke, 2003; Zhou et al., 2019).

1.4 Improving remote sensing datasets to assess SUHI and overheating at the local scale

Reconstructing missing data and enhancing the spatial resolution of remotely sensed images has received much attention. Crosson et al. (2012) explored creating seamless MODIS LST products by fusing data from the MODIS Aqua and Terra satellites. Shen et al. (2016) investigated combining data from different sensors to fill gaps. Spatial and temporal interpolations using geostatistics have also been used to fill gaps (Weiss et al., 2014; Zhang et al., 2022). Other studies relied on the empirical relationships between LST and other parameters, such as NDVI, to determine LST for missing pixels (Fan et al., 2014; Ke et al., 2013). Approaches combining several methods have also been considered to fill gaps; for example, Li et al. (2018) used merging and spatio-temporal interpolation to create seamless LST for the contiguous United States, while Yao et al. (2021) used interpolation and the relationship between similar LST products. Metz et al. (2017) also combined interpolation with emissivity and elevation information to predict missing data in MODIS LST.

Generating synthetic LST products at higher resolution (also downscaling) has been proposed to make data available at desirable spatial and temporal resolution (Bechtel et al., 2012; Hutengs & Vohland, 2016; Pu & Bonafoni, 2023). Wang et al. (2020) downscaled 1 km MODIS LST to 100 m using a geographically weighted autoregressive model. Weng et al. (2014) fused Landsat and MODIS data to produce daily LST products without gaps. The Disaggregation procedure for radiometric surface temperature (DisTrad) and Thermal Sharpening (TsHARP) are popular in generating high-resolution synthetic LST (Agam et al., 2007; Huryna et al., 2019; Kustas et al., 2003). TsHARP and DisTrad are techniques that use

least-squares expressions built between LST and a vegetation cover parameter at a coarse resolution to generate LST at higher resolutions (Agam et al., 2007; Kustas et al., 2003).

The use of Machine Learning (ML) models to fill gaps and downscale coarse-resolution LST data has gained much attention lately (Bisquert et al., 2016; Hutengs & Vohland, 2016; Li et al., 2019; Sarafanov et al., 2020). The models rely on the relationships between LST and various predictors. The relationships are often challenging to model using conventional statistical models. Osborne & Alvares-Sanches (2019) preferred ML models to evaluate the relationship between LST and landscape compositions over statistical models. Li et al. (2019) found ML algorithms superior to the widely used TsHARP algorithm in down-scaling coarse-resolution LST. Using an ML-based geostatistical method, Xu et al. (2020) downscaled a 90-m resolution ASTER LST to 10 m.

Most studies have focused on filling gaps or downscaling LST for a single day or a few days to prove an algorithm's suitability. Little has been done to generate seamless daily LST datasets covering heatwave periods, with which SUHI and urban overheating could be evaluated. Furthermore, attempts using ML models limit model evaluation to a few regression evaluation metrics, such as Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Bias Error (MBE). The explanation of how the models arrive at estimations and the suitability of the resulting product for different land cover types have received little to no attention.

1.5 Microscale assessment of urban overheating for heat action planning

Air temperature is a standard metric to evaluate urban overheating using sensor networks or mobile measurements at microscales (Crank et al., 2020; Nazarian et al., 2022; Oke et al., 2017). High-resolution satellite LST has also been used at the microscale (Schwaab et al., 2021; Zhou et al., 2020). Quantifying the impacts of overheating on urban dwellers at the microscale is essential. Air temperature and LST are inadequate because they do not explicitly represent the heat load on a person. Air temperature measurements do not account for the radiative exchange between a person outdoors and their surroundings. Satellite LST is limited to the horizontal plane and objects in a satellite's instantaneous field of view; the surface temperatures of rooftops are not beneficial in evaluating pedestrian thermal comfort (Nazarian et al., 2022; Voogt & Oke, 2003).

Shade distribution and T_{MRT} are essential parameters for assessing outdoor pedestrian thermal comfort (Dzyuban et al., 2022; Tabatabaie et al., 2019). Reduced UV exposure, thermal discomfort, and heat stress are all benefits of shade (Aleksandrowicz et al., 2020; Middel et al., 2021; Schwaab et al., 2021; Tabatabaie et al., 2019). According to Handy et al. (2002), tree shade influences people's physical activity and improves the aesthetic appeal of streetscapes. Many municipalities worldwide have focused their heat mitigation plans around street trees to increase the thermal comfort of pedestrians (Jamei & Rajagopalan,

2017; Speak et al., 2021; Wallenberg et al., 2022). For instance, the Victorian government aims to increase the percentage of trees in the City North (Melbourne) from 14 to 40% by the year 2050 (Jamei & Rajagopalan, 2017). Likewise, Brisbane and Bendigo, Australia, implemented the "*shadeways*" concept, which considers street trees crucial to improving pedestrian thermal (Butt et al., 2019; Deilami et al., 2020).

T_{MRT} quantifies the human body's heat load and is significant in assessing outdoor thermal comfort in cities (Guo et al., 2020; Hardy & Stoll, 1954; Vanos et al., 2010). T_{MRT} represents all radiative fluxes (short and longwave) (Höppe, 1992; Lindberg et al., 2008; Nazarian et al., 2022). It is a fundamental parameter for calculating biometeorological indices such as Physiologically Equivalent Temperature (PET) and the Universal Thermal Comfort Index (UTCI) (Crank et al., 2020; Krüger et al., 2014). A set of three net radiometers with sensors pointing upward, downward, and in four cardinal directions, or globe thermometers, can be used to measure T_{MRT} (Aviv et al., 2022; Guo et al., 2020; Middel et al., 2016; Middel & Krayenhoff, 2019; Vanos et al., 2021). Although direct measurements give valuable insight into a person's heat load in a particular space at a given time, spatially explicit and continuous T_{MRT} and shade maps are needed to optimize site selection for municipal heat action plans.

Following Höppe (1992), Lindberg et al. (2008) developed the SOlar Long-Wave Environmental Irradiance Geometry (SOLWEIG) model, which simulates high-resolution, spatially precise shadow patterns, radiative fluxes, and T_{MRT} for large areas. SOLWEIG has been used in numerous geographic contexts to estimate T_{MRT} in heterogeneous urban settings (Acero & Arrizabalaga, 2018; Aleksandrowicz et al., 2020; Gál & Kántor, 2020; Kong et al., 2022). In previous studies, SOLWEIG has performed better than other radiation flux models (Chen et al., 2014; Gál & Kántor, 2020; Szucs et al., 2014). Observational data used in previous validation studies were limited to a single day or several days with only a few sites (Aminipouri et al., 2019; Chen et al., 2014; Gál & Kántor, 2020). The model's estimates under different exposure types and urban morphology have not been thoroughly validated.

1.6 Research questions

This thesis generates thermal and shade distribution datasets to evaluate urban overheating at local and micro urban climate scales.

The thesis investigates the following questions at the local scale:

- How can ML algorithms such as Random Forest (RF) reconstruct remote sensing data and generate synthetic products to assess heatwaves effectively?
- Which predictors are influential in reconstructing satellite LST and producing synthetic NDVI and LST with ML, and what relationships do they have with the target variables?
- How valuable is a synthetic NDVI product in predicting synthetic LST?

Articles I and II addressed the effectiveness of the random forest algorithm in reconstructing and generating (synthetic) remotely sensed data. **Article I** focused on filling gaps in MODIS LST over heatwave periods in Estonia. A random forest model was built using NDVI, surface type fractions, spatial parameters (elevation, distance to the coast, and centroid coordinates), and temporal variables (day-of-the-year and year) as predictors. **Article I** also evaluated the performance of the gap-filling random forest model with *gapfill,* an established R-based gap-filling algorithm. **Article II** used RF models to generate synthetic daily LST and NDVI products for Tallinn, Estonia. Two NDVI models were created to predict MODIS-like and Landsat 8-like NDVI at 500 m and 100 m resolutions using LULC fractions and day-of-the-year (DOY) as predictors. The synthetic LST products were created using the synthetic NDVI and other features following the approach in **Article I**. **Articles I and II** investigated the importance of predictors in making estimations and the relationships between predictors and target variables using Partial Dependence Plots (PDPs). **Article II** evaluated the suitability of synthetic NDVI to predict LST at target resolutions. Furthermore, **Article II** evaluated the synthetic products to identify how accurate the models were in making estimations for different LULC types.

Regarding the datasets at the microclimatic scale, the thesis addressed the following questions:

- How accurately does the SOLWEIG model estimate T_{MRT} for various urban forms and exposure types?
- How can the shade available to pedestrians in a large metropolis be quantified for heat mitigation plans?

Both issues were addressed in **Article III** using the cities of Tempe and Phoenix in the Phoenix metropolitan area, USA, as case studies. The suitability of SOLWEIG to estimate T_{MRT} for different exposure types (sun-exposed, shaded by trees, and buildings) and urban forms was assessed with the “gold standard” six-directional setup observations made at 60 unique sites. The City of Phoenix intends to invest in street trees to reach 25% tree canopy coverage by 2030. In addition, the City has started a "Cool Corridor" program as part of its Climate Action Plan Framework, with a target of 100 cool corridors networked across its districts. To provide valuable data for the program, **Article III** produced high-resolution hourly T_{MRT} and shade maps for a hot, dry, calm summer day to inform the City and quantified shade coverage on sidewalks in Phoenix and Tempe. The article separated shades from trees and buildings to identify how they contributed to shading in the area.

2. MATERIALS AND METHODS

2.1 Enhancing remote sensing products to assess heatwaves in Estonia at local scales

The Estonian Weather Service (EWS) has two heatwave classes. A timeframe of three or more consecutive days with daily maximum temperatures ≥ 27 ºC or daily averages ≥ 20 ºC is classified as Level 1. A similar timeframe would be Level 2 if the daily maximum temperatures are ≥ 30 ºC or daily averages are ≥ 25 ºC.

The daily maximum SAT from 25 automatic weather stations collected over the summers (May-August) of 2010, 2014, and 2018 were analyzed to identify heatwaves in Estonia. This thesis defines a heatwave period as the time lapse from the first day of the first three consecutive days meeting the Level 1 criteria at any of the EWS stations to the last day of such. Table 1 shows the identified heatwave periods for the three years considered in this thesis.

The following subchapters present a gap-filling approach to secure seamless daily 1 km MODIS LST products to study heatwaves across Estonia **(Article I)**. A two-part downscaling approach that yields seamless daily NDVI and LST at 500 m and 100 m resolutions for Tallinn, Estonia's capital, is also presented **(Article II)**. The choice of target resolutions was informed by the scales at which Statistics Estonia (SE) maps population distribution.

Table 1. Identified heatwave periods in Estonia in 2010, 2014, and 2018. Source: **Article I**, Table 1.

Year	Period	Days
2010	05.07–17.08	44
2014	06.07–11.08	36
2018	12.07–10.08	30

2.1.1 Overview of the random forest algorithm

The RF algorithm proposed by Breiman (2001) operates with an ensemble of decision trees to overcome overfitting – when a model performs well on the training data but underperforms on unseen data (Bargagli Stoffi et al., 2022; Shalev-Shwartz & Ben-David, 2013). Each decision tree in the ensemble represents a model, and the final output of an RF model is an average (regression, Figure 1a) or a majority vote (classification, Figure 1b) (Moisen, 2008). Each tree grows by splitting the instances at the starting node into two based on one feature. After that, the instances in each node are further split based on one of the features in a similar fashion. This binary recursive portioning continues until an acceptable margin of error is attained at a terminal node (Moisen, 2008). Some measures for determining the margin of error are the mean square error for regression and

the Gini index for classification (Ethem, 2010; Moisen, 2008). Most ML packages allow users to define hyperparameters like the number of trees in the forest, the maximum depth to which trees can grow, and others that ensure a model does not overfit. As a rule of thumb, several combinations of hyperparameters should be experimented with to find the most appropriate. The algorithm's robustness has made it mainstream for remote sensing applications (Belgiu & Drăgu, 2016).

An essential part of any ML project is the interpretation of the model. The RF algorithm is interpretable using model-agnostic tools (Molnar, 2020). These post-training tools explain a model in general or/and make specific explanations (Molnar, 2020). Using these tools makes an RF model explainable.

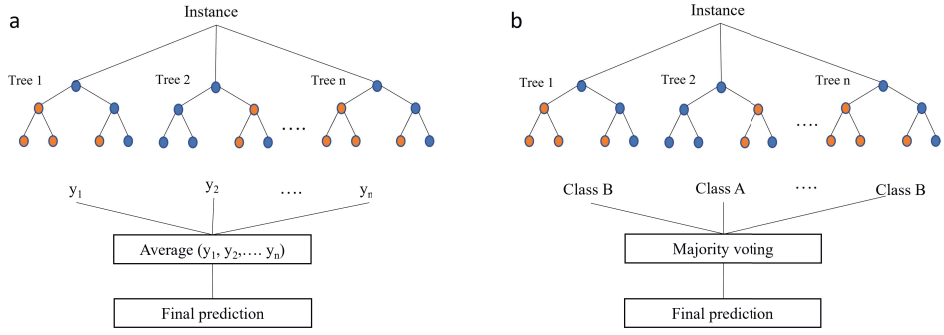

Figure 1. a) Prediction process in random forest regression; b) Prediction process in random forest classification.

2.1.2 Seamless LST products for countrywide evaluation of heatwaves

The MODIS instruments onboard the Terra and Aqua satellites are established LST sources for assessing urban heat at local scales (Chakraborty et al., 2020; Cheval et al., 2022; Roth, 2013). The MODIS instrument on each satellite provides a daytime and nighttime LST product daily at a 1 km resolution. While the temporal and spatial resolutions are appropriate for countrywide analysis, missing data due to cloud contamination is a limitation (Li et al., 2018; Liu et al., 2020). Gaps in MODIS LST data covering Estonia were filled using an RF model to obtain a set of seamless products appropriate for evaluating the variability of surface temperatures during the heatwave periods. Figure 2 shows the gap-filling approach.

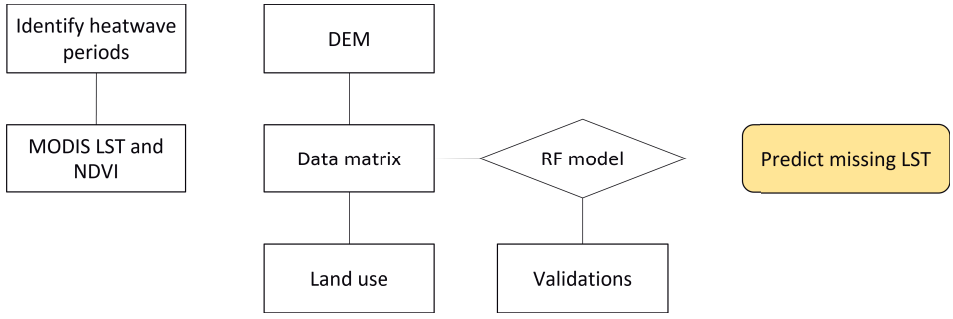

Figure 2. Gap-filling workflow.

Data

Vegetation abundance, LULC characteristics, and other spatial properties like elevation and distance to the sea or tidal bodies influence the LST of a place (Asgarian et al., 2015; Bernales et al., 2016; Guo et al., 2015; Osborne & Alvares-Sanches, 2019; Peng et al., 2020). Therefore, parameters representing these properties were employed as features for the ML model.

LST and NDVI from MODIS were downloaded using Google Earth Engine (GEE). LST composites for each day of the heatwave periods were generated in GEE using the daytime data from the Aqua (MYD11A1) and Terra (MOD11A1) satellites. The corresponding daily NDVI (MODIS/006/MCD43A4_006_NDVI) products are generated from MODIS surface reflectance composites (Asadi et al., 2019). The LULC data from the Estonian Topography Database (ETAK) was used to represent land cover information in the model. ETAK is highly detailed, with 55 LULC classes. The LULC was reclassified into 15 broad categories: agricultural facilities, built-up commercial use, built-up mixed use, open space-commercial use, open space mixed use, roads, croplands, extraction sites, forests, reeds and bushes, water, urban greenery, wetlands, grassland, and other. In addition, the objects in the database,except water were classified into surface types: building, impervious, and pervious.

Using the 1 km population grids from SE, the shares of reclassified land use and surface types in each grid cell were calculated. Using simple spatial averaging, the mean elevation for each grid cell was calculated by resampling a 5 m resolution Digital Elevation Model (DEM) from the Estonian Land Board. Finally, the distance from each grid cell's center to the Estonian coastline was calculated. In the final data matrix, each grid cell (instance/row) had 15 columns for land use fractions, three columns for surface type fractions, and a column each for elevation, distance to the coast, X centroid coordinates, Y centroid coordinates, LST, NDVI, DOY, and Year. Because the lengths of the heatwave periods are unequal, the final data matrix was sampled to ensure a balance in the data before modeling.

Modeling

The relationships between the features and LST were explored using a correlation matrix. A subset (n = 3104) of the data was set aside to validate the final model with a gap-filling algorithm. The remaining data matrix was split into two parts in the ratio of 60: 40. 60% of the data was used in training two linear regression baseline models using Ordinary Least Squares (OLS). The first baseline model used the 15 land use fractions and the rest of the features. The second baseline model used the three surface type fractions, the fraction of water, and the rest of the features. The land use and surface type fractions were not used together in the baseline models because the latter is a derivative of the former. The baseline models were tested using the remaining 40 % of the data using the RMSE metric. The RMSEs from both models were similar. The combination that included the surface type fractions was selected for the main model built with the RF algorithm because of its simplicity and *explainability*. The LST of a grid cell is expressed as a function of the predictors in eq. 1.

$$LST_i = f(NDVI_i, \propto LS_i, CC_i, DC_i, ME_{i,}DOY, Y) \quad (1)$$

LST_i is the LST of cell i, $NDVI_i$ is the NDVI of cell $i, \propto LS_i$ represents the fractions (shares in percentages) of land surface types and water in cell i, CC_i are the centroid coordinates of cell i, DC_i is the distance to the coast from cell i, $ME_{i,}$ is the mean elevation in cell i, DOY is the day-of-the-year, and Y is the year.

Before building the main model, k-fold grid search cross-validation was done to identify the best combination of hyperparameters. In this cross-validation, the training data is divided into k-folds, and the algorithm is fit to k-1 folds and tested on the remaining fold; the process is done k times, with each fold being used as the test data once (Anguita et al., 2012; Raschka, 2015; Siji George & Sumathi, 2020). The k value used in this task was 5. Finally, the RF algorithm was fitted to the training data using the best combination of hyperparameters.

Model evaluation and gap-filling

The performance of the final model was evaluated using the RMSE metric and the coefficients of correlation (R) and determination (R^2). Additionally, the model's response to variations in each feature was investigated using PDPs, a model-agnostic interpretation tool (Hastie et al., 2008). PDPs highlight the marginal influence of predictor values on the target and reveal linear relationships (Friedman, 2001; Molnar, 2020). The model's performance was compared with an R-based gap-filling algorithm that predicts missing data by applying quantile regression to the data using spatio-temporal features (Gerber et al., 2018). The gaps in the MODIS LST collected for the heatwave periods identified in Table 1 were filled with the final model after the evaluation.

2.1.3 Synthetic NDVI and LST for city-wide evaluation of heatwaves

While a 1 km-resolution LST is appropriate for assessing SUHI, there could be significant geometric and structural variability within a 1 km cell. LST at relatively higher resolutions is essential to evaluating cities' surface urban heat. This section aimed to generate LST products at 500 m and 100 m resolutions for Tallinn and its immediate environs (Figure 3) for the summer of 2018. In this thesis, Tallinn is considered the Central Urban District (CUD). The CUD and its suburbs form the Tallinn Urban Agglomeration (TUA). The SE maps the population in the CUD at 100 m and 500 m in the TUA. The set limits for the study site included settlements within the TUA and those within a 5 km buffer. The CUD is predominantly a large low-rise Local Climate Zone (LCZ), and the surrounding areas are dominated by dense trees and low plants LCZ (Demuzere et al., 2019)

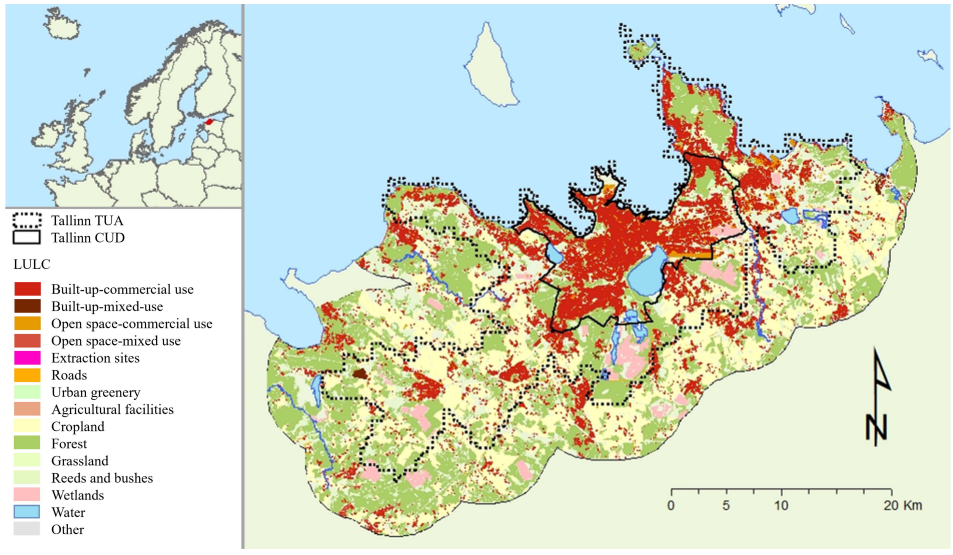

Figure 3. A map of the study area showing LULC distribution. The insert map shows the study area's location (red mark) relative to other parts of Europe. Source: **Article II**, Figure 1.

Currently, no satellite mission provides daily LST products at these resolutions over the area of interest. From **Section 2.1.2**, NDVI was significant in predicting LST. However, daily NDVI products at these resolutions are currently unavailable.

Hence, a two-part approach (Figure 4a) was proposed to generate the LST at the desired resolutions. The first part dealt with predicting NDVI at these resolutions. The second part used the predicted NDVI and other features to predict LST. The NDVI and LST modeling were executed using the schematic in Figure 4b. Each process began with using an RF to model the relationship between the target and the predictors at a coarse resolution of 1 km (Figure 4b). Next, the models were used to predict the target features (NDVI, LST) at the desired resolutions

using the predictors at the same resolutions. The premise was that relationships between the target and the predictors are similar at coarse and high resolutions (Bechtel et al., 2012). Table 2 summarizes the data and processes used in this section.

Table 2. Summary of datasets used, Source: **Article II**, Table 2

Dataset	Resolution	Processing
Landsat 8 NDVI	30 m	Resampled to 1000 m, 500 m, and 100 m
MODIS NDVI	1000 m	Target in MODIS-NDVI model and a predictor in the LST model
Landsat 8 LST	30 m	Resampled to 500 m and 100 m
MODIS LST	1000 m	Target feature in LST model
DEM	25 m	Resampled to 1000 m, 500 m, and 100 m
Surface-air temperature (SAT)	Hourly	Mean daily temperature for daytime MODIS overpass window (10:00 am – 2:00 pm)

* The local weather station in Tallinn is located at 59.3981 N and 24.6029 E

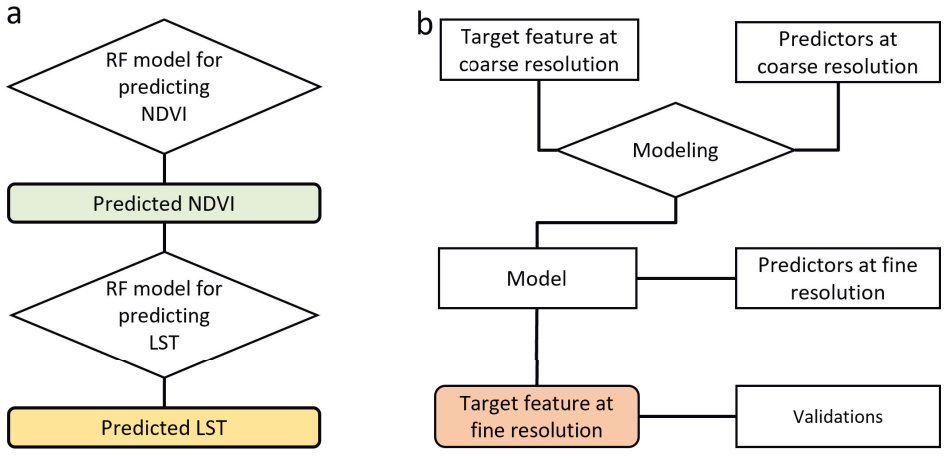

Figure 4. a) A schematic of the two-step approach to generating synthetic products; b) The general workflow adopted for modeling at each process step in (a). Source: **Article II**, Figure 2.

Generating synthetic NDVI

The land cover types in a pixel influence a pixel's NDVI value. Hence, NDVI has been employed in research to classify land cover (Defries & Townshend, 1994; Jeevalakshmi et al., 2016; Sasidhar et al., 2019). This thesis assumed that, given the proportions of different land cover types in a pixel, the NDVI could be determined. Two models to predict NDVI at the desired resolutions were created

using the RF algorithm. One model predicted a Landsat 8-like NDVI and the other predicted a MODIS-like NDVI. Figure 5 shows the detailed approach to predicting NDVI at the desired resolutions.

Landsat 8 images for the study area were downloaded to create the Landsat 8 NDVI-based model. Landsat 8 provides an image every 16 days. After discarding cloud-contaminated images, 19 images were used. Details for the images are provided in **Article II**, Table 1.

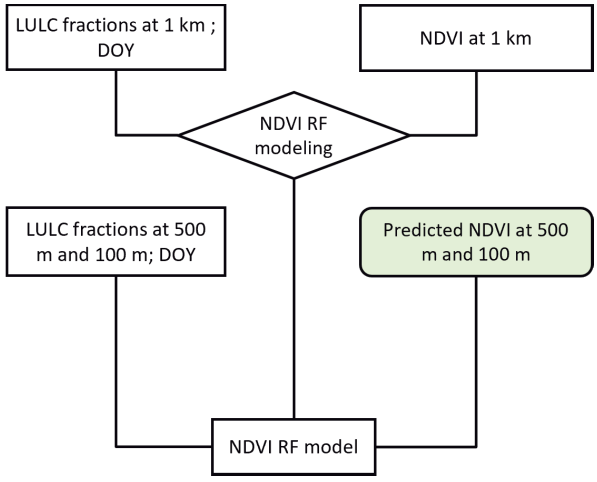

Figure 5. The detailed approach to generating synthetic NDVI at the desired resolutions (500 m and 100 m). Source: **Article II**, Figure 3a.

Images for two days (the 17th of May and the 2nd of June) were reserved for validating the models. The Landsat 8 NDVI was calculated using the formula in eq. 2.

$$NDVI = \frac{(\rho_{nir} - \rho_r)}{(\rho_{nir} + \rho_r)} \quad (2)$$

ρ_{nir} and ρ_r are the surface reflectance in the near-infrared band and red band, respectively.

Using the 1 km grid from SE for the study area, each grid cell's LULC type fractions were calculated using the reclassified LULC in **Section 2.12**. Two data matrices were created afterward. The first consists of 15 columns for LULC fractions, one for Landsat 8 NDVI, and one for the DOY on which the images were sensed. In the second matrix, each instance had 15 columns for LULC fractions, one column for MODIS NDVI, and the DOY for the MODIS NDVI product. The MODIS NDVI product for 2018 from **Section 2.12** was used here.

Landsat 8 has a native spatial resolution of 30 m; therefore, the NDVI was resampled to 1 km using spatial averaging. In addition, the Landsat 8 NDVIs for the validations were resampled to 500 m and 100 m using spatial averaging. LULC fractions were also calculated at 500 m and 100 m to predict the NDVI.

Each data matrix was split into 60% for training and 40% for testing. Three baseline models were run on each data matrix using OLS linear regression, ridge

regression, and RF with *Scikit-learn*'s default parameters. RF stood out as the best-performing algorithm when the baseline models were evaluated. Hence, an RF model was created with each data matrix after performing hyperparameter tuning, as presented in **Section 2.1.2**. The NDVI is expressed as a function of the predictors in eq. 3. The Landsat 8-like and MODIS-like daily NDVI at 500 m and 100 m resolutions were generated using their respective models. The mean predicted NDVI values from both models were also calculated.

$$NDVI_i = f(\propto LULC_i, DOY) \tag{3}$$

$NDVI_i$ is the NDVI of cell i, $\propto LULC_i$ represents the fractions (shares in percentages) of LULCs present in cell i, and DOY is the day-of-the-year.

The models were evaluated by calculating the R, R^2, and RMSE between the predictions and the Landsat 8 NDVI from the reserved images. In addition, grid cells with underestimated or overestimated predictions and their predominant LULC were examined to assess the performance of the models for different LULCs. Cells where the difference between the predicted NDVI and the validation was $\geq +0.3$ were marked as overestimated. Cells where the difference was ≤ -0.3 were marked as underestimated.

Generating synthetic LST

An RF model at a spatial scale of 1 km was built to predict the daily LST at 500 m and 100 m following the procedure outlined in **Section 2.1.2**. The differences between both models lie in the number of features and instances. Since this section focused on predicting daily LST for the summer of 2018, the "Year" feature was excluded from the data matrix. Figure 6 shows the modeling procedure and the approach employed to predict the LST at the desired resolutions.

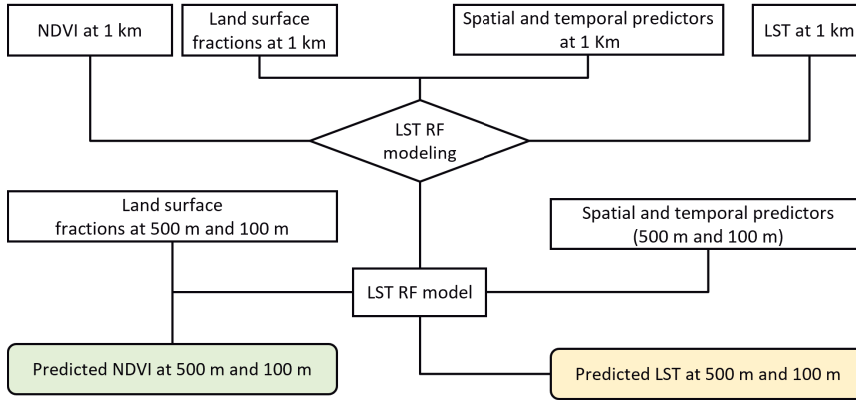

Figure 6. The workflow used in generating synthetic LST at the desired resolutions (500 m and 100 m). The NDVI used at the modeling stage is the MODIS NDVI product used in Figure 5. The predicted NDVIs at 500 m and 100 m are the mean predictions from Figure 5. Source: **Article II**, Figure 3b.

The surface type fractions and values for the other features for the 500 m and 100 m grids were calculated following the approach used in **Section 2.1.2**. The mean predicted NDVI at 500 m and 100 m from both NDVI models were used as input features to predict the LST at the respective resolutions.

LSTs for the Landsat 8 images reserved for validation were calculated by applying the single-channel algorithm proposed by Cristóbal et al. (2018) on the thermal bands (Band 10). Finally, the LSTs were resampled to 500 m and 100 m resolutions using spatial averaging. Further analyses were done to identify the predominant LULC in grid cells where the model overestimated or underestimated the LST. Grid cells with a difference $\geq +3$ °C were flagged as overestimated, and those with a difference ≤ -3 °C were underestimated. The thresholds were set following Malamiri et al. (2018). The trend of SAT measurements from a local weather station was compared with those of the predicted products.

2.2 Microscale thermal data and shade distribution for heat action planning

The human body's exposure to radiation fluxes and how complex urban forms and shade distribution affect thermal comfort are essential for heat mitigation planning. This section validates the SOLWEIG T_{MRT} estimates and generates T_{MRT} and shade distribution maps for the Phoenix-Tempe area. The section also evaluates hourly shade coverage on sidewalks in Phoenix and Tempe. Figure 7 presents the workflow used in this section.

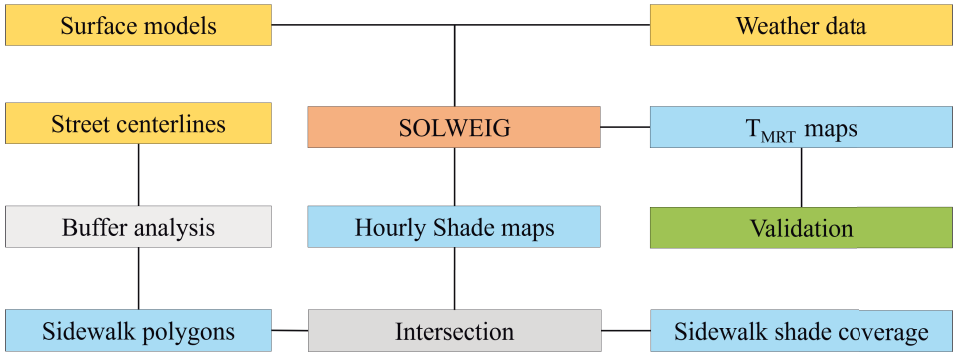

Figure 7. The workflow of the study including input data (yellow), GIS operations (grey), validation with MaRTy observations (green), and output data (blue). Source: **Article III**, Figure 2.

2.2.1 Study sites

The Phoenix metropolitan area is an amalgamation of 24 cities in Maricopa County, Arizona, USA (Figure 8). The area is located at 33°25'00.8'' N 111°58'41.5'' W. The cities of Phoenix and Tempe make up the urban core with an area of 1051 km^2 and 442 census tracts (population of 4,837,000), which accounts for about 29% of all census tracts in Arizona. The area is in the Sonoran Desert and has a subtropical desert climate (Köppen Climate Classification subtype Bwh). The maximum air temperature reaches or exceeds 38 °C between June and August. T_{MRT} can be as high as 76.2 °C in sun-exposed sites in the summer (Middel et al., 2021). Phoenix is mainly an open low-rise LCZ with an open high-rise downtown area. Tempe is primarily an open low-rise LCZ with an open mid to high-rise downtown area (Wang et al., 2018).

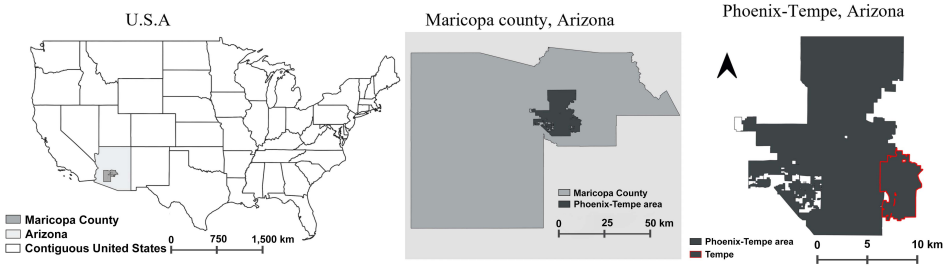

Figure 8. A map of the study area in the Southwestern US: Phoenix and Tempe in the Phoenix metropolitan area. Maricopa County, Arizona. Source: **Article III**, Figure 1.

2.2.2 Mean radiant temperature modeling with SOLWEIG

SOLWEIG requires spatial and meteorological inputs to simulate T_{MRT}. The spatial inputs include a surface model of buildings, a DEM, and a vegetation canopy height model. Air temperature, wind speed, relative humidity, and global solar radiation are the required meteorological inputs. Users can include direct and diffuse shortwave radiation if available. Otherwise, they can be estimated on-the-fly using global solar radiation.

The surface models were created using a high-resolution LiDAR point cloud from the United States Geological Survey. The points are pre-classified and collected in 1 km × 1 km overlapping tiles. A 1 m resolution surface model representing the elevation of the ground and all above-ground features, hereinafter referred to as the Full Digital Surface Model (FDSM), was created using triangulated irregular networks (Franklin et al., 2006). The FDSM for each tile was modeled after eliminating all points classified as noise. A DEM for each tile was created using ground points only. A Building Surface Model (BSM) was generated for each tile by extracting pixels in the FDSM masked by buildings using building footprint data from the Arizona State University library. The BSM was finished by filling the pixels outside the building footprint (no-data pixels) with corresponding data from the DEM. The Canopy Digital Surface Models (CDSM)

for all the tiles were created by subtracting the DEM from the FDSM for pixels outside the building footprints; pixels covered by building footprints were set to zero. Pixels with values less than or equal to 1.3 m were set to 0 in the CDSM because they were regarded as not high enough to provide shade for a standing person with an average height.

SOLWEIG estimates the T_{MRT} at a location by resolving the mean radiant flux (R_{str}). R_{str} is a function (eq. 5) of short and longwave radiation in six directions, angular factors, and a representation of human posture, as suggested by Höppe (1992). View factors control radiation exchange in urban environments (Lindberg et al., 2008). The model calculates the view factors using the surface models (BSM, CDSM, and DEM).

$$R_{str} = \zeta_k \sum_{i=1}^{6} K_i F_i + \varepsilon_p \sum_{i=1}^{6} L_i F_i \tag{5}$$

K_i represents shortwave radiation fluxes and L_i are the directional longwave radiation fluxes. F_i are the angular factors between a person and the surrounding surfaces. ζ_k is the absorption coefficient for shortwave radiation, ε_p is the emissivity of the human body. F_i was set as 0.22 in the cardinal directions and 0.06 in the upward and downward directions to represent a standing person. ζ_k and ε_p were set as 0.7 and 0.97, respectively (Lindberg et al., 2008; VDI, 1998). T_{MRT} is calculated in degree Celsius from R_{str} using the Stefan-Boltzmann law (eq. 6).

$$T_{MRT} = \sqrt[4]{(R_{str}/(\varepsilon_p \sigma))} + 273.15 \tag{6}$$

σ is the Stefan-Boltzmann constant (5.67×10^{-8} $Wm^{-2}K^{-4}$).

2.2.3 Validation of the SOLWEIG model

SOLWEIG T_{MRT} estimates were validated with T_{MRT} measured directly with MaRTy – a mobile human-biometeorological station that measures radiative fluxes in six directions (Middel & Krayenhoff, 2019). The measurements were made between 07:00 h and 21:00 h local standard time over nine hot summer days in 2016, 2018, and 2019 (Middel et al., 2021). A total of 763 measurements were made at 60 unique sites distributed across the Arizona State University Tempe campus and in Kiwanis Park, Tempe (Figure 9). The measurements were obtained under trees, in building canyons, and at open sites.

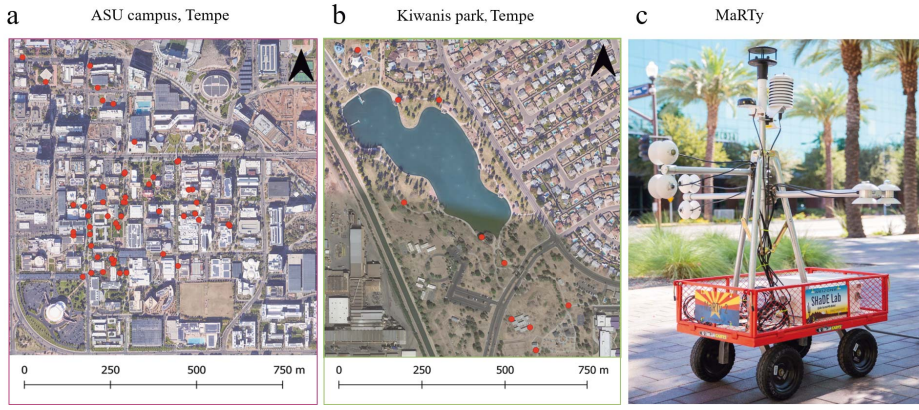

Figure 9. a) Model validation sites on Arizona State University's Tempe campus; b) model validation sites in Kiwanis Park; c) MaRTy, the biometeorological cart (Middel and Krayenhoff, 2019). Source **Article III**, Figure 3.

The T_{MRT} for the campus and park was calculated using SOLWEIG for the days and times of the validation data. Meteorological data from the Arizona Meteorological Network's Encanto Park weather station (AZMET, 2022) was used in forcing the model. The model's performance was evaluated based on exposure (sun-exposed or shaded; this classification is time-dependent for each site) and exposure type (open, shaded by trees, shaded by buildings). Linear relationships between the SOLWEIG-estimated and MaRTy-observed T_{MRT} were identified by calculating R and R^2. Willmott's index of agreement (d), a bounded and non-dimensional metric (Willmott, 1982), was computed for both sets of T_{MRT}. The evaluation metrics were compared with thermal stress standards (ISO7726) set by the International Organization for Standardization (ISO, 1998). After the validation, hourly 1 m resolution T_{MRT} maps for the 27th of June 2012 (07:00–20:00), a hot, dry, calm summer day, were generated for the Phoenix metropolitan area using the respective forcing data from AZMET.

2.2.4 Sidewalk shade assessment

SOLWEIG's shadow function was used to generate hourly shade distribution maps for the study site for June 21, 2019, from 7:00 h to 19:00 h. The function is based on the shadow-casting algorithm for DEMs originally developed by Ratti et al. (1999). Lindberg & Grimmond (2011) enhanced the algorithm when they considered the influence of vegetation. The algorithm is an iterative process in which a surface model is continuously shifted with a reduction in height at each step based on the Sun's altitude to get a shadow volume. The surface model is subtracted from the total shadow volume, and the result is reduced to a binary map where 0 indicates shade and 1 otherwise. The total shade is a function (eq. 7) of the transmissivity of foliated vegetation for shortwave radiation ($\boldsymbol{\tau}$) and shadows from buildings and tree canopy.

$$Sh_i = sh_{building} - \left(1 - sh_{canopy}\right) * (1 - \tau) \quad (7)$$

Sh_i is the total shade at location i, $sh_{building}$ is the shade cast by buildings at location i and sh_{canopy} is the shade cast by tree canopy at location i. $\boldsymbol{\tau}$ was set to 0.03 (3%). The total shade provided by buildings and trees was calculated for all time steps considered. Shade maps of buildings and vegetation were generated separately to evaluate the fractions of total shade provided by buildings and vegetation.

Using street width information provided in the city's guidelines for street design (City of Phoenix, 2009), sidewalk centerlines were identified by creating buffers around arterial and local street centerlines (Figure 7). Polygons representing the sidewalk were generated by creating buffers around the sidewalk centerlines using half of their corresponding design width as the buffer sizes.

The total shade coverage on each sidewalk and the contribution of shade from buildings and vegetation at a 1 m spatial scale were determined by intersecting the sidewalk polygons and the shade maps. The Maricopa Association of Governments' Active Transportation Plan advises that a 20-minute pedestrian route should have a minimum acceptable shade coverage of 20% (Maricopa Association of Governments, 2020). This threshold was used as the benchmark for evaluating the walkability of the area's sidewalks.

3. RESULTS

3.1 Enhancing remote sensing products to assess heatwaves in Estonia at local scales

The RF algorithm was used to fill gaps in MODIS LST products and was implemented in a two-part approach to create seamless synthetic daily NDVI and LST products. The findings, which include performance metrics, feature importance, and the partial dependence of the target variables on predictors, are presented in the following subchapters.

3.1.1 Seamless LST products for countrywide evaluation of heatwaves

Building and impervious surface fractions had a moderately positive correlation (**Article I**, Figure 3). The fractions of water and pervious surfaces had a moderate negative correlation. The same relationship was noticed between the Y centroid coordinates and the distance to the coast (**Article I**, Figure 3). The rest of the correlations were low.

The feature importance values reveal a feature's influence in predicting the LST. The importance (I) of a feature lies between 0 and 1. The sum of all feature importance values is 1. DOY was the model's most important feature in predicting LST (I = 0.51), followed by NDVI (I = 0.11). The year had an importance of 0.1. The rest of the features, including the surface type fractions, spatial features (elevation, centroid coordinates, and distance to the coast), and water, all had importance values of < 0.1 (Figure 10). The closeness of the RSMEs (training = 0.9 °C, test = 1.4 °C) for the training and test data suggests that the model did not overfit. There is a strong agreement between the predicted and actual data given the R (training = 0.96, test = 0.92) and R^2 (training = 0.93, test = 0.85) values. The residuals mainly lie within the +5 °C and –5 °C (**Article I**, Figure 5). In a few cases, the predictions were off by as much as 20 °C (**Article I**, Figure 5).

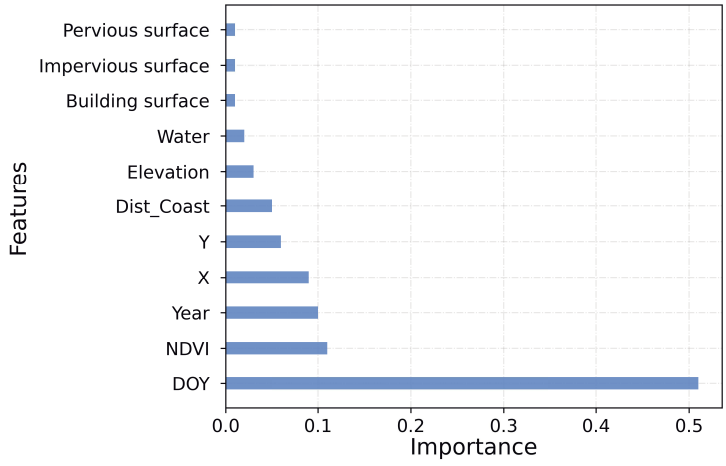

Figure 10. Feature Importance in the gap-filling model.

The feature importance values do not give any insight into what an estimate would be given a specific predictor value. The PDPs (Figure 11) show the model's average predicted LST, given a particular value of the predictors. Figure 11 suggests that when a cell's water fraction is < 20%, the model predicts an LST of about 26 °C. The predicted LST gradually reduces as the cell's water fraction increases. A cell entirely covered by water would have a predicted LST of about 25 °C. The predicted LST is the same for different values of all three surface type fractions. The NDVI PDP shows that if the NDVI of a cell is less than 0.5, the model predicts the same LST (> 29 °C). The predicted LST decreases as the NDVI of the cell increases from 0.5 and stabilizes at 25 °C as the NDVI approaches 1.0. The influence of distance to the coast is perceivable between 0 and a few kilometers; the predicted LST is relatively stable as the distance increases. A similar effect is noticed in the plot for elevation, where cells with a height of less than 30 meters have a slightly lower predicted LST, while those above 30 meters have almost the same predictions. The variations of the predicted LST along the East-West and North-South directions are captured in the X and Y coordinates' PDPs. There is more variation in the X-coordinate than in the Y-coordinate. The general temporal variation of LST across the three heatwave periods is represented in the DOY plot.

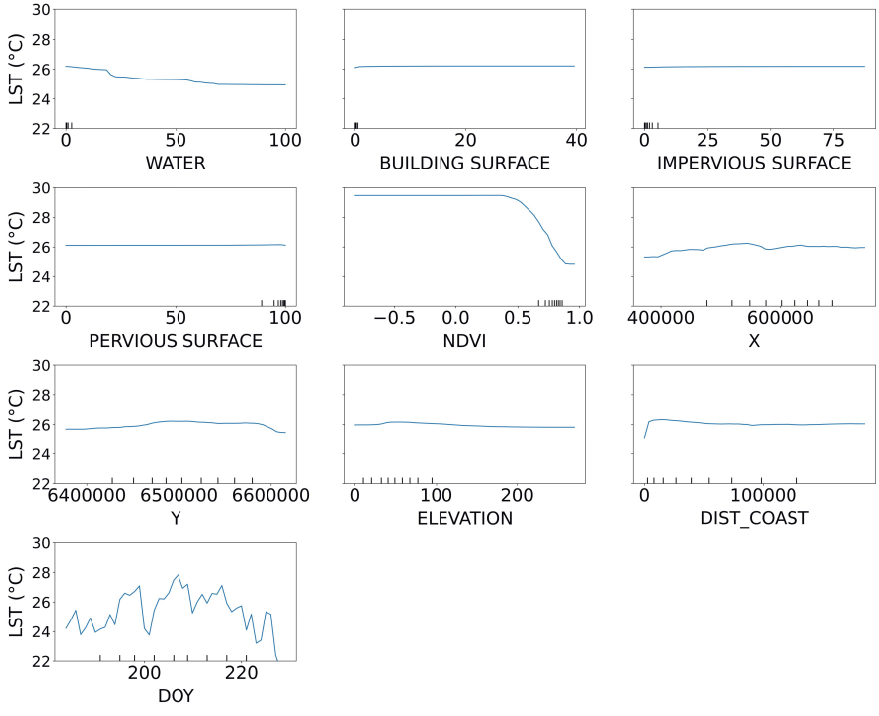

Figure 11. Partial dependence plots for the gap-filling model. Source: **Article I**, Figure 4.

The RF model (RMSE = 1.5 °C) outperformed *gapfill* – the R-based algorithm (RMSE = 2.6 °C) on the holdout data. The RF model's RMSEs on the holdout and test data are similar, indicating the model's ability to generalize well. Figure 12 shows some original images and their respective filled versions.

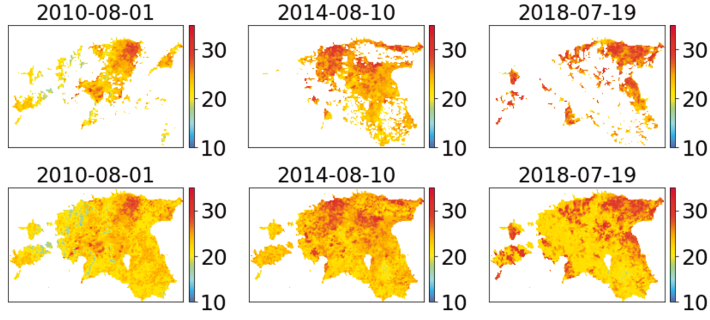

Figure 12. Original MODIS LST (upper row) and RF model filled (lower row), Source: **Article I**, Figure 6.

3.1.2 Synthetic NDVI and LST for city-wide evaluation of heatwaves

This subchapter presents the results of the two-part approach that generates synthetic NDVI and LST at 500 m and 100 m resolutions. The performances of the two RF models based on Landsat 8 and MODIS NDVI products are presented. The output of the RF model used to generate the LST at the desired resolutions is also presented here.

NDVI models

Before the baseline modeling, the MODIS and Landsat 8 NDVI were compared. Comparing the two products at a 1 km resolution revealed a strong positive correlation ($R = 0.88$, $R^2 = 0.78$, RMSE = 0.07). The evaluation of the baseline models for both products revealed that three features (built-up commercial, agricultural facilities, and grasslands) were of no importance (feature importance = 0). Hence, these features were not included in the final RF models.

Figure 13 shows the importance of each LULC type fraction for both models in estimating a cell's NDVI. The fraction of water in a cell was the most important feature (I = 0.41) in determining the NDVI for the Landsat 8 model, while for the MODIS model, it was the share of forest with an importance of 0.23. For the Landsat 8 model, the percentage of forest was the second most relevant feature (I = 0.19), while the proportion of water in a cell took the same position in the MODIS model (I = 0.16). The DOY had similar importance in both models (Landsat 8 model = 0.15, MODIS model = 0.19). The remaining features had importance values of < 0.1 and were almost the same for some features in both models. However, croplands as well as reeds and bushes had higher values in the MODIS NDVI model than in the Landsat 8 NDVI model. The feature importance values are given in **Article II,** Table 3.

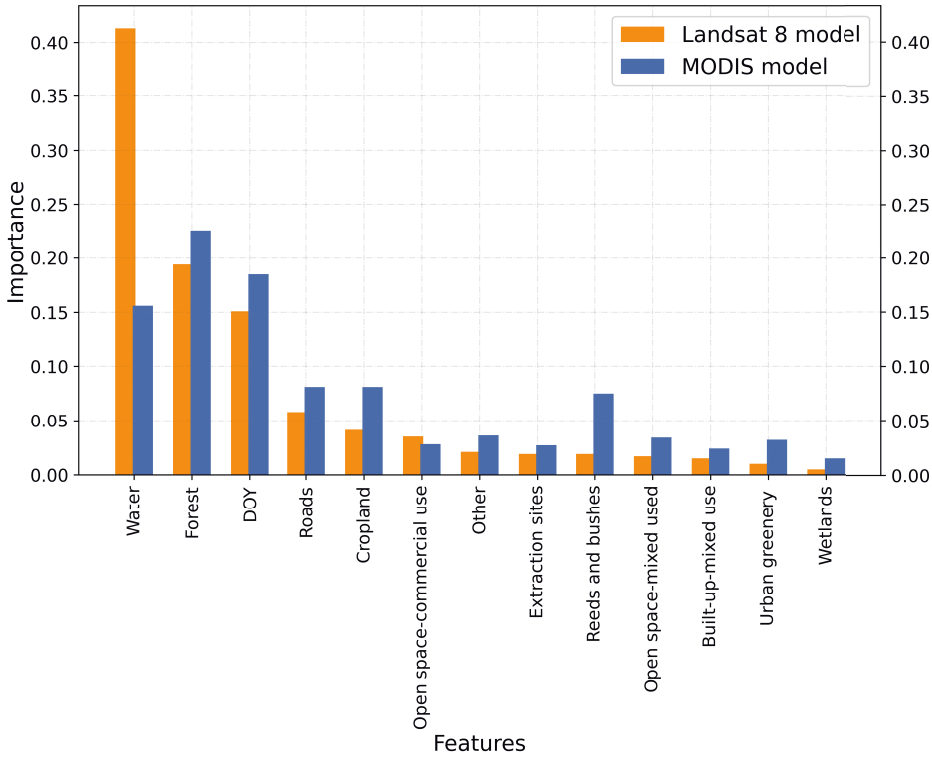

Figure 13. Feature importance for the models used in generating synthetic NDVI.

PDPs were employed here, as in **Section 3.1.1**. At first glance, the plots (Figure 14) show that the models "think" alike; both models predict lower NDVI values as the water fraction in a cell increases. The models estimate higher NDVI values with an increase in the share of reeds and bushes. The estimates from the models were mostly stable for different fractions of forests, wetlands, built-up mixed-used, and open space commercial. The striking difference in the PDPs is the range of predictions from both models. The MODIS model makes NDVI estimations from 0.4 to 0.8 (Figure 14b), while the Landsat 8 model's estimations range from –0.2 to 0.7 (Figure 14a). The ranges result from what the models learned from their respective training datasets.

Table 3 summarizes the performance of the models. Both models did not overfit, as indicated by the RMSEs from the training and test data. The levels of agreement between the predicted and measured NDVI for the test data are high for both models.

Table 3. Performance metrics for NDVI models. Source: **Article II,** Table 3.

	R		R^2		RMSE (°C)	
Model	Landsat 8	MODIS	Landsat 8	MODIS	Landsat 8	MODIS
Training	0.98	0.99	0.97	0.99	0.03	0.01
Test	0.94	0.99	0.89	0.99	0.05	0.02

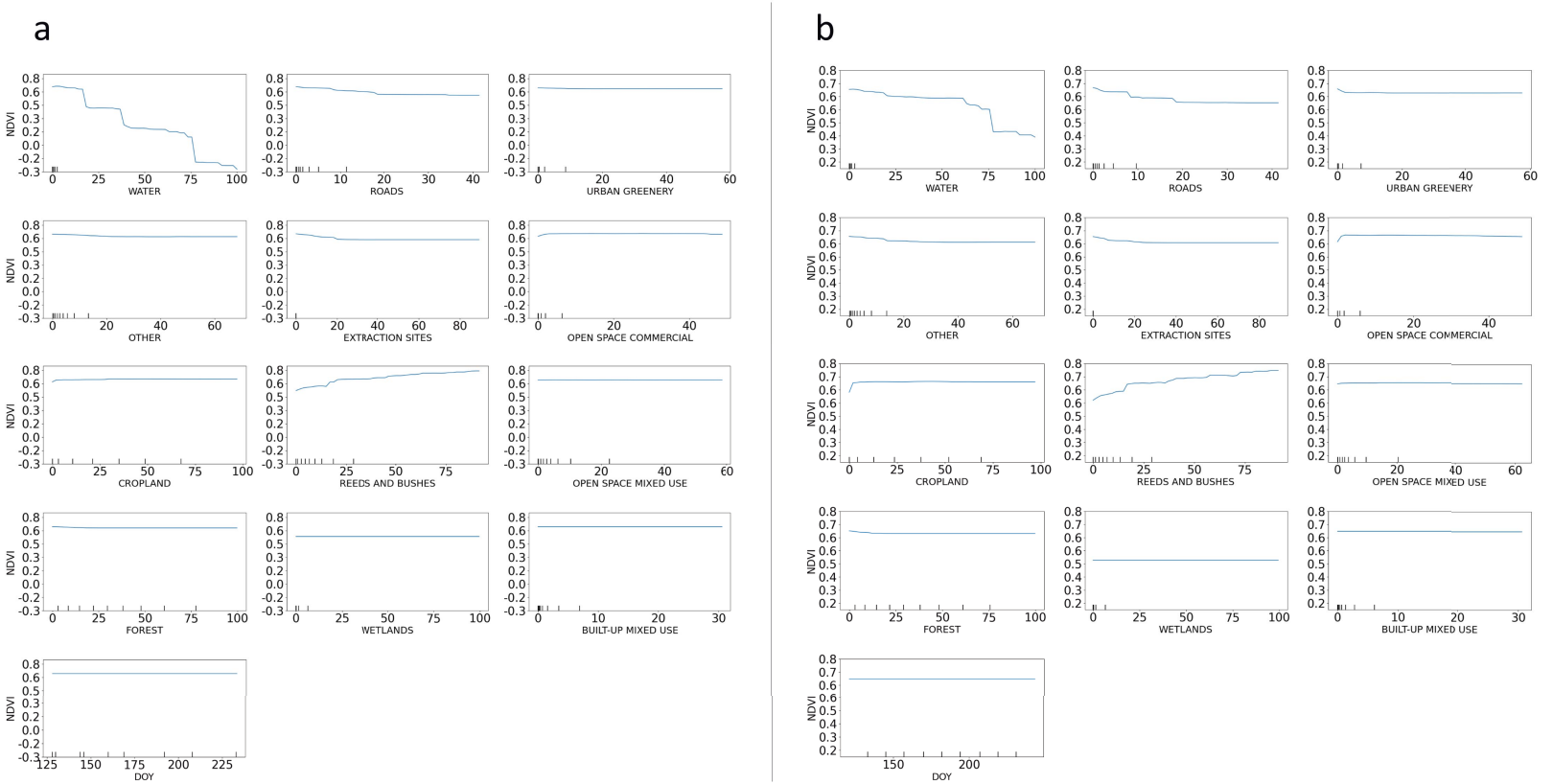

Figure 14. a) Partial dependence plots for the Landsat 8 NDVI model; b) Partial dependence plots for the MODIS NDVI model. The horizontal axis for the LULC classes represents the possible percentages per grid cell based on the data. Source: **Article II**, Figure 4.

LST model

In Figure 15, the DOY was the LST model's most relevant feature (I = 0.73), just as in **Section 3.1.1**. NDVI was the next with a feature importance of 0.06, followed by the centroid coordinates X (I = 0.06) and Y (I = 0.05). The fractions of the surface types were the least important features (I < 0.05). The feature importance values are given in **Article II,** Table 4.

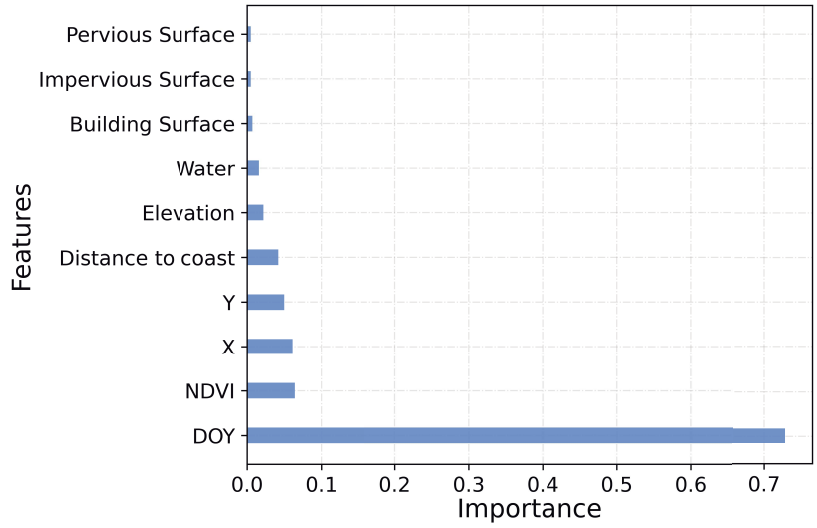

Figure 15. Feature importance for the model used in generating synthetic LST.

The PDP generated for this LST model (Figure 16) is identical to that in **Section 3.1.1**. The significant difference lies in the range of LST predictions. The model in **Section 3.1.1** had estimations ranging from 20 °C to 30 °C, whereas the estimates ranged from 15 to 30 °C for this model. The difference in the ranges can be attributed to the temporal extent. Instances (rows in training data) for the model in **Section 3.1.1** cover three heatwave periods (2010, 2014, and 2018), while those for this model cover the summer of 2018. A strong agreement was noticed between the original LST and the predicted LST in the training (R = 0.97, R^2 = 0.96) and test data (R = 0.95, R^2 = 0.91). The RMSE on the training data was 0.9 °C and 1.3 °C on the test data.

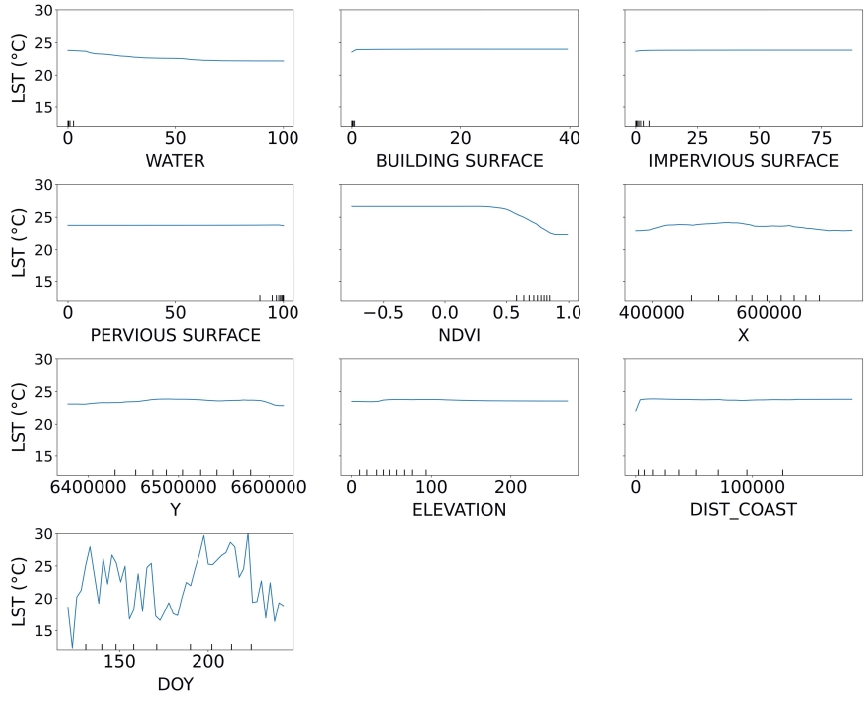

Figure 16. Partial dependence plots for the model used in generating synthetic LST. Source: **Article II**, Figure 5.

Synthetic NDVI products

The predictions from both MODIS and Landsat 8 NDVI models were similar, but a single NDVI product was required to determine the LST at the desired resolutions. The mean prediction for each grid cell was computed and used in the subsequent steps. Figure 17 shows the resampled Landsat 8 NDVI calculated for validation, the mean predicted NDVI, and the difference between the validation and the mean predictions. The differences were more significant at 100 m compared to 500 m (Figure 17). Table 4 shows that over 95% of the predictions from both models had differences (Δ) within the set range (for absolute figures, see **Article II,** Table 5). The mean product improves this statistic, suggesting that calculating the mean yields more estimations within the set range.

The predicted NDVI from both models showed high agreement with the validation NDVI, with over 70% of the variance explained at 500 m resolution (Table 4). At 100 m resolution, the proportions of variance explained (MODIS model = 63% and Landsat 8 model = 59 %) were relatively low despite the strong correlation between the validations and predictions. The correlation and the variance explained were enhanced in the mean product at both resolutions, even though the RMSEs did not significantly improve (Table 4).

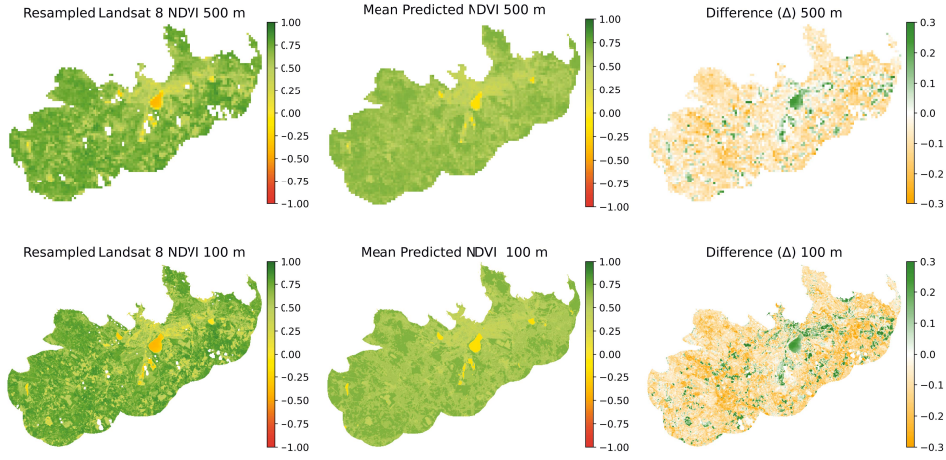

Figure 17. Resampled Landsat 8 NDVI, mean predicted NDVI, and difference maps for 17-05-2018. The difference maps are the results of subtracting the resampled Landsat NDVI from the mean predicted NDVI. Source: **Article II**, Figure 6a.

Table 4. Evaluation metrics for synthetic NDVI products. Sources: **Article II**, Tables 5 and 6.

	Estimations within range		R		R^2		RMSE	
	500 m	100 m	500 m	100 m	500 m	100 m	500 m	100 m
MODIS model products	98.4 %	97.0 %	0.85	0.79	0.73	0.63	0.09	0.13
Landsat 8 model products	99.1 %	96.4 %	0.84	0.77	0.72	0.59	0.11	0.15
Mean predictions	99.3 %	97.4 %	0.87	0.79	0.76	0.63	0.09	0.14

Synthetic LST products

The predicted LST and the LST estimated from the Landsat 8 images reserved for validation were compared. Figure 18 shows the resampled Landsat 8 LST at 500 m and 100 m resolutions, the predicted LST, and the difference between the resampled validation and the predictions. The predicted LST was mainly overestimated. However, remarkable similarities exist between the validations and the predictions, especially in the CUD area (Figure 18). The agreement metrics were high at both spatial scales (R; 500 m = 0.78 and R; 100 m = 0.72). The percentages of variance explained were moderate (500 m = 61% and 100 m = 52%). The observed RMSEs were 4.3 °C at 500 m and 100 m resolutions. The estimations within the set range were less than 30% at both resolutions (500 m = 26% and 100 m = 29%)

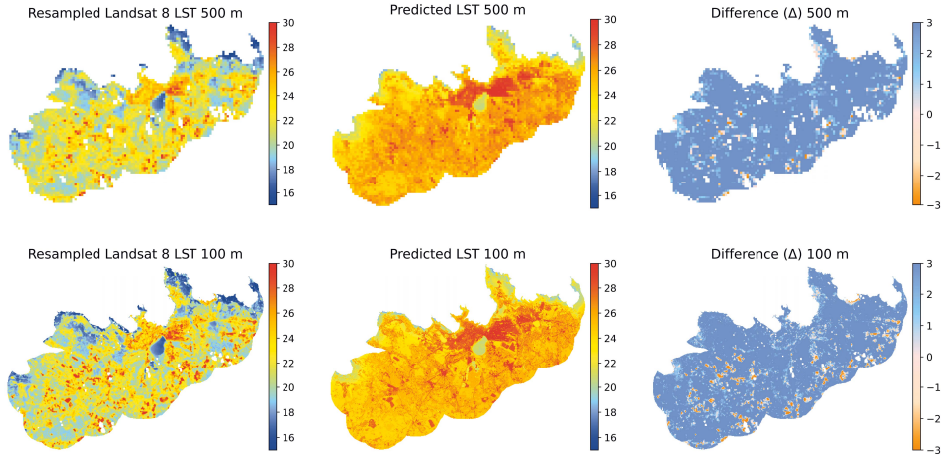

Figure 18. Resampled Landsat 8 LST, predicted LST, and difference maps for 17-05-2018. The difference maps are the results of subtracting the resampled Landsat LST from the predicted LST. Source: **Article II**, Figure 7a.

Temporal analysis of the predicted LST and SAT

The SAT measurements from the local weather station were compared to the predicted LST at both resolutions. The comparison demonstrated a significant temporal correlation between the field measurements and the predicted LST (Figure 19). Correlation coefficients of 0.68 and 0.70 were observed between the SAT measurements and the predicted LST at 500 m and 100 m, respectively. In both cases, a p-value < 0.005 was observed, indicating significance.

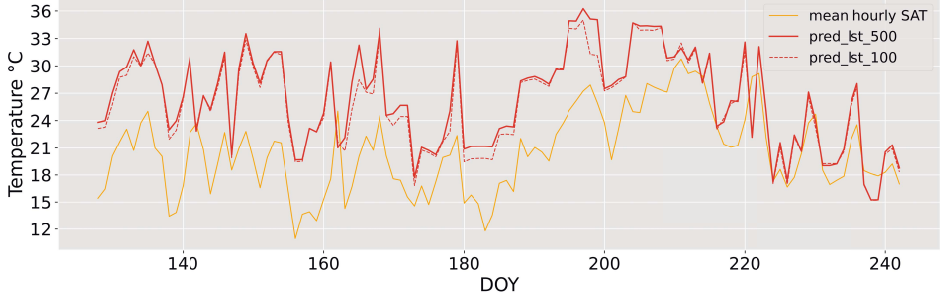

Figure 19. SAT versus predicted LST trends. Source: **Article II**, Figure 9.

Assessment of synthetic NDVI and LST based on LULC

The predicted products (LST and mean NDVI) were evaluated to identify the predominant LULC type in the cells in which the LST or NDVI were underestimated or overestimated. There were more misestimations at the 100 m resolution.

NDVI was mostly overestimated in cells predominantly covered by water, although there were relatively few water-dominated cells where NDVI was underestimated (Figure 20a). Overestimations were also noticed in cells primarily covered by croplands, extraction sites, open space commercial land use, and other LULC types.

Grid cells predominantly covered by forests, croplands, reeds, and bushes had their LSTs overestimated (Figure 20b). A relatively negligible number of cells covered by the open space commercial LULC type had their LST underestimated.

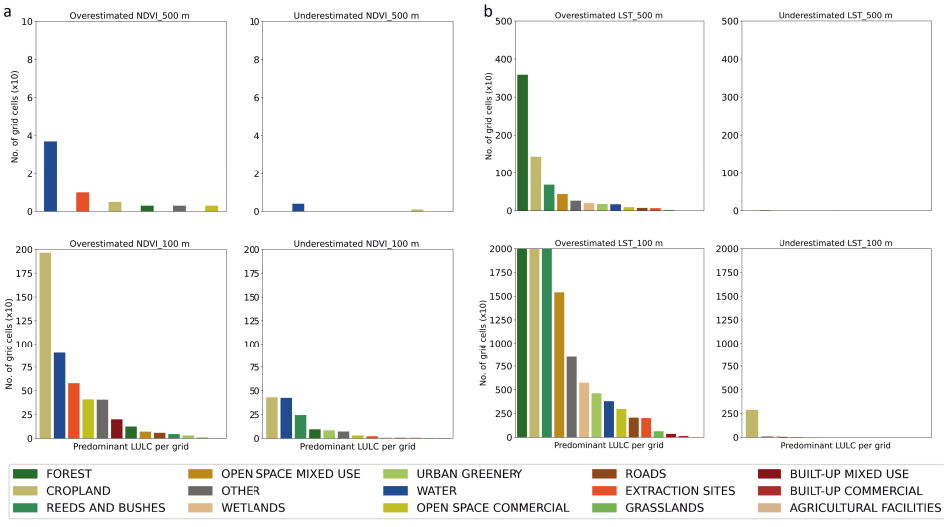

Figure 20. a) Predominant LULC in overestimated and underestimated grid cells for the mean NDVI products for all validation days; b) Predominant LULC in overestimated and underestimated grid cells for the LST products for all validation days. Overestimated NDVI ($\Delta \geq 0.3$); underestimated NDVI ($\Delta \leq -0.3$). Overestimated LST ($\Delta \geq 3$ °C); underestimated LST ($\Delta \leq -3$ °C). Source: **Article II**, Figure 8.

3.2 Microscale thermal data and shade distribution for heat action planning

This subchapter covers the findings from evaluating the T_{MRT} estimations from SOLWEIG under different conditions with field observations. The subchapter also presents sidewalk shade coverage analysis results for Phoenix and Tempe.

3.2.1 Validation of the SOLWEIG model

A strong linear relationship (R = 0.91, d = 0.95) was observed between the SOLWEIG T_{MRT} estimates and the MaRTy-observed $T_{MRT,}$ with 83% of the variance in the MaRTy-observed data explained. The overall RMSE and MAE were 5.59 °C and 4.60 °C, respectively (Table 5). Overall, the model overestimates, as can be gathered from the MBE value of 1.59 °C (Table 5).

High positive correlations were observed between SOLWEIG estimates and MaRTy-observed data for different exposure conditions, regardless of exposure type or urban form. The R-value for shaded locations was 0.75 and 0.70 at sun-exposed sites, respectively. Despite the high correlation values, the percentages of variance explained were not significantly high (shaded sites = 56% and sun-exposed sites= 49%). The RMSEs at shaded and sun-exposed sites were 5.12 °C and 6.60 °C, respectively, and the MAEs were 4.47 °C and 4.93 °C. These figures indicate that the model functions better in shaded conditions than in sun-exposed conditions. From Table 5, the model is expected to overestimate T_{MRT} in shaded areas (MBE = 3.21 °C) and underestimate it in sun-exposed areas (MBE = –2.28 °C).

The model's performance under trees was the lowest among the three categories of urban form and exposure types. A positive correlation (R = 0.75) was observed between the estimated and the observed data, with 56 % of the variance explained (Table 5). The RMSE under trees was 5.44 °C. As the MBE of 4.16 °C (Table 5) demonstrates, the model overestimates T_{MRT} under trees. Under trees, most of the errors were in the –4 °C to 11 °C range, with half in the –2 °C to –7 °C range (Figure 21). Outliers for trees were as high as 10 °C (Figure 21). The estimates at open sites were strongly correlated with the MaRTy-observed data (R = 0.92 and 85% of the variance explained) despite the high RMSE of 6.23 °C. The model underestimates T_{MRT} at these locations, as the MBE of –1.78 °C suggests (Table 5). The error range for open sites was the highest (–15 °C to 10 °C), with half of the sites having an error between –5 °C and 3 °C (Figure 21). The best performance of the model was noticed in building canyons (RMSE = 4.38 °C) with the strongest linear relationship between SOLWEIG and MaRTy-observed T_{MRT} (R = 0.94, and R^2 = 0.88). The model underestimates T_{MRT} (MBE = –0.24 °C) in building canyons, with errors ranging from –9 °C to 10 °C with half in the – 3 °C to 2 °C bracket **(**Figure 21**)**.

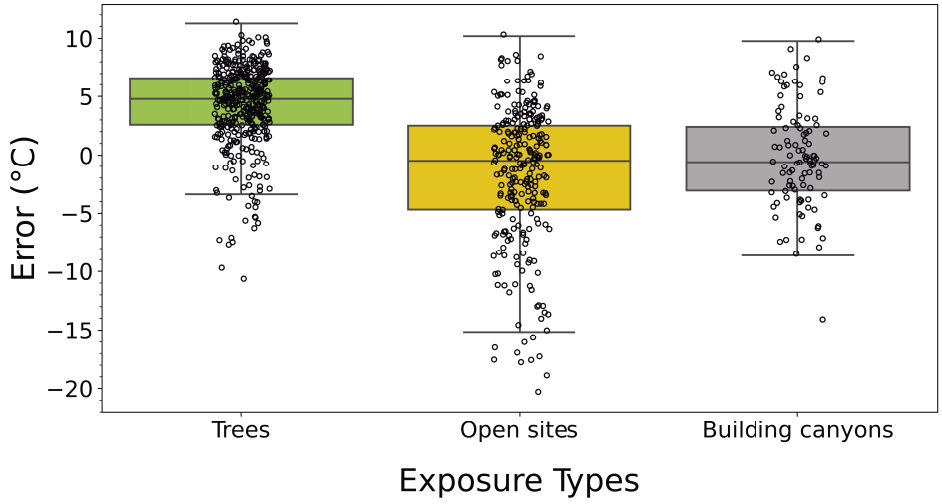

Figure 21. Distribution of the modeling errors for different exposure types. Source: **Article III**, Figure 4.

Table 5. Summary of SOLWEIG's performance under evaluated conditions: Overall, under shade, at sun-exposed locations, under trees, at open sites, and in building canyons. Source: **Article III**, Table 1.

Metric	Overall	Shade	Sun-exposed	Trees	Open sites	Building canyons
R	0.91	0.75	0.70	0.75	0.92	0.94
R^2	0.83	0.56	0.49	0.56	0.85	0.88
d	0.95	0.9	0.97	0.86	0.97	0.97
RMSE (°C)	5.59	5.12	6.60	5.44	6.23	4.38
RMSEz (°C)	2.86	3.3	2.72	4.3	1.84	1.71
RMSEu (°C)	4.81	3.91	6.01	3.33	5.95	4.03
MAE (°C)	4.60	4.47	4.93	4.90	4.61	3.45
MBE (°C)	1.59	3.21	–2.28	4.16	–1.78	–0.24

3.2.2 Hourly mean radiant temperature and shade maps

Hourly T_{MRT} and shade maps were generated after the validation using the respective forcing data. The highest T_{MRT} was 83 °C at sun-exposed sites and 41 °C at shaded locations (**Article III,** Figure 5a). An illustration of an hourly shade distribution map (combined shade from trees and buildings) used to determine the amount of sidewalk shade is shown in Figure 22. Downtown areas had more shade coverage. Due to the low sun elevation angle, shade coverage is most significant immediately following sunrise and just before sunset.

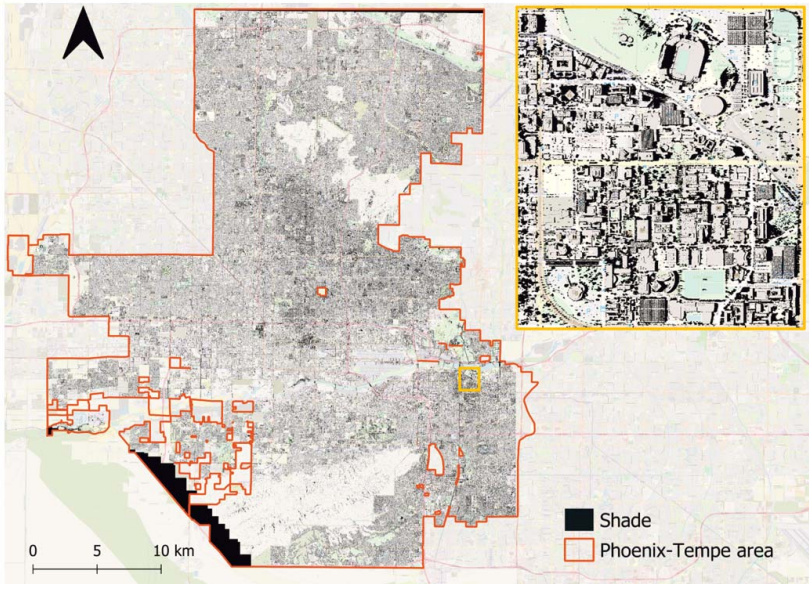

Figure 22. Shade distribution map of the Phoenix-Tempe area at 16:00 h local time on June 21, 2019, base map source: Open Street Maps. The dark patches in the southwest and north result from no data. Source: **Article III**, Figure 5b.

3.2.3 Sidewalk shade assessment

The hourly shade maps were intersected with sidewalk polygons to determine the amount of shade available on the sidewalks in the area. Sidewalk shade coverage at crucial times of the day: 07:00 h and 18:00 h (people commute to and from work or school), 16:00 h (maximum air temperature), and 12:00 h (Sun reaches the highest altitude) were examined further.

In the morning, North-South-oriented sidewalks were more shaded than those with other orientations (Figure 23a). At midday, the shade coverage was remarkably low for all sidewalks, with the most coverage at less than 20% (Figure 23b). At peak solar altitude, trees with large enough crowns accounted for the shade on sidewalks, which met the coverage requirements. Because of low solar elevations, the shade coverage increased at the hottest time of the day due to the long shadows cast by buildings and trees. In the evening, shade coverage for most sidewalks increased dramatically, even at locations with minimal vegetation. Buildings primarily provided this shade.

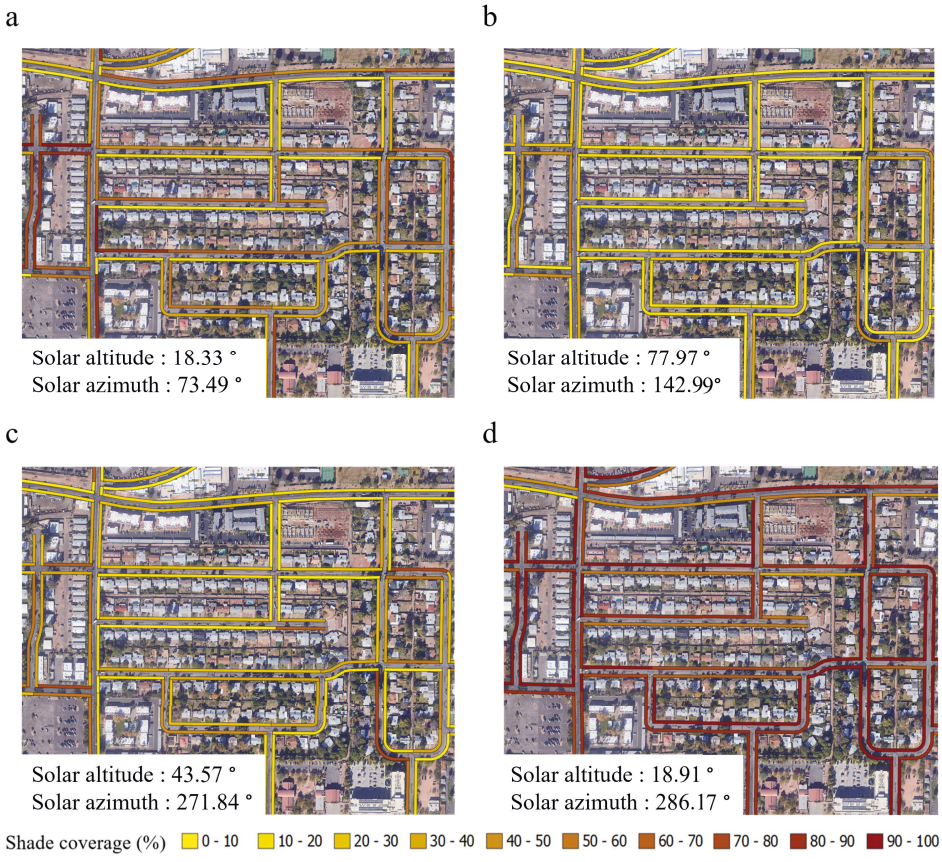

Figure 23. Total shade coverage for selected sidewalk polygons in Tempe on June 21, 2019, a coverage at 07:00 h; b 12:00 h; c 16:00 h; d 18:00 h. Sidewalk widths are enlarged in the figure for clarity Source: **Article III**, Figure 6.

Only 8% of all sidewalks in the study site achieved the 20% shade requirement throughout the day. Just before sunset at 19:00 h, most sidewalks exceed the 20% shadow coverage threshold. Less than half of the sidewalks fulfill the minimum criteria between 08:00 h and 17:00 h. The overall sidewalk shade coverage and shade contributions from trees and buildings per census tract were estimated for the times of interest listed above to enable a neighborhood-level shade availability analysis (Figure 24). Tracts with adequate shade coverage are mainly in central Phoenix and Tempe. Although all tracts achieved the minimum shade target for at least one hour in the day, only two tracts (< 1% of the total) met the requirement at all times. Tracts with a median total shade coverage of 20% or higher numbered 84, making up 23% of all tracts. Most tracts had their maximum shade coverage at 19:00 h and their minimum at either 12:00 h or 13:00 h. Vegetation shade accounts for over 50% of total shade in most tracts, especially at midday. The proportion of building shade in the study area is more significant in the mornings and evenings when the Sun's elevation angle is low.

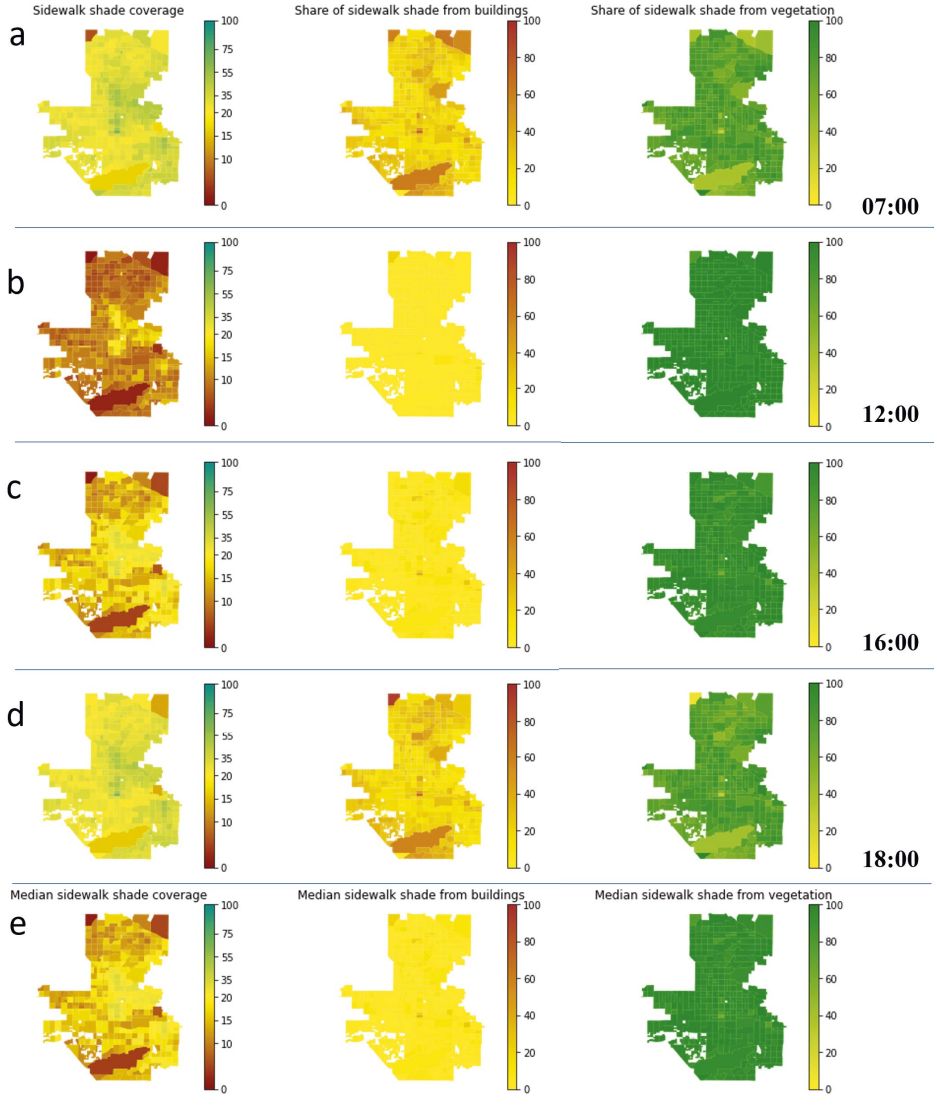

Figure 24. Total sidewalk shade and shares contributed from buildings and vegetation per census tract in the Phoenix-Tempe area on June 21, 2019, a) coverage at 07:00 h; b) 12:00 h; c) 16:00 h; d) 18:00 h; e) median coverage per tract for the entire day. Note that the outlier census tracts (top left and right, elongated tract to the south) are mountain preserves with few residential areas. Source: **Article III**, Figure 8.

4. DISCUSSION

4.1 Enhancing remote sensing products to assess heatwaves in Estonia at local scales

This subchapter discusses the approaches used to fill gaps in the MODIS LST and generate synthetic LST and NDVI for local scale assessment of heatwaves in this thesis. The subchapter also covers the suitability of the random forest algorithm, the output products, and their limitations.

4.1.1 Gap-filling satellite LST

Gaps resulting from cloud contamination (Crosson et al., 2012; Li et al., 2018) hinder the utilization of satellite LST to evaluate the spatial variability of surface temperatures at local climate scales over large spatial extents (regional, country-wide, and continental).

This thesis filled gaps in MODIS LST using an approach that combined compositing and empirical relationships. The compositing focused on combining MODIS Terra and Aqua daytime products. The assessment of the composites revealed that, while the method is relatively simple, it does not guarantee seamless products. In some composites, the area with missing data was less than 10% of the total area; for some, it was about 90% (**Article I,** Figure 2). This effect is also seen in Crosson et al. (2012), where the compositing method did not yield seamless products despite the appreciation of data coverage by about 24 to 26%. Therefore, compositing alone may not be appropriate for some seasons and regions, like Estonia, where cloud contamination is common.

Spatial and temporal interpolation have been identified to yield seamless products. For example, Zhang et al. (2022) reported an RMSE of 1.88 °C using their spatial and temporal interpolation approach. In this thesis, the temporal features were significant in making LST predictions (Figure 10). Hence, it could be hypothesized that the temporal features alone could yield predictions with acceptable accuracies. While Zhang et al. (2022) had an adequate error margin, the LST of a place is not only influenced by spatial and temporal properties (Bernales et al., 2016; Feng et al., 2019).

This study filled the remaining gaps based on empirical relationships between LST and features like NDVI, spatial and temporal features, and surface type fractions after compositing. These features have been identified to influence LST (Feng et al., 2019; Osborne & Alvares-Sanches, 2019). The RMSE on the holdout data from the model was better by 1 °C compared to that of the *gapfill* algorithm, which uses a spatio-temporal approach. The empirical method Yao et al. (2021) used also performed better than *gapfill* and other spatial and temporal algorithms for filling gaps. The results of this research indicate the advantage of an empirical approach over a spatio-temporal method.

4.1.2 Generating synthetic NDVI and LST products

Studies focusing on the dynamics of terrestrial ecosystems have relied on synthetic remote sensing products generated at various spatial resolutions to compensate for data unavailability due to low temporal and spatial resolutions (Atkinson, 2013; Gu & Wylie, 2015; Nomura & Oki, 2021). In this thesis, two synthetic remote sensing products (NDVI and LST) widely used for evaluating and understanding the urban thermal environment were generated using RF.

Synthetic NDVI

The comparison of NDVI from both sensors revealed a strong correlation between both products, which is also reported by Albarakat & Lakshmi (2019) and Ke et al. (2015). Discrepancies in NDVI derived from both sensors may result from differences in sensor properties and the processing algorithms utilized in atmospheric corrections for the respective bands (Franke et al., 2006; Gastellu-Etchegorry et al., 2012; Ju et al., 2012; Y. Ke et al., 2015; Steven et al., 2003; Trishchenko et al., 2002). Although the MODIS-like and Landsat 8-like NDVI products had strong relationships with the validation data, the mean product was preferred in predicting the LST at the target resolutions. Some biases (Figure 14) inherent in the models were assumed to be reduced after calculating the mean predictions.

The MODIS model overestimated the NDVI in grid cells covered mainly by water. The cause for this overestimation is visible in Figure 14, where the MODIS model does not make estimates below +0.4, although water bodies are known to have lower NDVI values. The Landsat 8 model also misestimated water-dominated cells. The NDVI of water bodies can be impacted by in situ flora and fauna. As a result, it is possible that in cases where a cell is entirely covered by water, the information on the proportion of water in a grid cell won't be sufficient to estimate its NDVI. The inaccurate predictions for croplands could be caused by variations in crop types and growing seasons – information not captured in the training data. In the mean predictions, there was a significant reduction in the count of misestimations for specific LULC categories, indicating the advantage of the mean product. The prevalent LULCs in the overestimated and underestimated grid cells at 500 m were the same as those at 100 m resolution (Figure 20a).

Despite being good, the RMSEs for the synthetic NDVI products did not match those obtained by Filgueiras et al. (2020), which ranged between 0.035 and 0.06. The differences in the predictors could be a reason for the difference in RMSEs. This study employed LULC fractions as predictors. Filgueiras et al. (2020) used surface reflectance for different bands as predictors, which have also been used in other works (Gao et al., 2017). To obtain NDVI for each day, as was intended in this study, predictors with the same temporal and spatial resolutions must be used. As a result, using surface reflectance may be difficult, especially when satellite missions are the primary sources. As Bonafoni et al. (2016) and Hassan-Esfahani et al. (2017) illustrate, this difficulty may be overcome by using

drone imagery. The downside of using drone imagery to overcome this challenge is the significant planning and fieldwork the data collection requires. Radiative modeling to obtain sensor-like spectral reflectance is an alternative (Gastellu-Etchegorry et al., 2012).

Synthetic LST

A strong correlation was noticed when the predicted LST was compared with the LST used for validation. The difference in sensing times is a potential reason for the disparities in the LSTs. The model was created using MODIS LST products, which are daily composites from two observations made at separate times that differ from those of the validations. Furthermore, LST varies diurnally (Hu et al., 2020; Sharifnezhadazizi et al., 2019; Weng & Fu, 2014); thus, it was not anticipated that the predicted LST would be close to the validation. Despite this, some researchers produced better RMSEs using data from different sensors and features (Hutengs & Vohland, 2016; Xu et al., 2020). The relatively high RMSE for this thesis could be attributed to the significant presence of forest and cropland patches in the study area. On these patches, the model failed to perform well.

The evaluation of the synthetic products with respect to the LULC is critical to appreciating their applicability (Pan et al., 2018). Bartkowiak et al. (2019) noticed varying error sizes when they estimated the LST for different vegetation types. The model in this thesis misestimates LST for vegetated surfaces. LSTs for predominantly forested grid cells were the most difficult to estimate, as the model overestimated at both resolutions. These misestimations may result from not distinguishing forest and crop types in the modeling phase. In this thesis, NDVI was the only feature to serve this purpose because of its suitability (Fan et al., 2021); however, the results indicate that it was not good enough. Therefore, additional predictors like canopy and crop height, which are obtainable with terrestrial laser scanning and valuable in distinguishing crops and forest types, could be considered to improve the predictions (Farhadur Rahman et al., 2022; Tilly et al., 2013).

Moreover, ETAK lacks information on clear-cut patches, which may affect the model's performance. These areas remain vegetated in the database, even though they are not. In such cases, the LST will not match the data supplied to the model.

Despite these constraints, the model works well for largely built-up grid cells (RMSE_500 m = 2.8 °C and RMSE_100 m = 3.0 °C). The model's performance for built-up cells is critical because the primary goal was to provide LST suitable for analyzing overheating in the CUD and TUA, where most people live. The synthetic products are acceptable for assessing intra-urban SUHI and the spatial variability of heatwaves inside the CUD. The high correlations between the surface-air temperature from the local weather station and predicted LST trends indicate the synthetic products' usefulness in evaluating the temporal variations of LST in the CUD.

4.1.3 Evaluation of the models and resulting products

The results of this thesis highlight the well-documented applicability of the RF algorithm to reconstruct or create synthetic remote sensing products (Holloway-Brown et al., 2020; Jing et al., 2016; Li et al., 2019; Xu et al., 2020). Furthermore, with RF, parameters influencing LST besides spatial and temporal information could be incorporated into a model to fill gaps and overcome the limitations associated with geostatistical and temporal interpolation methods (Sarafanov et al., 2020; Siabi et al., 2022).

Zhao & Duan (2020) demonstrated that building an RF model with spectral indices like NDVI, Enhanced Vegetation Index, Normalized Difference Water Index, surface albedo, and topographic information yields reconstructed LST with acceptable accuracies. While this research used the NDVI and other features to fill the gaps, the results are comparable to those Zhao & Duan (2020) achieved. The advantage of the gap-filling approach used in this research is that it does not rely on multiple spectral indices, reducing pre-modeling computations. Furthermore, the gap-filled products generated in this thesis will allow for assessing the spatio-temporal variation of LST during the identified heatwave periods.

In validating the synthetic LST and NDVI, the validation metrics were like those obtained when NDVI and LST from both missions were compared at coarse resolutions. Therefore, the RF algorithm can retain relationships well (Li et al., 2019). However, more studies focusing on different scenarios would be needed to confirm this based on the suggestions of Dong et al. (2020). While most researchers have validated the suitability of the algorithm to reconstruct or downscale remote sensing products in various contexts, their explanation of the models has been mostly limited to feature importance (Bartkowiak et al., 2019; Hutengs & Vohland, 2016; Li et al., 2019; Zhao & Duan, 2020). This research explored the *explainability* of RF models with model-agnostic tools like the PDP. The plots give further insight into the range of predictions a model will likely make and the relationship between each predictor and the target variables.

4.1.4 Limitations and prospects

Although this research's gap-filled and synthetic products are considered good and fit for purpose, they are not free from limitations. During baseline modeling, the surface type fractions model had a similar RMSE to that of the LULC types. Hence, the surface type model was preferred because it was a "simpler" model (i.e., a model with fewer features). While this ML heuristic based on Occam's Razor (Bargagli Stoffi et al., 2022) helped explain the model, it cost the model's accuracy when predicting LST for vegetated surfaces. This limitation was evident when the errors for the synthetic products were evaluated based on the predominant LULC types. The accuracy of the models in predicting LST for vegetated surfaces is relatively low compared to built-up surfaces. Reclassifying all vegetation classes as pervious surface types, a broad category that includes non-vegetated pervious surfaces, may have contributed to this limitation. Future works

should consider splitting this category into two categories, perhaps vegetated and non-vegetated pervious surfaces.

The LST prediction approach relied on a single spectral index. Additional spectral indices, such as the normalized difference sand index (Pan et al., 2018) and the normalized difference built-up index (Xu et al., 2020), are also helpful. However, as Bonafoni et al. (2016) demonstrated, some surface materials, such as roofs, are not adequately represented by spectral indices; hence, the intended use of the final products must be considered beforehand. In addition, other parameters available at daily temporal resolutions that were not considered here, like solar radiation and wind speed, could be evaluated as potential features to improve the LST predictions. For example, the sky-view factor is an excellent metric that can be used to quantify solar radiation and simulate LST (Hutengs & Vohland, 2016; Scarano & Mancini, 2017).

Considering the study site, the Estonian LULC database was used for this research. However, LULC data from Open Street Map or LCZ data from the World Urban Database and Access Portal Tools (WUDAPT) would be more appropriate to scale this approach to cover areas from different geographic regions. The LCZ from WUDAPT is based on a classification system that considers the thermal and physical characteristics of surface features (Bechtel et al., 2019; Demuzere et al., 2019; Dutta et al., 2021; Stewart & Oke, 2012).

4.2 Microscale thermal data and shade distribution for heat action planning

This subchapter discusses the performance of the SOLWEIG model under the three evaluated urban forms and exposure types. The evaluated metrics are compared with the accuracy standards set for estimating T_{MRT}. In addition, the quality of shading in the Phoenix-Tempe area is discussed.

4.2.1 Validation of the SOLWEIG model

The ISO recommends that T_{MRT} estimates should be within ± 5 ºC of observations in the ISO7726 standard (ISO, 1998). Considering this yardstick, the model's overall RMSE of 5.59 °C and MAE of 4.60 °C are satisfactory. Although the RMSE is 0.59 °C higher than the standard, the model's estimations are reasonable because outliers could have influenced this value. Earlier research yielded comparable results; for example, Gál & Kántor (2020), RMSE = 5.02 °C, and Lindberg et al. (2008), RMSE = 4.80 °C. Furthermore, the high index of agreement (d) indicates that the predictions match the observed data (Table 5). Except for trees, systematic errors (all < 5 °C) were less than unsystematic errors, indicating that the model has a minimal bias.

For shaded sites, the model tends to overestimate T_{MRT}. This limitation was also reported by Szucs et al. (2014) and Gál and Kántor (2020). Gál and Kántor

(2020) indicate that using domain-wide surface temperature in estimating longwave radiation and the theoretical approach the model follows in calculating fluxes from sun-exposed and shaded walls account for the overestimations in the shade. In this study, an imbalance in the dataset (a greater number of shade/tree observations) presents the model as one that overestimates. A balanced observational dataset would be needed to validate this hypothesis.

The validation demonstrates that SOLWEIG underestimates T_{MRT} at sun-exposed sites with an RMSE above the ISO limit by 1.6 °C. The bias noticed at these sites is mainly associated with underestimated downwelling radiation. Gál and Kántor (2020) also observed that SOLWEIG underestimates the lateral shortwave and longwave estimations for sun-exposed areas. Shade mismatch is an issue in estimating T_{MRT} under trees because the model relies on shade distribution to estimate radiative fluxes. Generating an appropriate tree geometry in the CDSM is challenging because it depends on the point cloud's density. Therefore, LiDAR data with a high point density should be preferred, especially in densely forested cities, to ensure that the CDSM accurately represents the geometries of tree crowns. The MaRTy observations used in this study were taken under various tree species with differing characteristics (leaf area density, canopy size, height, etc.). Azcarate et al. (2021) also show that tree transmissivity varies with foliation density. Although different transmissivities were investigated to find the optimum value for the entire area, a single transmissivity value is not valid for all trees. As a result, the T_{MRT} estimations do not match the MaRTy observations beneath some trees.

Despite the limited information on wall properties and the assumption of a uniform albedo for all sites, the model performed well for building canyons. According to ISO criteria, the building canyon RMSE is well within range, giving credence to the model. Again, shortwave radiation is underestimated, an observation also made by Lau et al. (2016).

Information on the land cover was not incorporated into the modeling process for this study. Prior research found that ground surface features have less influence on T_{MRT} than shading, despite being relevant (Lindberg et al., 2016; Middel & Krayenhoff, 2019). Including ground-surface features would reduce inaccuracies in upwelling longwave radiation and reflected shortwave radiation.

4.2.2 Sidewalk shade coverage

Most of the Phoenix-Tempe area is not walkable in the summer, according to the percentage of sidewalks that attain the required shade coverage throughout the day. Tracts in the downtown parts of Phoenix and Tempe had desirable shade coverage due to a combination of trees and tall buildings that enhances shade distribution in cities (Sabrin et al., 2021). According to Jamei & Rajagopalan (2017), tall buildings could offer more shade and lower T_{MRT} in downtown areas than trees. However, the sidewalk shade coverage is low outside the downtown area because the residential neighborhoods feature low-rise structures, broad roads,

and high sky view factors (LCZ 6). The share of tree shade is significantly greater than that of buildings in the residential areas.

Planting street trees could significantly improve thermal comfort in tracts with poor shading (Aminipouri et al., 2019; Sabrin et al., 2021; Tan et al., 2016). According to studies, adding more street trees could appreciably lower PET and T_{MRT} (Aminipouri et al., 2019; Gál & Kántor, 2020; Lachapelle et al., 2023; Sabrin et al., 2021). According to a study done in a residential area of Freiburg, trees lower T_{MRT} by 6.6 ºC, which results in a 3.0 ºC decrease in PET (Lee et al., 2016). There is potential for the residential areas in Phoenix and Tempe to achieve a similar impact, but the trees would need to be mature (Sabrin et al., 2021) and may require a lot of water, which is a trade-off in hot, dry climates (Middel et al., 2021). Planners could use advanced computational tools to simulate scenarios involving certain tree species and other conditions to maximize tree-planting efforts and increase pedestrian thermal comfort (Azcarate et al., 2021; Tan et al., 2016; Wallenberg et al., 2022). In the planning and optimization stages, metrics like the Shade Index employed by Aleksandrowicz et al. (2020) could be used. Other urban shade sources that have proven to reduce T_{MRT} comparably, particularly during the day (Middel et al., 2021), may be used as temporary solutions while trees grow. As Bäcklin et al. (2021) demonstrated, the shade analyses carried out for sidewalks may be reproduced for public spaces like parks and playgrounds for children with the output maps to assess the thermal comfort at these places.

4.2.3 Limitations and prospects

The lack of readily available sidewalk polygon data could be problematic for implementing this method in other regions. Given the street width data, this can be resolved by applying the buffer analysis used in this study. Alternatively, ML-based models could be used to extract sidewalk footprints from high-resolution aerial images.

Combining T_{MRT} with individual travel data will help evaluate pedestrian heat perception in dynamic situations (Dzyuban et al., 2022; Lau et al., 2019; Li et al., 2023). Further research should concentrate on pedestrian travel patterns and the heat exposure that results from incorporating fine-scale T_{MRT} data into trip simulation models.

5. CONCLUSIONS

Urban overheating is a significant challenge in the 21st century. Heat-related illnesses and deaths are preventable if the location of vulnerable populations and local biometeorological conditions are known to target heat mitigation interventions. Spatially explicit thermal data is essential for developing heat mitigation plans and preparing for extreme heat events. This thesis addresses challenges associated with thermal data for evaluating urban overheating at local and microscales.

Remotely sensed thermal data helps identify hotspots and evaluate urban overheating at the local climate scale. However, its utilization is hindered by missing data due to cloud contamination. Furthermore, there is a trade-off between spatial and temporal resolutions, given the limitations associated with sensor characteristics. This study addressed these challenges by filling gaps in satellite LST and generating synthetic products for evaluating urban overheating at the local scale. The thesis used RF models to fill gaps in 1 km MODIS LST and create synthetic NDVI and LST products at 500 m and 100 m resolutions. The results validate RF as a suitable algorithm to fill gaps resulting from cloud contamination and generate synthetic products at various resolutions. The RF models reconstruct and generate synthetic remote sensing data based on empirical relationships that are challenging to represent using statistical and spatial interpolation methods. This thesis advances the applicability of ML to enhance remotely sensed products by exploring the relationship between target variables (LST and NDVI) and predictors to reveal how the models arrive at an estimation using PDPs. In addition to showing how the models make estimations, the PDPs help explain the key influencers of overheating in urban spaces to policymakers and planners with actual figures. The approaches used in this thesis to enhance remote sensing products for assessing overheating at local scales are easy to implement; however, the limitations discussed must be considered beforehand.

T_{MRT} is a critical parameter for the microscale evaluation of urban overheating and the development of human-centric heat mitigation plans. This thesis advances human thermal exposure research by extensively validating the SOLWEIG model using six-directional field observations. The study shows that T_{MRT} estimates from SOLWEIG are acceptable following the ± 5 °C criteria required for ISO7726. To assist local municipalities with Cool Corridor Planning, the thesis calculated hourly shade coverage on sidewalks for a hot summer day in the Phoenix-Tempe area. The calculated shade coverage allows the identification of neighborhoods that do not meet the minimum shade coverage recommendation outlined in the Maricopa Association of Government's Active Transportation Plan. This research provides essential information on current shade coverage previously unavailable to local governments to make better-informed decisions on human-centric cooling strategies. The results will inform municipal plans to optimize shade presence, use, and effectiveness.

Combined with socio-economic and demographic data, the datasets generated in this thesis will help identify hotspots and inform heat action plans to make cities in the respective study sites heat-ready.

REFERENCES

Acero, J. A., & Arrizabalaga, J. (2018). Evaluating the performance of ENVI-met model in diurnal cycles for different meteorological conditions. *Theoretical and Applied Climatology*, *131*(1–2), 455–469. https://doi.org/10.1007/s00704-016-1971-y

Agam, N., Kustas, W. P., Anderson, M. C., Li, F., & Colaizzi, P. D. (2007). Utility of thermal sharpening over Texas high plains irrigated agricultural fields. *Journal of Geophysical Research*, *112*(D19), D19110. https://doi.org/10.1029/2007JD008407

Albarakat, R., & Lakshmi, V. (2019). Comparison of normalized difference vegetation index derived from landsat, MODIS, and AVHRR for the mesopotamian marshes between 2002 and 2018. *Remote Sensing*, *11*(10). https://doi.org/10.3390/rs11101245

Aleksandrowicz, O., Zur, S., Lebendiger, Y., & Lerman, Y. (2020). Shade maps for prioritizing municipal microclimatic action in hot climates: Learning from Tel Aviv-Yafo. *Sustainable Cities and Society*, *53*(August 2019), 101931. https://doi.org/10.1016/j.scs.2019.101931

Aminipouri, M., Knudby, A. J., Krayenhoff, E. S., Zickfeld, K., & Middel, A. (2019). Modelling the impact of increased street tree cover on mean radiant temperature across Vancouver's local climate zones. *Urban Forestry and Urban Greening*, *39*(January), 9–17. https://doi.org/10.1016/j.ufug.2019.01.016

Anguita, D., Ghelardoni, L., Ghio, A., Oneto, L., & Ridella, S. (2012). The 'K' in K-fold cross validation. *ESANN 2012 Proceedings, 20th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*, *April*, 441–446. http://www.i6doc.com/en/livre/?GCOI=28001100967420

Asadi, H., Shahedi, K., Jarihani, B., & Sidle, R. C. (2019). Rainfall-runoff modelling using hydrological connectivity index and artificial neural network approach. *Water (Switzerland)*, *11*(2). https://doi.org/10.3390/w11020212

Asgarian, A., Amiri, B. J., & Sakieh, Y. (2015). Assessing the effect of green cover spatial patterns on urban land surface temperature using landscape metrics approach. *Urban Ecosystems*, *18*(1), 209–222. https://doi.org/10.1007/s11252-014-0387-7

Atkinson, P. M. (2013). Downscaling in remote sensing. *International Journal of Applied Earth Observation and Geoinformation*, *22*, 106–114. https://doi.org/10.1016/j.jag.2012.04.012

Aviv, D., Gros, J., Alsaad, H., Teitelbaum, E., Voelker, C., Pantelic, J., & Meggers, F. (2022). A data-driven ray tracing simulation for mean radiant temperature and spatial variations in the indoor radiant field with experimental validation. *Energy and Buildings*, *254*, 111585. https://doi.org/10.1016/j.enbuild.2021.111585

Azcarate, I., Acero, J. Á., Garmendia, L., & Rojí, E. (2021). Tree layout methodology for shading pedestrian zones: Thermal comfort study in Bilbao (Northern Iberian Peninsula). *Sustainable Cities and Society*, *72*, 102996. https://doi.org/10.1016/J.SCS.2021.102996

AZMET. (2022). *The Arizona Meteorological Network*. https://ag.arizona.edu/azmet/

Bäcklin, O., Lindberg, F., Thorsson, S., Rayner, D., & Wallenberg, N. (2021). Outdoor heat stress at preschools during an extreme summer in Gothenburg, Sweden – Preschool teachers' experiences contextualized by radiation modelling. *Sustainable Cities and Society*, *75*, 103324. https://doi.org/10.1016/J.SCS.2021.103324

Banholzer, S., Kossin, J., & Donner, S. (2014). Reducing disaster: Early warning systems for climate change. In Z. Zommers & A. Singh (Eds.), *Reducing Disaster: Early Warning Systems for Climate Change* (Vol. 9789401785, pp. 21–49). Springer. https://doi.org/10.1007/978-94-017-8598-3

Bargagli Stoffi, F. J., Cevolani, G., & Gnecco, G. (2022). Simple Models in Complex Worlds: Occam's Razor and Statistical Learning Theory. *Minds and Machines*, *32*(1), 13–42. https://doi.org/10.1007/s11023-022-09592-z

Bartkowiak, P., Castelli, M., & Notarnicola, C. (2019). Downscaling land surface temperature from MODIS dataset with random forest approach over alpine vegetated areas. *Remote Sensing*, *11*(11), 1–19. https://doi.org/10.3390/rs11111319

Basara, J. B., Basara, H. G., Illston, B. G., & Crawford, K. C. (2010). The Impact of the Urban Heat Island during an Intense Heat Wave in Oklahoma City. *Advances in Meteorology*, *2010*, 1–10. https://doi.org/10.1155/2010/230365

Bechtel, B., Alexander, P. J., Beck, C., Böhner, J., Brousse, O., Ching, J., Demuzere, M., Fonte, C., Gál, T., Hidalgo, J., Hoffmann, P., Middel, A., Mills, G., Ren, C., See, L., Sismanidis, P., Verdonck, M. L., Xu, G., & Xu, Y. (2019). Generating WUDAPT Level 0 data – Current status of production and evaluation. *Urban Climate*, *27*(October 2018), 24–45. https://doi.org/10.1016/j.uclim.2018.10.001

Bechtel, B., Zakšek, K., & Hoshyaripour, G. (2012). Downscaling Land Surface Temperature in an Urban Area: A Case Study for Hamburg, Germany. *Remote Sensing*, *4*(10), 3184–3200. https://doi.org/10.3390/rs4103184

Belgiu, M., & Drăgu, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, *114*, 24–31. https://doi.org/10.1016/J.ISPRSJPRS.2016.01.011

Bernales, A. M., Antolihao, J. A., Samonte, C., Campomanes, F., Rojas, R. J., Dela Serna, A. M., & Silapan, J. (2016). Modelling the relationship between land surface temperature and landscape patterns of land use land cover classification using multi linear regression models. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences – ISPRS Archives*, *41*(July), 851–856. https://doi.org/10.5194/isprsarchives-XLI-B8-851-2016

Bisquert, M., Sánchez, J. M., & Caselles, V. (2016). Evaluation of Disaggregation Methods for Downscaling MODIS Land Surface Temperature to Landsat Spatial Resolution in Barrax Test Site. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *9*(4), 1430–1438. https://doi.org/10.1109/JSTARS.2016.2519099

Breiman, L. (2001). Random forests. In R. E. Schapire (Ed.), *Machine Learning* (45th ed., pp. 5–32). Kluwer Academic Publishers. https://doi.org/10.1201/9780367816377-11

Bruse, M., & Fleer, H. (1998). Simulating surface–plant–air interactions inside urban environments with a three dimensional numerical model. *Environmental Modelling & Software*, *13*(3–4), 373–384. https://doi.org/10.1016/S1364-8152(98)00042-5

Butt, A., Amati, M., Sun, C., & Deilami, K. (2019). Planning Shadeways in Bendigo: An Example of Digital Planning to Adapt to Extreme Heat'. In G.: Williams (Ed.), *TREENET Proceedings of the 20th National Street Tree Symposium 2019* (pp. 65–75).

Chakraborty, T., Hsu, A., Manya, D., & Sheriff, G. (2020). A spatially explicit surface urban heat island database for the United States: Characterization, uncertainties, and possible applications. *ISPRS Journal of Photogrammetry and Remote Sensing*, *168*, 74–88. https://doi.org/10.1016/j.isprsjprs.2020.07.021

Chen, F., Kusaka, H., Bornstein, R., Ching, J., Grimmond, C. S. B., Grossman-Clarke, S., Loridan, T., Manning, K. W., Martilli, A., Miao, S., Sailor, D., Salamanca, F. P., Taha, H., Tewari, M., Wang, X., Wyszogrodzki, A. A., & Zhang, C. (2011). The integrated WRF/urban modelling system: development, evaluation, and applications to urban environmental problems. *International Journal of Climatology*, *31*(2), 273–288. https://doi.org/10.1002/joc.2158

Chen, Y.-C., Lin, T.-P., & Matzarakis, A. (2014). Comparison of mean radiant temperature from field experiment and modelling: a case study in Freiburg, Germany. *Theoretical and Applied Climatology*, *118*(3), 535–551. https://doi.org/10.1007/s00704-013-1081-z

Cheval, S., Dumitrescu, A., & Bell, A. (2009). The urban heat island of Bucharest during the extreme high temperatures of July 2007. *Theoretical and Applied Climatology*, *97*(3–4), 391–401. https://doi.org/10.1007/s00704-008-0088-3

Cheval, S., Dumitrescu, A., Irașoc, A., Paraschiv, M. G., Perry, M., & Ghent, D. (2022). MODIS-based climatology of the Surface Urban Heat Island at country scale (Romania). *Urban Climate*, *41*, 101056. https://doi.org/10.1016/J.UCLIM.2021.101056

City of Phoenix. (2009). *Street Planning and Design Guidelines*. https://www.phoenix.gov/streetssite/documents/d_039263.pdf

Crank, P. J., Middel, A., Wagner, M., Hoots, D., Smith, M., & Brazel, A. (2020). Validation of seasonal mean radiant temperature simulations in hot arid urban climates. *Science of the Total Environment*, *749*, 141392. https://doi.org/10.1016/j.scitotenv.2020.141392

Cristóbal, J., Jiménez-Muñoz, J. C., Prakash, A., Mattar, C., Skoković, D., & Sobrino, J. A. (2018). An improved single-channel method to retrieve land surface temperature from the landsat-8 thermal band. *Remote Sensing*, *10*(3). https://doi.org/10.3390/rs10030431

Crosson, W. L., Al-Hamdan, M. Z., Hemmings, S. N. J., & Wade, G. M. (2012). A daily merged MODIS Aqua–Terra land surface temperature data set for the conterminous United States. *Remote Sensing of Environment*, *119*, 315–324. https://doi.org/10.1016/J.RSE.2011.12.019

Defries, R. S., & Townshend, J. R. G. (1994). NDVI-derived land cover classifications at a global scale. *International Journal of Remote Sensing*, *15*(17), 3567–3586. https://doi.org/10.1080/01431169408954345

Deilami, K., Kamruzzaman, M., & Liu, Y. (2018). Urban heat island effect: A systematic review of spatio-temporal factors, data, methods, and mitigation measures. *International Journal of Applied Earth Observation and Geoinformation*, *67*(November 2017), 30–42. https://doi.org/10.1016/j.jag.2017.12.009

Deilami, K., Rudner, J., Butt, A., MacLeod, T., Williams, G., Romeijn, H., & Amati, M. (2020). Allowing users to benefit from tree shading: Using a smartphone app to allow adaptive route planning during extreme heat. *Forests*, *11*(9), 1–15. https://doi.org/10.3390/f11090998

Demuzere, M., Bechtel, B., Middel, A., & Mills, G. (2019). Mapping Europe into local climate zones. *PLoS ONE*, *14*(4), 1–2. https://doi.org/10.1371/journal.pone.0214474

Dong, P., Gao, L., Zhan, W., Liu, Z., Li, J., Lai, J., Li, H., Huang, F., Tamang, S. K., & Zhao, L. (2020). Global comparison of diverse scaling factors and regression models for downscaling Landsat-8 thermal data. *ISPRS Journal of Photogrammetry and Remote Sensing*, *169*(August), 44–56. https://doi.org/10.1016/j.isprsjprs.2020.08.018

Dutta, K., Basu, D., & Agrawal, S. (2021). Evaluation of seasonal variability in magnitude of urban heat islands using local climate zone classification and surface albedo. *International Journal of Environmental Science and Technology*. https://doi.org/10.1007/S13762-021-03602-W

Dzyuban, Y., Hondula, D. M., Coseo, P. J., & Redman, C. L. (2022). Public transit infrastructure and heat perceptions in hot and dry climates. *International Journal of Biometeorology*, *66*(2), 345–356. https://doi.org/10.1007/s00484-021-02074-4

Dzyuban, Y., Ching, G. N. Y., Yik, S. K., Tan, A. J., Banerjee, S., Crank, P. J., & Chow, W. T. L. (2022). Outdoor thermal comfort research in transient conditions: A narrative literature review. *Landscape and Urban Planning*, *226*, 104496. https://doi.org/10.1016/j.landurbplan.2022.104496

Ethem, A. (2010). *Introduction machine Learning* (2nd ed.). The MIT Press.

Fan, D., Su, X., Weng, B., Wang, T., & Yang, F. (2021). Research Progress on Remote Sensing Classification Methods for Farmland Vegetation. *AgriEngineering*, *3*(4), 971–989. https://doi.org/10.3390/agriengineering3040061

Fan, X.-M., Liu, H.-G., Liu, G.-H., & Li, S.-B. (2014). Reconstruction of MODIS land-surface temperature in a flat terrain and fragmented landscape. *International Journal of Remote Sensing*, *35*(23), 7857–7877. https://doi.org/10.1080/01431161.2014.978036

Farhadur Rahman, M., Onoda, Y., & Kitajima, K. (2022). Forest canopy height variation in relation to topography and forest types in central Japan with LiDAR. *Forest Ecology and Management*, *503*, 119792. https://doi.org/10.1016/j.foreco.2021.119792

Feng, Y., Gao, C., Tong, X., Chen, S., Lei, Z., & Wang, J. (2019). Spatial patterns of land surface temperature and their influencing factors: A case study in Suzhou, China. *Remote Sensing*, *11*(2). https://doi.org/10.3390/rs11020182

Fichera, C. R., Modica, G., & Pollino, M. (2012). Land Cover classification and change-detection analysis using multi-temporal remote sensed imagery and landscape metrics. *European Journal of Remote Sensing*, *45*(1), 1–18. https://doi.org/10.5721/EuJRS20124501

Filgueiras, R., Mantovani, E. C., Fernandes-Filho, E. I., da Cunha, F. F., Althoff, D., & Dias, S. H. B. (2020). Fusion of MODIS and landsat-like images for daily high spatial resolution NDVI. *Remote Sensing*, *12*(8), 11–13. https://doi.org/10.3390/RS12081297

Fortuniak, K., Kłysik, K., & Wibig, J. (2006). Urban–rural contrasts of meteorological parameters in Łódź. *Theoretical and Applied Climatology*, *84*(1–3), 91–101. https://doi.org/10.1007/s00704-005-0147-y

Franke, J., Heinzel, V., & Menz, G. (2006). Assessment of NDVI – Differences Caused by Sensor Specific Relative Spectral Response Functions. *2006 IEEE International Symposium on Geoscience and Remote Sensing*, 1138–1141. https://doi.org/10.1109/IGARSS.2006.294

Franklin, R., Inanc, M., & Xie, Z. (2006). Terrain Elevation Data Structure Operations. *Autocarto 2006*. https://cartogis.org/docs/proceedings/2006/franklin_inanc_xie.pdf

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, *29*(5), 1189–1232. https://doi.org/10.1214/aos/1013203451

Gál, C. V., & Kántor, N. (2020). Modeling mean radiant temperature in outdoor spaces, A comparative numerical simulation and validation study. *Urban Climate*, *32*(December 2019), 100571. https://doi.org/10.1016/j.uclim.2019.100571

Gao, F., Anderson, M. C., Zhang, X., Yang, Z., Alfieri, J. G., Kustas, W. P., Mueller, R., Johnson, D. M., & Prueger, J. H. (2017). Toward mapping crop progress at field scales through fusion of Landsat and MODIS imagery. *Remote Sensing of Environment*, *188*, 9–25. https://doi.org/10.1016/j.rse.2016.11.004

García-León, D., Casanueva, A., Standardi, G., Burgstall, A., Flouris, A. D., & Nybo, L. (2021). Current and projected regional economic impacts of heatwaves in Europe. *Nature Communications*, *12*(1), 1–10. https://doi.org/10.1038/s41467-021-26050-z

Gastellu-Etchegorry, J. P., Grau, E., & Lauret, N. (2012). DART: A 3D Model for Remote Sensing Images and Radiative Budget of Earth Surfaces. In *Modeling and Simulation in Engineering* (Issue March). InTech. https://doi.org/10.5772/31315

Gerber, F., De Jong, R., Schaepman, M. E., Schaepman-Strub, G., & Furrer, R. (2018). Predicting missing values in spatio-temporal remote sensing data. *IEEE Transactions on Geoscience and Remote Sensing*, *56*(5), 2841–2853. https://doi.org/10.1109/TGRS.2017.2785240

Giridharan, R., & Emmanuel, R. (2018). The impact of urban compactness, comfort strategies and energy consumption on tropical urban heat island intensity: A review. *Sustainable Cities and Society*, *40*, 677–687. https://doi.org/10.1016/j.scs.2018.01.024

Gu, Y., & Wylie, B. (2015). Downscaling 250-m MODIS Growing Season NDVI Based on Multiple-Date Landsat Images and Data Mining Approaches. *Remote Sensing*, *7*(4), 3489–3506. https://doi.org/10.3390/rs70403489

Guo, G., Wu, Z., Xiao, R., Chen, Y., Liu, X., & Zhang, X. (2015). Impacts of urban biophysical composition on land surface temperature in urban heat island clusters. *Landscape and Urban Planning*, *135*, 1–10. https://doi.org/10.1016/J.LANDURBPLAN. 2014.11.007

Guo, H., Aviv, D., Loyola, M., Teitelbaum, E., Houchois, N., & Meggers, F. (2020). On the understanding of the mean radiant temperature within both the indoor and outdoor environment, a critical review. *Renewable and Sustainable Energy Reviews*, *117* (October 2019), 109207. https://doi.org/10.1016/j.rser.2019.06.014

Haddad, S., Ulpiani, G., Paolini, R., Synnefa, A., & Santamouris, M. (2020). Experimental and Theoretical analysis of the urban overheating and its mitigation potential in a hot arid city–Alice Springs. *Architectural Science Review*, *63*(5), 425–440. https://doi.org/10.1080/00038628.2019.1674128

Handy, S. L., Boarnet, M. G., Ewing, R., & Killingsworth, R. E. (2002). How the built environment affects physical activity: Views from urban planning. *American Journal of Preventive Medicine*, *23*(2S), 64–73. https://doi.org/10.1016/S0749-3797(02)00475-0

Hardy, J. D., & Stoll, A. M. (1954). Measurement of radiant heat load on man in summer and winter Alaskan climates. *Journal of Applied Physiology*, *7*(2), 200–211. https://doi.org/10.1152/jappl.1954.7.2.200

Hastie, T., Tibshirani, R., & Friedman, J. (2008). *The Elements of Statistical Learning Data Mining, Inference, and Prediction* (Springer (ed.); 2nd ed.).

Holloway-Brown, J., Helmstedt, K. J., & Mengersen, K. L. (2020). Stochastic spatial random forest (SS-RF) for interpolating probabilities of missing land cover data. *Journal of Big Data*, *7*(1), 55. https://doi.org/10.1186/s40537-020-00331-8

Höppe, P. (1992). A new procedure to determine the mean radiant temperature outdoors. *Wetter Leben*, *44*, 147–151.

Hu, L., Sun, Y., Collins, G., & Fu, P. (2020). Improved estimates of monthly land surface temperature from MODIS using a diurnal temperature cycle (DTC) model. In *ISPRS Journal of Photogrammetry and Remote Sensing* (Vol. 168, pp. 131–140). https://doi.org/10.1016/j.isprsjprs.2020.08.007

Huryna, H., Cohen, Y., Karnieli, A., Panov, N., Kustas, W. P., & Agam, N. (2019). Evaluation of TsHARP utility for thermal sharpening of Sentinel-3 satellite images using Sentinel-2 visual imagery. *Remote Sensing*, *11*(19). https://doi.org/10.3390/rs11192304

Hutengs, C., & Vohland, M. (2016). Downscaling land surface temperatures at regional scales with random forest regression. *Remote Sensing of Environment*, *178*, 127–141. https://doi.org/10.1016/j.rse.2016.03.006

Imhoff, M. L., Zhang, P., Wolfe, R. E., & Bounoua, L. (2010). Remote sensing of the urban heat island effect across biomes in the continental USA. *Remote Sensing of Environment*, *114*(3), 504–513. https://doi.org/10.1016/j.rse.2009.10.008

International Organization for Standardization. (1998). *Ergonomics of the thermal environment — Instruments for measuring physical quantities*. International Organization for Standardization.

Jamei, E., & Rajagopalan, P. (2017). Urban development and pedestrian thermal comfort in Melbourne. *Solar Energy*, *144*, 681–698. https://doi.org/10.1016/j.solener.2017.01.023

Järvi, L., Grimmond, C. S. B., & Christen, A. (2011). The Surface Urban Energy and Water Balance Scheme (SUEWS): Evaluation in Los Angeles and Vancouver. *Journal of Hydrology*, *411*(3–4), 219–237. https://doi.org/10.1016/j.jhydrol.2011.10.001

Jeevalakshmi, D., Reddy, S. N., & Manikiam, B. (2016). Land cover classification based on NDVI using LANDSAT8 time series: A case study Tirupati region. *2016 International Conference on Communication and Signal Processing (ICCSP)*, 1332–1335. https://doi.org/10.1109/ICCSP.2016.7754369

Jing, W., Yang, Y., Yue, X., & Zhao, X. (2016). A Comparison of Different Regression Algorithms for Downscaling Monthly Satellite-Based Precipitation over North China. *Remote Sensing*, *8*(10), 835. https://doi.org/10.3390/rs8100835

Ju, J., Roy, D. P., Vermote, E., Masek, J., & Kovalskyy, V. (2012). Continental-scale validation of MODIS-based and LEDAPS Landsat ETM+ atmospheric correction methods. *Remote Sensing of Environment*, *122*, 175–184. https://doi.org/10.1016/j.rse.2011.12.025

Kapwata, T., Gebreslasie, M. T., & Wright, C. Y. (2022). An analysis of past and future heatwaves based on a heat-associated mortality threshold: towards a heat health warning system. *Environmental Health: A Global Access Science Source*, *21*(1), 1–12. https://doi.org/10.1186/s12940-022-00921-4

Ke, L., Ding, X., & Song, C. (2013). Reconstruction of Time-Series MODIS LST in Central Qinghai-Tibet Plateau Using Geostatistical Approach. *IEEE Geoscience and Remote Sensing Letters*, *10*(6), 1602–1606. https://doi.org/10.1109/LGRS.2013.2263553

Ke, Y., Im, J., Lee, J., Gong, H., & Ryu, Y. (2015). Characteristics of Landsat 8 OLI-derived NDVI by comparison with multiple satellite sensors and in-situ observations. *Remote Sensing of Environment*, *164*, 298–313. https://doi.org/10.1016/j.rse.2015.04.004

Keellings, D., & Moradkhani, H. (2020). Spatiotemporal Evolution of Heat Wave Severity and Coverage Across the United States. *Geophysical Research Letters*, *47*(9), 1–9. https://doi.org/10.1029/2020GL087097

Kong, F., Chen, J., Middel, A., Yin, H., Li, M., Sun, T., Zhang, N., Huang, J., Liu, H., Zhou, K., & Ma, J. (2022). Impact of 3-D urban landscape patterns on the outdoor thermal environment: A modelling study with SOLWEIG. *Computers, Environment and Urban Systems*, *94*, 101773. https://doi.org/10.1016/J.COMPENVURBSYS.2022.101773

Krüger, E. L., Minella, F. O., & Matzarakis, A. (2014). Comparison of different methods of estimating the mean radiant temperature in outdoor thermal comfort studies. *International Journal of Biometeorology*, *58*(8), 1727–1737. https://doi.org/10.1007/s00484-013-0777-1

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

Lachapelle, J. A., Scott Krayenhoff, E., Middel, A., Coseo, P., & Warland, J. (2023). Maximizing the pedestrian radiative cooling benefit per street tree. *Landscape and Urban Planning*, *230*(October 2022), 104608. https://doi.org/10.1016/j.landurbplan.2022.104608

Lau, K. K.-L., Shi, Y., & Ng, E. Y.-Y. (2019). Dynamic response of pedestrian thermal comfort under outdoor transient conditions. *International Journal of Biometeorology*, *63*(7), 979–989. https://doi.org/10.1007/s00484-019-01712-2

Lee, H., Mayer, H., & Chen, L. (2016). Contribution of trees and grasslands to the mitigation of human heat stress in a residential district of Freiburg, Southwest Germany. *Landscape and Urban Planning*, *148*, 37–50. https://doi.org/10.1016/J.LANDURBPLAN.2015.12.004

Li, D., & Bou-Zeid, E. (2013). Synergistic interactions between urban heat islands and heat waves: The impact in cities is larger than the sum of its parts. *Journal of Applied Meteorology and Climatology*, *52*(9), 2051–2064. https://doi.org/10.1175/JAMC-D-13-02.1

Li, R., Chester, M. V., Middel, A., Vanos, J. K., Hernandez-Cortes, D., Buo, I., & Hondula, D. M. (2023). Effectiveness of travel behavior and infrastructure change to mitigate heat exposure. *Frontiers in Sustainable Cities*, *5*. https://doi.org/10.3389/frsc.2023.1129388

Li, W., Ni, L., Li, Z. L., Duan, S. B., & Wu, H. (2019). Evaluation of machine learning algorithms in spatial downscaling of modis land surface temperature. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *12*(7), 2299–2307. https://doi.org/10.1109/JSTARS.2019.2896923

Li, X., Zhou, Y., Asrar, G. R., & Zhu, Z. (2018). Creating a seamless 1 km resolution daily land surface temperature dataset for urban and surrounding areas in the conterminous United States. *Remote Sensing of Environment*, *206*(February 2017), 84–97. https://doi.org/10.1016/j.rse.2017.12.010

Lindberg, F., & Grimmond, C. S. B. (2011). The influence of vegetation and building morphology on shadow patterns and mean radiant temperatures in urban areas: Model development and evaluation. *Theoretical and Applied Climatology*, *105*(3), 311–323. https://doi.org/10.1007/s00704-010-0382-8

Lindberg, F., Holmer, B., & Thorsson, S. (2008). SOLWEIG 1.0 – Modelling spatial variations of 3D radiant fluxes and mean radiant temperature in complex urban settings. *International Journal of Biometeorology*, *52*(7), 697–713. https://doi.org/10.1007/s00484-008-0162-7

Lindberg, F., Onomura, S., & Grimmond, C. S. B. (2016). Influence of ground surface characteristics on the mean radiant temperature in urban areas. *International Journal of Biometeorology*, *60*(9), 1439–1452. https://doi.org/10.1007/s00484-016-1135-x

Liu, H., Lu, N., Jiang, H., Qin, J., & Yao, L. (2020). Filling Gaps of Monthly Terra/MODIS Daytime Land Surface Temperature Using Discrete Cosine Transform Method. *Remote Sensing*, *12*(3), 361. https://doi.org/10.3390/rs12030361

Malamiri, H. R. G., Rousta, I., Olafsson, H., Zare, H., & Zhang, H. (2018). Gap-filling of MODIS time series land surface temperature (LST) products using singular spectrum analysis (SSA). *Atmosphere*, *9*(9). https://doi.org/10.3390/atmos9090334

Maricopa Association of Governments. (2020). *MAG Active Transportation Plan*.

Matzarakis, A., Rutz, F., & Mayer, H. (2010). Modelling radiation fluxes in simple and complex environments: basics of the RayMan model. *International Journal of Biometeorology*, *54*(2), 131–139. https://doi.org/10.1007/s00484-009-0261-0

Meehl, G. A., & Tebaldi, C. (2004). More intense, more frequent, and longer lasting heat waves in the 21st century. *Science*, *305*(5686), 994–997. https://doi.org/10.1126/science.1098704

Metz, M., Andreo, V., & Neteler, M. (2017). A new fully gap-free time series of land surface temperature from MODIS LST data. *Remote Sensing*, *9*(12), 1333. https://doi.org/10.3390/rs9121333

Middel, A., Alkhaled, S., Schneider, F. A., Hagen, B., & Coseo, P. (2021). 50 Grades of Shade. *Bulletin of the American Meteorological Society*, 1–35. https://doi.org/10.1175/bams-d-20-0193.1

Middel, A., & Krayenhoff, E. S. (2019). Micrometeorological determinants of pedestrian thermal exposure during record-breaking heat in Tempe, Arizona: Introducing the MaRTy observational platform. *Science of The Total Environment*, *687*, 137–151. https://doi.org/10.1016/J.SCITOTENV.2019.06.085

Middel, A., Lukasczyk, J., Zakrzewski, S., Arnold, M., & Maciejewski, R. (2019). Urban form and composition of street canyons: A human-centric big data and deep learning approach. *Landscape and Urban Planning*, *183*(October 2017), 122–132. https://doi.org/10.1016/j.landurbplan.2018.12.001

Middel, A., Selover, N., Hagen, B., & Chhetri, N. (2016). Impact of shade on outdoor thermal comfort—a seasonal field study in Tempe, Arizona. *International Journal of Biometeorology*, *60*(12), 1849–1861. https://doi.org/10.1007/s00484-016-1172-5

Moisen, G. G. (2008). Classification and regression trees. In S. E. Orgensen & B. D. Fath (Eds.), *Encyclopedia of Ecology* (1st ed., pp. 582–588). Routledge. https://doi.org/10.1201/9781315139470

Molnar, C. (2020). *Interpretable Machine Learning A Guide for Making Black Box Models Explainable*. Lean Publishing. https://christophm.github.io/inter%5C%0Apretable-ml-book

Nazarian, N., Krayenhoff, E. S., Bechtel, B., Hondula, D. M., Paolini, R., Vanos, J., Cheung, T., Chow, W. T. L., de Dear, R., Jay, O., Lee, J. K. W., Martilli, A., Middel, A., Norford, L. K., Sadeghi, M., Schiavon, S., & Santamouris, M. (2022). Integrated Assessment of Urban Overheating Impacts on Human Life. *Earth's Future*, *10*(8). https://doi.org/10.1029/2022EF002682

Nazarian, N., & Lee, J. K. W. (2021). Personal assessment of urban heat exposure: A systematic review. *Environmental Research Letters*, *16*(3). https://doi.org/10.1088/1748-9326/abd350

Nomura, R., & Oki, K. (2021). Downscaling of MODIS NDVI by Using a Convolutional Neural Network-Based Model with Higher Resolution SAR Data. *Remote Sensing*, *13*(4), 732. https://doi.org/10.3390/rs13040732

Oke, T., Mills, G., Christen, A., & Voogt, J. A. (2017). *Urban Climates*. Cambridge University Press. https://doi.org/10.1017/9781139016476

Oke, T. R. (1973). City size and the urban heat island. *Atmospheric Environment (1967)*, *7*(8), 769–779. https://doi.org/10.1016/0004-6981(73)90140-6

Osborne, P. E., & Alvares-Sanches, T. (2019). Quantifying how landscape composition and configuration affect urban land surface temperatures using machine learning and neutral landscapes. *Computers, Environment and Urban Systems*, *76*(April), 80–90. https://doi.org/10.1016/j.compenvurbsys.2019.04.003

Pan, X., Zhu, X., Yang, Y., Cao, C., Zhang, X., & Shan, L. (2018). Applicability of Downscaling Land Surface Temperature by Using Normalized Difference Sand Index. *Scientific Reports*, *8*(1), 1–14. https://doi.org/10.1038/s41598-018-27905-0

Peng, X., Wu, W., Zheng, Y., Sun, J., Hu, T., & Wang, P. (2020). Correlation analysis of land surface temperature and topographic elements in Hangzhou, China. *Scientific Reports*, *10*(1), 10451. https://doi.org/10.1038/s41598-020-67423-6

Perkins-Kirkpatrick, S. E., & Lewis, S. C. (2020). Increasing trends in regional heatwaves. *Nature Communications*, *11*(1), 1–8. https://doi.org/10.1038/s41467-020-16970-7

Perkins, S. E., & Alexander, L. V. (2013). On the measurement of heat waves. *Journal of Climate*, *26*(13), 4500–4517. https://doi.org/10.1175/JCLI-D-12-00383.1

Pu, R., & Bonafoni, S. (2023). Thermal infrared remote sensing data downscaling investigations: An overview on current status and perspectives. *Remote Sensing Applications: Society and Environment*, *29*(December 2022), 100921. https://doi.org/10.1016/j.rsase.2023.100921

Rajulapati, C. R., Gaddam, R. K., Nerantzaki, S. D., Papalexiou, S. M., Cannon, A. J., & Clark, M. P. (2022). Exacerbated heat in large Canadian cities. *Urban Climate*, *42* (January), 101097. https://doi.org/10.1016/j.uclim.2022.101097

Rao, P. K. (1972). Remote sensing of urban "heat islands" from an environmental satellite. *Bulletin of the American Meteorological Society*, *53*, 647–648.

Raschka, S. (2015). Python machine learning: unlock deeper insights into machine learning with this vital guide to cutting-edge predictive analytics. In *Python machine learning :* (1st ed.). Packt Publishing. www.packtpub.com

Ratti, C., Di Sabatino, S., & Britter, R. (1999). Urban Texture Analysis with Image Processing Techniques Carlo. *CAADFutures99*.

Roth, M. (2013). Urban Heat Islands. In H. J. Fernando (Ed.), *Handbook of Environmental Fluid Dynamics, Volume Two: Systems, Pollution, Modeling, and Measurements* (1st ed., pp. 143–158). CRC Press. https://doi.org/10.1201/b13691

Sabrin, S., Karimi, M., Nazari, R., Pratt, J., & Bryk, J. (2021). Effects of Different Urban-Vegetation Morphology on the Canopy-level Thermal Comfort and the Cooling Benefits of Shade Trees: Case-study in Philadelphia. *Sustainable Cities and Society*, *66*(September 2020), 102684. https://doi.org/10.1016/j.scs.2020.102684

Sarafanov, M., Kazakov, E., Nikitin, N. O., & Kalyuzhnaya, A. V. (2020). A Machine Learning Approach for Remote Sensing Data Gap-Filling with Open-Source Implementation: An Example Regarding Land Surface Temperature, Surface Albedo and NDVI. *Remote Sensing*, *12*(23), 3865. https://doi.org/10.3390/rs12233865

Sasidhar, T. T., K., S., M.T., V., V., S., & K.P., S. (2019). Land Cover Satellite Image Classification Using NDVI and SimpleCNN. *2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, 1–5. https://doi.org/10.1109/ICCCNT45670.2019.8944840

Scarano, M., & 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. https://doi.org/10.1080/01431161.2017.1368099

Schwaab, J., Meier, R., Mussetti, G., Seneviratne, S., Bürgi, C., & Davin, E. L. (2021). The role of urban trees in reducing land surface temperatures in European cities. *Nature Communications*, *12*(1), 6763. https://doi.org/10.1038/s41467-021-26768-w

Schwarz, N., Lautenbach, S., & Seppelt, R. (2011). Exploring indicators for quantifying surface urban heat islands of European cities with MODIS land surface temperatures. *Remote Sensing of Environment*, *115*(12), 3175–3186. https://doi.org/10.1016/j.rse.2011.07.003

Seneviratne, S. I., Nicholls, N., Easterling, D., Goodess, C. M., Kanae, S., Kossin, J., Luo, Y., Marengo, J., Mc Innes, K., Rahimi, M., Reichstein, M., Sorteberg, A., Vera, C., Zhang, X., Rusticucci, M., Semenov, V., Alexander, L. V., Allen, S., Benito, G., … Zwiers, F. W. (2012). Changes in climate extremes and their impacts on the natural physical environment. *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation: Special Report of the Intergovernmental Panel on Climate Change*, *9781107025*, 109–230. https://doi.org/10.1017/CBO9781139177245.006

Shalev-Shwartz, S., & Ben-David, S. (2013). *Understanding machine learning: From theory to algorithms* (Vol. 9781107057). https://doi.org/10.1017/CBO9781107298019

Sharifnezhadazizi, Z., Norouzi, H., Prakash, S., Beale, C., & Khanbilvardi, R. (2019). A Global Analysis of Land Surface Temperature Diurnal Cycle Using MODIS Observations. *Journal of Applied Meteorology and Climatology*, *58*(6), 1279–1291. https://doi.org/10.1175/JAMC-D-18-0256.1

Shen, H., Huang, L., Zhang, L., Wu, P., & Zeng, C. (2016). Long-term and fine-scale satellite monitoring of the urban heat island effect by the fusion of multi-temporal and multi-sensor remote sensed data: A 26-year case study of the city of Wuhan in China. *Remote Sensing of Environment*, *172*, 109–125. https://doi.org/10.1016/j.rse.2015.11.005

Siabi, N., Sanaeinejad, S. H., & Ghahraman, B. (2022). Effective method for filling gaps in time series of environmental remote sensing data: An example on evapotranspiration and land surface temperature images. *Computers and Electronics in Agriculture*, *193*, 106619. https://doi.org/10.1016/j.compag.2021.106619

Siji George, C. G., & Sumathi, B. (2020). Grid search tuning of hyperparameters in random forest classifier for customer feedback sentiment prediction. *International Journal of Advanced Computer Science and Applications*, *11*(9), 173–178. https://doi.org/10.14569/IJACSA.2020.0110920

Speak, A., Montagnani, L., Wellstein, C., & Zerbe, S. (2021). Forehead temperatures as an indicator of outdoor thermal comfort and the influence of tree shade. *Urban Climate*, *39*(March), 100965. https://doi.org/10.1016/j.uclim.2021.100965

Steuri, B., Bender, S., & Cortekar, J. (2020). Successful user-science interaction to co-develop the new urban climate model PALM-4U. *Urban Climate*, *32*, 100630. https://doi.org/10.1016/j.uclim.2020.100630

Steven, M. D., Malthus, T. J., Baret, F., Xu, H., & Chopping, M. J. (2003). Intercalibration of vegetation indices from different sensor systems. *Remote Sensing of Environment*, *88*(4), 412–422. https://doi.org/10.1016/J.RSE.2003.08.010

Stewart, I. D., & Oke, T. R. (2012). Local Climate Zones for Urban Temperature Studies. *Bulletin of the American Meteorological Society*, *93*(12), 1879–1900. https://doi.org/10.1175/BAMS-D-11-00019.1

Szucs, Á., Gál, T., & Andrade, H. (2014). Comparison of measured and simulated mean radiant temperature. Case study in Lisbon (Portugal). *Finisterra*, *49*(98), 95–111. https://doi.org/10.18055/Finis6469

Tabatabaie, S., Litt, J. S., & Carrico, A. (2019). A study of perceived nature, shade and trees and self-reported physical activity in Denver. *International Journal of Environmental Research and Public Health*, *16*(19). https://doi.org/10.3390/ijerph16193604

Tan, Z., Lau, K. K. L., & Ng, E. (2016). Urban tree design approaches for mitigating daytime urban heat island effects in a high-density urban environment. *Energy and Buildings*, *114*, 265–274. https://doi.org/10.1016/J.ENBUILD.2015.06.031

Tilly, N., Hoffmeister, D., Cao, Q., Lenz-Wiedemann, V., Miao, Y., & Bareth, G. (2013). Precise plant height monitoring and biomass estimation with Terrestrial Laser Scanning in paddy rice. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *II-5/W2*, 295–300. https://doi.org/10.5194/isprsannals-II-5-W2-295-2013

Trishchenko, A. P., Cihlar, J., & Li, Z. (2002). Effects of spectral response function on surface reflectance and NDVI measured with moderate resolution satellite sensors. *Remote Sensing of Environment*, *81*(1), 1–18. https://doi.org/10.1016/S0034-4257(01)00328-5

Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, *8*(2), 127–150. https://doi.org/10.1016/0034-4257(79)90013-0

Vanos, J. K., Rykaczewski, K., Middel, A., Vecellio, D. J., Brown, R. D., & Gillespie, T. J. (2021). Improved methods for estimating mean radiant temperature in hot and sunny outdoor settings. *International Journal of Biometeorology*, *65*(6), 967–983. https://doi.org/10.1007/s00484-021-02131-y

Vanos, J. K., Warland, J. S., Gillespie, T. J., & Kenny, N. A. (2010). Review of the physiology of human thermal comfort while exercising in urban landscapes and implications for bioclimatic design. *International Journal of Biometeorology*, *54*(4), 319–334. https://doi.org/10.1007/s00484-010-0301-9

VDI. (1998). *Methods for the Human-Biometeorological Assessment of Climate and Air Hygiene for Urban and Regional Planning. Part 1: Climate. DI 3787*.

Voogt, J. A., & Oke, T. R. (2003). Thermal remote sensing of urban climates. *Remote Sensing of Environment*, *86*(3), 370–384. https://doi.org/10.1016/S0034-4257(03)00079-8

Wahba Tadros, S. N., Wellenstein, A., Das, M. B., Palmarini, N., D'Aoust, O. S., Singh, G., Restrepo Cadavid, P., Goga, S., Terraza, H. C., Lakovits, C., Baeumler, A. E. N. ., & Gapihan, A. T. (2021). *Demographic Trends and Urbanization*. World Bank Group. https://doi.org/10.1596/35469

Wallenberg, N., Lindberg, F., & Rayner, D. (2022). Locating trees to mitigate outdoor radiant load of humans in urban areas using a metaheuristic hill-climbing algorithm - introducing TreePlanter v1.0. *Geoscientific Model Development*, *15*(3), 1107–1128. https://doi.org/10.5194/gmd-15-1107-2022

Wang, C., Middel, A., Myint, S. W., Kaplan, S., Brazel, A. J., & Lukasczyk, J. (2018). Assessing local climate zones in arid cities: The case of Phoenix, Arizona and Las Vegas, Nevada. *ISPRS Journal of Photogrammetry and Remote Sensing*, *141*, 59–71. https://doi.org/10.1016/j.isprsjprs.2018.04.009

Wang, S., Luo, X., & Peng, Y. (2020). Spatial downscaling of MODIS land surface temperature based on geographically weighted autoregressive model. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *13*, 2532–2546. https://doi.org/10.1109/JSTARS.2020.2968809

Weiss, D. J., Atkinson, P. M., Bhatt, S., Mappin, B., Hay, S. I., & Gething, P. W. (2014). An effective approach for gap-filling continental scale remotely sensed time-series. *ISPRS Journal of Photogrammetry and Remote Sensing*, *98*, 106–118. https://doi.org/10.1016/j.isprsjprs.2014.10.001

Weng, Q. (2009). Thermal infrared remote sensing for urban climate and environmental studies: Methods, applications, and trends. *ISPRS Journal of Photogrammetry and Remote Sensing*, *64*(4), 335–344. https://doi.org/10.1016/j.isprsjprs.2009.03.007

Weng, Q., & Fu, P. (2014). Modeling diurnal land temperature cycles over Los Angeles using downscaled GOES imagery. In *ISPRS Journal of Photogrammetry and Remote Sensing* (Vol. 97, pp. 78–88). https://doi.org/10.1016/j.isprsjprs.2014.08.009

Weng, Q., Fu, P., & Gao, F. (2014). Generating daily land surface temperature at Landsat resolution by fusing Landsat and MODIS data. *Remote Sensing of Environment*, *145*, 55–67. https://doi.org/10.1016/j.rse.2014.02.003

Willmott, C. J. (1982). Some comments on the evaluation of model performance. *Bulletin of the American Meteorological Society*, *63*(11), 1309–1313. https://doi.org/10.1175/1520-0477(1982)063<1309:SCOTEO>2.0.CO;2

Xu, J., Zhang, F., Jiang, H., Hu, H., Zhong, K., Jing, W., Yang, J., & Jia, B. (2020). Downscaling ASTER land surface temperature over urban areas with machine learning-based area-to-point regression kriging. *Remote Sensing*, *12*(7). https://doi.org/10.3390/rs12071082

Yao, R., Wang, L., Huang, X., Sun, L., Chen, R., Wu, X., Zhang, W., & Niu, Z. (2021). A Robust Method for Filling the Gaps in MODIS and VIIRS Land Surface Temperature Data. *IEEE Transactions on Geoscience and Remote Sensing*, *59*(12), 10738–10752. https://doi.org/10.1109/TGRS.2021.3053284

Zhang, T., Zhou, Y., Zhu, Z., Li, X., & Asrar, G. R. (2022). A global seamless 1 km resolution daily land surface temperature dataset (2003–2020). *Earth System Science Data*, *14*(2), 651–664. https://doi.org/10.5194/essd-14-651-2022

Zhao, W., & Duan, S.-B. (2020). Reconstruction of daytime land surface temperatures under cloud-covered conditions using integrated MODIS/Terra land products and MSG geostationary satellite data. *Remote Sensing of Environment*, *247*, 111931. https://doi.org/10.1016/j.rse.2020.111931

Zhou, B., Kaplan, S., Peeters, A., Kloog, I., & Erell, E. (2020). "Surface," "satellite" or "simulation": Mapping intra-urban microclimate variability in a desert city. *International Journal of Climatology*, *40*(6), 3099–3117. https://doi.org/10.1002/ joc.6385

Zhou, D., Xiao, J., Bonafoni, S., Berger, C., Deilami, K., Zhou, Y., Frolking, S., Yao, R., Qiao, Z., & Sobrino, J. A. (2019). Satellite remote sensing of surface urban heat islands: Progress, challenges, and perspectives. *Remote Sensing*, *11*(1), 1–36. https://doi.org/10.3390/rs11010048

SUMMARY IN ESTONIAN

Erinevas mõõtkavas soojusandmete genereerimine linnade ülekuumenemise hindamiseks kasutades kaugseiret, masinõpet ja kiirgusvoo modelleerimist

Inimtegevus, näiteks fossiilkütuste põletamine ja industrialiseerimine, on põhjustanud maakera kliima soojenemise. Selle nähtuse üheks väikesemahuliseks ilminguks on linna soojussaare (Urban Heat Island, UHI) efekti tugevnemine. UHI on olukord, kus linnas on õhutemperatuur kõrgem kui linna ümbritsevatel aladel, mida põhjustab mitmesugune inimtegevus. Kuumalainete ajal on linnas UHI piirkonnas õhutemperatuur eriti kõrge. Viimasel ajal on teadlased hakanud kasutama mõistet linnade ülekuumenemine, mis kirjeldab olukorda, kus kuumalaines tõuseb linna temperatuur tasemeni, kus see ohustab linnaelanike tervist ja eluolu. Näiteks registreeriti 2003. aasta suvise kuumalaine ajal üle 30 000 kuumusega seotud surmajuhtumi Euroopas. Praegused kliimaprognoosid näitavad, et kuumalained sagenevad, on pikemad ja intensiivsemad. Seetõttu on hädavajalik valmistuda sellisteks sündmusteks ja muuta linna, et vähendada ülekuumenemise negatiivset mõju.

Enamiku linnavalitsuste suvise kuuma leevendamise plaanide elluviimist takistab ruumiliste soojusandmete puudumine. Doktoritöö keskendub sellele, kuidas kaugseire andmeid, masinõpet ja kiirgusvoo modelleerimist kasutada temperatuuri andmekogu koostamiseks, et hinnata linnade ülekuumenemist ja teha leevendusplaane. Sealjuures käsitleti kahte linnakliima mastaapi: kohalik mastaap, mis koosneb elamukvartalitest, tänavavõrgustikust või homogeensest linnaosast, mida planeerijad kujutavad maakasutusklassina, ja mikroskaala, mille moodustavad tänav ja sellega külgnevad hooned.

Maapinna temperatuur (LST), mis mõõdetakse Maa ümber tiirlevate satelliitide kaugandurite abil, on laialt kasutatud linnade ülekuumenemise hindamisel. Seda eelistatakse ilmajaamade andmetele, sest satelliidipiltidel on lisaväärtus ruumilise katvuse näol. Vaatamata sellele eelisele takistavad satelliidi LST kasutamist pilved, mis varjavad anduri vaadet Maale, põhjustades andmetes lünki. Teine probleem on seotud kaugseire kujutiste ruumilise (väikseim kauganduriga tajutav ala) ja ajalise (kujutist on samas kohas saadavuse sagedus) lahutusvõimega. Kasutades kaugseire andmeid tuleb leida kompromiss ruumilise ja ajalise lahutusvõime vahel, et mõista linnade ülekuumenemist ja soojasaasrte efekti. Kuumalained esinevad lühikese aja jooksul, mille jälgimiseks on pidevad ja sujuvad LST andmed hädavajalikud. Väitekiri käsitleb seda probleemi, kasutades masinõppe otsustusmetsa algoritmi (Random Forest, RF), et täita lünki satelliidi MODIS LST kujutistes Eesti alal (artikkel I) ja koostada linna taimestikuindeksi NDVI ning LST andmestikku lahutusvõimega 500 m ja 100 m kaart (artikkel II).

Artiklis I on esitatud otsustusmetsa mudel, mis põhineb LST ning ruumiliste ja ajaliste tegurite vahelisel empiirilisel seosel, võimaldades täita MODISe LST

lünki kolme kuumalaine jooksul. Mudeli keskmine ruutviga (RMSE) oli treenimis- ja testandmetel vastavalt 0,9 °C ja 1,3 °C. Võrreldes lünkade täitmise algoritmiga gapfill toimis artiklis I RF-mudel paremini. Mudeli RMSE oli 1,5 °C, samas kui gapfill algoritmi korral oli see 2,6 °C.

Artikli II eesmärk oli genereerida LST 500 m ja 100 m lahutusvõime kaart, millega Statistikaamet kaardistab linnade rahvastikku. Artikli I tulemuste põhjal märgiti, et NDVI ennustab LSTd. LST andmete saamiseks 500 m ja 100 m lahutusega töötati välja kaheosaline lähenemine. Esimene keskendus NDVI genereerimisele 500 m ja 100 m eraldusvõimele, mida seejärel kasutati teises osas ennustamistunnusena LST arvutamiseks samadel lahutusvõimetel. Loodi kaks otsustusmetsa mudelit, et ennustada MODISe ja Landsati NDVI-d, kasutades ennustajatena maakasutust ja maakatte fraktsioone. Kui MODISe ja Landsati NDVI-de täpsus oli piisav, peeti keskmisi ennustusi paremaks ja seetõttu kasutati neid LST ennustamisel. NDVI RMSE-d olid lahutusvõime 500 m ja 100 m puhul vastavalt 0,09 ja 0,14. LST ennustamiseks kasutati sama lähenemist, mida artiklis I. Sünteetiliste LST väärtuste RMSE-d olid mõlema võrgulahutuse korral 4,3 °C. Kuid asustatud aladel jääb prognoositava LST RMSE vahemikku ± 3 °C. Mudel ei suutnud täpselt ennustada enamiku taimkattega pindade LST-d, eriti metsade ja põllumaade. Seetõttu tuleb piirangut eelnevalt arvestada, kui kasutada seda metoodikat taimestunud pindade analüüsimisel.

Artiklid I ja II püüdsid selgitada, kuidas masinõppemudelid „mõtlevad“, et jõuda tulemusteni, hinnates olulisi omadusi ning uurides seoseid ennustajate ja sihtmuutujate vahel. Sihtmärgi ja ennustaja suhete väljendamiseks kasutati osalise sõltuvuse diagramme (Partial Dependence Plots, PDP). Planeerijad saavad soojuse leevendamise plaanides toetuda PDP-dele kui ühele paljudest teabeallikaist.

Kui LST sobib linna ülekuumenemise hindamiseks, siiski ei anna see palju infotinimese kokkupuutest väliskeskkonna või tänavatasandi soojusega. Artiklis III uuriti, kuidas keskmist kiirgustemperatuuri (Mean Radiant Temperature, TMRT), inimkeha kogusoojuskoormust, saab modelleerida suurtes linnades suure ruumilise eraldusvõimega. Lisaks uuriti artiklis, kuidas saab kvantifitseerida kõnnitee varju katvust jalakäijatele sobiva varjuarvutamisel. Puude või muude objektide vari on leevendab kuuma keskkonda. Artiklis III uuriti seda probleemi, genereerides Phoenixi suurlinnapiirkonnas LiDAR-i punktipilvedest 1-meetrise eraldusvõimega varju ja TMRT kaardid, kasutades päikese pikalainelise kiirgustiheduse geomeetria (SOlar LongWave Environmental Irradiance Geometry, SOLWEIG) mudelit, et aidata omavalitsusi „jahedate koridoride“ planeerimisel. TMRT hinnanguid valideeriti 763 vaatlusega, kasutades mobiilset inimese biometeoroloogilist 6-suunalist mõõteaparatuuri. Valideerimine näitas, et SOLWEIGi üldine RMSE oli 5,6 °C, kusjuures see oli 6,2 °C avatud aladel, 5,4 °C puude all ja 4,4 °C majadevahelistes tänavakanjonites. Igatunnised TMRT- ja varjukaardid ajavahemikul kell 7.00–20.00 koostati 27. juuniks 2012, mis oli tüüpiline selge, kuiv ja tuulevaikne suvepäev. Kõrgeim TMRT oli 83 °C kell 16.00 päikese käes avatud kohtades ja 41 °C varulistes kohtades. Kõnniteede päikese eest varjatust hinnati vastava plaani alusel. See näitas, et ainult 8% kõnniteedest vastas soovi-

tatavale 20%-lisele minimaalsele varjutusele. Kriteeriumile vastas vähem kui 50% kõnniteedest, mis näitab, et linnapiirkond ei ole tippkuumuseajal inimesele talutav. Uuringu tulemused annavad teavet omavalitsusele jahedate koridoride planeerimisel, et optimeerida kuumuse leevendamist.

Väitekiri käsitleb probleeme, mis on seotud ruumiliste temperatuuri andmete hankimisega, et välja töötada kohalikke kuumuse leevendamise strateegiaid linnas. Koos sotsiaalmajanduslike ja demograafiliste andmetega annavad kõik uurimuses genereeritud andmestikud teavet soojuse tegevuskavade koostamiseks. Tulevased tööd peaksid uurima nähtusi, mis võivad parandada LST-mudeli ennustustäpsust taimkattega aladel. Valdkonna edasised uuringud peaksid keskenduma jalakäijate liikumisharjumustele ja kuumaga kokkupuutele, mis tuleneb TMRT täpsete andmete lisamisest reisisimulatsiooni mudelitesse.

ACKNOWLEDGEMENTS

My name is on the cover of this thesis, but it took a village to make that happen. Three elders in my household oversaw my activities and ensured I did the right things. I thank Dr. Valentina Sagris and Prof. Jaak Jaagus for their patience, guidance, and supervision. They provided the resources to make my studies successful. Halfway into my doctoral studies, I encountered Dr. Ariane Middel of the SHaDE lab at Arizona State University (ASU), who hosted me for a year in her lab. The lessons learned, and the mentorship I received from working with Dr. Middel cannot go unappreciated; thank you, Dr. Middel, for all you did during and after my time at ASU.

In the village are elders outside my household and peers who had a hand in the success of my studies. Although I had no supervision from him, Prof. Ülo Mander played a vital role in my studies by putting me on his grant, and I appreciate that. Prof Evelyn Uuemaa and Dr. Alexander Kmoch made valuable contributions to my studies, and I am thankful. Dr. kmoch inspired and guided me to prepare my first workshop, gave me constructive feedback on my seminar presentations, and reviewed the technical aspects of my machine learning projects. Prof. Uuemaa reviewed my manuscripts before submission, helped me with the workshops I made, and invited me to several social events she organized for her lab, even though I was not a part of her team. In four years, some people became good friends and helped me stay on track with my schoolwork. To Bruno Montibeller, Desalew Meseret Moges, Holger Virro, Iuliia Burdun, and Oleksandr Karasov, thanks for your help, discussions, and positive energy. Even when you graduated, and I was the last of the pack in the program, you kept in touch.

My village has a great community and support system. When the demands of schooling and academic stress had the best of me, these people ensured I experienced love and care. I am grateful for these families who welcomed me into their homes, helped me survive as an international student, and cared for me during the holidays when I could not visit my biological family. To the Ackleys, Faulkners, Lays, Mandres, Moseses, and Turners, I appreciate everything you did for me throughout my time as a student in Estonia and the USA. I am thankful for the group of Ghanaian folks in Tartu who have been supportive throughout. I thank my parents and siblings for their support, understanding, and encouragement. Despite the great distance between us, my family always reached out to me to ensure I was doing well, empathized when I discussed the low points and celebrated the little achievements with me.

I want to thank all my friends who kept checking on me, asking the most annoying question: “When are you finishing?”. It reminded me that there was a job to finish. I can mention the names of friends and family who have encouraged me, but that would require a few more pages. I appreciate everything you have done for me.

Finally, and most importantly, to the one who makes the words of *Jeremiah 29:11* true, I am grateful for this one, Lord.

PUBLICATIONS

CURRICULUM VITAE

Name Isaac Newton Kwasi Buo
Date of birth 30.09.1990
E-mail isaac.buo@ut.ee

Education

2019–present	PhD in Geoinformatics, University of Tartu, Estonia
2017–2019	MSc in Geoinformatics for Urbanized Societies, University of Tartu, Estonia
2009–2013	BSc in Geomatic Engineering, University of Mines and Technology, Ghana

Work experience

2021–2023	Junior researcher, University of Tartu, Estonia
2015–2017	Geomatic engineer, Joeamah Geomatics Consult
2013–2014	Teaching assistant, University of Mines and Technology

Awards

2023	International Congress of Biometeorology Travel fund
2021–2022	Dora Plus mobility grant for doctoral studies
2021	Dora Plus short-term mobility grant

Supplementary Education

2022	Bochum Urban Climate Summer School, Ruhr-University Bochum, Germany
2021–2022	Visiting Research Scholar, SHaDE Lab, Arizona State University, USA
2021	Discrete Anisotropic Radiative Transfer Model Training (DART), Paul Sabatier University, Toulouse, France
2018–2019	Exchange student, University of Salzburg, Austria

Publications

Buo, Isaac; Sagris, Valentina; Jaagus, Jaak; Middel, Ariane (2023). High-resolution thermal exposure and shade maps for cool corridor planning. Sustainable Cities and Society, 104499. DOI: 10.1016/j.scs.2023.104499.

Li, Rui; Chester, Mikhail V.; Middel, Ariane; Vanos, Jennifer K.; Hernandez-Cortes, Danae; **Buo, Isaac**; Hondula, David M. (2023). Effectiveness of travel behavior and infrastructure change to mitigate heat exposure. Frontiers in Sustainable Cities, 5. DOI: 10.3389/frsc.2023.1129388.

Buo, Isaac; Sagris, Valentina; Jaagus, Jaak (2022). Generating synthetic daily remote sensing products suitable for surface heat island and heatwaves assessments at urban scale. International Journal of Environmental Science and Technology. DOI: 10.1007/s13762-022-04510-3.

Buo, Isaac; Sagris, Valentina; Jaagus, Jaak (2022). Gap-Filling Satellite Land Surface Temperature Over Heatwave Periods with Machine Learning. IEEE Geoscience and Remote Sensing Letters, 19. DOI: 10.1109/LGRS.2021.3068069.

Buo, Isaac; Sagris, Valentina; Burdun, Iuliia; Uuemaa, Evelyn (2021). Estimating the expansion of urban areas and urban heat islands (UHI) in Ghana: a case study. Natural Hazards, 105 (2), 1299−1321. DOI: 10.1007/s11069-020-04355-4.

ELULOOKIRJELDUS

Nimi Isaac Newton Kwasi Buo
Sünniaeg 30.09.1990
E-mail isaac.buo@ut.ee

Haridustee

2019–present	Doktorikraad geoinformaatikas, Tartu Ülikool, Eesti
2017–2019	Magistrikraad geoinformaatikas linnastunud ühiskondade jaoks, Tartu Ülikool, Eesti
2009–2013	Bakalaureusekraad geomaatilises inseneriteaduses, Kaevandus- ja tehnoloogiaülikool, Ghana

Töökogemus

2021–2023	Nooremteadur, Tartu Ülikool, Eesti
2015–2017	Geomaatiline insener, Joeamah Geomaatika Konsult, Ghana
2013–2014	Õpetamisassistent, Kaevandus- ja tehnoloogiaülikool, Ghana

Auhinnad

2023	Rahvusvaheline Biomeetoroloogia Kongressi reisitoetus
2021–2022	Dora Plus liikuvustoetus doktoriõpinguteks
2021	Dora Plus lühiajaline liikuvustoetus

Täiendav haridus

2022	Bochumi linnakliima suvekool, Ruhri ülikool Bochum, Saksamaa
2021–2022	Külalisuurimisscholaar, SHaDE Lab, Arizona Osariigi Ülikool, USA
2021	Diskreetse anisotroopse kiirguse ülekande mudeli koolitus (DART), Paul Sabatieri ülikool, Toulouse, Prantsusmaa
2018–2019	Vahetusüliõpilane, Salzburgi Ülikool, Austria

Publikatsioonid

Buo, Isaac; Sagris, Valentina; Jaagus, Jaak; Middel, Ariane (2023). High-resolution thermal exposure and shade maps for cool corridor planning. Sustainable Cities and Society, 104499. DOI: 10.1016/j.scs.2023.104499.

Li, Rui; Chester, Mikhail V.; Middel, Ariane; Vanos, Jennifer K.; Hernandez-Cortes, Danae; **Buo, Isaac**; Hondula, David M. (2023). Effectiveness of travel behavior and infrastructure change to mitigate heat exposure. Frontiers in Sustainable Cities, 5. DOI: 10.3389/frsc.2023.1129388.

Buo, Isaac; Sagris, Valentina; Jaagus, Jaak (2022). Generating synthetic daily remote sensing products suitable for surface heat island and heatwaves assessments at urban scale. International Journal of Environmental Science and Technology. DOI: 10.1007/s13762-022-04510-3.

Buo, Isaac; Sagris, Valentina; Jaagus, Jaak (2022). Gap-Filling Satellite Land Surface Temperature Over Heatwave Periods with Machine Learning. IEEE Geoscience and Remote Sensing Letters, 19. DOI: 10.1109/LGRS.2021.3068069.

Buo, Isaac; Sagris, Valentina; Burdun, Iuliia; Uuemaa, Evelyn (2021). Estimating the expansion of urban areas and urban heat islands (UHI) in Ghana: a case study. Natural Hazards, 105 (2), 1299−1321. DOI: 10.1007/s11069-020-04355-4.

DISSERTATIONES GEOGRAPHICAE UNIVERSITATIS TARTUENSIS

1. **Вийви Руссак.** Солнечная радиация в Тыравере. Тарту, 1991.
2. **Urmas Peterson.** Studies on Reflectance Factor Dynamics of Forest Communities in Estonia. Tartu, 1993.
3. **Ülo Suursaar.** Soome lahe avaosa ja Eesti rannikumere vee kvaliteedi analüüs. Tartu, 1993.
4. **Kiira Aaviksoo.** Application of Markov Models in Investigation of Vegetation and Land Use Dynamics in Estonian Mire Landscapes. Tartu, 1993.
5. **Kjell Weppling.** On the assessment of feasible liming strategies for acid sulphate waters in Finland. Tartu, 1997.
6. **Hannes Palang.** Landscape changes in Estonia: the past and the future. Tartu, 1998.
7. **Eiki Berg.** Estonia's northeastern periphery in politics: socio-economic and ethnic dimensions. Tartu, 1999.
8. **Valdo Kuusemets.** Nitrogen and phosphorus transformation in riparian buffer zones of agricultural landscapes in Estonia. Tartu, 1999.
9. **Kalev Sepp.** The methodology and applications of agricultural landscape monitoring in Estonia. Tartu, 1999.
10. **Rein Ahas.** Spatial and temporal variability of phenological phases in Estonia. Tartu, 1999.
11. **Эркки Таммиксаар.** Географические аспекты творчества Карла Бэра в 1830–1840 гг. Тарту, 2000.
12. **Garri Raagmaa.** Regional identity and public leaders in regional economic development. Tartu, 2000.
13. **Tiit Tammaru.** Linnastumine ja linnade kasv Eestis nõukogude aastatel. Tartu, 2001.
14. **Tõnu Mauring.** Wastewater treatment weltlands in Estonia: efficiency and landscape analysis. Tartu, 2001.
15. **Ain Kull.** Impact of weather and climatic fluctuations on nutrient flows in rural catchments. Tartu, 2001.
16. **Robert Szava-Kovats.** Assessment of stream sediment contamination by median sum of weighted residuals regression. Tartu, 2001.
17. **Heno Sarv.** Indigenous Europeans east of Moscow. Population and Migration Patterns of the Largest Finno-Ugrian Peoples in Russia from the 18th to the 20th Centuries. Tartu, 2002.
18. **Mart Külvik.** Ecological networks in Estonia — concepts and applications. Tartu, 2002.
19. **Arvo Järvet.** Influence of hydrological factors and human impact on the ecological state of shallow Lake Võrtsjärv in Estonia. Tartu, 2004.
20. **Katrin Pajuste.** Deposition and transformation of air pollutants in coniferous forests. Tartu, 2004.

21. **Helen Sooväli.** *Saaremaa waltz*. Landscape imagery of Saaremaa Island in the 20th century. Tartu, 2004.
22. **Antti Roose.** Optimisation of environmental monitoring network by integrated modelling strategy with geographic information system — an Estonian case. Tartu, 2005.
23. **Anto Aasa.** Changes in phenological time series in Estonia and Central and Eastern Europe 1951–1998. Relationships with air temperature and atmospheric circulation. Tartu, 2005.
24. **Anneli Palo.** Relationships between landscape factors and vegetation site types: case study from Saare county, Estonia. Tartu, 2005.
25. **Mait Sepp.** Influence of atmospheric circulation on environmental variables in Estonia. Tartu, 2005.
26. **Helen Alumäe.** Landscape preferences of local people: considerations for landscape planning in rural areas of Estonia. Tartu, 2006.
27. **Aarne Luud.** Evaluation of moose habitats and forest reclamation in Estonian oil shale mining areas. Tartu, 2006.
28. **Taavi Pae.** Formation of cultural traits in Estonia resulting from historical administrative division. Tartu, 2006.
29. **Anneli Kährik.** Socio-spatial residential segregation in post-socialist cities: the case of Tallinn, Estonia. Tartu, 2006.
30. **Dago Antov.** Road user perception towards road safety in Estonia. Tartu, 2006.
31. **Üllas Ehrlich.** Ecological economics as a tool for resource based nature conservation management in Estonia. Tartu, 2007.
32. **Evelyn Uuemaa.** Indicatory value of landscape metrics for river water quality and landscape pattern. Tartu, 2007.
33. **Raivo Aunap.** The applicability of gis data in detecting and representing changes in landscape: three case studies in Estonia. Tartu, 2007.
34. **Kai Treier.** Trends of air pollutants in precipitation at Estonian monitoring stations. Tartu, 2008.
35. **Kadri Leetmaa.** Residential suburbanisation in the Tallinn metropolitan area. Tartu, 2008.
36. **Mare Remm.** Geographic aspects of enterobiasis in Estonia. Tartu, 2009.
37. **Alar Teemusk.** Temperature and water regime, and runoff water quality of planted roofs. Tartu, 2009.
38. **Kai Kimmel.** Ecosystem services of Estonian wetlands. Tartu, 2009.
39. **Merje Lesta.** Evaluation of regulation functions of rural landscapes for the optimal siting of treatment wetlands and mitigation of greenhouse gas emissions. Tartu, 2009.
40. **Siiri Silm.** The seasonality of social phenomena in Estonia: the location of the population, alcohol consumption and births. Tartu, 2009.
41. **Ene Indermitte.** Exposure to fluorides in drinking water and dental fluorosis risk among the population of Estonia. Tartu, 2010.
42. **Kaido Soosaar.** Greenhouse gas fluxes in rural landscapes of Estonia. Tartu, 2010.

43. **Jaan Pärn.** Landscape factors in material transport from rural catchments in Estonia. Tartu, 2010.
44. **Triin Saue.** Simulated potato crop yield as an indicator of climate variability in Estonia. Tartu, 2011.
45. **Katrin Rosenvald.** Factors affecting EcM roots and rhizosphere in silver birch stands. Tartu, 2011.
46. **Ülle Marksoo.** Long-term unemployment and its regional disparities in Estonia. Tartu, 2011, 163 p.
47. **Hando Hain.** The role of voluntary certification in promoting sustainable natural resource use in transitional economies. Tartu, 2012, 180 p.
48. **Jüri-Ott Salm.** Emission of greenhouse gases CO_2, CH_4, and N_2O from Estonian transitional fens and ombrotrophic bogs: the impact of different land-use practices. Tartu, 2012, 125 p.
49. **Valentina Sagris.** Land Parcel Identification System conceptual model: development of geoinfo community conceptual model. Tartu, 2013, 161 p.
50. **Kristina Sohar.** Oak dendrochronology and climatic signal in Finland and the Baltic States. Tartu, 2013, 129 p.
51. **Riho Marja.** The relationships between farmland birds, land use and landscape structure in Northern Europe. Tartu, 2013, 134 p.
52. **Olle Järv.** Mobile phone based data in human travel behaviour studies: New insights from a longitudinal perspective. Tartu, 2013, 168 p.
53. **Sven-Erik Enno.** Thunderstorm and lightning climatology in the Baltic countries and in northern Europe. Tartu, 2014, 142 p.
54. **Kaupo Mändla.** Southern cyclones in northern Europe and their influence on climate variability. Tartu, 2014, 142 p.
55. **Riina Vaht.** The impact of oil shale mine water on hydrological pathways and regime in northeast Estonia. Tartu, 2014, 111 p.
56. **Jaanus Veemaa.** Reconsidering geography and power: policy ensembles, spatial knowledge, and the quest for consistent imagination. Tartu, 2014, 163 p.
57. **Kristi Anniste.** East-West migration in Europe: The case of Estonia after regaining independence. Tartu, 2014, 151 p.
58. **Piret Pungas-Kohv.** Between maintaining and sustaining heritage in landscape: The examples of Estonian mires and village swings. Tartu, 2015, 210 p.
59. **Mart Reimann.** Formation and assessment of landscape recreational values. Tartu, 2015, 127 p.
60. **Järvi Järveoja.** Fluxes of the greenhouse gases CO_2, CH_4 and N_2O from abandoned peat extraction areas: Impact of bioenergy crop cultivation and peatland restoration. Tartu, 2015, 171 p.
61. **Raili Torga.** The effects of elevated humidity, extreme weather conditions and clear-cut on greenhouse gas emissions in fast growing deciduous forests. Tartu, 2016, 128 p.
62. **Mari Nuga.** Soviet-era summerhouses On homes and planning in post-socialist suburbia. Tartu, 2016, 179 p.

63. **Age Poom.** Spatial aspects of the environmental load of consumption and mobility. Tartu, 2017, 141 p.
64. **Merle Muru.** GIS-based palaeogeographical reconstructions of the Baltic Sea shores in Estonia and adjoining areas during the Stone Age. Tartu, 2017, 132 p.
65. **Ülle Napa.** Heavy metals in Estonian coniferous forests. Tartu, 2017, 129 p.
66. **Liisi Jakobson.** Mutual effects of wind speed, air temperature and sea ice concentration in the Arctic and their teleconnections with climate variability in the eastern Baltic Sea region. Tartu, 2018, 118 p.
67. **Tanel Tamm.** Use of local statistics in remote sensing of grasslands and forests. Tartu, 2018, 106 p.
68. **Enel Pungas.** Differences in Migration Intentions by Ethnicity and Education: The Case of Estonia. Tartu, 2018, 142 p.
69. **Kadi Mägi.** Ethnic residential segregation and integration of the Russian-speaking population in Estonia. Tartu, 2018, 173 p.
70. **Kiira Mõisja.** Thematic accuracy and completeness of topographic maps. Tartu, 2018, 112 p.
71. **Kristiina Kukk.** Understanding the vicious circle of segregation: The role of leisure time activities. Tartu, 2019, 143 p.
72. **Kaie Kriiska.** Variation in annual carbon fluxes affecting the soil organic carbon pool and the dynamics of decomposition in hemiboreal coniferous forests. Tartu, 2019, 146 p.
73. **Pille Metspalu.** The changing role of the planner. Implications of creative pragmatism in Estonian spatial planning. Tartu, 2019, 128 p.
74. **Janika Raun.** Mobile positioning data for tourism destination studies and statistics. Tartu, 2020, 153 p.
75. **Birgit Viru.** Snow cover dynamics and its impact on greenhouse gas fluxes in drained peatlands in Estonia. Tartu, 2020, 123 p.
76. **Iuliia Burdun.** Improving groundwater table monitoring for Northern Hemisphere peatlands using optical and thermal satellite data. Tartu, 2020, 162 p.
77. **Ingmar Pastak.** Gentrification and displacement of long-term residents in post-industrial neighbourhoods of Tallinn. Tartu, 2021, 141 p.
78. **Veronika Mooses.** Towards a more comprehensive understanding of ethnic segregation: activity space and the vicious circle of segregation. Tartu, 2021, 161 p.
79. **Johanna Pirrus.** Contemporary Urban Policies and Planning Measures in Socialist-Era Large Housing Estates. Tartu, 2021, 142 p.
80. **Gert Veber.** Greenhouse gas fluxes in natural and drained peatlands: spatial and temporal dynamics. Tartu, 2021, 210 p.
81. **Anniki Puura.** Relationships between personal social networks and spatial mobility with mobile phone data. Tartu, 2021, 144 p.
82. **Alisa Krasnova.** Greenhouse gas fluxes in hemiboreal forest ecosystems. Tartu, 2022, 185 p.

83. **Tauri Tampuu.** Synthetic Aperture Radar Interferometry as a tool for monitoring the dynamics of peatland surface. Tartu, 2022, 166 p.
84. **Najmeh Mozaffaree Pour.** Urban Expansion in Estonia: Monitoring, Analysis, and Modeling. Tartu, 2022, 169 p.
85. **Bruno Montibeller**. Evaluating human-induced forest degradation in different biomes using spatial analysis of satellite-derived data. Tartu, 2022, 112 p.
86. **Holger Virro**. Geospatial data harmonization and machine learning for large-scale water quality modelling. Tartu, 2022, 138 p.
87. **Azadeh Rezapour**. The impact of climate change on fine root trait responses of deciduous and coniferous trees. Tartu, 2023, 108 p.