Optimizing a Sensor Network's Granularity to Mitigate Urban Heat Island Effect at 2032 Brisbane Olympics

Celine Schoe
Leibniz University Hannover
schoe@iwi.uni-hannover.de

Matthias Tuczek
Leibniz University Hannover
tuczek@iwi.uni-hannover.de

Kenan Degirmenci
Queensland University of Technology
kenan.degirmenci@qut.edu.au

Michael H. Breitner
Leibniz University Hannover
breitner@iwi.uni-hannover.de

Abstract

In anticipation of the 2032 climate-positive Olympic Games in Brisbane, we address the Urban Heat Island (UHI) effect optimizing the granularity of a sensor network in the Northshore Hamilton Priority Development Area (PDA), i.e., the location of the Athletes' Village, for efficient environmental monitoring and the provision of a Green Information System (IS). We use spatio-temporal sensor data and leverage advanced interpolation techniques to optimize both, temporal and spatial granularity settings. Results and findings from our granularity analysis reveal an optimal temporal granularity at one-hour intervals, providing the optimal trade-off balance between computational efficiency and sufficient detail for urban planning. Finer temporal resolutions do not significantly enhance prediction accuracy. Spatial analysis further helps decision makers to balance trade-offs between economic costs and prediction accuracy, eliminating unnecessary sensors in the network.

Keywords: Sensor Network, Information Granularity, Urban Heat Island Effect, Olympic Games, Spatio-temporal Analysis

1. Introduction

In July 2021 Brisbane was selected as host for the 2032 Olympic Games (IOC, 2021), while contractually obliged to deliver a climate-positive event placed in a broader framework of environmental awareness (IOC, 2020). Northshore Hamilton PDA is declared as the location for the Brisbane Athletes' Village focusing on renewable energy use, long-term urban planning and flexibility, where UHI mitigation strategies are of significant importance (Brennan, 2022). The UHI effect refers to higher temperatures experienced in urban areas compared to their surrounding rural regions, influencing energy efficiency goals and environmental accounting of major events like the Olympics (Oke, 1973; State of Queensland, 2022; Tuczek et al., 2022). To solve such problems of global warming, Green IS can help by providing information that enables and motivates economic and behaviorally driven solutions, by collection and analysis of specific data. The starting point for this analysis in the context of Green IS is the Integrated Framework proposed by Watson et al. (2010), integrating three types of technology into a single system. The *flow network*, consisting of interconnected transport components facilitating the transfer of continuous substances or discrete entities, the *sensor network*, that is recording data, and the *sensitized object*, that is owned by a particular consumer. The IS ties the elements together and gives a complete solution for the specific issue. A major problem for future research, named by Watson et al. (2010) is the choice of the optimum level of information granularity during the design stage of the sensor network. This includes finding the optimal level of detail and frequency of the recorded data from the sensors. Determining the optimal level of granularity leads to a trade-off process between costs and benefits, where both, the costs of a denser network, including the purchase and installation of sensors, and the computational costs must be evaluated against the benefits resulting from, e.g., improved prediction accuracy (Kools & Phillipson, 2016).

The sensor network in our work is located in Northshore Hamilton PDA in Brisbane, the location of the Athletes' Village. In the foundation of mitigating UHI effects and urban planning, a functioning and accurate sensor network that records micro-climate variables in a specific area can be extremely useful

URI: https://hdl.handle.net/10125/108954
978-0-9981331-8-8

for environmental decision support (Deilami et al., 2018). The recorded variables can provide a detailed picture of the microclimate and localization of heat islands (Ho et al., 2016). Additionally, we also aim to investigate whether heat islands can occur within smaller areas and not just in the typical urban-rural comparisons. This precise understanding of the microclimate enables urban planners to take measures to adapt the urban environment to changing climatic conditions and improve the well-being of residents (Watkins et al., 2007). These include, e.g., the placement of green spaces, trees, green roofs and other green infrastructure elements to reduce heat stress (Gunawardena et al., 2017). Therefore, recorded data from the sensors in Northshore Hamilton PDA can be used for building Green IS for the strategic integration of actions for considering the UHI effect and for achieving energy efficiency goals. We address this practical challenge and further contribute to IS research (Watson et al., 2010) by addressing the following research question:

What is the optimal granularity of sensor networks to balance cost and accuracy in modeling microclimate conditions for mitigating the UHI effects?

Initially, granularity concepts in the literature are identified and analyzed and an introduction to the factors influencing the development of UHI is given. Based on this, the granularity levels relevant to this work are defined and the available sensor network is analyzed concerning the UHI influencing factors. A model for temperature modeling is developed and tested with different granularity combinations. After that we discuss the results, implications, and recommendations of the model and conclude by providing the limitations of the study, proposing further research and giving a conclusion.

2. Background

2.1. Literature Search

The literature review is based on the well-founded methods proposed by Vom Brocke et al. (2015) and Webster and Watson (2002). We extended the established procedure by modern AI-based literature search tools and graphical illustrations (Watson & Webster, 2020). Its objective is to underscore the particular types of granularity emphasized in the literature, along with the underlying concepts and methodologies in building distinct granularities. To initiate the literature search, the following selection of research databases are used as a first fundamental: ACM Digital Library, AISeL, EBSCOhost Business Source Elite, IEEE Xplore Digital Library, INFORMS PubsOnline, ScienceDirect, Wiley Online Library, SpringerLink and Taylor & Francis Online. The keywords and operators used in the database search where *data granularity* OR *information granularity*. Duplicates are removed and a backward and forward search is performed. To narrow down the selection of papers to just high-quality papers, any articles not published in Scimago Q1-listed journals are excluded. To perform an intelligent and efficient literature search, an additional selection of papers is then made using the AI-based graphical tool Connected Papers as well as AI-based search tool Consensus. The research procedure ends up in a final sample of 102 papers.

2.2. Granularity Concepts

With the identified database of papers, the concept-centric approach is applied and four different granularity types, three different frameworks, gradations and properties of the granulation processes can be derived (Figure 1).

Granularity type	Temporal granularity	Spatial granularity	Object granularity	Theoretical granularity
Frameworks	Interval set	Fuzzy set	Rough set	
Concept Gradation	Linguistic granularity		Hierarchical granularity	
Properties	Numerical		Non-numerical	

Figure 1: Granularity concepts

Among the different types, **Temporal granularity** refers to the level of detail or resolution at which time-related data is collected and analyzed, including aggregation of data or different choices of temporal resolution (Al-Hmouz et al., 2015; Khan et al., 2022). Different time resolutions can differ, e.g., between daily, hourly or every few seconds. **Spatial granularity** describes the level of detail or resolution at which spatial or geographic data is analyzed, e.g., by dividing spatial data into sub-regions for tasks like analyzing electricity distribution in more detail (Bargiela & Pedrycz, 2003; Knirsch et al., 2016). Furthermore, the **Object granularity** type deals with how individual objects or entities are categorized based on common attributes, such as analyzing energy use by individual appliances or groups of appliances instead of measuring the consumption of the entire house (Eibl & Engel, 2014). In contrast, the concept of information granules, as outlined by Song et al. (2023), does not fit a specific granularity type but rather embodies an approach of **Theoretical granularity**. They can be adjusted and refined according to the context or specific needs (Pedrycz, 2018). The information granules formed in this theoretical construct are aggregates of entities grouped by common traits or functions, signifying levels of abstraction. Key frameworks for forming these granules include interval sets, fuzzy sets, and rough sets (Song & Wang, 2016; Song et al., 2019).

Interval sets, as the most popular framework, represent data as ranges, characterized by upper and lower bounds, with granularity defined by the narrowness or broadness of these ranges (Gacek & Pedrycz, 2012; Song & Liu, 2021). **Fuzzy sets** in contrast permit partial membership, typically described through linear or triangular membership functions (Song & Wang, 2016). By differentiating between definite and potential memberships, **Rough sets** tackle the challenges of data vagueness using upper and lower approximations along with boundary regions (Y. Yao, 1998). Two additional concept gradations, **Linguistic granularity** and **Hierarchical granularity**, can be identified in the literature. Linguistic granularity interprets sets like interval or fuzzy sets using linguistic terms such as "slightly low", "low", and "high" to describe the information granule (Chen & Chen, 2015). The concept of hierarchical granularity involves the multi-layered granularity representation, forming a tree-like structure that represents data at multiple detail and abstraction levels (Huang & Li, 2018; M. X. Yao, 2019; Zhu et al., 2020). Finally, the categorization distinguishes between **Numerical** and **Non-numerical** properties of the results of post-granulation processes, whereby granular models produce either abstract, non-numerical outcome or provide numerical prediction results (Zhu et al., 2019).

After classifying the collected articles according to the various concepts discussed, we have gained a comprehensive understanding of how granularity is employed in the literature. Over 60% of these studies focus on the theoretical aspects of granularity and the formation of information granules, with the interval set being the predominant framework used, closely followed by fuzzy sets. Approximately one-third of the articles explore temporal granularity within the context of numerical time series, whereas spatial granularity is less frequently addressed. Moreover, the concurrent analysis of temporal and spatial granularity is uncommon in the existing literature, however, many processes and phenomena in the real world have both temporal and spatial dimensions, which are interlinked. For the consideration of the UHI effect and the associated analysis of the micro-climate, the joint consideration of temporal and spatial granularity is extremely important. Heat islands are spatially heterogeneous and can vary depending on local conditions and urban characteristics. Similarly, the intensity and extent of heat islands can change throughout the day or year. Understanding and mitigating urban heat pollution necessitates an optimal consideration of spatial and temporal granularity, which is why we focus on these two types in the following (Deilami et al., 2018).

2.3. Urban Heat Island Effect

First, a rough overview of the factors influencing the formation of UHI is to be given to better understand the spatial and temporal components for determining the optimal granularity. Building density and height are predominant influences of UHI effects, where denser urban areas increase surface radiation and heat storage within structures. This effect is particularly pronounced after sunset, contributing to elevated night-time temperatures (Ho et al., 2016; Steeneveld et al., 2014). Additionally, the design of narrow streets with tall buildings can create the so-called *Canyon effect*, intensifying temperatures within the enclosed space (Steeneveld et al., 2014). Moreover, the design of urban areas, including narrow streets flanked by tall buildings, not only traps heat but also significantly reduces wind flow, thus limiting the natural cooling effect (Watkins et al., 2007). Another significant influencing factor is the absence of vegetation. Watkins et al. (2007), Zhang et al. (2017) and Ghosh and Das (2018) name different types of vegetation including green areas, trees or green roofs that can minimize the UHI effect, creating a cool island effect in small areas, by converting some of the solar energy into latent heat energy through the evaporation of water. Beyond this, trees can intercept solar radiation high above street level. Further key factor for intensity of UHI is the absorption of solar radiation on roofs and road surfaces. Dark-colored surfaces with low albedo value absorb more solar radiation, thereby intensifying UHI effects (Watkins et al., 2007). Employing materials with high reflectivity, known for their high albedo values, can significantly counteract this heat accumulation (Watkins et al., 2007; Zhang et al., 2017). Water bodies add another layer of temperature regulation through their high heat capacity and the cooling effects of evaporation. These bodies not only absorb and store vast amounts of heat but also reflect a significant portion of incoming solar radiation, reducing energy absorption and subsequent heat generation (Ghosh & Das, 2018; Steeneveld et al., 2014). Thermal images at Northshore Hamilton PDA visually confirm these influencing factors, showing higher temperatures on asphalt compared to vegetated or open areas, with variations exceeding up to 30 $^{\circ}C$ Figure 2.

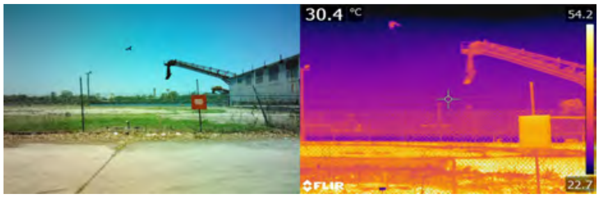

Figure 2: Thermal illustration

3. Spatio-temporal Sensor Network Analysis

The considered sensor network in Northshore Hamilton PDA consists of four sensor locations with sensors of the Atmos and Netvox type, while the four sites reflect different environmental characteristics that capture distinct named spatial influencing factors for the origin of the UHI (Figure 3). The Northshore Hamilton PDA is defined by the yellow boundary and the respective four sites are marked in red.

Figure 3: Sensor network - Northshore Hamilton PDA

The sensors record microclimate variables at 15-minute intervals, with measurements starting on November 8, 2019. For this analysis, sensor data only from the year 2020 is considered. For our analysis, it is crucial to identify differences in the locations with regard to the UHI effect and to investigate whether the distinct locations have different temperature profiles. This investigation is particularly important for the selection of the granularity parameters and the modeling of the temperature. The detailed illustrations of the four sensor locations can be found in Figure 4.

The sensors at Site 1 are strategically positioned at the corner of Theodore St & Cullen Avenue on a green area near a side road surrounded by trees. There are few buildings near the sensors. The greenery and minimal presence of nearby buildings suggest a potentially mitigating effect on the UHI phenomenon, as vegetation provides shade and evaporative cooling, thereby moderating local temperatures. Located atop a building within the car park at 92 Macarthur Avenue, Site 2 experiences conditions conducive to UHI intensification. The dark-colored roof of the structure contributes to a low albedo, facilitating the absorption and retention of solar radiation. Surrounded predominantly by asphalt and lacking significant green spaces, this site is susceptible to elevated temperatures, accentuated by the heat-retaining properties of urban surfaces. Sensors at Site 3 are located at Bincote Street next to Eat Street near the Brisbane River on gray asphalt ground. There are virtually no large green spaces or trees in the immediate vicinity except for the nearby Maritime Green Northshore, an event location with an integrated park. Despite the nearby Maritime Green Northshore, the limited greenery in the immediate vicinity provides minimal cooling effects. The predominance of impermeable surfaces suggests the potential for heightened UHI intensity due to reduced evaporative cooling and increased heat retention. The last sensors at Site 4 are located at the Pop-Up Park directly near the Brisbane River. There are also a few trees in the area. The site is surrounded by several buildings. The presence of trees within the park, the vicinity of the water and the nearby buildings indicates a mix of cooling and heating influences.

Figure 4: Sensor placement locations

Heatmaps for the year 2020 are created to compare the four locations in terms of their temperature patterns (Figure 5). It can be noted that Site 2 shows particular differences to the other sites. The site records comparatively colder months at the beginning and warmer months at the end of the year. No major differences between the other three locations can initially be identified. After the analysis on a daily basis, individual hours of the day are considered at the finest level of temporal granularity (Figure 6. In the first half of the year in particular, represented by cut-off dates in March, June, September and December (each representing one of the four seasons in Brisbane) it can be seen that Site 2 overall records lower temperatures

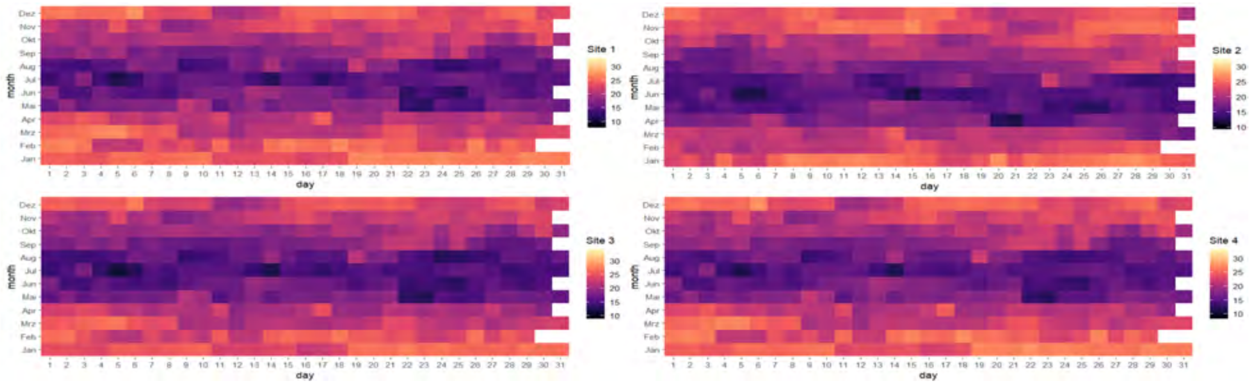

Figure 5: Heatmaps of temperature by sites

at day and notably at night than the other sites. In the months of the second half of the year, represented by cut-off dates in September and December, higher temperatures are especially recorded in the evening and early morning hours, which indicates a comparatively reduced cooling behavior at night. According to the explanations in the previous sections, absorbing surfaces release heat mainly three to five hours after sunset, which can be linked to these observations.

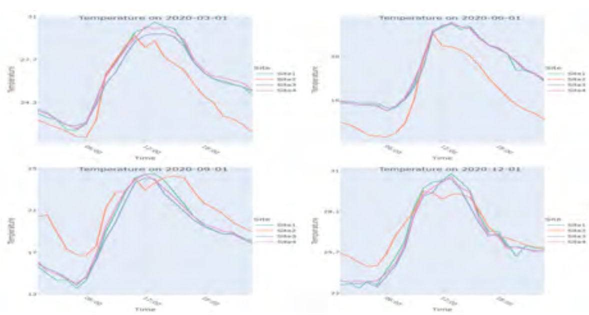

Figure 6: Hourly temperature trajectories

For statistical analysis of significant differences in locations, a t-test is performed separately for each hour of the dataset and for each season to identify significant differences in the temperature time series of the sites. For this purpose, the temperature differences for each site combination and each hour are calculated separately. It can be noted that Site 2 is statistically significantly different from the other temperature time series at all hours and seasons. Also, differences between the other sites have become more apparent. Site 3 and Site 4 show the same behavior over the night and different behavior over the day. Differences in temperature patterns of Sites 1 and 3 are particularly pronounced during the day, but after sunset until midnight, temperatures appear to behave similarly in summer and spring. Site 1 and Site 4 show almost the same temperature curves over the entire day. Just in winter and spring, the two sites differ slightly from each other.

4. Spatio-temporal Temperature Modeling

This section presents the methodology for modeling temperatures and optimizing the sensor network's information granularity. Common spatio-temporal interpolation methods used in the literature include Kriging and Inverse Distance Weighting (IDW). Recently, machine learning algorithms, such as the Random Forest (RF) algorithm, have become popular for spatio-temporal prediction due to their flexibility, lack of assumptions about data distribution and their ability to model nonlinear complex relationships (Hengl et al., 2018; Sekulić et al., 2020; Zhan et al., 2018). We want to compare a common methodology, like Kriging or IDW with a machine learning algorithm to evaluate the optimal granularity of the sensor network. To choose an appropriate spatio-temporal interpolation method, the relationship between spatial and temporal data is considered. The data is spatially sparse in comparison to the huge amount of temporal measurements. For this reason, the Kriging method cannot be used for temperature interpolation, as the spatial autocorrelation for variogram analysis cannot be adequately calculated (Rusu & Rusu, 2006). This is negligible for the IDW algorithm, which is why it is chosen as the common method. The two methodologies will now be briefly presented formally.

Let $(s_i, t_i), i = 1, 2..., n$ be a set of spatio-temporal locations, with s_i as a geographic location i and t_i a point in time i. It can be summarized in the domain $Z(s, t)$, which includes the repeated measurements at different points in time at multiple spatial locations (Montero et al., 2015).

Formally, the IDW can be expressed as

$$\hat{Z}(s_0, t_0) = \sum_{i=1}^{n} \lambda_i Z(s_i, t_i), \tag{1}$$

where $\hat{Z}(s_0, t_0)$ is the predicted value at location s_0 and at time t_0, n is the number of nearest known points that are included in the calculation and λ_i are the assigned weights to each of the known points $Z(s_i, t_i)$. The weights λ_i are calculated by a smoothing parameter p that influences the weighting of the predicted points (Nielsen, 2009).

As part of the RF approach, the algorithm is trained iteratively by generating several decision trees based on bootstrap samples. The final predictions are the average of the predictions of the individual trees based on

$$\hat{Z}^B(s_0, t_0) = \frac{1}{B} \sum_{b=1}^{B} t_b^*(s_i, t_i), \tag{2}$$

where b represents an individual bootstrap sample, B denotes the total number of trees, and t_b^* is the individual decision tree comprising pairs of values for the target variable and the covariates. The RF algorithm is performed using the computationally fast *ranger* command in R (Hengl et al., 2018).

For building covariates for both methods, the Northshore Hamilton PDA area is classified with the help of ArcGIS Pro, according to the factors mentioned influencing the formation of heat islands. The spatial factors are considered constant over the period under review. This was confirmed with the help of an image comparison of satellite images during the observation period. For this purpose, classes are identified by building polygons within the image file via pixel recognition, which are used as training data for the image classification algorithm. The support vector machine algorithm, which uses non-linear radial basis functions as a kernel, is used to classify the area. Therefore the categories Tree (dark green), Grass (light green), Buildings (brown), Asphalt (light grey), Dark area (dark grey) and Sandy subsoil (light yellow) are defined. The classification results are summarized in Figure 7. An additional covariate that maps the effect of water in the vicinity is created. Therefore all locations that are up to 150 meters away from the Brisbane River are classified as locations close to the water. To achieve this, the coordinates of all grid layer centroids located within a 150-meter distance from the Brisbane river are extracted and labeled.

To perform the spatio-temporal prediction of the temperatures, a 100-meter x 100-meter grid layer with

Figure 7: Support vector machine classification

centroids is placed over the map to define the locations for the estimation. A spatial join is used to determine to which of the previously defined classes the centroid is assigned to. Based on the class assignment, seven dummy variables are created and included as predictors in the estimated temperature model. The IDW model is designed for the spatio-temporal case and thus includes distance-based and time-based covariates, which is why only covariates for the location characteristics (X_l) are included additionally. To take into account the temporal and spatial factors in the RF model, temporal (X_t) and spatial (X_s) factors are included in the model as covariates. Descriptive analyses in section 3 show temporal seasonal effects and fluctuations within the day that are relevant for modeling UHI. Therefore in addition, the day of the year, to model seasonal effects, and the hour of the day, to model daily fluctuations, are included in the model as well as geographical distances between the predicted and observed locations. Temperature is therefore modeled as a function of the time, geographical distances and location-specific covariates:

$$\hat{Z}(s, t) = f(X_l, X_t, X_s). \tag{3}$$

5. Granularity Optimization

To identify the optimal granularity of a UHI sensor network, different combinations of spatial and temporal granularity are used to predict the temperatures for unknown locations in the entire Northshore Hamilton PDA space. Differences between the sites were recognizable on an hourly temporal granularity. Accordingly, the granularity analysis identifies whether finer granular temporal resolutions than one hour lead to an improvement in the modeling. For this reason, 30-minute and 15-minute intervals are further considered. To achieve 30-minute and 60-minute temporal resolution, mean values are calculated over two (for 30-minute) and four (for 60-minute) temperature measurements, respectively. To incorporate spatial granularity, all possible sensor combinations are tested. In each case, individual sensors are omitted, which means that the associated environmental characteristics of the sites are no longer included in the modeling. The temperature data of the omitted sensors can be used to calculate the root mean square error (RMSE) and determine the importance of individual sensors for the estimation. Based on the various temporal and spatial granularities, 33 distinct spatio-temporal granularities are created for evaluation (Table 1).

Table 1: Temporal and spatial granularity variants

Temporal Granularity	Spatial Granularity
A_1{15 minutes}	B_1{S1; S2; S3; S4}
A_2{30 minutes}	B_2{S1; S2; S3}
A_3{60 minutes}	B_3{S1; S2; S4}
	B_4{S2; S3; S4}
	B_5{S1; S3; S4}
	B_6{S1; S2}
	B_7{S1; S3}
	B_8{S1; S4}
	B_9{S2; S3}
	B_{10}{S2; S4}
	B_{11}{S3; S4}

The two algorithms of IDW and RF are tested across all granularity combinations and it is found that increasing the temporal granularity of the data to 30-minutes or 15-minutes intervals did not result in any significant improvements in prediction accuracy. In the IDW application, the RMSEs increase continuously from 60 minutes to 15 minutes, while in the RF application, there are no strongly recognizable differences in the RMSEs between the individual temporal resolutions. If the temporal and spatial factors are considered simultaneously, the result of the combination $\{A3; B9\}$ with an RMSE of 1.48 is the most optimal combination in terms of prediction accuracy of the temperatures. Thus, the best results are achieved when the data from Site 1 and Site 4 are not included in the estimates. Concerning the comparison of the algorithms the analysis also showed that the RF algorithm delivers significantly more precise predictions than the IDW algorithm. This can be demonstrated in particular by the significantly lower error values of RMSE. While the IDW RMSEs reach up to 4.7, the RMSEs within the RF estimate are limited to a maximum of 2.22. To test the impact of the different variables included in the model, we generate a variable importance plot. It can be deduced that the hour of the day and the day of the year are the most significant variables in the model. In addition, dark areas, i.e. low albedo values, are the next most important variable for predicting the temperature outcome. Due to these results of the MSE values, only the estimates of the RF are visualized in Figure 8 choosing the cut-off dates December, 1 and June, 1 of the year 2020 at three different day-times for the optimal granularity combination $\{A3; B9\}$.

Once the optimum granularity for the sensor network has been identified, the second step is to use temperature modeling to identify the times and locations at which heat islands occur. After analyzing the heat islands across the four distinct seasons in Brisbane and examining the average temperatures categorized by environmental classification, both during the day and at night, some patterns emerge regarding the distribution of heat islands. During winter, temperatures exhibit minimal divergence among various site characteristics, maintaining relatively uniform conditions throughout the day and night. However, in summer and autumn, asphalt regions distinctly manifest as heat hotspots. Here, temperatures are approximately 0.2 degrees Celsius higher during the day compared to other areas, escalating further by up to 0.35 degrees Celsius during the night. Moreover, in spring, the prominence of black areas as heat hotspots becomes evident. These locales exhibit temperatures averaging 0.4 degrees Celsius higher during the day and peaking up to 0.6 degrees Celsius higher during the night compared to other classified areas. Additionally, green spaces show comparatively high temperatures throughout the day but cool down significantly overnight, although they do not demonstrate the expected significantly positive effects of green spaces mentioned in the literature. Waterfront regions, on the other hand, are characterized by comparatively lower temperatures and cool down significantly at night.

6. Discussion, Implications and Recommendations

In our study, we evaluate the optimal spatio-temporal granularity of a sensor network for temperature modeling using two algorithms, taking into account the trade-off between costs and benefits. A key objective of our research is to build a Green IS in a micro area, based on the efficient sensor network, that gives environmental decision support for the 2032 Olympics in Brisbane. Regarding the temporal component, the analysis indicates that a finer temporal granularity level than one hour does not provide relevant additional information about the temperature trends for modeling heat islands. An hourly resolution thus enables an efficient balance between computational costs and prediction accuracy, as the data complexity and the amount of data to be

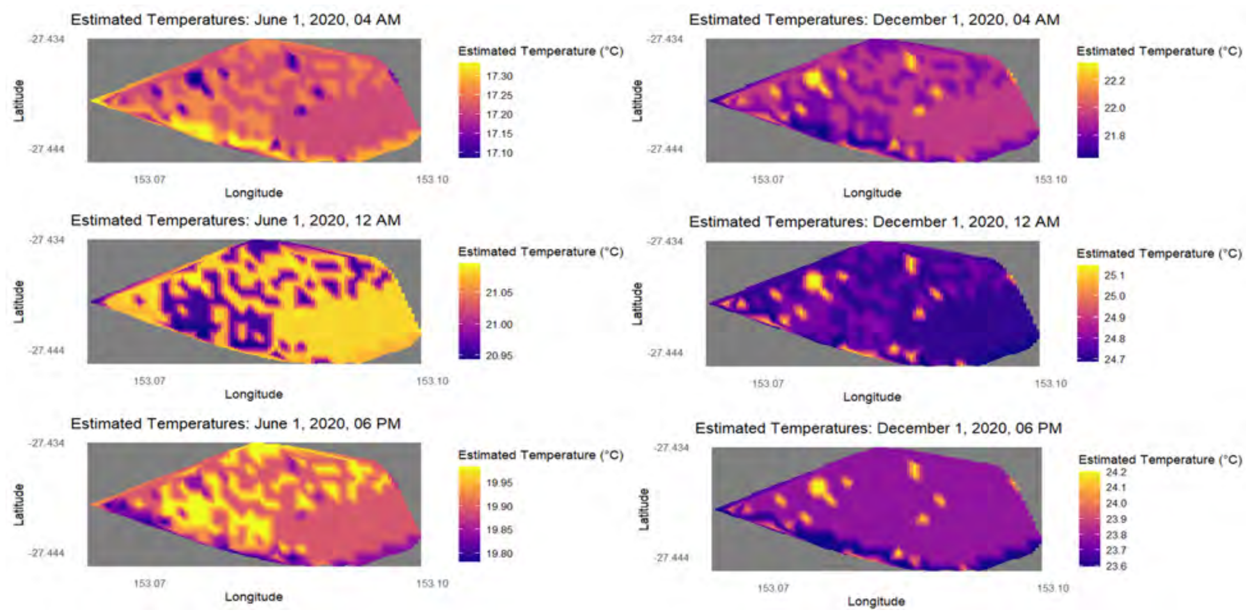

Figure 8: Temperature modeling for June and December

processed can be reduced. This also helps to avoid information overload for UHI stakeholders. Looking at the spatial component, it became clear that the location selection of the sensors plays a crucial role in optimizing the prediction performance for temperature modeling. Our investigation revealed redundancies between Site 1 and Site 4, whereas Site 2 (low albedo value) and Site 3 (sealed surfaces and water proximity) proved to be key spatial components for temperature mapping in the area studied. A conspicuously low albedo value is the strongest identifiable factor for temperature anomalies. To include the green areas in the modeling, just one sensor (Site 1 or Site 4) should be eliminated from the sensor network. By omitting sensors from the sensor network, monetary costs arising by investment, installation and maintenance of the sensors can be saved while at the same time, the benefits of higher prediction accuracy of heat hotspot detection can be increased. By optimizing the complex interaction of spatial and temporal factors in a micro-space, that influence the formation of heat islands, we offer implications for research and practice. We address the scientific lack of sensor network's granularity research in the IS field mentioned by Watson et al. (2010) as well as the gap in the joint consideration of temporal and spatial factors identified in our literature review. For implementation in practice, we supply solid database for UHI stakeholders, to efficiently implement UHI mitigation strategies. The model gives information about the spatio-temporal interactions, particularly emphasizing that different spatial factors have different influences at different times of the year. It can be seen that a blacktop, as a surface with a low albedo, only leads to increased temperatures in spring and summer and is therefore only significantly responsible for the formation of heat islands at these times. Our analysis has shown that an optimal granularity, which considers both spatial and temporal aspects, is necessary to adequately represent the complex interaction for heat island detection in a micro area.

7. Limitations, Further Research and Conclusions

With the optimization of the sensor network granularity, we propose an efficient model to illustrate the microclimate for heat island detection and mitigation by balancing costs and prediction accuracy. Our results and findings can help by deriving specific recommendations for constructing the Athletes' Village in future research. A limitation of our study is the localized application of the model, currently confined to the Northshore Hamilton PDA. However, the optimized sensor network for UHI detection can be readily extended to other regions within the same Köppen Climate Classification as Brisbane. Future research could adapt this model for different climate zones, enabling its application across diverse regions. This approach would also facilitate comparative climate analysis in the topic of UHI across various zones. This also includes the evaluation of the model on external data. At present, our model merely provides a data basis for decision support for the city planners responsible for the 2032 Olympic Games in Brisbane. Future research can offer more in-depth site-specific recommendations based on the optimized sensor network and temperature estimates. This requires a detailed analysis of heat hotspots. The temperature modeling and heat island identification can, e.g., be incorporated into a graphical user interface, providing urban planners and policymakers with environmental decision support for optimal area design for the Olympic Games. In addition to the technical and environmental aspects, future research could also explore the social and behavioral factors that may influence the effectiveness of UHI mitigation strategies. Another limitation is the consideration of a relatively short time horizon. As time progresses, additional temporal data could be integrated into the model to increase prediction accuracy.

Our results and findings underscore the importance of meticulously choosing the spatio-temporal granularity of sensor networks for environmental modeling in the context of heat islands. An optimally configured sensor network not only enhances the accuracy and reliability of heat island detection but also provides a solid basis for decision-making in urban planning. This is especially important for the 2032 Olympic Games in Brisbane and the development of the Athletes' Village, where understanding and mitigating UHI effects are crucial for ensuring sustainable and comfortable living conditions.

8. References

Al-Hmouz, R., Pedrycz, W., & Balamash, A. (2015). Description and prediction of time series: A general framework of granular computing. *Expert Systems with Applications*, *42*(10), 4830–4839.

Bargiela, A., & Pedrycz, W. (2003). Recursive information granulation: Aggregation and interpretation issues. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, *33*(1), 96–112.

Brennan, E. (2022). *Apartment project for brisbane 2032 olympic athletes' village revealed*. https://www.insidethegames.biz/articles/1127085/apartment-project-brisbane

Chen, M.-Y., & Chen, B.-T. (2015). A hybrid fuzzy time series model based on granular computing for stock price forecasting. *Information Sciences*, *294*, 227–241.

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*, 30–42.

Eibl, G., & Engel, D. (2014). Influence of data granularity on smart meter privacy. *IEEE Transactions on Smart Grid*, *6*(2), 930–939.

Gacek, A., & Pedrycz, W. (2012). A characterization of electrocardiogram signals through optimal allocation of information granularity. *Artificial Intelligence in Medicine*, *54*(2), 125–134.

Ghosh, S., & Das, A. (2018). Modelling urban cooling island impact of green space and water bodies on surface urban heat island in a continuously developing urban area. *Modeling Earth Systems and Environment*, *4*(2), 501–515.

Gunawardena, K. R., Wells, M. J., & Kershaw, T. (2017). Utilising green and bluespace to mitigate urban heat island intensity. *Science of the Total Environment*, *584*, 1040–1055.

Hengl, T., Nussbaum, M., Wright, M. N., Heuvelink, G. B., & Gräler, B. (2018). Random forest as a generic framework for predictive modeling of spatial and spatio-temporal variables. *PeerJ*, *6*, #e5518.

Ho, H. C., Knudby, A., Xu, Y., Hodul, M., & Aminipouri, M. (2016). A comparison of urban heat islands mapped using skin temperature, air temperature, and apparent temperature (humidex), for the greater vancouver area. *Science of the Total Environment*, *544*, 929–938.

Huang, B., & Li, H. (2018). Distance-based information granularity in neighborhood-based granular space. *Granular Computing*, *3*, 93–110.

IOC. (2020). Olympic games to become climate positive from 2030. https://olympics.com/ioc/news/olympic-games-to-become-climate-positive-from-2030

IOC. (2021). Ioc elects brisbane 2032 as olympic and paralympic host. https://olympics.com/ioc/news/ioc-elects-brisbane-2032-as-olympic-and-paralympic-host

Khan, W., Liao, J. Y., Walker, S., & Zeiler, W. (2022). Impact assessment of varied data granularities from commercial buildings on exploration and learning mechanism. *Applied Energy*, *319*, 119281.

Knirsch, F., Eibl, G., & Engel, D. (2016). Error-resilient masking approaches for privacy preserving data aggregation. *IEEE Transactions on Smart Grid*, *9*(4), 3351–3361.

Kools, L., & Phillipson, F. (2016). Data granularity and the optimal planning of distributed generation. *Energy*, *112*, 342–352.

Montero, J.-M., Fernández-Avilés, G., & Mateu, J. (2015). *Spatial and spatio-temporal geostatistical modeling and kriging*. John Wiley & Sons.

Nielsen, A. A. (2009). Geostatistics and analysis of spatial data. *Informatics and Mathematical Modelling, Technical University of Denmark, DTU*, 7–12.

Oke, T. R. (1973). City size and the urban heat island. *Atmospheric Environment (1967)*, *7*(8), 769–779.

Pedrycz, W. (2018). Granular computing for data analytics: A manifesto of human-centric computing. *IEEE/CAA Journal of Automatica Sinica*, *5*(6), 1025–1034.

Rusu, C., & Rusu, V. (2006). Radial basis functions versus geostatistics in spatial interpolations. *IFIP International Conference on Artificial Intelligence in Theory and Practice*, 119–128.

Sekulić, A., Kilibarda, M., Heuvelink, G. B., Nikolić, M., & Bajat, B. (2020). Random forest spatial interpolation. *Remote Sensing*, *12*(10), 1687.

Song, M., Hu, L., Feng, S., & Wang, Y. (2023). Feature ranking based on an improved granular neural network. *Granular Computing*, *8*(1), 209–222.

Song, M., Jing, Y., & Pedrycz, W. (2019). Granular neural networks: A study of optimizing allocation of information granularity in input space. *Applied Soft Computing*, *77*, 67–75.

Song, M., & Liu, Y. (2021). A development framework of granular prototypes with an allocation of information granularity. *Information Sciences*, *573*, 154–170.

Song, M., & Wang, Y. (2016). A study of granular computing in the agenda of growth of artificial neural networks. *Granular Computing*, *1*, 247–257.

State of Queensland. (2022). Northshore hamilton scheme amendment. www.statedevelopment.qld.gov.au

Steeneveld, G. J., Koopmans, S., Heusinkveld, B. G., & Theeuwes, N. E. (2014). Refreshing the role of open water surfaces on mitigating the maximum urban heat island effect. *Landscape and Urban Planning*, *121*, 92–96.

Tuczek, M., Degirmenci, K., Desouza, K. C., Watson, R. T., Yigitcanlar, T., & Breitner, M. H. (2022). Mitigating urban heat with optimal distribution of vegetation and buildings. *Urban Climate*, *44*, 101208.

Vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., Plattfaut, R., & Cleven, A. (2015). Standing on the shoulders of giants: Challenges and recommendations of literature search in information systems research. *Communications of the association for information systems*, *37*(1), 9.

Watkins, R., Palmer, J., & Kolokotroni, M. (2007). Increased temperature and intensification of the urban heat island: Implications for human comfort and urban design. *Built Environment*, *33*(1), 85–96.

Watson, R. T., Boudreau, M.-C., & Chen, A. J. (2010). Information systems and environmentally sustainable development: Energy informatics and new directions for the is community. *MIS quarterly*, *34*(1), 23–38.

Watson, R. T., & Webster, J. (2020). Analysing the past to prepare for the future: Writing a literature review a roadmap for release 2.0. *Journal of Decision Systems*, *29*(3), 129–147.

Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS quarterly*, *26*(2), xiii–xxiii.

Yao, M. X. (2019). Granularity measures and complexity measures of partition-based granular structures. *Knowledge-Based Systems*, *163*, 885–897.

Yao, Y. (1998). A comparative study of fuzzy sets and rough sets. *Information sciences*, *109*(1-4), 227–242.

Zhan, Y., Luo, Y., Deng, X., Zhang, K., Zhang, M., Grieneisen, M. L., & Di, B. (2018). Satellite-based estimates of daily no2 exposure in china using hybrid random forest and spatiotemporal kriging model. *Environmental science & technology*, *52*(7), 4180–4189.

Zhang, N., Chen, Y., Luo, L., & Wang, Y. (2017). Effectiveness of different urban heat island mitigation methods and their regional impacts. *Journal of Hydrometeorology*, *18*(11), 2991–3012.

Zhu, X., Pedrycz, W., & Li, Z. (2019). Development and analysis of neural networks realized in the presence of granular data. *IEEE Transactions on Neural Networks and Learning Systems*, *31*(9), 3606–3619.

Zhu, X., Pedrycz, W., & Li, Z. (2020). A development of hierarchically structured granular models realized through allocation of information granularity. *IEEE Transactions on Fuzzy Systems*, *29*(12), 3845–3858.