12th International Workshop on Simulation for Energy, Sustainable Development & Environment, 008
21th International Multidisciplinary Modeling & Simulation Multiconference

2724-0029 © 2024 The Authors.
doi: 10.46354/i3m.2024.sesde.008

Pedestrians' urban thermal comfort: A machine learning assessment through transect walks

José-Antonio Rodríguez-Gallego[1,2*], Eduardo Diz-Mellado[1], Tomás Chacón Rebollo[2], Carmen Galán-Marín[1] and Carlos Rivera-Gómez[1]

[1]Departamento de Construcciones Arquitectónicas I, Escuela Técnica Superior de Arquitectura, Universidad de Sevilla, Avda. Reina Mercedes, 2. 41012, Seville, Spain

[2]Departamento de Ecuaciones Diferenciales y Análisis Numérico, Facultad de Matemáticas, Universidad de Sevilla, Avda. Reina Mercedes, 41012, Seville, Spain

[*]Corresponding author. Email address: jrodriguez14@us.es

Abstract

As the effect of climate change increases, combined with the urban exodus, the relevance of thermal comfort becomes more and more evident, becoming an urgent need for all. This research presents a methodology for data collection of personal, microclimatic, and morphological variables relevant to the evaluation of thermal comfort in an urban environment. A dataset was collected and processed for a total of 200 surveys in 6 different transect walks in Seville, Spain, half of them in the historical city centre and the other half outside of it. Following the visualization of the data, a descriptive analysis of the main variables was carried out, showing the differences between the UTCI index and the real perceived thermal comfort, as well as the improvement associated to vegetation and spatiality. Additionally, the relevance of both air temperature and radiation on perceived thermal comfort was stated, showing how these two variables are intertwined regarding personal preferences. After this analysis, 3 reliable machine learning algorithms were used to predict the expected comfort: Random Forest, XGBoost and a Multilayer Perceptron. This was achieved with the R language, obtaining results of 0.6 accuracy for thermal comfort prediction in all cases, showing therefore the complexity of the problem.

Keywords: Climate resilience; Machine learning algorithms; Outdoor thermal comfort; Perceived thermal comfort surveys

1. Introduction

Thermal comfort is a critical issue in today's changing environment (Lai et al., 2020). With the urban exodus phenomenon and Urban Heat Island (UHI) (Kim and Brown, 2021) becoming increasingly relevant step by step not only for the use of outdoor urban areas, but even for survival purposes, the consideration of urban resilience is mandatory for the future of the cities. The increase in temperature, both globally (Lindsey and Dahlman, 2020) and in urban areas as a consequence of the UHI (Kim and Brown, 2021) has transformed from an inconvenience to a real and dire hazard to life, especially considering children and elders (Cleland et al., 2023).

In recent years, for the analysis of this phenomenon, many studies have been carried out to gather useful data in a variety of ways, from climate walks with the occasional collaboration of citizens (Peng et al., 2022) to fast measurements with bikes or other transports (Young et al., 2022). This approach can prove to be difficult, as specialized equipment is required as well as many collaborators if the objective is to gather an appropriate amount of data. However, notice that this action is essential, as there is a generalized research gap on this topic.

1.1. Thermal comfort

One of the main objectives when talking about thermal comfort nowadays -and since the origins of the area- is to accurately assess the perceived thermal comfort associated with a certain set of conditions, which are usually related to the microclimatic situation, the personal perception, or the influence of the urban environment, especially when speaking about the outdoor environment. As one could expect, this aim is challenging, on the one hand, due to the vast number of variables involved in the phenomenon, and on the other due to the lack of available and reliable data. Even if advanced machine learning models are suitable for this task, such as the Random Forest (Ho, 1995) and the XGBoost (Chen and Guestrin, 2016), and could work relatively well in problems of this nature, this particular situation proves to be especially demanding and complex for these models.

1.2. Research issues and aim

With the following research, innovative results for climate walks are presented, firstly presenting a methodology for collecting data while, at the same time, involving the citizens. With the gathered data a descriptive analysis is presented for the different transects, as well as for the relationship between some of the usually avoided variables. Subsequently, the application of three different machine learning models for the prediction of thermal comfort follows, trying to develop a fast tool for the assessment of this perceived variable in different outdoor conditions, taking into account not only microclimatic variables but also variables related to both the particular person and the urban morphology, expanding the set of input variables over similar works as (Chaudhuri et al., 2018; Farhan et al., 2015). With this, it would be possible to know the effect certain variables have over the perceived outdoors comfort.

The research is structured in different sections. After discussing the state of the art in Section 2, in Section 3, the methodology used for data collection in the different selected neighbourhoods of the city of Seville is presented, as well as the machine learning tools which will be employed. Subsequently, in Section 4, an extended analysis of the collected data, novel variables and comparison with the UTCI outdoor comfort index is presented, as well as the presentation of the results using three different machine learning techniques, namely Random Forest, XGBoost, and a Multi-Layer Perceptron, all of them fine-tuned. This work provides additional information on commonly overlooked variables, as well as the relationship between them and the accuracy that common and accurate machine learning models can provide for the prediction of thermal perception.

2. State of the art

The evaluation of outdoor thermal comfort plays a crucial role in urban planning and design, especially given the increasing urbanization and effects of climate change (Hurlimann et al., 2021; Lindsey and Dahlman, 2020). Through technological progress and interdisciplinary studies, this field has undergone substantial development, incorporating computational simulations, data analysis (Lai et al., 2020), and creative design approaches to enhance outdoor spaces with a focus on human well-being. Important elements such as temperature, humidity, wind speed and sunlight are closely examined (Bröde et al., 2012) to understand their influence on how people perceive thermal conditions, leading to the establishment of comprehensive methodologies for assessing outdoor thermal comfort.

In the current ever-changing environment, the focus on creating sustainable outdoor spaces has become crucial. Stakeholders are utilizing big data (Kharrazi et al., 2016), machine learning, and adaptive design approaches to improve thermal comfort and support climate resilience (Lai et al., 2020). Various initiatives, such as green infrastructure (Monteiro et al., 2020) and Climate adaptive architecture, are being implemented to combat the effect of urban heat islands (Kim and Brown, 2021) and develop inviting, healthy and pleasant outdoor areas that prioritize the welfare of urban populations in the face of climate change.

2.1. Outdoor thermal comfort

When considering outdoor comfort, people in urban environments (Chan and Chau, 2021) often assess thermal comfort through direct exposure to the outdoor environment, leading to a pleasant experience for urban dwellers (Mayer and Höppe, 1987). This perception is subject to seasonal variations, whether it is cold or warm, as well as factors such as activity levels and clothing choices (Matzarakis et al., 1999). In recent years, the development of various outdoor thermal indices has been seen. The evaluation of outdoor thermal comfort has been approached through various methods, including numerical simulation (Berkovic et al., 2012), survey-based assessments or measurements, and the exploration of novel models (Coccolo et al., 2016). These indices serve as tools to assess the outdoor thermal environment concerning its occupants. Despite the creation of new thermal comfort indices for both indoor and outdoor spaces, only a select few are currently utilized. This is due to considerations of complexity, comprehensiveness, and adaptability to diverse climatic scenarios, leading to a reduction in the number of outdoor comfort indices in practical use (Potchter et al., 2018).

Several outdoor comfort indices are available to estimate user comfort, the most notable example being the Universal Thermal Climate Index (UTCI) (Höppe, 1999), which emerged as the most widely used index among meteorologists worldwide. After the intensive development of thermophysiological models by Fiala et al. (Bröde et al., 2012), 20 years ago the UTCI was presented as an equivalent temperature to that in a reference environment, making it easy and fast to assess the expected comfort under certain conditions. One of the main advantages of this index is that it only requires the measurements of four microclimatic variables to provide its classification of the expected comfort, using under the hat Fiala's multinodal thermal regulation model (Fiala et al., 2012) to predict heat across a standard body. However, by doing this, some variables are summarized or ignored, such as sex, weight, or mood (Bröde et al., 2012), therefore losing information that could be decisive in predicting the expected comfort.

Recent studies have introduced Metamatrix Thermal Comfort (Höppe, 1999) to facilitate the selection of comfort indices, showing UTCI as one of the most comprehensive indices. Furthermore, recent investigations (Ma et al., 2021) have incorporated coupling models employing comfort indices based on experimental values such as pavement porosity (Liu et al., 2022), humidity levels, etc.

2.2. Climate walks

The analysis of thermal comfort relies deeply on the availability of clean data, which for this task can be something uncommon. Even if one could develop tools for the assessment of the perceived thermal comfort in certain conditions, the reality is that those models — such as Random Forest or XGBoost— require big datasets for the training step. Therefore, collecting data for different conditions and individuals is key for the expansion of the area.

To address this difficulty, a new alternative has appeared in recent years: climate walks (Peng et al., 2022), a way of collecting data while at the same time making people invested in the analysis of their surrounding reality. Different ways of doing this include counting with the help of voluntary organizations, collectives, or bystanders formed for this particular purpose. A climate walk is therefore considered as an organized event in which the participants walk a prepared route to raise awareness about climate change, as well as to discuss environmental subjects and action plans.

2.3. Machine learning algorithms for comfort prediction

Machine learning algorithms are increasingly popular for predicting outdoor thermal comfort (Fard et al., 2022), which can help design environments that promote human well-being. Among these algorithms, the Random Forest model is particularly effective due to its ability to handle complex datasets and non-linear relationships. By combining multiple decision trees, Random Forest can accurately capture the complex interactions between various environmental factors, such as temperature, humidity, wind speed, and solar radiation, to predict levels of thermal comfort. Its ensemble approach improves the accuracy of predictions, making it an ideal choice for outdoor comfort evaluations. In recent years, it has been extensively used in the context of thermal comfort, but mainly for the analysis of interiors, as in the work by (Aparicio-Ruiz et al., 2023) for offices or (Diz-Mellado et al., 2021) for semi-outdoor spaces like courtyards in the city of Seville.

In the same vein, XGBoost, a superior version of gradient boosting, exhibits outstanding results in predicting outdoor thermal comfort. Thanks to its ability to efficiently process sparse data and apply regularization methods, XGBoost skillfully tackles the complexities brought about by varied environmental conditions. Its collection of weak learners progressively improves predictions, leading to sturdy models that can detect delicate trends in outdoor thermal behaviour. Even if it has been proven to be precise and reliable in many situations, this particular model has not been used for the prediction of thermal comfort, except for a few papers, such as (Wu et al., 2024). Finally, the Multilayer Perceptron (MLP), a fundamental component of neural network structures, provides an additional method for precise predictions. The deep learning abilities of MLP allow it to understand intricate feature representations from complex environmental data, enabling a detailed understanding of changes in outdoor comfort. Its adaptability to modelling nonlinear relationships and to adjusting to different scenarios makes it an essential tool in the search for ideal outdoor thermal comfort solutions. Being the first and smallest neural network model, in the last year it has attracted researchers from the thermal comfort area, but again, mainly for the comfort in the home (Irshad et al., 2020).

3. Materials and Methods

As the work entailed a complete analysis, the need was not only of available and useful data, but also of software for its exploitation. Therefore, the preparation of an adequate setting for the investigation was mandatory, making clear the origin and size of the data and how to be used afterwards.

3.1. Transect walks

A transect walk is a systematic walk along a defined path (transect) across a specific city area. In this work, the selection of transects was determined to achieve a comprehensive coverage of the city's landscape and morphology. Geographical distribution played a key role, as the study had to be within the city limits, being a good sample of the city's behaviour. Moreover, the decision to include transects with varying population densities was made so that a broad spectrum of perspectives could be captured. By incorporating densely populated areas alongside neighborhoods with potentially different demographic profiles, the survey was ensured to accurately reflect the city's diverse populace. As a result, this provides not only a wide snapshot of the city dynamics, but also a variety of data for the following analysis, which returns a higher accuracy in terms of thermal comfort prediction.

Furthermore, the choice of paths included both core and peripheral regions of the city to enable a comparative study of views and experiences, as follows:

- C1, Santa Cruz: Embracing the traditional Andalusian style, Santa Cruz features narrow, urban canyons with traditional materials such as ceramic bricks in the pavements.
- C2, El Arenal: This historical neighbourhood, with winding, narrow streets, presents small hubs for local activity and events, as well as compact parks.
- C3, Alameda de Hércules: Situated in the heart of Seville, this route flows from Plaza Nueva up to Resolana Street, changing from narrow streets to the open square that is the Alameda de Hércules, filled with people at all hours.
- E1, Juan XXIII: Providing modern living spaces within Seville's urban landscape, Juan XXIII features apartments with clean lines, ample windows, and communal green areas.
- E2, El Carmen: Neighborhood with medium and high-rise buildings, with variable road sections including squares and urban parks.
- E3, Huerta del Carmen: This residential neighborhood, situated in the bounds of the city, was built in the 1950s, and is filled with small squares with scattered vegetation.

By including transects such as C1, C2, and C3, which traverse central locales, alongside transects like E1, E2, and E3, representing non-central, residential neighbourhoods, the objective was to discern potential disparities between the city centre and its outskirts. This tactical method improves the way the survey data delivers a thorough understanding of the city's dynamics, providing viewpoints on urban living across various geographic and socio-economic settings. In Table 1, the exact number of surveys carried out in all six neighbourhoods is displayed, together with their corresponding codes.

Table 1. Surveys per route and codification of the route. 23/09 and 21/10 of 2023 for C-routes and E-routes, respectively.

Route	Neighbourhood	Centre?	Number of surveys
C1	Santa Cruz	✓	39
C2	El Arenal	✓	44
C3	Alameda de Hércules	✓	21
E1	Juan XXIII	×	26
E2	El Carmen	×	17
E3	Huerta del Carmen	×	51

3.2. Climate walks

The way the data collection step was realized was through climate walks in collaboration with the Climathon event (Climathon, 2024), the participants being the actors of these surveys, asking people on their respective transect questions about thermal comfort.

With the implication of bystanders as well as the participants of the abovementioned event both the reach of the research and its diffusion were expanded. At the same time, the discussion about their reality and the effect some variables could have on the perceived comfort proved to be useful for the task of choosing the most relevant variables for their analysis.

3.3. Surveys

The process of collecting data and preparing them for the following steps was complex in itself, as there is no standard form of outdoor thermal comfort, and the number of variables typically considered in this subject ranges from a couple of dozen to up to 70 (Lai et al., 2020). In this particular case, the focus was not only on measuring personal data regarding perception but also on the morphological and microclimatic components of the phenomenon and its impact on personal perception.

To address this challenge, a standardized survey has been developed, based on previous work such as (Nikolopoulou and Lykoudis, 2006), so that this information can be easily obtained with the aid of Google Forms. To include a wider range of data, two different dates were chosen, one month apart, during the Autumn of Seville, making measurements in both the historical centre, whose thermal resilience has been extensively studied (Sánchez Ramos et al., 2022; Diz-Mellado et al., 2023), and in other neighbourhoods, aiming at providing a complex snapshot of morphologically different areas.

Each form was divided into three distinct sections, one for personal information and perception, one for the particular microclimatic situation of the environment, and one for the morphological characteristics of the rounding area. Notice that the second and third items were common for all surveys obtained in a particular stop, as it is assumed that the microclimatic conditions do not change under such small spatio-temporal variations. With the merging of the different sections, a form for a particular set of conditions for a certain person is generated, with more than 60 variables available in total.

The study areas were chosen trying to obtain the representation of both the historical inner part of the city and of that of their surroundings, working with a total of six transects, three for the historical centre, and another three for other neighbourhoods. On each transect, between five and seven stops were selected for the measurements and the execution of the surveys, trying to capture morphologically different areas.

3.4. Software for data processing

Within the area of urban planning and environmental evaluation, it is vital to precisely forecast thermal comfort to ensure the well-being and satisfaction of inhabitants. Machine learning methods offer a promising way to improve this prediction process, handling large amounts of data, and helping to find patterns that might not be obvious otherwise. By using machine learning, it is possible to create models that adapt to different situations, making outdoor environments more comfortable and sustainable for everyone.

3.4.1. Random Forest

A Random Forest, presented originally in (Ho, 1995), is an ensemble learning method that builds several decision trees during training, with each tree trained on random subsets of features and data to reduce overfitting. This randomness improves the robustness and precision of the model. In the prediction phase, every tree makes an independent class prediction, with the ultimate prediction being decided by either a majority vote (in classification) or by taking the average (in regression) from all the trees. The popularity of random forests stems from their simplicity, ability to scale, and proficiency in managing intricate high-dimensional data interactions.

3.4.2. XGBoost

XGBoost (Extreme Gradient Boosting) is a highly optimized and scalable implementation of gradient boosting algorithms (Chen and Guestrin, 2016). Sequentially, it builds a collection of first learners, typically decision trees, by adapting them to the residuals of prior iterations. It utilizes regularization methods to avoid overfitting and parallel processing to boost computational speed. Its novelty is found in the dual optimization of the loss function and the regularization component, which results in quicker convergence and improved model efficacy. XGBoost is widely used in machine learning competitions and real-world applications because of its high predictive accuracy and efficiency.

3.4.3. Multilayer Perceptron

A Neural Network (Haykin, 1994) is a computational model inspired by the human brain, consisting of interconnected neurons organized in layers. Through training with algorithms such as backpropagation, they learn to map input data to output predictions by adjusting neuron weights. They excel in learning intricate patterns from large datasets, offering flexibility and the ability to capture nonlinear relationships for both classification and regression tasks. More precisely, in this case, a Multi-Layer Perceptron (MLP) is deployed (Haykin, 1994), using different widths, i.e., the number of neurons in the only layer of the net. This is done classically, but using a bit of fine-tuning so that an appropriate value for the hyperparameter can be chosen.

3.4.4. Data processing

For the necessary cleaning and processing steps over the collected data, the R language was used, a programming language specially designed for data visualization (Wickham, 2016) and, recently, for the extensive development of machine learning and artificial intelligence algorithms (Kuhn and Wickham, 2020). More precisely, to do this, the guidelines given by Wickham followed, tidying up the data through using the tidyverse framework for data reading, sorting, recording and, in general, cleaning.

After preparing a "tidy" dataset, in the sense used by Wickham, the tidymodels framework (Kuhn and Wickham, 2020) was used to quickly deploy three different machine learning algorithms, more precisely a Random Forest (RF), a boosted tree via XGBoost, and a Multilayer Perceptron (MLP), the simplest instance of neural network. For a fast explanation of each of those models, one can check (Rigatti, 2017) for the Random Forest, (Chen and Guestrin, 2016) for the XGBoost, and (Ruck et al., 1990) for the Multilayer Perceptron. These particular models were chosen because their utility and reach have been proven one time and another for many different situations. In addition to their accuracy and adaptability to many scenarios, they were chosen due to their availability within the tidymodels framework.

4. Results and Discussion

In this section, the main results of the analysis are exposed, ranging from the review of the surveys performed, to the descriptive analysis of some of the variables involved and their relationship with thermal comfort, to end with the use of the machine learning models and their accuracy for the assessment of the thermal comfort associated with certain conditions.

4.1. Surveys sample

A standard questionnaire was designed, and 198 surveys were carried out in 6 different neighbourhoods, both in the city's historic center and outside of it. The surveys were performed on the morning of days 23rd of September and the 21st of October, between 11:00 and 13:00 hours in the first case and 13:00 and 15:00 hours in the second. The exact number of surveys performed in each of the six neighborhoods, as well as the codification of the following figures, can be seen in Table 1.

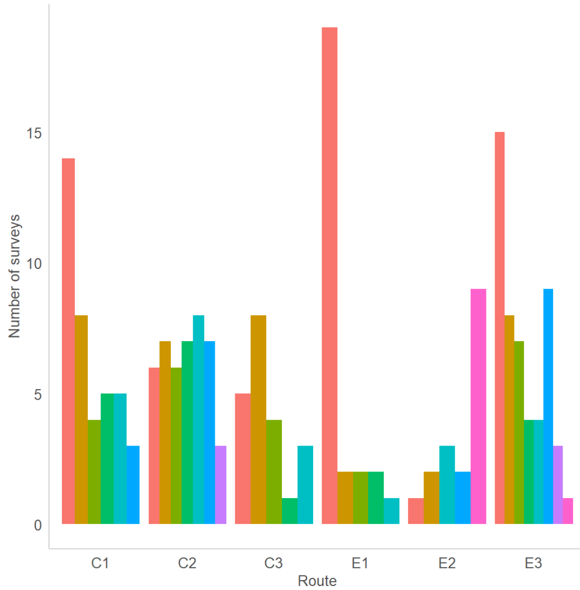

Figure 1. Survey distribution per neighborhood.

Due to the variability in the number of people throughout the transects, the number of surveys changes from one stop to another, as well as between transects. To see the distribution of surveys per stop, one can look at Figure 1, in which it can be seen how many surveys were taken in each stop, ranging from 1 to more than 15. One can see in the figure that for most stops the number of surveys was around 6, becoming 3 for noncentric transects 1 and 2. In general, however, almost the same number of surveys were taken in and outside the city historical center. Notice, however, that there were certain locations with a higher pedestrian traffic, as it is the case with the first stop at route E1, Plaza Nueva, a common meeting place and transit area. The opposite behaviour was found on other stops on routes E1 and E2, where there was little transit.

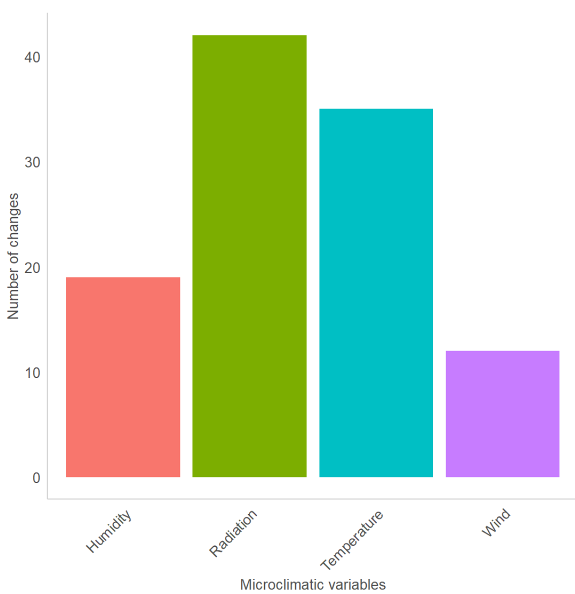

Figure 2. Would like to make changes in microclimatic variables.

One of the questions asked to the participants in the survey was about the microclimatic variables that they would like to change, from a list of the four considered to compute the UTCI index. In Figure 2 it is possible to see the total count of people who answered asking to change one or more microclimatic variables. One of the insights given by the figure is that both radiation and temperature are the most prominent variables in this aspect, which could be explained by the high temperature for the chosen days, even in autumn. It is interesting, additionally, to see how little people cared about the wind in this situation, being less important than humidity, which was already less relevant than temperature and radiation.

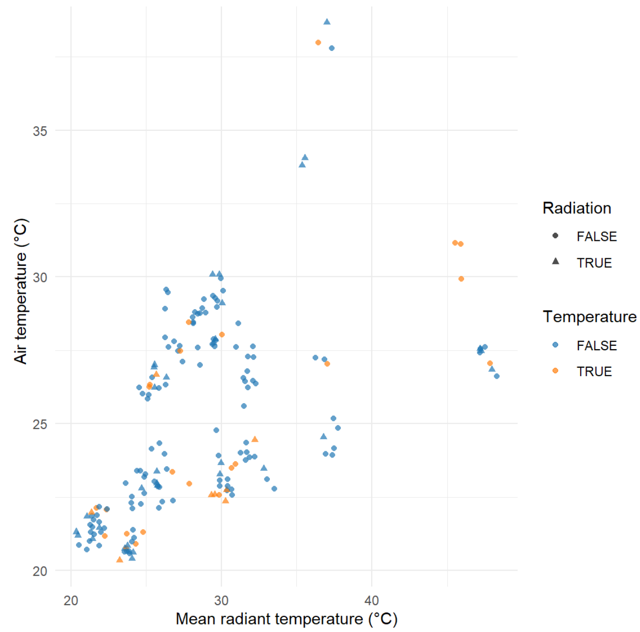

Figure 3. Would change temperature versus would change radiation, including a small jitter for a better visualization.

After the results given in Figure 2 the question arises on whether these two variables, radiation and temperature, are connected regarding personal perception. To quickly check this hypothesis Figure 3 was made, plotting the air temperature (ºC) against the mean radiant temperature (ºC), and using colour and shape to represent whether people wanted to change temperature and/or radiation. It can be seen that the behaviour of the personal preferences are similar when changing the temperatures, pointing out that there is no clear relation between the perception of these variables.

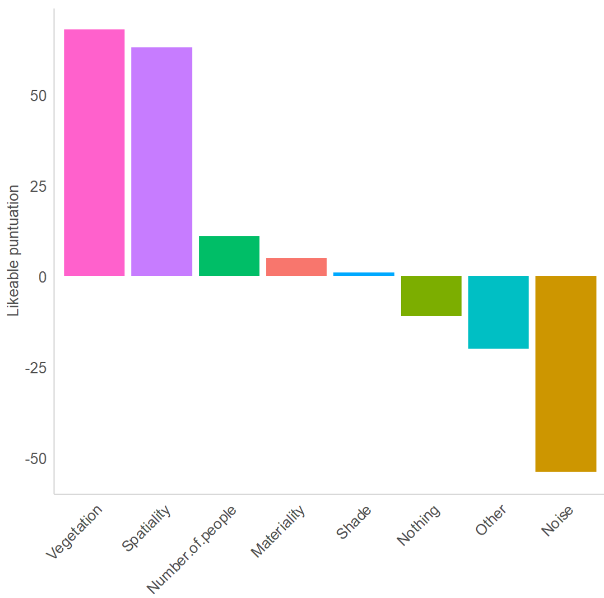

Figure 4. Likeable and dislikeable elements

When asked about the morphological aspects of the environment, a list of possible elements was given, making the participants answer whether they found the element likeable or not. From those answers a likeable punctuation was developed, measuring how likeable were those elements across the whole campaign. To compute the punctuation one must simply, for each element, add 1 for each person that liked that element, and -1 if they disliked it, with neutral answers adding 0. This value represents in a fast manner the mean behaviors, as can be seen in Figure 4, in which is clear that both vegetation and spatiality are the main liked elements in the surveyed areas, while noise seems to be the most unlikeable. This last result is coherent with the fact that the number of people is also one of the most liked elements, being completely opposed to the perceived noise.

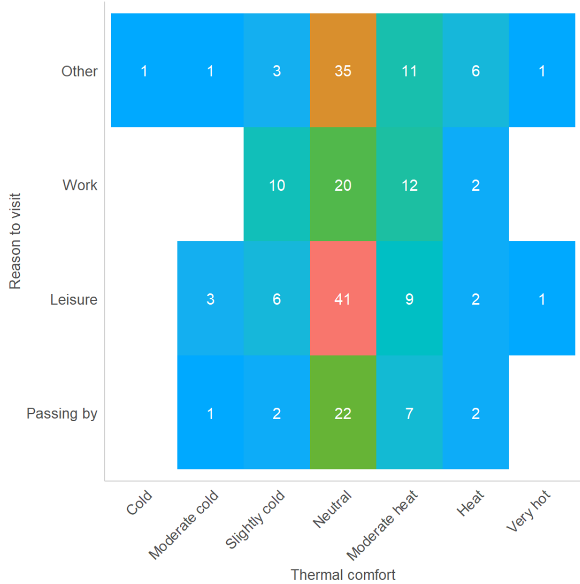

Figure 5. Reason of visit and perceived comfort

In Figure 5 the distribution of answers for the reason to visit the area is presented, as well as the perceived thermal comfort, split by reason. This variable is not usually considered, and it was expected to affect the thermal perception, but from this plot it, can be understood that the thermal comfort doesn't have such a dependence on the reason of the visit. However, future tests should be taken to properly verify this fact. Notice that "Passing by" and "Work"/"Leisure" are considered different categories, the first being for cases in which people were moving to a different place, no matter the reason; while the second was used only if the surveyed person was actually either working or spending some leisure time at the site of the survey.

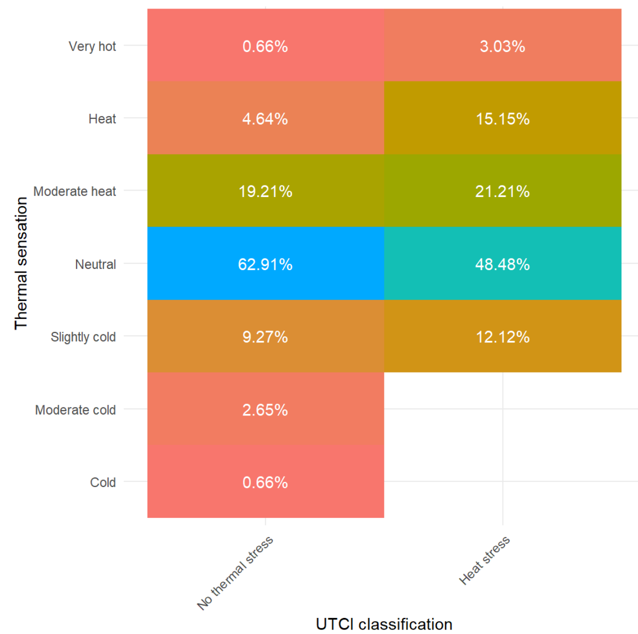

Figure 6. UTCI classification versus real thermal sensation. Percentages by UTCI classification.

One of the addressed objectives was to verify the accuracy of the UTCI index, by measuring the perceived thermal comfort in the same scales. When making a comparison between the perceived and the expected comfort, as seen in Figure 6, one can see that the UTCI classification predicts heat stress in a great number of cases in which the participant was neutral or even slightly cold, losing the expected accuracy of this index.

This reduced performance could potentially be attributed to the development of the index, as it was created as a global tool, with our data coming from a single city in Spain. If we consider Fiala's multinodal model (Fiala et al., 2012), a bias can be found on the original data, as there is mainly representation of American people. That bias could have been inherited by multinodal model and, then, by the UTCI index, therefore explaining the lack of accuracy working with Sevillian people.

4.2. Machine learning algorithms

For the first two cases a bootstrap cross-validation estimate of the metrics was used with 25 sets prepared over the 198 entries of the dataset, while for the multilayer perceptron cross-validation was employed, with 80% of the full dataset in the training subset and the rest in the test subset. This was done due to limitations in the deployed code.

Table 2. Hyperparameters for the different models.

Model	Parameter	Value	Description
RF	num. trees	100	Number of trees
RF	mtry	25	Number of regressors
XGBoost	mtry	27	Number of regressors
XGBoost	trees	100	Number of trees in the ensemble
XGBoost	tree_depth	6	Depth of the tree
MLP	hidden_units	20	Width of the net

The value of the hyperparameters obtained through fine-tuning can be seen on Table 2. The results for the associated classification are given in Table 3 for different metrics, followed by an example of classification, as the one presented for the UTCI index, for the Random Forest model, in Figure 7.

Table 3. Metrics for different models.

Model	Parameter	Value
RF	Accuracy	0.59
RF	Sensitivity	0.34
RF	F_1 score	0.47
XGBoost	Accuracy	0.54
XGBoost	Sensitivity	0.31
XGBoost	F_1 score	0.36
MLP	Accuracy	0.59
MLP	Sensitivity	0.2
MLP	F_1 score	0.74

RF stands for Random Forest, and MLP for Multilayer Perceptron.

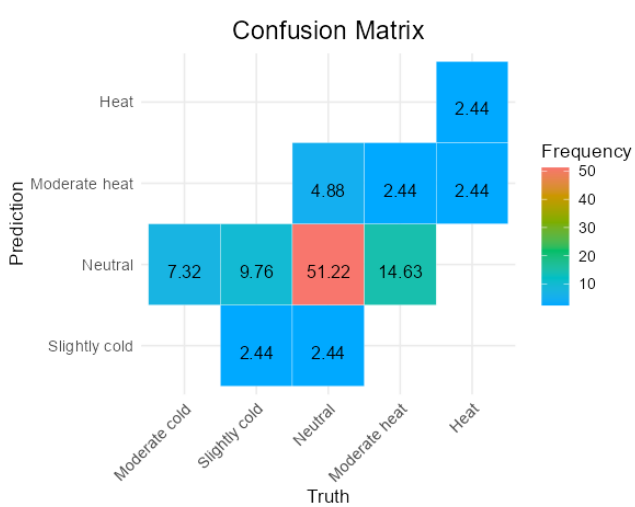

Figure 7. Real classification versus predictions for the test set using the Random Forest model.

From Table 3 it is possible to extract the following:

- For all models, the accuracy is below 0.6, indicating that the models are not performing very well in terms of overall correctness. RF and MLP have the highest accuracy at 0.59, while XGBoost has a slightly lower accuracy at 0.537.
- Regarding sensitivity or true positive rate, the number of positive cases that were correctly identified by the model, it can be said that RF has the highest sensitivity at 0.34, followed by XGBoost at 0.31 and MLP at 0.2. This suggests that RF is better at correctly identifying positive cases compared to the other models.
- As a summary of the previous metrics, the F1 score can be used, being the harmonic mean of both and being appropriate for cases in which the class distribution is not completely balanced. MLP has the highest F1 score at 0.74, followed by RF at 0.47 and XGBoost at 0.36. This indicates that MLP has the best balance between precision and recall among the three models.

Overall, based on these results, it can be said that RF and MLP perform similarly in terms of accuracy, but MLP outperforms RF in terms of the F1 score. XGBoost has the lowest performance in all metrics among the three models. MLP has the highest F1 score, suggesting it might be the best choice if a balance between precision and sensitivity is desired. Notice, however, that these results are lower than expected for datasets of this size, especially in terms of accuracy. This could be related to errors in the dataset, as well as the personal noise associated to the preferences of the participants.

5. Conclusions

As exposed, the analysis of outdoor thermal comfort can be truly challenging, as it is a complex phenomenon composed of many intertwined processes. To address this task with machine learning tools an adequate amount of data is essential, a difficult requirement in this area due to the general lack of available and reliable datasets. This is mainly attributed to both the need for special equipment and the large number of variables involved in this situation, some of them subjective and difficult to process correctly.

With these challenges in mind, a standardized survey was developed, and 198 surveys were conducted in 6 different neighbourhoods, both in the city's historic centre and out of it, measuring more than 60 variables in total for each answer, made up of variables related to personal, microclimatic, and morphological factors. Surveys were carried out as part of a climate walk, therefore involving the citizens of Seville in the analysis of their surroundings and providing information on the relevant variables.

Analysis of the variables by themselves provided information about the expected behaviours of people in this environment and relating them to perceived thermal comfort gave additional information on the main issues people find regarding this subject. One of the most prominent results was the real differences between the predicted thermal comfort, given by the UTCI index, and the real thermal comfort experienced, a discrepancy that needs further analysis. Additionally, it was observed, in a descriptive way, that vegetation and spatiality were the most liked aspects of the environment, compared to noise, the most disliked. It was also noted that, under the studied conditions, both air temperature and radiation are perceived as variables to change, even if the relationship between the two regarding perception is still unclear.

After the descriptive analysis, and addressing the main objective of the investigation, machine learning techniques were used to develop a tool for the evaluation of outside thermal comfort given a particular set of conditions. After data preparation, three different well-known and reliable data-driven models were used: Random Forest, XGBoost, and a Multilayer Perceptron. In all cases, the accuracy of the models ranged from 0.54 to 0.59, below the results for similar classification tasks, which shows the complexity of the studied phenomenon and the need for additional analysis. Even if the accuracy values were between acceptable ranges for this task, the obtained sensitivity was much lower, with a maximum value of 0.34, which points out how the trained models tend to predict positives, *i.e.*, comfort over discomfort.

As further work expands the one presented here, it would be useful to try other techniques to detect and assess the importance of the different variables, making it easier to collect data and, therefore, increasing the total size of the dataset. With an expanded dataset, the results given by these machine learning algorithms should improve, as they are based heavily on the quality and quantity of data. Additionally, applying the methodology for data collection to different thermal environments, such as Summer or during a heat wave would improve their prediction capabilities. Finally, expanding the research to other areas, in and outside of Seville, would further improve the representation of the results and the reach of the chosen models.

6. Funding

This work has been supported by the project PID2021-124539OB-I00 funded by MCIN/AEI/10.13039/501100011033 and by "ERDFA way of making Europe", and the project TED2021-129347B–C21 funded by MCIN/AEI/10.13039/501100011033 and by the "European Union NextGenerationEU/PRTR".

7. Acknowledgements

The authors would like to thank AEMET (Agencia Estatal de Meteorología) for the data provided, and the students of the Master in Sustainable Architecture of the University of Seville and participants in CLIMATHON Seville 2023 for their collaboration in the monitoring campaigns.

8. References

Aparicio-Ruiz, P., Barbadilla-Martín, E., Guadix, J., and Nevado, J. (2023). Analysis of variables affecting indoor thermal comfort in Mediterranean climates using machine learning. *Buildings*, 13(9):2215.

Berkovic, S., Yezioro, A., and Bitan, A. (2012). Study of thermal comfort in courtyards in a hot arid climate. *Solar Energy*, 86(5):1173–1186.

Bröde, P., Fiala, D., Błażejczyk, K., Holmér, I., Jendritzky, G., Kampmann, B., Tinz, B., and Havenith, G. (2012). Deriving the operational procedure for the universal thermal climate index (utci). *International journal of biometeorology*, 56:481–494.

Chan, S. Y. and Chau, C. K. (2021). On the study of the effects of microclimate and park and surrounding building configuration on thermal comfort in urban parks. *Sustainable Cities and Society*, 64:102512.

Chaudhuri, T., Zhai, D., Soh, Y. C., Li, H., and Xie, L. (2018). Random forest based thermal comfort prediction from gender-specific physiological parameters using wearable sensing technology. *Energy and Buildings*, 166:391–406.

Chen, T. and Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, pages 785–794.

Cleland, S. E., Steinhardt, W., Neas, L. M., West, J. J., and Rappold, A. G. (2023). Urban heat island impacts on heat-related cardiovascular morbidity: A time series analysis of older adults in us metropolitan areas. *Environment International*, 178:108005.

Climathon. (2024, May 10). https://climathon.climate-kic.org/

Coccolo, S., Kämpf, J., Scartezzini, J.-L., and Pearlmutter, D. (2016). Outdoor human comfort and thermal stress: A comprehensive review on models and standards. *Urban Climate*, 18:33–57.

Diz-Mellado, E., Nikolopoulou, M., López-Cabeza, V. P., Rivera-Gómez, C., and Galán-Marín, C. (2023). Cross-evaluation of thermal comfort in semi-outdoor spaces according to geometry in southern spain. *Urban Climate*, 49:101491.

Diz-Mellado, E., Rubino, S., Fernández-García, S., Gómez-Mármol, M., Rivera-Gómez, C., and Galán-Marín, C. (2021). Applied machine learning algorithms for courtyards thermal patterns accurate prediction. *Mathematics*, 9(10):1142.

Fard, Z. Q., Zomorodian, Z. S., and Korsavi, S. S. (2022).

Application of machine learning in thermal comfort studies: A review of methods, performance and challenges. *Energy and Buildings*, 256:111771.

Farhan, A. A., Pattipati, K., Wang, B., and Luh, P. (2015). Predicting individual thermal comfort using machine learning algorithms. In 2015 IEEE International Conference on Automation Science and Engineering (CASE), pages 708–713. IEEE.

Fiala, D., Havenith, G., Bröde, P., Kampmann, B., and Jendritzky, G. (2012). Utci-fiala multi-node model of human heat transfer and temperature regulation. *International journal of biometeorology*, 56:429–441.

Haykin, S. (1994). Neural networks: a comprehensive foundation. *Prentice Hall PTR*.

Ho, T. K. (1995). Random decision forests. In Proceedings of 3rd international conference on document analysis and recognition, volume 1, pages 278–282. IEEE.

Höppe, P. (1999). The physiological equivalent temperature–a universal index for the biometeorological assessment of the thermal environment. *International journal of Biometeorology*, 43:71–75.

Hurlimann, A., Moosavi, S., and Browne, G. R. (2021). Urban planning policy must do more to integrate climate change adaptation and mitigation actions. *Land Use Policy*, 101:105188.

Irshad, K., Khan, A. I., Irfan, S. A., Alam, M. M., Almalawi, A., and Zahir, M. H. (2020). Utilizing artificial neural network for prediction of occupants thermal comfort: A case study of a test room fitted with a thermoelectric air-conditioning system. *IEEE Access*, 8:99709– 99728.

Kharrazi, A., Qin, H., and Zhang, Y. (2016). Urban big data and sustainable development goals: Challenges and opportunities. *Sustainability*, 8(12):1293.

Kim, S. W. and Brown, R. D. (2021). Urban heat island (UHI) intensity and magnitude estimations: A systematic literature review. *Science of the Total Environment*, 779:146389.

Kuhn, M. and Wickham, H. (2020). Tidymodels: a collection of packages for modeling and machine learning using tidyverse principles.

Lai, D., Lian, Z., Liu, W., Guo, C., Liu, W., Liu, K., and Chen, Q. (2020). A comprehensive review of thermal comfort studies in urban open spaces. *Science of the Total Environment*, 742:140092.

Lindsey, R. and Dahlman, L. (2020). Climate change: Global temperature. *Climate*. gov, 16.

Liu, Y., Ma, H., Zhang, C., and Luo, X. (2022). Watering on porous pavement for improvement of environmental human thermal comfort in an ecological community in arid area: A case study in lanzhou, china. *Sustainable Cities and Society*, 85:104081.

Ma, H., Zhang, C., Jia, J., Hou, C., and Wang, G. (2021). Investigation on human thermal comfort of the ecological community in arid area of lanzhou, china. *Sustainable Cities and Society*, 72:103069.

Matzarakis, A., Mayer, H., and Iziomon, M. G. (1999). Applications of a universal thermal index: physiological equivalent temperature. *International journal of biometeorology*, 43:76–84.

Mayer, H. and Höppe, P. (1987). Thermal comfort of man in different urban environments. *Theoretical and applied climatology*, 38:43–49.

Monteiro, R., Ferreira, J. C., and Antunes, P. (2020). Green infrastructure planning principles: An integrated literature review. *Land*, 9(12):525.

Nikolopoulou, M. and Lykoudis, S. (2006). Thermal comfort in outdoor urban spaces: analysis across different european countries. *Building and environment*, 41(11):1455–1470.

Peng, Z., Bardhan, R., Ellard, C., and Steemers, K. (2022). Urban climate walk: A stop-and-go assessment of the dynamic thermal sensation and perception in two waterfront districts in Rome, Italy. *Building and Environment*, 221:109267.

Potchter, O., Cohen, P., Lin, T.-P., and Matzarakis, A. (2018). Outdoor human thermal perception in various climates: A comprehensive review of approaches, methods and quantification. *Science of the Total Environment*, 631:390–406.

Rigatti, S. J. (2017). Random forest. *Journal of Insurance Medicine*, 47(1):31–39.

Ruck, D. W., Rogers, S. K., and Kabrisky, M. (1990). Feature selection using a multilayer perceptron. *Journal of neural network computing*, 2(2):40–48.

Sánchez Ramos, J., Toulou, A., Guerrero Delgado, M., Palomo Amores, T. R., Castro Medina, D., and Álvarez Domínguez, S. (2022). Thermal resilience of citizens: comparison between thermal sensation and objective estimation in outdoor spaces: a case study in Seville, Spain. *Applied Sciences*, 12(22):11676.

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

Wu, C., Pan, H., Luo, Z., Liu, C., and Huang, H. (2024). Multi-objective optimization of residential building energy consumption, daylighting, and thermal comfort based on bo-xgboost-nsga-ii. *Building and Environment*, 254:111386.

Young, E., Kastner, P., Dogan, T., Chokhachian, A., Mokhtar, S., and Reinhart, C. (2022). Modeling outdoor thermal comfort along cycling routes at varying levels of physical accuracy to predict bike ridership in cambridge, ma. *Building and Environment*, 208:108577.