**OPEN**

# Using deep-learning to forecast the magnitude and characteristics of urban heat island in Seoul Korea

Jin Woo Oh[1], Jack Ngarambe[1], Patrick Nzivugira Duhirwe[1], Geun Young Yun[1]* & Mattheos Santamouris[1,2]

Urban heat island (UHI), a phenomenon involving increased air temperature of a city compared to the surrounding rural area, results in increased energy use and escalated health problems. To understand the magnitude and characteristics of UHI in Seoul and to accommodate for the high temporal variability and spatial heterogeneity of the UHI which make it inherently challenging to analyze using conventional statistical methods, we developed two deep learning models, a temporal UHI-model and a spatial UHI model, using a feed-forward deep neural network (DNN) architecture. Data related to meteorological elements (e.g. air temperature) and urban texture (e.g. surface albedo) were used to train and test the temporal UHI-model and the Spatial UHI-model respectively. Also, we develop and propose a new metric, UHI-hours, that quantifies the total number of hours that UHI exists in a given area. Our results show that UHI-hours is a better indicator of seasonal UHI than the commonly used index, UHI-intensity. Consequently, UHI-hours is likely to provide a better measure of the cumulative effects of UHI over time than UHI-intensity. UHI-hours will help us to better quantify the effect of UHI on, for example, the overall daily productivity of outdoor workers or heat-related mortality rates.

The world's population is increasing dramatically; new cities are being built, while existing cities are getting over-populated. Currently, more than 50% of the world population resides in cities[1]. This increased development in urban areas has resulted in changes in urban morphology and surfaces, as well as an increase in the amount of anthropogenic heat released to the atmosphere[2]. The combined effect of such changes leads to higher air temperature in urban areas than in the surrounding rural or suburban areas, hence the formation of heat islands[3]. These heat islands, especially in summers, considerably decrease the outdoor air quality and trigger heat-related diseases and deaths in urban areas[4,5]. Moreover, extreme temperatures have been reported to be the leading cause of weather-related deaths[6]. This is particularly prevalent among the elderly (people of 65 years of age and above) and those with cardiovascular and respiratory health issues. For example, Paravantis et al.[7] analyzed the impact of temperature and heat waves on the number of deaths caused by cardiovascular and respiratory health issues in people above 65 years of age in Athens, Greece. They reported a U-shaped exposure-response curve, indicating reduced mortality rates at moderate temperatures and 20% and 35% increase in mortality rates at very low and very high temperatures, respectively. Furthermore, due to the increase in urban temperatures, there has been an increase in air-conditioning system usage, which in turn has led to high electricity demands and overall increased building energy consumption[8]. At the same time, passive cooling methods, such as natural and night ventilation, have become ineffective for the thermal comfort of building occupants[9].

Due to the effects of urban heat island (UHI), the UHI issue has been receiving increased attention from many academic disciplines. A great part of this attention has been directed to the development of data-driven UHI-predictive models that can forecast UHI at a given time in the future. Such models are helpful in the development of UHI-mitigation strategies and in the drafting of relevant countermeasure policies. In addition, data-driven methods offer reduced computational load in the assessment of UHI than computer simulation programs such as UMEP-SOLWEIGH[10], ENVI-MET[11] etc., that are often used to model UHI behavior. Moreover, the process involved in conducting computer simulations is highly dependent on expert knowledge of the issue at hand. This makes the use of computer simulations in the analysis of UHI susceptible to errors in instances when the modeler lacks sufficient expert knowledge. Additionally, attaining an accurate model of complex urban

[1]Department of Architectural Engineering, Kyung Hee University, Yongin, 17104, Republic of Korea. [2]High Performance Architecture, School of Built Environment, University of New South Wales, Sydney, NSW, 2052, Australia. *email: gyyun@khu.ac.kr

geometries through analytical modeling requires a bulk of urban details that may not always be available to the casual user or difficult to obtain. In such cases, data-driven models provide a quick way to assess UHI behavior based on information that is commonly available and easily obtained (e.g. meteorological data). Moreover, data-driven models can be easily deployed in a system consisting of already programmed sensors (e.g. temperature sensors) to automatically and continuously display UHI levels at a given time based on prevailing environmental conditions and without human intervention.

Some of the common approaches used in developing UHI-predictive models include conventional regression methods. For instance, Chun and Guldmann[12] explored urban UHI determinants using regression analysis and spatial data obtained from 2D- and 3D-aerial satellite images. Similarly, Su et al.[13] demonstrated the relations among land-cover types (e.g., built-up, water, paddy field, and vegetation) using weighted-regression modeling. However, most regression methods are unable to capture complex nonlinear relations, which are also highly variable and spatially heterogeneous, between UHI determinants and UHI. Additionally, although there exist statistical methods that can be used to map non-linear data (e.g. non-linear regression methods), they often require iterative processes of fitting different types of mathematical functions to determine one that best fits a given dataset.

Consequently, advanced non-linear analysis techniques such as artificial neural networks (ANNs) have been extensively applied in studying the UHI phenomenon. With ANN algorithms, a model is automatically generated through the adjustment of weights and biases from one layer to another. For example, Mihalakakou et al.[14] developed a backpropagation ANN model to estimate diurnal and nocturnal UHI intensity fluctuations in the Greater Athens area. This model is based on hourly temperature and solar radiation data collected from 23 automatic weather stations (AWSs) developed over a period of 2 years. The study identified urban air temperature as the predominant factor explaining nighttime and daytime UHI. In a different study, Mihalakakou et al.[15] developed backpropagation ANN models based on synoptic-scale atmospheric circulation, ambient temperature, maximum daily solar radiation, and mean daily wind speed. Their conclusions emphasized the impact of synoptic-scale atmospheric circulation on UHI magnitude. Similarly, Gobakis et al.[16] compared the performance of three types of ANN architectures (Cascade, Elman, and Feedforward) for UHI prediction using ambient air temperature and global solar radiation data collected from 14 meteorological stations in Athens. The study results indicated Elman ANN as the optimum neural-network architecture for predicting UHI.

While ANN algorithms generally provide better fitted models than conventional statistical techniques, mainly owing to their ability to learn and mimic highly complex non-linear data, existing studies have mostly employed shallow ANNs (i.e., neural networks with a single hidden layer) to study UHI. However, ANN are prone to overfitting and often face problems related to generalizability[17]. To address this issue, we developed two types of deep neural network (DNN) models that predict and quantify the UHI effect in Seoul; a temporal and a spatial model. Spatial-pattern modeling attempts to explain the UHI phenomenon using geographic indicators (e.g., land-cover types), whereas temporal modeling uses meteorological indicators (e.g., air temperature) to explain UHI distribution. Spatial patterns have been used to explain several issues in urban environments[12,13,18,19]. The main difference between the temporal UHI-model and the spatial UHI-model developed in the current study is that the spatial UHI-model explains how urban factors affect the manifestation of UHI in urban areas whereas the temporal UHI-model explains how meteorological elements affect the manifestation of UHI in urban areas.

There are also several studies that report on the UHI phenomenon in Seoul specifically[20–23]. Although the analysis of UHI intensity in Seoul has been reported previously, the existing studies used either single-year meteorological data or data collected from only a few AWSs. As such, analyses from previously developed UHI-predictive models for Seoul, such as those reported by Lee et al.[20] and Kim and Baik[21], are likely to be limited, given that they did not use multiyear meteorological data, which are essential in developing accurate UHI-predictive models. The studies that did use multiyear meteorological data in model development, such as Kim and Baik[19], did so with data collected from relatively few AWSs, which also limited the breadth of their analysis. Consequently, to overcome the shortcomings of previous studies concerning UHI in Seoul, we used multi-year meteorological data collected over a period of 9 years (2009–2017) from 54 AWSs located across Seoul. The use of multi-year datasets in the analysis of UHI is likely to provide better results than using single-year datasets because data from a single particular year is unlikely to be a proper representative of typical weather conditions within a region. There are certain climatic events (i.e. increased heat waves) that occur in a particular year but which are not common occurrences in the region of interest. Moreover, data from 26 of the 54 AWSs is new data being used to study UHI in Seoul for the first time. It is also important to realize that UHI behavior in a given area is likely to change over time due to factors such as development and changes in land use. Therefore, the analysis of UHI in a given area should be based on recent meteorological and land-use data. Thus, unlike previous studies concerning UHI in Seoul, particularly those involving the development of predictive models, our study uses the most recent meteorological and land-use data, and therefore, provides insights into the current status of the UHI phenomenon in Seoul. Furthermore, UHI effects are commonly assessed using UHI intensity as an indicator[24]. Although UHI intensity is a useful and accurate indicator of UHI variations within a time frame, it has been reported to underestimate UHI levels at peak solar time or during the daytime[25]. Moreover, UHI intensity is a measure of UHI at a specific time and therefore inadequate to assess the cumulative effects of UHI. Consequently, we develop and propose the use of a new metric, UHI-hours, that computes the total number of hours that UHI exists in a given area. Such a measure is useful in quantifying the cumulative effects of UHI over a prolonged period of time. For instance, UHI-hours can be used to show the potential correlations between UHI and reduced productivity of outdoor workers, mortality rates in summertime, annual building energy consumption etc.

Conclusively, the aims of the current study are twofold; The first aim was to develop and propose a new metric to assist in quantifying the cumulative effects of UHI. The second aim was to develop data-driven models that can be used to predict UHI levels using either temporal data (e.g. meteorological factors) or data related to the spatial characteristics of a given area (e.g. land use data).
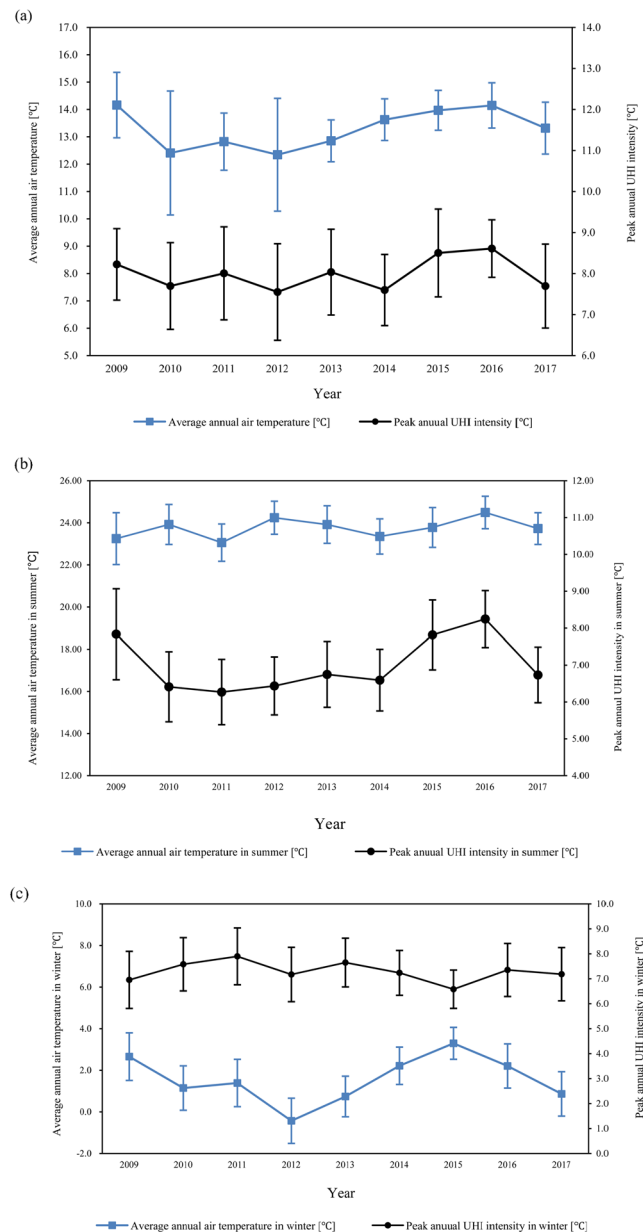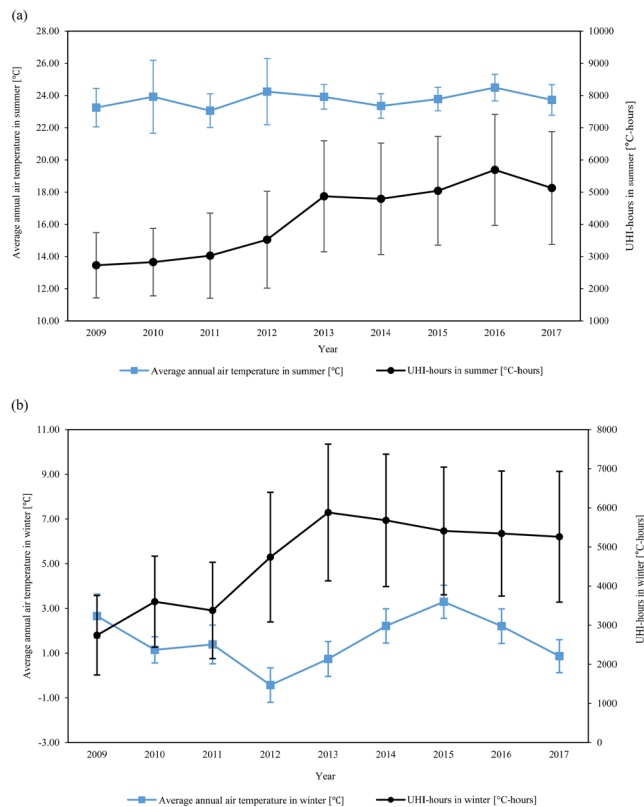
**Figure 1.** Changes in average ambient temperature and seasonal peak UHI intensity: (**a**) Average annual air temperature and peak annual UHI intensity. (**b**) Average ambient temperature and seasonal peak UHI intensity in summer (**c**) Average ambient temperature and seasonal peak UHI intensity in winter; (the error bars indicate standard deviations).

## Results

**Relations among UHI intensity, UHI-hours, and air temperature.** The trends in our data can be described in two distinct phases; changes in UHI intensity and air temperature over the years and correlations between UHI intensity and air temperature over the years. As shown in Fig. 1(a), our results show no clear trends in air temperature over the years. The air temperature tended to increase and decrease over the years in no particular order. For example, while the average air temperature was 14.2 °C in 2009, it was 12.3 °C in 2012 and 14.1 °C in 2016. This same trend is observed for UHI intensity. However, air temperature itself was correlated with UHI intensity. The years with high average air temperatures are the same years which experienced high average UHI intensity and vice-versa. For a detailed analysis, the relation between the average ambient measured temperature and peak measured UHI intensity was determined for two seasons: summer and winter. Based on existing empirical studies[26], the summer season in Seoul is considered to be from April 30 to September 7, and the winter is considered to be from November 1 to February 28.

Furthermore, as presented in Fig. 1(b,c), the average seasonal ambient temperature and peak seasonal UHI intensity were not, in general, proportional to each other, except during certain periods. From 2013 to 2017, the average ambient temperature during the summer season decreased from 23.9 °C in 2013 to 23.3 °C in 2014;
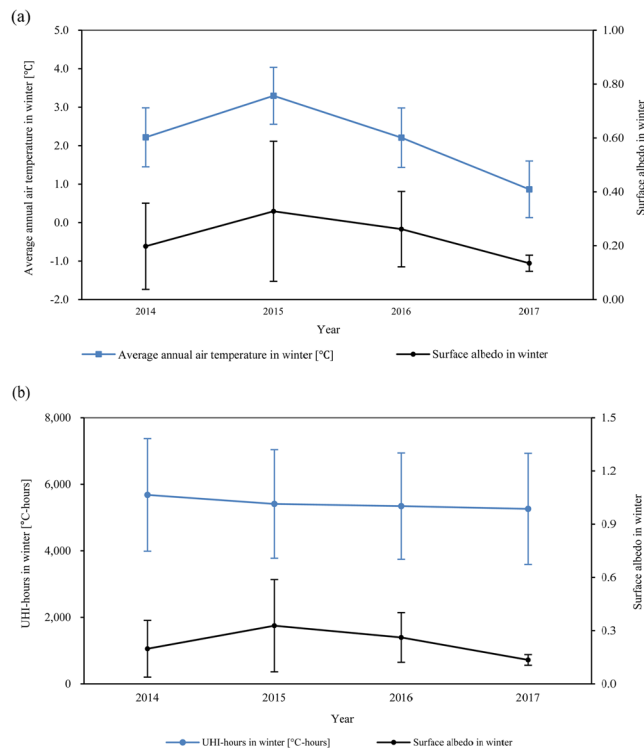
**Figure 2.** Changes in seasonal average ambient temperature and seasonal peak UHI-hours: (**a**) summer, (**b**) winter; (the error bars indicate standard deviations).

then, it steadily increased to 24.5 °C in 2016 and decreased again to 23.8 °C in 2017. During the same period, the summer peak UHI intensity decreased from 6.8 °C in 2013 to 6.6 °C in 2014; then, it increased to 8.3 °C in 2016 and decreased again to 6.7 °C in 2017. Thus, for the period 2013–2017, the summer peak measured UHI intensity was proportional to the average ambient temperature in the summer season. However, during 2009–2010 and 2012–2013, while the average summer ambient temperatures varied within 23.2–23.9 °C and 24.2–23.9 °C, respectively, the summer peak UHI intensity changed from 7.8 to 6.4 °C and from 6.4 to 6.8 °C, showing no clear relation between temperature and UHI for the periods 2009–2010 and 2012–2013.

In the winter season during 2010–2013, the average ambient temperature increased from 1.1 °C in 2010 to 1.4 °C in 2011, and then decreased to −0.4 °C in 2012 and increased again to 0.7 °C in 2013. The same trend of a repeated increase and decrease was observed in the peak UHI intensity, which changed from 7.6 °C to 7.9 °C, to 7.2 °C, and to 7.7 °C, thereby showing a correlated relation between air temperature and UHI intensity during winters. During 2014–2017, the average winter ambient temperature increased from 2.2 °C to 3.3 °C, and then decreased to 1.1 °C. During the same period, the peak UHI intensity repeatedly increased and decreased, from 7.2 °C to 6.6 °C, to 7.4 °C, and to 7.2 °C. Because we found no clear relation between air temperature and UHI intensity, and since UHI-hours is a better representation of cumulated heat islands, we conducted further analyses to see if there was any relationship between seasonal air temperature and UHI-hours.

Earlier, we explained that UHI-hours is the sum of hours when the UHI intensity is above 0 deg C for a certain period of time. Comparing UHI-hours and UHI-intensity therefore, our results showed that the average seasonal temperature is more correlated with UHI-hours than UHI intensity. For instance, when the summer UHI-hours increased from 4972 °C-h to 5888 °C-h between 2014 and 2016, the average ambient measured temperature also increased as indicated in Fig. 2(a). However, when UHI-hours decreased to 5289 °C-h in 2017, the average summertime measured ambient temperature also tended to decrease. In winter, except for the two years from 2014 to 2015, the average wintertime ambient temperature showed a correlated relation with wintertime UHI-hours as shown in Fig. 2(b).

**Relations among spectral radiance, surface albedo, and UHI-hours.** We analyzed the relations among spectral radiance, surface albedo, and seasonal air temperature as well as those among spectral radiance, surface albedo, and seasonal UHI-hours for the period between 2014 and 2017 (Fig. 3). During 2014–17, the results show that there was a potential correlation between the average winter surface albedo both the average ambient temperature as shown in Fig. 3(a) and UHI-hours as shown in Fig. 3(b). The winter surface albedo increased from 0.20 to 0.33 and then decreased to 0.13. This trend was similar to that of the average winter ambient temperature in the same period. In addition, the winter UHI-hours during the same period steadily decreased from 5680 °C-h in 2014 to 5260 °C-h in 2017. During 2015–2017, the changes in UHI-hours in the winter season

**Figure 3.** Wintertime changes in surface albedo with (**a**) average ambient temperature and (**b**) UHI-hours; (the error bars indicate standard deviations).

were correlated to those in winter surface albedo. Besides conveying simple weather information over time, surface albedo is also a surrogate for land-surface changes over time. As such, our results suggest that both the seasonal average ambient temperature and UHI-hours are influenced by the changes in urban morphology (e.g., changes in land-surface type). However, we found no substantial relation between spectral radiance and average ambient temperature or peak UHI-hours.

**Temporal UHI model.** We developed a temporal UHI model to use with Seoul AWS data, showing maximum, minimum, and median temperatures (AWS-17, AWS-35, and AWS-45, respectively), using a deep-learning method. Here, model performance results are presented for both the training and the validation datasets. The prediction models for UHI variations for the three AWSs showed that, for the training dataset, the model for AWS-35 had the highest $R^2$ (0.98) and lowest RMSE (0.26), the model for AWS-45 had the next highest $R^2$ (0.97) and next lowest RMSE (0.27), and that for AWS-17 had the lowest $R^2$ (0.96) and the highest RMSE (0.29). In case of the validation dataset, the models for AWS-17 and AWS-35 had the same $R^2$ (0.94) and that for AWS-45 had $R^2 = 0.93$. The model developed using data from AWS-17 had the lowest RMSE (0.29), followed by the model for AWS-35 (RMSE = 0.41), and that for AWS-45 (RMSE = 0.44). Also, we used the combined training and validation datasets to predict the UHI distribution for each location selected based on the actual UHI intensity obtained through the inverse-distance-weighted interpolation method. Generally, the prediction model developed with data from AWS-17, which also showed the maximum annual UHI-hours, produced the highest $R^2$ and lowest RMSE values, whereas that developed with data from AWS-45, which was associated with median annual UHI-hours, produced the lowest $R^2$ and highest RMSE (Fig. 4).

**Spatial UHI model.** We also developed a deep-learning model based on urban spatial information to predict UHI distribution during the winter and summer seasons. This model was developed to predict UHI intensity and UHI-hours for the entire area of Seoul and to quantitatively analyze the effect of urban spatial elements on UHI and UHI-hours distribution. Our results showed that prediction models for UHI-hours generally had a higher $R^2$ value than those for UHI intensity. Both UHI intensity and UHI-hours models showed a higher $R^2$ value and a lower RMSE value in summer than in winter. In summer, the models showed $R^2$ values of 0.98 and 0.99 for UHI intensity and UHI-hours prediction, respectively. In winter, the model had the same $R^2$ value of 0.99 for both UHI intensity and UHI-hours prediction. Furthermore, the model for predicting UHI intensity had an RMSE value of 0.08 in summer and 0.09 in winter, while that for UHI-hours had an RMSE value 145.22 in summer and 176.68 in winter. Table 1 shows the performance of the developed models for the prediction of UHI intensity and UHI-hours according to each season.

*Relative importance of predictive factors for UHI.* We analyzed the relative importance of physical and environmental factors affecting UHI using a deep-learning approach and UHI-hours as an indicator. We applied
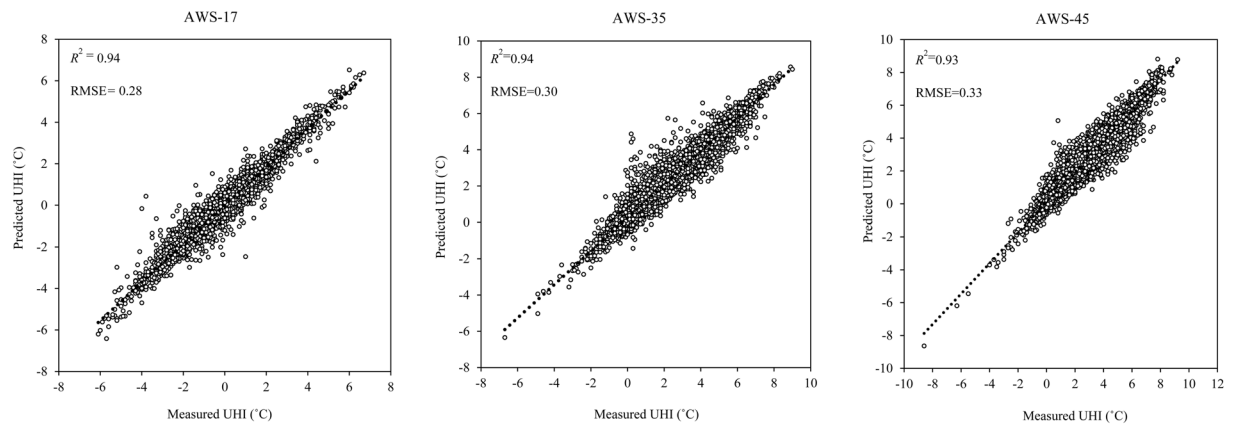
**Figure 4.** Performance of DNN temporal UHI model. [AWS, automatic weather station; $R^2$, coefficient of determination; RMSE, root-mean-square error; UHI, urban heat intensity].

| Spatial model | Season | $R^2$ | RMSE |
|---|---|---|---|
| UHI intensity UHI intensity | Summer | 0.98 | 0.08 |
| | Winter | 0.99 | 0.09 |
| UHI-hours UHI-hours | Summer | 0.99 | 145.52 |
| | Winter | 0.99 | 170.68 |

**Table 1.** Performance of deep-learning prediction models based on spatial data. $R^2$, coefficient of determination; RMSE, root-mean-square error; UHI-hours, urban heat island hours; UHI intensity, urban heat island intensity.

the Gedeon method[26], which uses functional analysis of the weighted matrix of a trained neural-network model to determine the behavioral significance of the model features in the hidden neurons. We found that in summer, the green-area, road-area, cropland-area, and bare-area ratios were the most important factors influencing UHI-hours, whereas spectral radiance and surface albedo had less impact. In winter, the green-area and road-area ratios were the most important factors affecting UHI-hours.

We also found that in both summer and winter, the green-area, road-area, and bare-area ratios were the most important factors affecting UHI-hours, whereas the sky-view factor (SVF) and the footprint were the least important factors affecting UHI-hours. However, the total floor-area ratio had a higher impact on UHI-hours than the footprint ratio.

Factors such as water-area ratio, spectral radiance, and surface albedo showed different degrees of influence on UHI-hours according to season. In summer, the water-area ratio was an important factor affecting UHI-hours. However, its influence was less significant during winter. In addition, the impact of spectral radiance and surface albedo on UHI-hours was more pronounced in winter than summer.

## Discussion

The present study provides a comprehensive analysis of the UHI phenomenon in Seoul city. We use a large dataset of meteorological elements (i.e. 9-year climatic data) and land use data (i.e. 3-year data) collected at short time intervals. Thus, because of the large amount of data collected and the intervals at which they were collected, including the data used for the first time in the analysis of UHI in Seoul, our study provides a deeper understanding of which factors are more likely to influence UHI behavior in an urban environment. For example, the data used in previous studies for UHI prediction were mostly collected at relatively long time intervals (3 or 6 h) compared to the data used in the present study. For instance, Kim and Baik[23] used data from a single AWS collected at 6-h intervals to predict UHI variations in Seoul. Similarly, Pakarnseree et al.[27] used climate data collected over a 5-year period at 3-h intervals to analyze UHI patterns in Bangkok. In our study, however, we used climate data collected over a 9-year period at 1-h intervals, thus allowing for a more-resolved analysis of UHI distribution in Seoul. This is because using data recorded in short time intervals (e.g. 1 hour) over that recorded in long time intervals (e.g. 3 hours), especially when dealing with dynamic physical elements such as air temperature, enables us to capture the sudden changes that occur in very short time spans which might be quite influential when predicting UHI.

Similarly, compared to previous studies, the data used in the present study was collected from an extensive network of stations. For example, Tan et al.[28] used data from 11 AWSs to examine the UHI effect on human health in Shanghai. Similarly, Li et al.[29] used data from four AWSs to analyze the relation between UHI intensity and summertime air pollution in Berlin. However, in the present study, we used data from 54 AWSs spread across Seoul city. The use of data collected from an extensive network of stations (i.e. spread over Seoul city) is important because it ensures that the models are trained on a heterogeneous dataset of the factors affecting the manifestation of UHI and consequently resulting in a better understanding of how different factors influence UHI behavior. For
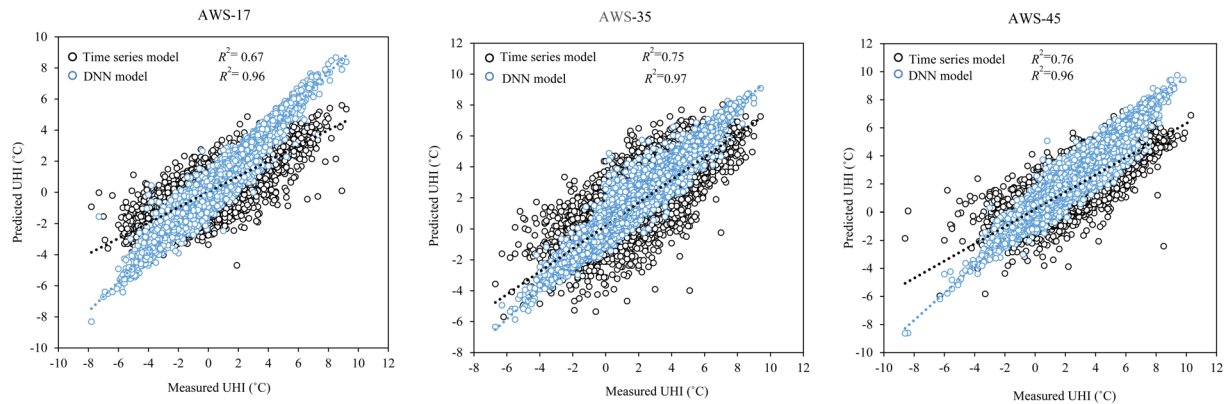
**Figure 5.** Comparative performance between the UHI-predictive models developed using the DNN algorithms and performance of UHI-predictive models developed using time-series algorithm models.

example, the 54 AWSs are spread across spots/areas with different characteristics. These different characteristics impact on the manifestation of UHI differently, especially those related to land use. By using data from an extensive network of stations therefore we ensured that the values of model variables are represented in diverse ranges and thus making the developed models more generalizable compared to the models developed in previous studies. To compare the coverage of the AWS used in our study and those from previous studies, we calculated and compared the area represented by each AWS between the abovementioned studies and our study. We found that the areas of the previously studied cities of Shanghai[28] and Berlin[29] and of Seoul (i.e. our study) are approximately 6340 km$^2$, 900 km$^2$, and 605 km$^2$, respectively. As such, while the surface area represented by each AWS in these two previous studies was 576 km$^2$ (Shanghai) and 225 km$^2$ (Berlin), the surface area represented by each AWS in our study was 11 km$^2$. This implies that the models developed in the present study are more representative of our subject area (i.e. Seoul) than those developed by other studies.

Another contribution of our study deals with quantifying and measuring UHI variation. To quantify UHI effects, many researchers have used UHI intensity as a metric. Although UHI intensity is accurate in representing urban heating during the nighttime and morning hours, it underrepresents urban heating at peak solar hours or in periods dominated by solar heating[25]. Consequently, we developed and proposed the use of a new metric, UHI-hours, to quantify UHI. UHI-hours allows for the cumulative quantification of UHI patterns in a small local area or UHI patterns for a specific season. Our results indicated that UHI-hours is a more accurate metric than UHI intensity for assessing seasonal UHI behavior. Consequently, UHI-hours is more likely to provide a better measure of the cumulative effects of UHI within an urban environment. For example, UHI-hours is more likely to be correlated with the overall reduced productivity of outdoor workers than UHI-intensity. Similarly, UHI-hours is likely to be better correlated with mortality rates due to heat stress in the summertime than UHI-intensity. This is primarily because the effect of heat exposure on the body builds up over time which in turn gradually starts affecting different physiological processes of the body. As such, UHI should be quantified as a cumulative element rather than a distinct value such as in the case of UHI-intensity.

Our research also improves on the predictive models previously developed to study and predict UHI behavioral patterns. The earlier studies mostly applied ANN and time-series modeling techniques to predict UHI intensity distribution. For example, Lee et al.[20] developed an ANN model to predict UHI intensity in Seoul using climate data from 28 AWSs and urban spatial characteristics. Kim and Baik[21] also developed an ANN model to predict UHI intensity in Seoul, which they based on the four factors of wind speed, relative humidity, cloudiness, and UHI intensity from the previous day. However, these two studies developed shallow ANN models, which are prone to overfitting and are not generally conducive to generalization[17]. To address this issue, we developed DNN models, which are less prone to overfitting and generally perform better on new data as compared to shallow ANN models. This implies that the models developed in the present study can be successfully used to predict UHI in other parts of Korea.

Furthermore, many studies have often used time-series models to study UHI[30–32]. Consequently, we developed a time-series model, based on ARIMA, for use as a benchmark model to assess the performance of the developed temporal UHI model relative to previously developed models. The algorithm used to develop the time-series models is based on autoregressive integrated moving average (ARIMA) modeling[30,31], which allows for the prediction of future data based on past data and the relations among the error factors. Our results showed that DNN models had a better predictive ability than the time-series models. The overall $R^2$ value was significantly higher for the DNN model than for the time-series model (Fig. 5); the difference in the $R^2$ value ranged from 0.19 and 0.29. As such, our results suggested that DNN-based models are likely to be more beneficial (i.e. in terms of predictive accuracy) in studies concerning UHI than models based on time series modelling techniques.

We also used the developed spatial model to determine the relative importance of the physical and environmental factors influencing UHI formation. The results of its application indicated that green- and road-area proportions are the main factors that influence UHI formation in Seoul in both the summer and winter seasons. This finding is consistent with those of previous studies, such as Li et al.[33], who reported that UHI was highly correlated with the type of land surfaces present in an area, and Weng et al.[34], who revealed the effect of green

areas on the control of surface temperatures. In addition, our results indicate that the total-floor-area ratio is an important factor influencing UHI distribution. This finding is also consistent with reports from previous studies, such as that by Chun and Guldmann[12], who reported the influence of building and road density on UHI in central urban areas. Consequently, the results obtained in the present study suggest that increasing the green areas and reducing impervious surfaces in urban areas are essential in combating adverse UHI effects. Also, the proposed method allows for the understanding of the impact of greenery under variable boundary conditions and for different synoptic systems. Additionally, the information provided here regarding the hierarchical ranking of the land use factors that are most likely to affect UHI can be leveraged by policy makers when drafting policies to mitigate the effects of UHI in urban climates.

While the present study presents useful information regarding the current status of the UHI phenomenon in Seoul city and subsequently develops accurate UHI-predictive models, it still faced certain limitations. One of the limitations deals with the variety of variables used to analyze UHI. Previous studies have shown that UHI can be influenced by a range of variables such as solar radiation[35], cloud cover[36], rainfall[37] etc. However, data related to such variables for Seoul city was not available from the AWS at the time of this study. Also, data related to the spatial characteristics of Seoul city was obtained from publicly available government databases which were also limited in variety. As such, we were unable to consider the contribution of many variables in our analysis. Future studies therefore should assess the contribution of several other variables on UHI.

Another limitation of our study deals with the radius at which the land use variations were assessed. The variations in land use were assessed in a 1-km radius within each AWS. However, using shorter distances (100 m–500 m) is likely to provide better results than those presented in the current study. Future studies therefore should consider using shorter distances when analyzing variations in land use.

Also, the data used to assess the influence of surface albedo and spectral radiance on UHI was based on data collected over a 4-year period (2013–2014). This is because the satellite used to collect "surface albedo" and "spectral radiance" data only started operating in 2013. To reveal further correlations among the said variables therefore, we suggest that future research studies use a larger dataset containing spectral radiance and surface albedo data collected over a longer period of time.

Furthermore, as we explained earlier, we develop two separate models for predicting UHI; a temporal UHI-predictive model and a spatial UHI-predictive model. The purpose for developing two separate models was so that we provide a model (i.e. spatial model) that explains how urban features affect the manifestation of UHI in urban areas and a separate model (i.e. temporal model) that explains how meteorological elements affect the manifestation of UHI. However, a model that explains the behavior of UHI using a combination of urban factors and meteorological elements is still necessary. Such a model would require that all the variables used in the training process be collected over the same time period, especially if supervised learning methods are to be employed in model development. This is a major reason that prevented the current study from developing a model that considered both the temporal and spatial factors influencing UHI; the available spatial and temporal data were recorded over different periods. Future research studies should endeavor to obtain spatial and temporal data collected over the same period of time and subsequently develop models able to explain the behavior of UHI based on a combination of both the meteorological and urban texture elements.

## Conclusion

UHI causes several issues related to health, energy consumption, comfort, etc., in urban areas. Also, the increase in urban development projects as well as the effects of climate change are likely to result in increased UHI in the future. Consequently, to mitigate the effects of UHI, proper and accurate tools to analyze and predict UHI behavior are of paramount importance. In the current study, we show that deep learning methods are useful tools that can be used to accurately analyze and predict UHI behavior in urban areas. We also show to what extent different urban texture elements are correlated with UHI. Such information can be leveraged by policymakers when drafting mitigation plans for UHI. Furthermore, the new index, UHI-hours, developed in the present paper provides a more accurate measure of the cumulative effects of UHI than other commonly used indices such as UHI intensity.

## Methodology

**Meteorological data.**  The meteorological data used in this study were obtained from 54 AWSs in Seoul, of which 28 are operated by the KMA and 26 are operated by the Seoul city office as presented Fig. 6. The AWSs collect meteorological data in 1-min intervals. For the present study, we gathered air-temperature, wind-direction, and wind-velocity data measured at the 54 AWSs for a period of 9 years.

**Physical and environmental factors.**  The factors that contribute to UHI effects have been reported in several previous studies[19,38–49]. These factors can be grouped into two categories, physical and environmental. In our study, data for both environmental and physical factors were gathered within a 1-km radius of each AWS.

Physical factors include sky view factor[38], total floor areas of entire buildings[39], area covered with green vegetation[40], building footprints, area covered with water and crops, and bare land area[41–43]. In our study, we obtained information related to physical factors from publicly available data on land-use patterns and building-management systems from a Geographic Information System (GIS) data from ministry of Land, Infrastructure and Transport Korea[38,39]. Environmental factors include spectral radiance[44,45] and surface albedo[45–47]. We collected data related to environmental factors from the metadata of the Landsat 8 remote-sensing satellite[46,50].

Landsat 8 (Operational Land Imager and Thermal Infrared Sensor) captures images with an approximate extent of 170 km north–south by 185 km east–west[50]. It takes about 15 days to circumnavigate the entire globe, and thus, returns to the same area about every 16 days[51]. Data from the Landsat satellite are easy to obtain and have been reported to be accurate[52]. Therefore, we used data collected by the Landsat 8 satellite for the period between 2014 and 2017 in 32-day intervals. However, the Landsat 8 satellite metadata for Seoul contain data for

**Figure 6.** Locations of the 54 Automatic Weather Stations (AWS) in Seoul city. (Figure generated using ArcMap, version 10.6.1, https://desktop.arcgis.com/en/arcmap/).

the entire Seoul area, whereas we needed spectral-radiance which is the amount of radiant flux from the sun per unit surface on land considering the spectrum of light from the sun[53] and surface-albedo data for a 1-km radius from each AWS. To obtain such urban radiance data corresponding to each AWS, we used a formula provided by the USGS[52]:

$$L_\lambda = L_{\min(\lambda)} + \left(L_{\max(\lambda)} - L_{\min(\lambda)}\right) * Q_{DN}/Q_{max} \tag{1}$$

where $L_\lambda$ is the spectral radiance received by the sensor, $Q_{max}$ is the maximum DN of the IR bnad and $L_{\max(\lambda)}$ and $L_{\min(\lambda)}$ are the maximum and minimum detected spectral radiance[53].

Surface albedo was computed from Landsat metadata, after converting DN value to the Top of atmosphere(TOA) reflectance values, the values of surface albedo ranging from 0 to 1

$$\rho\,\lambda' = M_\rho Q_{cal} + A_\rho \tag{2}$$

TOA reflectance was computed from Landsat digital number(DN) data. $\rho\lambda'$ is TOA planetary reflectance, $M_\rho$ is band-specific multiplicative rescaling factor, $Q_{cal}$ is quantized and calibrated standard product pixel values, $A_\rho$ is band-specific additive rescaling factor[54].

$$\alpha = \frac{((0.356 * \rho2) + (0.130 * \rho4) + (0.373 * \rho5) + (0.085 * \rho6) + (0.072 * \rho7) - 0.018)}{1.016} \tag{3}$$

**Concept of UHI-hours.** We developed a new assessment metric to quantify the UHI effect. The proposed metric, UHI-hours, follows the same concept as the "degree days" coefficient[55] used to quantify the energy demand in buildings. Degree days are calculated as the absolute difference between a reference temperature and the daily average temperature and represent a versatile climatic indicator commonly used in building energy performance analysis[56]. UHI-hours indicates to what degree the average temperature differs from the expected temperature (i.e., design temperature), and consequently, how much energy is demanded. Degree days can also be used to measure differences in regional temperature profiles[57] or to UHI-hours estimate to what degree the ambient-temperature difference between a given urban area and a rural area differs from 0. In our study, we selected the reference rural area based on a previous study on UHI in Seoul (i.e. Neunggok station). Neunggok is at the same latitude with Seoul and meets the criteria dictated by the World Meteorological Organization (WMO)[58] for choosing reference rural environments in UHI studies. Ultimately, the UHI-hours concept can also be used to gauge cumulative UHI intensity from season to season UHI-hours, thus allowing for a more efficient quantification of UHI for a small area or for a given season. UHI-hours is calculated by
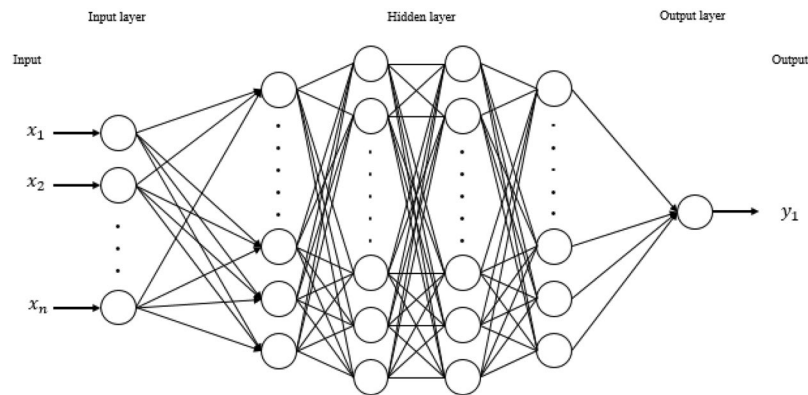
**Figure 7.** Schematic diagram of an Artificial Neural Network (ANN) architecture.

$$UHI-hours = \sum_{i=1}^{n} t_k (T_{urban} - T_{rural}) hour \tag{4}$$

where $T_{urban}$ and $T_{rural}$ are the average air temperatures of urban and rural environments, respectively, and constant term $t_k = 1$ if $(T_{urban} - T_{rural}) > 0$ and $t_k = 0$ if $(T_{urban} - T_{rural}) \leq 0$.

**Deep-learning methods.** We developed two deep-learning models that predict UHI distribution—a temporal and a spatial UHI model. The difference between the two lies in the input variables used to explain UHI in each model. The input variables used to develop the temporal UHI model are mainly time-dependent meteorological factors (i.e. date, time, air temperature, wind speed, and wind direction) whereas the spatial UHI model is based on the physical and environmental factors (i.e. sky view factor, total floor areas of entire buildings, area covered with green vegetation, building footprints, area covered with water and crops, bare land area, spectral radiance and surface albedo). The deep-learning algorithm used to develop the models is based on a multilayer feedforward ANN concept[59]. A typical multilayer feedforward ANN model is composed of several layers of interconnected units, called neurons (Fig. 7). In the prediction model, the weighted combination of input signals (represented by $\omega_i$ and $x_i$) is aggregated, and then, the output signal $f(\alpha)$ is transmitted by the connected basic neuron, as

$$\alpha = \sum_{i=1}^{n} \omega_i x_i + b \tag{5}$$

where $\omega_i$ is the weight, $x_i$ is the dependent variable, $b$ is a constant, and $\alpha$ is the output variable.

In the feedforward system, the bias units are included in each of the non-output layers of the network. The weights that connect neurons and bias with other neurons affect the output of the whole network. Learning occurs when these weights are adjusted to minimize errors in the results. In our models, the loss function is based on the mean square error value (Eq. (5)), and the objective in each training example ($j$ in Eq. (5)) is to minimize the loss function. Basically, $W_i$ denotes the weight matrix connecting layers $i$ and $i + 1$ for a network of $N$ layers. $B_i$ also denotes the column vector of biases for layer $i + 1$. $t^{(i)}$ and $o^{(j)}$ are the predicted and actual values, respectively, for training example $j$. The activation function used is the rectifier function with dropout; its role is to ensure that each neuron chooses the largest output from separate channels with its own weights and bias values. During the process of model training, we used several activation functions to see which one provided optimum model performance. Our results showed the rectifier function to provide the best model performance. This outcome is also in line with previous studies[60] which report the rectifier function as being superior to other activation functions.

For the temporal model, hourly data of ambient temperature, wind direction, and wind speed measured during 2009–2017 were used as features in the development of the prediction model. The data from 2014 to 2015, when the land-cover data were also determined, were used as the validation dataset, and the remaining data (2009–2013 and 2016–2017) were used as the training dataset. While developing the model, the number of layers and neurons were consistently increased until there was no further significant change in $R^2$ and root-mean-square error (RMSE) values. Thus, a total of five hidden layers were used, and the number of neurons was between 200 and 400 per layer. For the spatial model, the urban spatial data used to train the model (i.e., physical and environmental factors) was collected in 2014, and the environmental data (i.e., surface albedo and spectral radiance) derived from the same dataset. This prediction model was constructed by identifying the numbers of layers and neurons that produced the maximum possible $R^2$ and the lowest possible RMSE value, as was also done for the temporal UHI model discussed above. Accordingly, a total of three layers were used and the number of neurons was between 200 and 400 per layer.

$$L(W, B|j) = \frac{1}{2} \left\| t^{(i)} - o^{(j)} \right\|_2^2 \tag{6}$$

$$L'(W, B|j) = L(W, B|j) + \lambda_1 R_1(W, B|j) + \lambda_2 R_2(W, B|j) \tag{7}$$

To avoid overfitting, we used regularization variables, $\ell_1$ (L1: Lasso) and $\ell_2$ (L2: Ridge), to modify the loss function, and thus, minimize loss. For $\ell_1$ regularization (Eq. (6)), $R_1(W, B|j)$ is the sum of all $\ell_1$ norms for the weights and biases in the network, and $\ell_2$ regularization $R_2(W, B|j)$ is the sum of squares of all weights and biases in the network. The network weights for $\ell_2$ regularization are scaled toward 0. In addition, the algorithms share the same global parameters but learn different models through each training session. Therefore, we used the dropout function to allow an exponentially large number of models to be averaged as an ensemble. This reduces the likelihood of overfitting and improves model generalization.

**Time-series modeling.** To compare the previous studies[61,62], which used time-series modeling to predict UHI, with our study, we used the same data to develop a time-series model. We applied an autoregressive integrated moving average (ARIMA), a class of statistical models used for forecasting time-series data. In ARIMA, the variable of interest is assumed to be a function of its past observations and random errors. The model is categorized as ARIMA ($p, d, q$), where $p$ denotes the autoregressive parts of the dataset, $d$ refers to integrated parts of the dataset, and $q$ represents moving-average parts of the dataset. A typical ARIMA model[32] can be represented as

$$
\begin{aligned}
y_t =\ & \theta_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} \\
& + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q},
\end{aligned} \tag{8}
$$

where $y_t$ is our variable of interest; $\varepsilon_t$ is a random error at specific time $t$; $\phi_1, \phi_2, \ldots \phi_p$ and $\theta_0, \theta_1, \theta_2, \ldots \theta_q$ are model parameters; $p$ and $q$ are all non-negative integers; and errors $\varepsilon_t$ are assumed to have a mean of zero and constant variance.

## Data availability

The datasets generated and analysed during the current study are available from the corresponding author on reasonable request and with permission of Korea Meteorological Administration and Seoul Metropolitan Government.

## References

1. Jansson, Å. Reaching for a sustainable, resilient urban future using the lens of ecosystem services. *Ecological Economics* **86**, 285–291 (2013).
2. Bernard, J., Musy, M., Calmet, I., Bocher, E. & Keravec, P. Urban heat island temporal and spatial variations: Empirical modeling from geographical and meteorological data. *Building and Environment* **125**, 423–438 (2017).
3. Santamouris, M. Analyzing the heat island magnitude and characteristics in one hundred Asian and Australian cities and regions. *Science of The Total Environment* **512–513**, 582–598 (2015).
4. Basu, R. Relation between Elevated Ambient Temperature and Mortality: A Review of the Epidemiologic Evidence. *Epidemiologic Reviews* **24**, 190–202 (2002).
5. Doyon, B., Bélanger, D. & Gosselin, P. The potential impact of climate change on annual and seasonal mortality for three cities in Québec, Canada. *International Journal of Health Geographics* **7**, 23 (2008).
6. Basu, R. High ambient temperature and mortality: a review of epidemiologic studies from 2001 to 2008. *Environmental Health* **8** (2009).
7. Paravantis, J., Santamouris, M., Cartalis, C., Efthymiou, C. & Kontoulis, N. Mortality Associated with High Ambient Temperatures, Heatwaves, and the Urban Heat Island in Athens, Greece. *Sustainability* **9**, 606 (2017).
8. Santamouris, M. *et al.* On the impact of urban climate on the energy consumption of buildings. *Solar Energy* **70**, 201 (2001).
9. Doulos, L., Santamouris, M. & Livada, I. Passive cooling of outdoor urban spaces. The role of materials. *Solar Energy* **77**, 231–249 (2004).
10. Lindberg, F. *et al.* Urban Multi-scale Environmental Predictor (UMEP): An integrated tool for city-based climate services. *Environmental Modelling & Software* **99**, 70–87 (2018).
11. Taleghani, M., Sailor, D. & Ban-Weiss, G. A. Micrometeorological simulations to predict the impacts of heat mitigation strategies on pedestrian thermal comfort in a Los Angeles neighborhood. *Environmental Research Letters* **11**, 024003 (2016).
12. Chun, B. & Guldmann, J.-M. Spatial statistical analysis and simulation of the urban heat island in high-density central cities. *Landscape and Urban Planning* **125**, 76–88 (2014).
13. Su, Y.-F., Foody, G. M. & Cheng, K.-S. Spatial non-stationarity in the relationships between land cover and surface temperature in an urban heat island and its impacts on thermally sensitive populations. *Landscape and Urban Planning* **107**, 172–180 (2012).
14. Mihalakakou, G., Santamouris, M., Papanikolaou, N., Cartalis, C. & Tsangrassoulis, A. Simulation of the Urban Heat Island Phenomenon in Mediterranean Climates. *Pure and Applied Geophysics* **161**, 429–451 (2004).
15. Mihalakakou, G., Flocas, H. A. & Santamouris, M. The impact of synoptic-scale atmospheric circulation on the urban heat island effect over athens, Greece. 6.
16. Gobakis, K. *et al.* Development of a model for urban heat island prediction using neural network techniques. *Sustainable Cities and Society* **1**, 104–115 (2011).
17. Mhaskar, H., Liao, Q. & Poggio, T. When and Why Are Deep Networks Better than Shallow Ones? 7.
18. Zakšek, K. & Oštir, K. Downscaling land surface temperature for urban heat island diurnal cycle analysis. *Remote Sensing of Environment* **117**, 114–124 (2012).
19. Zhang, H. *et al.* Analysis of land use/land cover change, population shift, and their effects on spatiotemporal patterns of urban heat islands in metropolitan Shanghai, China. *Applied Geography* **44**, 121–133 (2013).
20. Lee, Y. Y., Kim, J. T. & Yun, G. Y. The neural network predictive model for heat island intensity in Seoul. *Energy and Buildings* **110**, 353–361 (2016).
21. Kim, Y.-H. & Baik, J.-J. Maximum Urban Heat Island Intensity in Seoul. *Journal of Applied Meteorology* **41**, 651–659 (2002).
22. Kim, Y.-H. & Baik, J.-J. Spatial and Temporal Structure of the Urban Heat Island in Seoul. *Journal of Applied Meteorology* **44**, 591–605 (2005).
23. Kim, Y.-H. & Baik, J.-J. Daily maximum urban heat island intensity in large cities of Korea. *Theoretical and Applied Climatology* **79**, 151–164 (2004).

24. Rizwan, A. M., Dennis, L. Y. C. & Liu, C. A review on the generation, determination and mitigation of Urban Heat Island. *Journal of Environmental Sciences* **20**, 120–128 (2008).
25. Memon, R. A., Leung, D. Y. C. & Liu, C.-H. An investigation of urban heat island intensity (UHII) as an indicator of urban heating. *Atmospheric Research* **94**, 491–500 (2009).
26. Gedeon, T. D. Data Mining of Inputs: Analysing Magnitude and Functional Measures. *International Journal of Neural Systems* **08**, 209–218 (1997).
27. Pakarnseree, R., Chunkao, K. & Bualert, S. Physical characteristics of Bangkok and its urban heat island phenomenon. *Building and Environment* **143**, 561–569 (2018).
28. Tan, J. *et al*. The urban heat island and its impact on heat waves and human health in Shanghai. *International Journal of Biometeorology* **54**, 75–84 (2010).
29. Li, H. *et al*. Interaction between urban heat island and urban pollution island during summer in Berlin. *Science of The Total Environment* **636**, 818–828 (2018).
30. Wang, C., Grozev, G. & Seo, S. Decomposition and statistical analysis for regional electricity demand forecasting. *Energy* **41**, 313–325 (2012).
31. Chen, Y., Tan, H. & Berardi, U. Day-ahead prediction of hourly electric demand in non-stationary operated commercial buildings: A clustering-based hybrid approach. *Energy and Buildings* **148**, 228–237 (2017).
32. Zhang, G. P. Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing* **50**, 159–175 (2003).
33. Li, J. *et al*. Impacts of landscape structure on surface urban heat islands: A case study of Shanghai, China. *Remote Sensing of Environment* **115**, 3249–3263 (2011).
34. Weng, Q., Lu, D. & Schubring, J. Estimation of land surface temperature–vegetation abundance relationship for urban heat island studies. *Remote Sensing of Environment* **89**, 467–483 (2004).
35. Li, H., Harvey, J. T., Holland, T. J. & Kayhanian, M. The use of reflective and permeable pavements as a potential practice for heat island mitigation and stormwater management. *Environmental Research Letters* **8**, 015023 (2013).
36. Schatz, J. & Kucharik, C. J. Seasonality of the Urban Heat Island Effect in Madison, Wisconsin. *Journal of Applied Meteorology and Climatology* **53**, 2371–2386 (2014).
37. Dixon, P. G. & Mote, T. L. Patterns and Causes of Atlanta's Urban Heat Island–Initiated Precipitation. *Journal of Applied Meteorology* **42**, 1273–1284 (2003).
38. Kolokotroni, M. & Giridharan, R. Urban heat island intensity in London: An investigation of the impact of physical characteristics on changes in outdoor air temperature during summer. *Solar Energy* **82**, 986–998 (2008).
39. Giridharan, R., Ganesan, S. & Lau, S. S. Y. Daytime urban heat island effect in high-rise and high-density residential developments in Hong Kong. *Energy and Buildings* **36**, 525–534 (2004).
40. Alves, E. Seasonal and Spatial Variation of Surface Urban Heat Island Intensity in a Small Urban Agglomerate in Brazil. *Climate* **4**, 61 (2016).
41. Quan, J. *et al*. Multi-temporal trajectory of the urban heat island centroid in Beijing, China based on a Gaussian volume model. *Remote Sensing of Environment* **149**, 33–46 (2014).
42. Coseo, P. & Larsen, L. How factors of land use/land cover, building configuration, and adjacent heat sources and sinks explain Urban Heat Islands in Chicago. *Landscape and Urban Planning* **125**, 117–129 (2014).
43. Zheng, B., Myint, S. W. & Fan, C. Spatial configuration of anthropogenic land cover impacts on urban warming. *Landscape and Urban Planning* **130**, 104–111 (2014).
44. Chen, X.-L., Zhao, H.-M., Li, P.-X. & Yin, Z.-Y. Remote sensing image-based analysis of the relationship between urban heat island and land use/cover changes. *Remote Sensing of Environment* **104**, 133–146 (2006).
45. Tran, H., Uchihama, D., Ochi, S. & Yasuoka, Y. Assessment with satellite data of the urban heat island effects in Asian mega cities. *International Journal of Applied Earth Observation and Geoinformation* **8**, 34–48 (2006).
46. Stathopoulou, M. & Cartalis, C. Daytime urban heat islands from Landsat ETM+ and Corine land cover data: An application to major cities in Greece. *Solar Energy* **81**, 358–368 (2007).
47. Liang, S. Narrowband to broadband conversions of land surface albedo I Algorithms. *Remote Sensing of Environment* **26** (2000).
48. He, T. *et al*. Evaluating land surface albedo estimation from Landsat MSS, TM, ETM +, and OLI data based on the unified direct estimation approach. *Remote Sensing of Environment* **204**, 181–196 (2018).
49. Liang, S. *et al*. Narrowband to broadband conversions of land surface albedo: II. Validation. *Remote Sensing of Environment* **84**, 25–41 (2003).
50. Roy, D. P. *et al*. Landsat-8: Science and product vision for terrestrial global change research. *Remote Sensing of Environment* **145**, 154–172 (2014).
51. Li, J. & Roy, D. A Global Analysis of Sentinel-2A, Sentinel-2B and Landsat-8 Data Revisit Intervals and Implications for Terrestrial Monitoring. *Remote Sensing* **9**, 902 (2017).
52. Storey, J., Choate, M. & Lee, K. Landsat 8 Operational Land Imager On-Orbit Geometric Calibration and Performance. *Remote Sensing* **6**, 11127–11152 (2014).
53. Abutaleb, K. *et al*. Assessment of Urban Heat Island Using Remotely Sensed Imagery over Greater Cairo, Egypt. *Advances in Remote Sensing* **04**, 35–47 (2015).
54. Hamoodi, M. N., Corner, R. & Dewan, A. Thermophysical behaviour of LULC surfaces and their effect on the urban thermal environment. *Journal of Spatial Science* **64**, 111–130 (2019).
55. van Asseldonk, M. A. P. M. Insurance against weather risk: Use of heating degree-days from non-local stations for weather derivatives. *Theoretical and Applied Climatology* **74**, 137–144 (2003).
56. De Rosa, M., Bianco, V., Scarpa, F. & Tagliafico, L. A. Heating and cooling building energy demand evaluation; a simplified model and a modified degree days approach. *Applied Energy* **128**, 217–229 (2014).
57. Sailor, D. J. Simulations of annual degree day impacts of urban vegetative augmentation. *Atmospheric Environment* **32**, 43–52 (1998).
58. Oke, T. R. Instruments and Observing Methods Report No. 8. 51.
59. He, K., Zhang, X., Ren, S. & Sun, J. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. in *2015 IEEE International Conference on Computer Vision (ICCV)* 1026–1034, https://doi.org/10.1109/ICCV.2015.123 (IEEE, 2015).
60. Schmidhuber, J. Deep learning in neural networks: An overview. *Neural Networks* **61**, 85–117 (2015).
61. Van Weverberg, K., De Ridder, K. & Van Rompaey, A. Modeling the Contribution of the Brussels Heat Island to a Long Temperature Time Series. *Journal of Applied Meteorology and Climatology* **47**, 976–990 (2008).
62. Mathew, A., Sreekumar, S., Khandelwal, S., Kaul, N. & Kumar, R. Prediction of surface temperatures for the assessment of urban heat island effect over Ahmedabad city using linear time series model. *Energy and Buildings* **128**, 605–616 (2016).

## Acknowledgements

### Author contributions

G.Y.Y. and M.S. designed the study; J.W.O. and G.Y.Y. collected the data; J.W.O., J.N., G.Y.Y. and M.S. analyzed the data; J.W.O., J.N. and P.N.D. made the figures; J.W.O., J.N., P.N.D., G.Y.Y. and M.S. drafted and revised the paper; all authors approve the final version of the manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

**Correspondence** and requests for materials should be addressed to G.Y.Y.

**Reprints and permissions information** is available at www.nature.com/reprints.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.