

# Explainable heat-related mortality with random forest and SHapley Additive exPlanations (SHAP) models

Yesuel Kim, Youngchul Kim\*

*KAIST Urban Design Lab, Department of Civil and Environmental Engineering, Korea Advanced Institute of Science and Technology, 291 Daehak-ro Yuseong-gu, Daejeon 34141, Republic of Korea*

## ARTICLE INFO

**Keywords:**  
Climate change  
Heat waves  
Human health  
Random forest  
Grid search  
SHAP

## ABSTRACT

The heat increase caused by climate change has worsened the urban heat environment and damaged human health, which has led to heat-related mortality. One of the most important ways to respond to heat-related damage is to develop effective forecasting tools. However, accurately predicting heatwave damage is difficult in regions in a city with different conditions. Damage due to extreme heat can be evaluated differently in each region, as climatic, demographic and socioeconomic sectors are diversely distributed across local areas. In this study, we develop a random forest-based model for estimating the occurrence of heat-related mortality in a detailed spatial unit within a city. Through hyperparameter optimization, the model yielded accuracy, F1-score and AUC values of 90.3%, 94.75%, and 86%, respectively. The estimation results of the model were interpreted from the global and local perspectives by introducing the latest SHAP method. As a result of interpretation, demographic, socioeconomic and climatic sectors were determined to contribute the most to the estimation process. This is the first study of partial scenarios through the development and interpretation of a spatial unit machine learning-based occurrence estimation model for heat-related mortality.

## 1. Introduction

The acceleration of urbanization has led to environmental degradation and rapid changes in the urban climate (Ali, Bakhsh & Yasin, 2019; Yang et al., 2020a). Heat waves occur as a result of climate change, and their frequency and intensity are expected to continuously increase (Keramitsoglou et al., 2013; van Loenhout et al., 2021; Yang, Yu, Jørgensen & Vejre, 2020b). Heatwaves are likely to increase globally due to the acceleration of climate change, which poses potential risks to the environment, society and public health (Hatvani-Kovacs, Belusko, Skinner, Pockett & Boland, 2016; Zhang et al., 2020). Therefore, finding a balance between urban expansion and the quality of the heat environment is key to ensuring sustainable urban formation (Yang et al., 2020a). The extreme heat environment caused by climate change in cities is important because it directly affects human health and is particularly fatal for vulnerable groups, even leading to death (Green, Andrews, Armstrong, Bickler & Pebody, 2016). The effect of heat waves on mortality and morbidity in urban populations due to enhanced climate change has become significant in the public health sector (Bakhsh, Rauf & Zulfiqar, 2018; Kim, Deo, Park, Lee & Lee, 2019a; Royé, Codesido, Tobías & Taracido, 2020; Samaniego et al., 2017). The

proportional increase in mortality due to excessive heat has gradually and nonlinearly increased (Oechslin & Buechley, 1970). Therefore, effective predictions of the health effects of heat waves can help protect the public and reduce heat-related health damage. However, knowledge and readiness within health care and health systems to address the effects of heat waves are currently lacking in most countries around the world (Ebi & Schmier, 2005). To build an adaptation system for heatwaves, it was important to analyze the determinants of impacts on people's cognition (Rauf et al., 2017). In addition, recent heat wave warning systems lack verification and have limitations in providing accurate health hazard warnings (Li, Ding, Sun, Zhang & Kinney, 2016). Moreover, the existing studies on health impact prediction in the field of heat waves have provided insufficient accuracy due to the uncertainty of modeling parameters (Samaniego et al., 2017).

Considering social issues in the urban heat environment, a number of previous studies have analyzed the effects of various indicators related to heat waves on health damage (Anderson & Bell, 2011; Faye, Dème, Diouf, 2021; Fouillet et al., 2008; Royé et al., 2020; Smoyer, 1998). In particular, Faye et al. (2021) used a Poisson generalized additive model (GAM) and a distributed-lag nonlinear model (DLNM) to explore the effects of several heat-related indicators on the total number of heat-related deaths by age and sex. They considered various

\* Corresponding author.

E-mail addresses: [yesuel21@kaist.ac.kr](mailto:yesuel21@kaist.ac.kr) (Y. Kim), [youngchulkim@kaist.ac.kr](mailto:youngchulkim@kaist.ac.kr) (Y. Kim).



Nomenclature		
<i>MT</i>	Maximum temperature	$x_i$ Vector of input variables
<i>RH</i>	Relative humidity	<i>AUC</i> Area under the curve
<i>HI</i>	Heat index	<i>TP</i> True positive cases
<i>YE</i>	Young and elderly population	<i>FP</i> False positive cases
<i>AR</i>	Aging rate	<i>TN</i> True negative cases
<i>EAR</i>	Percentage of elderly individuals living alone (elderly individuals living alone rate)	<i>FN</i> False negative cases
<i>RBL</i>	Recipients of a basic livelihood	<i>f</i> Estimation (prediction) model
<i>PD</i>	Population density	<i>g</i> Linear function of binary features
<i>BD</i>	Building density	$z'_i$ Observed ( $z'_i = 1$ ) or unknown ( $z'_i = 0$ ) features
<i>HS</i>	Heat shelter	<i>M</i> Number of simplified input features
<i>EMI</i>	Emergency medical institution	$\phi_i$ Contribution of each feature
<i>MOR</i>	Heat-related mortality	$f_{S \cup \{i\}}$ Model when $i$ is included
$P_i$	Probability corresponding to class $i$	$f_S$ Model when $i$ is excluded
$l_i$	Class labels	$x_S$ Specific model input
		$F$ Set of all features
		$S$ Set of all features excluding $i$

demographic indicators in a region but did not consider other urban components, such as socioeconomic factors. Thus, analyses and estimations of the vulnerability of human health considering various urban parameters, such as socioeconomic factors, were rare in previous studies (Wang et al., 2019c).

Some studies have predicted the health damage caused by heat waves through various urban composition indicators by combining statistical or machine learning techniques. Park, Jung, Lee and Park (2020) used a random forest (RF) model to predict health damage. They macroscopically designated a test area of 17 administrative districts composed of 8 municipalities and 9 provinces in South Korea. Zhang, Li, Schwartz and O'Neill (2014) used an RF-based machine learning technique to identify the weather variable that had the greatest influence on the prediction of heat-related mortality in 4 cities in the United States. Kim et al. (2019a) developed a model for predicting the number of deaths using several statistical techniques, and various climatic and demographic indicators were used. Wang et al. (2019c) conducted a study to predict the occurrence of heatstroke using a machine learning algorithm. This study introduced various urban indicators, and the results were compared to those of other studies. Lee et al. (2018) analyzed the effect of summer weather types on daily excess mortality using data from 1991 to 2010 for 14 major cities in Korea. They developed a mortality prediction model using multiple stepwise regression techniques. Kodera et al. (2019) identified factors affecting the number of heat-related patients using the integrated computational technique of thermodynamics and thermoregulation in three major cities in Japan. They proposed a formula for estimating the number of daily patients with heat-related disease. Ikeda and Kusaka (2021) developed a prediction model of patients due to heatstroke using statistical and machine learning methods with 11 variable sets (temperature, relative humidity, wind speed, etc.) as inputs. To develop a high-accuracy model, they examined which combination of input variables performed most effectively in prediction. Although these studies have proposed a model for estimating the damage caused by heat waves, their analyses were only performed at the entire city or individual human scale and did not consider regions within a city (Ikeda & Kusaka, 2021; Kim et al., 2019a; Kodera et al., 2019; Lee et al., 2018; Wang et al., 2019c; Zhang et al., 2014). These studies are meaningful in that they predicted damage from heat waves but are limited in that they used a macroscopic city-scale approach and did not consider specific spatial areas within cities (Kim et al., 2019a; Wang et al., 2019c; Zhang et al., 2014). Ruttan, Stoltz, Jackson-Vance, Parks and Keim (2013) developed a model to predict the probability of at least one heat death occurring by performing logistic regression; they assessed the performance of the model (AUC of approximately 0.84), but the model was limited because factors other

than meteorological variables were not considered in the scope of the study. Hirano et al. (2021) performed mortality prediction for heat-related illnesses by applying machine learning. They used data from Japan's "Heatstroke study" to collect data for and study patients with heat-related diseases. These patient data consisted of 24 variables, such as age and sex, for individual populations and were used as inputs for the prediction model. This study is important because it applied a machine learning method for heat wave-related mortality estimation for the first time.

However, since that study was conducted based on individuals, generalized insight related to the occurrence of mortality and the corresponding variations in local spatial units, even within cities, was limited. Although some previous studies have provided predictions for cities and individuals in the health, forecasting and heat wave prediction fields, no study has attempted to estimate heat-related mortality based on urban indicators for various sectors in spatial units within a city.

In terms of urban planning and design, the heat distribution spatially varies in a city. Local climatic events have been reported to pose a serious threat to the built environment in cities (Pyrgou, Castaldo, Pisello, Cotana & Santamouris, 2017). In fact, according to the distribution of thermal environments in a city reported by Lee and Hong (2008), the temperature is generally higher close to the city center than in other areas, and the local temperature differences among adjacent areas are often large. Accordingly, even within a city, the number of deaths due to heat may vary locally, and a detailed spatial unit approach is needed beyond the comprehensive consideration of entire cities. Similarly, since the distributions of urban composition indicators that affect heat vulnerability in a city, such as socioeconomic and infrastructure variables, vary, it is necessary to consider the influence of various factors in different spatial units. Nishimura et al. (2021) proposed a framework for estimating the number of heat-related patients in "wards", a detailed administrative district unit of Japan, in Nagoya city by applying machine learning. Although they performed a regional approach to a detailed local administrative district unit in Nagoya City to predict the damage from the heat wave sector, they did not perform a comprehensive consideration of various components, such as the physical environment and social and economic sectors, influencing the degree of heat-related risk by region. Therefore, we developed a heat-related mortality estimation model based on various urban components related to heatwaves to assess detailed administrative districts in a city at a high resolution. Furthermore, in addition to an estimation model based on the RF algorithm, we adopted the SHapley Additive exPlanations (SHAP) technique, which supports various interpretations based on the contribution of each input variable in the model system. In research using machine learning techniques, model interpretation is

very important, but some models can be difficult to apply (Parsa, Movahedi, Taghipour, Derrible & Mohammadian, 2020). Recently, to interpret machine learning models, the SHAP model (Lundberg & Lee, 2017b), which is a unified framework, has been used (Futagami, Fukazawa, Kapoor & Kito, 2021; Mangalathu, Hwang & Jeon, 2020; Mokhtari, Higdon & Başar, 2019; Zeng, Davoodi & Topaloglu, 2020). Mokhtari et al. (2019) analyzed the priority of the contributions of all features from a global perspective through the SHAP values of a prediction model with a support vector machine (SVM) classifier. Zeng et al. (2020) performed a local interpretation of the estimation results of an RF prediction model; they provided force plots for several instances among the estimated results to provide an explanation of which index was dominant. As such, the SHAP method supports both the global analysis of estimations and the local analysis of detailed instances. In this study, a similar approach to the SHAP method is introduced, as it is considered suitable for developing index-based scenarios that can be practically referenced when evaluating heat wave damage in detailed spatial units and establishing management plans by region. This is the first study to demonstrate the application of SHAP techniques for the interpretation of machine learning estimation models developed in the heat wave field.

In summary, the human health damage related to excessive heat and differences in local heat conditions within a city can vary among neighboring areas. Several previous works in the heat wave field focused on comprehensive forecasting for entire cities as single large units. Since spatial models of heat-related mortality estimation mainly focused on regional-scale areas, it was difficult to identify local-scale characteristics about the occurrence of heat-related mortality within a city. Additionally, while some studies (e.g., Zhang et al., 2014) predicted damage in by heat wave, few studies have interpreted modeling results. Thanks to state-of-the-art techniques, such as SHAP, it is possible to interpret modeling results. Thus, to fill up this gap, this study seeks to develop an estimation model of heat-related mortality with relatively high-resolution spatial units and relevant model interpretation technique.

Therefore, in this study, an RF model that can estimate the occurrence of heat-related mortality in a detailed spatial unit within a city is developed, and various city sector indicators are included as input features to provide interpretations from various perspectives; then, this model is combined with the most recent SHAP model. Hyperparameter optimization is performed with the GridSearchCV method, and the performance of the optimally developed RF model is verified through a comparison with other RF classification algorithms in the same field. Following model development, an explanatory SHAP model compatible with the RF model is applied to provide a consistent explanation for the proposed heat-related mortality estimation model for spatial units. Each result is based on global and local interpretations. Based on the relevant contributions, the global analysis provides insight into the importance of, dependence on, and decision process for each feature of the model, and a SHAP summary plot, SHAP decision plot, and SHAP dependence plot are produced. In the local analysis, we determine which features contributed the most to each instance, and we obtain a SHAP force plot.

## 2. Materials and methods

### 2.1. Research scope

In this study, a spatial-level RF classification model was developed to estimate the occurrence of heat-related mortality using various urban indicators. We performed a validation study of a heat-related damage prediction model that predicts the probability of heat-related deaths in Daegu city, South Korea. We collected datasets of urban indicators related to heat vulnerability and heat-related mortality during the summer period of June to September from 2016 to 2018. We used an RF to model the probability of one or more heat deaths occurring in the summer season using various urban parameters as the model predictor.

The study procedure is shown in Fig. 1. First, the data were pre-processed. Data were grouped into training and testing datasets at proportions of 75% and 25%, respectively. Next, an RF algorithm was introduced for model building, and a grid search method was introduced to derive accurate model parameters. The final step consisted of model evaluation and interpretation. A performance evaluation was performed based on the confusion matrix and ROC curve results. Related indicators, such as accuracy, precision, recall, F1 score and AUC, were utilized to quantify the evaluation results. Finally, in the interpretation process, the concept of SHAP was introduced, and global and local analyses were performed.

### 2.2. Study area

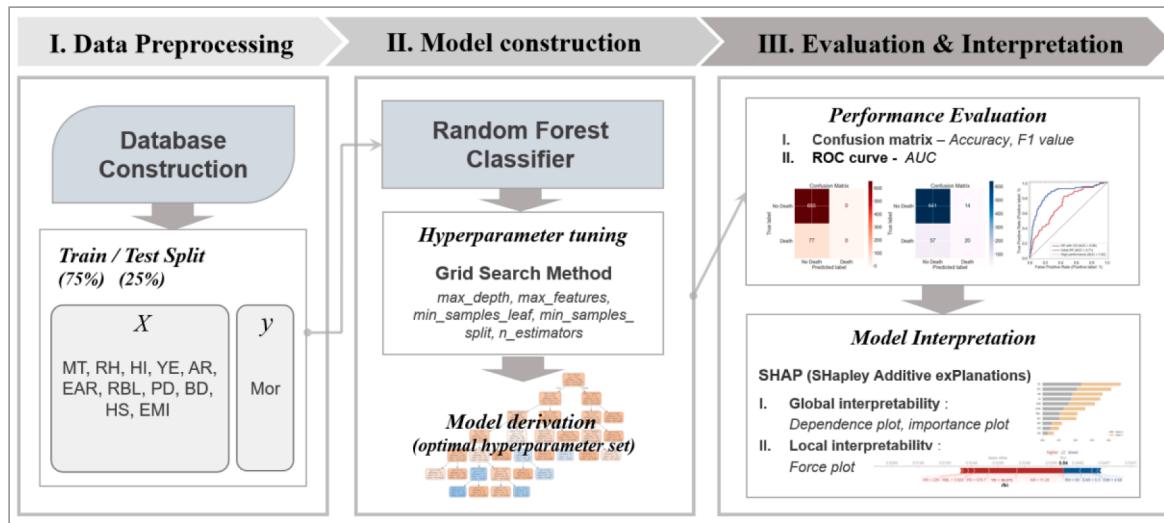
Daegu city in South Korea was designated as the test area for this study. Daegu is a representative metropolitan city located at 38°15' latitude and 128°34' longitude (an area of 885.56 km<sup>2</sup> and population of 2.43 million) in Fig. 2(a). The city is surrounded by mountains, which create an inland basin-type climate that is cold in winter and hot in summer. There is a large difference in temperature by season, and due to the basin-type topography, air circulation in the city is poor. Due to these geographical and environmental factors, Daegu is one of the most vulnerable cities to heat in South Korea. Notably, extreme heat waves were recorded in 2016 and 2018, and they caused many deaths (Lee, Min, Bae & Cha, 2020a). Fig. 2(b) shows the yearly average temperature and number of heat-related mortalities in Daegu from 2016 to 2018. According to the plot, in 2018, when the annual average temperature was the most extreme, the damage to human health was the most severe. In addition, the regression line over three years (dotted line) confirms that there is a positive correlation between climate and the number of deaths, reflecting heat-related mortalities. Fig. 2(c) shows the distribution of total heat-related mortalities between 2016 and 2018. According to Fig. 2(c), heat damage varies in detailed spatial units within the city.

### 2.3. Data collection

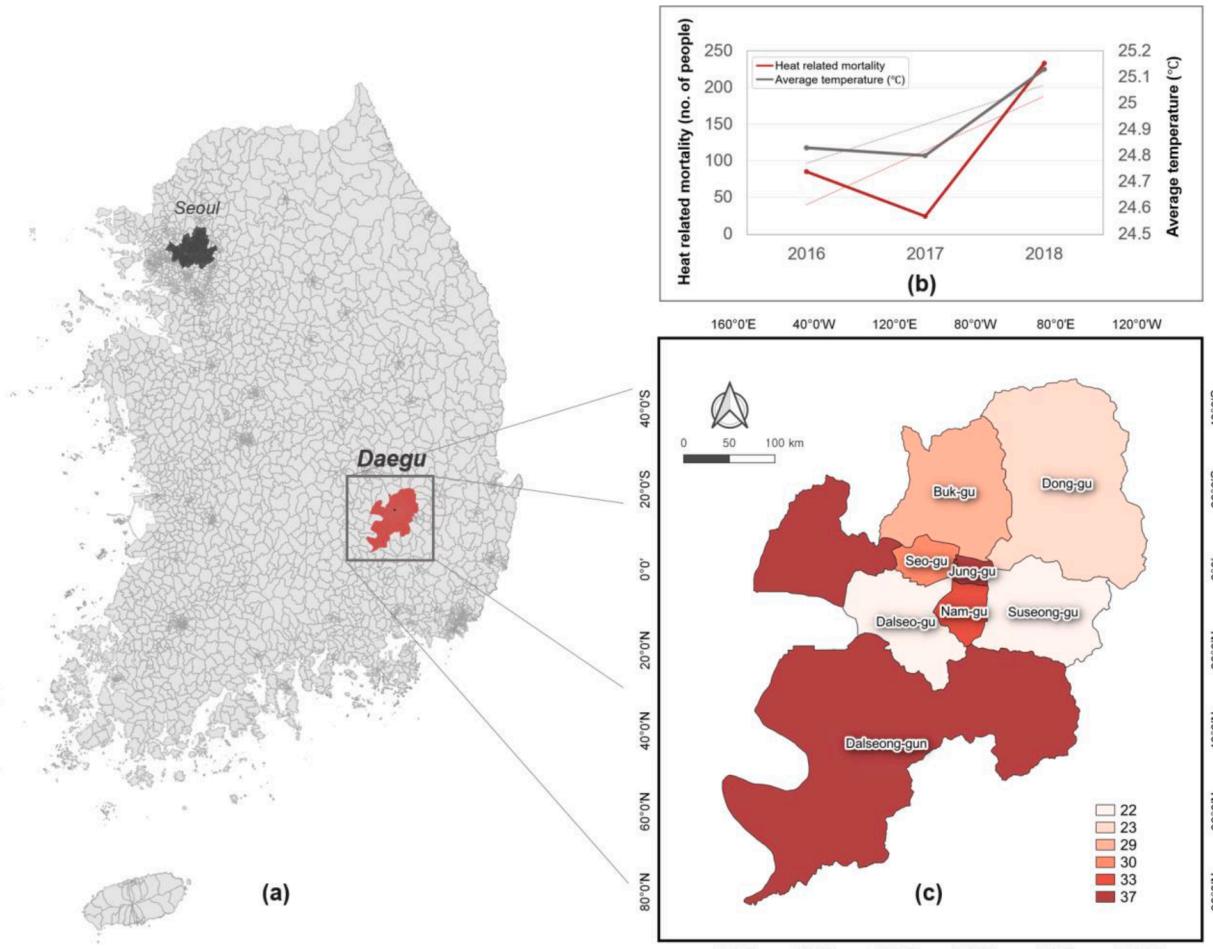
The training data for model estimation were derived from various open-source materials provided by public institutions in South Korea. All data were collected between 2016 and 2018. Since 2018 was the most recent year available in South Korea to choose as indicators for model estimation, we chose a time range of recent three years from 2016 to 2018 to construct data set. The spatial units of the data include 8 administrative districts in Daegu city, and a total of 2928 records with daily data were obtained for the summer season (June to September) in each administrative district. We chose June, July, August and September as the experimental period corresponding to summer season because the months from June to September in 2016, 2017 and 2018 showed the highest monthly average temperature according to the monthly average temperature provided by the Korea Meteorological Administration (KMA). Data for 12 features corresponding to the meteorological, demographic, socioeconomic, infrastructure, and human health sectors were used for model development. Heat-related mortality was used as the output variable, and the other 11 features were used as input variables. All the features generated and used in this work and brief descriptions of these features are given in Table 1.

#### 2.3.1. Meteorological features

We used the daily maximum temperature (MT), relative humidity (RH) and heat index (HI) as indicators for the meteorological sector. Heat-related diseases are induced in extremely hot and humid environments (Hirano et al., 2021). RH contributes to health damage from heat waves as much as climate factors do (Montero, Miron, Criado, Linares & Díaz, 2013). In addition, the heat wave index is defined as a measure of discomfort that combines temperature and RH and has been used to define heat waves in other studies (Xu, FitzGerald, Guo, Jalaludin & Tong, 2016). Therefore, we utilized weather data (the air



**Fig. 1.** Research flowchart. RH, relative humidity; YE, young and elderly; HI, heat index; AR, aging ratio; MT, max. temp.; EMI, emergency medical institute; EAR, elderly alone rate; RBL, recipients of a basic livelihood; BD, building density; PD, population density; HS, heat shelter; MOR, heat-related mortality.



**Fig. 2.** Location of and heat-related mortality trends in the study area. (a) Location of Daegu city (red area), South Korea. (b) Summer seasonal average temperature and total mortality count in Daegu city from 2016 to 2018. The straight line depicts the linear regression of the mean temperature data over different years. (c) Distribution of total heat-related mortalities in 2018.

temperature and RH were extracted) provided by the Korea Meteorological Administration (KMA). The time unit of the raw data was UTC (Coordinated Universal Time, Korean Standard Time –9), and the data

were reprocessed according to the situation in South Korea. The HI was calculated and used based on a formula consisting of temperature and humidity (Rothfusz & Headquarters NWS Southern Region, 1990).

**Table 1**  
Descriptions of the model variables and data sources.

Features	Abbreviations	Units	Data sources
<b>Input</b>			
<b>Meteorological features</b>			
Max. temperature	M T	°C	Korea Meteorological Administration
Relative humidity	RH	%	Korea Meteorological Administration
Heat index	HI	°C	Korea Meteorological Administration
<b>Demographic and socioeconomic features</b>			
Young and elderly population	YE	Person	Daegu statistical information
Aging rate	AR	%	Daegu statistical information
Elderly living alone rate	EAR	%	Daegu statistical information
Recipients of a basic livelihood	RBL	Person	Korean statistical information service
Population density	PD	Person/m <sup>2</sup>	Daegu statistical information
<b>Infrastructure features</b>			
Building density	BD	m	National Spatial Data Infrastructure Portal
Heat shelter	HS	Count	Data Portal, Ministry of the Interior and Safety (MOIS)
Emergency medical institution	EMI	Count/Person	National emergency medical center
<b>Target/Output Human health features</b>			
Heat-related mortality	MOR	Count	Micro Data Integrated Service (MDIS)

Based on the temperature (t) and relative humidity (r), the HI is given by Eq. (1).

$$HI(\text{heat index}) = (-42.379) + (2.04901523 \times t) + (10.14333127 \times r) - (0.22475541 \times t \times r) \\ - ((6.83783 \times 10^{-3}) \times t^2) - ((5.481717 \times 10^{-2}) \times r^2) + ((1.22874 \times 10^{-3}) \times t^2 \times r) + \\ ((8.5282 \times 10^{-4}) \times t \times r^2) - ((1.99 \times 10^{-6}) \times t^2 \times r^2) \quad (1)$$

### 2.3.2. Demographic and socioeconomic features

We used the young and elderly population (YE), aging rate (AR), elderly living alone rate (EAR), recipients of a basic livelihood (RBL), and population density (PD) features for the demographic and socioeconomic sector. Population density is a measure of how many people live in a spatial unit and is an indicator related to climate change vulnerability (Busby, Cook, Vizy, Smith & Bekalo, 2014; Kim & Jung, 2020). The degree of risk due to heat waves varies according to socioeconomic characteristics, such as age (Toloo et al., 2014). In particular, social isolation among the elderly population living in urban areas is associated with heat-related mortality (Kim, Lee, Kim & Cho, 2020). Raw demographic and socioeconomic sector data were used, as shown in Table 1, and open-source data from the Daegu statistical information and Korean statistical information services were obtained from a Korean public portal site. YE was determined by extracting the number of children under 5 years old and adults over 65 years old from resident registration demographic data provided by the Daegu statistical information service. AR was calculated using the same raw data as YE, and this feature was defined as the ratio of the population over 65 to the total

population. PD was defined as the total population divided by the land area ( $\text{m}^2$ ) of each administrative district. EAR was based on regional elderly population data provided by the Daegu statistical information service and was defined as the ratio of elderly individuals living alone to the total population in each administrative district. Finally, RBL was based on statistics regarding the basic livelihood recipients provided by the Korean statistical information service (National statistical portal).

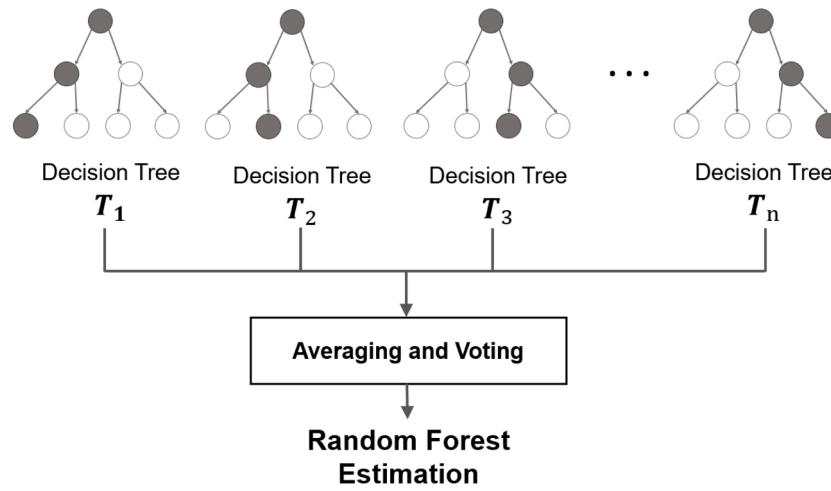
### 2.3.3. Infrastructure features

We used the building density (BD), heat shelter (HS), and emergency medical institution (EMI) features corresponding to the infrastructure sector. Buildings increase the risk of heat waves in cities. In fact, the relationship between heat-related indicators and the BD in cities has been explored (Song et al., 2020). Likewise, BD is often used as an indicator in risk assessments for heat waves (Kim, Kwon, Lee & Kim, 2019b). Integrated information-master data provided by the National Spatial Data Infrastructure Portal of the Ministry of Land, Infrastructure and Transport were used to calculate the BD in this study. The final BD data were processed by dividing the sum of the building gross floor area by the total land area of each administrative district. Some cities provide heat shelters as part of a response to heat waves (Guo & Hendel, 2018). In some cases, 24-hour shelter, nutrition and nursing services are provided (Applegate et al., 1981). We have defined the following types of facilities as HSs: financial institutions, senior facilities, community centers, welfare centers, public health centers, district offices, religious facilities, and other facilities classified as local shelters. We collected current HS location information from a domestic public data portal provided by the Ministry of the Interior and Safety (MOIS). A final HS dataset was constructed for each detailed study area. The number of days with excessive heat is correlated with emergency medical service accidents and the healthcare system load (Zottarelli, Sharif, Xu & Sunil, 2021). The number of emergency medical institutions per capita has been used as an evaluation index in the heat wave field (Jung, Kim & Park, 2018). Accordingly, we introduced the EMI indicator as an input feature, and the Emergency Medical Statistical Annual Report provided

by the National Emergency Medical Center was used to obtain raw data. The total counts of facilities that provide emergency medical services for each administrative district were extracted from the raw data, and the final EMI data were derived by dividing the resulting values by the total population of each administrative district.

### 2.3.4. Human health features

We used the number of heat-related deaths as a target feature in the estimation model developed in this study to represent damage from heat waves. High temperatures, which are associated with heat waves, affect the morbidity and mortality of cardiovascular diseases (CVDs) and the occurrence of thermal stress (Shin & Lee, 2014; Tian, Li, Zhang, Jaakkola & Guo, 2012). The effects of heat waves vary widely depending on the type of disease, and CVDs are particularly affected (Dong, Zeng, Ma, Li & Pan, 2016). Therefore, in this study, data related to mortality from hypertensive diseases, ischemic heart diseases, other heart diseases, and cerebrovascular diseases, which are subtypes of cardiovascular diseases, were considered. The raw dataset of the mortality data was obtained from the 'Demographic trend data' provided by the Micro Data Integrated Service (MDIS) online portal after processing.



**Fig. 3.** Diagram of a random forest composed of multiple decision trees.

#### 2.4. Model constructions

##### 2.4.1. Random forest classification

In this study, we estimated the occurrence of heat-related mortality ("death"/"no death" classes) using an RF model. Although various classification techniques exist such as Support Vector Machine (SVM), Naïve Bayes Classifier (NBC), and Back Propagation Neural Network (BPNN), the RF model demonstrated the best performance (Dong, Li & Xie, 2014; Liu, Wang, Wang & Li, 2013). In particular, unlike BPNN and SVM models, RF has the advantage of being good at dealing unbalanced, multiclass, and small sample data without data preprocessing process (Liu et al., 2013). The RF method is an ensemble learning technique that was developed by Breiman (2001). It is a very powerful machine learning method for classification and regression based on multiple decision trees (Hu et al., 2017; Mohammady, Pourghasemi & Amiri, 2019).

An RF generates multiple small decision trees (DTs) for arbitrary subsets and functions in Fig. 3. Each DT considers only a subset of the data, and each is responsible for a different trend, thus providing an individually biased predictive model. The final model estimate is determined by an average voting scheme among individual trees. Let  $h$  be a learning sample, where  $h = \{(x_1, l_1), (x_2, l_2) \dots (x_i, l_i)\}$ ; in this case,  $x_1, x_2 \dots x_i$  are values from the measurement vector, and  $l_1, l_2 \dots l_i$  are class labels.  $x_1$  can be viewed as a vector of input variables, and feature division is based on one of these variables. If  $P_i$  is a probability corresponding to class  $l_i$ , it is expressed as shown in Eq. (2) (Tangirala, 2020).

$$P_i = \frac{l_i}{h} \quad (2)$$

In the classification process, the attribute selection criteria support the selection of the optimal condition for each tree, and they include the Gini index and entropy-based information gain. In this study, the Gini index was used. The Gini index is a value between 0 and 1, with 0 indicating the most important information and 1 indicating the least-important information (Christo, Nehemiah, Brighty & Kannan, 2020). The Gini index is expressed as shown in Eq. (3).

$$\text{Gini index} = 1 - \sum_{i=1}^j (P_i)^2 \quad (3)$$

Each individual tree is individually a weak learner. However, an RF of trees is a strong learner. RF models provide low prediction bias compared to other ensemble methods because sampling is performed in the feature space (Alshraideh, Castillo & Gil Del Val, 2020). Moreover, the RF approach provides a measure of variable importance that reflects the predictive strength of each variable in the model itself. This

consideration yields results that are more interpretable than those normally produced by neural networks (Friedman, Hastie & Tibshirani, 2001; Hu et al., 2017). Random forests have displayed better predictive performance than other classification and regression tree methodologies and can automatically handle missing values and combinations of binary and continuous variables. In addition, the RF method considers complex interactions among highly correlated predictors and can reduce prediction errors better than conventional regression tree methods (Breiman, 2001; Zhang et al., 2014).

The development of the RF model was performed using Python 3.8 programming language. The implementation of the RF model was performed using the scikit-learn library, which is an open-source machine learning tool. The RF classifier included in the ensemble module of the scikit-learn library was used for all experimental processes. The `train_test_split()` method from the scikit-learn library was used to divide the datasets into training and testing sets (training:testing ratio of 75:25). The performance of the estimation models was compared in terms of the testing accuracy, F1-score and AUC. The scikit-learn library was also used to derive various model performance plots for the RF model. The entire process was implemented on a personal computer (Intel Core i7-8700 CPU with a 3.20 GHz processor and 16 GB of RAM).

##### 2.4.2. Hyperparameter optimization: GridSearchCV

We obtained the parameters corresponding to the estimated model features by applying an exhaustive parameter search algorithm called GridSearchCV. Python's GridSearchCV function helps find parameters with the best model estimation accuracy (Lu et al., 2021). We derived the optimal RF hyperparameter combination by specifying the values and ranges of the hyperparameters of the RF estimation model and applying the GridSearchCV function. The RF with the GS method provides high estimation accuracy when `n_estimators` is set to 100, `max_features` is set to 2, `min_samples_leaf` is set to 1, `min_samples_split` is set to 9 and `max_depth` is set to 10. Table 2 shows the tuning parameters considered in GridSearchCV.

**Table 2**  
Tuning parameters considered in GridSearchCV.

Tuning parameters	Values and ranges
Bootstrap	True, False
Max_depth	[1, 10]
Max_features	[1, 10]
Min_samples_leaf	[1, 10]
Min_samples_split	[1, 10]
N_estimators	100, 200, 300, 500, 700, 1000

**Table 3**  
Confusion matrix.

Conditions for each sector		Predicted values	
		Positive	Negative
Actual values	Positive	True positive (TP)	False negative (FN)
	Negative	False positive (FP)	True negative (TN)

## 2.5. Model performance evaluation

In general, the most commonly used criterion for model evaluation is accuracy, which may not be the best metric depending on the characteristics of the data. In terms of performance evaluation, it is better to introduce other scoring criteria, such as the area under the curve (AUC), which are less prone to overfitting (Gómez-Ramírez, Ávila-Villanueva & Fernández-Blázquez, 2020). Therefore, in this study, three measures were introduced for model evaluation: accuracy, F1-score and AUC. Upon completion of the classification process, each sample was assigned to one of four categories: true positive (TP), false positive (FP), true negative (TN), or false negative (FN). The four classes are expressed in a confusion matrix, as shown in Table 3, where TP represents the number of positive and correctly identified cases and FN represents the number of positive cases that are misclassified as negative. TN represents the number of cases identified as negative and correctly identified, and FP represents the number of negative cases that are misclassified as positive.

Accuracy is the most commonly used metric to evaluate classification performance (Maeda-Gutiérrez et al., 2020). Accuracy indicates the ratio of samples classified as true to all sample numbers and can be defined as shown below. Through the accuracy index, it is possible to determine the percentage of samples that are correctly classified based on Eq. (4):

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \times 100 (\%) \quad (4)$$

In this case, precision is an indicator of correct classification and can be used to evaluate the predictive power of an algorithm. Precision is a measure of how many of the values estimated to be positive by a model are actually positive and is calculated using Eq. (5).

$$\text{Precision} = \frac{TP}{TP + FP} \times 100 (\%) \quad (5)$$

Recall indicates the accuracy of positive samples and reflects how many positive classes are correctly labeled, as shown in Eq. (6).

$$\text{Recall} = \frac{TP}{TP + FN} \times 100 (\%) \quad (6)$$

The F1-score can be calculated as the harmonic mean of the precision and recall indicators (see Eq. (7)) and is used to evaluate the quality of a classification model. This indicator is focused on the positive class, and the larger the value is, the better the model performs for the positive class.

$$\text{F1-score} = \frac{\text{Recall} \times \text{Precision} \times 2}{\text{Recall} + \text{Precision}} (\%) \quad (7)$$

Finally, the area under the ROC curve (AUC) indicator was obtained through ROC graph analysis. The ROC graph was obtained by plotting all possible true positive and false positive rates to find the corresponding threshold (Wang, Gong, Li & Qiu, 2019b). In the ROC curve, the false positive rate (FPR) on the x-axis and the true positive rate (TPR) on the y-axis are calculated through Eq. (8) and Eq. (9), respectively. In previous studies, it was confirmed that the AUC indicator can be used to effectively assess classification performance, such as by indicating how well the positive and negative classes are classified (Bradley, 1997). The AUC indicator is a measure of discrimination that maximizes predictions to minimize false positives and is dependent on both true and false

positives (Ford, Dirmeyer & Benson, 2018). AUC values range from 0 to 1, and values > 0.8 reflect good model performance (Radinger et al., 2017; Wang et al., 2019b).

$$TPR = \frac{TP}{TP + FN} \times 100 (\%) \quad (8)$$

$$FPR = \frac{FP}{TN + FP} \times 100 (\%) \quad (9)$$

## 2.6. Model interpretation with the SHAP method

This study introduces the SHAP method proposed by Lundberg and Lee (2017b) for RF model interpretation. SHAP supports the interpretation of machine learning models through Shapely values. When a model is formed with an RF, each node in each DT provides a condition for splitting the dataset. There are two criteria used to measure the quality of a classification. The criteria commonly used to select the optimal condition in the classification process include the Gini index and entropy-based information gain. With these indicators, it is possible to obtain information on how each feature contributes to an average decrease in error in the model classification process. As such, the average of the data collected from all trees in the forest is a measure of feature importance. However, the feature importance ranking obtained based on the error decrease in data division is insufficient for explaining individual predictions (Meng, Yang, Qian & Zhang, 2021). For example, the contribution of each feature to the output is unknown; e.g., in the studied case, the indicators that may increase mortality (MOR) are not known.

Therefore, we introduce the SHAP concept to determine which features contribute to the occurrence of mortality and how important the contributions are. SHAP uses the concept of game theory developed by Shapley (1953) to calculate the importance of individual independent variables (Seyrfar, Ataei, Movahedi & Derrible, 2021). Ribeiro, Singh and Guestrin (2016) combined an estimation (prediction) model  $f$  with an explanation model  $g$  through local methods. Lundberg and Lee (2017b) defined an additive feature attribution method (see Eq. (10)) with a linear function explanatory model  $g$  for binary variables.

$$g(z') = \phi_0 + \sum_{i=1}^M \phi_i z'_i \quad (10)$$

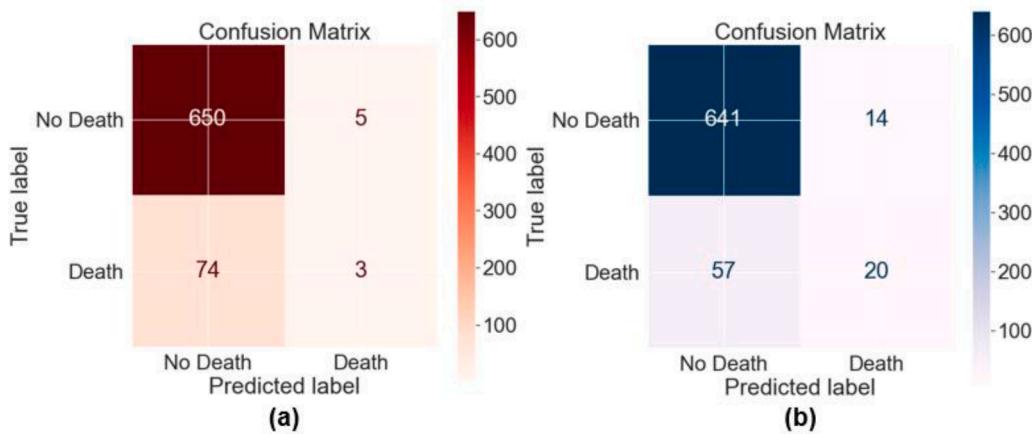
where  $g$  = a linear function of binary features;  $z'$  = observed ( $z'_i = 1$ ) or unknown ( $z'_i = 0$ ) features;  $M$  = the number of simplified input features; and  $\phi_i$  = the contribution of each feature. A system is developed to assess the contribution of each input feature to the value produced by the model. The Shapely value satisfies all three characteristics (local accuracy, missingness, and consistency) that the additive feature attribution technique should have and is the value most consistent with human intuition (Lee, Oh & Kim, 2020b). Shapely values are a measure of the contribution of each predictor (feature) in a machine learning model.

To examine the effect of a specific feature  $i$  on the model, two models are trained:  $f_{S \cup \{i\}}$  when  $i$  is included and  $f_S$  when it is not. For a specific input  $x_S$ , the difference in output derived through these two models indicates the effect of feature  $i$  on the model. Based on this theoretical concept, the Shapely value representing the contribution of each feature is finally calculated as the weighted average of all possible differences, as shown in Eq. (11) (Mokhtari et al., 2019).

$$\phi_i = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|!(|F| - |S| - 1)!}{|F|!} [f_{S \cup \{i\}}(x_{S \cup \{i\}}) - f_S(x_S)] \quad (11)$$

where  $f$  = the prediction/estimation model;  $F$  = the set of all features; and  $S$  = the set of all features excluding  $i$ .

With the contributions derived in this way as a measure, the SHAP method can explain the output of a machine learning model through global and local analyses. Global interpretability reflects whether each



**Fig. 4.** Confusion matrices of the RF models. (a) Test set evaluation for the RF initial model; (b) test set evaluation for the RF model with GS. RF, random forest; GS, grid search.

feature contributes positively or negatively to the output variable. Local analysis yields a unique SHAP value for each case or instance, indicates why a case is derived as a specific output and supports assessments of the contributions of features in each instance (Urista et al., 2020). In Python, a practical package can be used to calculate SHAP values by combining various techniques, including the LightGBM, GBoost, CatBoost, XgBoost, and scikit-learn tree models (Lundberg & Lee, 2017a). In this study, we implemented SHAP using the Python shap package for RF model interpretation.

### 3. Results

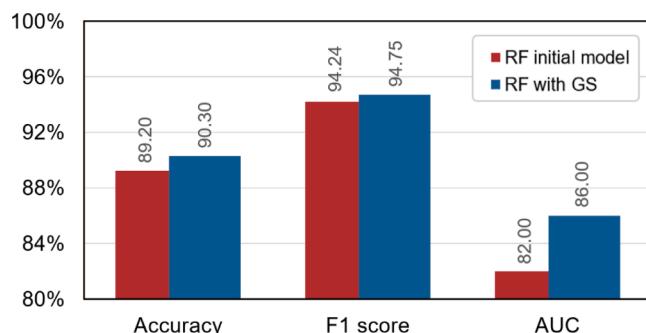
#### 3.1. Model evaluation

##### 3.1.1. Confusion matrix

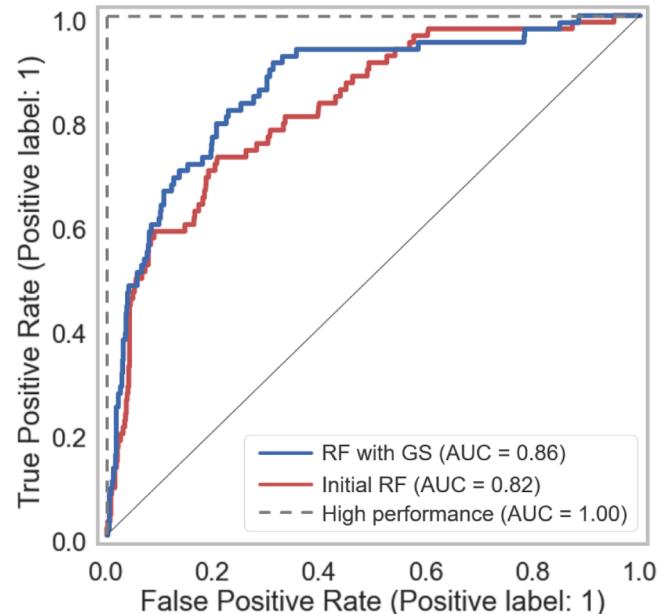
The confusion matrix between the predictions and observations is shown in Fig. 4. Based on the confusion matrix, we performed a performance comparison of the estimation results for each training and testing dataset. In the case of the initial RF model, the accuracy and F1-score were 89.2% and 94.24%, respectively, for estimations with the test dataset. In the case of the RF model with the grid search method, the accuracy and F1-score measures were 90.3% and 94.75%, respectively, for estimations with the test dataset. As a result of comparing the performance of the models based on the confusion matrix, it was confirmed that the RF model with the grid search method yielded the best estimation performance (see Fig. 5).

##### 3.1.2. ROC curve

Fig. 6 is a ROC curve comparison plot generated based on the TPRs



**Fig. 5.** Comparison of the RF model and the RF model with GS (each indicator is expressed as a unified percentage). RF, random forest; GS, grid search; AUC, area under the curve.



**Fig. 6.** Receiver operating characteristic curves illustrating the performance of the models.

RF, random forest; GS, grid search; AUC, area under the curve.

and FPRs of the initial model (red line) and RF model with GS (blue line). The AUC was 0.71 for the initial model and 0.86 for the RF model with GS. In the case of the RF model with the GS method, the AUC was consistently 0.8 or higher, and it displayed the best performance. Fig. 5 is a graph comparing the initial RF model and the RF model with GS based on the accuracy, F1-score, and AUC. For all three indicators, it was confirmed that the latter model performed better.

##### 3.1.3. Comparison of the accuracy achieved in various studies

Table 4 shows a comparison of the evaluation results of previous studies that predicted urban environment disasters using various intelligent machine learning algorithms, including RF models. In the considered research field, no studies had established detailed spatial unit classification models of heat-related mortality, which is an indicator of heat wave damage, for various areas of a city; therefore, a direct heat wave damage comparison could not be included. Similarly, it was difficult to verify many previous models used in the studied field because machine learning has rarely been applied in heat wave analysis.

**Table 4**

Comparison of the performance of various models.

No.	Research (Year)	Model target	Algorithms	Evaluation indicators		
				AUC	Accuracy (%)	F1-score
1	Apostolakis, Girtsou, Kontoes, Papoutsis and Tsoutsos (2021)	Wildfire	RF	–	–	0.93
2	Dikshit, Pradhan and Alami (2020)	Drought	RF	0.84	–	–
3	Vafakhah, Mohammad Hasani Loor, Pourghasemi and Katebikord (2020)	Flood susceptibility	FR, ANFIS, RF	0.71	–	–
4	Mohammady et al. (2019)	Land subsidence susceptibility	RF	0.77	–	–
5	Lee, Kim, Jung, Lee and Lee (2017)	Flood susceptibility	RF, BT	–	79.18	–
6	Proposed (2021)	Heat mortality	RF-GS	0.86	90.30	0.95

Note: RF refers to random forest; FR refers to frequency ratio; ANFIS refers to adaptive neurofuzzy inference system; LR refers to logistic regression; RF-GS refers to random forest with grid search; BT refers to boosted tree.

Unfortunately, we could not conduct the verification of direct estimation performance for heat and mortality with RF-based algorithms because no previous study estimated human casualties in a detailed local unit using the RF classification model with similar indicators used in this study. However, an indirect comparison with a study that established a disaster-related prediction/estimation model with an RF algorithm was considered, despite the difference in field. More generalized verification may be possible if various predictive models in the heat wave field are developed under more diverse conditions in the future and comparisons based on them are performed.

In comparisons of two or more algorithms other than RF models in Table 4 (e.g., no. 3 and no. 5), the performance of the best derived model was considered. Although the performance evaluation indexes used in each comparative study were different, in this study (Table 4; no. 6), three measures were used for evaluation; therefore, the comparison was performed based on commonly used measures. The AUC of 0.86, accuracy of 90.3, and F1-score of 0.95 for the model proposed in this study were higher than those of models introduced in other studies.

### 3.2. Model interpretation

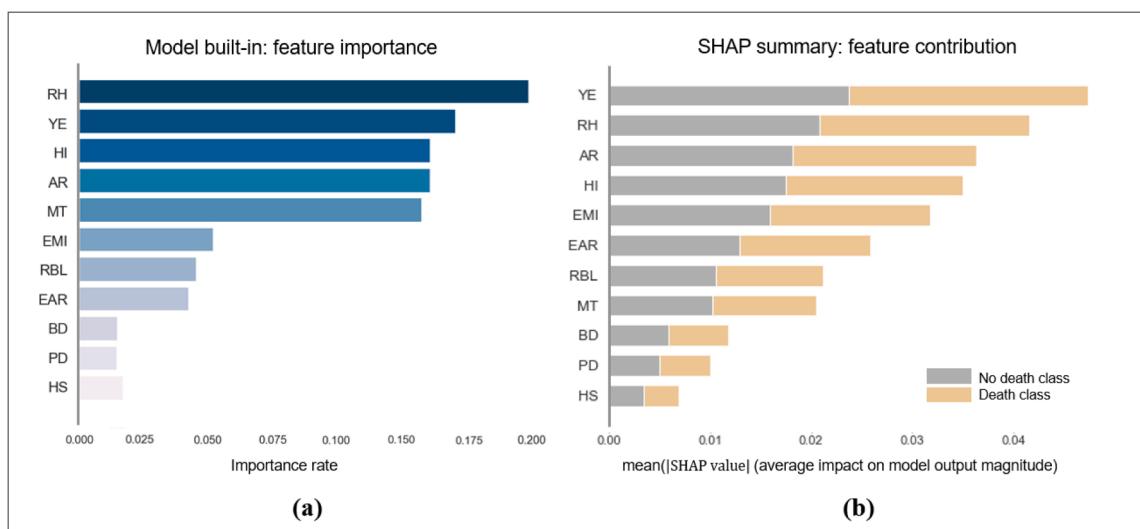
#### 3.2.1. Global feature interpretation

**3.2.1.1. Analysis of the importance and contributions of model features.** Feature importance can be assessed based on various methods (Lai, Cai & Tan, 2019). Some tree-based models, such as RFs, automatically assign feature importance, and the assignment scheme influences the results. However, feature importance is different from feature contributions. Feature importance reflects which features have the greatest

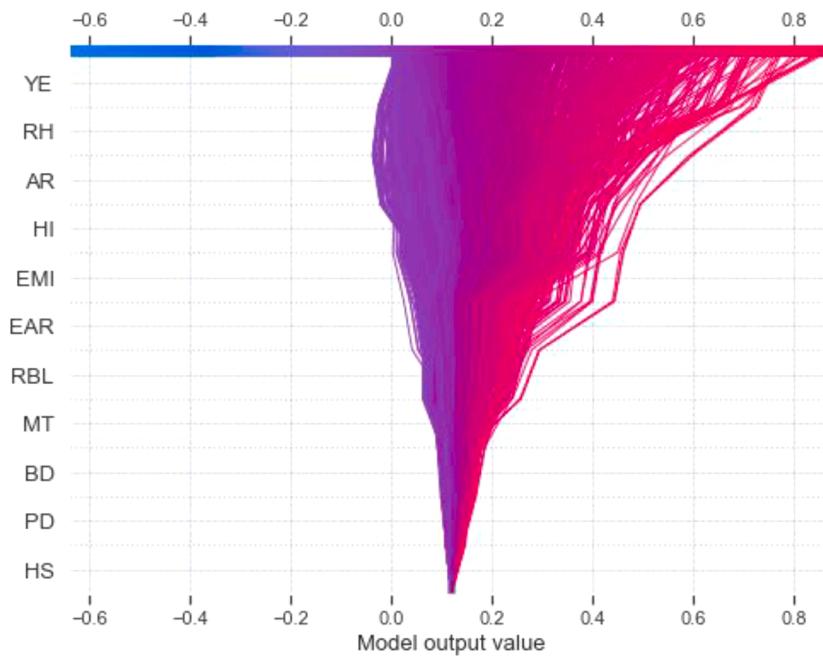
influence on the performance of a model. According to Meng et al.'s study (2021), feature contributions provide an intuitive explanation for the considered output (death or no death) beyond the identification of influential factors. In this study, two analyses were introduced to determine how important each feature is and the corresponding contribution to the model estimation: the model-based feature importance approach and the feature importance based on the SHAP summary plot approach.

First, the importance of each feature in the model was analyzed. The importance plot shown in Fig. 7(a) focuses on how well the model classifies samples based on the Gini index and illustrates the importance of reducing Gini indexes in the classification process (Gómez-Ramírez et al., 2020). The sum of the Gini index for all features in the model equals 1. The four variables with the highest importance scores are RH, YE, HI, and AR, which are all meteorological sector and socioeconomic sector variables. However, the Gini index-based result does not indicate how much each feature contributes to a specific model output.

Therefore, SHAP summary analysis was additionally introduced for in-depth model analysis. Through the SHAP summary plot, we derived a quantitative value that aggregated Shapely values, reflecting the contributions of variables to the model (see Fig. 7(b)). The x-axis of the graph represents the contribution of each feature to the prediction result of classification through the average of the sum of the absolute values of SHAP values. As a result of analyzing the contribution of each input feature to the model, the four variables with the highest contributions were derived in the following order: YE, RH, AR, and HI. These variables from the meteorological sector and the socioeconomic sector are the same indicators identified in the analysis of model importance, as shown in Fig. 7(a). However, in assessing the model contributions, it was found



**Fig. 7.** Analysis of feature importance and contributions to the RF model. (a) Gini index-based importance plot; (b) SHAP summary plots for mortality (MOR) predictors. RH, relative humidity; YE, young and elderly; HI, heat index; AR, aging ratio; EMI, emergency medical institute; EAR, elderly alone rate; RBL, recipients of a basic livelihood; MT, max. temp.; BD, building density; PD, population density; HS, heat shelter.



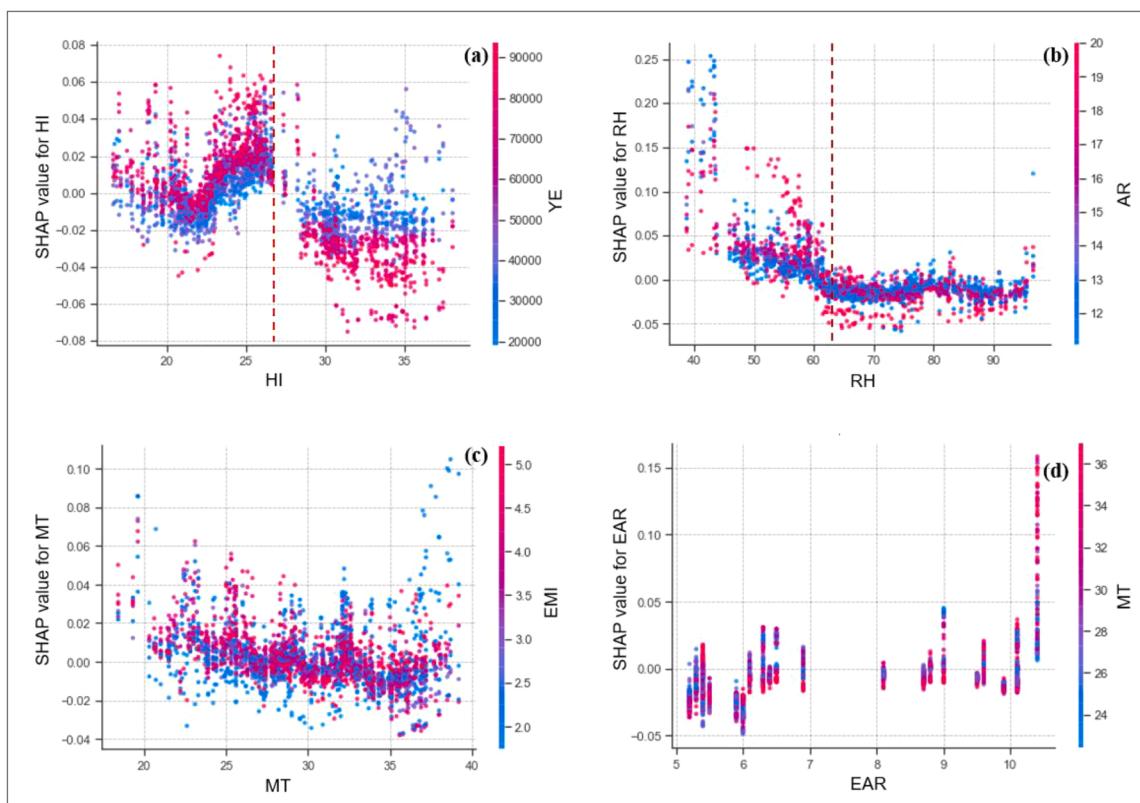
**Fig. 8.** SHAP decision plot for the RF-GS model.

RH, relative humidity; YE, young and elderly; HI, heat index; AR, aging ratio; MT, max. temp.; EMI, emergency medical institute; EAR, elderly alone rate; RBL, recipients of a basic livelihood; BD, building density; PD, population density; HS, heat shelter.

that the rankings of the socioeconomic sector indicators YE and AR increased. Thus, the socioeconomic sector makes an important contribution to the model output, that is, the mortality (MOR) estimation.

#### 3.2.1.2. Analysis of the global decision-making process. The decision plots

of the SHAP method show how total estimates change during the decision-making process (see Fig. 8). On the Y-axis, features sorted according to their contributions are listed. The X-axis represents the output of the model. As in the SHAP summary plot, the SHAP value of each feature is summed for the base value of the model while moving from the



**Fig. 9.** SHAP dependence plot for the RF-GS model. HI, heat index; YE, young and elderly; RH, relative humidity; AR, aging rate; MT, max. temp.; EAR, elderly alone rate.

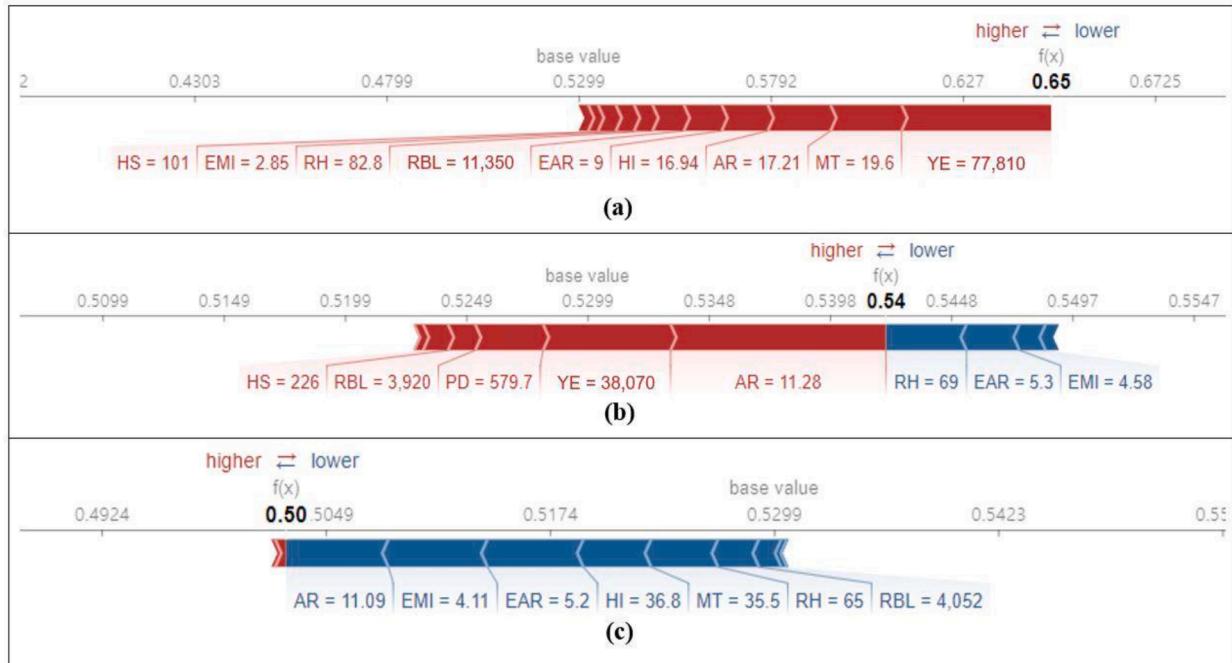
bottom to the top of the plot to derive the final output value. Therefore, it is possible to determine how much each feature contributes to the output during the overall estimation process. Based on the results of this analysis, as observed for the SHAP importance plot, the ranking of the indicators with the greatest contributions to the overall estimation is as follows: YE, RH, AR, and HI.

**3.2.1.3. Feature dependence analysis.** The SHAP dependency plot was analyzed to observe the interactions among input variables used in RF model estimation in terms of their contributions (see Fig. 9). As shown in Fig. 9, the most prominent trend among the SHAP dependence analysis results was associated with the graphs for HI, YE, RH, and AR, which were ranked high in the contribution rankings based on the SHAP summary plot (see Fig. 9(a) and (b)). The scatter plots of the red and blue points show the variations in the HI and the SHAP values of HI. In the graph, the HI and HI SHAP value trends significantly change in the vicinity of 27 °C (vertical red line). At this point, it can be confirmed that the SHAP value is relatively high at 27 °C and decreases at higher temperatures, and it rises remarkably from approximately 20–27 °C. Accordingly, the influence of the corresponding index significantly contributes to the MOR estimation at 27 °C and lower; conversely, at 27 °C and higher, the result is evenly distributed in a relatively wide range. Thus, it can be confirmed that the effects of other variables on the MOR increase at an HI of 27 °C, and the contribution of HI may be affected. In particular, for YE, as expressed by the blue and red dots in the section above 27 °C, the SHAP value of HI decreases as the dot color approaches red; additionally, the SHAP value of HI is highly distributed for YE dots with colors closer to blue. Hence, as YE increases, the contribution of HI is reduced due to the influence of the YE variable on the model. Without considering the differences between the red and blue colors of points in the graph in Fig. 9(b), the scatter plots of red and blue points show the variations in the RH and the SHAP values of RH. In the graph, the trend significantly changes in the vicinity of an RH of 65% (red line), and the SHAP value is remarkably reduced in the right-downward direction at 65% or less compared to when the RH is higher than 65%. Accordingly, it can be confirmed that in the case of an RH of 65% or less, the smaller the RH value is, the greater the contribution to the MOR estimation.

Conversely, above 65%, relatively uniformly low SHAP values are distributed. It can be inferred that the influence of other input variables on the model increases when the RH contribution increases at 65% or higher. Additionally, for AR, as reflected by the size of the dots in the plot, the SHAP value of RH increases as the dots become closer to blue in the vicinity of an RH of 40%, where the RH SHAP value is the largest. Moreover, the closer the dots are to red, the lower the SHAP value of RH. This result confirms that the contribution of the RH to the proposed estimation model is maximized when AR values are low. In Fig. 9(c), it is confirmed that the contribution slightly decreases as the MT increases, but the trend does not exhibit a clear pattern compared to those in Fig. 9(a) and (b). In Fig. 9(d), EAR does not significantly affect the model; however, a sharp upward trend and the highest contribution can be observed at an EAR of approximately 10.5%. In addition, the EAR SHAP value also increases as the points become closer to red than blue in color. Overall, it is confirmed that the contribution of EAR to the model increases when MT is high.

### 3.2.2. Local feature interpretation

**3.2.2.1. Individual sample analysis through a SHAP force plot.** Fig. 10 shows the SHAP explanatory force plot for three instances randomly selected from the actual estimation results. In the graph, the base value represents the average of the model estimation results for the training set. If the output of the model is to the right of the base value (i.e., higher than the base value), MOR occurs (death). Conversely, if the output of the model is to the left of the base value (i.e., lower than the base value), MOR does not occur (no death). Red arrows indicate the degree of influence of input variables that increase the occurrence of MOR (increased possibility of death). Blue arrows indicate the influence of input variables that suppress the occurrence of MOR (increased possibility of no death). The size (amount) of the area occupied by the variables in each arrow expresses the degree of influence of the variable. Fig. 10(a) shows the case with the highest mortality risk prediction among the three cases. At this time, the three indicators that could potentially contribute to the largest increases are YE, MT, and AR. In particular, the significantly higher YE index compared to other indexes



**Fig. 10.** Explanation of evaluation results based on the SHAP tree explainer. RH, relative humidity; YE, young and elderly; HI, heat index; AR, aging ratio; MT, max. temp.; EMI, emergency medical institute; RBL, recipients of a basic livelihood; BD, building density; PD, population density; HS, heat shelter.

reflects the higher expected occurrence of mortality. Although AR and YE rank high in the order of influence in both cases in Fig. 10(a) and (b), the estimation of MOR occurrence is lower in Fig. 10(b) than that in Fig. 10(a). The AR and YE values in Fig. 10(b) are lower than the corresponding values in Fig. 10(a). Fig. 10(c) illustrates the lowest risk of mortality among the three instances in Fig. 10(a)-(c), and this reduced AR effect is considered the most notable. Another implication that can be obtained from the comparative analysis of the graphs in Fig. 10(a) and (c) is that the increase in EMI, which is related to risk reduction, suppresses the occurrence of MOR (increased possibility of no death). According to Fig. 10(c), EMI ranks as the second largest influencing factor that decreases the MOR estimate, and AR ranks first. Thus, EMI also partially affects MOR.

#### 4 Discussion

In this study, a spatial unit RF-based classification model was developed to estimate the occurrence of heat-related mortality using various urban indicators. Although a time series prediction per city unit was performed based on an RF method to investigate damage caused by heat waves in Korea (Park et al., 2020), this study is the first RF-based heat-related mortality estimation model for multiple spatial units within a city. Additionally, it is important to build a model with the optimal combination of parameters, as the prediction of heat waves can introduce parametric uncertainty (Samaniego et al., 2017). To derive hyperparameters that optimized model performance, we used the GridSearchCV method. According to the result of comparing the RF model before and after using the GridSearchCV method, it was confirmed that the accuracy, F1 score and AUC increased by 1.1%, 0.51%, and 4%, respectively, compared to those without the GridSearchCV method. It was an important task to develop a model with the best accuracy when the estimation model output was an index indicating human health damage (Iwendi et al., 2020). Similarly, the process of finding the hyperparameter combination of the model with the best performance is essential in this study according to the results of estimating the MOR with improved performance by incorporating the GridSearchCV. We compared the performance of the proposed final model with the RF-based models developed in several previous studies in the urban disaster field, and verified the model by confirming that our model has higher performance than others (see Table 4).

In some previous studies, RF models were developed to predict the health damage caused by heat waves (Park et al., 2020; Wang et al., 2019c). However, few studies have gone beyond evaluating the performance of prediction or estimation models in the heat wave field and have performed in-depth model analysis. Understanding the predictions of machine learning models is important for evaluating the accuracy of models in many applications; however, the nature of black box models makes it difficult to interpret the model accurately in most cases (Alaa & van der Schaar, 2019). Among the various analysis techniques that can be applied to interpret a model with black box characteristics, the SHAP and LIME methods can be used (Ahmed et al., 2020). The SHAP technique can explain the output of a developed machine learning model and provides theoretical guarantees of consistency and local accuracy based on game theory (Rathi, 2019). Therefore, in this study, we adopted the SHAP value to provide an in-depth explanation of the individual MOR predictions for the RF model. RF model interpretation in this study was largely divided into global and local analyses. First, in the global analysis, the contributions of model features were analyzed through SHAP summary plots, decision plots, and dependence plots (see Figs. 7(b), 8 and 9). Based on the analyses of the SHAP summary plot and decision plot, the feature sectors with the largest contributions to the results were the meteorological, demographic and socioeconomic sectors. Among them, the YE, RH, AR, and HI indicators displayed the highest contributions in the estimation process of the model for MOR estimation. According to the common results derived from the two analyses, YE was the most significant indicator in the damage estimation model of heat

waves more than the other indicators. Additionally, considering that AR was the third-most-important contributing factor in the two analyses, it was confirmed that indicators related to the older population are significantly correlated with MOR estimates. In fact, studies have reported that extremely high temperatures increase the rate of cardiovascular disease-based mortality (Yin & Wang, 2018). It can be inferred that considering the number of elderly individuals and the distribution of elderly individuals in various spatial areas in a city is important when developing strategies to reduce the effects of heat waves. SHAP dependence plots were also used in the global analysis, and it was confirmed that HI, YE, RH, AR, MT, and EAR displayed a co-influence in the model estimation process (see Fig. 9). It was confirmed that the contribution of HI decreases as the value of YE increases when the HI is 27 °C or higher (see Fig. 9(a)). These results confirmed that YE contributes significantly more to estimates of heat-related MOR than does the HI value in a certain area, assuming that HI is continuously increasing. Finally, we introduced the SHAP force plot for local analysis (see Fig. 10). High YE and AR index values were the cause of the increase in MOR occurrence (death) (see Fig. 10(a)-(b)). Additionally, in Fig. 10(c), the lower YE and higher EMI values than in the other cases contributed significantly to the low MOR (no death). To establish a plan to reduce human casualties due to heat environment, it is necessary to interpret the results of model by identifying demographic and socioeconomic characteristics of each spatial unit. As vulnerable groups, such as the elderly and socially isolated people, have been reported to have an increased mortality rate during periods of extreme heat (Kravchenko, Abernethy, Fawzy & Lyerly, 2013), the YE and AR indicators rank at the highest contribution in the estimation model of this study. Thus, to reduce such heat-related damage, a strategy to appropriately plan and establish adaptation facilities such as EMI at a local level is necessary according to the indicator-based spatial characteristics of various sectors including demographic and socioeconomic sectors.

This study provides implications for heat-related damage in a city. First, this method can be used to estimate future trends in heat-related damage according to various changing scenarios in the future (e.g., climate change, population change, infrastructure change, etc.). Second, the proposed model is helpful to develop heat wave adaptation plans because the distribution of input characteristics demonstrate spatial characteristics within a city. Because the indicators in the proposed model encompass urban climate environmental and spatial characteristics, it is possible to identify specific environment and spatial characteristics for making a sustainable urban heat environment. Finally, the urban variables can be determined which are significant to assess human health damage related to urban heat. By using RF model results and detailed SHAP interpretation technique, it is possible to identify appropriate variables that affect local health damage due to heat waves.

#### 5. Conclusion

To the best of our knowledge, this is the first study to develop and interpret a detailed spatial unit estimation model of heat-related mortality within a city based on an RF classifier. Unlike previous studies that mainly focused on explaining the importance of features, this study performed an in-depth interpretation of the estimation model based on comprehensive and detailed feature contributions.

According to the findings from RF model estimation and SHAP analysis, climate, demographic and socioeconomic sectors contributed the most to MOR estimation. Among the various features of the demographic and socioeconomic sectors, the YE and AR indicators displayed the highest contributions; the factors reflect the status of the socially vulnerable class, the distribution of which varies within cities. In particular, because the YE displayed higher contribution than any others including the climate sector, demographic and socioeconomic sectors should be considered to establish an adaption strategy by heat-wave in a city.

Limited by the experimental conditions, this study only included

three years of data. Although many open-source datasets were explored for database construction, there were material and labor limitations that prohibited some long-term data from being used. Nevertheless, from a long-term perspective, the estimation model and corresponding methodologies proposed in this study can be actively used as a basis for developing a control or mitigation system to reduce the effects of climate change and heat waves in the future. Simultaneously, due to time and labor limitations associated with data collection, we included only four indicators in the scope of this study for the demographic and socioeconomic sectors; these factors were related to the socially disadvantaged and displayed the highest contributions to the model estimates. Therefore, future research should focus on collecting more diverse meteorological, demographic and socioeconomic sector features over a long period of time and rerunning the proposed model.

## Role of the funding source

The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Author contributions

**Yesuel Kim:** Conceptualization, Data Analysis, Methodology, Technique application, Resources, Writing - Original Draft; **Youngchul Kim:** Conceptualization, Formal analysis, Methodology, Resources, Funding acquisition, Methodology, Project administration, Supervision, Validation, Writing - Review & Editing

## Declaration of Competing Interest

No potential conflict of interest was reported by the authors.

## Acknowledgements

### Funding

This work is supported by the Korea Agency for Infrastructure Technology Advancement (KAIA) grant funded by the Ministry of Land, Infrastructure and Transport of Korea (22NSPS-B149840-05, 22UMRG-C158194-03 and 'Innovative Talent Education Program for Smart City').

## References

- Ahmed, J., do Nascimento, A., Kilinc, C., Pan, D., i Riu, J. R., & Gustafsson, J. (2020). Using blackbox ML techniques to diagnose QoI problems for an IPTV service. In *NOMS 2020 - 2020 IEEE/IFIP Network Operations and Management Symposium* (pp. 1–8). IEEE. <https://doi.org/10.1109/NOMS47738.2020.9110375>.
- Alaa, A. M., & van der Schaar, M. (2019). Demystifying black-box models with symbolic metamodels. *Adv. Neural Inf. Process. Syst.*, *32*, 11304–11314.
- Ali, R., Bakhsh, K., & Yasin, M. A. (2019). Impact of urbanization on CO<sub>2</sub> emissions in emerging economy: Evidence from Pakistan. *Sustainability Cities Society*, *48*, Article 101553. <https://doi.org/10.1016/j.scs.2019.101553>
- Alshraideh, H., Castillo, E. D., & Gil Del Val, A. (2020). Process control via random forest classification of profile signals: An application to a tapping process. *J. Manufacturing Processing*, *58*, 736–748. <https://doi.org/10.1016/j.jmapro.2020.08.043>
- Anderson, G. B., & Bell, M. L. (2011). Heat waves in the United States: Mortality risk during heat waves and effect modification by heat wave characteristics in 43 U.S. communities. *Environmental Health Perspectives*, *119*, 210–218. <https://doi.org/10.1289/ehp.1002313>
- Apostolakis, A., Girtsoi, S., Kontoes, C., Papoutsis, I., & Tsoutsos, M. (2021). Implementation of a random forest classifier to examine wildfire predictive modelling in greece using diachronically collected fire occurrence and fire mapping data. *MultiMedia Modeling*, 318–329. [https://doi.org/10.1007/978-3-030-67835-7\\_27](https://doi.org/10.1007/978-3-030-67835-7_27)
- Applegate, W. B., Runyan, J. W., Brasfield, L., Williams, M. L., Konigsberg, C., & Fouche, C. (1981). Analysis of the 1980 heat wave in memphis. *Journal of the American Geriatrics Society*, *29*, 337–342. <https://doi.org/10.1111/j.1532-5415.1981.tb01238.x>
- Bakhsh, K., Rauf, S., & Zulfiqar, F. (2018). Adaptation strategies for minimizing heat wave induced morbidity and its determinants. *Sustain. Cities Soc.*, *41*, 95–103. <https://doi.org/10.1016/j.scs.2018.05.021>
- Bradley, A. P. (1997). The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition*, *30*, 1145–1159. [https://doi.org/10.1016/s0031-3203\(96\)00142-2](https://doi.org/10.1016/s0031-3203(96)00142-2)
- Breiman, L. (2001). Random forests. *Machine Learning*, *45*, 5–32. <https://doi.org/10.1023/a:1010933404324>
- Busby, J. W., Cook, K. H., Vizy, E. K., Smith, T. G., & Bekalo, M. (2014). Identifying hot spots of security vulnerability associated with climate change in Africa. *Climatic Change*, *124*, 717–731. <https://doi.org/10.1007/s10584-014-1142-z>
- Christo, V. R. E., Nehemiah, H. K., Brighty, J., & Kannan, A. (2020). Feature selection and instance selection from clinical datasets using co-operative co-evolution and classification using random forest. *IETE Journal of Research*, *1*–14. <https://doi.org/10.1080/03772063.2020.1713917>
- Dikshit, A., Pradhan, B., & Alamri, A. M. (2020). Short-term spatio-temporal drought forecasting using random forest model at New South Wales. *Australia. Applied Sciences*, *10*, 4254. <https://doi.org/10.3390/app10124254>
- Dong, L., Li, X., & Xie, G. (2014). Nonlinear methodologies for identifying seismic event and nuclear explosion using random forest, support vector machine, and naive Bayes classification. *Abstr. Applied Analysis*, *2014*, Article 459137. <https://doi.org/10.1155/2014/459137>
- Dong, W., Zeng, Q., Ma, Y., Li, G., & Pan, X. (2016). Impact of heat wave definitions on the added effect of heat waves on cardiovascular mortality in Beijing, China. *International Journal of Environmental Research on Public Health*, *13*, 933. <https://doi.org/10.3390/ijerph13090933>
- Ebi, K. L., & Schmierer, J. K. (2005). A stitch in time: Improving public health early warning systems for extreme weather events. *Epidemiologic Reviews*, *27*, 115–121. <https://doi.org/10.1093/epirev/mxi006>
- Faye, M., Dème, A., Diougue, A. K., & Diouf, I. (2021). Impact of different heat wave definitions on daily mortality in Bandaressi, Senegal. *PloS one*, *16*, Article e0249199. <https://doi.org/10.1371/journal.pone.0249199>
- Ford, T. W., Dirmeyer, P. A., & Benson, D. O. (2018). Evaluation of heat wave forecasts seamlessly across subseasonal timescales. *NPJ Climate and Atmosphere Science*, *1*, 1–9. <https://doi.org/10.1038/s41612-018-0027-7>
- Fouillet, A., Rey, G., Wagner, V., Laaidi, K., Empereur-Bissonnet, P., & Le Tertre, A. (2008). Has the impact of heat waves on mortality changed in France since the European heat wave of summer 2003? A study of the 2006 heat wave. *International Journal of Epidemiology*, *37*, 309–317. <https://doi.org/10.1093/ije/dym253>
- Friedman, J., Hastie, T., & Tibshirani, R. (2001). *The elements of statistical learning*. Berlin, Germany: Springer.
- Putagami, K., Fukazawa, Y., Kapoor, N., & Kito, T. (2021). Pairwise acquisition prediction with SHAP value interpretation. *Journal Finance Data Science*, *7*, 22–44. <https://doi.org/10.1016/j.jfds.2021.02.001>
- Gómez-Ramírez, J., Ávila-Villanueva, M., & Fernández-Blázquez, M.Á. (2020). Selecting the most important self-assessed features for predicting conversion to mild cognitive impairment with random forest and permutation-based methods. *Science Reports*, *10*, 1–15. <https://doi.org/10.1038/s41598-020-77296-4>
- Green, H. K., Andrews, N., Armstrong, B., Bickler, G., & Pebody, R. (2016). Mortality during the 2013 heatwave in England – How did it compare to previous heatwaves? A retrospective observational study. *Environmental Research*, *147*, 343–349. <https://doi.org/10.1016/j.envres.2016.02.028>
- Guo, X., & Hendel, M. (2018). Urban water networks as an alternative source for district heating and emergency heat-wave cooling. *Energy*, *145*, 79–87. <https://doi.org/10.1016/j.energy.2017.12.108>
- Hatvani-Kovacs, G., Belusko, M., Skinner, N., Pockett, J., & Boland, J. (2016). Heat stress risk and resilience in the urban environment. *Sustainability Cities Society*, *26*, 278–288. <https://doi.org/10.1016/j.scs.2016.06.019>
- Hirano, Y., Kondo, Y., Ifumi, T., Yokobori, S., Kanda, J., & Shimazaki, J. (2021). Machine learning-based mortality prediction model for heat-related illness. *Science Reports*, *11*, 1–8. <https://doi.org/10.1038/s41598-021-88581-1>
- Hu, X., Belle, J. H., Meng, X., Wildani, A., Waller, L. A., & Strickland, M. J. (2017). Estimating PM2.5 concentrations in the conterminous United States using the random forest approach. *Environmental Science & Technology*, *51*, 6936–6944. <https://doi.org/10.1021/acs.est.7b01210>
- Ikeda, T., & Kusaka, H. (2021). Development of Models for Predicting the Number of Patients with Heatstroke on the Next Day Considering Heat Acclimatization. *Journal Meteorology Society Japan*. <https://doi.org/10.2151/jmsj.2021-067>
- Iwendi, C., Bashir, A. K., Peshkar, A., Sujatha, R., Chatterjee, J. M., & Pasupuleti, S. (2020). COVID-19 patient health prediction using boosted random forest algorithm. *Front. Public Health*, *8*, 357. <https://doi.org/10.3389/fpubh.2020.00357>
- Jung, E. H., Kim, C. W., & Park, J. H. (2018). Design of climate change vulnerability assessment database system for heat wave and drought. *Journal of Korea Institute Electronic Communication Sciences*, *13*, 813–818.
- Keramitsoglou, I., Kiranoudis, C. T., Maiheu, B., De Ridder, K., Daglis, I. A., Manunta, P., et al. (2013). Heat wave hazard classification and risk assessment using artificial intelligence fuzzy logic. *Environmental Monitoring and Assessment*, *185*, 8239–8258. <https://doi.org/10.1007/s10661-013-3170-y>
- Kim, D. W., Deo, R. C., Park, S. J., Lee, J. S., & Lee, W. S. (2019a). Weekly heat wave death prediction model using zero-inflated regression approach. *Theoretical and Applied Climatology*, *137*, 823–838. <https://doi.org/10.1007/s00704-018-2636-9>
- Kim, G., & Jung, H. (2020). Climate justice assessment considering regional inequality: A focus on heatwave. *Journal of Climate Change Research*, *11*, 621–628. <https://doi.org/10.15531/kscr.2020.11.6.621>
- Kim, J., Kwon, T. H., Lee, J. M., & Kim, Y. (2019b). Identifying major components of extreme heatwave risk assessment indexes in urban areas. *KIEAE Journal*, *19*, 5–10. <https://doi.org/10.12813/kiae.2019.19.5.005>
- Kim, Y. O., Lee, W., Kim, H., & Cho, Y. (2020). Social isolation and vulnerability to heatwave-related mortality in the urban elderly population: A time-series multi-community study in Korea. *Environment International*, *142*, Article 105868. <https://doi.org/10.1016/j.envint.2020.105868>

- Kodera, S., Nishimura, T., Rashed, E. A., Hasegawa, K., Takeuchi, I., Egawa, R., et al. (2019). Estimation of heat-related morbidity from weather data: A computational study in three prefectures of Japan over 2013–2018. *Environment International*, 130, Article 104907. <https://doi.org/10.1016/j.envint.2019.104907>
- Kravchenko, J., Abernethy, A. P., Fawzy, M., & Lyerly, H. K. (2013). Minimization of heatwave morbidity and mortality. *American Journal of Preventive Medicine*, 44(3), 274–282. <https://doi.org/10.1016/j.amepre.2012.11.015>
- Lai, V., Cai, Z., & Tan, C. (2019). Many faces of feature importance: Comparing built-in and post-hoc feature importance in text classification. arXiv preprint arXiv: 1910.08534.
- Lee, D. G., Kim, K. R., Kim, J., Kim, B. J., Cho, C. H., Sheridan, S. C., et al. (2018). Effects of heat waves on daily excess mortality in 14 Korean cities during the past 20 years (1991–2010): An application of the spatial synoptic classification approach. *International Journal of Biometeorology*, 62(4), 575–583. <https://doi.org/10.1007/s00484-017-1466-2>
- Lee, H. D., Min, K. H., Bae, J. H., & Cha, D. H. (2020a). Characteristics and comparison of 2016 and 2018 heat wave in Korea. *Atmosphere*, 30, 1–15.
- Lee, K., & Hong, W. H. (2008). A study on the urban heat environment pattern analysis and alleviation plan. *Journal of Architecture Institute Korea*, 24, 253–260.
- Lee, S., Kim, J. C., Jung, H. S., Lee, M. J., & Lee, S. (2017). Spatial prediction of flood susceptibility using random-forest and boosted-tree models in Seoul metropolitan city. *Korea Geomaterial Natural Hazards Risk*, 8, 1185–1203. <https://doi.org/10.1080/19475705.2017.1308971>
- Lee, Y. G., Oh, J. Y., & Kim, G. (2020b). Interpretation of load forecasting using explainable artificial intelligence techniques. *Transactions of the Korean Institute of Electrical Engineers*, 69, 480–485. <https://doi.org/10.5370/kiee.2020.69.3.480>
- Li, T., Ding, F., Sun, Q., Zhang, Y., & Kinney, P. L. (2016). Heat stroke internet searches can be a new heatwave health warning surveillance indicator. *Science Reports*, 6, 1–6. <https://doi.org/10.1038/srep37294>
- Li, M., Wang, M., Wang, J., & Li, D. (2013). Comparison of random forest, support vector machine and back propagation neural network for electronic tongue data classification: Application to the recognition of orange beverage and Chinese vinegar. *Sensor. Ctuat. B-Chemistry*, 177, 970–980. <https://doi.org/10.1016/j.snb.2012.11.071>
- Lu, J., Zhang, Y., Chen, M., Wang, L., Zhao, S., Pu, X., et al. (2021). Estimation of monthly 1km resolution PM2.5 concentrations using a random forest model over “2 + 26” cities, China. *Urban Climate*, 35, Article 100734. <https://doi.org/10.1016/j.ulclim.2020.100734>
- Lundberg, S. M., & Lee, S. I. (2017a). Consistent feature attribution for tree ensembles. arXiv preprint arXiv:1706.06060.
- Lundberg, S. M., & Lee, S. I. (2017b). A unified approach to interpreting model predictions. *Advances in neural information processing systems* (pp. 4765–4774). New York, NY: Association for Computing Machinery.
- Maeda-Gutiérrez, V., Galván-Tejada, C. E., Zanella-Calzada, L. A., Celaya-Padilla, J. M., Galván-Tejada, J. I., Gamboa-Rosales, H., et al. (2020). Comparison of convolutional neural network architectures for classification of tomato plant diseases. *Applied Sciences*, 10, 1245. <https://doi.org/10.3390/app10041245>
- Mangalathu, S., Hwang, S. H., & Jeon, J. S. (2020). Failure mode and effects analysis of RC members based on machine-learning-based SHapley Additive exPlanations (SHAP) approach. *Engineering Structures*, 219, Article 110927. <https://doi.org/10.1016/j.engstruct.2020.110927>
- Meng, Y., Yang, N., Qian, Z., & Zhang, G. (2021). What makes an online review more helpful: An interpretation framework using XGBoost and SHAP values. *Journal of Theoretical Application of Electronics Commercial Research*, 16, 466–490. <https://doi.org/10.3390/jtaer16030029>
- Mohammady, M., Pourghasemi, H. R., & Amir, M. (2019). Land subsidence susceptibility assessment using random forest machine learning algorithm. *Environmental Earth Sciences*, 78, 503. <https://doi.org/10.1007/s12665-019-8518-3>
- Mokhtari, K. E., Higdon, B. P., & Başar, A. (2019). Interpreting financial time series with SHAP values. In *Proceedings of the 29th Annual International Conference on Computer Science and Software Engineering IBM Corp* (pp. 166–172).
- Montero, J. C., Miron, I. J., Criado, J. J., Linares, C., & Diaz, J. (2013). Difficulties of defining the term, “heat wave”, in public health. *International Journal of Environmental Health Research*, 23, 377–379. <https://doi.org/10.1080/09603123.2012.733941>
- Nishimura, T., Rashed, E. A., Kodera, S., Shirakami, H., Kawaguchi, R., Watanabe, K., et al. (2021). Social implementation and intervention with estimated morbidity of heat-related illnesses from weather data: A case study from Nagoya City. *Japan Sustainability Cities Society*, 74, Article 103203. <https://doi.org/10.1016/j.scs.2021.103203>
- Oechsli, F. W., & Buechley, R. W. (1970). Excess mortality associated with three Los Angeles September hot spells. *Environmental Research*, 3, 277–284. [https://doi.org/10.1016/0013-9351\(70\)90021-6](https://doi.org/10.1016/0013-9351(70)90021-6)
- Park, M., Jung, D., Lee, S., & Park, S. (2020). Heatwave damage prediction using random forest model in Korea. *Applied Sciences*, 10, 8237. <https://doi.org/10.3390/app10228237>
- Parsa, A. B., Movahedi, A., Taghipour, H., Derrible, S., & Mohammadian, A. (2020). Toward safer highways, application of XGBoost and SHAP for real-time accident detection and feature analysis. *Accident Analysis and Prevention*, 136, Article 105405. <https://doi.org/10.1016/j.aap.2019.105405>
- Pyrgou, A., Castaldo, V. L., Pisello, A. L., Cotana, F., & Santamouris, M. (2017). On the effect of summer heatwaves and urban overheating on building thermal-energy performance in central Italy. *Sustainability Cities Society*, 28, 187–200. <https://doi.org/10.1016/j.scs.2016.09.012>
- Radinger, J., Essl, F., Höller, F., Horký, P., Slavík, O., & Wolter, C. (2017). The future distribution of river fish: The complex interplay of climate and land use changes, species dispersal and movement barriers. *Global Change Biology*, 23, 4970–4986. <https://doi.org/10.1111/gcb.13760>
- Rathi, S. (2019). Generating counterfactual and contrastive explanations using SHAP. arXiv preprint arXiv:1906.09293.
- Rauf, S., Bakhsh, K., Abbas, A., Hassan, S., Ali, A., & Kächele, H. (2017). How hard they hit? Perception, adaptation and public health implications of heat waves in urban and peri-urban Pakistan. *Environmental Sciences Pollution R*, 24, 10630–10639. <https://doi.org/10.1007/s11356-017-8756-4>
- Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). Why should i trust you?” Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1135–1144). New York, NY: ACM.
- Rothfusz, L. P., & Headquarters NWS Southern Region. (1990). *The heat index equation (or, more than you ever wanted to know about heat index)*. Fort Worth, Texas: National Oceanic and Atmospheric Administration, National Weather Service, Office of Meteorology.
- Royé, D., Codesido, R., Tobías, A., & Taracido, M. (2020). Heat wave intensity and daily mortality in four of the largest cities of Spain. *Environmental Research*, 182, Article 109027. <https://doi.org/10.1016/j.enrenv.2019.109027>
- Ruttan, T., Stoltz, U., Jackson-Vance, S., Parks, B., & Keim, S. M. (2013). Validation of a temperature prediction model for heat deaths in undocumented border crossers. *Journal Immigration Minor Health*, 15, 407–414. <https://doi.org/10.1007/s10903-012-9619-1>
- Samaniego, L., Thober, S., Kumar, R., Rakovec, O., Wood, E., Sheffield, J., et al. (2017). Drought and heatwaves in Europe: Historical reconstruction and future projections. *19th EGU general assembly, EGU2017* (p. 5659). Austria: Vienna.
- Seyrfar, A., Ataei, H., Movahedi, A., & Derrible, S. (2021). Data-driven approach for evaluating the energy efficiency in multifamily residential buildings. *Practical Periodical on Structural Design and Construction*, 26, Article 04020074. [https://doi.org/10.1061/\(asce\)sc.1943-5576.0000555](https://doi.org/10.1061/(asce)sc.1943-5576.0000555)
- Shapley, L. S. (1953). 17. A value for n-person games. *Contributions to the theory of games (AM-28), volume ii* (pp. 307–318). Princeton: Princeton University Press.
- Shin, H., & Lee, S. (2014). Development of a climate change vulnerability index on the health care sector. *Journal Environment Policy*, 13, 69–93. <https://doi.org/10.17330/joep.13.1.201403.69>
- Smoyer, K. E. (1998). A comparative analysis of heat waves and associated mortality in St. Louis, Missouri - 1980 and 1995. *International Journal of Biometeorology*, 42, 44–50. <https://doi.org/10.1007/s004840050082>
- Song, J., Chen, W., Zhang, J., Huang, K., Hou, B., & Prishchepov, A. V. (2020). Effects of building density on land surface temperature in China: Spatial patterns and determinants. *Landscape Urban Planning*, 198, Article 103794. <https://doi.org/10.1016/j.landurbplan.2020.103794>
- Tangirala, S. (2020). Evaluating the impact of GINI index and information gain on classification using decision tree classifier algorithm. *International Journal Advance Computer Science Application*, 11, 612–619. <https://doi.org/10.14569/ijacs.2020.0110277>
- Tian, Z., Li, S., Zhang, J., Jaakkola, J. K., & Guo, Y. (2012). Ambient temperature and coronary heart disease mortality in Beijing, China: A time series study. *Environmental Health*, 11, 1–7. <https://doi.org/10.1186/1476-069x-11-56>
- Toloo, G. S., Guo, Y., Turner, L., Qi, X., Aitken, P., & Tong, S. (2014). Socio-demographic vulnerability to heatwave impacts in Brisbane, Australia: A time series analysis. *Australian and New Zealand Journal of Public Health*, 38, 430–435. <https://doi.org/10.1111/1753-6405.12253>
- Urista, D. V., Carrué, D. B., Otero, I., Arrasate, S., Quevedo-Tumaili, V. F., Gestal, M., et al. (2020). Prediction of antimalarial drug-decorated nanoparticle delivery systems with random forest models. *Biology*, 9, 198. <https://doi.org/10.3390/biology9080198>
- Vafakhab, M., Mohammad Hasani Loor, S., Pourghasemi, H. R., & Katebikord, A. (2020). Correction to: Comparing performance of random forest and adaptive neuro-fuzzy inference system data mining models for flood susceptibility mapping. *Arabian Journal of Geosciences*, 13, 1–16. <https://doi.org/10.1007/s12517-020-05637-8>
- van Loenhout, J. A. F., Vanderplanken, K., da Almeida, M. M., Kashibadze, T., Giuashvili, N., & Gamkrelidze, A. (2021). Heatwave preparedness in urban Georgia: A street survey in three cities. *Sustainability Cities Society*, 70, Article 102933. <https://doi.org/10.1016/j.scs.2021.102933>
- Wang, X., Gong, G., Li, N., & Qiu, S. (2019b). Detection analysis of epileptic EEG using a novel random forest model combined with grid search optimization. *Frontiers of Human Neuroscience*, 13, 52. <https://doi.org/10.3389/fnhum.2019.00052>
- Wang, Y., Song, Q., Du, Y., Wang, J., Zhou, J., Du, Z., et al. (2019c). A random forest model to predict heatstroke occurrence for heatwave in China. *The Science of the Total Environment*, 650, 3048–3053. <https://doi.org/10.1016/j.scitotenv.2018.09.369>
- Xu, Z., FitzGerald, G., Guo, Y., Jalaludin, B., & Tong, S. (2016). Impact of heatwave on mortality under different heatwave definitions: A systematic review and meta-analysis. *Environment International*, 89–90, 193–203. <https://doi.org/10.1016/j.envint.2016.02.007>
- Yang, G., Yu, Z., Jørgensen, G., & Vejre, H. (2020b). How can urban blue-green space be planned for climate adaption in high-latitude cities? A seasonal perspective. *Sustainability Cities Society*, 53, Article 101932.
- Yang, J., Wang, Y., Xiu, C., Xiao, X., Xia, J., & Jin, C. (2020a). Optimizing local climate zones to mitigate urban heat island effect in human settlements. *Journal of Cleaner Production*, 275, Article 123767. <https://doi.org/10.1016/j.jclepro.2020.123767>
- Yin, Q., & Wang, J. (2018). A better indicator to measure the effects of meteorological factors on cardiovascular mortality: Heat index. *Environmental Science Pollution Research*, 25, 22842–22849. <https://doi.org/10.1007/s11356-018-2396-1>

- Zeng, W., Davoodi, A., & Topaloglu, R. O. (2020). Explainable DRC hotspot prediction with random forest and SHAP tree explainer. In *2020 Design, Automation & Test in Europe Conference & Exhibition. IEEE* (pp. 1151–1156).
- Zhang, K., Li, Y., Schwartz, J. D., & O'Neill, M. S. (2014). What weather variables are important in predicting heat-related mortality? A new application of statistical learning methods. *Environmental Research*, 132, 350–359. <https://doi.org/10.1016/j.envres.2014.04.004>
- Zhang, Y., Mao, G., Chen, C., Lu, Z., Luo, Z., & Zhou, W. (2020). Population exposure to concurrent daytime and nighttime heatwaves in Huai River Basin. *China Sustainability Cities Society*, 61, Article 102309. <https://doi.org/10.1016/j.scs.2020.102309>
- Zottarelli, L. K., Sharif, H. O., Xu, X., & Sunil, T. S. (2021). Effects of social vulnerability and heat index on emergency medical service incidents in San Antonio, Texas, in 2018. *Journal of Epidemiology and Community Health*, 75, 271–276. <https://doi.org/10.1136/jech-2019-213256>