International Journal of Computer Applications (0975 – 8887)
Volume 184– No.16, June 2022

# Heat Wave Prediction using Machine Learning Techniques: A Review

F.S. Sourjah
Department of Computer Science,
Faculty of Applied Sciences,
University of Sri Jayewardenepura

W.P.J. Pemarathne
Department of Computer Science,
Faculty of Computing,
General Sir John Kotelawala Defence University

## ABSTRACT

Heat waves have become one of the major calamities in the present world because of global warming and the rise in temperature. Exceptional temperatures that are recorded over a period can be known as extreme heat events or heat waves. Even though heat waves are less exciting than other disasters like floods, tornados and earthquakes, they are an equivalent hazard to the existence of life on earth. It is critical to forecast heatwaves throughout the year and take the necessary precautions to avoid losses. Due to the advancement in machine learning (ML) techniques, ML can become handy in predicting heat waves. Our main objective is to study different existing ML algorithms that can be used to predict extreme heat events. The main approach to this study is to review the existing studies of machine learning techniques like the use of regression algorithm, K-Nearest Neighbour algorithm, Deep Learning algorithms and other significant ML algorithms that are being used to predict heatwaves and heatwave-related predictions. This research study will help to compare and contrast different ML algorithms to help this research derive the best ML model that can be developed for future heat wave predictions.

## General Terms

Machine Learning, Algorithms

## Keywords

Heat Waves, Global Warming, Hazard, Machine Learning (ML), Prediction

## 1. INTRODUCTION

Heatwaves have become a rising threat to life on earth. The recent report on climate science by the U.N indicates that the extreme heat wave that used to strike once every 50 years is now experienced very frequently due to human-induced climate change[1]. Also, the report predicts that the temperature will continue to increase with global warming. The earth's temperature is governed by how much heat energy from the sun is directed at it and how much of that energy is contained by the greenhouse gases in the earth’s atmosphere. Global warming has very adverse effects on human life. Drought, hunger and local disputes over resources killed hundreds of thousands of people in some parts of Africa. [2].

Extreme hot weather, along with high humidity, is known as a heat wave, which is a consequence of global warming. These heatwaves are hard to explain in terms of natural climatic changes without human-made climate variations. The severe 2003 heat wave, according to the World Health Organization (WHO) and different national studies, resulted in around 70,000 more fatalities, especially in France and Italy. Heat waves cause widespread agricultural loss, several wildfires and diseases[3]. Heat waves can affect the quality of the water encouraging the growth of hazardous algae that can produce toxins in drinking water, bodies, increasing the population's danger of poisoning[4] 2016.

It is very important to measure the heatwaves to estimate the danger it brings upon the human life. A threshold must be established to determine whether temperature levels above the average should be deemed a heat wave. The threshold includes not only the ambient temperature but also the thermal sensation encountered by people, relative humidity and human activities as factors[4]. The most well-known heat index has been recommended by the United States National Oceanographic Administration (NOAA). But the main drawback to the heat index is the lack of access to relative humidity information. These researchers discovered that basic indices based only on temperature may be the most useful for use in alarm systems, but that for regional analysis, all of the relevant temperature thresholds should be incorporated[5].

Heat waves can have a greater impact on highly populated places than on other locations. Preventive measures can be taken against heatwaves by creating sustainable cities and communities. That is, establishing settlements that are resilient, safe and sustainable as an adaptation to extreme heat[6].

Heat waves due to global warming and climate change have become a global issue. The negative impact of heatwaves on human life and society is tremendous. It is very important to predict extreme temperatures in the future to mitigate the consequences. Machine learning (ML) algorithms like artificial neural networks can be used to build climate models to identify heat waves[7]. The average temperature over a period can be used as data to the machine learning models to predict the thermal waves.

With the use of the implications of the foremost literature, this study presumed that the application of machine learning techniques can be used to predict the heatwaves. Therefore, this study has deduced two objectives.

1) To investigate the use of ML techniques to predict the heatwaves to mitigate the hazard caused to the human life and environment.

2) To discover means to inspire the climate scientist and researchers to initiate their work on heatwave prediction and climate changes.

The review conducted in this study has assessed previous literature available on heat wave prediction using ML techniques using a group of keywords namely heatwaves, weather changes, climate changes, global warming, heatwave prediction using ML and machine learning techniques. This study which exploits these keywords to review the proposed and implemented methods, models and applications in predicting the heatwaves using ML techniques.

# 2. LITERATURE REVIEW

## 2.1 Heat wave prediction using Regression Algorithm

Global warming and the temperature rise have caused severe heat events to occur in several parts of the world. Heat waves can cause severe damage to human life. Predicting the occurrence of heat waves is crucial in preventing human deaths from privations and sickness. A statistical model can be proposed to predict the heat waves in a chosen area. A Quantile Regression model was proposed to predict the heat waves in Pakistan. The dataset used to train this model contains the daily temperature data from Princeton's Global Meteorological Forcing (PGF). The variables used to train the model was selected from data collected from National Centers for Environmental Prediction (NCEP). The model was able forecast the days of occurrence of heat events and the departure dates in three months May, June and July with an accuracy of about ±5 days. The evaluation of the variables used in the training of the model revealed that the relative humidity and the wind were two determining factors of the heat in Pakistan. Furthermore, the performance of the Quantile Regression model (QRF) was compared to that of the Random Forest model (RF). QRF produced better predictions compared to RF, proving its efficacy [8].

The [9] publication from the EGU Assembly held in April 2015 in Vienna shows that there is a huge impact of urbanization on extreme temperatures, leading to severe heat events occurring in metropolitan areas. To mitigate temperature in the urban areas, this impact has drawn much attention. The research study [10] integrates spatial analysis (satellite measurements), ground based measurements and machine learning to create a model that can predict the urban areas with extreme heat waves. Using sensors mounted to automobiles with a global positioning system (GPS), temperature data was collected from a few urban cities on days with temperatures in the 90th percentile of historic averages. A Random Forest Regression model was used to do predict temperatures with high predictive power and low Root mean square error (RMSE). An evaluation technique of cross validation was done to avoid the over fitting of data. 70 percent of the data was used for training the model and 30 percent was used for testing [10].

A quantile regression model is trained to predict the urban city temperature. The dataset used combines sensor data from different sources, data from spatial and meteorological predictors. All this data was collected between June 25 and June 30, 2019, and a total of 85,942 data points were used. The data collected (16 per cent) on $30^{th}$ June was used to test the model. The model was evaluated in terms of RMSE. The objective of this study is to create high resolution temperature maps of urban areas to see extreme heat events and avoid the serious ill effects that can occur[11].

An overpopulated metropolitan city of Seoul in South Korea was chosen as the area of the study to create a model that can predict extreme heat events. From 2013 to 2017, the next-day maximum and lowest air temperatures were acquired from 25 Automatic Weather Stations (AWSs) run by KMA in Seoul throughout July and August. The current day's maximum and lowest air temperatures were calculated using hourly air temperature data from the 25 (AWS)s. Some of the feature variables used to train the model were Latitude, longitude, elevation, slope, and solar radiation. Machine learning (ML) models were used to create prediction models on the data. Random Forest model which is the best model for multi-class classification and regression problems[12][13] was trained. Second model, an Artificial neural network (ANN) which has a networked structure that mimics the functions and connections of actual neurons in the human brain was used [14][15]. Also, a Support Vector Regression algorithm (SVR) was used to result prediction with minimum errors. Finally, a Multi Model Ensemble was used after comparing the results of other ML models. MME combines multiple ML models to achieve better accuracy and resilience [16][17].

Global warming has caused a high heat flow in the region of Antarctica. High Geo Thermal (GT) heat flow can thwart the glaciers in the arctic region [18]. Melting of Ice in the northern region can result in numerous adverse effects like sea level rising, impact on climate to cause extreme weather events, extinction of vulnerable animal species, etc. A Radient Boosted Regression Tree algorithm is used to predict GT heat flow in Antarctica. Dataset was created by collecting several reliable geophysical and geological global datasets. The normalized root-mean-square error (RMSE/mean) of the test set, as well as the model score and $R^2$ score, are calculated in the same way as [19]. In the test set, the established prediction model has an $R^2$ score of 0.44 and a relative RMSE of 29%. For the model evaluation the predicted values are compared with the actual values [20].

Land Surface Temperature (LST) observed by satellites is an important factor in predicting heat waves. But due to the disturbances from the clouds, the clarity of the LST recorded can be poor, which can affect the poor prediction of the heatwaves [26]. A regression model was developed to predict the LST, which will be used in accurate heatwave prediction. The LST predicted using the ML model will fill in the gap between the satellite predicted LST and the actual LST[21]. The heatwave days per annum in Iran were predicted by extracting the daily temperature from 27 points and trained using Ada-Boost Regression and decision tree (ABR-DT). The feature variables used to train the data were extracted for four different pressure levels. The model was evaluated by a grid-point based evaluation by obtaining a corelation coefficient of about 0.86 and a mean squared error of 6.929. The research work identified the best set of feature variables to predict the heatwave days in a year [22].

There are few existing systems that can predict the number of heat strokes. Most of these systems are developed only using meteorological variables, and none of the socio-economic variables are in cooperation, which leads to a lack of verification of these models and poor accuracy[23]. A machine learning model was developed using both meteorological and economic variables to predict heat stroke to reduce the impact of the extreme heat events that occur in the cities of China. The daily heat stroke occurrences of certain cities in China in 2012- 2014 were collected. The socio-economic variables were collected from the statistical year books of the cities in China. The novelty of this study is a search index factor is used in the model. The term “Heat Stroke” was used as the key word. The network search data was compiled using China's most popular search engine. The search index was provided daily as a public data source, and it was proven to have a greater connection with other meteorological characteristics such as temperature and relative humidity, demonstrating that the search index is a reasonable predictor [23]. A Random Forest model was used to train the data. The Boruta algorithm was used to filter the variables and maybe shorten the time it took to calculate the follow-up random forest model establishment [24]. 90 per cent of data was trained and 10 per cent was used for testing and validation. The model performance was analysed using the linear fitting approach, and the model fitness was measured using linear $R^2$[25].

Occurrence and the impact of heatwaves is higher in urban areas than the rest. The extreme temperature in the urban areas can greatly affect the vulnerable population like the elderly and the infants [26][27][28]. The fatality rates due to heatwaves were greater than 4000 in the urban city of Paris in the year 2003 [29]. Predicting higher air temperatures can be helpful in designing early warning systems to alert the population of extreme heat events. Crowd sourced data of higher air temperatures in the city of Berlin was used to train ML models like Random Forest (RF), Stochastic Gradient Boosting, and Model Averaged Neural Network comparativel. RF model performed well with a $R^2$of 0.512. One of the novelties of this work is that it is done using open source datasets and therefore can be adapted to predict extreme temperatures in other parts of the world[30].

## 2.2 Heat Wave Prediction using K-Nearest Neighbour Algorithm

A prediction tool was designed integrating the K-Nearest Neighbors model and weighted moving average algorithm to predict the flood and heatwaves that can occur using the weather data obtained from the city of Mangalore, India. A heat wave is recognized by looking for abnormal temperature behaviour in weather data over a long period of time. The model was evaluated by comparing the predicted temperature against the actual average temperature [31].

Extreme climatic events and heatwaves can destroy cultivation due to the lack of irrigation and finally can result in droughts. Drought forecasting is extremely useful in providing early warning and preparing the most vulnerable areas for the worst effects of drought. Three ML models the K-Nearest Neighbours, Support-Vector machine and Artificial Neural Network were used to create a model to forecast droughts. The data set consist of data obtained from National Centres for Environmental Prediction/National Centre for Atmospheric Research (NCEP/NCAR) reanalysis database. A unique feature selection technique called Recursive Feature Elimination (RFE) was also employed for the first time in drought modelling to determine optimum sets of predictors. In validation, KNN, which was utilized for the first time in building drought models, performed poorly when compared to SVM and ANN-based drought models. The main limitation of this study is the use of one dataset can lead to some uncertainties of prediction [32].

## 2.3 Prediction of Heatwaves using Deep Learning Algorithms

Extremely hot summers and heatwaves pose a serious threat to human life, even though people are not as conscious of heatwaves as they are of other calamities. There were so many reported incidents of high death tolls due to extreme heatwaves over many parts of the world, like over Russia in 2010[33], [34]or during the summer of 2021 in North America[35] or in western Europe during the summer of 2003[36]. Therefore, predicting heatwaves is important to mitigate the losses. A physics driven approach to predict heatwaves need a significant amount of computational cost, time and resources. Deep neural network models can overcome these challenges and perform well in heatwave predictions [37]. A Convolution Neural Network (CNN) was used to train a very large dataset consisting of 1000 years of climate model data with the use of Transfer Learning. The training and the test datasets were split into 900 and 100 years, respectively, by random sampling. This was a supervised classification problem trained using a CNN with a 4 layered architecture. Under sampling techniques were used to overcome the challenge of having imbalanced classes in the training set to predict the rare events.[38]. In predicting high intensity heatwaves Transfer Learning was used since under sampling can degrade the performance [39].

Marine Heat waves (MHW) can be a huge threat to the aquatic fauna, flora and millions of humans whose main livelihood depends on the sea. Marine heatwaves can be a great impact to the biodiversity in and out of the ocean. The extreme heat event occurred in 2011 destroyed hundred kilometres of underwater ecosystem wiping out very rare species of aquatic plants like kelp forests, seagrass meadows, and coral reefs in the western coast of Australia [40]. It is important to forecast these extreme under water events to prevent the aquatic life. The research work suggests to train several modified CNN like Spherical CNNs, CNN-LSTMs, CNN-AEs and GANs and GNNs on Spatiotemporal data.The evaluated CNNs were integrated to achieve a better MHW prediction [41]. LSTM a recurring neural network was integrated with the existing convolution neural network to train on satellite temperature data to predict the Ocean heat waves in Korea. The prediction of the model was evaluated by the RMSE between the predicted and observed values [42].

## 2.4 Prediction of Heatwaves using other ML Algorithms

The industrialization and and the high emission of harmful gases into the atmosphere have caused a rise in the temperature of the earth. A predictive model was created using cellular automata (CA) and artificial neural networks (ANN) to predict the surface temperatures. The land satellite images from five-year intervals from 1995-2020 were used to create a dataset to train the models to forecast temperature. Study computed overall accuracy, user accuracy, producer accuracy, and Kappa statistics. Regression analysis was used to check the accuracy. One of the limitations of this study was the CA model has limited ability to identify the explicit relationship between the influential variables [43].

Predicting the heat wave days in a year is important to get prepared for the losses that may have to be faced in the future. Many of the heat wave prediction models perform poor due to the constant and inevitable changes in the climate [44]. The change in the ocean-atmospheric variable can affect the performance of these forecast models in higher levels. Therefore, a constant update in the models is vital to maintain the models reliable. A heat wave day (HWS) prediction model was developed using the Atmospheric Research reanalysis data to predict the occurrence of heatwaves in Pakistan. The machine learning models used in the work are Support Vector Machines (SVM), random forest and artificial neural networks. The SVM model was selected comparatively based on its best performance, with an $R^2$ value of 0.86, to predict the heat wave days. An approach of updating the link between the predictor and predictor was used to update the model constantly to keep up with climate change[45].

Extreme weather conditions like heatwaves can be forecasted by proper prediction of temperature [46]. Time Series Time-Delay Neural Network, a feedforward network was trained on the temperature dataset obtained from the Morocco meteorological administration to predict accurate temperature. The model predicted results with a corelation coefficient R of about 0.99. The model predicted temperatures were evaluated against the actual temperature values [47].

## 2.5 Prediction of Heat Wave Related Factors

The climate issue has resulted in high-intensity heatwaves that have wreaked havoc on human existence and caused innumerable deaths. It is very important to take actions to prevent and minimize the damages caused by these heat events. Depending merely on the human instincts in taking measures to disaster can lead to unproductive decision-making. Artificial intelligence-based analytical models can be used to make effective solutions during catastrophe response. A Random Forest model, a strong ML model that can be applied both for classification and regression [48] was used to predict the heat wave related damages. The dataset was created using the statistical, meteorological, and floating population data from South Korea in the years 2015-2018. The Random Forest model was evaluated by comparing other traditional regression models in terms of the coefficient of determination, mean absolute error, root mean squared error, root mean squared logarithmic error, and mean absolute error ($R^2$). The suggested model has an R2 value of 0.804 in a comparison with observed data. The loss functions were used to evaluate the results with the actual values. The loss functions of the random forest model used in this study were the Mean Absolute Error (MAE) and the Mean Square Error (MSE) [49]

While predicting heatwaves are important its vital to predict the heat deaths to understand the intensity of the heatwaves. Heat deaths can occur by exposure to the excessive heat. Predicting the number of heat deaths can give us an idea about how much the heatwaves going to affect the human life. A zero-inflated regression model was trained on Korean statistical data to predict the number of heat deaths in Korea. Some of the main parameters on which the model was developed are the weekly mean of the daily temperature, the number of elderly people in the area, the number of heatwave days and the vulnerability of people to heatwaves. The model was evaluated by comparing the predicted number of heat deaths with the actual number of heat deaths [50]. Another regression model was used to predict the number of deaths due to heat disorders (NDHD). This model assesses the direct impact of heatwaves on the human fatality rates. The model was trained on the National Statistics database on deaths due to heatwaves in the period 1994–2012. The model produced a high $R^2$ value of 0.96 and a very low mean square error of 1.4 deaths. The model was evaluated by using scatter and time series plots to compare the estimated and observed fatalities. Also, the prediction results from this model show that there will be extreme heatwaves in the year 2050 that will cause about 250 deaths[51].

# 3. RESEARCH METHODOLOGY

A systematic literature review has been conducted to study and review the recent literature on heatwaves and heatwave-related predictions with the aim of discovering how much work has been done using ML techniques in heat wave prediction. Searching online research databases such as Google Scholar, IEE Explorer and Science Direct yielded high-quality research articles. Specific keywords were used to search both conference and journal articles in the database. Only recent research studies (from 2016 to 2021) were chosen for the review. After the selection of relevant research work, a filtration process was carried out to select the best quality and most relevant articles.

The online databases were searched with the keywords "Heatwave prediction using machine learning" and "Extreme temperature prediction using machine learning". When searching the databases, there were filters applied to only show the research articles that had been done after the year 2016. Only the articles related to subject areas Environmental science or Computer Science were selected. 31 articles which are most relevant and most recent were selected for thescreening process. The acquired research articles were filtered with the following inclusion and exclusion factors.

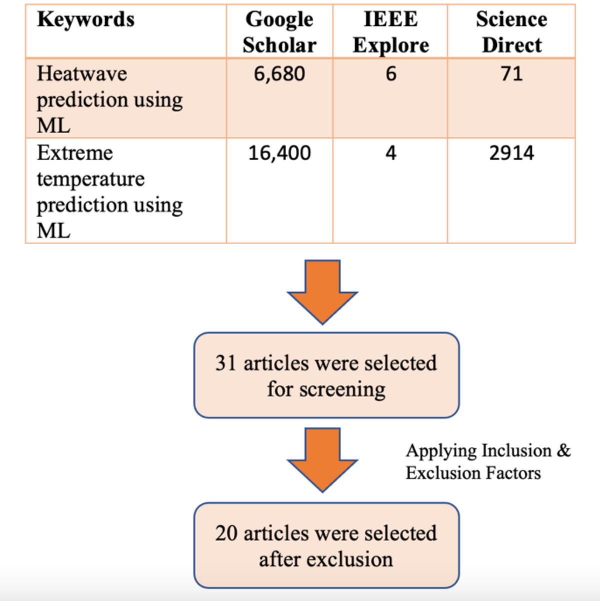
Inclusion Factors

- Topics related to heatwave and heatwave related predictions.
- Only the predictions using machine learning techniques.
- Only the articles technically addressing the heatwave prediction.

Exclusion Factors

- Physics-driven heatwave prediction articles.
- Literature reviews on the heatwave predictions.

After the exclusion of the 31-research article used for the screening process, finally 20 articles were selected for the review of this study. Figure 1 shows the systematic filteringprocess of the research works.



| Keywords | Google Scholar | IEEE Explore | Science Direct |
|---|---|---|---|
| Heatwave prediction using ML | 6,680 | 6 | 71 |
| Extreme temperature prediction using ML | 16,400 | 4 | 2914 |



**Fig 1: Systematic filtering process of the research works**

## 4. RESULTS

A deep analytical literature review was done on 17 out of 20 (85%) papers which included work on heatwave predictions (HWP). When reviewing this category, the research works were categorized according to the ML algorithm used by the research work. 7 out of 17 (41%) papers were reviewed under the topic “Heatwave predictions using Regression models”. 2 papers out of 17 (12%) were reviewed under the topic “Heatwave predictions using Random Forest models”. 2 papers (12%) were reviewed under the topic “Heatwave predictions using K-Nearest Neighbour models”. 3 out of 17 papers (18%) were under the topic “Heatwave predictions using Deep Learning models”. The rest of the 3 papers were studied under the topic “Heatwave predictions using other ML models”. This approach of study allows us to investigate the application of different ML algorithms in predicting the heatwaves.

Apart from the heatwave predictions its crucial to predict the heatwave related factors too. Identification of how much damage and deaths can an extreme weather like heatwaves can cause allow us to predict the intensity of the heatwaves. This also will help in taking necessary measures in protecting human lives and properties. Therefore 3 such research works were reviewed under the topic “Heatwave Related Predictions”.

## 5. DISCUSSION.

In this study, the research articles were reviewed to understand the key points such as the ML algorithm used, the dataset used, the evaluation of the model, and the limitations of the work. This kind of study can give a complete idea of how much work has been done in heatwave prediction and can pave the way for future developments in this area.

Few of the research papers reviewed in this literature review used multiple algorithms or the integration of multiple models to forecast heatwaves. But when grouping according to the used ML algorithms, the model that is primarily used for the study is used. This study not only reviews the heatwave prediction work but also considers the work related to the heatwaves. Predicting the damage and deaths caused by heatwaves is as important as predicting the heatwaves themselves to take necessary prevention measures.

The databases like IEEE and Science Direct yielded very few results when the keyword "Heatwave prediction using ML" was used to search. When the same keyword was used in other databases like Microsoft academic and Taylor and Francis online, it yielded no results at all. That is one reason the authors had to search a social research platform like Google Scholar. When the same keyword was searched in the database Google Scholar, it produced considerably large results. But the relevance of the work was not compatible. So, a great effort was made to select the most recent and most relevant work (31 articles) for the screening process. Since there were low results for the keyword "heatwave prediction using ML," another related keyword, "extreme temperature prediction using ML," was used.

The result of this study shows that 41% of the literature used regression models for prediction. Although there are so many advancements in the ML field, like the latest deep learning and transfer learning techniques, a considerably low amount of work has been done using those models. Even when searching the databases, the results of literature using deep learning for heat wave prediction were low. This work emphasizes the fact that more work can be done to predict heatwaves in the future using advanced ML techniques like deep learning.

## 6. CONCLUSION, LIMITATIONS, AND FUTURE DEVELOPMENTS.

This systematic literature review has been done by deeply reviewing recent studies on heatwave predictions. For the review, the research works were filtered and methodically analyzed. The results of this show how much work has been done in heatwave predictions using different ML models. According to the results, regression models are widely used for heatwave predictions, and the amount of work that has been done using the other ML models is considerably low. Significantly, the literature using advanced ML techniques like deep learning models is low.

One of the limitations of this work is that there is not adequate recent literature available in the databases. The amount of work done in the field of heatwave prediction is very low, even though heatwaves are a major climate threat. A greater effort was put into selecting research papers needed for the screening process. After the exclusion of papers, limited number studies are eligible for the review.

Since there are so many recent advancements in the use of ML techniques, a better heatwave prediction system can be built using deep learning models and transfer learning with better accuracy and performance as the future development of this literature review.

## 7. REFERENCES

[1]X. Gu *et al.*, “Attribution of Global Soil Moisture Drying to Human Activities: A Quantitative Viewpoint,” *Geophys. Res. Lett.*, vol. 46, no. 5, pp. 2573–2582, Mar. 2019, doi: 10.1029/2018GL080768.

[2]M. Munasinghe, “Rising temperatures, rising risks,” *Finance Dev.*, vol. 45, Mar. 2008.

[3]W. Marx, R. Haunschild, and L. Bornmann, “Heat waves: a hot topic in climate change research,” *Theor. Appl. Climatol.*, vol. 146, no. 1–2, pp. 781–800, Oct. 2021, doi: 10.1007/s00704-021-03758-y.

[4]P. Martinez and E. R. Bandala, “Heat Waves: A Growing Climate Change-related Risk,” p. 4.

[5]M.-A. P. F and B. E. R., “Heat Waves: Health Effects, Observed Trends and Climate Change,” in *Extreme Weather*, P. J. Sallis, Ed. InTech, 2018. doi: 10.5772/intechopen.75559.

[6]R. Ortlepp and D. Schiela, “Perspectives of climate change adaptation of building areas against heat waves,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 615, no. 1, p. 012004, Oct. 2019, doi: 10.1088/1757-899X/615/1/012004.

[7]J. Safieh, R. Dara, and J. Forough, “Effect of climate change on heat waves in the South Sea of Iran,” *Ukr. J. Ecol.*, vol. 10, pp. 87–93, Oct. 2020, doi: 10.15421/2020_211.

[8]N. Khan, S. Shahid, L. Juneng, K. Ahmed, T. Ismail, and N. Nawaz, “Prediction of heat waves in Pakistan using quantile regression forests,” *Atmospheric Res.*, vol. 221, pp. 1–11, Jun. 2019, doi: 10.1016/j.atmosres.2019.01.024.

[9]M. Wang, X. Yan, J. Liu, and X. Zhang, “The contribution of urbanization to recent extreme heat events and a potential mitigation strategy in the Beijing–Tianjin–Hebei metropolitan area,” *Theor. Appl. Climatol.*, vol. 114, no. 3–4, pp. 407–416, Nov. 2013, doi: 10.1007/s00704-013-0852-x.

[10]V. Shandas, J. Voelkel, J. Williams, and J. Hoffman, “Integrating Satellite and Ground Measurements for Predicting Locations of Extreme Urban Heat,” *Climate*, vol. 7, no. 1, p. 5, Jan. 2019, doi: 10.3390/cli7010005.

[11]M. Zumwald, B. Knüsel, D. N. Bresch, and R. Knutti, “Mapping urban temperature using crowd-sensing data and machine learning,” *Urban Clim.*, vol. 35, p. 100739, Jan. 2021, doi: 10.1016/j.uclim.2020.100739.

[12]S. Lee *et al.*, “Arctic Sea Ice Thickness Estimation from CryoSat-2 Satellite Data Using Machine Learning-Based Lead Detection,” *Remote Sens.*, vol. 8, no. 9, p. 698, Aug. 2016, doi: 10.3390/rs8090698.

[13]S. Yoo, J. Im., and J. E. Wagner, “Variable selection for hedonic model using machine learning approaches: A case study in Onondaga County, NY,” *Landsc. Urban Plan.*, vol. 107, no. 3, pp. 293–306, Sep. 2012, doi: 10.1016/j.landurbplan.2012.06.009.

[14]R. Özçelik, M. J. Diamantopoulou, J. R. Brooks, and H. V. Wiant, “Estimating tree bole volume using artificial neural network models for four species in Turkey,” *J. Environ. Manage.*, vol. 91, no. 3, pp. 742–753, Jan. 2010, doi: 10.1016/j.jenvman.2009.10.002.

[15] S. Tiryaki and A. Aydın, “An artificial neural network model for predicting compression strength of heat-treated woods and comparison with a multiple linear regression model,” *Constr. Build. Mater.*, vol. 62, pp. 102–108, Jul. 2014, doi: 10.1016/j.conbuildmat.2014.03.041.

[16]P. J. L. Adeodato, A. L. Arnaud, G. C. Vasconcelos, R. C. L. V. Cunha, and D. S. M. P. Monteiro, “MLP ensembles improve long term prediction accuracy over single networks,” *Int. J. Forecast.*, vol. 27, no. 3, pp. 661–671, Jul. 2011, doi: 10.1016/j.ijforecast.2009.05.029.

[17]D. Cho, C. Yoo, J. Im, and D. Cha, “Comparative Assessment of Various Machine Learning- Based Bias Correction Methods for Numerical Weather Prediction Model Forecasts of Extreme Air Temperatures in Urban Areas,” *Earth Space Sci.*, vol. 7, no. 4, Apr. 2020, doi: 10.1029/2019EA000740.

[18]R. Dziadek, F. Ferraccioli, and K. Gohl, “High geothermal heat flow beneath Thwaites Glacier in West Antarctica inferred from aeromagnetic data,” *Commun. Earth Environ.*, vol. 2, no. 1, p. 162, Dec. 2021, doi: 10.1038/s43247-021-00242-3.

[19]S. Rezvanbehbahani, L. A. Stearns, A. Kadivar, J. D. Walker, and C. J. van der Veen, “Predicting the Geothermal Heat Flux in Greenland: A Machine Learning Approach,” *Geophys. Res. Lett.*, vol. 44, no. 24, p. 12,271-12,279, 2017, doi: 10.1002/2017GL075661.

[20]M. Lösing and J. Ebbing, “Predicting Geothermal Heat Flow in Antarctica With a Machine Learning Approach,” *J. Geophys. Res. Solid Earth*, vol. 126, no. 6, Jun. 2021, doi: 10.1029/2020JB021499.

[21]I. Buo, V. Sagris, and J. Jaagus, “Gap-Filling Satellite Land Surface Temperature Over Heatwave Periods With Machine Learning,” *IEEE Geosci. Remote Sens. Lett.*, pp. 1–5, 2021, doi: 10.1109/LGRS.2021.3068069.

[22]S. B. H. S. Asadollah, N. Khan, A. Sharafati, S. Shahid, E.-S. Chung, and X.-J. Wang, “Prediction of heat waves using meteorological variables in diverse regions of Iran with advanced machine learning models,” *Stoch. Environ. Res. Risk Assess.*, Oct. 2021, doi: 10.1007/s00477-021-02103-z.

[23]T. Li, F. Ding, Q. Sun, Y. Zhang, and P. L. Kinney, “Heat stroke internet searches can be a new heatwave health warning surveillance indicator,” *Sci. Rep.*, vol. 6, no. 1, p. 37294, Dec. 2016, doi: 10.1038/srep37294.

[24]M. B. Kursa and W. R. Rudnicki, “Feature Selection with the **Boruta** Package,” *J. Stat. Softw.*, vol. 36, no. 11, 2010, doi: 10.18637/jss.v036.i11.

[25]Y. Wang *et al.*, “A random forest model to predict heatstroke occurrence for heatwave in China,” *Sci. Total Environ.*, vol. 650, pp. 3048–3053, Feb. 2019, doi: 10.1016/j.scitotenv.2018.09.369.

[26]R. S. Kovats and S. Hajat, “Heat Stress and Public Health: A Critical Review,” *Annu. Rev. Public Health*, vol. 29, no. 1, pp. 41–55, 2008, doi: 10.1146/annurev.publhealth.29.020907.090843.

[27]K. M. A. Gabriel and W. R. Endlicher, “Urban and rural mortality rates during heat waves in Berlin and Brandenburg, Germany,” *Environ. Pollut.*, vol. 159, no. 8, pp. 2044–2050, Aug. 2011, doi: 10.1016/j.envpol.2011.01.016.

[28]D. Scherer, U. Fehrenbach, T. Lakes, S. Lauf, F. Meier, and C. Schuster, “Quantification of heat-stress related mortality hazard, vulnerability and risk in Berlin, Germany,” *ERDE – J. Geogr. Soc. Berl.*, vol. 144, no. 3–4, Art. no. 3–4, 2013, doi: 10.12854/erde-144-17.

[29]B. Dousset *et al.*, “Satellite monitoring of summer heat waves in the Paris metropolitan area,” *Int. J. Climatol.*, vol. 31, no. 2, pp. 313–323, 2011, doi: 10.1002/joc.2222.

[30]S. Vulova, F. Meier, D. Fenner, H. Nouri, and B. Kleinschmit, “Summer Nights in Berlin, Germany: Modeling Air Temperature Spatially with Remote Sensing, Crowdsourced Weather Data, and Machine Learning,” *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 13, pp. 5074–5087, 2020, doi: 10.1109/JSTARS.2020.3019696.

[31] C. S. Bangera, P. S. Kotian, C. Dias, T. Divya, and G. Aithal, “Flood and Heat Wave Prediction using Weighted Moving Average, Anomaly Detection and K-Nearest Neighbours for the City Of Mangalore,” in *2018 IEEE Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER)*, Mangalore (Mangaluru), India, Aug. 2018, pp. 93–97. doi: 10.1109/DISCOVER.2018.8674113.

[32]N. Khan, D. A. Sachindra, S. Shahid, K. Ahmed, M. S. Shiru, and N. Nawaz, “Prediction of droughts over Pakistan using machine learning algorithms,” *Adv. Water Resour.*, vol. 139, p. 103562, May 2020, doi: 10.1016/j.advwatres.2020.103562.

[33]D. Barriopedro, E. M. Fischer, J. Luterbacher, R. M. Trigo, and R. García-Herrera, “The Hot Summer of 2010: Redrawing the Temperature Record Map of Europe,” *Science*, vol. 332, no. 6026, pp. 220–224, Apr. 2011, doi: 10.1126/science.1201224.

[34]F. E. L. Otto, N. Massey, G. J. van Oldenborgh, R. G. Jones, and M. R. Allen, “Reconciling two approaches to attribution of the 2010 Russian heat wave,” *Geophys. Res. Lett.*, vol. 39, no. 4, 2012, doi: 10.1029/2011GL050422.

[35]S. Y. Philip *et al.*, “Rapid attribution analysis of the extraordinary heatwave on the Pacific Coast of the US and Canada June 2021,” *Earth Syst. Dyn. Discuss.*, pp. 1–34, Nov. 2021, doi: 10.5194/esd-2021-90.

[36]R. García-Herrera, J. Díaz, R. Trigo, J. Luterbacher, and E. Fischer, “A Review of the European Summer Heat Wave of 2003,” *Crit. Rev. Environ. Sci. Technol.*, vol. 40, pp. 267–306, Mar. 2010, doi: 10.1080/10643380802238137.

[37]J. A. Weyn, D. R. Durran, and R. Caruana, “Can Machines Learn to Predict Weather? Using Deep Learning to Predict Gridded 500-hPa Geopotential Height From Historical Weather Data,” *J. Adv. Model. Earth Syst.*, vol. 11, no. 8, pp. 2680–2693, 2019, doi: 10.1029/2019MS001705.

[38]J. Johnson and T. Khoshgoftaar, “The Effects of Data Sampling with Deep Learning and Highly Imbalanced Big Data,” *Inf. Syst. Front.*, vol. 22, Oct. 2020, doi: 10.1007/s10796-020-10022-7.

[39]V. Jacques-Dumas, F. Ragone, P. Borgnat, P. Abry, and F. Bouchet, “Deep Learning-based Extreme Heatwave Forecast,” *ArXiv210309743 Phys.*, Oct. 2021, Accessed: Nov. 29, 2021. [Online]. Available: http://arxiv.org/abs/2103.09743

[40]N. J. Holbrook *et al.*, “Keeping pace with marine heatwaves,” *Nat. Rev. Earth Environ.*, vol. 1, no. 9, pp. 482–493, Sep. 2020, doi: 10.1038/s43017-020-0068-4.

[41]D. Ning, V. Vetrova, K. Bryan, and S. Delaux, “Deep Learning for Spatiotemporal Anomaly Forecasting: A Case Study of Marine Heatwaves,” presented at the ICML 2021 Workshop on Tackling Climate Change with Machine Learning, Jul. 2021. Accessed: Nov. 29, 2021. [Online]. Available: https://www.climatechange.ai/papers/icml2021/87

[42]S. Jung, Y. J. Kim, S. Park, and J. Im, “Prediction of Sea Surface Temperature and Detection of Ocean Heat Wave in the South Sea of Korea Using Time-series Deep-learning Approaches,” *Korean J. Remote Sens.*, vol. 36, no. 5_3, pp. 1077–1093, Oct. 2020, doi: 10.7780/KJRS.2020.36.5.3.7.

[43]A.- Al Kafy *et al.*, “The operational role of remote sensing in assessing and predicting land use/land cover and seasonal land surface temperature using machine learning algorithms in Rajshahi, Bangladesh,” *Appl. Geomat.*, Aug. 2021, doi: 10.1007/s12518-021-00390-3.

[44]B. Wang *et al.*, “Rethinking Indian monsoon rainfall prediction in the context of recent global warming,” *Nat. Commun.*, vol. 6, no. 1, p. 7154, Nov. 2015, doi: 10.1038/ncomms8154.

[45]N. Khan, S. Shahid, T. B. Ismail, and F. Behlil, “Prediction of heat waves over Pakistan using support vector machine algorithm in the context of climate change,” *Stoch. Environ. Res. Risk Assess.*, vol. 35, no. 7, pp. 1335–1353, Jul. 2021, doi: 10.1007/s00477-020-01963-1.

[46]Tushar Kanti Routh *et al.*, “Artificial neural network-based temperature prediction and its impact on solar cell,” in *2012 International Conference on Informatics, Electronics & Vision (ICIEV)*, Dhaka, Bangladesh, May 2012, pp. 897–902. doi: 10.1109/ICIEV.2012.6317369.

[47]A. Kabbori, J. Antari, R. Iqdour, and Z. E. Abidine El Morjani, “Temperature Prediction using Time Series Time-Delay Neural Networks,” in *2019 7th International Renewable and Sustainable Energy Conference (IRSEC)*, Agadir, Morocco, Nov. 2019, pp. 1–4. doi: 10.1109/IRSEC48032.2019.9078299.“randomforest2001 .pdf.”

[49]M. Park, D. Jung, S. Lee, and S. Park, “Heatwave Damage Prediction Using Random Forest Model in

Korea," *Appl. Sci.*, vol. 10, no. 22, p. 8237, Nov. 2020, doi: 10.3390/app10228237.

[50]D.-W. Kim, "Weekly heat wave death prediction model using zero-inflated regression approach," p. 16.

[51]D.-W. Kim, R. C. Deo, J.-H. Chung, and J.-S. Lee, "Projection of heat wave mortality related to climate change in Korea," *Nat. Hazards*, vol. 80, no. 1, pp. 623–637, Jan. 2016, doi: 10.1007/s11069-015-1987-0.