ISSN No:-2456-2165

Urban Heat Island Prediction Using ANN

Gitika Kaur[1], Garima Sharma[2], Dr. Meenu Vijarania[3], Swati Gupta[4]
[1]Amity University Haryana ,122413, India

Abstract:- A urban heat island (UHI) is an urban region or territory that is fundamentally hotter than its encompassing rural zones because of man practices. The UHI can be explained as either the surface temperature difference or air temperature difference between the urban and the rural regions. The project is based on the study of "Urban Heat Island prediction using ANN" grounded on the values of various test results done on an individual technique. The technique used for this project is the Artificial Neural Network (ANN) algorithm and Time Series analysis to correctly predict a region as urban heat island. The data of six weather stations (WS) was collected each having 12929 instances to determine whether a particular region forms UHI or not. The ANN model was implemented on all the six weather stations and after examining the ANN model, the two weather stations WS4 and WS6 proved out to be the best in terms of correlation between dependent and independent variable that was evaluated using MAE and R2 score. The two weather stations WS6 and WS4 having R2 score or accuracy as 79.6 and 79.3 respectively was further chosen for time series analysis. In time series analysis, we just have one variable i.e., time. We can analysis the time series data in order to extract meaningful insights and other features. Time series is a set of observation taken at specified times at equal interval.

Keywords:- *Urban Heat Island Prediction, ANN, Time Series Analysis.*

I. INTRODUCTION

Temperatures in urban zones have for quite a while been concentrated as a limit of metropolitan morphology, land use and anthropogenic activities than just focus meteorological boundaries. The wonder of urban heat island is currently an ongoing issue of worry in numerous urban areas of created as well as creating countries of the world.

An urban heat island (UHI) is a metropolitan region or territory that is fundamentally hotter than its encompassing rural zones because of man practices. The word heat island is used to reflect any area that is comparatively hotter than the surrounding, but generally describe human-disturbed areas. The temperature difference is generally bigger around evening time than throughout the day, and is most obvious when breeze [1]

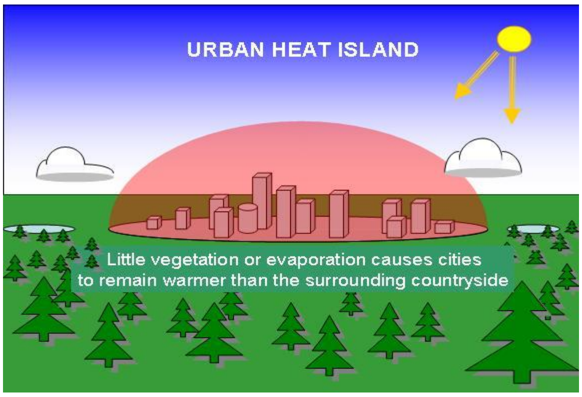

Figure 1: An illustration of an urban heat island.

1.2 Why does this happen?

The primary reason of the UHI impact is from the change of land surfaces. There are a few reasons for an UHI; for instance, dark surfaces retain essentially progressively sunlight-based radiation, which makes urban groupings of streets and structures heat extra than rural and provincial territories throughout the day.

Materials regularly utilized in urban regions for asphalt and rooftops, for example, cement and black-top, have fundamentally extraordinary thermal mass properties and surface radiative properties than the encompassing rural regions.

1.3 Machine Learning and UHI

Machine learning (ML) is the process of making our machine or model learn the correlations between various input and output features and based on what the model learns it is capable of making predictions on a new data. [2]

Figure 2: Machine Learning

ML has seen a tremendous growth in predictions of UHI because of its high accuracy and large datasets. Number of machines learning algorithms are being used by researchers in the prediction of UHI and estimating the important parameters causing UHI.

1.3 Deep Learning (DL)

Deep learning is an artificial intelligence that imitate the functions of the human mind in processing data and pattern creation. DL is a subset of machine learning which is based on ANN (artificial neural networks). So, deep learning is a kind of mimic of human brain.

The whole concept of deep learning is to mimic how human brain works. It involves an input, hidden and output layer. We use activation functions in neural networks. Usually, rectifier function is used in the hidden layers and for output layers we use sigmoid.[3]

We have parameters like batch size and epochs which can be changed so as to get much higher accuracies.

Deep learning algorithms are:

- **Artificial Neural Networks (ANN)** – used for regression and classification.
- Convolutional Neural Networks (CNN) – used for dealing with images.
- Recurrent Neural Networks (RNN) – applicable in time series analysis.

1.5 Artificial Neural Network

The whole perception of ANN is to mimic how human brain works. ANN involves an input layer, hidden layer(s), and an output layer. Each of these layers have a number of neurons and synapses which are responsible for propagating input data through our network. In the ANN we have the unit of calculation which are called neurons. These neurons are attached by the weight values called as synapses. This means given a number, a neuron will perform some calculations with the help of function known as activation function.

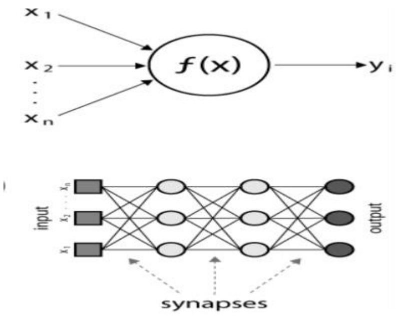

Figure 3: Artificial Neural Network

In the project, the Feed Forward ANN has been implemented. In this, the flow of information is unidirectional. A unit sends data to other unit from which it does not get any data. There is no feedback allowed. They have fixed number of inputs.

This paper is ordered as follows: section 2 related work reviews about the great work done in the area of urban heat island detection model, section 3 deals with methods and materials used, section 4 carries out the actual implementation work done followed by section 5 contains results and analysis of the model, section 6 concludes with some directions for future work.

II. RELATED WORK

A lot of work has been done on this particular topic of urban heat island prediction. This section survey some of the related work already done in this field.

In recent years the UHI phenomenon has become one of the most growing environmental problems in the urban areas due to the land cover, urbanization and population density [4]. UHI can be assessed by applying thermal remote sensing to perform land cover classifications and thermal behavior of various urban surfaces which varies in response to the surface energy balance [5]. Besides that, UHI detection can also be performed using fixed-station or meteorological station and car traverse measurement to detect UHI intensity and the air temperature in the regions [6].

Several research studies have been performed on the UHI phenomenon. The UHI intensity differs between seasonal climate and tropical cities. In the seasonal climate, the UHI are strongest in the summer or winter season. According to [7,8], the UHI intensity is positively correlated with solar radiation and relative humidity during summer while on the other hand it is negatively correlated with wind speed and relative humidity during winter.

The urban heat island (UHI) phenomenon serves as a trap for atmospheric pollutants, deteriorates the quality of life and has a socio-economic impact in the urbanized areas [9,10]. Important research has been accomplished over the last hundred years to quantify its impact on the urban climate [11,12]. Various heat island studies have been performed in Europe during the last 15 years [13]. Urban heat island and increased urban temperatures [14,15,16], exacerbate the cooling load of buildings, increase the peak electricity demand for cooling and decrease the efficiency of air conditioners [17].

Moreover, the urban agglomeration has a negative impact on the cooling effectiveness of natural and night ventilation [18]. [19], was the first to relate UHI intensity to meteorological elements such as cloudiness, wind speed, temperature, and absolute humidity using a multiple linear regression method. He showed that cloudiness and wind speed parameters are negatively correlated with the UHI intensity in Uppsala, Sweden, and that the total variance explained by the regression model is larger in the nighttime than in the daytime.

Consequently, the prediction of the urban heat island behavior has gained a significant attention. Although a number of modelling approaches for urban heat island do exist [20], the complexity of the phenomenon, the bulk of urban details required to attain an accurate urban model and the increased cost and computational time of analytical modelling approaches has led to the exploration of other prediction methods. Artificial neural networks (ANNs) have been used in a number of prediction studies that involves atmospheric time series data. [21], predicted daily maximum ozone levels in Texas metropolitan areas with a standard three-layer ANN model with nine inputs and four hidden nodes and found it to be superior to statistical methods

III. METHOD AND MATERIALS

Based on the previous work done by all the researchers, this research-based project will provide with ANN algorithm and time series analysis for the prediction of UHI.

The dataset is collected from different areas of Rajasthan and the dataset consist of four input parameters and an output or dependent parameter. Dataset we choose is different from the existing datasets as in the existing ones we have more parameters on the basis of which the calculation was performed whereas we took the 4 major parameters for our calculations. The input parameters are the maximum temperature, minimum temperature, wind velocity, relative humidity and the dependent that is the output parameter is the solar radiations. The dataset totally comprises of 6 different areas having all the input parameters in each of the workstation. The data consist of 12929 instances. The collected data is a day wise collection which starts from 01-01-1979 to 31-07-2014.

Inside the training and testing directory data belonging to each class is present. Training and testing images are present in separate directories inside the main directory. The research work used libraries part of Anaconda4 framework.

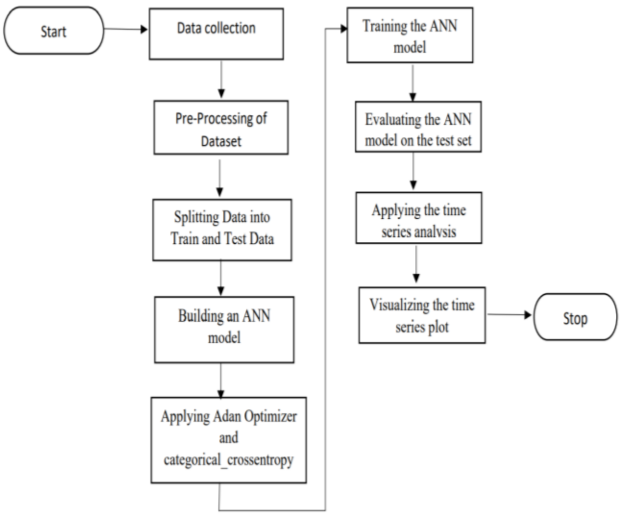

Figure 4: Flow chart of the model

GridSearchCV – This is a technique used for finding out the best values of hyper-parameters. The GridSearchCV has been used to find out the best number of hidden layers and their respective number of nodes.

K-Fold Cross Validation – This is a model evaluation technique. This consists of dividing the training dataset into "K" folds and then training the model on "K-1" folds and testing is done on the remaining one-fold. 10-Fold Cross Validation was used to find out the optimal values of hyper-parameters in each of the classification algorithms implemented.

Both of these techniques helped in getting a much higher accuracy over the test data.

IV. IMPLEMENTATION

This section describes the dataset used for the project, data preprocessing phase and a brief of all the classification algorithms used.

4.1 Data Preprocessing

This phase involved dealing with missing values, Dummy variables, standardization, etc. Work done on the above-mentioned dataset:

- Checked for any missing values.
- Dummy variables were created to deal with categorical values.
- StandardScaler object was applied on the dataset.

4.2 Build an ANN Model

In the project, the ANN model comprises of input layer, 3 hidden layer, and 1 output layer. The total quantity of nodes in the hidden layers was set to (40,40,30) respectively and epoch was set to 30. The activation function for the hidden layer was set as a rectifier function (relu) and sigmoid function for the output layer.

The ANN model was used for data classification using keras. Compiling the model is done by Adam optimizer and loss is "categorical_crossentropy.

The architecture of our model is:

```
model = Sequential ()
model.add(Dense(40, input_shape=(4,), activation = 'relu'))
model.add(Dense(40, activation='relu'))
model.add(Dense(30, activation='relu'))
model.add(Dense(1,))
model.compile(loss='mean_absolute_error',
optimizer='adam', metrics=['mean_absolute_error'])
history = model.fit(X_train, y_train, epochs = 20, verbose = 1)
```

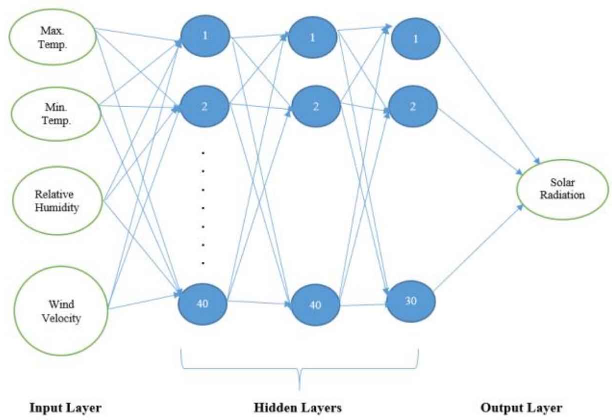

Figure 5: ANN model

The implementation steps are as follows:

a. Data collection and importing of dataset. The dataset totally comprises of 5 different parameters in each of the workstation. Each work station consists of 12999

ISSN No:-2456-2165

instances.

b. The input parameters are Maximum temperature, minimum temperature, wind velocity, relative humidity.
c. Output parameters is the solar radiations.
d. Pre-processing the dataset, dealing with missing values, Dummy variables, standardization, normalization, etc.
e. Forming training and test set, using the train_test_split() method to split training and testing data
f. Building an ANN model, for each Work station. We Build 6 different ANN models.
g. The ANN model has been implemented on the real time data of 6 WS which consists of the inputs (independent variables) as Maximum Temperature, Minimum Temperature, Relative Humidity and Wind velocity. The output (dependent variable) is **Solar Radiation.**
h. Compiling the model is done by Adam optimizer and loss is "categorical_crossentropy". The activation function for the hidden layer was set as a **rectifier function (relu)** and **sigmoid function** for the output layer.
i. Subsequent to building the model design, we train the model using model.fit(). The model performs better After 20 epochs.
j. GridSearchCV and K-Fold Cross Validation helped in getting a much higher accuracy over the test data.
k. After implementing the ANN model for all the 6 WS, it has been found that the **WS4 and WS6** shows the best correlation between the solar radiation and the input parameters based on the MAE and R2 score which are used as model evaluation tools.
l. Applying the time series analysis on WS6 and WS4 in order to find out among then which workstation is more likely to form an Urban heat island (UHI).
m. Many statistics techniques like Rolling statistics and Dickey Fuller test have been applied on WS 6 and WS4 on parameters wind velocity and solar radiations, in order to extract meaning full insights from the data.
n. Visualizing the time series plot. Autocorrelation & partial correlation graphs of solar radiation and Autocorrelation & partial correlation graphs of wind velocity for WS6 as well as WS4 is plotted.

V. RESULTS AND ANALYSIS

5.1 ANN Results:

The aim of the ANN model is to find the WS which shows the best correlation between the solar radiation and the input parameters. And resulted weather stations has been used for time series analysis to predict the UHI.

After implementing the ANN model for all the 6 WS, it has been found that the **WS4 and WS6 shows the best correlation** between the solar radiation and the input parameters based on the MAE and R2 score which are used as model evaluation tools.

Table 1. Model Evaluation Results

Weather Stations	MSE	Train Accuracy (R2score)	Test Accuracy (R2 Score)
1	1.8570	0.782	0.761
2	1.8380	0.797	0.773
3	1.8210	0.789	0.768
4	1.7687	0.772	**0.796**
5	1.7964	0.808	0.778
6	1.8120	0.808	**0.793**

Table 1 shows the values of MAE and R2 score calculated for each weather stations.

5.2 Effect of Solar Radiation and wind velocity on UHI

Many researches have shown that the direct solar radiation has a positive effect on UHI and the wind velocity has a negative outcome on the UHI. It means that if solar radiation increases then UHI may form on that region and similarly, if the wind velocity decreases then the UHI may not form on that region.

5.3. Time Series Analysis Results

The time series analysis plot of solar radiation v/s time and wind velocity v/s time has been plotted for WS4 and WS6. After visualizing the plots, it has been found that the plot corresponding to the WS6 i.e (Wind v/s time) shows that wind velocity decreases over the time and (Solar radiation v/s time) shows that solar radiation significantly increases over the time.

So, we can say that the region corresponding to the WS6 forms the UHI over the time.

5.5 Graphical Representation of Results A. Plot of WS 6

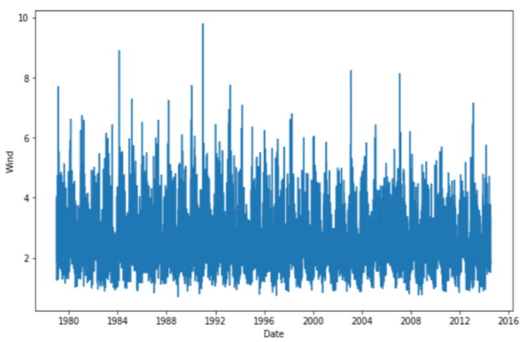

Figure 6: Graphical representation of wind velocity dataset.

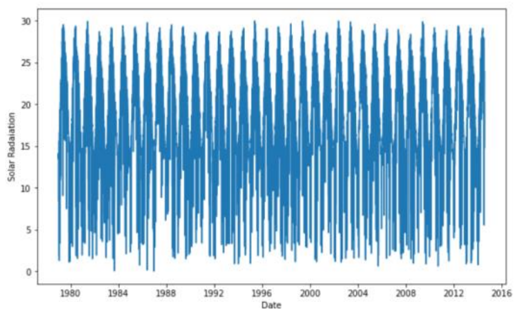

Figure 7: Graphical representation of solar radiation dataset.

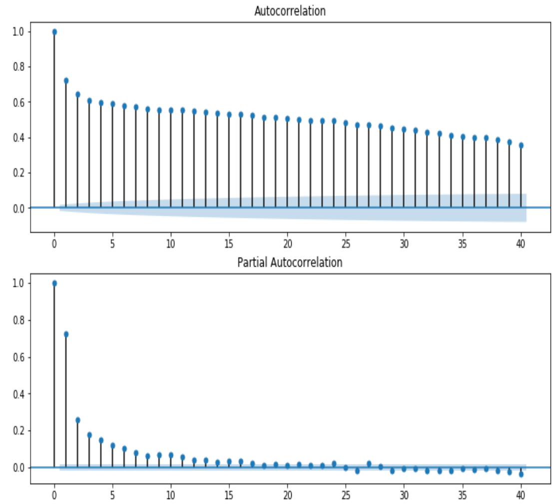

Figure 8: Autocorrelation and partial correlation graphs of solar radiation.

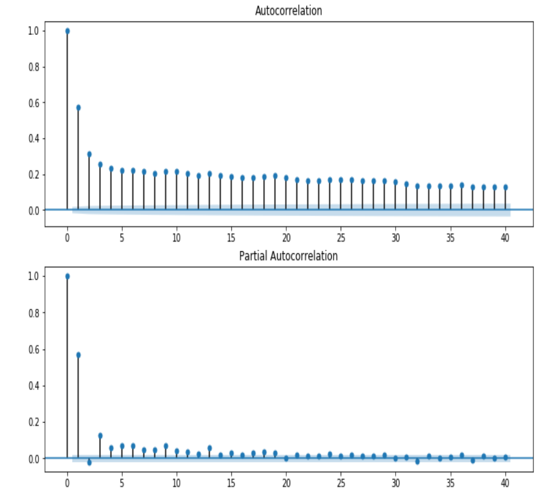

Figure 9: Autocorrelation and partial correlation graphs of wind velocity.

B. Plot of WS4

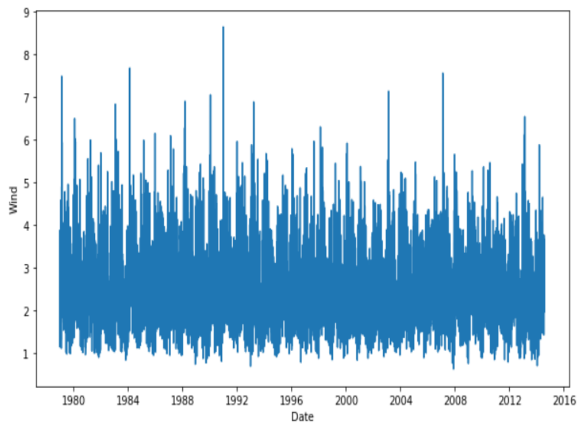

Figure 10: Graphical representation of wind velocity dataset.

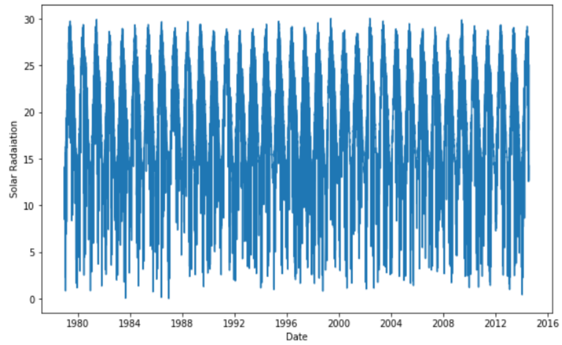

Figure 11: Graphical representation of solar radiation dataset.

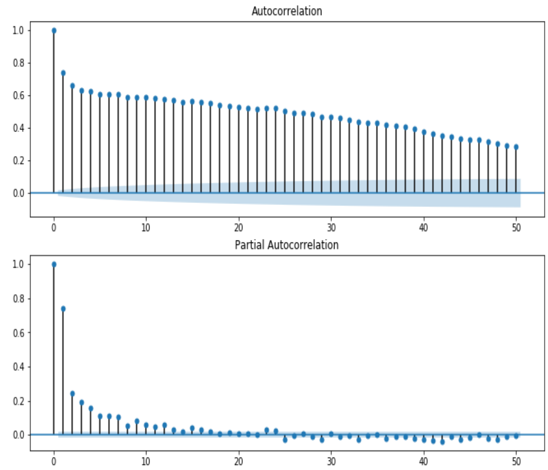

Figure 12: Autocorrelation and partial correlation graphs of solar radiation.

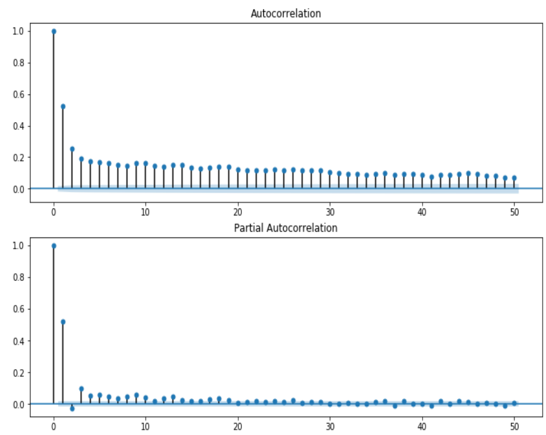

Figure 13: Autocorrelation and partial correlation graphs of wind velocity

As we can see the autocorrelation graphs for both the work stations (WS) 4 and 6, In work station 4 there is a decrease in solar radiation with increase in lags whereas in WS 6 solar radiation the decreasing factor is not to that extend. So, by this we can predict that WS6 has more chances of being an urban heat island.

On the other hand, when it comes to wind velocity, we can see in the auto correlation graphs that it is high WS 4 than in WS 6. Thus, by all the graphs we can conclude that WS 6 can be predicted to be an urban heat island.

VI. CONCLUSION

In conclusion, this research-based project dealt with ANN algorithm and time series analysis for the prediction of UHI and comparing with other traditional methods. K-Fold Cross Validation and GridSearchCV were also applied so as to increase the accuracies of the models.

After examining the ANN model for 6WS, the two weather stations WS4 and WS6 proved out to be the best in terms of correlation between dependent and independent variable that was evaluated using MAE and R2, R2 score or accuracy as 79.6 and 79.3 respectively was further chosen for time series analysis. It has been concluded from the time series analysis plot that WS6 forms the UHI in comparison to WS4.

During the development process, we studied carefully and understanding the criteria for making it more effective with a high accuracy percentage. We also realized the importance of maintaining a token margin for error.

There are many places we can improve. The following are the features that we are about to implement in near future.

- Forecast the solar radiation or the wind velocity for the coming years.
- A technique to forecast in how many years an area will become UHI.
- Improve the efficiency of this algorithm.

VII. ACKNOWLEDGMENTS

The authors would like to thank Google for quick and efficient results on selected keywords, Google Collaboratory. Sincere thanks to mentors for guiding well and helping from design and development of deep learning model.

REFERENCES

[1]. https://www.nationalgeographic.org/encyclopedia/urban-heat-island/
[2]. https://www.ibm.com/cloud/learn/machine-learning
[3]. https://www.investopedia.com/terms/d/deep-learning.asp
[4]. Sarkar H., (2004), Study of Landcover and Population Density Influences on Urban Heat Island in Tropical Cities by Using Remote Sensing and GIS: A Methodological Consideration, 3rd FIG Regional Conf., Jakarta, Indonesia.
[5]. Chudnovsky A., Ben-Dor E., Saaroni H., (2004), Diurnal thermal behavior of selected urban objects using remote sensing measurement, Energy and Buildings, 36, 1063-1074.
[6]. Kolokotsa D., Psomas A. Karapidakis E., (2009), Urban heat island in southern Europe: The case study of Hania, Crete, Solar Energy, 83, 1871-1883.
[7]. Liu W., Ji C., Zhong J., Jiang X., Zheng Z., (2007), Temporal characteristics of the Beijing urban heat
[8]. island, Theoretical and Applied Climatology, 87, 213-221.
[9]. Papanastasiou D.K., Kittas C., (2012), Maximum urban heat island intensity in a medium-sized coastal Mediterranean city, Theoretical and Applied Climatology, 107, 407-416.
[10]. Santamouris, M., Paraponiaris, K., & Mihalakakou, G. (2007). Estimating the ecological footprint of the heat island effect over Athens, Greece. Climate Change, 80, 265–276.
[11]. Santamouris, M., Pavlou, K., Synnefa, A., Niachou, K., & Kolokotsa, D. (2007). Recent progress on passive cooling techniques advanced technological developments to improve survivability levels in low-income households. Energy and Buildings, 39, 859–866
[12]. Akbari, H., Konopacki, S., & Pomerantz, M. (1999). Cooling energy savings potential of reflective roofs for residential and commercial buildings in the United States. Energy, 24, 391–407.
[13]. Mihalakakou, P., Flokas, H., Santamouris, M., & Helmis, C. (2000). Application of neural networks to the simulation of the heat island over Athens Greece using synoptic types as a predictor. Journal of Applied Meteorology, 41, 519–527.
[14]. Santamouris, M.(2007). Heatisland research in Europe: The state ofthe art. Advances in Building Energy Research, 1, 123–150.
[15]. Livada, I., Santamouris, M., & Assimakopoulos, M. N. (2007). On the variability of summer air temperature during the last 28 years inAthens. Journal ofGeophysical Research, 112, D12103.
[16]. Livada, I., Santamouris, M., Niachou, K., Papanikolaou, N., & Mihalakakou, G. (2002). Determination of places in the great Athens area where the heat island effect is observed. Theoretical and Applied Climatology, 71(3–4), 219–230.
[17]. Mihalakakou, M., Santamouris, N., & Papanikolaou, C. (2004). Cartalis simulation of the urban heat island phenomenon in Mediterranean climates. Pure and Applied Geophysics, 161, 429–451.
[18]. Cartalis, C., Synodinou,A., Proedrou, M., Tsangrassoulis,A., & Santamouris, M.(2001). Modifications in Energy Demand in urban areas as a result of climate changes: An assessment for the southeast Mediterranean region. Energy Conversion and Management, 42(14), 1647–1656.
[19]. Geros, V., Santamouris, M., Karatasou, S., Tsangrassoulis, A., & Papanikolaou, N. (2005). On the cooling potential of night ventilation techniques in the urban environment. Energy and Buildings, 37(3), 243–257.
[20]. Sundborg, A., 1950: Local climatological studies of the temperature conditions in an urban area. Tellus, 2, 222–232.

[21]. Mirzaei, P. A., & Haghighat, F. (2010). Approaches to study urban heat island—Abilities and limitations. Building and Environment, 45(10), 2192–2201.
[22]. Yi, J., & Prybutok, V. R. (1996). A neural network model forecasting for prediction of daily maximum ozone concentration in an industrialized urban area. Environmental Pollution, 92(3), 349–357.