Estimation of near-surface Air temperature during day and night-time from MODIS over Different LC/LU Using machine learning methods in Berlin

*A **thesis** presented for the degree of **Dr. rer. nat**.*

by Forough Marzban

To the Department of Earth Sciences of Freie Universität Berlin

Berlin, 2020

1. Supervisor and Reviewer

Prof. Dr. Sahar Sodoudi[1], Freie Universität Berlin, Germany

2. Supervisor and Reviewer

Prof. Dr. Tim Conrad[2], Freie Universität Berlin, Germany

Date of Defense:

17st June 2020

Version: 0.1

[1]Dept. of Earth Sciences, Urban Climate and Health Working Group
[2]Dept. of Mathematics and Computer Science, Medical Bioinformatics Group

Statutory Declaration

I declare that I have developed and written the enclosed PhD thesis completely by myself, and have not used sources or means without declaration in the text. Any thoughts from others or literal quotations are clearly marked.

The PhD thesis was not used in the same or in a similar version to achieve an academic grading or is being published elsewhere.

Berlin (Germany), 17st June 2020

_____

Forough Marzban

Abstract

Urbanization is manifest by changes in the physical structure of the land surface, owing to extensive construction features such as buildings, street canyons, changes in the thermal structure because of materials of different thermal properties and also intensive human activities. Urban areas are generally also characterized by higher surface air temperatures as compared to the rural surroundings. This temperature excess can be up to 10-12°C and more and is referred to as the urban heat island(UHI)phenomenon. Since residents living in cities are especially affected by extreme temperature events, urban climate studies are gaining in importance. Currently, more than half of the world s population already lives in urban areas, which accentuates the major role agglomerations must play in mitigation and adaptation to climate change. Recommendations regarding behavioural patterns during heat stress situations and urban planning measures require a comprehensive understanding of the inner urban temperature distribution including the identification of thermal hot spots. Both very cold and very hot temperatures could affect the human health. Excessive exposure to heat is referred to as heat stress and excessive exposure to cold is referred to as cold stress. Urban temperature data (2 m temperature data) is very important for all investigations on the urban heat island (UHI) effect, human health. They are usually either based on remote sensing techniques or air temperature measurements or from models. Remote sensing data like infra-red surface temperature from airborne measuring instruments may have a very high spatial resolution and are presently available for many urban areas, but only in clear sky cases. This spatial resolution is appropriate to exhibit typical urban structures that are expected to cause the UHI effect. Nevertheless, information on surface temperature cannot replace air temperature data, since beside the problem that the former is typically only available for single days, there is no fixed relation between surface and air temperatures. Especially for systematic analyses of the relationship between urban structures and 2m temperatures for different weather situations a large data basis is desirable. Air temperature data can be obtained from mobile measurements and measurement at permanent or temporary weather stations. On the one hand, the use of weather stations provides high data accuracy using a well-known standard technology. On the other hand, the spatial representation of weather station data within the urban environment, which is characterized by the surface composition including buildings, infrastructure and different types of land use, is very limited. Consequently, since the beginning of the 20th century, many efforts have been made to identify temperature patterns in urban areas with high spatial resolution instead of only using single point information. In this regard, in this study Air temperature (T_{2m} or T_{air}) measurements from 20 ground weather stations in Berlin were used to estimate the relationship between air temperature and the remotely sensed land surface temperature (LST) measured by Moderate Resolution Imaging Spectroradiometer over different land-cover types (LCT). Knowing this relationship enables a better understanding of the magnitude and pattern of Urban Heat Island (UHI), by considering the contribution of land cover in the formation of UHI. In order to understand the seasonal behaviour of this relationship, the influence of the normalized difference vegetation index (NDVI) as an indicator of degree of vegetation on LST over

different LCT was investigated. Next to it, to evaluate the influence of LCT, a regression analysis between LST and NDVI was made. The results demonstrate that the slope of regression depends on the LCT. It depicts a negative correlation between LST and NDVI over all LCTs. Our analysis indicates that the strength of correlations between LST and NDVI depends on the season, time of day, and land cover. This statistical analysis can also be used to assess the variation of the LST–T_{2m} relationship during day- and night-time over different land covers. The results show that LST_{Day} and LST_{Night} are correlated significantly ($p = 0.0001$) with T_{2mDay}(daytime air temperature) and $T_{2mNight}$(night-time air temperature). The correlation (r) between LST_{Day} and T_{Day} is higher in cold seasons than in warm seasons. Moreover, during cold seasons over every LCT, a higher correlation was observed during daytime than during night-time. In contrast, a reverse relationship was observed during warm seasons. It was found that in most cases, during daytime and in cold seasons, LST is lower than T_{2m}. In warm seasons, however, a reverse relationship was observed over all land-cover types. In every season, LST_{Night} was lower than or close to $T_{2mNight}$. Air temperature (T_{air} or T_{2m}) is an important climatological variable for forest biosphere processes and climate change research. Due to the low density and the uneven distribution of weather stations, traditional ground-based observations cannot accurately capture the spatial distribution of T_{air}. Therefore, it is necessary to develop a method for the estimation of air temperature with reasonable accuracy and spatial and temporal resolution in the urban areas with low temperature gauge density. But since the estimation of meteorological variables using various statistical techniques (such as linear regression models or combined regression and kriging techniques for T interpolation) have been examined by many researchers and they came to conclusion that an appropriate machine learning technique could be a robust computational technique which has been used for the estimation of meteorological data as a function of the corresponding data of one or more reference stations. In this research, T_{air} in Berlin is estimated during the day and night-time over six land cover/land use (LC/LU) types by satellite remote sensing data over a large domain and a relatively long period (7 years). Aqua and Terra MODIS (Moderate Resolution Imaging Spectroradiometer) data and meteorological data for the period from 2007 to 2013 were collected to estimate T_{air}. Twelve environmental variables (land surface temperature (LST), normalized difference vegetation index (NDVI), Julian day, latitude, longitude, Emissivity31, Emissivity32, altitude, albedo, wind speed, wind direction and air pressure) were selected as predictors. Moreover, a comparison between LST from MODIS Terra and Aqua with daytime and night-time air temperatures (T_{day}, T_{night}) was done respectively and in addition, the spatial variability of LST and T_{air} relationship by applying a varying window size on the MODIS LST grid was examined. An analysis of the relationship between the observed T_{air} and the spatially averaged remotely sensed LST, indicated that 3 × 3 and 1 × 1 pixel size was the optimal window size for the statistical model estimating T_{air} from MODIS data during the day and night time, respectively. Three supervised learning methods (Adaptive Neuro Fuzzy Inference system (ANFIS), Artificial Neural Network (ANN) and Support vector machine (SVR)) were used to estimate T_{air} during the day and nighttime, and their performances were validated by cross-validation for each LC/LU. by applying each technique, a estimator model of air temperature had been generated.

The comparison between these methods has been done and finally we evaluated the accuracy of each model and choose the best one for the high-resolution temperature estimation. Moreover, tuning the hyper parameters of some models like SVR and ANN were investigated. For tuning the hyper parameters of SVR, Simulated Annealing (SA) was applied (SA-SVR model) and a multiple-layer feed-forward (MLF) neural networks with three layers and variable nodes in hidden layers had been applied with Levenberg-Marquardt back-propagation (LM-BP), in order to achieve higher accuracy in the estimation of T_{air}. Results indicated that the ANN model achieved better accuracy (RMSE=2.16°C, MAE =1.69°C, R^2 =0.95) than SA-SVR model (RMSE= 2.50°C, MAE =1.92°C, R^2 =0.91) and ANFIS model (RMSE=2.88°C, MAE=2.2°C, R^2 =0.89) over six LC/LU during the day and night time. The Q-Q diagram of SA-SVR, ANFIS and ANN show that all three models slightly tended to underestimate and overestimate the extreme and low temperatures for all LC/LU classes during the day and night-time. The weak performance in the extreme and low temperatures are a consequence of the small numbers of data in these temperatures. These satisfactory results indicate that this approach is proper for estimating air temperature and spatial window size is an important factor that should be considered in the estimation of air temperature. Moreover, for better understanding the relationship between LST and T_{air} in Berlin during day and night-time, over six land LC/LU types namely airport, agriculture, urban area, forest, industrial and needle leaf trees, two input variable selection methods were applied. Input variable selection is an essential step in environmental, biological, industrial and climatological applications. One approach which help us in better understanding data, decreasing computation effort, the impact of curse of dimensionality and improving the estimator performance. Through input variable selection the irrelevant or redundant variables will be to eliminated therefore a suitable subset of variables is identified as the input of a model. Meanwhile, the complexity of the model structure is simplified, and the computational efficiency is improved. In this work, the two input variable selection methods, including brute force search and greedy best search algorithm using artificial neural network (ANN) were considered for estimating of near surface air temperature from MODIS over six LC/LU types. The motivation behind this research was to formulate a more efficient way of choosing input variables using ANN models of environmental processes. Moreover, AIC, BIC and RMSE are considered for ranking the features and finding a subset of potential variables which improves the overall estimation performance. In this study, Aqua and Terra MODIS data and meteorological data for the period from 2007 to 2013 were collected to estimate T_{air}. Moreover, twelve environmental variables LST, normalized difference vegetation index (NDVI), Julian day, latitude, longitude, Emis31, Emis32, altitude, albedo, wind speed, wind direction and air pressure were selected as predictors. The results show that the LC/LU has a key factor in the relationship between T_{air} and LST. The results show that the effectiveness of optimal models in estimation T_{air} is varied in different LC/LU because of the specific heat capacities of different LC/LU. Air temperature mainly rely on the heat transfer process which was significantly affected by the local radiation budget. Generally, air is heated much quicker over barren land than forest because, barren land has lower heat capacity than forest. Vegetation can cause to latent heat flux, such as enhancing or reducing transpiration and cool the T_{air} in forests. In this

study, the cooling effect was not take into account because of roughly distribution of meteorological stations across different vegetation types. Therefore, it was difficult to consider the vegetation type in our models. However, land cover also affected land surface albedo, thus, the influence of LU/LC on estimating T_{air} was conditional and time dependent because different variables are selected for the same LU/LC during day and night time. Moreover, another issue that we tried to find an answer was, what is the pitfall of using the global model and what is the advantage of features selection? It has been debated that inferencing from a model with all the features which thought to be important is simple and avoid the complications of model selection.

Zusammenfassung

Urbanisierung stellt eine Veränderungen in der physikalischen Struktur der Landoberfläche durch umfangreiche Konstruktionsmerkmale, wie Gebäude und Straßenschluchten, dar. Die damit verbundenen Änderungen der thermischen Struktur durch Verwendung von Materialien mit unterschiedlichen thermischen Eigenschaften sowie intensive menschliche Aktivität spielen hierbei eine wichtige Rolle. Urbane Gebiete sind im Allgemeinen durch eine höhere Oberflächentemperatur im Vergleich zur ländlichen Umgebung gekennzeichnet. Der Temperaturüberschuss kann bis zu 10-12°C und mehr betragen und wird als Phänomen der städtischen Wärmeinsel (UHI) bezeichnet. Da in Städten lebende Menschen besonders stark von extremen Temperaturereignissen betroffen sind, gewinnen Studien zum urbanen Klima vermehrt an Bedeutung. Derzeit lebt mehr als die Hälfte der Weltbevölkerung in urbanen Gebieten. Dies unterstreicht die wichtige Rolle die Ballungsräume in Bezug auf Minderung und Anpassung an den Klimawandel darstellt. Empfehlungen bezüglich des Verhaltens während Hitzestresssituationen sowie städtebauliche Maßnahmen erfordern ein umfangreiches Verständnis der innerstädtischen Temperaturverteilung einschließlich der Identifizierung von thermischen Hotspots. Sehr kalte wie auch sehr heiße Temperaturen können gleichermaßen die menschliche Gesundheit beeinträchtigen. Übermäßige Hitzebelastung wird als Hitzestress bezeichnet, übermäßige Kältebelastung als Kältestress. Urbane Temperaturdaten (2m Temperaturdaten) sind wichtig für alle Untersuchungen bezüglich des urbanen Wärmeinseleffekts (UHI), der menschlichen Gesundheit. Normalerweise basieren die Daten entweder auf Fernerkundungstechniken oder auf Messungen oder Simulationen der Lufttemperatur. Fernerkundungsdaten, wie die der Infrarot-Oberflächentemperatur von satellitengestützen Messinstrumenten können eine sehr hohe räumliche Auflösung haben und sind gegenwärtig für viele urbane Gebiete zugänglich, doch nur im Fall von wolkenfreiem Himmel. Die räumliche Auflösung ist dafür geeignet typische urbane Strukturen zu erkennen, die den UHI-Effekt auslösen. Dennoch können Informationen der Oberflächentemperatur, die Lufttemperaturdaten nicht ersetzen, da neben dem Problem, dass die Oberflächentemperatur in der Regel nur für einzelne Tage zur Verfügung steht, es keinen festen Zusammenhang zwischen Oberflächentemperatur und Lufttemperatur besteht. Besonders für systematische Analysen der Zusammenhänge zwischen urbanen Strukturen und der 2m-Temperatur unterschiedlicher Wettersituationen ist eine hohe Datenbasis wünschenswert. Daten der Lufttemperatur können von mobilen Messungen und permanenten Messstationen oder von temporären Wetterstationen erhalten werden. Einerseits bieten die Wetterstationen eine hohe Datenqualität durch Verwendung von bekannten Standard-Technologien; andererseits ist die räumliche Verteilung der Wetterstationsdaten in der urbanen Umgebung, die durch die oberflächliche Komposition von Gebäuden, Infrastruktur und verschiedenen Landnutzungsklassen charakterisiert ist, sehr eingeschränkt. Seit Beginn des 20. Jahrhunderts konnten somit viele Vorzüge bei der Identifizierung von Temperaturmustern in urbanen Gebieten mit hoher räumlicher Auflösung erzielt werden anstatt nur einzelne Punktinformationen zu nutzen. Folglich werden für diese Studie Lufttemperatur (T_{2m}) oder T_{air} Messungen von 20 Bodenwetterstationen in Berlin verwendet, um den Zusammenhang zwischen Luft-

temperatur und Fernerkundungsdaten der Oberflächentemperatur (LST) gemessen vom Moderate Resolution Imaging Spectroradiometer (MODIS) über verschiedene Landnutzungstypen (LCT). Die Kenntnis über diesen Zusammenhang ermöglicht ein besseres Verständnis der Stärke und Muster von urbanen Wärmeinseln (UHI) durch Beachtung der Verteilung der Oberflächenbeschaffenheit bei Ausbildung von UHI. Um das saisonale Verhalten dieses Zusammenhangs zu verstehen, wurde der Einfluss des normalisierten Differenzvegetationsindex (NDVI) als ein Indikator für den Vegetationsgrad auf LST über verschiedene LCT untersucht. Darüber hinaus wurde eine Regressionsanalyse zwischen der LST und dem NDVI durchgeführt, um den Einfluss der LCT zu bewerten. Die Ergebnisse zeigen, dass die Steigung der Regressionsgeraden von der LST abhängt. Es besteht eine negative Korrelation zwischen LST und NDVI über alle LCTs. Unsere Analyse signalisiert, dass die Stärke der Korrelation zwischen LST und NDVI von der Jahreszeit, der Tageszeit sowie der Landnutzung abhängig ist. Die statistische Analyse kann auch verwendet werden, um die Variation der LST-T_{2m} Beziegung während der Tages- und Nachtzeit über verschiedene Bodenbedeckungen zu bewerten. Die Ergebnisse zeigen eine signifikante Korrelation ($p=0.0001$) von LST_{day} und LST_{night} mit der T_{2mDay} (Lufttemperatur tagsüber) und der $T_{2mNight}$(Lufttemperatur nachts). Zwischen LST_{day} und T_{day} ist die Korrelation (r) in der kalten Jahreszeit höher als in der warmen. Darüber hinaus wurde eine höhere Korrelation während der kalten Jahreszeit über alle LCTs am Tag beobachtet als in der Nacht. In der warmen Jahreszeit wurde im Gegensatz dazu ein umgekehrter Zusammenhang festgestellt. Es wurde beobachtet, dass in den meisten Fällen, tagsüber und in kalten Jahreszeiten, die LST niedriger ist als die T_{2m}. In warmen Jahreszeiten wurde jedoch ein umgekehrter Zusammenhang über alle Landbedeckungsarten beobachtet. In jeder Saison war die LST_{Night} niedriger oder fast gleich wie die $T_{2mNight}$. Die Lufttemperatur (T_{air} oder T_{2m})ist eine wichtige klimatologischen Variable für Prozesse der Waldbiosphäre und die Erforschung des Klimawandels. Aufgrund der geringen Dichte und der ungleichmäßigen Verteilung von Wetterstationen können herkömmliche bodengebundene Beobachtungen die räumliche Verteilung von T_{air} nicht genau erfassen. Daher ist es notwendig, eine Methode zur Abschätzung der Lufttemperatur mit angemessener Genauigkeit sowie räumlicher und zeitlicher Auflösung in urbanen Gebieten mit niedriger Temperaturmessdichte zu entwickeln. Da aber die Abschätzung meteorologischer Variablen mit verschiedenen statistischen Techniken (wie linearen Regressionsmodellen und kombinierten Regressions- und Krigingtechniken für die T-Interpolation) von vielen Forschern untersucht wurde, kamen sie zu dem Schluss, dass eine geeignete machine learning Technik eine robuste Rechentechnik sein könnte, die für die Abschätzung meteorologischer Daten in Abhängigkeit von entsprechenden Daten einer oder mehrerer Referenzstationen verwendete. In dieser Studie wird T_{air} in Berlin tagsüber sowie nachts über sechs Landbedeckungs/Landnutzungsarten (LC/LU) mittels Satelliten-Fernerkundungsdaten über einen großen Bereich und einem relativ langen Zeitraum (7Jahre) geschätzt. Daten des,Terra und Aqua MODIS (Moderate Resolution Imaging Spectroradiometer) und meteorologische Daten für den Zeitraum von 2007 bis 2013 wurden gesammelt, um T_{air} zu bestimmen. Als Prädikatoren wurden zwölf Umweltvariablen (Landoberflächentemperatur (LST), normalisierter Differenzvegetationsindex (NDVI), Julianischer Tag, Breitengrad, Längengrad, Emissi-

onsgrad 31, Emissionsgrad 32, Höhe, Albedo, Windgeschwindigkeit, Windrichtung und Luftdruck) ausgewählt. Drüber hinaus wurde ein Vergleich zwischen LST von MODIS Terra und Aqua mit Tages- und Nachtlufttemperaturen (T_{Day}, T_{Night}) durchgeführt bzw. zusätzlich die räumliche Variabilität des Zusammenhangs von LST und T_{air} durch Anwendung einer variierenden Fenstergröße auf das MODIS LST-Gitter untersucht. Eine Analyse der Beziehung zwischen der beobachteten T_{air} und dem räumlich gemittelten Fernerkundungs-LST ergab, dass die Größe 3 x 3 und 1 x 1 Pixel die optimale Fenstergröße für das statistische Modell war, das T_{air} aus den MODIS-Daten während Tages- bzw. Nachtzeit schätzte. Drei überwachte Lernmethoden (Adaptive Neuro Fuzzy Inference system (ANFIS),künstliches neuronales Netzwerk (ANN) und Support vector machine (SVR)) wurden verwendet, um T_{air} währen des Tages und der Nacht zu schätzen. Die Leistungen wurden durch Kreuzvalidierung für jede LC/LU validiert. Durch die Anwendung jeder Technik wurde ein Schätzmodell ausgewertet und das Beste für die hochauflösende Temperaturschätzung ausgewählt. Darüber hinaus wurde die Einstellung der Hyperparameter einiger Modelle wie SVR und ANN untersucht. Für die Einstellung der Hyperparameter von SVR wurde ‚Simulated Annealing (SA) angewendet (SA-SVR Modell). Mit der Levenberg-Marquardt Backpropagation (LM-BP) wurde ein mehrschichtiges Feed-forward(MLF) neuronales Netzwerk mit drei Schichten und variablen Knoten in versteckten Schichten angewendet, um eine höhere Genauigkeit bei der Schätzung von T_{air} zu erreichen. Die Ergebnisse zeigten, dass das ANN-Modell über sechs LC/LU, tags sowie nachts, eine höhere Genauigkeit erreichte (RMSE=2.16°C, MAE =1.69°C, R^2 =0.95) als das SA-SVR-Modell (RMSE= 2.50°C, MAE =1.92°C, R^2=0.91) und das ANFIS-Modell (RMSE=2.88°C, MAE=2.2°C, R^2=0.89). Das Q-Q-Diagramm von SA-SVR, ANFIS und ANN zeigt, dass alle drei Modelle die extrem hohen und niedrigen Temperaturen für alle LC/LU-Klassen tagsüber sowie nachts leicht unterschätzen und überschätzen. Die schwache Leistung bei extrem hohen und niedrigen Temperaturen ist eine Folge der geringen Datenmenge bei diesen Temperaturen. Um den Zusammenhang zwischen LST und T_{air} in Berlin bei Tag und Nacht besser verstehen zu können, wurden über sechs Land-LC/LU-Typen (Flughafen, Landwirtschaft, urbanes Gebiet, Wald, Industrie und Nadelblattbäume) zwei Auswahlmethoden für die Eingangsvariablen angewendet. Die Auswahl dieser Variablen ist ein wesentlicher Schritt in ökologischen, biologischen, industriellen und klimatologischen Anwendungen. Ein Ansatz, der uns hilft, Daten besser zu verstehen, den Rechenaufwand zu verringern, die Auswirkungen des Fluches der Dimensionalität und die Leistungsfähigkeit des Schätzers zu verbessern. Durch die Auswahl der Eingabevariablen werden die irrelevanten oder redundanten Variablen eliminiert, so dass eine geeignete Teilmenge von Variablen als Input eines Modells identifiziert wird. Zur gleichen Zeit wird die Komplexität der Modellstruktur vereinfacht und die Recheneffizienz verbessert. In dieser Studie wurden Aqua und Terra MODIS-Daten und meteorologische Daten für den Zeitraum von 2007 bis 2013 gesammelt, um T_{air} zu schätzen. Darüber hinaus wurden zwölf Umweltvariablen (LST, normalisierter Differenzvegetationsindex (NDVI), Julianischer Tag, Breitengrad, Längengrad, Emis31, Emis32, Höhe, Albedo, Windgeschwindigkeit, Windrichtung und Luftdruck) als Prädikatoren ausgewählt. Die Ergebnisse zeigen, dass die LC/LU einen Schlüsselfaktor in der Beziehung zwischen T_{air} und LST haben. Die Ergebnisse zeigen, dass die Effektivität

optimaler Modelle bei der Schätzung von Tair aufgrund der spezifischen Wärmekapizität verschiedener LC/LU in unterschiedlichen LC/LU variiert. Die Lufttemperatur hängt vor allem vom Wärmeübertragungsprozess ab, der maßgeblich vom lokalen Strahlungsbudget beeinflusst wurde. Im Allgemeinen wird die Luft viel schneller über unfruchtbarem Land erhitzt als über Wald, weil unfruchtbares Land eine geringere Wärmekapazität besitzt als Waldgebiete. Vegetation kann zu latenten Wärmeflüssen führen, z.B. zur Verbesserung oder Verringerung der Transpiration und zur Abkühlung der T_{air} in Wäldern. Wegen der groben Verteilung meteorologischer Stationen auf verschiedene Vegetationstypen wurde der Kühleffekt in dieser Studie nicht berücksichtigt. Aufgrund der geringen Datengrundlage war es schwierig, den Vegetationstyp in unseren Modellen zu berücksichtigen. Die Bodenbedeckung beeinflusste jedoch auch die Bodenoberflächenalbedo, so dass der Einfluss von LC/LU auf die Schätzung von T_{air} bedingt und zeitabhängig war, da für dieselbe LC/LU tagsüber und nachts unterschiedliche Variablen ausgewählt wurden. Darüber hinaus haben wir versucht, eine Antwort zu finden: Was sind die Fallstricke bei der Verwendung des globalen Modells und was ist der Vorteil der Auswahl von Merkmalen? Es wurde diskutiert, dass die Ableitung von einem Modell mit allen als wichtig erachteten Merkmalen einfach ist und die Komplikationen der Modellauswahl vermeidet.

Table of Contents

List of Tables

List of Figures

1 Introduction

The fifth assessment report of the Intergovernmental Panel on Climate Change (IPCC, the most comprehensive report on climate appraisal, referred to city area as the main region that must respond to climate change risks. Climate change research on urban scales has become an important topic. The international literature demonstrated the importance of cities in coping with climate change (Betsill et al., 2011, Lindseth et al., 2014, Grimm et al., 2018). A series of issues has been upraised by the rapid development of urbanization, such as the large numbers of greenhouse gas emissions of urban systems, the sprawling layout of urban spaces, and the disordered use of land, leading some to take climate issues more seriously. Therefore, urban systems need to respond to the major challenges of climate change.
As an introduction to urban heat islands, this chapter reviews some of the characteristics of urban climates and the important factors which leads to urban heat islands (UHI)and also its consequence. It is clear that, urban climates vary from those of rural surrounding, and the magnitudes of the differences can be totally large at times, depending on anthropogenic moisture, weather conditions, urban thermophysical, geometrical characteristics, and heat sources present in the area. In the research of Changnon(1976,1981), for northern hemisphere, was mentioned that, urban areas annually have an average of 12% less solar radiation, 10% more snowfall, 14% more rainfall, 8% more clouds, and 15% more thunderstorms than their rural area(Changnon 1976,1981). Generally, urban air temperatures are higher than their corresponding rural area in most mid-and high latitude cities. This phenomenon is called urban heat island which has been well documented (Chandler1960, Oke 1987, Oke 1988, Karl et al., 1989). A heat island can occur at different scales; it can present itself around a single building (Thurow 1989), a large portion of a city, or a small vegetative canopy (Taha et al.,1989, Taha et al.,1991).
This chapter provides an overview of how urban heat islands forms, factors that contribute to their development, study area, data description, and general description of what have been done in each chapter.

1.1 Urban Heat Islands, Climate Change, and Global Warming

Urban heat islands refer to the difference in temperatures measured inside and outside the city (Oke TR. 1973, Oke TR. 1995, Rizwan et al., 2008, Oke 1982, Huang et al., 2005, Arnfield 2003). Urban heat islands will be formed by development and the changes in irradiative and thermal properties of urban materials, specially, thermal emissivity, solar reflectance, and heat capacity, as they determine how the sun's energy is emitted, reflected and absorbed. Surface materials such as roofing, and paving are another important factor, which must consider in urban areas as compared

with rural area, because it has a lower albedo than those in rural settings. Therefore, built up area generally absorb more and reflect less of the sun's energy. This absorbed heat rises up surface temperatures and leads to the formation of surface and atmospheric urban heat islands. Although solar reflectance is the main determinant of a material's surface temperature, thermal emittance, or emissivity, also plays a role. Thermal emittance is a measure of a surface's ability to shed heat or emit long-wave(infra-red) radiation. All things equal, surfaces with high emittance values will stay cooler, because they will release heat more easily. As result, Albedo and emissivity are considered as radiative properties in our model estimation. Another key property in the formation of urban heat island is urban geometry specially at night. Urban geometry refers to the dimensions and spacing of buildings which influences energy absorption, wind flow, and a given surface's ability to emit long-wave radiation back to space.
In Down-town metropolitan, surfaces and structures are often obstructed by neighbouring tall buildings, therefore, large thermal masses cannot release their heat very easily because of these obstructions. The air above urban area is typically warmer than air over rural surrounding, especially at night. Night-time atmospheric heat islands can have serious health impact for urban residents during heat waves.
Reduced vegetation in urban areas, is another important property that influences heat island development. Generally, rural area is characterized by more vegetation, trees and open land which trees and vegetation provide shade, which leads to lower surface temperatures. Moreover, through evapotranspiration process, they also reduce air temperatures. In contrary, urban areas are dominated by dry, impervious surfaces, side-walks, and roads. Therefore, in cities, we can see less vegetation, and more surfaces which are paved or covered with buildings. The change in land cover results in less shade and moisture to keep urban areas cool and next to it, built up areas evaporate less water, which leads to increase air and surface temperatures. Furthermore, there are other additional factors which contribute to urban heat island formation such as geographic location, certain weather condition, time of day/season, and anthropogenic heat emissions. In addition, the UHI effect climate change. In recent years, global temperature has elevated significantly and probably, the average annual temperature will increase by several degrees during this century. The warming that consequence of UHI over small region such as cities is an example of local climate change. Basically, local climate changes which is consequence of UHI, differ from global climate changes, because, their impact are restricted to the local scale and decrease with distance from their source while climate change points out to any significant change in measures of climate such as wind, temperature, or precipitation, which lasting for an long period such as decades or longer and their impact are not locally or regionally confined. Climate change may outcome of:

- Natural drivers, such as slow changes in the Earth's orbit around the sun and changes in the sun's intensity.
- Natural processes within the climate system.
- Human activities that change the atmosphere's composition (e.g. burning fossil fuels) and the land surface(anthropogenic heat, deforestation, or urbanization).

Global warming is an average temperature increase of the earth's climate system, which can lead to changes in global climate patterns. Global warming may result from natural and human activities. Global warming can be considered as a part of global climate change along with changes in precipitation, sea level, etc. The UHI and global climate change (or global warming) often has similar impacts. For example, some communities may experience longer growing seasons due to either or both phenomena. UHI and global climate change can also increase greenhouse gas emissions and energy demand, etc. Generally, UHI contribute to climate warming, and elevates the intensity of high heat waves (MCCarthy et al.2010, Huang and Lu 2015). Urban warming, is also referred to as UHI, that is people who live in the urban area are exposed to more heat stress than those living in rural surrounding (Zhou and Shepherd 2010) which means that UHI has a great direct impact on human health (Kovats and Hajat 2008). Heat islands have been important in some heat wave events (Watkins et al, 2002) because extreme temperature leads to the clinical syndromes of heat exhaustion, heat stroke, and heat cramps (Kilbourne EM.1997). Many studies reveal that mortality is more sensitive to heat in urban areas compared with rural areas. The prevention of deaths caused by extreme high temperatures is now an issue of public health concern. Moreover, UHI cause serious environment (Sarrat et al., 2006; Roth 2007; Grimm et al., 2008), and energy problems (Kolokotroni et al., 2012). It is obvious that increased temperatures in urban area can have a direct impact on the energy required to heat and cool buildings (Crawley et al.,2008; Lu et al., 2010; Crawley et al., 2008).

1.2 The influence of land-cover type on the relationship between NDVI–LST and LST-T_{air}

Many studies have revealed a strong negative correlation between normalized difference vegetation index (NDVI) and land surface temperature (LST) (Goward, Cruickshanks, and Hope 1985; Hope et al. 1986; Smith and Choudhury 1991; Schultz and Halpert 1995;Churkina and Running 1998; Nemani et al. 2003; Julien, Sobrino, and Verhoef 2006; Sun and Kafatos 2007; Julien and Sobrino 2009; Kumar and Shekhar 2015; Tayyebi and Jenerette 2016; Zhou et al. 2014). Previous studies have determined the variability in the slope of the inverse LST–NDVI relationship, in association with local topographic and environmental conditions. Goward and Hope (1989) stated that the LST–NDVI slope and intercept are expected to vary from one day to another day based on the magnitude of incident solar radiation, advective atmospheric conditions, and surface moisture availability. Hope (1988) revealed that soil moisture potential affects the relationship between canopy temperature and NDVI. Yue et al. (2007) have investigated the mean LST and NDVI values associated with different land-use (LU) types in the city of Shanghai, China, and they found out that mean LST and NDVI values associated with different LU types are significantly different. Joshi and Bhatt (2012) stated that the areas with vegetation and water-body have lower temperature compared to the built-up areas. Sun and Kafatos (2007) found that the

correlation between LST and NDVI is positive for winter and negative during warm seasons. Knowing this relationship enables a better understanding of the magnitude and pattern of UHI, by considering the contribution of LC/LU type in the formation of UHI. In order to understand the seasonal behaviour of this relationship, the influence of the normalized difference vegetation index (NDVI) as an indicator of degree of vegetation on LST over different LCT was investigated. Moreover, for investigating the seasonal variation of the LST and NDVI relationship, a comprehensive comparison between mean LST and NDVI, based on different LU/LC types were compared and also a correlation analysis is carried out. In addition, in order to evaluate the influence of LCT, a regression analysis between LST and NDVI was made. More details on this topic was explained in the manuscript no.1 entitled "The influence of land-cover type on the relationship between NDVI-LST and LST-T_{air}" . The following major questions will be investigated within this study?

Question 1 How the relationship between T_{air} and Land surface temperature(LST) change seasonally during day and night time over six land cover types(LCT)if the land cover type can affect the relationship between land surface temperature and air temperature.

Question 2 How the relationship between T_{air} and the normalized difference vegetation index (NDVI) seasonally change during day and night time over six LCT if NDVI has different effects on nocturnal and daily air temperatures and how is this effect in different land cover types.

1.3 Estimation of the Near-surface Air Temperature during the Day and Night-time from MODIS in Berlin, Germany

Air temperature (T_{air} or T_{2m}) is an important climatological variable for forest biosphere processes and climate change research. Due to the low density and the uneven distribution of weather stations, traditional ground-based observations cannot accurately capture the spatial distribution of T_{air}. An accurate estimation of T_{air} and the mapping of its spatial distribution are useful for predicting ecological consequences of climate change. For example, climate warming will lead to higher temperatures and an increase of extreme weather conditions, which are associated with changes in wildfire regime (Westerling et al., 2006; Chen et al., 2011; Manzo-Delgado et al., 2009), forest biomass distribution (Reich et al., 2014) and crop yield (Ruane et al., 2014; Rosenzweig et al., 2014). The demand for accurate spatial T_{air} data over a large scale has continued to rise (Oyler et al., 2015; Beier et al., 2012). However, the spatial distribution of the weather stations in many parts of the world, is often limited which restricts the use of T_{air} measurements over a large spatial domain (Vancutsem et al., 2010). LST, but on the other hand, is measured in a global extent with significant higher spatial coverage (Jin and Dickinson, 2010). The US National Research Council

and the Intergovernmental Panel on Climate Change (IPCC) expressed the need for long-term remotely sensed LST data in global warming studies to overcome the limits of conventional surface T_{air} measurements (IPCC 2007, Jin 2004). Remote sensing data has great potential to estimate spatial-temporal patterns of T_{air} which can further our knowledge, on both the climate and terrestrial biological processes at regional and global scales (Benaliet al., 2012). Monitoring and understanding the trends of T_{air} and LST are crucial in the study of regional and global climate changes (Yoo et al., 2011). LST can be monitored and modelled from multiple daily satellite observations, such as the MODIS LST. Studies have shown that LST can be used for linear regression estimates of daily minimum and maximum T_{air} on a local scale (Mostovoy et al., 2006; Vancutsem et al., 2010; Zhang et al., 2011a; Yoo et al., 2011; Evrendilek et al., 2012; Benali et al., 2012; Zhu et al., 2013). Cresswell et al. (1999) found an over and underestimation of T_{air} during the day and at night, respectively, from Meteosat LST observations. They attempted to correct these errors and produce a proxy of T_{air} by applying a solar zenith angle correction on the Meteosat geostationary observations. They achieved an accuracy of 3°C for over 70% of the Meteosat temperatures. Similarly, Jin and Dickinson (2010) have studied the differences in the diurnal cycles of LST and T_{air} over a single site. Some studies (Florio et al., 2004) have used several statistical approaches that combined a simple AVHRR Spilt-Window Technique (SWT) with ground meteorological station measurements in the prediction of T_{air}. Other studies (Wloczyk et al., 2011) have used the Landsat LST data to derive T_{air}. They have attempted to assign the satellite-derived T_{air} to a certain height above the ground and have investigated the possibility of a simple correction for reference height. They also considered the link between T_{air} spatial pattern and the window-size of the Landsat LST pixels. Xu et al. (2012) used four empirical regression models to estimate the relationship between T_{air} measurements and the MODIS-Aqua LST and found different relationships between the two different LC types in their study. They also assessed the effect of the MODIS LST window-size on the agreement between the two variables and found that spatial averaging over multiple pixels improves the accuracy of T_{air} estimates. Zaksek and Schroedter-Homscheidt (2009) reviewed the types of methods commonly used to estimate T_{air} based on LST, dividing them into three distinct groups:

- Statistical approaches which are based on regression techniques, can be simple if only based on LST and T_{air}(e.g. Mostovoy et al., 2006; Vogt et al., 1997) or advanced, when more than one independent variable is used such as solar zenith angle (SZA), elevation, altitude, Julian day among others (Lin et al., 2012; Cresswell et al. 1999; Jang et al. 2004). Lin et al. (2012) used stepwise linear regression method to estimate daily maximum air temperature (T_{max}) and daily minimum air temperature (T_{min}) with MAE = 1.9, agreement index = 0.79 and MAE = 1.9°C, agreement index = 0.92, respectively, over east Africa. Fu et al. (2011) used linear regression between MODIS LST and T_{max} from stations on the northern Tibetan Plateau. In general, these methods perform well within the spatial and time frame they were developed, but the accuracy might decrease when extended in time and space (Stisen et al., 2007). Statistical methods generally perform well within the spatial and time frame they were derived in, but have limited generalization and

require large amounts of data to train the algorithms (Stisen et al., 2007).

- The second category is index-based such as Temperature Vegetation index (TVX). It is based on the assumption that for an infinitely thick canopy, the top-of-canopy temperature is the same as within the canopy (Czajkowski et al., 2000; Prihodko and Goward (1997), Nemani and Running et al., 1989; Nieto et al., 2011) and uses the Normalized Difference Vegetation Index (NDVI) as a key input variable. However, the assumption of linear and negative slope between LST and NDVI is not always applicable and is influenced by the seasons, the type of ecosystem and soil moisture variability (Sandholt et al., 2002; Vancutsem et al., 2010). Zhu et al. (2013) used the TVX method to estimate daily T_{max} with RMSE (the root mean square error) =3.709°C, MAE (the mean absolute error) = 3.03°C and r (correlation coefficient) = 0.83 in Xiangride River Basin of China. However, Vancutsem et al. (2010) found that TVX method did not adapt to different ecosystems over Africa because non-significant relationship between LST and NDVI in their study. Karnieli et al. (2003) found that the approaches based on this negative NDVI/LST relationship have minimal utility in energy-limited environments (e.g., high latitude and elevations) compared to moisture-limited environments because vegetation-expressed NDVI response is more related to available solar radiation than land surface conditions (e.g., soil moisture).
- The final approach uses surface energy balance parametrizations based on physically-based models (Sun et al., 2005). The sum of incoming net radiation is considered equal to the sum of the soil heat flux, sensible flux and latent heat flux (Zaksek and Schroedter-Homscheidt, 2009; Meteotest 2010; Sun et al., 2005). However these methods require large amounts of information that are usually not only from remote sensing (e.g., roughness, soil physical properties) (Benali et al., 2012, Mostovoy et al., 2006, Prince et al., 1998).

Most of the previous studies have focused on daily estimations or instantaneous T_{air}. The TVX method has been widely used for T_{air} estimation. Czajkowski et al. (2000) estimated Tavg for a weekly period with associated RMSE between 1.72 and 3.48°C and R^2=0.64. Stisen et al. (2007) and Prihodko and Goward (1997) estimated T_{air} with RMSE higher than 2.5°C and R^2 between 0.64 and 0.86. Cresswell et al. (1999) used a statistical method to derive instantaneous T_{air} with an associated RMSE below 3°C for more than 70% of the sampled data. Zaksek and Schroedter-Homscheidt (2009) used a more sophisticated method, which was based on the energy balance to estimate instantaneous T_{air} with an RMSE of 2°C. Vancutsem et al. (2010) used 1 km MODIS data to estimate weekly T_{min} and T_{max}. They reported correlations between LST and T_{min} ranging from 0.01 to 0.96 for several stations and T_{max} was estimated with R^2=0.92 and RMSE=1.83°C. Moreover, in previous studies, several variables were employed to estimate air temperature. For example, the variables used by Benali et al. (2012) included LST, Julian Day, elevation, and the distance to coast. Benali et al. (2012) used both weekly daytime LST data (LST_{day}) and night-time LST data (LST_{night}) to estimate the average, maximum and minimum weekly temperature. They found that there was a higher correlation between average weekly temperature and averaged weekly LST_{night}, which indicates the potential of LST_{night}in

estimating averaged weekly temperature. The variables used by Kim and Han (2013) included LST, NDVI, altitude, and solar zenith angle. The variables used by Cristóbal, Ninyerola and Pons (2008) included LST, NDVI, and albedo. The variables used by Zakšek and Schroedter-Homscheidt (2009) included LST, NDVI, solar zenith, albedo, solar radiation, and altitude. After comprehensive consideration of these variables, twelve variables were selected as the predictors for the modelling of air temperature during the day and night-time: LST, NDVI, Julian day, latitude, longitude, Emissivity31, Emissivity32, altitude, albedo, wind speed, wind direction and air pressure. The reasons for selecting these parameters are described in manuscript no.2 entitled "Estimation of the Near-Surface Air Temperature during the Day and Night-time from MODIS in Berlin, Germany". In this study, T_{air} in Berlin is estimated during the day and night time over six LC/LU types by satellite remote sensing data over a large domain and a relatively long period (7 years). Aqua and Terra MODIS (Moderate Resolution Imaging Spectro-radiometer) data and meteorological data for the period from 2007 to 2013 were collected to estimate T_{air}. First, this chapter presents the comparison of state-of-the-art remote sensing-based LST data from MODIS with T_{air} for the six LC/LU. Within this study, we compared the relationship between T_{air} and the Four LST products of MODIS over Berlin because of influence of the time of observation on the estimation of T_{air} which has been studied and discussed in several studies, which resulted in different conclusions. Benali et al.(2012) stated that the use of both aqua LST_{day} and LST_{night} could improve the estimation of T_{day} and T_{night}, respectively. In the other word, the aim is to analyse the agreement between LST from MODIS Terra and Aqua and $_{air}$ for the period of 2007 to 2013 based on different LC/LU, and then to investigate the spatial variability of LST and T_{air} relationship by applying a varying window size on the MODIS LST grid based on different LC/LU. The comparison is done by using statistical parameters such as the correlation coefficient, the slope and the intercept with the y-axis of the regression line, mean bias error (MBE), and normalized mean bias also known as bias. Secondly, this study was to develop a simplified parametrization model for estimating T_{air} during the day and night-time from MODIS LST products and auxiliary data over Berlin, using three supervised learning methods (Adaptive Neuro Fuzzy Inference system (ANFIS), Artificial Neural Network (ANN) and Support vector machine (SVR)), and their performances were validated by cross-validation for each LC/LU. Twelve environmental variables (land surface temperature (LST), normalized difference vegetation index (NDVI), Julian day, latitude, longitude, Emissivity31, Emissivity32, altitude, albedo, wind speed, wind direction and air pressure) were selected as predictors. Moreover, tuning the hyper parameters of some models like SVR and ANN were investigated. For tuning the hyper parameters of SVR, Simulated Annealing (SA) was applied (SA-SVR model) and a multiple-layer feed-forward (MLF) neural networks with three layers and different nodes in hidden layers are used with Levenberg-Marquardt back-propagation (LM-BP), in order to achieve higher accuracy in the estimation of T_{air}. The errors associated with T_{air} estimation based on remote sensing data are often large and strongly limit its applicability (e.g. Czajkowski et al., 2000; Vazquez et al., 1997; Vogt et al., 1997). One of the objectives of this work is to provide T_{air} estimations with an accuracy, which will potentate the future applications. The comprehensive description is provided on

chapter 3. The following major questions will be investigated within this chapter:

Question 3) How is the Spatio-temporal variability of LST-T_{air} relationship: the relationship between remotely sensed LST and T_{air}, is strongly influenced by the local surface heat fluxes, is analysed by overlaying a spatial window of varying size on the MODIS LST grid.

Question 4) How is the relationship between observed T_{air} and the four LST Products over Berlin.

Question 5) What is the best method to estimate air temperature from remote sensed land surface temperature and how can we find the best estimator?

1.4 Feature selection for estimating of near surface air temperature from MODIS over different LC/LU (Germany, Berlin)

The focus of this chapter is on feature set reduction. The problem is important because a high number of features in a dataset leads to model over-fitting, which in turn leads to poor results on the validation datasets. Additionally, constructing models from datasets with many features is more computationally demanding (Korn et al., 2011). All this leads researcher to propose many methods for feature set reduction. The reduction is performed through the processes of feature extraction (transformation) and feature selection. Feature extraction methods such as Principal Component Analysis (PCA, Linear Discriminant Analysis (LDA) and Multidimensional Scaling work by transforming the original features into a new feature set constructed from the original one based on their combinations, with the aim of discovering more meaningful information in the new set (Tang et al. 2014). The new feature set can then be easily reduced by taking into consideration characteristics such as dataset variance coverage. In this chapter, we focus on feature selection and provide an overview of the existing methods that are available for handling several different problems. The objective of this research is to describes the procedures of the input variable selection for estimating T_{air} during day and night time over six LC/LU using both Terra and Aqua MODIS LST products (daytime and night-time) and auxiliary data from 2007–2013. Two input variable selection methods were applied because predictor selection is an essential step in environmental, biological, industrial and climatological applications. Feature Selection helps in understanding data, reducing computation requirement, reducing the effect of curse of dimensionality and improving the predictor performance. Through input variable selection to eliminate the irrelevant or redundant variables, a suitable subset of variables is identified. Meanwhile, the complexity of the model structure is simplified and the computational efficiency is improved. In order to find out which variables, among 12 predefined variables as potential predictors, are the most effective parameter to describe the relationship between LST and T_{air}, brute-force search or exhaustive search (ES) and greedy best first search by using ANN were

applied in this study. In addition, there are several research's which estimated T_{air} but the most recently popular studies of T_{air} estimation using statistic approaches are shown in Table 3.1 in chapter 3. However, most of these studies have only used LST daytime and LST night-time solely for T_{air} maximum (T_{max}) and T_{air} minimum (T_{min}) estimation, respectively. In a recent study (Zeng et al 2015), both LST night-time and daytime were used for estimation of T_{max} and T_{min}. Moreover, Zaksek and Schroedter-Homscheidt (Zakšek et al. 2009) stated that T_{air} is driven more by LST than by direct solar radiation, meaning that LST is the most important variable for T_{air} estimation; in reference to previous studies (Vancutsem et al. 2010, Benali et al. 2012, zeng et al. 2015, Zhang et al. 2011, Jang and Viau 2004) also in consideration of all of the available data that potentially have an effect on the accuracy of T_{air} estimation of the study area. The 12 variables, including LST, emissivity31, emissivity32, Albedo, NDVI, altitude, relative humidity, latitude, wind speed, wind direction, air pressure and Julian day, as the potential variables (candidate variables) are considered for T_{air} estimation. Besides the LST variables, NDVI was selected because it influences the land surface vegetation properties. Elevation, latitude was chosen for capturing the variability of climatic conditions between different regions. We chose Julian day because it reflects seasonal variation in air temperature. In order to assess, rank the feature, select the best model through many candidate model, and find a subset of variables which improves the overall prediction performance several criterion, such as: Root mean squared error (RMSE), Bayesian Information Criterion (BIC), and Akaike Information Criteria (AIC).

The following major questions will be investigated within this study:

Question 6)

To rank the features base on using the Greedy Best-First Search and Brute-force search to find the optimal subset of feature which are influence the relationship between LST and T_{air} during day and night for different LC/LU using RMSE, BIC, and AIC.

Question 7)

Why not just use the global model? It has been argued that one should make inference from a model with all the factors thought to be important. This approach would seem to be simple and avoid the complications of model selection.

Question 8)

How can we select the best approximate model and evaluate them?

1.5 Structure of the Thesis

This thesis is organized into five main chapters and each chapter begins with an introduction, methods and conclusion section.In Chapter 1, is an overall introduction. In chapter 2, method and data are described. In chapter 3, The influence of land cover

type on the relationship between NDVI–LST and LST-T_{air} is briefly discussed. In chapter 4, Estimation of the Near-surface Air Temperature during the Day and Night-time from MODIS in Berlin using ANN, SA-SVR, and ANFIS are briefly explained. Furthermore, an optimization method namely, simulated annealing algorithm is investigated. In chapter 5, a comprehensive description of Feature selection for estimating of near surface air temperature from MODIS over different LC/LU is explained. In chapter 6 is dedicated to conclusions and discussions and final chapter is about outlook.

2 Data and Method

2.1 Study Area

Berlin is the capital city of Germany. It is located in the north-east of the country, covers an area of 892 km^2. Berlin is located on a mostly flat topography. Regarding land use patterns, Berlin is characterized by a significant amount of green areas and water bodies. Outside the inner-city, there is a relatively low building and population density, with many allotment gardens for private cultivation and recreation. There are a considerable number of urban brownfield sites, despite the slight trend of population growth in the last decade. Berlin consists of 45% water bodies and urban green spaces (forested and unforested,allotment gardens), almost 20% transport and infrastructural areas(streets and railways), and around 35% built-up areas (e.g. residential use). Table 2.1 shows the location and related land use of the weather stations in Berlin used in this study.

Table 2.1: Information about weather stations over Berlin, including their LC/LU, Latitude, Longitude and Elevation.

Station	LU	Lat	Long	Elevation(m)
Botanischer-Garten	Green urban area	52.45	13.30	46.88
Fasanenstraße	Industrial, commercial, public, military	52.51	13.33	34.08
Tegel-Forstamt	Forest	52.60	13.27	39.58
Gatow	Industrial, commercial, public, military	52.47	13.13	47.09
Marzahn	Green urban area	52.54	13.58	50.61
Pichelsdorf	Evergreen needle leaf tree	52.50	13.19	29.66
Wannsee	Evergreen needle leaf tree	52.43	13.18	40.77
Dahlem-FU	Industrial, commercial, public, military	52.45	13.31	67.50
Tegel	Airport	52.56	13.30	35.25
Schönefeld	Airport	52.38	13.53	45
Buch	Industrial, commercial, public, military	52.63	13.50	65.45
Marzahn	Green urban area	52.54	13.55	63.29
Kaniswall	Agriculture, semi-natural and wet area	52.40	13.73	32.57

2.2 Data Description

Three main datasets for the period of 2007-2013 according to the availability of meteorological station record and MODIS data were used:

1 Remotely sensed data.

2 Ground measurements from 20 meteorological stations in Berlin.
3 Digital elevation model (Berlin Digital Environmental Atlas).

2.2.1 MODIS Data

The LST product from MODIS has been used in previous studies to derive T_{air}(Benali et al., 2012, Vancutsem et al., 2010, Zhu et al., 2013). MODIS sensors were launched on board the National Aerodynamics and Space Administration (NASA) Observing System (EOS) Terra and Aqua satellites in December 1999 and May 2002, respectively (Zhu et al., 2013). Both sensors are on board sun-synchronous polar orbiting satellites. MODIS Terra data is available during 10:30 – 12:00 a.m. and p.m. (daytime/night-time) local time, while MODIS Aqua sensor collects the imagery during 1:00 – 3:00 a.m. and p.m. (daytime/night-time). All MODIS LST data used in this study were acquired from the U.S. Geological Survey (USGS) website (Piao et al., 2009). We used two MODIS LST products (v005, h18v03), MOD11A1 and MYD11A1 from Terra and Aqua satellites, respectively. The MODIS LST consists of daytime and night-time data at a spatial resolution of 1 km. Thus, in total there are four LST datasets: Aqua daytime, Aqua night-time, Terra daytime and Terra night-time.

2.2.1.1 Vegetation Index

Normalized difference vegetation index is the most common remote sensing index used to parametrize vegetation status (Zhu et al., 2013, Stow et al., 2004, Raynolds et al., 2008). The absorption and reflectivity of the vegetation cover are correlated with their structural properties, such as leaf area index (LAI), fractional vegetation cover (FVC), and their physiological condition (Bustos et al.2014, Raynolds et al., 2006). The values of NDVI vary between −1 and 1, where the range between 0.2 and 0.9 is mostly common in continuous vegetation cover (Bustos et al.2014). In this study, the NDVI was extracted from Terra (MOD13A2.005) and Aqua (MYD13A2.005) products with 16-day temporal and 1km resolution as mentioned in table 2.2.

2.2.2 Meteorological Data

Air temperature observations were obtained from 20 meteorological ground stations in the study area. The measurements included daily T_{air}, wind speed, wind direction, air pressure and relative humidity. The meteorological station records were obtained from the Deutscher Wetterdienst(ftp://ftp-cdc.dwd.de/pub/CDC) and from the Freie university Berlin meteorological station (http://mevis-www.met.fu-berlin.de/devel/mevis). The time selected for our study ranged from 2007 to 2013 according to the availability of meteorological station and MODIS data. Moreover, the accuracy of observation data in meteorological stations is as following:

1. 2m air temperature: $\pm 0.2K$
2. Wind speed: $\pm 0.3\%$ of measured value
3. Wind direction: $\pm 5C^{\circ}$
4. Relative humidity: $\pm 0.3\%$ up to $\pm 0.5\%$
5. Air pressure: $\pm 0.1hpa$

2.2.3 Auxiliary Data

In addition to MODIS products (LST, emissivity31, emissivity32), Albedo and NDVI, some auxiliary variables were used, including latitude, altitude, Julian day, air pressure, wind speed, wind direction and relative humidity. These auxiliary variables either have a known impact on T_{air} and LST or influence the relationship between T_{air} and LST. Latitude, Land cover/Land Use (LC/LU) and altitude were derived from the location of meteorological stations. Altitude was obtained from a 5m resolution digital elevation model(DEM) (downloaded from https://www.eea.europa.eu/data-and-maps/data/urban-atlas). Moreover, Julian day was also considered as proxies for the fraction of solar energy absorption during the day and emission during the night, influencing the diurnal amplitude of T_{air} throughout the year. Julian day is the continuous count of days from 1 January every year. In addition, the LC/LU of each meteorological station was extracted in terms of its position, and reclassified into urban, industrial, forest, airport, needle leaf trees and agriculture based on a 5m resolution map of LC/LU, which was downloaded from https://www.eea.europa.eu/data-and-maps/data/urban-atlas. In addition, all data (Auxiliary and MODIS data) were combined to create a single dataset for each LC/LU for day and night-time. The collinearity of independent variables was detected using variance inflation factor (VIF > 10) and pair wise correlation (r > 0.75) (Zurr et al., 2010, Dormann et al., 2013). More descriptions regarding, temporal matching of T_{air} to LST observations is presented in Chapter 4.

2.3 LST Pre-Processing

A certain number of pre-processing steps were required to convert the original LST product in HDF format to raster layers with a versatile projected coordinate system. Firstly, raster subsets of the LST product were extracted based on the boundary extent of the study area. LST L3 product is gridded in the global Sinusoidal projection, and the grid containing data for the study area is located at column 18(h18 and line 03 (v03). It is important to eliminate low quality data in the MODIS LST data because remote sensing based T_{air} estimates are strongly influenced by errors (e.g., errors caused by clouds and large sensor viewing angles, uncertainties in surface emissivity (Wan et al., 2004). Validation studies of MODIS LST show that under clear sky conditions the precision is approximately 1 K or less, but higher errors would be observed at large viewing angles and in semiarid regions (Wan et al., 2008). So only the pixels of the targeted land cover types that were flagged in the MODIS quality assurance data as cloud free and of high quality were retained.

Table 2.2: Data source and variables

Variable ID	Variable	Abbreviation	Source	Explanation
1	Land Surface Temperature	LST	MODIS	Land Surface temperature derived over the 2007 to 2013 time period using MYD11A1, MOD11A1 product
2	Julian Date	JD	Meteorological data	The continuous count of days was from 1 January to the last day every year.
3	Emissivity31	Emiss31	MODIS	Emissivity31 derived over the 2007 to 2013 time period using MYD11A2, MOD11A1 product
4	Emissivity32	Emiss32	MODIS	Emissivity32 derived over the 2007 to 2013 time period using MYD11A2, MOD11A1 product
5	Normalized Difference Vegetation Index	NDVI	MODIS	Vegetation index at 1km resolution from Terra (MOD13A2.005) and Aqua (MYD13A2.005) products with 16-day temporal resolution
6	Albedo	Albedo	MODIS	Albedo at 1Km resolution from Terra (MCD43B3.005) product with 16-day temporal resolution
7	Relative Humidity	RH	Meteorological data	The RH was extracted for each station during the year of 2007 to 2013
8	Altitude	Alt	DEM	The altitude extracted from a 5m resolution digital elevation model (DEM) according to the location of meteorological stations
9	Latitude	Lat	Meteorological data	The geographical location of meteorological stations was extracted from meteorological metadata
10	Wind Direction	WD	Meteorological data	The WD was extracted for each station during the year of 2007 to 2013.
11	Wind Speed	WS	Meteorological data	The WS was extracted for each station during the year of 2007 to 2013
12	Air Pressure	AP	Meteorological data	The AP was extracted for each station during the year of 2007 to 2013

2.3.1 Calculation of LST for each Weather Station

The LST for each meteorological station were determined as described in the study of Noi et al.,2017. It should be noted that, only the LST data under clear sky then the following steps were applied to retrieve the LST for each station:

- A total of 5110 MODIS files(MOD11A1 and MYD11A1, h18v03, Collection 5, from 1 January 2007 to 31 December 2013 over Berlin) in HDF (Hierarchical Data Format) format were re-projected to WGS-1984-UTM-zone-33N using the nearest neighbour re-sampling method.

The LST-Day-1km, LST-Night-1km, Daytime LST observation time, and Night-time LST observation time) were extracted.

- For each weather stations, LST data are retrieved from MODIS using nearest neighbor algorithm for the pixels in which the weather stations are located.
- Using the following equation, all LST data (DN value) were converted to Celsius temperature where 0.02 is the scale factor of the MODIS LST product.

$$T(C^{\circ}) = 0.02 * DN - 273.15 \tag{2.1}$$

- Removing low quality data: MODIS LST products are not available for a location (pixel), if clouds are present (Wan, 2008). To avoid this kind of data, only the pixels of the targeted land cover types that were flagged in the MODIS quality assurance data as cloud-free and of high quality were retained (Ackerman et al., 2008, Williamson et al., 2013).

3 The influence of land cover type on the relationship between NDVI–LST and LST-T_{air}

3.1 Introduction

Many studies have revealed a strong negative correlation between normalized difference vegetation index (NDVI) and land surface temperature (LST) (Goward, Cruickshanks, and Hope 1985; Hope et al. 1986; Smith and Choudhury 1991; Schultz and Halpert 1995;Churkina and Running 1998; Nemani et al. 2003; Julien, Sobrino, and Verhoef 2006; Sun and Kafatos 2007; Julien and Sobrino 2009; Kumar and Shekhar 2015; Tayyebi and Jenerette 2016; Zhou et al. 2014). Previous studies have determined the variability in the slope of the inverse LST–NDVI relationship, in association with local topographic and environmental conditions. Goward and Hope (1989) stated that the LST–NDVI slope and intercept are expected to vary from one day to another day based on the magnitude of incident solar radiation, advective atmospheric conditions, and surface moisture availability. Hope (1988) revealed that soil moisture potential affects the relationship between canopy temperature and NDVI. Using an LST–NDVI scatter plot, Price (1990) differentiated between areas of full vegetation cover, dry soils, and moist soil. Hope and McDowell (1992) used the LST and NDVI relationship to discriminate between burned and unburned surfaces. The LST and NDVI slope has been utilized in many applications associated with water and energy balance. Research has shown that this slope is related to moisture availability and canopy resistance, indicating vegetation stress and/or soil water stress. Other researchers (Nemani and Running 1989, 1993; Carlson, Gillies, and Perry 1994; Goetz1997) found that the slope is inversely correlated with the crop moisture index (CMI), developed by Palmer (1968) to assess short-term crop water conditions. CMI is based on the concept of abnormal evapotranspiration (ET) deficit, calculated as the dierence between computed actual ET and computed potential ET (i.e. expected or appropriate ET). Actual ET is based on the temperature and precipitation that occurs during a week and computed soil moisture in both topsoil and subsoil layers. In addition, the slope of LST and NDVI is also related to vegetation type, topography, and vegetation cover, respectively (Nemani et al. 1993). Goetz (1997) revealed that the slope varies with climate conditions, with steeper slopes associated with drier situations. Other studies have shown that the slope is related to the rate of ET from the surface. Prihodko and Goward (1997), Boegh et al. (1999), and Goward, Xue, and Czajkowski (2002) used NDVI and LST data to estimate near-surface air temperature. In addition, Weng, Lu, and Schering (2004) found that the vegetation fraction has a slightly stronger negative correlation with LST. Yue et al. (2007) have investigated the mean LST and NDVI values associated with dierent land-use (LU) types in the city of Shanghai, China, and they found out that mean LST and NDVI values associated with different LU types are significantly dierent. Joshi and Bhatt (2012) stated that

the areas with vegetation and water-body have lower temperature compared to the built-up areas. Sun and Kafatos (2007) found that the correlation between LST and NDVI is positive for winter and negative during warm seasons. Moreover, Gorgani, Panahi, and Rezaie (2013) applied a regression technique to obtain the correlation between LST and NDVI in the city of Mashhad, Iran. They found that the regression coefficient from NDVI to LST is negative. Moreover, a considerable amount of research has been conducted to investigate relationships between LST and urban LU/land cover (LC) (e.g. Weng, Lu, and Schering 2004; Yuan and Bauer 2007; Buyantuyev and Wu 2010; Zhou, Huang, and Cadenasso 2011, 2017; Huang and Cadenasso 2016) in Urban Heat Island (UHI) issue because the surface UHI relates directly to land surface characteristics (Oke 1995; Quattrochi and Luvall 1999; Voogt and Oke 2003). In addition, in order to investigate the driving mechanism of UHI more deeply, increasing emphases have been placed on the research of vegetation–LST relationship (Petropoulos, Griths, and Kalivas 2014; Li et al. 2009). Gallo and Owen (1999) assessed seasonal trends in LST and NDVI and found that dierences in NDVI and satellite-based surface temperature accounted for 40% of the variation in urban–rural temperature dierences. Zhou et al. (2014) investigated the seasonal variability of the relationships between LST and LU/LC variables and how the spatial and thematic resolutions of LU/LC variables aect these relationships. They developed 10 models to evaluate eects of spatial and thematic resolution of LU/LC data on the observed relationships between LST and LU/LC variables for each season. They found that the directions of the eects of LU/LC variables on estimation LST were consistent across seasons, but the magnitude of eects, varied by season, providing the strongest predictive capacity during summer and the weakest during winter. Moreover, the correlation between LST and NDVI was examined by Sruthi and Mohammed Aslam (2015) for monitoring and early warning system for the farmers regarding agricultural drought issue in the Study of Raichur District. The LST–NDVI relationships as an indicator of drought were considered by Karnieli et al. (2009). The relationships between NDVI and temperature have also been utilized in various studies to evaluate two variables – fractional vegetation cover and surface soil water content (Gillies, Kustas, and Humes 1997; Goward, Xue, and Czajkowski 2002; Carlson, Gillies, and Schmugge 1995; Gillies and Carlson 1995). Moreover, research on LST has shown that the partitioning of sensible, latent heat fluxes and surface radiant temperature response depends on varying surface soil water content and vegetation cover (Owen, Carlson, and Gillis 1998). A higher level of latent heat exchange was found in areas with higher levels of vegetation, while sensible heat exchange was favoured more by sparsely vegetated areas such as urban areas (Oke 1982). This finding has led to more research focusing on the relationship between LST and vegetation abundance (Carlson, Gillies, and Perry 1994; Gillies, Kustas, and Humes 1997; Gallo and Owen 1999; Gillies and Carlson 1995; Lo, Quattrochi, and Luvall 1997; Weng2001). However, the relationship between NDVI and LST also varies according to the seasons (Sun and Kafatos 2007; Bayarjargal et al. 2006) and surface moisture conditions. The correlation between NDVI and LST decreases when surface moisture increases (Lambin and Ehrlich 1996; Prihodko and Goward 1997; Moran et al. 1994; Sandholt, Rasmussen, and Andersen 2002). Other methods use this relationship to estimate surface ET and soil moisture

(Carlson 2007), coupling the Soil Vegetation Atmosphere Transfer (SVAT) model to the NDVI and LST relationship to study land–atmosphere interactions (Petropoulos, Carlson, and Wooster 2009). To account for seasonal variations, many researchers (Voogt and Oke 1998; Cresswell et al.1999) have used the Solar Zenith Angle (SZA) to correct the air temperature (T_{air}) retrieval. Cresswell et al. (1999) used a regression analysis technique between SZA and LST to retrieve T_{air} obtaining an accuracy within 3 K for 70% of the observations and an accuracy of 5 K and above for the remaining observations. Moreover, several methods and approaches have been developed to interpret the relationships between LST and NDVI, including:

1 The 'triangle' method using soil–vegetation–atmosphere transfer model (Gillies, Kustas, and Humes 1997; Carlson, Gillies, and Schmugge1995; Gillies and Carlson 1995).

2 In situ measurement method (Friedl and Davis 1994).

3 Remote sensing-based method (Betts et al. 1996).

3.2 Material and methods

3.2.1 The study area

Berlin is the capital city of Germany. It is located in the north-east of the country, covers an area of 892 km^2, with a population of 3.5 million. Berlin is located on a mostly flat topography. Regarding LU patterns, Berlin is characterized by a significant amount of green areas and water-bodies. Outside the inner city, there is a relatively low building and population density, with many allotment gardens for private cultivation and recreation. There are a considerable number of urban brownfield sites, despite the slight trend of population growth in the last decade. While some crucial local structural changes have taken place since the early 1990s (especially at the location of the former Berlin wall), the overall LU patterns have remained relatively constant over the past decade. Figure 1 shows the LU map of Berlin. Berlin consists of 45% water-bodies and urban green spaces (forested and unforested, allotment gardens), almost 20% transport and infrastructural areas (streets and railways), and around 35% built-up areas (e.g. residential use). Table 1 shows the location and related LU of the weather stations in Berlin used in this study.

3.2.2 Data description

In order to find the relationship between NDVI and LST, as well as LST and T_{air}, three main data sets for the period of 2007–2013 were used:

1 Remotely sensed data.

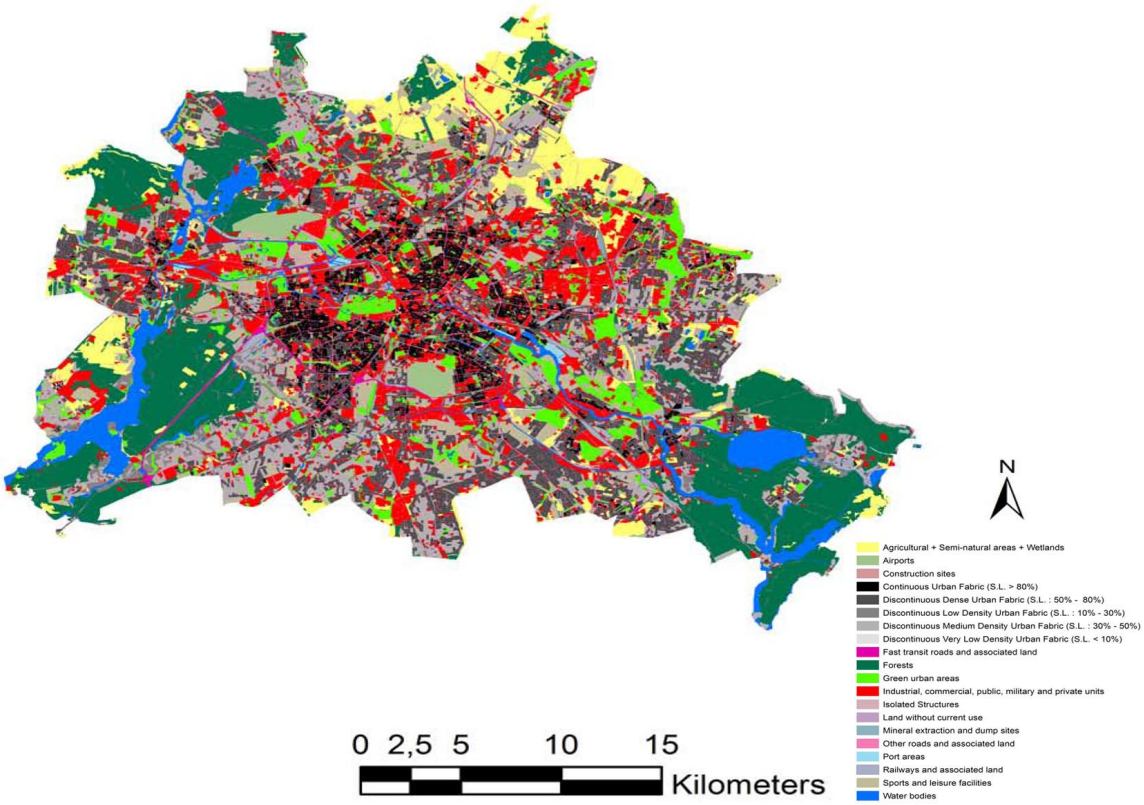

Figure 3.1: Land-use map of Berlin. S.L in Legend means soil sealing.

2 Ground measurements from 20 meteorological stations in Berlin.

3 Digital elevation model (Berlin Digital Environmental Atlas).

The meteorological stations in this study are located in dierent LU/LC types (see Table 3.1, making it possible to study the eect of different LU/LC types). In general, the meteorological stations are categorized into the following six dierent LU/LC types based on a 5 m resolution map of LU/LC which is downloaded from https://www.eea.europa.eu/data-and-maps/data/urban-atlas:

1 Urban areas (green urban areas, discontinuous low-density urban fabric, and discontinuous dense urban fabric)

2 Industrial (Industrial, commercial, public, and military)

3 Forest

4 Airport

Table 3.1: Information about weather stations over Berlin, including their Land cover, Lat, Long and Elevation

Station	Land cover	Latitude	Longitude	Elevation(m)
Botanischer-Garten	Green urban area	52.45	13.30	46.88
Fasanenstraße	Industrial,commercial,public,military	52.51	13.33	34.08
Tegel-Forstamt	Forest	52.60	13.27	39.58
Gatow	Industrial,commercial,public,military	52.47	13.13	47.09
Marzahn	Green urban area	52.54	13.58	50.61
Pichelsdorf	Everegreen needleleaf tree	52.50	13.19	29.66
Wannsee	Everegreen needleleaf tree	52.43	13.18	40.77
Dahlem-FU	Industrial,commercial,public,military	52.45	13.31	67.50
Tegel	Airport	52.56	13.30	35.25
Schönefeld	Airport	52.38	13.53	45
Buch	Industrial,commercial,public,military	52.63	13.50	65.45
Marzahn	Discontinues Dense urban fabric	52.54	13.55	63.29
Kaniswall	Agriculture,semi-natural and wet area	52.40	13.73	32.57
Tempelhof	Airport	52.46	13.40	47.74
Eiskeller	Agriculture,semi-natural and wet area	52.58	13.13	31.78
Kreuzberg	Industrial,commercial,public,military	52.49	13.40	34.91
Wannsee-meteo	Everegreen needleleaf tree	52.43	13.18	43.49
Aldershof	Industrial,commercial,public,military	52.42	13.52	35.15
Potsdam	Industrial,commercial,public,military	52.38	13.11	33.79
Insulaner	Discontinues Low Density urban fabric	52.45	13.35	43.75

5 Agriculture (agriculture and semi-natural)

6 Needle leaf trees area (evergreen needle leaf tree)

The second source of data is the satellite data. The satellite data used in this study are LSTs at 1 km spatial resolution derived from Moderate Resolution Imaging Spectroradiometer (MODIS) sensors on board Terra and Aqua satellites. Aqua has a 1:30 am/pm equator crossing time, while Terra has a 10:30 am/pm equator crossing time, meaning MODIS data are typically available on a daily basis. These data have been downloaded from the Land Processes Distributed Active Archive Center (https://lpdaac.usgs.gov/).The following MODIS products (version 5) were used:

1 MODIS daily LST at 1 km resolution, from Terra (MOD11A1.005)

2 MODIS daily LST at 1 km resolution, from Aqua (MYD11A1.005)

3 MODIS monthly vegetation index at 1 km resolution from Terra (MOD13A2.005)

4 MODIS monthly vegetation index at 1 km resolution from Aqua (MYD13A2.005) product

The MODIS monthly vegetation product contains both NDVI and the enhanced vegetation index (EVI), retrieved from blue, red, and near-infrared reflectance, centred at

469, 645, and 858 nm channels, respectively. The MODIS NDVI and EVI products are calculated from atmospheric bidirectional surface reflectance values that have been masked for clouds, cloud shadows, water, and heavy aerosols. The MODIS LST is derived from two thermal infrared band channels, 31 (10.78–11.28 mm) and 32 (11.77–12.27 mm), using the split window algorithm (Wan et al. 2002). This algorithm corrects for emissivity and atmospheric eects by using a look-up table based on global land surface emissivity in the thermal infrared (Snyder et al. 1998). Daily MODIS LST and monthly NDVI from Aqua and Terra were extracted at the nearest points to the stations during day- and night-time by considering quality assurance information in the MODIS product (Only values labelled as 'good data' or 'marginal data'). The data have a significant number of missing points due to clouds,heavy aerosols, gaps between satellite swaths, and failed retrievals under certain conditions. We combined the Aqua and Terra products by considering their overpass time, so as to increase the maximum of usable observations as Alcantara et al. (2013) did in his research. The third source of data was the URBAN ATLAS (www.eea.europa.eu/data-and maps/data/urban-atlas, 2012), used for the LU/LC classification.

3.2.3 Day/night analysis

Apart from spatial variations,observation time can aect the relationship between LST and T_{air} time series. To identify any variability in LST and T_{air} relationship in a diurnal basis, time series of both variables were separated based on the MODIS overpass times to produce two series over single pixel window from MODIS-Terra and MODIS-Aqua day and night. In addition, in order to analyse the seasonal eect on the relationship between NDVI and LST, and between LST and T_{air}, the seasonal time series were created for each LU/LC during day- and night-time. Moreover, for temporal matching of LST with the T_{air} from meteorological stations, the two approaches were considered: (1) For those meteorological stations that minutely data are available, the air temperature at over pass time of the satellite was collected. (2) But for other meteorological stations with hourly data, a linear equation was considered for calculating the T_{2m} at the overpass time of satellite.

3.2.4 Statistical methods

A simple linear relationship is often assumed between LST and T_{air} and also between LST and NDVI in the literature (Brunel 1989; Mostovoy et al. 2006). In view of this, a univariate linear regression analysis with the MODIS LST as the independent (or explanatory) and T_{air} or NDVI as the dependent (or response) variable was applied to analyse LST–T_{air} and also LST–NDVI relationships, respectively. The correlation coefficient, r, is reported as a quantitative measure to evaluate the strength of the agreement between them in the analysis. Significance levels (p values) are addressed in the results to explain how unlikely the given r values would occur if no relationship between the explanatory and response variables did exist, where the smaller the p

level, the more significant the relationship. Moreover, in order to investigate the LST relationship with air temperature at 2 m height (T_{2mDay}/$T_{2mNight}$), statistical parameters such as correlation coefficient (r), mean value, mean absolute error (MAE), and root mean square deviation (RMSD) were calculated for six LU/LC types from 2007 to 2013.

3.3 Relationships between NDVI and LST

Fractional vegetation cover depicts the amount and nature of vegetation cover and also quantises the proportions of vegetation and ground visible to a sensor. The dierences in radiative temperature between the vegetation canopy and the ground aect the measurement of LST (Sandholt, Rasmussen, and Andersen 2002). For areas without vegetation, LST measurements typically represent the radiometric temperatures of sunlit surfaces, such as bare soil. As the amount of vegetation cover increases, the radiative temperature recorded by a sensor more closely approximates the temperatures of green leaves and the canopy temperature at spectral vegetation maximum or complete canopy cover (Goward, Xue, and Czajkowski 2002). LST measurements are also subject to the influence of the lower atmosphere and the temperature dierence between the vegetation canopy and the soil background (Friedl 2002). Thermal responses of vegetation can also be highly dependent on the biophysical properties of the vegetation itself (Quattrochi and Ridd 1998). Some internal properties, such as thermal conductivity, heat capacity, and inertia, play significant roles in controlling the temperature of a body at equilibrium with its neighbourhood (Campbell 2002). These thermal properties dier with moisture content and soil type (Sandholt, Rasmussen, and Andersen 2002). As a result of relatively low thermal inertia, bare, dry, and low-density soils have been linked to the highest LST (Carnahan and Larson 1990). Soil emissivity is a function of soil density and moisture (Larson and Carnahan 1997). Therefore, for areas characterized by low vegetation cover, surface thermal properties can strongly aect LST, through the thermal processes of convection, radiation, and conduction. Moreover, the relationship between LST and vegetation indices, such as NDVI, has been investigated.

3.4 Results and discussion

3.4.1 Seasonal variation of the LST and NDVI relationship

In order to better understand the relationship between LST and NDVI over dierent LU/LC types, the thermal environment and green space signature of each LU/LC type must be considered. To achieve this goal, the mean values of LST and NDVI over different LU/ LC were compared. Figure 3.2(a,b) gives information about mean LST and NDVI values based on dierent LU/LC types between 2007 and 2013 for spring season in Berlin, respectively. As shown in Figure 2.2, the industrial, green urban area, and airport LUs show the highest mean LST values. These LUs include more impervious

surfaces primarily associated with transportation (street, highway, and parking lots) and a large amount of anthropogenic heat. This anthropogenic heat is less significant for industrial LU as they are also surrounded by vegetation cover. In contrast, the highest mean NDVI values can be seen in needle leaf trees, agriculture, and forest LUs. Spatial variation of NDVI is not only related to the amount of vegetation, but also aected by solar radiation availability, topography, slope, and other factors. Greater vegetation cover causes comparatively higher rates of ET and favours latent over sensible heat exchange between the surface and the atmosphere (Wilson et al. 2003). As Figure 3(b) shows, the mean ET value (free water evaporation) in all seasons for six LU/LC types depicted higher rates of ET for forest, agriculture, and needle leaf trees. The lowest mean LST, however, is neither in evergreen needle leaf trees (except in one case) nor in agriculture LUs. One reason could be that these areas are spread throughout urban areas and are highly aected by the UHI and ambient LU/LC types. Forest areas are far from the urban centre, with a low influence from the UHI eect and relatively homogeneous land surface material, leading to a low mean LST. This is primarily attributed to the fact that even during the conditions when maximum temperatures occur, forests are able to access water with their deep root systems and continue transpiration. A greater proportion of incoming solar radiation is partitioned to latent heat flux as a result of transpiring vegetation, thereby cooling the canopy surface temperature. Additionally, forests have deep, complex canopies that promote cooling through turbulent exchange. Overall, by considering and comparing each LU/LC type within each class, Figure 3 demonstrates a direct negative relationship between LST and NDVI values. It shows that a higher NDVI is related to lower LST values. However, difficulties still exist in the interpretation of the relationship between LST and NDVI. These results do not provide enough information to cover and analyse every aspect of this relationship because the NDVI measurements depend on the visible and near infrared reflectance from the plant canopy, reflectance of the same spectra from the soil, and the atmospheric reflectance and are subject to the impact of observational errors (Yang, Yang, and Merchant 1997). Plant species, soil background, leaf area, and shade can all contribute to NDVI variability (Jasinski 1990). The relationship between NDVI and other measures of vegetation abundance is known to be non-linear (Asrar et al. 1984; Small 2001; Small and Lu 2006). Moreover, these results may suggest that percent of vegetation or percent of tree canopies are better predictors of LST than NDVI (Tayyebi and Jenerette 2016; Zhou, Wang, and Cadenasso 2017). In order to better understand the seasonal variation of NDVI and LST over different LU/LC, a correlation analysis is carried out. As shown in Table 3.2, the correlation between LST and NDVI varies with the seasons and time of day. In general, during summer, the negative correlations between NDVI and LSTDay are much stronger than those between NDVI and LSTNight because vegetation cover has a key influence on maximum surface temperature (e.g. leaves, for example, are considerably more efficient at shedding absorbed energy than the soil surface, even when not transpiring. They also have a significantly cooler surface temperature than bare soil). Moreover, in winter, the correlation between NDVI and LST for some LU/LC such as industrial is as same as needle leaf trees, forest, and agriculture due to the fact that the NDVI values in winter were almost less (which indicated that there was very little actively

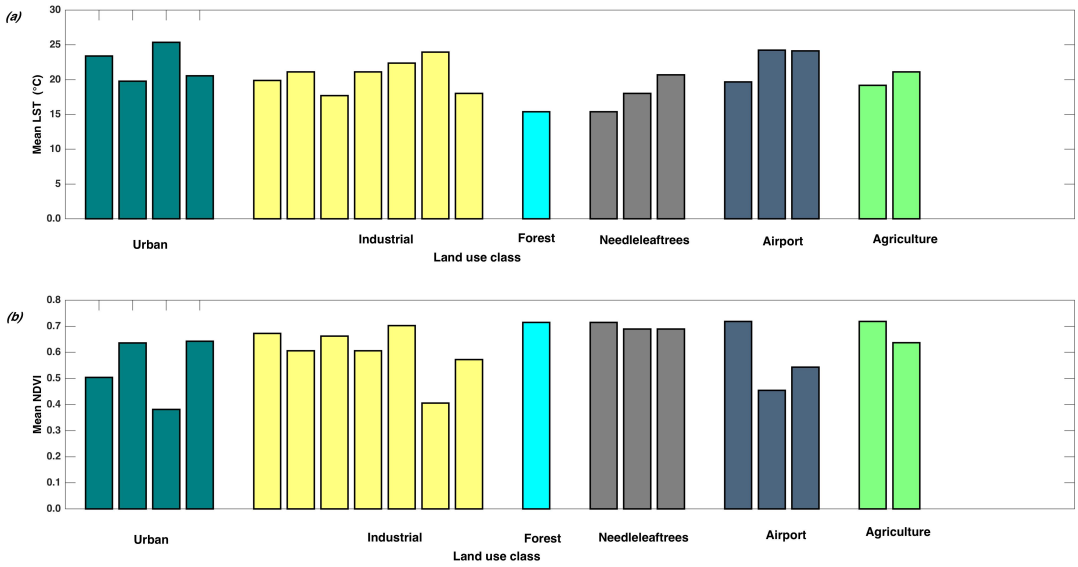

Figure 3.2: Each of the bars in this figure shows the mean LST and NDVI values in subplots(a) and (b)respectively, for each individual weather station in Berlin during 2007-2013.

photosynthesizing vegetation during this period). Consequently, plants play much less important role in regulating LST during winter, and variation in LST mostly is related to non-vegetated surfaces (i.e. impervious surfaces) in urban areas in that season (Yuan and Bauer 2007). In addition, during all seasons except summer, the positive correlations between NDVI and LSTDay/Night are observed for all LC types. This is in good agreement with the results in North America in the study of Kaufmann et al. (2003).

Table 3.2: Correlation coefficient between $LST_{Max/Min}$-NDVI during different season and LU/LC

	Corr($LST_{Max/Min}$-NDVI)			
LC/LU	**Winter**	**Spring**	**Summer**	**Fall**
Green urban area	0.15/0.19	0.10/0.13	-0.48/-0.33	0.65/0.46
Industrial	0.52/0.40	0.44/0.62	-0.12/0.20	0.66/0.51
Forest	0.11/0.10	0.53/0.56	0.29/0.14	0.45/0.43
Evergreen needleleaf trees	0.34/0.35	0.55/0.53	0.19/0.17	0.63/0.49
Airport	0.46/0.41	0.48/0.54	-0.46/-0.18	0.56/0.46
Agriculture	0.49/0.40	0.50/0.59	-0.37/-0.20	0.70/0.50

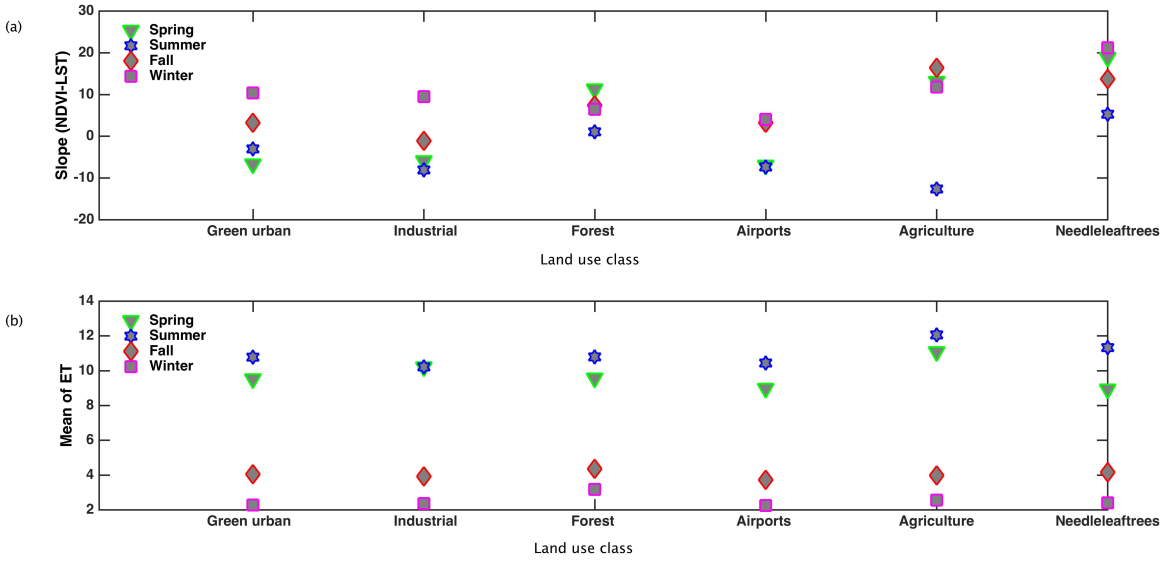

Figure 3.3: (a)Variation of slope between NDVI and LST and (b)mean ET values during different seasons for all LU/LC types.

3.4.2 The effects of LU/LC on the relationship between LST and NDVI

In order to investigate the influence of LU/LC types on the relationship between LST and NDVI, a regression analysis was utilized for each season and LU/LC. A number of studies have explored the relationship between LST and NDVI. Weng, Lu, and Schering (2004) used seven LU/LC types (industrial land, commercial, residential, cropland, grassland, pasture, forest, and water) at dierent scales; all the results showed a noticeable inverse correlation. Lo, Quattrochi, and Luvall (1997) used high-resolution thermal infrared imagery, revealing a strong negative correlation between NDVI and the radiant temperature of residential, agricultural, and vacant/transitional LC types in Huntsville, indicating that the irradiance of an LC type is greatly influenced by the amount of vegetation present. Sobrino and Raissouni (2000) used the inverse relationship between NDVI and radiant surface temperature measurements collected in multi-temporal Advanced Very High Resolution Radiometer (AVHRR) imagery to characterize LC dynamics in Morocco. In order to test the quantitative relations of NDVI to LST, the regression analysis in Equation (1) was considered for each LU/LC type and the results are shown in Table 3.3.

LST_{Day} and LST_{Night} are the daytime and night-time LST, respectively. a and b are regression coefficients. One of the goals of this work is to understand the eect vegetation (NDVI) has on LST. Regression analysis results reveal that there is an inverse correlation between LST and NDVI values for the dierent LU/LCs. Table 3.3 shows that the NDVI–LST slope is highly dependent on the arrangement of the natural

and artificial physical features of an area. There is a negative slope for areas with sparse vegetation cover, such as industrial, airports, and urban areas. For closed vegetation canopies such as agricultural areas, needle leaf trees areas, and forests, the slope is positive. It can be concluded that in some areas with greater vegetation cover (greater NDVI), a higher rate of ET is expected, causing an increase in latent heat exchange between the land surface and atmosphere. For this reason, the daily free water evaporation value (Fitzpatrick and Stern 1966) was calculated for each LU/LC type and then the mean ET value was calculated for each season. This is shown in Figure 3.3(b). As expected, a higher ET rate is in needle leaf trees, forest, and agriculture areas, in spring and summer, because shedding of leaves from deciduous trees in winter greatly reduces the canopy ET (latent heat flux) and thus reduces the capacity of trees in regulating surface temperature due to latent heat exchange. In addition, the cooler air temperature in winter also would greatly reduce the ET of plants. As shown in Figure 3.3(a), the LST–NDVI slope varies with the seasons. A noticeable decrease in slope occurs from spring to summer, especially for dense vegetation areas; this reduction is less clear in sparse vegetation areas such as airports. As expected, the highest slope is observed in needle leaf trees, forest, and agriculture areas during warm seasons. In contrast, the lowest slope is seen in industrial, airport, and urban areas. During the cold seasons, the highest slope is seen in needle leaf trees areas. Moreover, the results indicate that the slope is related to moisture availability (Nemani and Running 1989; Carlson, Gillies, and Perry 1994; Gillies, Kustas, and Humes 1997; Goetz 1997; Palmer 1968), canopy resistance indication vegetation stress, or soil water stress. They also show that greater NDVI can cause higher rates of ET and increase the latent heat exchange between the land surface and atmosphere. Therefore, a decrease in slope occurs in summer, and a notable increase in slope is seen from summer to autumn. To further investigate the influence of LU/LC on the relationship between LST and NDVI, a regression analysis was applied for each season during daytime and night-time for each individual station as shown in Table 3.1. Figure 3.4 show that daytime regression coefficient are higher than night-time during cold months. The variation of regression coefficient is more noticeable during warm months due to the duration of heat waves and the solar radiation (Zhou et al. 2014). The solar radiation that appears is the most dominant driver for the LST–NDVI relationship at the beginning and end of growing seasons (from spring to summer). A regression analysis of the NDVI–LST relationship during day- and night-time in different seasons is shown in Table 3.4. The regression coefficient indicate that increases in NDVI cause lower LST_{Day} in spring, summer, and autumn. During the warm months from spring to summer, the regression coefficient indicate that the cooling eect of vegetation is stronger during daytime than night-time (Sun and Kafatos 2007). Overall, the regression coefficient varies from cold seasons to warm seasons during day- and night-time. The coefficient for all LU/LC mimics almost the same pattern among cold months during day- and night-time. To understand how the relationship between NDVI and LST varies over different seasons, a scatter plot of 7 years (2007–2013) from winter to autumn over different LU/LC is shown in Figure 5. X and Y axes depict the NDVI and LST, respectively. As shown in Figure 5, the relationship between LST and NDVI is strongly affected by season. The negative relationship is observed in spring

(growing season), except in airport areas. This is similarly to the study of Kaufmann et al. (2003). To sum up, this study found that the relationship between LST and NDVI depends on seasonal changes and time of day. The correlation and regression coefficient from NDVI to LST are positive in cold seasons. The negative relationship is found in warm seasons.

Table 3.3: Linear regression equation and correlation coefficients for the relationship between NDVI and LST by LU/LC types

Land use type	Regression functions	r	Mean-NDVI
Green urban area	LST=9.06NDVI+14	0.57	0.63
Industrial	LST=-5.9NDVI+23.86	-0.26	0.67
Airport	LST=8.17NDVI+20.50	0.48	0.45
everegreen needleleaf tree	LST=18.82NDVI+23.57	0.55	0.71
Forest	LST=11.32NDVI+22.48	0.53	0.69
Agriculture	LST=13.10NDVI+12.75	0.50	0.63

Table 3.4: Linear regression coefficients A in Equation $LST_{Max/Min}$=a*NDVI+b during different seasons by LU/LC

	Regression slope$_{Max/Min}$ with NDVI			
LC/LU	**Winter**	**Spring**	**Summer**	**Fall**
Green urban area	10.36/6.92	-6.70/-2.08	-3.04/-0.13	3.22/3.39
Industrial	6.72/1.58	-5.91/-3.23	-8.06/1.04	-1.12/-0.84
Forest	6.45/5.61	11.32/10.27	7.06/6.04	7.35/4.76
Evergreen needleleaf trees	21.20/16.47	18.82/13.39	5.26/3.28	13.67/6.86
Airport	4.06/4.98	-6.92/-2.54	-7.39/-1	3.23/1.15
Agriculture	11.79/7.06	13.10/12.87	-12.65/-4.12	16.37/8.77

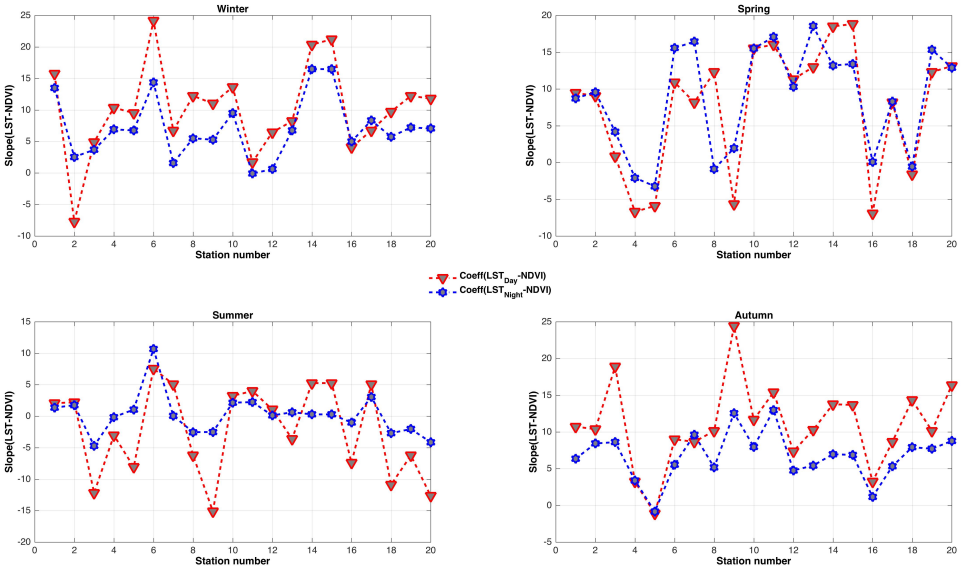

Figure 3.4: (a)Variation of slope between NDVI and LST for each meteorological station during all seasons are shown in subplot(a)in winter, (b)in spring,(c)summer, and (d) in autumn, respectively.

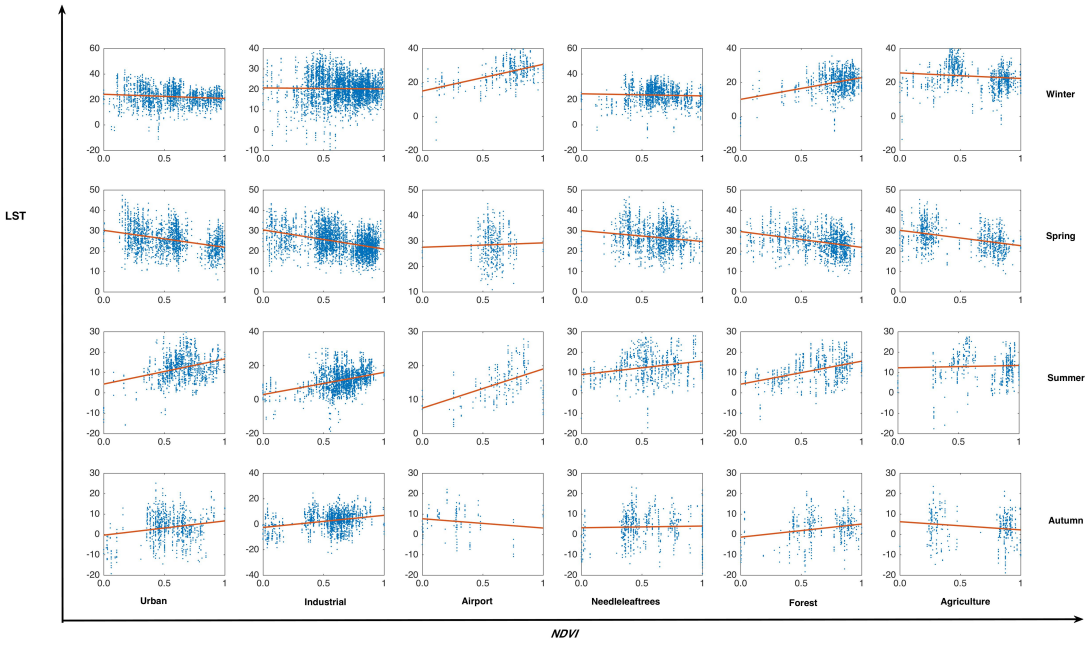

Figure 3.5: The 7 years scatter plots of LST-NDVI between 2007 and 2013 during different seasons and LU/LC axes depict the NDVI and LST, respectively.

3.4.3 LST relationship with air temperature at 2 m above ground

The standard meteorological T_{air} is measured in a shelter at 2 m height (Jin and Dickinson 2010). It is an important descriptor of terrestrial environmental conditions across the Earth (Peón, Carmen, and Javier 2014) and one of the most widely used climatic variables in global change studies. It plays an important role in multiple biological and physical processes among the hydrosphere, atmosphere, and biosphere (Stisen et al. 2007; Shamir and Georgakakos 2014; Benali et al. 2012). Monitoring and understanding the trends of T_{air} and LST are crucial in the study of regional and global climate change (Yoo et al. 2011). Some studies show that LST is usually higher than T_{2m} during daytime, while the opposite occurs at night-time (Cresswell et al. 1999; Benali et al.2012). Therefore, observation time is a key factor which has impact on the relationship between LST and T_{2m}. Tables 3.5 – 3.8 show the relationship between LST and air temperature ($T_{2mDay,2mNight}$) which is included some statistical parameters such as correlation coefficient, mean value, MAE, and RMSD for six LU/LC types from 2007 to 2013. A simple linear relationship is considered between LST and T_{2m}. A univariate linear regression analysis was applied, with the MODIS LST as the independent (or explanatory)and T_{2m} as the dependent (or response) variable, to investigate the relationship between LST and T_{2m}. The correlation coefficient, r, is considered as a quantitative measure to evaluate the strength of the agreement between LST and T_{2m} time series in different LU/LC. As shown in Tables 3.5–3.8 and Figure 3.6, the results are as follows:

1 The relationship between MODIS LST of daytime and night-time (LST_{Day} and LST_{Night}), with T_{2mDay} and $T_{2mNight}$, was estimated. As shown in Figure 3.6 and Tables 3.5–3.8, the correlation between LST and T_{2m} is smaller during the day ($0.42 \leq r \leq 0.81$, $p = 0.0001$) than during the night ($0.46 \leq r \leq 0.86$, $p = 0.0001$) during warm seasons (Zhang and Shen 2011; Benali et al.2012). The dierence (RMSD) between them is greater during the day than at night, as solar radiation does not aect the thermal infrared signal during the night (Vancutsem et al. 2010). The difference between LST and T_{2m} in the presence of radiation depends on the surface energy balance. In every season, a higher correlation between LST and T_{2m} was observed in forest areas, as the remotely sensed LST represents the tree canopy temperature, which is closely correlated to air temperature. The daytime correlation coefficient is higher during cold seasons ($0.77 \leq r \leq 0.89$) than in warm seasons ($0.57 \leq r \leq 0.81$), due to higher turbulence between land surface and atmosphere in the summer months (in comparison to winter months with lower incoming solar radiation) (Zhang and Seidel 2011). Moreover, for all LU/LC types during cold seasons, a closer correlation was observed during daytime ($0.77 \leq r \leq 0.89$) than night-time ($0.47 \leq r \leq 0.86$), while a reverse relationship was obtained for warm seasons. This is because the correlation is mainly driven by the radiative cooling at night, as radiative heating during daytime in the cold seasons is relatively small (Pérez Díaz et al.2015). It must be noted that only clear sky days with high radiative cooling were considered. LST was selected as an indicator for radiative cooling. In contrast, for all LU/LC types during warm seasons, a higher agreement was observed during night-time ($0.46 \leq r$

≤ 0.86) than daytime (0.42 ≤ r ≤ 0.81), due to the higher solar radiation in daytime.

2 As shown in Tables 3.5–3.8, LST is lower than T_{2m} in most cases, during daytime and in cold seasons. In warm seasons, a reverse relationship is observed for all LU/LC types. Generally, during night-time, and depending on the LU/LC, LST is either close or lower than T_{2m} (Haashemi et al. 2016).

3 For all seasons, as expected, higher dierences between T_{2mDay} and LST_{Day} are observed in industrial, urban areas, and airports, respectively. The areas covered by vegetation, bush, trees such as forest, agriculture, and needle leaf trees show lower differences between T_{2mDay} and LST_{Day} because dense vegetation can reduce the amount of heat stored in the soil and surface structures through transpiration. We are unable make conclusions on which LU/LC types have the highest RMSD. This results warrant further research on the cooling effectiveness of vegetation/trees and spatial configuration of trees in the study area.

4 These results show that LC types, seasons, and time of day help to create an undulating temperature surface across dierent LU/LC types and influence the relationship between LST and air temperature. Further investigation is needed to examine this relationship.

3.5 Summary and conclusions

This study investigates both LST–NDVI and LST–T_{2m} relationships over different LU/LC over Berlin, during different seasons at day and night. In the analysis of different LU/LC types, the results indicate that different LU/LC types have significantly different effects on LST and NDVI as measured by the MODIS in Berlin. The results reveal that NDVI values are dependent on seasonal variations. Furthermore, the correlation analysis between NDVI and LST for different seasons show that this relationship depends on the season, time of day, and LU/LC type. The results depict an inverse correlation between LST and NDVI over every LU/LC type. In order to evaluate the influence of LU/LC, a regression analysis between LST and NDVI was utilized and multiple comparisons were made. By considering the slope of regression function for different LU/LC and seasons, it was found that the regression coefficient is dependent on LU/LC type. This means that the NDVI–LST slope is highly dependent on the arrangement of the vegetation type, natural features, rate of ET from surface, and artificial physical features of an area. There is a negative slope for sparse vegetation covers such as industrial, airports, and urban area, but for closed vegetation canopies such as agriculture and needle-leaf trees areas, the slope is positive. The comparisons of mean LST and NDVI values by individual pairings of LU/LC types were made in the research area. The mean LST and NDVI within every LU/LC show a clear negative correlation

between LST and NDVI. In certain LU/LC type, it is observed that the relationship between LST and NDVI is varied. This reveals that the LST is affected mainly by the land surface materials and has a close relationship with the abundance of vegetation and the effect to its surrounding. The strength of relationships between LU/LC variables (e.g. NDVI and vegetation fraction) and LST, however, varies significantly from study to study due to the different measurement of variables and units of analyses (Huang, Guan and Ji 2012; Li et al. 2011). A statistical analysis was applied to determine the variation of the LST–T_{2m} relationship during day- and night-time, for different seasons and LU/LCs. It was found that in most cases, during daytime and in cold seasons, LST is lower than T_{2m}, while in warm seasons, a reverse relationship is observed for all LU/LC types. Generally, for all seasons during night-time, LST is lower than T_{2m}. Moreover, higher difference between $T_{2mNight}$ and LST_{Day} are observed among industrial, urban area, airports, and areas covered by vegetation, bush, and trees (such as forest, agriculture, and needle leaf trees). These results indicate that LU/LC types, seasons, and temporal variation influence this relationship. A linear regression analysis, with the MODIS LST as the independent and T_{2m} as the dependent variable, was applied to analyse this relationship seasonally. The results reveal that LSTDay and LSTNight are significantly correlated (p=0.0001) with T_{2mDay} and $T_{2mNight}$. The correlation between LST_{Day} and T_{2mDay} is higher during cold seasons ($0.77 \leq r \leq 0.89$) than in warm seasons ($0.57 \leq r \leq 0.81$). Moreover, for all LU/LC types during cold seasons, a higher agreement was observed during daytime ($0.77 \leq r \leq 0.89$) than night-time ($0.47 \leq r \leq 0.86$), while a reverse relationship was obtained for warm seasons. In addition, it would be interesting to examine whether and how the size of the analytical unit may affect the observed relationships between LST–NDVI and LST and T_{air} in future research, considering the inhomogeneity of land surface characteristics within a grid box.

Table 3.5: Statistical parameter between $LST_{Day/Night}$ and $T_{2mDay/2mNight}$

Winter

LU/LC	r	Mean (T)	Mean(LST)	MAE	RMSD
Green urban area	0.77/0.61	4.54/-0.90	3.39/-4.40	2.63/3.64	3.45/5.11
Industrial	0.79/0.55	4.31/-0.27	2.73/-4.68	2.64/4.67	3.56/6.15
Forest	0.89/0.46	5.83/-1.26	6.33/-2.81	1.78/3.01	2.30/4.72
Airport	0.81/0.61	3.94/-0.45	3.66/-4.62	2.43/4.36	3.08/5.59
Agriculture	0.86/0.62	4.83/-2.34	2.84/-4.90	2.50/3.33	3.20/4.90
Needle leaf trees	0.79/0.59	4.64/-0.54	3.93/-3.32	2.73/3.33	3.53/4.76

Table 3.6: Statistical parameter between $LST_{Day/Night}$ and $T_{2mDay/2mNight}$

Fall

LU/LC	r	Mean (T)	Mean(LST)	MAE	RMSD
Green urban area	0.81/0.70	13.53/5.94	12.28/4.01	2.27/2.41	2.99/3.38
Industrial	0.79/0.56	13.22/6.14	11.46/3.29	2.65/3.55	3.49/4.50
Forest	0.83/0.62	14.57/5.40	15.15/5.93	1.73/1.99	2.29/2.78
Airport	0.85/0.68	12.88/5.97	12.63/3.32	1.85/3.02	2.52/3.80
Agriculture	0.86/0.63	13.88/4.36	11.18/3.07	2.95/2.41	3.69/3.39
Needle leaf trees	0.86/0.73	14.42/6.84	13.03/5.08	2.29/2.30	2.89/3.12

Table 3.7: Statistical parameter between $LST_{Day/Night}$ and $T_{2mDay/2mNight}$

Spring

LU/LC	r	Mean (T)	Mean(LST)	MAE	RMSD
Green urban area	0.57/0.71	20.60/9.93	22.16/7.93	3.79/2.72	4.86/3.77
Industrial	0.66/0.65	21.42/10.52	20.29/7.41	3.16/3.75	4.12/4.71
Forest	0.81/0.86	20.73/7.74	25.50/9.84	4.97/2.56	5.62/3.01
Airport	0.72/0.76	19.48/9.90	23.36/7.48	4.48/3.01	5.27/3.75
Agriculture	0.71/0.81	22.07/7.69	20.17/6.97	3.12/1.99	4.16/2.64
Needle leaf trees	0.69/0.80	21.26/10.52	23.81/9.18	3.84/2.22	4.76/2.88

Table 3.8: Statistical parameter between $LST_{Day/Night}$ and $T_{2mDay/2mNight}$

Summer

LU/LC	r	Mean (T)	Mean(LST)	MAE	RMSD
Green urban area	0.51/0.50	25.18/15.64	25.53/13.87	3.41/2.51	4.44/3.57
Industrial	0.47/0.46	26.08/16.47	24.42/13.21	3.78/3.80	4.85/4.72
Forest	0.70/0.74	25.20/13.71	28.52/15.80	3.90/2.48	4.71/2.86
Airport	0.69/0.63	23.98/15.48	26.81/13.38	3.64/2.67	4.47/3.32
Agriculture	0.42/0.66	25.81/13.54	24.55/12.68	3.89/2.05	4.86/2.61
Needle leaf trees	0.60/0.67	26.12/16.45	27.13/14.86	3.17/2.27	4.03/2.91

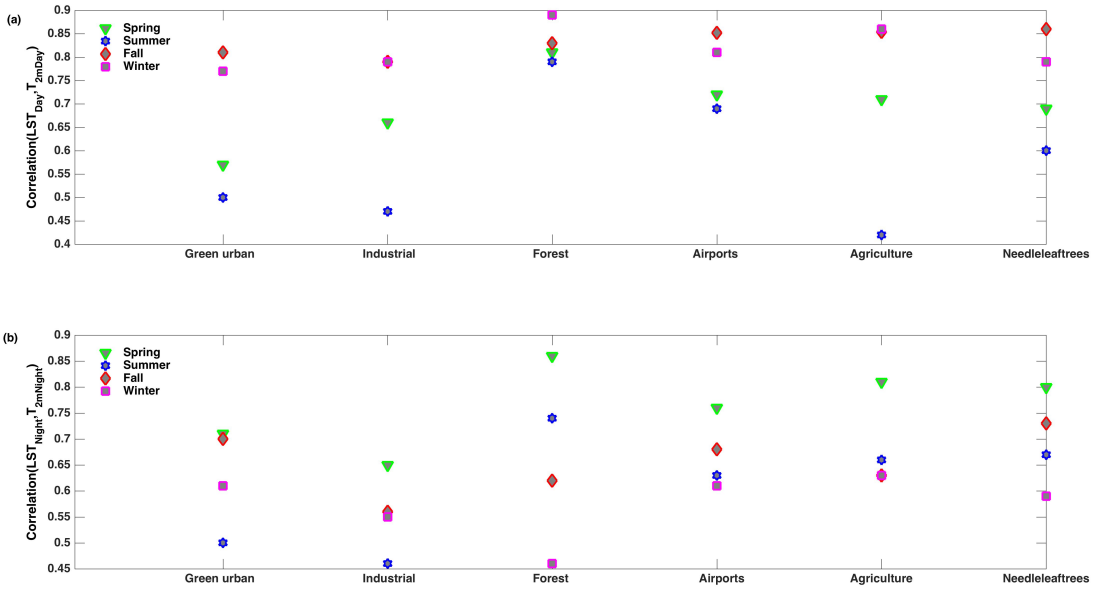

Figure 3.6: The plots of correlation coefficient between LST and $T_{2mDay/2mNight}$ during different seasons over different LU/LC during day and night-time are shown in subplots (a) and (b), respectively.

4 Estimation of the Near-surface Air Temperature during the Day and Night-time from MODIS in Berlin, Germany

4.1 Introduction

The standard meteorological T_{air} is measured in a shelter at 2m height (Brunel 1989; Jin and Dickinson (2010). It is an important indicator of terrestrial environmental conditions across the earth (Prihodko and Goward (1997); Peón et al. 2014) and one of the most widely used climatic variables in global change studies. It plays an important role in multiple biological and physical processes among the hydrosphere, atmosphere and biosphere (Stisen et al. 2007; Shamir et al. 2014; Benali et al. 2012). Regarding ecosystem, it influences the distribution of plant species (Cabrera 2002) and affects the dynamics of the soil–plant–water system (Chartzoulakis and Psarras 2005; Zavala 2004), being included in evapotranspiration models (Allen et al. 2006; Carlson et al. 1995) as well as hydrological models (Purkey et al. 2007; Yates et al. 2005). At the individual level, temperature affects plant growth and net primary productivity since photosynthetic and respiration rates depend on it. Moreover, T_{air} plays a critical role in vegetation distributions, phenology, and growth (Benavides et al. 2007; Stahl et al. 2006). The maximum temperature also shows significant relationship with the occurrence of wildfire on hot and sunny days (Aldersley et al., 2011; Litschert et al., 2012). Therefore, detailed knowledge of the spatial variability of air temperature is of interest for many research and management. In addition, T_{air} plays an important role in energy balance and is a key input in various environmental models and applications, such as crop evapotranspiration estimation (De Bruin et al. 2010), distributed hydrology (Gao et al. 2014) and climate change models (Lofgren et al. 2011). Moreover, the importance of temperature in urban area are related to heat stress and human health. Meteorological measurements provide accurate discrete T_{air} information for specific locations but have limited ability to describe its spatial heterogeneity over large areas (Benali et al., 2012; Willmott and Robeson 1995). The non-uniform spatial distribution of weather station locations within most networks and the complexity of the land surface conditions and patterns make it a challenge to get spatial-continuous T_{air} data. However, weather stations are usually sparsely distributed in mountainous regions, especially in high-elevation areas, and thus may not optimally represent all environments (Rolland 2003). Given the large spatial heterogeneity of T_{air} in complex terrain (Holden et al., 2011), it is difficult to accurately characterize the distribution of T_{air} over mountainous areas (Carrega 1995). Different interpolation methods have been used to generate spatially continuous T_{air} from point station measurements (Benavides et al., 2007; Dodson and Marks 1997; Duhan et al., 2013; Kurtzman and Kadmon 1999; Stahl et al., 2006). However, the performance of interpolation methods is highly dependent on the spatial density and distribution of weather stations (Chan and Paelinckx 2008; Vogt et al., 1997), which is not considered

satisfactory in mountainous areas. Satellite remote sensing observations from global imaging sensors, such as the Advanced Very High-Resolution Radiometer (AVHRR) and Moderate Resolution Imaging Spectroradiometer (MODIS), represent a potentially valuable alternative to characterize spatially-detailed T_{air} patterns across large areas. A split window technique was applied to AVHRR (Pinheiro et al. 2006), MODIS (Wan et al., 2002), and Meteosat (Atitar and Sobrino, 2009) thermal data to estimate Land Surface Temperature (LST). The science-grade quality of the LST data collected by MODIS has proven valuable for monitoring land surface dynamics over large areas (Benali et al., 2012, Mostovoy et al., 2006, Lin et al., 2012). The earth's surface is heated by solar radiation, while the atmosphere is mainly heated from the ground up through long-wave infrared radiation (Frederick et al., 2006). The relationship between Land Surface Temperature and T_{air} may vary with time and location, because the land surface energy balance is a complex phenomenon that depends on multiple factors (e.g., cloud cover, surface roughness, wind speed and soil moisture), whereas some of them (e.g., wind speed) are usually not available from satellite (Goward et al., 1997; Prince et al., 1998; Stisen et al., 2007). An accurate estimation of T_{air} and the mapping of its spatial distribution are useful for predicting ecological consequences of climate change. For example, climate warming will lead to higher temperatures and an increase of extreme weather conditions, which are associated with changes in wildfire regime (Westerling et al., 2006; Chen et al., 2011; Manzo-Delgado et al., 2009), forest biomass distribution (Reich et al., 2014) and crop yield (Ruane et al., 2014; Rosenzweig et al., 2014). The demand for accurate spatial T_{air} data over a large scale has continued to rise (Oyler et al., 2015; Beier et al., 2012). However, the spatial distribution of the weather stations in many parts of the world, is often limited which restricts the use of T_{air} measurements over a large spatial domain (Vancutsem et al., 2010). LST, but on the other hand, is measured in a global extent with significant higher spatial coverage (Jin and Dickinson, 2010). The US National Research Council and the Intergovernmental Panel on Climate Change (IPCC) expressed the need for long-term remotely sensed LST data in global warming studies to overcome the limits of conventional surface T_{air} measurements (IPCC 2007, Jin 2004). Remote sensing data has great potential to estimate spatial-temporal patterns of T_{air} which can further our knowledge, on both the climate and terrestrial biological processes at regional and global scales (Benaliet al., 2012). Monitoring and understanding the trends of T_{air} and LST are crucial in the study of regional and global climate changes (Yoo et al., 2011). LST can be monitored and modelled from multiple daily satellite observations, such as the MODIS LST. Studies have shown that LST can be used for linear regression estimates of daily minimum and maximum T_{air} on a local scale (Mostovoy et al., 2006; Vancutsem et al., 2010; Zhang et al., 2011a; Yoo et al., 2011; Evrendilek et al., 2012; Benali et al., 2012; Zhu et al., 2013). Cresswell et al. (1999) found an over and underestimation of T_{air} during the day and at night, respectively, from Meteosat LST observations. They attempted to correct these errors and produce a proxy of T_{air} by applying a solar zenith angle correction on the Meteosat geostationary observations. They achieved an accuracy of 3°C for over 70% of the Meteosat temperatures. Similarly, Jin and Dickinson (2010) have studied the differences in the diurnal cycles of LST and T_{air} over a single site. Some studies (Florio et al., 2004) have

used several statistical approaches that combined a simple AVHRR Spilt-Window Technique(SWT) with ground meteorological station measurements in the prediction of T_{air}. Other studies (Wloczyk et al., 2011) have used the Landsat LST data to derive T_{air}. They have attempted to assign the satellite-derived T_{air} to a certain height above the ground and have investigated the possibility of a simple correction for reference height. They also considered the link between T_{air} spatial pattern and the window-size of the Landsat LST pixels. Xu et al. (2012) used four empirical regression models to estimate the relationship between T_{air} measurements and the MODIS-Aqua LST and found different relationships between the two different LC types in their study. They also assessed the effect of the MODIS LST window-size on the agreement between the two variables and found that spatial averaging over multiple pixels improves the accuracy of T_{air} estimates. Zaksek and Schroedter-Homscheidt (2009) reviewed the types of methods commonly used to estimate T_{air} based on LST, dividing them into three distinct groups which are explained in the introduction chapter but most of the previous researches have focused on daily estimations or instantaneous T_{air}. The TVX method has been widely used for T_{air} estimation. Czajkowski et al. (2000) estimated T_{avg} for a weekly period with associated RMSE between 1.72 and 3.48°C and R^2=0.64. Stisen et al. (2007) and Prihodko and Goward (1997) estimated T_{air} with RMSE higher than 2.5°C and R^2 between 0.64 and 0.86. Cresswell et al. (1999) used a statistical method to derive instantaneous T_{air} with an associated RMSE below 3°C for more than 70% of the sampled data. Zaksek and Schroedter-Homscheidt (2009) used a more sophisticated method, which was based on the energy balance to estimate instantaneous T_{air} with an RMSE of 2°C. Vancutsem et al. (2010) used 1 km MODIS data to estimate weekly T_{min} and T_{max}. They reported correlations between LST and T_{min} ranging from 0.01 to 0.96 for several stations and T_{max} was estimated with an R^2=0.92 and RMSE=1.83°C. Moreover, in previous studies, several variables were employed to estimate air temperature. For example, the variables used by Benali et al. (2012) included LST, Julian Day, elevation, and the distance to coast. Benali et al. (2012) used both weekly daytime LST data (LST_{day}) and night-time LST data (LST_{night}) to estimate the average, maximum and minimum weekly temperature. They found that there was a higher correlation between average weekly temperature and averaged weekly LST_{night}, which indicates the potential of LST_{night} in estimating averaged weekly temperature. The variables used by Kim and Han (2013) included LST, NDVI, altitude, and solar zenith angle. The variables used by Cristóbal, Ninyerola and Pons (2008) included LST, NDVI, and albedo. The variables used by Zakšek and Schroedter-Homscheidt (2009) included LST, NDVI, solar zenith, albedo, solar radiation, and altitude. After comprehensive consideration of these variables, twelve variables were selected as the predictors for the modelling of air temperature during the day and night-time: LST, NDVI, Julian day, latitude, longitude, Emissivity31, Emissivity32, altitude, albedo, wind speed, wind direction and air pressure. First, this research presents the comparison of state-of-the-art remote sensing-based LST data from MODIS with T_{air} for the six LC/LU. Within this study, we compared the relationship between T_{air} and the Four LST products of MODIS over Berlin. In the other word, the aim of this research is to analyse the agreement between LST from MODIS Terra and Aqua and T_{air} for the period of 2007 to 2013 based on different land cover classes, and then to investigate

the spatial variability of LST and T_{air} relationship by applying a varying window size on the MODIS LST grid based on different land cover classes. The comparison is done by using statistical parameters such as the correlation coefficient, the slope and the intercept with the y-axis of the regression line, mean bias error (MBE), and normalized mean bias also known as bias. The MBE is calculated by the difference between LST and T_{air} divided by the amount of observed time steps. If the MBE is positive, the LST detects warmer temperatures than the measured T_{air}, and vice versa (Hachem et al., 2012). Secondly, the main objective of this study was to develop a simplified parametrization model for estimating T_{air} during the day and night-time from MODIS LST products and auxiliary data over Berlin for different land cover types, using Adaptive neuro fuzzy system (ANFIS), artificial neural network (ANN) and support vector machine (SVR). The accuracy of these models were assessed by a comparison of the observed air temperature data from weather stations and the cross validation (CV) approach, in order to find the best model with high accuracy during the day and night-time. The errors associated with T_{air} estimation based on remote sensing data are often large and strongly limit its applicability (e.g. Czajkowski et al., 2000; Vazquez et al., 1997; Vogt et al., 1997). One of the objectives of this work is to provide T_{air} estimations with an accuracy, which will potentiate the future applications. Moreover, tuning the hyper parameters of some models like SVR and ANN were investigated. To select the hyper parameters of SVR, Simulated Annealing (SA) was applied and a multiple-layer feed-forward (MLF) neural networks with three layers and different nodes in hidden layers are used with Levenberg–Marquardt back-propagation (LM-BP) in order to achieve higher accuracy in the estimation of T_{air} during the day and night-time over six LC/LU.

4.2 Methods

4.2.1 The Relationship between Observed T_{air} and the Four LST Products over Berlin

The influence of the time of observation on the estimation of T_{air} has been studied and discussed in several studies, which resulted in different conclusions. Benali et al.(2012) stated that the use of both aqua LST_{day} and LST_{night} could improve the estimation of T_{day} and T_{night}(T_{day} and T_{night} are not the maximum and minimum temperature of a day and night-time) respectively, because the MODIS Aqua overpass time is closer to the time of both T_{day} and T_{night} than Terra's. In contrast, Zhu et al. (2013)showed that both terra LST_{day} and LST_{night} were better than aqua LST_{day} and LST_{night} for T_{air} estimations in Xiangride River basin of China. In another study, Mostovoy et al. (2006) found that the difference between the satellite overpass (Terra and Aqua) had little impact on the estimation accuracy of T_{air}.

4.2.2 Temporal matching of T_{air} to LST observations

The temporal frequency of the MODIS LST L3 product is four observations per day (Terra passes over the equator at approximately 10:30 am, 10:30 pm each day, Aqua satellite passes over the equator at approximately 1:30 pm and 1:30 am) in cloud free conditions, which are derived from a composite of several MODIS overpasses with different view angles (Wan 1999; Zhu et al., 2013). Depending on the local longitude (which results in changes in the sensor's viewing angle) and latitude, the local solar observation times at each pixel can vary up to 120 minutes or more over a repeating cycle (16 days) of the MODIS twin sensors (Figure 4.1). Other than that, overpass times do not follow a regular period during the day and over the sensor's repeat cycle. On the other hand, T_{air} data from the weather stations are provided at an hourly and by the minute frequency in Berlin in standard time (MEZ and UTC). This complicates the matching of the MODIS observation times with T_{air} time-series. To overcome this issue, for those stations that are in minute temporal resolution, we only need to convert from MEZ to UTC, but for other stations that are in hourly temporal resolution, a linear equation was considered for the synchronizing of T_{air} with LST form HDF file. For the creation of a data set using during the day and night-time for each LC/LU separately, we need to consider the overpass time of MODIS over Berlin. Another point is that the data has a significant number of missing points due to clouds in our study area, therefore, in order to increase the maximum of usable observations, as did Alcantara et al. (2013) in his research, the Terra and Aqua data were considered. As shown in figure 4.1, the LST day from Terra and aqua in descending and ascending orbit were considered as a daytime series, respectively and only Terra in ascending orbit was considered as night-time series because of higher correlation which was observed between LST and T_{air} in this time.

4.2.3 Day/Night analysis

Apart from spatial variations, the observation time can affect the relationship between LST and T_{air} time-series. To identify any variability in LST and T_{air} relationship in a diurnal basis, time-series of both variables were separated, based on the MODIS overpass times to produce four series over a single pixel window from MODIS-Terra and MODIS-Aqua day and night overpasses (four in total) were used in this analysis.

4.2.4 Statistical methods

A simple linear relationship is often assumed between LST and T_{air} in literature (Brunel 1989; Mostovoy et al., 2006). In view of this, a univariate linear regression analysis with the MODIS LST as the independent (or explanatory) and T_{air} as the dependent variable was applied to analyse LST and T_{air} relationship. The correlation coefficient, r, is considered as a quantitative measure to evaluate the strength of the agreement between LST and T_{air} time series in different steps of the analysis. Significance levels

(p-values) are reported in the results to express how unlikely the given r values would occur if no relationship between the explanatory and response variables did exist. The smaller the p-level is, the more significant the relationship. Moreover, two other statistical measurements such as RMSD and Normalized mean Bias (Bias), were considered as following:

$$Normalizedmeanbias = \frac{\sum_{i=1}^{n}(M-O)}{\sum_{i=1}^{n}(O)} \tag{4.1}$$

Where n is the numbers of data, M is LST value and O is the temperature.

4.3 Theory and Methodology

In this section, a brief overview of SVR, Simulated annealing (SA), ANN and ANFIS will be discussed. In addition, the theoretical concepts for selecting input parameters were explained.

4.3.1 Support Vector Regression

Support vector machine (SVM) is a very promising artificial intelligence method applied extensively for solving the classification problems. Support Vector Regression (SVR) method is derived from the SVM, which is a powerful technique to solve a non-linear regression problem, but it has received less attention, due to the fact that SVR algorithm is sensitive to users' defined free parameters. The involved hyper parameters of the SVR model consist of penalty parameter C, insensitive loss function parameter ϵ, and the parameter σ for kernel function. The penalty parameter C control the degree of punishing the samples whose errors go beyond the given value. The value of ϵ can enhance the generalization capability; with the increase of ϵ, the number of support vectors will decrease, and the algorithmic computation complexity will also reduce. The bandwidth σ of the kernel function has a great influence on the performance of the learning machine. Recently, a number of new algorithms like genetic algorithm, grid search optimizing, cross-validation and particle swarm optimization (PSO) have been proposed for the optimization of the SVR parameters (Sartakhti et al., 2011; Ustün et al., 2005; Wang et al., 2016; Chen and Wang (2007); Hu et al., 2010; Keerthi 2002; Ito and Nakano (2005))but in this study, Simulated Annealing (SA) as an optimization method, was considered because inappropriate parameters in SVR can lead to over-fitting or under fitting problems and they greatly influence the regression accuracy and computation complexity of SVR(Schölkopf and Smola 1998). The RBF was applied in the study, which has the ability to universally approximate any distribution in the feature space. With an appropriate parameter, RBF usually provides a better estimation performance. Moreover, the basic SVR concept is concisely described in the research of Cristianini and Taylor 2000; Smola and Schölkopf 2004; Ito and Nakano 2005; Keerthi 2002).

4.3.2 Simulated Annealing Optimization Method

Simulated annealing is a local search algorithm capable of escaping from local optima. Its ease of implementation and convergence properties and its use of hill climbing moves to escape local optima have made it a popular technique over the past two decades. Survey articles that provide a good overview of simulated annealing's theoretical development and domains of application include (Eglese 1990; Fleischer 1995; Henderson et al., 2003; Koulamas et al., 1994; Romeo et al., 1991; Anily and Federgruen (1987); Suman and Kumar (2006); Abramson et al., 1999; Ben-Ameur 2004; Aarts and Korst (1989); van Laarhoven and Aarts (1988); Aarts and Lenstra (1997)). This study proposed SA-based approach for parameter tuning in the SVR. For convenience, the SVR model with SA is referred to as a SA-SVR method. The idea is to find the parameters that minimize the generalization error of the algorithm at hand. This error can be estimated on some data which has not been used for learning. To achieve this aim, the three basic decision variables as mentioned before must be tuned in proper manner. We propose here a methodology for automatically tuning multiple parameters for the SVR. The process of SA-SVR algorithm approach is briefly summarized as follows:

Algorithm: Simulated Annealing Algorithm:

1 Step 1:

Solution space X

Object function F

Neighbourhood structure N

2 Step 2:

Current = An initial solution, among all possible state (X)

S-optimal=Current

T_0=INFINITY

T=T_0

Iteration=MAX-Iter

Epoch=1

Select temperature reduction function alpha, 0.8 ≤ alpha ≤ 0.99

3 Step 3: Repeat

Next= randomly selected from N (Current)

$\triangle F = F(Next) - F(Current)$

If $\triangle F > 0$ Then

Current = Next

Else

r=rand(0, 1)

if $r < e^{(\frac{-\triangle F}{T})}$ Then Current=Next

Until Epoch <= Iteration

$T = alpha * T$

If F [Current] < F [S-optimal] Then S-optimal =Current

Until stop condition is met

4 Step 4: Return S-optimal as an approximation to the global minimum solution

The proposed parameter values of SA-SVR approach were set as follows: Iteration = 200, T_0 was set to a sufficiently large number, while the set of hyper parameters are initialized in the given range as following:$C \in [0, 10000]$, $\sigma^2 \in [0, 2]$, and $\epsilon \in [0, 0.0001]$. Where optimization method (SA) is to seek the global optimal solutions. The best solution among these possible solutions is then selected as the optimal solution in the SA-SVR. According to research of Ustün and Melssen (2005), the general range of C, σ^2, and ϵ has been considered.

4.3.3 Adaptive Neuro-Fuzzy Inference System

An ANFIS is a combination of an adaptive ANN and a fuzzy inference system (FIS). The parameters of the FIS are determined by the neural network learning algorithms. Since this system is based on the FIS, reflecting amazing knowledge, an important aspect is that the system should be always interpretable in terms of fuzzy IF-THEN rules (Abraham. 2005). ANFIS can approximate any real continuous function on a compact set of parameters to any degree of accuracy (Jang et al., 1997; Jang and Shing (1991, 1993)). ANFIS identifies a set of parameters through a hybrid learning rule combining back propagation gradient descent error digestion and a least-squared error method. There are mainly two approaches for fuzzy inference systems, namely the approaches of Mamdani (Mamdani and Assilian, 1975) and Sugeno (Sugeno and Tanaka 1992; Takagi and Sugeno, 1985). The differences between the two approaches arise from the consequent part where Mamdani's approach uses fuzzy membership functions, while linear or constant functions are used in Sugeno's approach. The neuro-fuzzy model used in this study implements the Sugeno's fuzzy approach with input variables and air temperature values as output variable. More information about ANFIS theory, can be found in the study of Jang (1993) and Jang et al., (1997). Here, ANFIS was considered as a universal estimator (Jang, Sun and Mizutani 1997).

4.3.4 Artificial Neural Network

Artificial neural network models are universal approximations with the ability to generalize through learning non-linear relationships between provided variables of input(s) and output(s) (Hájek and Olej 2012). ANN are organized and interconnected collections of processing units (neurons or nodes), whose operation is analogue to a neural structure (Müller and Fill (2003)). ANN extract its computational power from its solid parallel distribution structure and ability to learn/generalize, allowing the resolution of complex propositions in many known areas (Haykin 2001). ANN execution is inspired on the human brain (Haykin 2001) and has been used in many applications with success. In agreement with Galvão et al. (1999), by the reason of its non-linear structure, the ANN can acquire more complex data characteristics, which are not always possible using traditional statistical techniques (Maier et al., 2010; Razavi and Tolson 2011). ANN is a robust computational technique which is primarily used for pattern recognition, classification, and prediction (Bose and Liang, 1996; Haykin 1999; Panchal et al., 2011). The use of ANNs in meteorological applications includes the prediction of ozone concentration, sulfur dioxide concentration, tornadoes, storms, solar radiation, carbon dioxide, pollutants, and monsoon rainfall (Gardner and Dorling 1998), monthly and year precipitation levels (Bodri and Cermak 2000), tide charts (Steidley et al., 2005), wave heights (Wedge et al., 2005), flash floods (Luk et al., 2000), and air temperature (Jain et al., 2003; Smith et al., 2006; Maqsood et al., 2004; Abdel-Aal 2004), estimation of dew point temperature (Mittal and Zhang 2003; Shank et al., 2008). Bilgili and Sahin (2010) used ANN for predicting long-term monthly temperature and rainfall in Turkey. Kisi and Shiri (2011) introduced new hybrid wavelet-AI models for precipitation forecasting. Smith et al. (2005) developed an enhanced ANN for air temperature prediction by including information on seasonality and modifying parameters of an existing ANN model.

4.3.4.1 Determining hidden node

Many researchers put their best effort in analysing the solution to the problem that how many neurons are kept in hidden layers in order to get the best results (Rivals I. and Personnaz L. 2000; F. Fnaiech.et al., 2001; Kortmann-Unbehauen 1988; Onoda 1995; Md.Islam and Murase (2001); Asthana and Bhujade (2011); Kazuhiro Shinike 2010; Doukim et al., 2010; Yuan et al., 2003; Wu and Hong 2010; Panchal et al., 2011; Hunter et al., 2012; Shuxiang et al., 2008; Ke and Liu (2008)), but unfortunately no one succeeded in finding the optimal formula for calculating the number of neurons that the neural network training time can be reduced and also accuracy in determining the target output can be increased. Usually some rule-of-thumb methods are used for determining the number of neurons in the hidden nodes.

- The number of hidden layer neurons are 2/3 (or 70% to 90%) of the size of the input layered. If this is insufficient then the number of output layer neurons can be added later on (Boger and Guterman 1997).

- The number of hidden layer neurons should be less than twice of the number of neurons in input layer (Berry and Linoff 1997).
- The size of the hidden layer neurons is between the input layer size and the output layer size (Blum 1992).

But the above three methods are not considered to be always true because not only the input layer and the output layer decide the size of the hidden layer neurons, but also the complexity of the activation function applied on the neurons, the neural network architecture, the training algorithm, and most important the training samples of the database on which the neural network is designed to execute. In this work, we decided to use the cross-validating approach in the 3-layers MLP in the following simulations, in order to select the number of hidden nodes in the second layer. The 3-layer MLP with an input layer, one hidden layer with non-linear transfer functions and an output layer with linear transfer functions and Back Propagation (BP as training algorithm) is considered. Moreover, in order to have a configuration that minimizes the RMSE in the test phase while keeping an eye on over fitting and the train set error.

4.3.5 Theoretical concepts for selecting input parameters

Physical processes control the spatial and temporal variation in temperature. LST is a function of incoming solar radiation, cooling factor by wind, land cover, temperature inversion and other effects at some regions in space and time. The temperature patterns differ between day and night-time while during the night-time temperature patterns are mainly determined by air humidity, land cover and proximity to water bodies and/or soil moisture (van Leeuwen et al., 2011). In urban and industrial areas, temperature is often locally higher due to heat emissions from industrial activities or heating (see e.g. Cheval and Dumitrescu 2009). Moreover, near-surface air temperature is driven more by land surface temperature than by direct solar radiation (Zaksek and Schroedter-Homscheidt 2009), which making LST an important variable for estimating T_{air}. Other parameters, such as vegetation cover, soil moisture, solar radiation, and albedo also have some influence on air temperature. In previous studies, several variables were employed to estimate air temperature. For example, Kim and Han (2013)applied LST, NDVI, altitude, and solar zenith angle. The variables used by Cristóbal, Ninyerola, and Pons (2008) included LST, NDVI, and albedo. Benali et al. (2012) used LST, Julian Day, elevation, and distance to coast. The variables used by Zakšek and Schroedter-Homscheidt (2009) included LST, NDVI, solar zenith, albedo, solar radiation, and altitude. After comprehensive consideration of these variables, twelve variables were selected as the predictors for modelling air temperature: LST, NDVI, latitude, longitude, altitude, albedo, wind speed, wind direction, emissivity31, emissivity32, relative humidity and Julian day. The reasons for selecting these variables as input to our model for estimating the air temperature are summarized as follows:

- The latitude, longitude and elevation were selected as an input parameter to model because the incoming solar radiation can be globally derived as a function of this factors. Moreover, latitude, longitudes and elevation are always the underlying effect relative to temperature (Zhao and Cheng, 2005, Samanta, et al., 2012, Stahl et al., 2006.

- Emissivity is important, because all objects at temperatures above absolute zero emit thermal radiation. However, for any particular wavelength and temperature the amount of thermal radiation emitted depends on the emissivity of the object's surface. Emissivity is defined as the ratio of the energy radiated from a material's surface to that radiated from a black-body (a perfect emitter) at the same temperature and wavelength and under the same viewing conditions. The emissivity of a surface depends not only on the material but also on the nature of the surface. The emissivity also depends on the temperature of the surface as well as wavelength and angle. Knowledge of surface emissivity is important both for accurate non-contact temperature measurement and for heat transfer calculations. Moreover, Surface emissivity is a measure of inherent efficiency of the surface in converting heat energy into radiant energy above the surface (Sobrino et al., 2001). Therefore, land surface emissivity is critical for determining the thermal radiation of the land surface (Caselles et al., 1995). The emissivity of a surface is controlled by some factors such as water content, chemical composition, structure, roughness, and the observation conditions (i.e. wavelength, pixel resolution and observation angle) (Snyder et al., 1998). For these reasons, in our study, due to considering six different LC/LU, the land surface emissivity also considered as an input parameter.

- LST is the radiative temperature of the land surface (Ghent et al., 2010)which is influenced by albedo, vegetation cover and soil moisture (Land Surface Temperature | Copernicus Global Land Service). The surface can include snow and ice, bare soil, grass, or the roofs of buildings (Land Surface Temperature: Global Maps, 2016. Available on-line). Near-surface air temperature is a measurement of the average kinetic energy of the air near the surface of the Earth (Near Surface Air Temperature—GES DISC-Goddard Earth Sciences Data and Information Services Center, 2016). Usually LST is measured by remote sensing whereas air temperature is measured 1–2 m above the ground. Near-surface air temperature is a consequence of complex effects of the turbulent heat transports produced by nearby heated surfaces (Unger, et al., 2009). The advantage of using MODIS LST is that, they account for small differences in temperature that are due to different land cover, moisture content which cannot model with constant physical parameters such as elevation, latitude, longitudes.

- The Julian day is proxies for the fraction of solar energy absorption during the day and emission during the night, influencing the diurnal amplitude of T_{air} throughout the year. The Julian day included the information of vegetation cover changes with seasons.

- The NDVI and Albedo reflect the seasonal variation of land cover.

- The Relative humidity (RH) is the ratio of the partial pressure of water vapour to the equilibrium vapor pressure of water at a given temperature. Relative humidity depends on temperature and the pressure of the system of interest. It requires less water vapour to attain high relative humidity at low temperatures; more water vapour is required to attain high relative humidity in warm or hot air. (Perry, R.H. and Green, D.W, Perry's Chemical Engineers' Handbook, McGraw-Hill).
- Moreover, Seasonal variation in some parameters such as relative humidity, wind speed, wind direction and air pressure contribute to explaining seasonal variation air temperature over six LC/LU.
- The MODIS LST can be used to improve spatial prediction of ground-measured values.

4.3.6 Assess predictive performance of models

In a real application, cross-validation is a model assessment technique (Allen 1974; Stone 1974; Geisser 1975) used to evaluate a machine learning algorithm's performance in making predictions on new datasets which has not been trained on. This is done by partitioning a dataset and using a subset to train the algorithm and the remaining data for testing. Because cross-validation does not use all of the data to build a model, it is a commonly used method to prevent over fitting during training. Each round of cross-validation involves, randomly partitioning the original dataset into a training set and a testing set. The training set is then used to train a supervised learning algorithm and the testing set is used to evaluate its performance. This process is repeated several times and the average cross-validation error is used as a performance indicator (Hastie et al., 2009; Yang 2007b). Common CV techniques include, k-fold, holdout, leave-out, repeated random sub-sampling, Stratify, Substituting. In this work, we apply K-fold CV (with k=4) techniques, to test how well our model can be trained by some data and then to estimate the data it hasn't seen before and then to select the best model.

4.3.7 Data normalization

Before computing, data of both input and output variables were normalized. In this study, data of all variables used were normalized into the range [0,1] with:

$$X_{\text{norm}} = \frac{(X_{\text{i}} - X_{\text{min}})}{(X_{\text{max}} - X_{\text{min}})} \tag{4.2}$$

where X_{norm} is the normalized value, X_i is the original value, and X_{max} and X_{min} are the maximum and minimum values out of the sample of X_i. This was due to the eliminating influence of different dimensions of data and to the avoidance of overflows of the model during calculations, as a result of very large or small weights towards

a maximization of model parsimony with considering computational effort. After the computation, output values were transformed back to the real prediction data.

4.3.8 Model calibration and validation

Cross-validation was used to evaluate the generalizability of a model for estimating the air temperature with the LST data. The observations were randomly divided into two parts. 70% of the observations were used for model calibration, and the rest were used as test dataset for model validation. The accuracy of the estimated air temperature obtained from three estimating models, ANFIS, NN and SA-SVR, have been assessed by a set of statistic measures, including: Root Mean-square Error (RMSE), coefficient of determination R-squared (R^2), Mean Bias Error (MBE) and Mean Absolute Error (MAE), respectively. The RMSE (was mainly used in the development process of the model and represents residual errors, which gives a global perspective of the differences between the observed and estimated values (Sousa et al. 2007; Zheng et al. 2013; Willmott et al. 2005). The RMSD is calculated similarly to RMSE. This goodness of fit criteria are expressed as following equations:

$$R = \frac{\sum_{i=1}^{M}(O_i - \overline{O})(S_i - \overline{S})}{\sqrt{\sum_{i=1}^{M}(O_i - \overline{O})^2}\sqrt{\sum_{i=1}^{M}(S_i - \overline{S})^2}} \tag{4.3}$$

$$MBE = \frac{1}{M}\sum_{i=1}^{M}(O_i - S_i) \tag{4.4}$$

$$MAE = \frac{1}{M}\sum_{i=1}^{M}|(O_i - S_i) \tag{4.5}$$

$$RMSE = \sqrt{\frac{1}{M}\sum_{i=1}^{M}(O_i - S_i)^2} \tag{4.6}$$

where, M is the total number of the observation data, O and S are the average of the observed and estimated T_{2m}, and O_i and S_i are the observed and estimated T_{2m} of the i^{th} data, respectively. In addition, graphical goodness-of-fit criteria such as quantile-quantile (Q-Q) diagram, bar plot of RMSE in train and test phases were applied for the comprehensive evaluation of simulation results. Although, the R^2 criteria is a measure of goodness-of-fit of the model and higher values are indicative that the predictive model fits the data in a better way. By definition, R^2 is the proportional measure of variance of one variable that can be predicted from the other variable. Thus, ideally the values of R^2 to approach one is always desirable. However, a high R^2 tells you that the curve came very close to the points, but in reality, it does not always indicate the model quality (Maddala 2001). In order to have a reliable statistical comparison between the models, both the MAE and RMSE can be used together to ascertain the

variation in errors in a given set of estimation. It should be noted that in MAE, all the individual errors have equal weight on the average, making it a linear score, but the RMSE has a quadratic error rule, where the errors are squared before being averaged. As a result, a relatively high weight is given to large errors. This could be useful when large errors are undesirable in a statistical model (Chai and Draxler (2014); Armstrong 2002).

4.4 Results and discussion

4.4.1 MODIS LST versus T_{air} time-series over a single pixel

Before analysing the effects of MODIS window size, the daily variability of LST and T_{air} relationship was examined by using separate LST series (over 1x1 window) (Diurnal differences). In this section, LST series used in this analysis is a composite time series which includes four daily LST observations (except for cloudy days) from both the MODIS Terra and Aqua day and night overpasses (approximately at 1:30, 10:30, 13:30, 22:30) supplied in the LST L3 product. The comparison between MODIS LST data and the T_{air} observations shows that LST_{day} and LST_{night} from both Terra and Aqua, with the mean relative bias above and under zero tended to overestimate T_{day} and underestimate T_{night} according to Tables(4.3 - 4.5)respectively which is in agreement with Cresswell et al. (1999) result. As shown in the table, a higher relative RMSD and bias values were seen for the Aqua $LST_{daytime}$ than the Terra $LST_{daytime}$ which might be given to the fact that more solar radiation has been received at the time of the Aqua MODIS overpass later in the day. Considering the scatter-plots of LST_{night} and T_{night} from Aqua for the industrial LC type, has higher scattering than daytime observations which are more spread around the 1:1 line (fig.4.5). This indicates the urban heat island (UHI) phenomena with RMSD=4.21°C. Both Aqua and Terra LST_{night} underestimated the T_{night} as well except for forest. Moreover, according to RMSD from Tables(4.3-4.5) and MODIS LST from Terra, a higher RMSDs is found for industrial and airport LC types during night time which indicates the UHI phenomena (with RMSD= 4.57°C and 4.32°C respectively). Moreover Tables(4.1-4.3) show that, correlations between the MODIS LST from Terra data are generally stronger from the daytime series compared with those from the night series, except for needle leaf trees. The needle leaf tree type showed more complex correlation patterns from day and night observations. The possible reason for this, is that the values of LST recorded by MODIS observation on this particular LC type is not exactly a representative of the skin temperature of the soil, but rather affected by the temperature near the top of the trees (canopy temperature). In addition, LST and T_{air} are correlated to a certain degree, with some drawbacks depending on factors, such as land cover type (Jin et al., 2010, Mildrexler et al., 2011). In general, figures (4.2-4.7) show that the time-series of the MODIS LST over six LU/ LC classes were correlated individually during the day and night-time. They are highly correlated with $R^2 > 0.80$. Moreover, figures (4.3-4.7)show that, during the warm months the LST_{day} is higher than T_{day} due to strong radiation,

while as expected during the cold months LST_{day} is lower than T_{day} for almost all LC/LU. Moreover, almost for all LC/LU, the LST_{night} is close to T_{night}. As due to long wave, radiation from surface LST and T_{air} at night are closer. Both the Terra and Aqua LST products were compared with the ground-based T_{air} as shown in the figure (4.8) and (4.9), the night-time LST datasets (MOD_{night} and MYD_{night}) and the observed T_{air} are more linearly concentrated along the fitting line than the daytime datasets. Strong correlations were observed between the night-time LST and T_{night} with minimal bias ($0.81< R^2 <0.89$, RMSE < 4.80 and MBE < 2.91°C). Specifically, the MYD_{night} tends to be more accurate for the estimation of T_{air} with lower intercepts, smaller RMSD and MBE than MOD_{night}. For T_{day}, the MYD_{day} had good agreement than MOD_{day} with lower intercept. This is most likely because the Aqua overpass time (1:30 the time when maximum temperature was recorded). However, LST from Aqua and Terra seems to be best for estimation of T_{day} among the LST products. To sum up, the relationship between LST and T_{air} may vary with time and location because the land surface energy balance is a complex phenomenon that depends on multiple factors (e.g., cloud cover, surface roughness, wind speed and soil moisture). In addition, the LST and T_{air} are different in principle. The satellite remotely sensed LST is a measure of the surface radiation. LST was calculated from the emissivity's surface, which is sensitive to LC, especially during daytime and another reason is the heat capacity or specific heat of LC. However, the specific heat varies significantly from one LC to another. The variation of the difference between LST_{day} and T_{day} may be due to the different heat capacities or specific heats of LC types. The heat capacity changes with temperature, which may result in different relations at the different times even over the same LC. Our results showed that the MODIS LST correlates best with T_{air} measurement during the daytime. To some extent, this outcome was contradictory to the other works in the literature (e.g., Zhang et al., 2011a; Benali et al., 2012) where they have reported a stronger correlation at night-time compared to daytime. It must be noted, though, that Benali et al. (2012) used MODIS-Terra but not MODIS-Aqua observations. Variations in the MODIS, overpasses time in its 16-day repeated cycle which enabled us to reconstruct the diurnal LST profile over a 7-year period. Although, many studies have shown a higher agreement between LST and T_{air} at night (Zhang et al. 2011a; Benali et al. 2012), this is not the case for all hours of the day or night. During some hours of the night the LST-T_{air} relationship is weaker than some hours during the day. These differences could be understood, as not only being the time of observation, but also geographical location affecting the relationship between LST product and T_{air} and therefore, affecting the accuracy estimation of T_{air} based on LST products.

4.4.2 Multiple LST window size

The relationship between the observed $T_{night} – T_{day}$ and LST is not limited to a single pixel, because the temperature of the near-surface air mass in a given area, is influenced by many factors such as energy exchanges with the land surface over a larger area. On the other hand, the T_{air} is impressed by both the local radiation

budget and air advection from the surrounding areas, thus, for better understanding of the spatial variability in LST-T_{air} relationship, a spatial window with a varying size is examined to discover the optimal spatial extent over which LST agrees best with the T_{air} measurements. To describe the effects of LST window-size on the LST and T_{air} relationship better, firstly, the time-series of LST from a single pixel (1x1 window size) overlapping each weather station were retrieved from the MODIS LST grid and then the LST of 3x3, 5x5, 7x7, 9x9, 11x11, 13x13 and 15x15 pixels were generated, respectively. Secondly, to determine the proper spatial window size for estimating air temperature, correlation coefficient analysis was made for different LC/LU. As shown in figure (4.10), the correlations were improved very slightly when the window size was increased from 1 to 3 pixels for daytime. The highest correlation values were achieved with 3x3 window for all LC/LU during the daytime and at the 1x1 during the night-time. Significance levels of all correlations were found to be at $\propto = 0.01$ which can be interpreted from p-values (all p-values <0.01). According to these results, the window size was selected for all LC/LU prior to model development for day and night-time data set.

4.4.3 Discussion

Three different methods namely SA-SVR, ANN and ANFIS were employed to estimate T_{air} during the day and night-time in Berlin by using the twelve variables as predictors. The performance of the three models was assessed using cross-validation with k=4 fold over different LC/LU during the day and night-time. All samples from each LC/LU were used in turn as the validation data set to test the model, while the remaining samples were used as the training data set to fit the model. RMSE, R^2, MBE and MAE were calculated from the measured and estimated T_{air} values to assess model performance. As shown in tables (4.4-4.9), ANN model with three layers structure, has higher adjusted R^2 value ranged from 0.93 to 0.97, RMSE ranged from 1.83°C to 2.53°C and MAE ranged from 1.53°C to 1.94°C in test phases for all LC/LU for estimating T_{day}. The results showed that all models have similar capability in the training phase for estimating T_{night} but the ANN has a higher adjusted R^2 which ranged from 0.89 to 0.93, RMSE and ranged from 2.13°C to 2.35°C and also MAE ranged from 1.54°C to 1.84°C values in the test phase in comparison to ANFIS and SA-SVR. The bar plots of RMSE for the three methods on testing data for each LC/LU are shown in figures (4.11) to (4.13), respectively. As shown in Tables(4.3-4.8)and figures(4.11) to (4.13), the three models SA-SVR, ANFIS and ANN have satisfactory been able to capture the relationship between the process variables. The bar plots depicted the performance of the ANN model on the testing data which was better than those of ANFIS and SA-SVR models for the whole of LC/LU during the day and night-time, but again we applied a CV approach to assess the model's performance in the test phase for the three mentioned models. As can be seen from figure 4.14 (b), the SA-SVR and ANN models are more robust and stable than ANFIS model regarding their SD values (ranged from 0.03 to 0.08) during the day and night-time and we can say that, these two models are more reliable than the ANFIS model. Moreover, in order to find the

optimum number of neurons in hidden layer, various numbers of neurons are used in the MLP and the optimum number of hidden neurons is determined using the CV approach to get the configuration that minimizes the RMSE in the test phase. Figure 4.14 (a) shows that after a certain number of hidden neurons are added, the model will start over fitting our data and give bad estimates on the test set. This indicates that over fitting starts to occur when the number of neurons is greater than 30, and in this point the model has lowest RMSE, and obviously we can conclude that the optimal number of hidden neurons should be 30, but if we consider the error bar which is the indicator of standard deviation, the less variation was observed at point 40, and then we can say that the model is more stable at this point as compared to point 30 (which is the number of neurons). Moreover, figure 4.15 shows Q-Q diagram of SA-SVR (left), ANFIS (middle) and ANN (right) models. Q-Q diagrams are often used to determine whether the model could extract the behaviour of the observed data (Chambers et al., 1983). As shown in figure 4.15, the models cannot estimate the high temperature for all LC/LU during the day and night-time. The weak performance of all models at high temperature are a consequence of a small numbers of data in these temperatures, and this is also highly related to the study area condition (Berlin) which has a short summer and has only a few numbers of high temperatures. In these cases, the learning algorithm of the three mentioned models have the tendency to underestimate the temperature. Therefore, the generalization of these models for the high temperature is reduced.

4.4.4 Conclusions

In this study, the comparison between the LST and T_{air} observations were done. The comparison shows that LST_{day} and LST_{night} from both Terra and Aqua, with the mean relative bias above and under zero tended to overestimate T_{day} and underestimate T_{night} respectively. In addition, a higher relative RMSD and bias values were seen for the Aqua $LST_{daytime}$ than the Terra $LST_{daytime}$which might be given the fact that more solar radiation has been received at the time of the Aqua MODIS overpass later in the day. The scatter-plots of LST_{night} and T_{night} from Aqua for industrial LC/LU has higher scattering than daytime observations which are more spread around the 1:1 line figure (4.4). This indicates UHI phenomena with RMSD =4.21°C. Moreover, according to RMSD from Tables(4.1-4.3) and MODIS LST from Terra, a higher RMSDs is found for industrial and airport LC/LU types during the night-time which indicated the UHI phenomena (with RMSD= 4.57°C and 4.32°C respectively). The results show that, the correlations between the MODIS LST from Terra data are generally stronger from the daytime series compared with those from the night-time series except for needle leaf trees. The needle leaf tree type showed a more complex correlation pattern from day and night observations. The reason is that the values of LST recorded by MODIS observation on this particular LC type is not exactly a representative of the skin temperature of the soil, but rather affected by the temperature near the top of the trees. In general, the results showed that the time-series of the MODIS LST over six LC/LU classes were correlated individually during the day and night-time. They are

highly correlated with r > 0.80. Moreover, for almost all LC/LU, the LST_{night} is close to T_{night}. Overall, the relationship between LST and T_{air} is varied with time and location because the land surface energy balance is a complex phenomenon that depends on multiple factors (e.g., cloud cover, surface roughness, wind speed and soil moisture). In the other words, The LST–T_{air} relationship is mainly controlled by the surface energy balance, but it also depends on factors that are closely linked to energy processes (Prince et al., 1998; Zhang et al., 2015). Moreover, in this chapter, the air temperature during the day and night-time in the period from 2007 to 2013 was estimated for the Berlin area over six LC/LU, using 1 km Aqua and Terra/MODIS data. The correlation coefficient between observed T_{air} and remotely sensed LST shows an increasing trend, with a spatial window size increasing from 1 km × 1 km to 3 km × 3 km, and subsequently decreasing slightly at window sizes larger than 3 km × 3 km for the daytime, but for the night-time this correlation coefficient between observed T_{air} and LST showed a decreasing trend, with spatial window size from 1 km × 1 km to 13 km × 13 km, and subsequently decreasing slightly at window sizes larger than 1 km × 1 km . These window sizes were therefore used to spatially average five satellite-derived environmental variables, (NDVI, Albedo, Emissivity31 and Emissivity32) which were used as predictors of T_{air} in the three models. In addition, a difficult task with ANN involves choosing the hidden nodes' number. Here, the ANN with one layer was used and the hidden nodes' number was determined using error and trials. For the ANFIS model, Gaussian membership function (MF) and 250 iterations were used. Different number of membership functions were tested and the best of which gave the minimum RMSE and was selected, which was 4 MFs for each variable. For the adjustment, the parameter in SVR model, the simulated annealing was applied. The ANN, ANFIs and SA-SVR models are compared in the test phase based on Tables(4.4-4.9). The ANN model, among the six LC/LU during the day and night-time performed better than the two other models with RMSE which ranged from 1.83°C to 2.53°C and from 2.13°C to 2.35°C during the day and night-time, respectively. The RMSE of SA-SVR model is ranged from 2.07 to 2.79°C during the day and night-time over six LC/LU and also the highest RMSE was observed in the ANFIS model with a range from 2.64°C to 3.70°C during the day and night-time over six LC/LU. These results indicated that the ANN model out performs the SA-SVR and ANFIS models for almost whole LC/LU during the day and night-time but based on figure (4.14) and the cross-validation results, the SA-SVR and ANN models out performs the ANFIS model. Moreover, the results showed that there was a high similarity between the training and testing tables, which demonstrates that the over-fitting has not been occurred in the SA-SVR, ANFIS and ANN. The Q-Q diagram of SA-SVR, ANFIS and ANN shows that all three models slightly tended to underestimate and overestimate the extreme and low temperature for all LC/LU during the day and night-time. The weak performance in the extreme and low temperature are a consequence of a small numbers of data in these temperatures. In these cases, the generalization of these models reduces for estimating the high and low temperature. In addition, despite moderate to high correlations between LST and T_{air}, LST cannot be directly used for estimating air temperature due to the large difference in MBE (Tables(4.1-4.3)), while by applying some additional parameters, in three models (Tables 4.4-4.9), It can be seen that the MBE was reduced notably, in all LC/LU during

day and night-time. Moreover, prediction of long-term monthly air temperature using ANFIS and ANN had been done in the study of Kisi and Shiri (2014). They applied station latitude, longitude and altitude values as input variable to predict the long-term monthly temperature values. They found that the ANN models generally performed better than the ANFIS model in the test period. The ANN models generally performed better than the ANFIS model in the test period and they found that for the ANN model, the maximum and minimum determination coefficient values were between 0.921 and 0.995. The maximum and minimum determination coefficient values were found as 0.99 and 0.876 for the ANFIS model in different stations. Testing results of the ANN and ANFIS models in the study of Kisi and Shiri (2014) show the RMSE values range from 0.1.53 to 4.20°C and 1.18°C to 9.25°C for each station, respectively. Furthermore, in the study of Xu. et al. (2014), they applied spatially averaged values of LST, NDVI, modified normalized difference water index (MNDWI), latitude, longitude, distance to ocean, altitude, albedo and solar radiation as predictors of T_{air} in linear regression and random forest models for estimating T_{air} in summer periods from 2003 to 2012. In their study, prior to model development, they also investigated the window size effect on the relationship between LST and T_{air}. The Cross-validation results of their study show that the random forest model (MAE = 2.02°C, R^2 = 0.74) outperforms the linear regression model (MAE = 2.41°C, R^2 = 0.64) and the distribution of residuals from the random forest model slightly overestimates T_{air}, with a mean residual value of 0.09°C. To sum up, in our study, instead of estimation monthly air temperature and only using the geographical input data, we estimate air temperature during day and night. Moreover, different parameters such as NDVI, Albedo, relative humidity, wind speed, wind direction and Julian day have been taking into consideration, which are representative of seasonal changes. The satisfactory results suggested that this modelling approach is appropriate for estimating air temperature in Berlin over six different LC/LU. In addition, the results indicate that MODIS time series of LST can be successfully combined with ground measurements of temperature to produce accurate and more detailed predications of temperature during day and night time. Although the air temperature estimated from satellites tends to be higher than ground-based measurement, the use of satellite remote sensing data can help to overcome the spatial problem of estimating T_{air} particularly in areas with low station density using satellite-based land surface temperature estimation and ground-based relationship between LST and air temperature. To reduce the biases in satellite-estimated air temperature, it can be effective to use retrieval method based on land surface heat budget (e.g. Kato and Yamaguchi, 2005) in future work.

Table 4.1: Statistical analyses between MODIS LST products and T_{air} observation from automatic meteorological stations. MOD_{day}, MOD_{night} , MYD_{day} and MYD_{night} are representative of MOD11A1 LST_{day}, MOD11A1 LST_{night}, MYD11A1 LST_{day} and MYD11A1 LST_{night} from Terra and Aqua respectively for urban and industrial LCT.

Dataset	**Urban**				**Industrial**			
	R^2	RMSD	MBE	Bias	R^2	RMSD	MBE	Bias
MOD_{day},T_{day}	0.88	3.72	0.256	0.015	0.86	3.69	-0.585	-0.035
MOD_{night},T_{night}	0.87	3.63	-1.896	-0.212	0.80	4.57	-2.651	0.288
MYD_{day},T_{day}	0.87	4.22	1.574	0.093	0.86	3.81	0.239	0.013
MYD_{night},T_{night}	0.88	2.94	-1.500	-0.217	0.80	4.21	-2.553	-0.356

Table 4.2: Statistical analyses between MODIS LST products and T_{air} observation from automatic meteorological stations. MOD_{day}, MOD_{night}, MYD_{day} and MYD_{night} are representative of MOD11A1 LST_{day}, MOD11A1 LST_{night}, MYD11A1 LST_{day} and MYD11A1 LST_{night} from Terra and Aqua respectively for agriculture and needle leaf trees LCT

Dataset	**Agriculture**				**Needleleaftrees**			
	R^2	RMSD	MBE	Bias	R^2	RMSD	MBE	Bias
MOD_{day},T_{day}	0.91	2.89	-1.420	-0.084	0.85	4.21	0.490	0.028
MOD_{night},T_{night}	0.85	3.23	-0.820	-0.117	0.87	3.61	-1.935	-0.204
MYD_{day},T_{day}	0.92	3.05	-0.222	-0.012	0.83	4.41	1.111	0.060
MYD_{night},T_{night}	0.85	2.83	-0.431	-0.090	0.87	3.15	-1.590	-0.213

Table 4.3: Statistical analyses between MODIS LST products and T_{air} observation from automatic meteorological stations. MOD_{day}, MOD_{night}, MYD_{day} and MYD_{night} are representative of MOD11A1 LST_{day}, MOD11A1 LST_{night}, MYD11A1 LST_{day} and MYD11A1 LST_{night} from Terra and Aqua respectively for Airport and Forest LCT

Dataset	**Airport**				**Forest**			
	R^2	RMSD	MBE	Bias	R^2	RMSD	MBE	Bias
$MODD_{day}$,T_{day}	0.90	3.99	1.656	0.100	0.89	4.43	2.460	0.132
MOD_{night},T_{night}	0.86	4.32	-2.925	-0.313	0.83	3.69	0.734	0.089
MYD_{day},T_{day}	0.89	4.80	2.721	0.158	0.88	4.82	3.543	0.178
MYD_{night},T_{night}	0.86	3.71	-2.413	-0.329	0.88	2.85	1.157	0.181

Table 4.4: statistic indices between estimated T_{day} values obtained by SA-SVR and measured value from meteorological station over six LCT in test phase.

LCT	RMSE	MAE	MBE	R^2
Agriculture	2.62	2	0.13	0.92
Forest	2.31	1.84	0.06	0.91
Industrial	2.79	2.12	0.09	0.91
Urban	2.46	1.89	0.13	0.92
Airport	2.41	1.87	0.02	0.92
Needleleaf trees	2.42	1.85	0.09	0.93

Table 4.5: statistic indices between estimated T_{night} values obtained by SA-SVR and measured value from meteorological station over six LCT in test phase.

LCT	RMSE	MAE	MBE	R^2
Agriculture	2.62	1.87	0.26	0.88
Forest	2.42	1.67	0.20	0.88
Industrial	2.54	1.84	0.16	0.89
Urban	2.56	1.87	0.16	0.89
Airport	2.07	1.49	0.16	0.92
Needleleaf trees	2.28	1.61	0.18	0.91

Table 4.6: statistic indices between estimated T_{day} values obtained by ANFIS and measured value from meteorological station over six LCT in test phase.

LCT	RMSE	MAE	MBE	R^2
Agriculture	2.85	2.21	0.12	0.91
Forest	2.64	2.08	0.69	0.90
Industrial	3.70	2.78	0.17	0.88
Urban	2.74	2.03	0.35	0.90
Airport	2.75	2.08	0.36	0.90
Needleleaf trees	2.64	2.06	0.24	0.90

Table 4.7: statistic indices between estimated T_{night} values obtained by ANFIS and measured value from meteorological station over six LCT in test phase.

LCT	RMSE	MAE	MBE	R^2
Agriculture	3.15	2.40	-0.06	0.84
Forest	1.98	1.45	-0.04	0.92
Industrial	2.68	1.91	0.19	0.88
Urban	2.38	1.72	-0.18	0.90
Airport	2.33	1.69	-0.16	0.90
Needleleaf trees	2.28	1.65	0.50	0.92

Table 4.8: statistic indices between estimated T_{day} values obtained by ANN and measured value from meteorological station over six LCT in test phase.

LCT	RMSE	MAE	MBE	R^2
Agriculture	2.28	1.82	0.05	0.97
Forest	1.83	1.54	-0.29	0.97
Industrial	2.53	1.94	-0.102	0.93
Urban	2.13	1.62	-0.07	0.97
Airport	2.14	1.69	-0.06	0.95
Needleleaf trees	2.08	1.57	0.14	0.95

Table 4.9: statistic indices between estimated T_{night} values obtained by ANN and measured value from meteorological station over six LCT in test phase.

LCT	RMSE	MAE	MBE	R^2
Agriculture	2.15	1.59	0.16	0.90
Forest	2.15	1.54	-0.01	0.89
Industrial	2.34	1.73	0.03	0.91
Urban	2.35	1.81	-0.06	0.93
Airport	2.35	1.84	0.10	0.92
Needleleaf trees	2.13	1.55	0.06	0.92

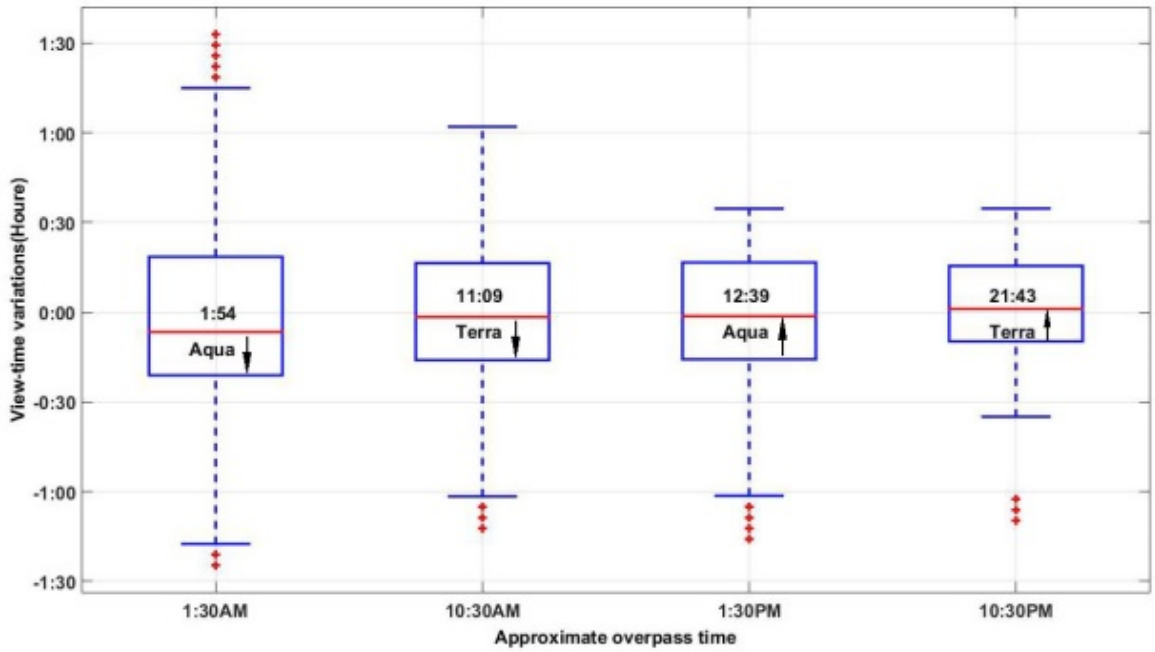

Figure 4.1: Average viewing times (local solar) and overpass nodes (shown as labels and arrows), maximum variations from the mean observation times (in hour shown by lower and upper caps of whiskers), median times (middle line), lower (25th) and upper (75th) quartiles of all observation times (lower and upper edges of boxes) of four overpasses of MODIS (on-board Terra and Aqua, two overpasses each) over the study area of 7 years (2007 to 2013). The mean local solar observation time of each overpass is subtracted from the series (scaled to zero) but is labelled on each box.

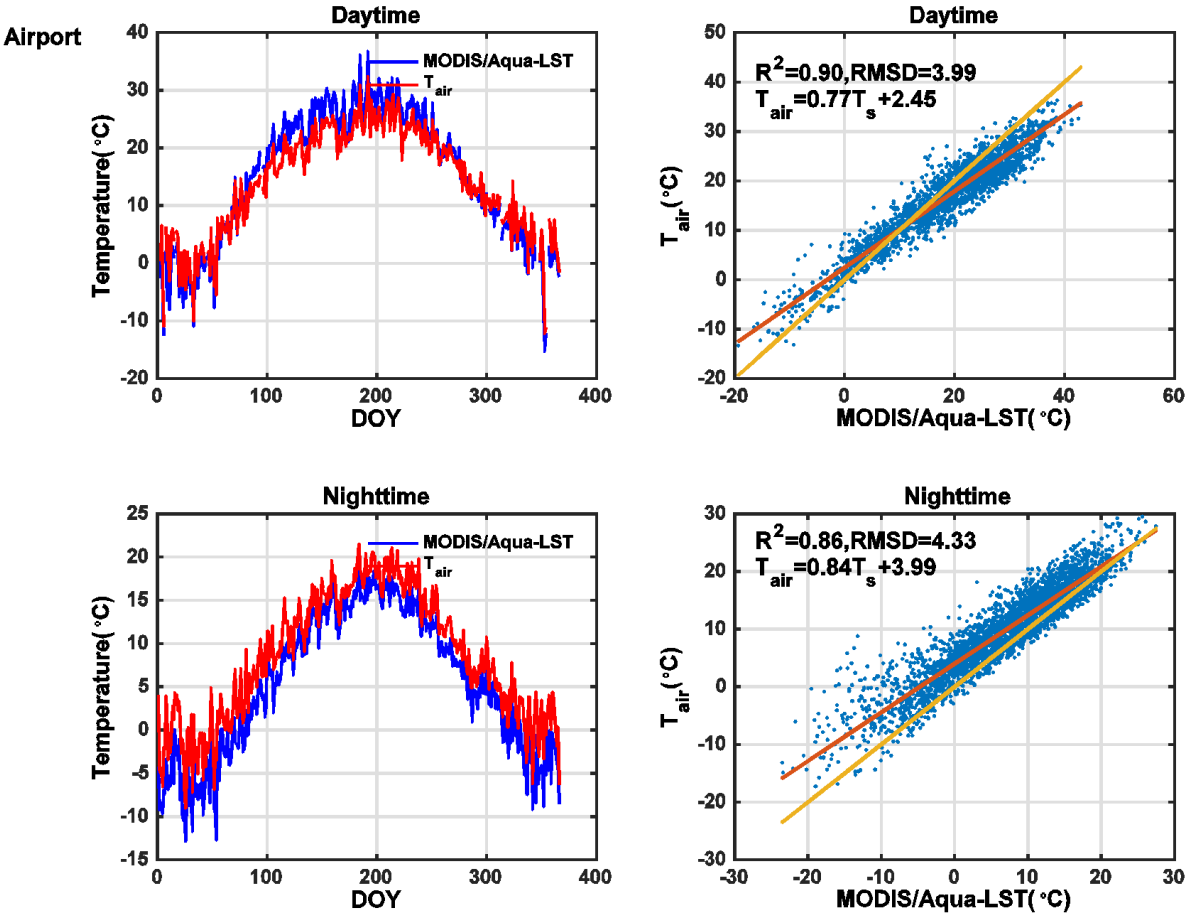

Figure 4.2: Correlations between LST and T_{air} time-series separated based on approximate overpass times of MODIS-Aqua, where each scatter-plot shows MODIS-Aqua daytime (right-up) and MODIS-Aqua night-time (right-down) and also MODIS-Aqua day and night-time observations (left up and down plots) plotted against T_{air} measurements at the corresponding times for airport LCT with P-value<0.01.

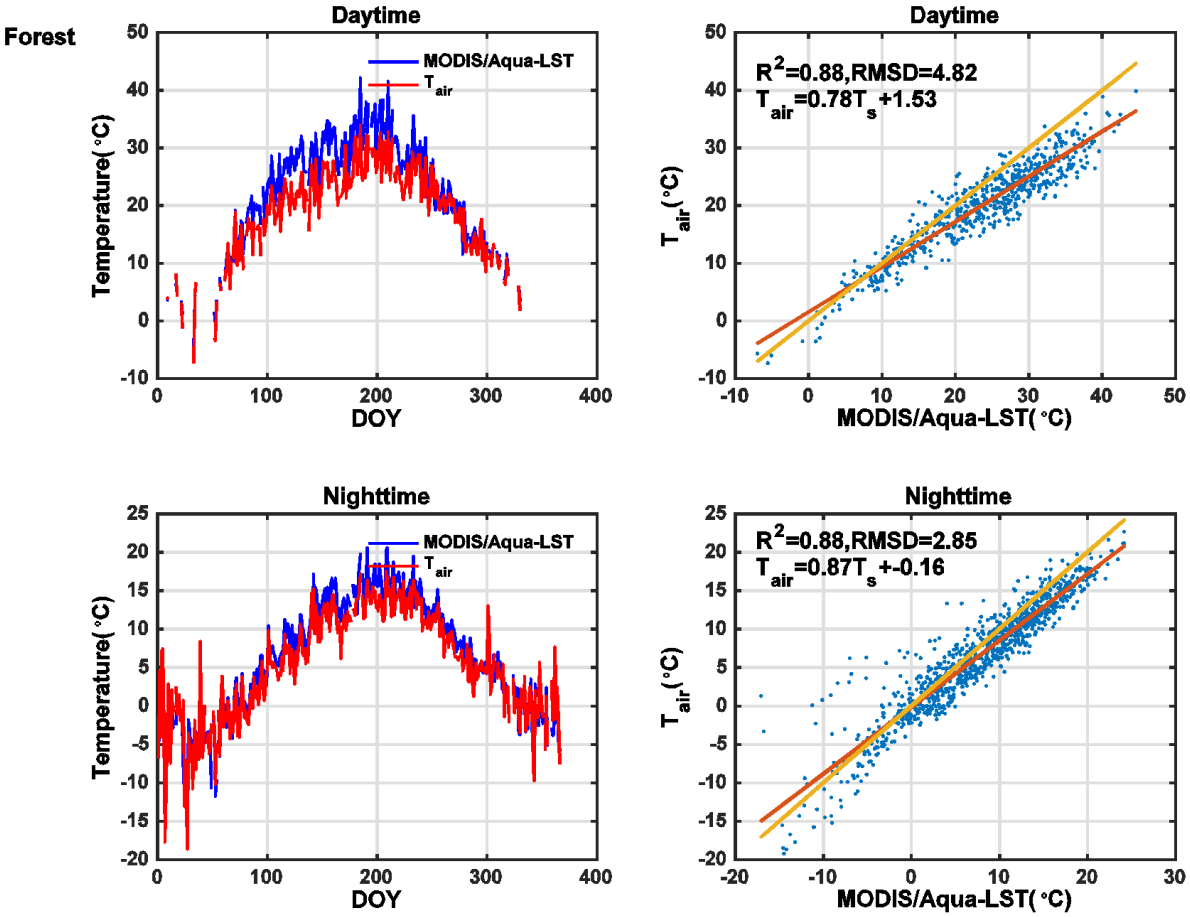

Figure 4.3: Correlations between LST and T_{air} time-series separated based on approximate overpass times of MODIS-Aqua, where each scatter-plot shows MODIS-Aqua daytime (right-up) and MODIS-Aqua night-time (right-down) and also MODIS-Aqua day and night-time observations (left up and down plots) plotted against T_{air} measurements at the corresponding times for forest LCT with P-value<0.01.

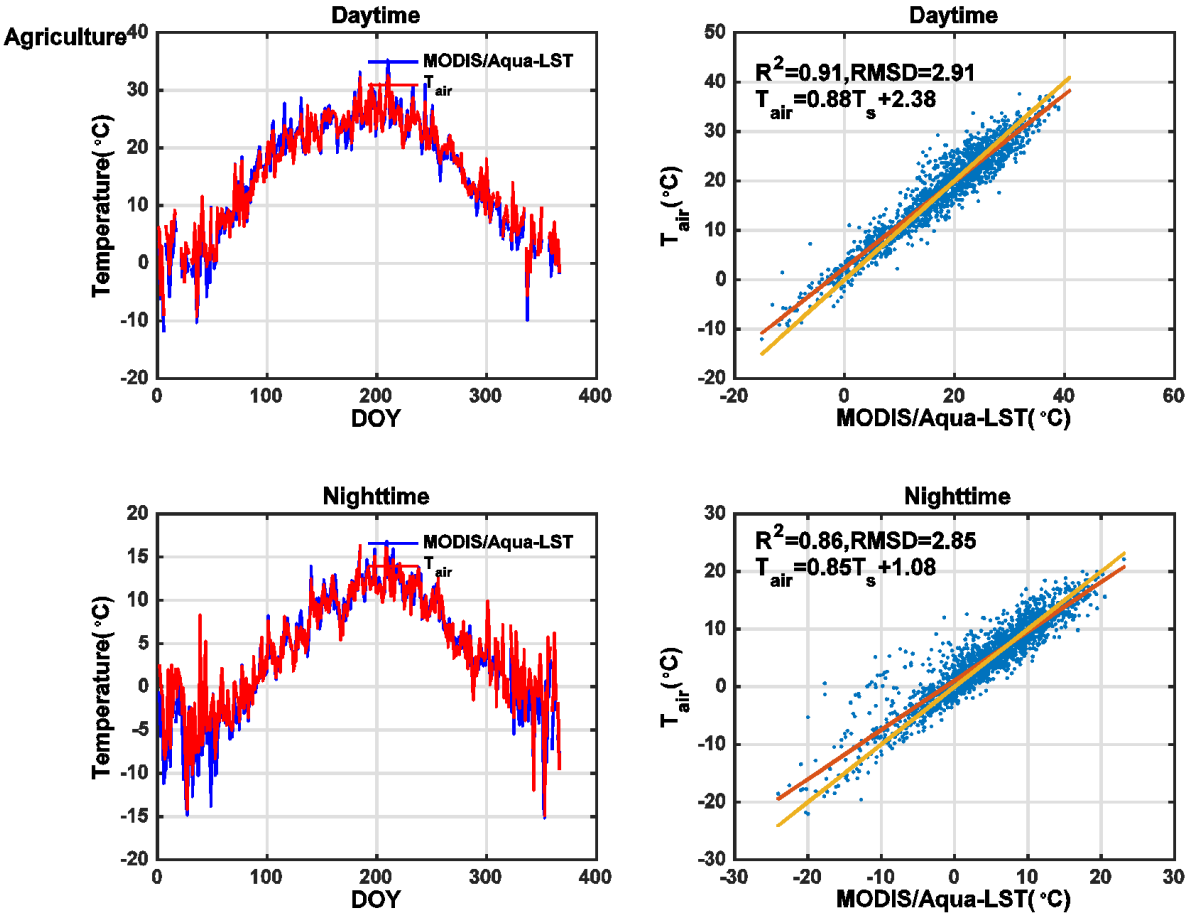

Figure 4.4: Correlations between LST and T_{air} time-series separated based on approximate overpass times of MODIS-Aqua, where each scatter-plot shows MODIS-Aqua daytime (right-up) and MODIS-Aqua night-time (right-down) and also MODIS-Aqua day and night-time observations (left up and down plots) plotted against T_{air} measurements at the corresponding times for agriculture LCT with P-value<0.01.

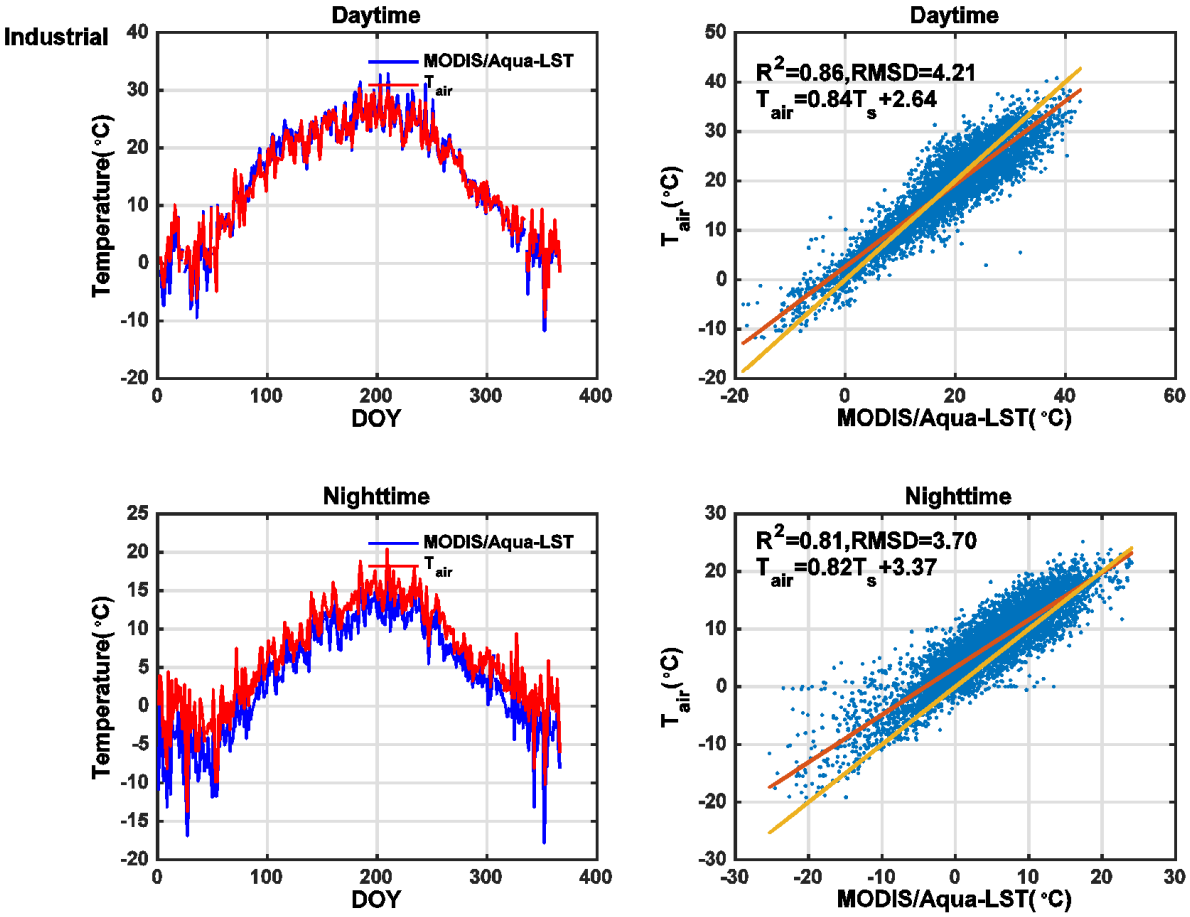

Figure 4.5: Correlations between LST and T_{air} time-series separated based on approximate overpass times of MODIS-Aqua, where each scatter-plot shows MODIS-Aqua daytime (right-up) and MODIS-Aqua night time (right-down) and also MODIS-Aqua day and night-time observations (left up and down plots) plotted against T_{air} measurements at the corresponding times for industrial LCT with P-value<0.01.

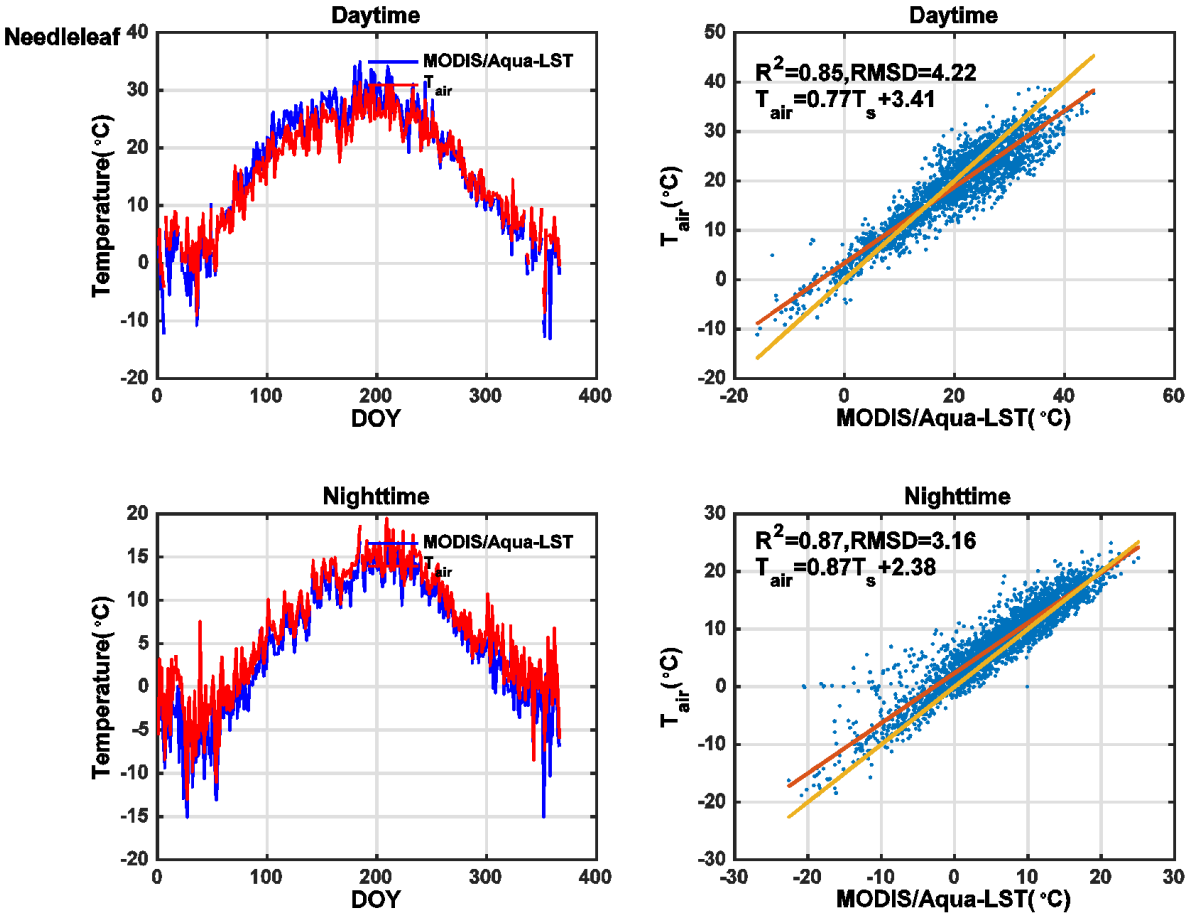

Figure 4.6: Correlations between LST and T_{air} time-series separated based on approximate overpass times of MODIS-Aqua, where each scatter-plot shows MODIS-Aqua daytime (right-up) and MODIS-Aqua night-time (right-down) and also MODIS-Aqua day and night-time observations (left up and down plots) plotted against T_{air} measurements at the corresponding times for needle leaf trees LCT with P-value <0.01.

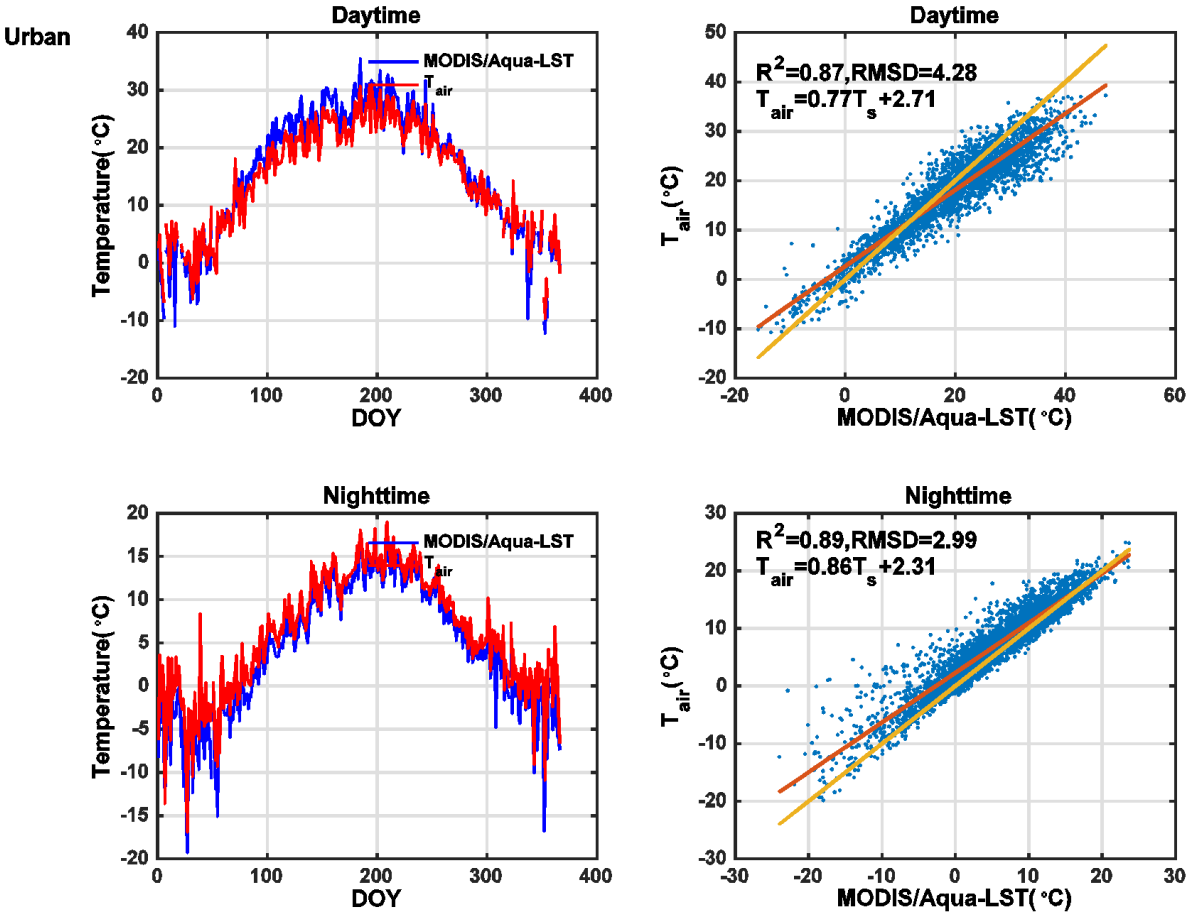

Figure 4.7: Correlations between LST and T_{air} time-series separated based on approximate overpass times of MODIS-Aqua, where each scatter-plot shows MODIS-Aqua daytime (right-up) and MODIS-Aqua night time (right-down) and also MODIS-Aqua day and night-time observations (left up and down plots) plotted against T_{air} measurements at the corresponding times for urban LCT with P-value<0.01.

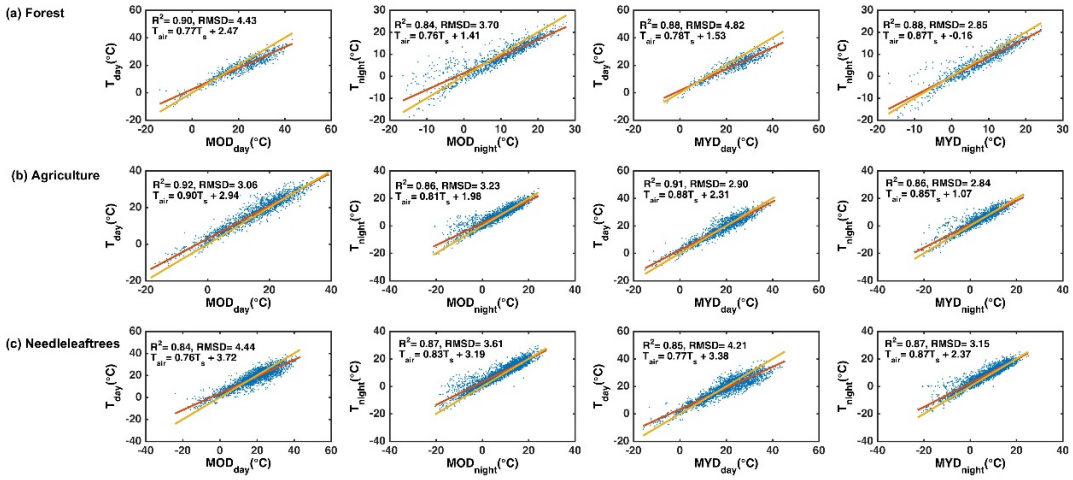

Figure 4.8: Scatter plots between observed T_{air} (T_{day} and T_{night}) and LST from Four MODIS products (MOD_{day},MOD_{night}, MYD_{day}, MYD_{night}) which are represented for three different LC/LU. T_s is land surface temperature.

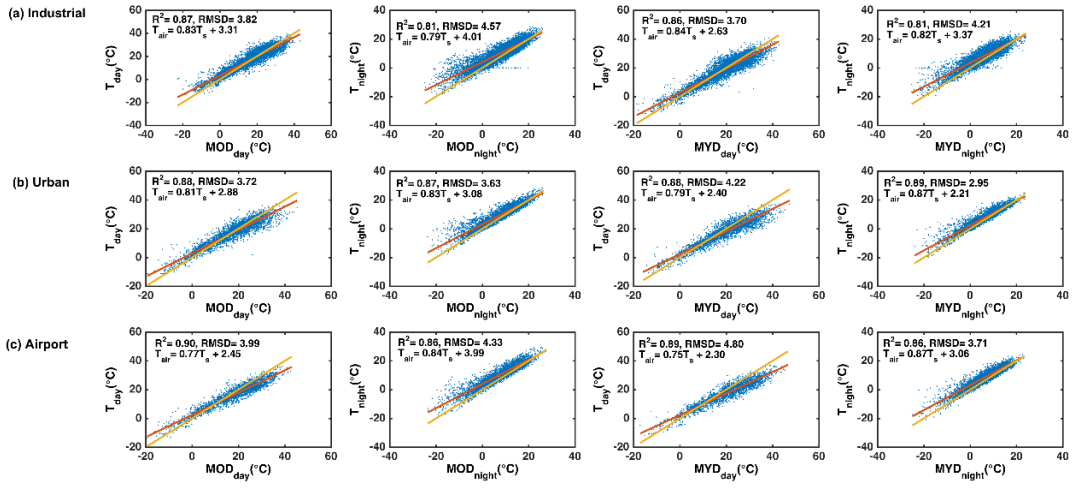

Figure 4.9: Scatter plots between observed T_{air} (T_{day} and T_{night})and LST from Four MODIS products (MOD_{day},MOD_{night}, MYD_{day}, MYD_{night}) which are represented for three different LC/LU. T_s is land surface temperature.

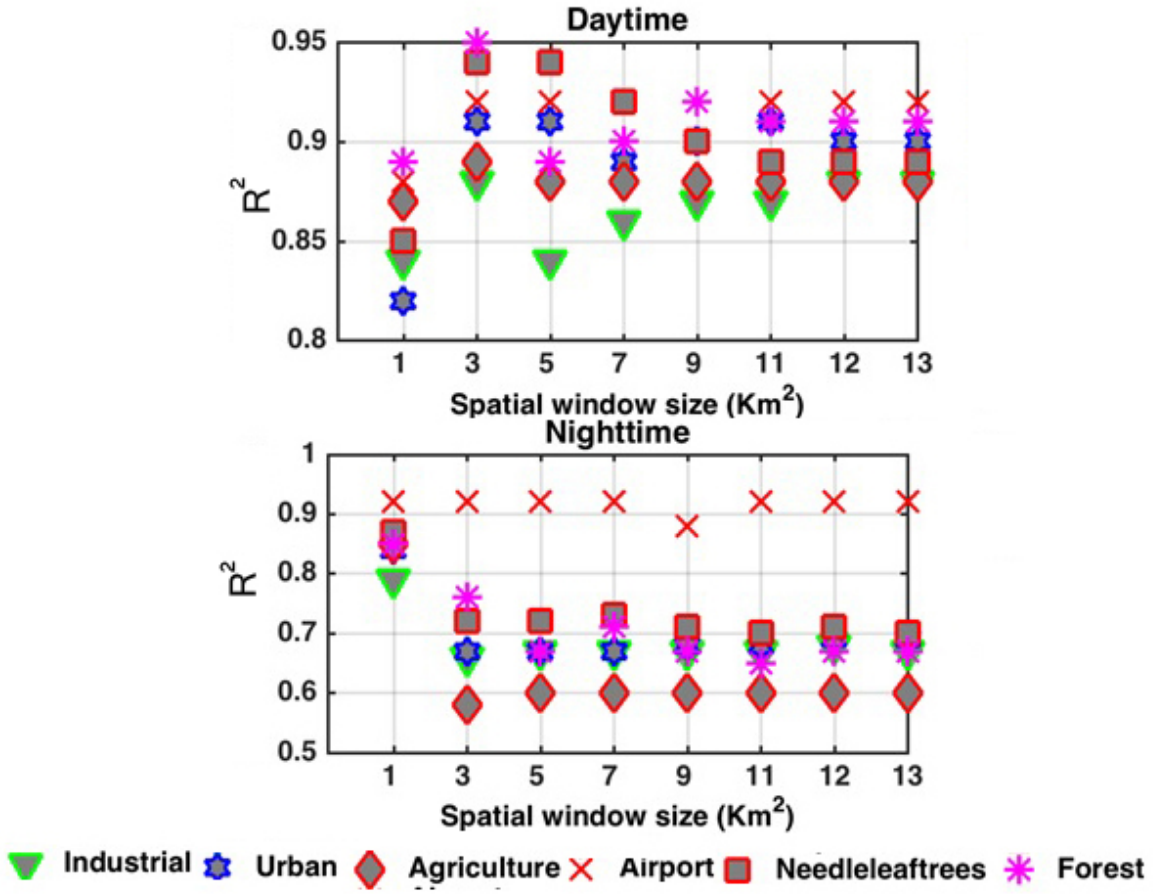

Figure 4.10: Variation of the correlation coefficient between T_{day}, T_{night} and LST with the varying spatial window size over six LCT.

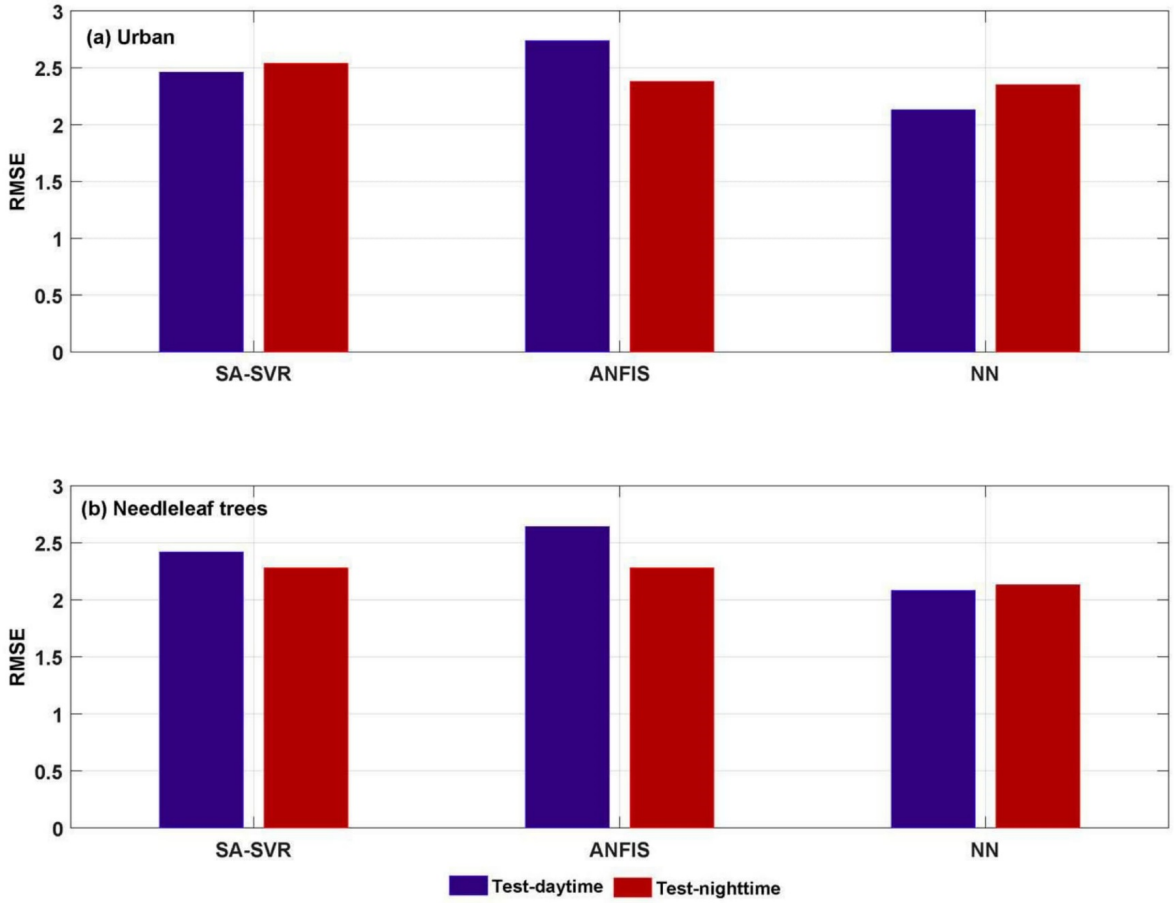

Figure 4.11: Bar plot of estimated T_{2m} versus measured temperature during day and night-time using SA-SVR, ANFIS and NN for (a) urban and (b) Needle leaf trees LCT.

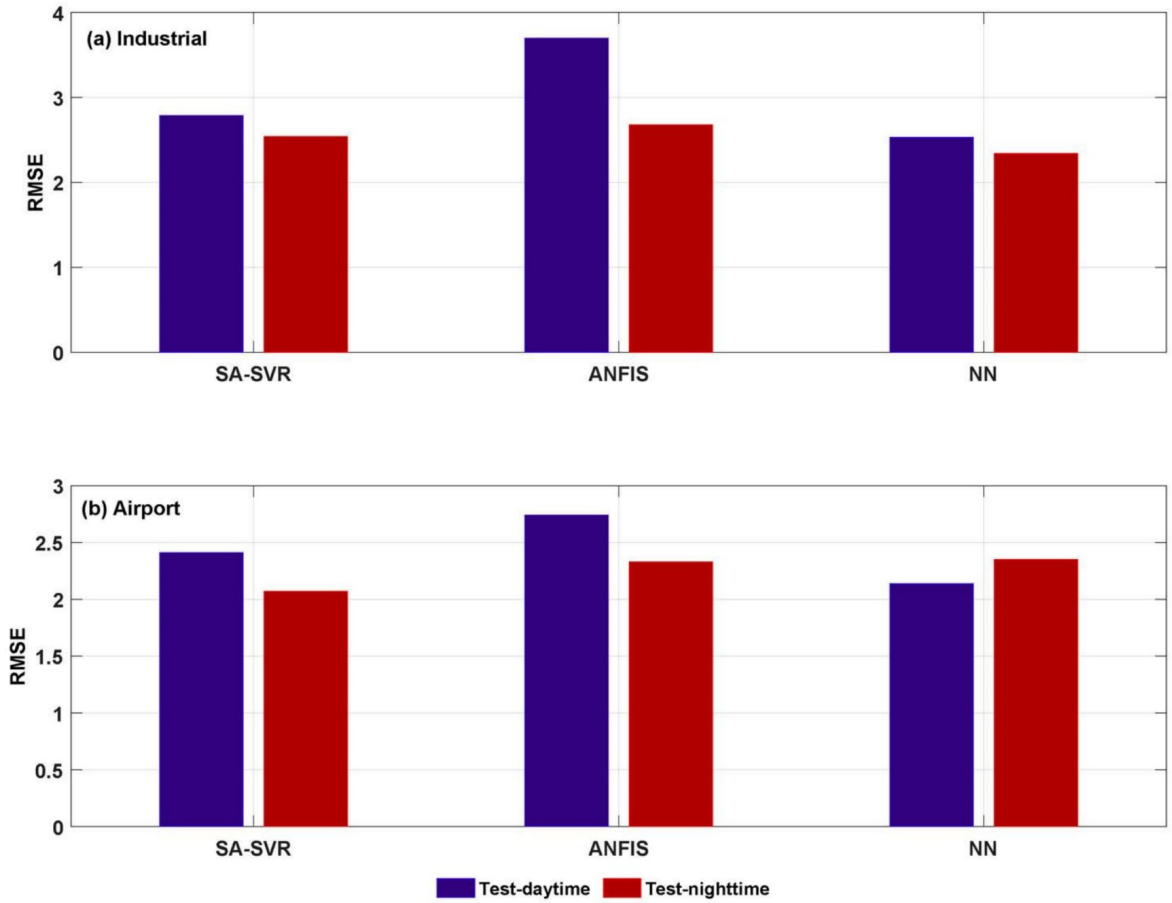

Figure 4.12: Bar plot of estimated T_{2m} versus measured temperature during day and night-time using SA-SVR, ANFIS and NN for (a) Industrial and (b) Airport LCT.

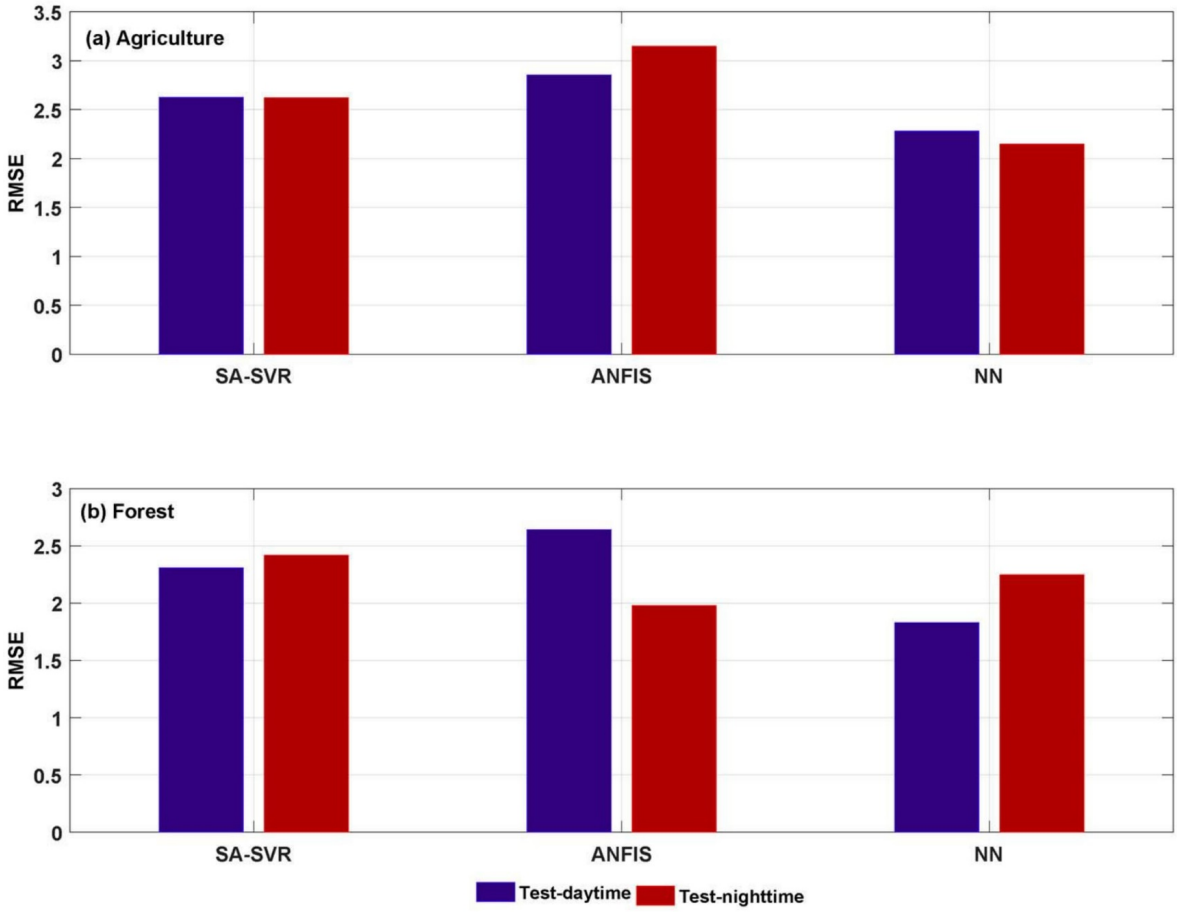

Figure 4.13: Bar plot of estimated T_{2m} versus measured temperature during day and night-time using SA-SVR, ANFIS and NN for (a) Agriculture and (b) Forest LCT.

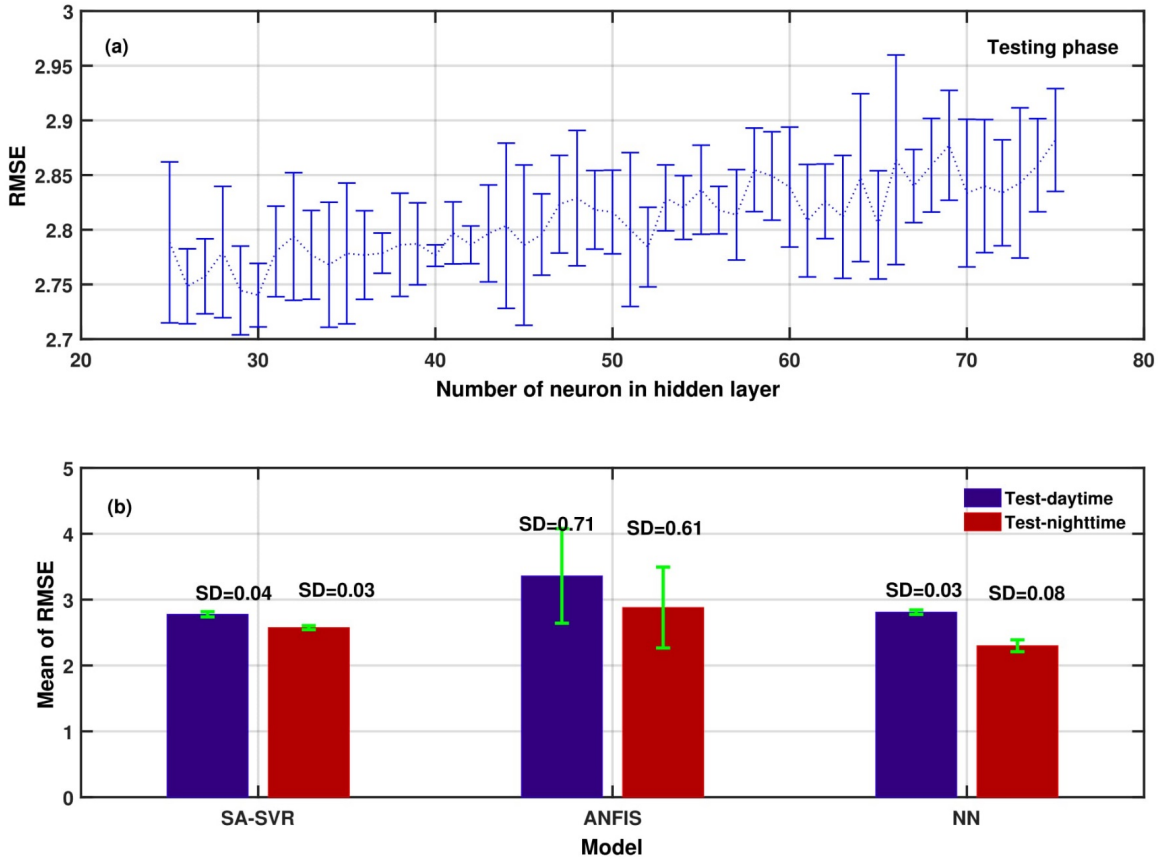

Figure 4.14: These two sub-plots show the effect of K-fold cross validation (with k=4) in three models. In sub-plot (a) x and y-axes show the average of cross validation error (RMSE) and number of nodes in hidden layer in testing phase respectively. In sub-plot (b), x and y-axes show the type of model and the average of cross validation error (RMSE)in three models respectively.

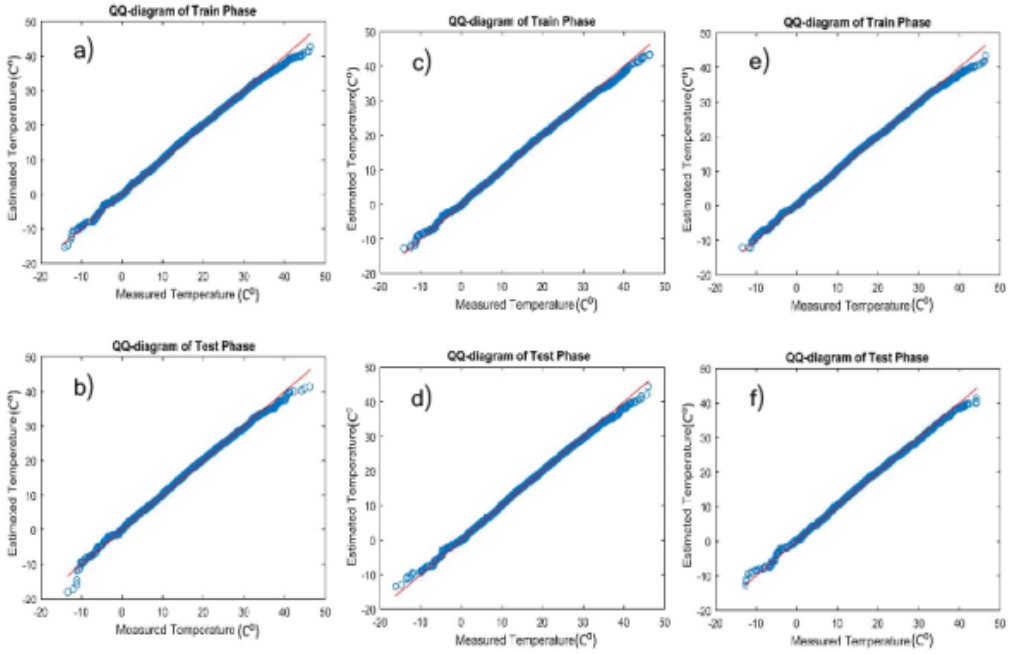

Figure 4.15: Q-Q diagram of estimated T_{2m} versus measured temperature during daytime for Industrial LCT using SA-SVR, ANFIS and ANN in testing phase.

5 Feature Selection for Estimating of near Surface Air Temperature from MODIS over Different LC/LU

5.1 Introduction

Nowadays, due to data abundance in datasets, development of algorithms are needed for discovering meaningful information. Depending on the data mining tasks, data models are made in the areas of clustering, regression and classification. For two main reasons, the pre-processing of the datasets must be done:

- Dimension reduction in dataset in order to achieve more effective analysis
- Adaptation of the dataset to best fit the selected analysis method.

Moreover, dataset size reduction can be done in one of the two ways:

- Feature set reduction
- Sample set reduction

The focus of this chapter is on feature reduction which is an important issue in data mining, because a model over-fitting is consequence of a high number of parameters in a dataset, comparable to or higher than the number of samples, which in turn cause to poor results on the validation datasets. Furthermore, building a model from datasets with many parameters is needed more computational time (Korn et al., 2011). All this reasons motivate researchers to make an effort to work on different methods for feature set reduction. (Korn et al., 2011).The reduction is done via:

- Feature selection
- Feature extraction

Feature extraction methods transform the original features into a new feature set which are made from the original one that depend on their combinations, with the goal of discovering more important information in the new set (such as Linear Discriminant Analysis (LDA), Multidimensional Scaling, and Principal Component Analysis (PCA)(Tang et al. 2014). The new feature set can be reduced by considerings characteristics such as dataset variance coverage. In this work, we focus on feature selection and provide an overview of the existing methods that are available for handling several different classes of problems. Generally, a feature is usually categorized as:

- Strongly relevant
- Weakly relevant, but not redundant-irrelevant
- Irrelevant

- Redundant (Yu and H. Liu 2004, Alelyani et al., 2013

In order to have an optimal feature subset, a strongly relevant feature is always essential. Depend on specific conditions, the weakly relevant feature may be considered but Irrelevant features are not essential to take into consideration. Redundant features are those that are weakly relevant but can be completely used as an alternative to a set of other features such that the target is not disturbed at all. The goal of feature selection is to find a feature subset including of only relevant features in order to maximize relevance and minimize redundancy(Yu and H. Liu 2004). The main objective of this research is doing feature selection for estimating T_{air} during day and night time over different LC/LU using both Terra and Aqua MODIS LST products (daytime and night-time) and auxiliary data from 2007–2013. Two input variable selection methods were applied because predictor selection is an essential step in environmental, biological, industrial and climatological applications. Feature Selection helps in understanding data, reducing computation requirement, reducing the effect of curse of dimensionality and improving the predictor performance. Through input variable selection to eliminate the irrelevant or redundant variables, a suitable subset of variables is identified. Meanwhile, the complexity of the model structure is simplified and the computational efficiency is improved. This work describes the procedures of the input variable selection for estimating of T_{air} during day and night-time from MODIS over six LC/LU types in order to figure out which parameters, among 12 candidate parameters can described the relationship between LST and T_{air} and has important effect on their relationship. Therefore, the brute-force search or exhaustive search (ES) and greedy best first search using artificial neural network (ANN) were applied in this study. In addition, there are several research's which estimated T_{air} using statistic approaches. Daily minimum and maximum air temperature are estimated in the research of Shen and Leptoukh.,2011 which considered the central and eastern Eurasia as their case study. They had reached out to the accuracy of 2.4–3.2°C and 3.0°C (MAE) for minimum and maximum air temperature respectively. T_{air} estimation was done by Vancutsem et al.2010 on Africa and the accuracy of their model was 2.1–2.76°C (RMSE). In the study of Xu et al., maximum temperature was estimated for Western Canada with accuracy of 2.02 (MAE). Zeng et al.2015 was also used the statical approach for estimating the maximum and minimum air temperature on Corn Belt over U.S and their model was achieved the accuracy of 2.15–4.27°C and 1.75–5.13°C (RMSE) for daily minimum and maximum air temperature respectively. Zhu et al.2013 applied the TVX method for estimating the daily minimum and maximum air temperature on Xiangride River Basin of China region and the accuracy of 2.97 and 3.79°C (RMSE) were achieved by them for minimum and maximum air temperature respectively. The M5 model tree method was applied by Emamifar et al.2016 on the south-west of Iran and they estimated the daily mean air temperature with the accuracy of 2.3°C (RMSE). However, most of these studies have only used LST daytime and LST night-time solely for T_{air} maximum and T_{air} minimum estimation. In a recent study (Zeng et al 2015), both LST night-time and daytime were used for T_{air} maximum and for T_{air} maximum estimation.

Zaksek and Schroedter-Homscheidt (Zaksek et al. 2009) in their research have found

that T_{air} is driven more by LST than by direct solar radiation, meaning that LST is the key variable for T_{air} estimation. We have collected 12 variables, including LST, emissivity31, emissivity32, Albedo, NDVI, altitude, relative humidity, latitude, wind speed, wind direction, air pressure and Julian day, as the potential variables for T_{air} estimation with reference to previous works (Vancutsem et al. 2010, Benali et al. 2012, zeng et al. 2015, Zhang et al. 2011, Jang and Viau 2004). Moreover, Elevation, latitude were take into account for considering the variability of climatic conditions between different regions. The Julian day and NDVI were considered because, it reflects seasonal variation in air temperature. However, the whole process of the feature subset selection is consists of four basic steps: 1) to generate the subset, 2) to evaluate the subset, 3) to consider a stopping criterion, and 4) to validate the results (Liu and Yu 2005). Dependent on the state space search method, the feature subset will be generated. After that each generated feature subset will be evaluated using an evaluation criterion in step 2. After repeating the steps 1 and 2 for depending on the stopping criterion, the best feature subset will be selected, then this subset will be validated on an unseen dataset. In order to assess, rank the feature, select the best model through many candidate model, and find a subset of variables which improves the overall prediction performance several criterion, such as: Root mean squared error (RMSE), Bayesian information criterion (BIC), Akaike information Criteria (AIC) and adjusted R-squared are considered.

5.2 Materials and Methods

5.2.1 Classification of feature selection methods

Feature selection methods can be categorized mainly into filters, wrappers, embedded, and hybrid methods (Hoque et al., 2014, Tang et al., 2014).

A. Filter Methods

Filter methods are generally used as a preprocessing step. Filter methods can be categorized into two categories: Univariate filter methods and multivariate filter methods.The univariate filter methods are the type of methods where individual features are ranked according to specific criteria and The selection of features is independent of any machine learning algorithms. Different types of ranking criteria which can be used for univariate filter methods have been enlisted below:

- Information gain
- Fisher score
- Variance threshold
- Chi-square test
- Correlation coefficient

The redundant features may be selected by the univariate filter methods, because the relationship between individual features is not considered while making decisions and it is one of drawback of this methods but univariate filter methods are ideal for removing constant and quasi-constant features from the data. While the multivariate filter methods are capable of removing redundant features from the data since because the mutual relationship between the features are taken into account. Moreover, multivariate filter methods can be used to remove duplicate and correlated features from the data. However, depends on tasks, the filters are also classified into: classification, regression or clustering, because not all the filter features can be used for all classes of data mining tasks. While there are many filter methods but common search strategies are described briefly in the studies of Hoque et al. 2014,. Liu and Motoda 1998, Bradley and Mangasarian 1998, Liu and H. Motoda 1998, Maldonado et al.2014, Tang et al. 2014, Kim et al. 2002, Cortizo and Giraldez 2006, Alelyani et al. 2013, Liu et al. 2014, Benoît et al. 2013, Alelyani et al. 2013, Guyon and Elisseeff, 2003, Sandri and Zuccolotto 2006).

B. Wrapper Methods

In the wrappers method, feature subsets is considered by the quality of the performance on a modelling algorithm, which is taken as a black box evaluator. Therefore, a wrapper will evaluate subsets based on the classifier performance for classification tasks (e.g. Naïve Bayes or SVM) (Bradley and Mangasarian, 1998; Maldonado et al., 2014), while for clustering, a wrapper will assess subsets based on the performance of a clustering algorithm (e.g. K-means) (Kim et al., 2002). For each subset, the evaluation is repeated, and the subset generation is dependent on the search strategy, in the same way as with filters method. Due to dependency on the resource demands of the modelling algorithm, wrappers are much slower than filters in finding adequately good subsets. The feature subsets are also biased towards the modelling algorithm on which they were evaluated (even when using cross-validation). Therefore, for a reliable generalization error estimate, it is necessary that both an independent validation sample and another modelling algorithm are used after the final subset is found. On the other hand, it has been empirically proven that wrappers obtain subsets with better performance than filters because the subsets are evaluated using a real modelling algorithm. Some common examples of wrapper methods are forward feature selection, backward feature elimination, recursive feature elimination, etc. Practically any combination of search strategy and modelling algorithm can be used as a wrapper, but wrappers are only feasible for greedy search strategies and fast modelling algorithms such as Naïve Bayes (Cortizo and Giraldez, 2006), linear SVM (Liu et al., 2014), and Extreme Learning Machines (Benoit et al., 2013).

C. Greedy Forward Search

The greedy forward selection starts with one variable in the model. In each iteration, the variable will be added then the the classifier is evaluated with the

new variable. The new variable is only kept if there is a notable increase in performance and improves the model's accuracy (Guyon and Elisseeff, 2003).

D. Exhaustive Search

In this chapter, a brute force approach is explored as decsriebed in (Guyon and Elisseeff, 2003). This means that go through all possible combination of parameters extensively in order to find which possible combination of features gives the best result. In this study, the 12 features are considered, therefore, there are $2^{12} = 4096$ different possibilities. Every single possible combination of a dataset were taken into account for selecting the best candidate model through the 4096-possible state.

E. Embedded and Hybrid Methods

In embedded methods, the feature selection is performed during the modelling algorithm's execution. Therefore, these methods are embedded either as its normal or extended functionality, in the algorithm. Common embedded methods are involved different types of decision tree algorithms: CART, random forest (Sandri and Zuccolotto, 2006), C4.5, but also other algorithms (e.g. multinomial logistic regression and its variants (Cawley et al., 2007). Some embedded methods perform feature weighting based on regularization models with objective functions that minimize fitting errors and, at the same time, force the feature coefficients to be small or to be exact zero. These methods are either based on Lasso (Ma and Huang, 2008) or Elastic Net (Zou and Hastie, 2005) which usually work with linear classifiers (SVM or others) and have penalties to features that do not contribute to the model. Hybrid methods were proposed to combine the best properties of filters and wrappers. First of all, in order to reduce the feature space dimension space, a filter method is used possibly providing several candidate subsets (Das, 2001). Then, a wrapper is applied to find the best candidate subset. Hybrid methods usually achieve high accuracy that is characteristic to wrappers and high efficiency characteristic to filters. While practically any combination of filter and wrapper can be used for constructing the hybrid methodology, several interesting methodologies were recently proposed, such as: fuzzy random forest-based feature selection (Cadenas et al., 2013), hybrid ant colony optimization (Ali and Shahzad, 2012), hybrid genetic algorithms (Oh et al., 2004) , or mixed gravitational search algorithm (Sarafrazi and Nezamabadi-pour, 2013).

5.2.2 Model Selection Criteria

In this chapter, a full brute force approach and greedy best first search were explored for six LC/LU during day and night time with ANN approximation. After that, we somehow choose a pool of candidate models with different subset size of the features from our original search space, some questions are arising such as:

A. Fist Question:

How we can the select the best model among this alternative model which somehow, in case of Exhaustive Search algorithm, they look like a nested model?

B. Second Question:

Can we choose the best model only with lowest RMSE and AIC or BIC values and how much is it reliable?

D. Third Question:

How we can find the most important features?

These questions will be answered in sections.

5.2.2.1 AIC Overview

Before engaging in the construction of a model either a linear regression or any generalized linear model, we must accept that there are no true models, indeed, models only approximate reality. The question then is to find which model would best approximate reality with the given data. Basically,what we are trying to do is to have a model which minimizing the loss of information. Kullback and Leibler (1951) investigated such issues and they developed a measure, which is well known to the Kullback-Leibler information. The Kullback-Leibler information represent the information lost when approximating reality (i.e., a good model minimizes the loss of information). A few decades later, Akaike (Kullback and Leibler, 1951; Akaike,1974; Burnham and Anderson, 2001) proposed using Kullback-Leibler information for model selection. They established a relationship between the maximum likelihood, which is an estimation method used in many statistical analyses, and the Kullback-Leibler information. In essence,they developed an information criterion to estimate the Kullback-Leibler information, Akaike's information criterion (AIC), which is expressed as following:

$$AIC = -2(loglikelihood) + 2k \tag{5.1}$$

Where K is the number of estimated parameters included in the model. The log-likelihood of the model reflects the overall fit of the model (smaller values indicate worse fit). For normally distributed errors where arbitrary constants have been deleted, the AIC with the following equation can be considered in cases where analyses are based on more conventional least squares regression

$$AIC = log\frac{SSE}{T} + \frac{2m}{T} \tag{5.2}$$

Where $SSE = \sum_{i=1}^{T}(y_i - \widehat{y}_i)^2$ and T is the sample size. where y_i is the i^{th} value of the variable to be predicted, and $\widehat{y}$ is the predicted value of y_i. It is important to note that, in the count of parameters (m) the estimation of variance must be taken into account. The goodness-of-fit of the model to the data will be measured by the first part in equation (5.2) but the second part is considered as a penalty for model

over-parametrization and over-fitting. The optimal subset will be selected when AIC is minimized. The first term of equation will be smaller as the model becomes more complex, but the second term will be larger(Min et al.,2001). The AIC penalizes for the additional parameters. Therefore, AIC selects a model that fits well but has a minimum number of parameters.In case of small sample sizes (i.e., n/m < 40), the second-order AICC can be applied:

$$AIC = log\frac{SSE}{T} + \frac{2m}{T-m-1} \tag{5.3}$$

Where T is the sample size. As sample size increases, the last term of the AICC approaches zero, and the AICC tends to lead the same conclusions as the AIC (Burnham and Anderson. 2002, Burnham et al., 2011).

Comparing Models

For comparison of model, two measures associated with the AIC such as the delta AIC and Akaike weights can be considered. These are easy to compute and interpret, as calculations remain the same regardless of whether the AIC or AICC is used. The delta AIC ($\triangle i$), is a measure of each model relative to the best model, which is calculated as:

$$DeltaAIC = \triangle_{\text{i}} = AIC_{\text{i}} - minAIC \tag{5.4}$$

Where AIC_{i} is the AIC value for model i, and min AIC is the AIC value of the best model. As a rule of thumb, a $\triangle_{\text{i}} < 2$ suggests substantial evidence for the model, values between 3 and 7 indicate that the model has considerably less support, whereas a $\triangle_{\text{i}} > 10$ indicates that the model is very unlikely (Burnham and Anderson, 2002 and 2004). Akaike weights (wi) provide another measure of the strength of evidence for each model, and represent the ratio of delta AIC values for each model relative to the whole set of R candidate models:

$$AICweight = w_{\text{i}} = \frac{exp(-\frac{1}{2}\triangle_{\text{i}})}{\sum_{r=1}^{R} exp(\frac{1}{2}\triangle_{\text{r}})} \tag{5.5}$$

In fact, we are simply changing the scale of the $\triangle_{\text{i}}$ to compare them on a scale of 1 (i.e. so that the sum of the w_{i} equals 1). The interpretation of Akaike weights (w_{i}) is straightforward and they indicate the probability that the model is the best among the whole set of candidate models. For example, an Akaike weight of 0.75 for a model, says that given the data, it has a 75% chance of being the best one among those considered in the set of candidate models (Symonds and Moussalli, 2011; Link and Barker, 2006; Grueber et al., 2011). In addition, one can compare the Akaike weights of the « best » model and competing models to determine to what extent it is better than another. These are termed evidence ratios and are calculated as:

$$Evidence - ratio = ER_{\text{i}} = \frac{w_{\text{j}}}{w_{\text{i}}} = 1.375 \tag{5.6}$$

Where model j is compared against model i. For instance, an evidence ratio of would indicate that model j is only 1.375 more likely than model i to be the best, given the set of R candidate models and the data. This suggests that the rank of model j might change if we were to take a series of independent samples of identical size (Burnham and Anderson, 2002). Moreover, there is also possibility to compare the evidence ration of the best model with another alternative model as following:

$$Evidence - ratio = ER_{\text{i}} = \frac{w_{\text{best}}}{w_{\text{i}}} \tag{5.7}$$

Where w_{best} is the weight of the best model and w_{i} is the weight of the other individual models. The weight, w_{i} is considered the weight of evidence in favour of a model being the actual best model for given data, given that one of the models must be the best model. Note that the weights of all models summed together is equals one. The evidence ratios, ER_{i}, is the relative likelihood of pair of models, representing the evidence about fitted models as to which is better in an information criteria sense. In addition, If the primary objective of modelling is to evaluate the relative importance of many potential predictor variables (Relative variable importance), then summing Akaike model weights across all models that include that variable can be a useful approach (Burnham and Anderson, 2002). In other words, there would be a high degree of uncertainty regarding the best model. Akaike weights are also useful to give a measure of the relative importance of a variable: one simply sums the w_{i} of the models including the variable and compares it to those that do not. However, a better approach is to obtain a model-averaged estimate for the variable across all models. The greatest strength of the AIC is its potential in model selection (i.e. variable selection), because it is independent of the order in which models are computed. In the case, where there are many models ranked highly based on the AIC, we can incorporate model uncertainty to obtain robust and precise estimates, and confidence intervals. Moreover, AIC may not be the best criterion for a specific case because often leads to a model with unnecessarily large number of parameters(Gooijer and Kumar, 1992).

5.2.2.2 Bayesian Information Criteria

The Bayesian information criterion (BIC), was proposed by Schwarz (1978). The difference between the BIC and the AIC is the greater penalty imposed for the number of additional parameters by the BIC. Burnham and Anderson provide theoretical arguments in favour of the AIC, particularly the AICC over the BIC (Burnham and Anderson, 2002). Moreover, in the case of multivariate regression analysis, Yang explains why AIC is better than BIC in model selection (Yang, 2005). The BIC is computed as follows:

$$BIC = log(SSE) + \frac{mlog(T)}{T} \tag{5.8}$$

there is similarity between Eq.(5.8) and Eq.(5.3) because both equations are composed of two parts but the difference is in the penalty term. For model complexity, BIC consider

bigger penalty than AIC. It means that, using of the BIC criteria in model selection could lead in a model whose number of variable is no bigger than that chosen by AIC. Schwarz (1978) and Riassanen (1978) had developed Eq. (5.8). A consistent estimate of the order of auto regressive model is given by BIC. Hence, BIC is often applied in real applications for model selection rather than AIC because it is a more reliable criterion. Regarding the proper penalty term both AIC and BIC, there are some discussions. The linear function of the number of parameters in the penalty term is argumentative issue (De Gooijer et al., 1985). De Gooijer and Kumar 1992 and Granger (Granger, 1993) had proposed, one extensions to these criteria which can be seen in equation (5.9):

$$BIC = log(SSE) + \frac{m^{\mathrm{d}} log(T)}{T} \tag{5.9}$$

Where d is a constant but for non-linear model usually is set to d > 1.

5.3 Results and discussion

Artificial neural networks (ANN) was considered as a universal approximation due to strong ability to represent complex, non-linear behaviour in comparison to more conventional modelling techniques. In recent year, many researchers had applied ANNs for environmental modelling. Finardi et al., 2008, 2008; Pires et al., 2008; Al-Alawi et al., 2008; Ionescu and Candau, 2007; Dutot et al., 2007; Sousa et al., 2007 had used ANNs for real-time forecasting of air quality. Glesias et al., 2007; Shanmuganathan et al., 2006, Maier and Dandy, 2000, Maier, 2006 had used ANNs in ecological modelling and remote sensing. Modelling and control of waste water processes, water quality forecasting within rivers and and distribution systems using ANNs are investigated by Raduly et al., 2007; Machon et al., 2007, Alp and Cigizoglu, 2007, Serodes et al., 2001; Rodriguez and Serodes, 1999, Maier and Dandy, 2000; Dawson and Wilby, 2001; Bowden, 2003; Kingston, 2006. However, the selection of an appropriate set of input variables during ANN development is important for obtaining high-quality models. In this research, two methods of variable selection namely exhaustive search and greedy best first search are applied. Moreover, the AIC, BIC, RMSE and theil's adjusted R^2 are considered as model selection criteria.

5.3.1 Implementation of Input Variable Selection using Brute Force Search-ANN

The Brute force search - ANN algorithm was run for each LC/LU dataset. The variables were selected according some criteria like AIC, BIC and RMSE. As mentioned before, the models are ranked by AIC, with the best approximating model being the one with the lowest AIC value. AIC thus considers how well the model fits the data, but models with greater numbers of fitted parameters will have higher AIC values, all other

things being equal. In other words, models with fewer parameters will be favoured. To illustrate with an example, we applied an analysis of feature ranking in different LC/LU with brute force search approaches. The AIC values for every possible combination of features were calculated for each LC/LU during day and night time. The results of this method are only shown for urban LC/LU during daytime in Table 5.1 Moreover, determining a reduced candidate set of models is a a complex issue as discussed in the studies of Dochtermann and Jenkins, 2010 and Burnham et al., 2010. Moreover, making inference based on the best approximating model alone may not be desirable and it is one of the main purposes of calculating AIC in this study is to present a range of models and their relative AIC scores which by comparing the different models, we can measure how much better the best approximating model is compared to the next best models. The simplest way of doing this is to calculate the difference ($\triangle_i$ or $\triangle AIC_i$) between the AIC value of the best model and the AIC value for each of the other models. The Akaike weight is between 0 and 1, with the sum of Akaike weights of all models in the candidate set being 1, and can be considered as analogous to the probability that a given model is the best approximating model (although there are some who disagree with this, e.g. Link and Barker, 2006; Bolker, 2008; Richards, 2005). Thus, in Table 5.1, the best model has a wi of 0.0885, which can be interpreted as meaning that there is 8.85% chance that it really is the best approximating model describing the data given the candidate set of models considered. With this low Akaike weight of models which is presented in Table 5.1, we cannot be certain that this model is the best. That is to say, there exists model selection uncertainty. Alternatively, the Akaike weights (w_i) or model weight can also be used to estimate the relative importance of variables under consideration. This is done by summing the Akaike weights for each model in which that variable appears. In our study, the variable Air pressure (AP) has an $Akaikeweight = 0.0883 + 0.0875 + 0.0877 = 0.2635$ and so on down the complete list of models. If a particular predictor appears in all of the top models, then its summed Akaike weight will tend towards 1. If that predictor only appears in the very unlikely models, its weight will tend towards zero. As with the Akaike model weight (w_i), the predictor weight can be interpreted as equivalent to the probability that that predictor is a component of the best model. Similarly, these summed weights can be used to rank the various predictors in terms of importance (Burnham and Anderson, 2002, 2004). Summary of model selection procedure is presented in Table 5.1 only for urban LC/LU and illustrated in Figure for all LC/LU during day and night-time. The best model selected using AIC and BIC has different results for different LC/LU during day and night-time. The summary of results is listed for each LC/LU as follow:

A. Needle Leaf Trees

From Figure 5.1(a), the AIC and RMSE chooses nine variables as the best model, whereas a model with four input variables is selected by the BIC during daytime. As shown in Figure 5.2(a), for the night time the AIC and RMSE choose nine variables as the best model, whereas a model with six input variables is selected by the BIC for ES algorithm. Moreover, Figure 5.3(a), shows different parameters have different relative importance weight during day and night time except LST which its Akaike weight is almost 1. In addition, Figure 5.2(a), shows that some

variables have same ranking during day and night time namely included LST and WS but for other parameters like NDVI, RH and WD there is big differences between day and night time. During daytime (WD, AP) and (RH, Emis32) have almost the same Akaike weight but for night-time (NDVI, Lat, WD) have a same Akaike weight. Moreover, it can be seen that the Emis31, Emis32 and Lat have very low Akaike weight which is less than 0.2 which can be representative of insignificant influence on the relationship between LST and T_{2m} in this LC/LU.

B. Agriculture

From Figure 5.1(b), the AIC and RMSE chooses eight variables as the best model, whereas a model with five input variables is selected by the BIC during daytime. As shown in Figure 5.2(b), for the night-time the AIC, RMSE and BIC choose six, eight and five variables as the best model, respectively. In addition, Figure 5.3(b) shows that in general, LST and WD are the most important parameter during day and night time which it's Akiak weight doesn't change too much but for other parameters like WD, Emis31 and NDVI there is big differences between day and night time. During daytime (RH, Alt) and (Emis31, Emis32, NDVI, WS) have a same Akiak weight score but for night-time (AP, WD) and (Emis32, JD) have a same Akiak weight importance. Moreover, the NDVI has very low Akiak weight which is less than 0.2 which can be representative insignificant influence on the relationship between LST and T_{2m} in night-time in this LC/LU but as expected the Akiak weight of NDVI is more than 0.7 which shows that it is a key factor during daytime.

C. Forest

From figure 5.1(c), the AIC and RMSE choose seven variables as the best model, whereas a model with 5 input variables is selected by the BIC during daytime. As shown in figure 5.2(c), for the night-time the AIC, RMSE and BIC choose six, six and three variables as the best model, respectively. In addition, figure 5.3(c), shows that in general, LST, JD, RH and Lat are the parameters which it's Akaike weight doesn't change too much during day and night-time but for other parameters like Emis31, NDVI, WD, WS and AP there is big differences between day and night-time. During daytime (RH, JD, NDVI) and (Albedo, AP) have a same Akaike weight but for night-time (Alt, Lat) and (NDVI, Emis32) have a same Akaike weight importance. Moreover, during night-time, the Emis31 and WD have very low Akaike weight which is less than 0.2 which demonstrates that these parameters have less contribution than the other variables to the measurement of T_{2m} in this LC/LU but as expected the Akaike weight for NDVI is more than 0.65 for daytime. It can be seen from Figure 5.3(a) and 5.3(c), that the Albedo has the same Akaike weight in forest and needle leaf trees LC/LU because these LC/LU are covered with tree.

D. Urban

From Figure 5.1(d), the AIC and RMSE choose nine variables as the best model, whereas a model with five input variables is selected by the BIC during daytime. As shown in Figure 5.2(d), for the night-time the AIC, RMSE and BIC choose

eight, eight and four variables as the best model, respectively. In addition, Figure 5.3(d) shows that in general, LST is only the parameter which it's Akaike weight doesn't change too much during day and night time but for other parameters, there is big changes between day and night-time. During daytime (Alt, Lat, WD), (AP, Emis31) and (Emis32, WS) have a same Akaike weight but for night-time (Alt, WD) and (JD, Lat) have a same Akaike weight importance. Moreover, during night-time, the Emis32 and WD have very low Akaike weight which is less than 0.2 which can be representative insignificant influence on the relationship between LST and T_{2m} in this LC/LU.

E. Airport

From Figure 5.1(e), the AIC and RMSE choose eight variables as the best model, whereas a model with 5 input variables is selected by the BIC during daytime. As shown in Figure 5.2(e), for the night-time the AIC, RMSE and BIC choose six, six and five variables as the best model, respectively. In addition, Figure 5.3(e) shows that in general, LST, Emis32 and Lat are the only the parameters which it's Akaike weight doesn't change too much during day and night time but for some parameters, like JD, RH, WS there is big changes between day and night time. During daytime (Alt, WS, Emis32, Albedo), (Lat, WD) and (Emis31, RH) have a same Akaike weight but for night time (WS, Lat) and (Emis31, Emis32, NDVI, JD) have a same Akaike weight importance. Moreover, during daytime, Albedo, Emis32, Alt and WS are not dominant parameters for estimating T_{2m}.

F. Industrial

From Figure 5.1(f), the AIC and RMSE choose eleven variables as the best model, whereas a model with 7 input variables is selected by the BIC during daytime. As shown in Figure 5.2(f), for the night-time the AIC, RMSE and BIC choose ten, ten and seven variables as the best model, respectively. In addition, Figure .5(f) shows that in general, LST, NDVI, RH, Emis32, WS and AP are the parameters which it's Akaike weight doesn't change too much during day and night time but for some parameters, like JD and WD there is big changes between day and night time. During daytime (NDVI, WS), (Emis31, Albedo) and (Alt, WD) have a same Akaike weight but for night-time (AP, RH) and (Emis31, NDVI, WS) have a same Akaike weight importance. Moreover, during day and night-time, Albedo, Emis31 and Emis32 are not dominant parameters for estimating T_{2m}.

Overall, according to Tables 5.1, 5.2 and Figure 5.3, the most important parameters for estimating the T_{2m} over different LC/LU which it's AIC weight importance is greater than 0.5 are namely included LST, JD, WD and RH for day time and LST, WD, RH and AP for night time respectively. Moreover, Figure 5.1 shows that, BIC always select less variable compare to AIC and RMSE because BIC penalizes larger models more heavily and so will tend to prefer smaller models in comparison to AIC. In addition, the selection results demonstrate LST has more contribution than the other variables to the measurement of T_{2m} during day and night over different LC/LU. Table 5.5 and 5.6 show the final selected model during day and night time for six considered LC/LU based on AIC and BIC criteria. It has been found that AIC and BIC criteria

tend to select the variable with Akaike weight importance of more than 0.4 and o.5, respectively. It means that BIC select variables which have strong and moderately strong relationships between a predictor variable and the response variable were associated with Akaike weight importance from 0.5 to 1. Moreover, Figure 5.1 shows the change of AIC, BIC and RMSE score when a new variable is added. As Figures 5.1(a)–(f) shown, the AIC, BIC and RMSE scores have a decreasing trend until a new variable were selected so that the termination criterion was reached. The variables with minimum value in Figures 5.1(a)–(f) are the final selected variables for the six LC/LU. In addition, it has found that the sum weight of wind speed has significant changes during day and night time on airport LC/LU because of impervious surface of this LC/LU in all season. Many factors influence wind speed, but the single factor that distinguishes wind over a forest from wind over more open terrain, such as an airport, is the presence of tree cover. The wind speed tends to decrease after sunset because at night the surface of the Earth cools much more rapidly than does the air above the surface. Because of this difference in cooling ability, it doesn't take long for the ground to become colder than the air above it. The air in close contact with the ground - say in the lowest 300 feet of the atmosphere - then becomes colder than the air above it. This circumstance leads to the development of what is known as a temperature inversion. Inversions dramatically reduce the amount of mixing that occurs between different vertical layers of the atmosphere. Therefore, once the inversion sets up (after sunset), it is much harder for fast-moving air above the ground to mix down to the surface, where it could appear as a gust of wind. In addition, by considering Tables (5.1- 5.4) and Figures 5.1,5.2, 5.5 and 5.6 we found that, those parameters which have influence on the estimation of T_{2m} during day and night time were selected in both algorithms. Both algorithms tend to select the parameters which their sum weight of Akaike were greater than 0.5. Moreover, compared with the exhaustive search, forward selection is much cheaper. However, forward selection may suffer because of its greediness. For example, if X (1) is the best individual feature, it does not guarantee that either X (1), X (2) or X (1), X (3) must be better than X (2), X (3). Therefore, a forward selection algorithm may select a feature set different from that selected by exhaustive searching. With a bad selection of the input features, the prediction may be significantly different from the true estimation. Our experiment demonstrates that Exhaustive Search (ES) is prohibitively time-consuming. ES is far more expensive than the Forward Selection algorithm (FS), while it is not significantly more accurate than FS. However, the features selected by FS may differ from the result of ES. That is because some of the input features are not mutually independent. Referring to figures 5.1,5.2, 5.5 and 5.6, it can be seen that the accuracy of ES and greedy best search for whole considered LC/LU are almost the same. Moreover, it has been found that LC/LU plays an important role in influencing the relationship between T_{2m} and LST. Our results showed that the effectiveness of optimal models in predicting T_{2m} varied in different land cover types. This variation could be introduced by the specific heat capacities of different land covers. Air temperature mainly depended on the heat transfer process which was strongly influenced by the local radiation budget (Lin et al., 2016). In general, barren land has lower heat capacity than forest. Hence, air was heated much quicker over barren land than forest. Vegetation could also

change latent heat flux, such as enhancing or reducing transpiration (Zeng et al., 2015; Kaufmann et al., 2003), and cool the T_{2m} in forests (Jeong et al., 2009; Pouteau et al., 2011; Van et al., 2013). The cooling effect was not considered in our models due to uneven distribution of meteorological stations across different vegetation types. The meteorological station was too scarce in some vegetation types. Thus, it was hard to take consider of vegetation type in our models. However, land cover also affected land surface albedo, thus, the influence of land cover on estimating T_{air} was conditional and time dependent. Moreover, it has been found the sum weight of RH in all LC/LU during night are higher that night-time the reason is the relative humidity depends not only upon the amount of water vapour present in the air but also on the air temperature. In fact, the relative humidity indicates how much is the percentage of saturation of air (with water vapour) for a particular temperature. If it is fully saturated, then the relative humidity is 100 percent. If air holds some amount of water vapour at a particular temperature and is unsaturated, then, at a lower temperature, the same amount of water vapour may be able to saturate it. Hence, for the same amount of moisture content in the air, the relative humidity may be less for a higher temperature and more for a lower temperature. Therefore, naturally the relative humidity is less during daytime and more during night-time. Finally, in this paper we have tried to introduce feature selection techniques. The literature on feature selection techniques is very vast encompassing the applications of machine learning and pattern recognition. Comparison between feature selection algorithms can only be done using a single dataset since each underlying algorithm will behave differently for different data. Feature selection techniques show that more information is not always good in machine learning applications. We can apply different algorithms for the data at hand and we can select a final feature selection algorithm. For the application at hand, a feature selection algorithm can be selected based on the following considerations: simplicity, stability, number of reduced features, classification accuracy, storage and computational requirements. Overall, applying feature selection will always provide benefits such as providing insight into the data, better classifier model, enhance generalization and identification of irrelevant variables. For the results in this paper, we use the classifier accuracy and the number of reduced features to compare the feature selection techniques. We have also successfully used feature selection for improving predictor performance.

5.3.2 Summary and Conclusion

The feature selection problem has been studied by the statistics and machine learning communities for many years. It has received more attention recently because of enthusiastic research in data mining. The aim is to construct a model that predicts/estimate well or explains the relationships in the data. Though AIC and BIC are the two model selection criteria s, but they are not the same. The AIC can be designated as a measure of the goodness of fit of any estimated statistical model. The BIC is a type of model selection among a class of parametric models with different numbers of parameters. For additional parameters, BIC penalizes free parameters more strongly

than AIC. AIC generally tries to find unknown model that has high dimensional reality. On the other hand, the BIC comes across only true models. It can also be said that BIC is consistent whereas AIC. However, AIC is susceptible to over-fitting the data, whereas BIC is susceptible to under-fitting the data. The reason is that they penalize the free parameters differently. Though BIC is more tolerant when compared to AIC, it shows less tolerance at higher numbers. AIC is good for making asymptotically equivalent to cross-validation. On the contrary, the BIC is good for consistent estimation (Aho et al., 2014, Bolker 2008). An advantage of brute force approach is that all possible combinations of variables are given a chance to appear together. An obvious disadvantage is that the computation time. Moreover, in this study, we tried to answer, why not just use the global model? It has been argued that one should make inference from a model with all the factors thought to be important. This approach can be easily apply and prevent the complications of model selection. The first serious drawback here is the lack of precision in the estimated parameters. A given data set has only a finite amount of information; each time a parameter estimate is made, the information left is reduced. Increasing the number of parameters eventually makes the fitted model unstable and uninformative. The probability of finding factors that are spurious increases. New parameters are estimated but with increasing uncertainty phenomenon is an aspect of the Principle of Parsimony. A second serious pitfall arises, when as is common, the global model has many parameters in it. One has to resort to analysing the set of resultant parameter estimates, as if reduced-dimension models to the set of (poorly estimated and correlated) global parameter estimates. It is very demanding to do this efficiently and validly. Indeed, the proper way to proceed is to fit the corresponding reduced models (as special cases of the global model) to the original data and do proper multi model inference. This latter approach facilitates understanding of the information in the data; fitting only a large global model generally fails as a strategy for effective inference. In addition, in this chapter, we applied ES feature selection method, which is to exhaustively evaluate all possible combinations of the input features, and then find the best subset. Obviously, the computational cost of exhaustive search is prohibitively high, with considerable danger of over-fitting. Hence, people resort to greedy methods, such as forward selection. Our investigation shows that the greediness of the feature selection algorithms greatly improves the efficiency, while does not corrupt the correctness of the selected feature set so that the prediction accuracy using the selected features remains satisfactory. The main motivation for this research was to investigate the practical applicability of AIC and BIC criteria in two different feature selection approach using supervised training of ANN. theil's adjusted R^2 is another criteria that is applied to select the number of repressors. Though the determination coefficient (R^2) measures the goodness-of-fit of a model, it almost always increases and never decreases with the number of repressors. Therefore, if we consider R^2 as model selection criteria, then it would always favour larger number of variable. The adjusted R^2 corrects the problem with an adjustment to the degrees of freedom. The two other measurements namely sum of squared errors (SSE) or residual variance can be used for the goodness-of-fit. It should have been considered that minimizing the estimated residual variance is equal to maximizing R^2, therefore residual variance is redundant in cases that R^2 is applied. Cameron (1993)

debates that R^2 is not an effective tool for the prevention of data mining because it will rise on the addition of any variable whose t ratio is greater than one when entered into the model, although R^2 is used as a model selection criterion.

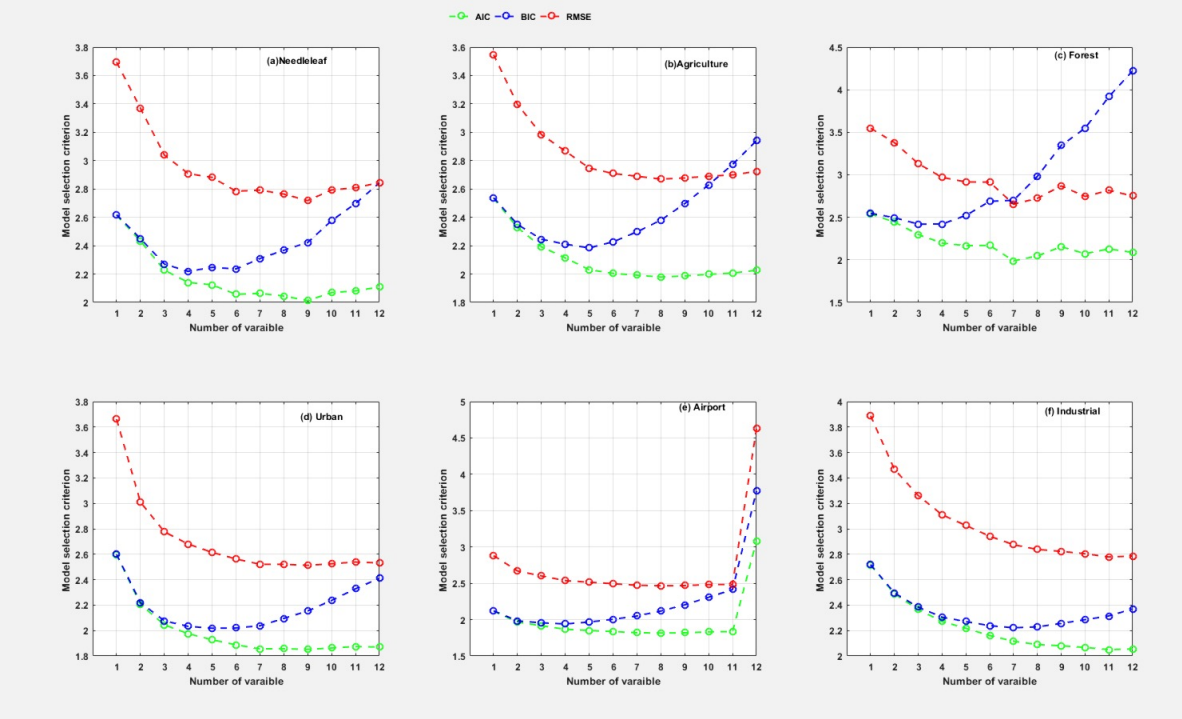

Figure 5.1: The sub-plots (a) to (f) shows the AIC, BIC and RMSE values for estimating T_{2m} for different LC/LU during daytime for the Exhaustive search algorithm.

Table 5.1: Summary of the model selection procedure applied to the urban LC/LU using Brute force search. For each candidate models, we reported RMSE, AIC, BIC, MAPE, MBE, MSE, Akaike weight Wi, Δ i = AICi – min AIC, ER and R^2 in Celsius degree

Number of variables	Candidate model	BIC	RMSE	MAPE	MBE	MAE	MSE	AIC	Δi	wi	ER	R^2
1	LST	2.599	3.661	7.691	0.357	2.825	13.409	2.596	0.745	0.061	1.451	0.840
2	LST, Emis32	2.219	3.009	6.301	0.399	2.312	9.057	2.205	0.353	0.0742	1.193	0.895
3	LST, JD, Lat	2.076	2.776	5.679	0.2959	2.0897	7.706	2.045	0.193	0.080	1.101	0.911
4	LST, JD, Emis32, Lat	2.032	2.679	5.492	0.232	2.024	7.180	1.975	0.123	0.0832	1.063	0.915
5	LST, JD, RH, Alt, Lat	2.018	2.614	5.418	0.529	1.981	6.838	1.927	0.075	0.0852	1.038	0.921
6	LST, JD, Albedo, RH, Alt, WD	2.020	2.562	5.335	0.612	1.946	6.567	1.888	0.036	0.086	1.018	0.925
7	LST, JD, RH, Alt, WD, WS, AP	2.037	2.521	5.258	0.625	1.917	6.356	1.856	0.004	0.088	1.002	0.927
8	LST, JD, Emis32, NDVI, RH, Alt, WD, WS	2.095	2.520	5.304	0.605	1.935	6.354	1.857	0.005	0.088	1.002	0.928
9	LST, JD, Emis31, NDVI, Albedo, RH, Lat, WD, WS	2.015	2.512	5.280	0.593	1.927	6.314	1.851	0	0.088	1	0.927
10	LST, JD, Emis32, NDVI, Albedo, RH, Alt, Lat, WD, WS	2.237	2.525	5.244	0.619	1.912	6.377	1.862	0.010	0.088	1.005	0.926
11	LST, JD, Emis31, Emis32, NDVI, RH, Alt, Lat, WD, WS, AP	2.328	2.538	5.303	0.765	1.926	6.445	1.874	0.022	0.087	1.011	0.928
12	LST,JD,Emis31,Emis32,NDVI,Albedo,RH,Alt,Lat, WD, WS, AP	2.411	2.531	5.338	0.807	1.937	6.410	1.870	0.018	0.087	1.009	0.930

Table 5.2: Selected models for different LC/LU based on AIC and BIC according to Brute force search algorithm during daytime

LC/LU	Industrial		Urban		Agriculture		Needle leaf		Forest		Airpoet	
Model selection criteria	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC
LST	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
JD	✓	✓	✓	✓	✓	✓	✓	✓	✓	-	✓	✓
Emis31	✓	-	-	-	-	-	-	-	-	-	✓	-
Emis32	✓	-	✓	-	-	-	-	-	-	-	-	-
NDVI	✓	-	✓	-	✓	-	✓	-	✓	✓	✓	✓
Albedo	✓	-	-	-	-	-	✓	-	✓	-	-	-
RH	✓	✓	✓	✓	✓	✓	✓	-	✓	✓	-	-
Alt	✓	✓	✓	✓	✓	✓	✓	✓	-	-	✓	-
Lat	✓	✓	✓	✓	-	-	-	-	-	-	✓	✓
WD	✓	✓	✓	✓	✓	✓	✓	-	✓	✓	✓	✓
WS	-	-	✓	✓	-	-	✓	✓	-	✓	-	-
AP	✓	✓	-	-	✓	-	✓	-	✓	-	✓	-
Total selected parameter	11	7	9	7	7	5	9	4	7	5	8	5

Table 5.3: Selected models for different LC/LU based on AIC and BIC according to Brute force search algorithm during night-time

LC/LU	**Industrial**		**Urban**		**Agriculture**		**Needle leaf**		**Forest**		**Airport**	
Model selection criteria	**AIC**	**BIC**	**AIC**	**BIC**	**AIC**	**BIC**	**AIC**	**BIC**	**AIC**	**BIC**	**AIC**	**BIC**
LST	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
JD	✓	✓	✓	✓	✓	✓	✓	✓	✓	-	✓	✓
Emis31	✓	-	-	-	-	-	-	-	-	-	✓	-
Emis32	✓	-	✓	-	-	-	-	-	-	-	-	-
NDVI	✓	-	✓	-	✓	-	✓	-	✓	✓	✓	✓
Albedo	✓	-	-	-	-	-	✓	-	✓	-	-	-
RH	✓	✓	✓	✓	✓	✓	✓	-	✓	✓	-	-
Alt	✓	✓	✓	✓	✓	✓	✓	✓	-	-	✓	-
Lat	✓	✓	✓	✓	-	-	-	-	-	-	✓	✓
WD	✓	✓	✓	✓	✓	✓	✓	-	✓	✓	✓	✓
WS	-	-	✓	✓	-	-	✓	✓	-	✓	-	-
AP	✓	✓	-	-	✓	-	✓	-	✓	-	✓	-
Total selected parameter	11	7	9	7	7	5	9	4	7	5	8	5

Table 5.4: Selected models for different LC/LU based on AIC and BIC according to Greedy best first search algorithm during daytime

LC/LU	**Industrial**		**Urban**		**Agriculture**		**Needel leaf**		**Forest**		**Airport**	
Model selection criteria	**AIC**	**BIC**	**AIC**	**BIC**	**AIC**	**BIC**	**AIC**	**BIC**	**AIC**	**BIC**	**AIC**	**BIC**
LST	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
JD	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Emis31	-	-	-	-	-	-	-	-	-	-	✓	✓
Emis32	-	-	-	-	-	-	-	-	-	-	-	-
NDVI	-	-	-	-	✓	✓	✓	✓	✓	✓	-	-
Albedo	-	-	-	-	-	-	-	-	✓	✓	-	-
RH	✓	✓	✓	✓	✓	✓	-	-	✓	✓	✓	-
Alt	-	-	✓	✓	-	-	✓	✓	-	-	-	-
Lat	✓	✓	-	-	✓	✓	✓	-	-	-	-	-
WD	✓	✓	✓	✓	✓	✓	✓	-	✓	✓	✓	-
WS	✓	-	-	-	-	-	-	-	-	-	-	-
AP	✓	✓	✓	✓	✓	-	✓	-	✓	-	✓	✓
Total selected parameter	7	6	6	6	7	6	7	4	7	6	6	4

Table 5.5: Selected models for different LC/LU based on AIC and BIC according to Greedy best first search algorithm during night-time

LC/LU	Industrial		Urban		Agriculture		Needel leaf		Forest		Airport	
Model selection criteria	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC
LST	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
JD	✓	-	✓	✓	-	-	✓	✓	✓	-	-	-
Emis31	-	-	-	-	✓	✓	-	-	-	-	-	-
Emis32	-	-	-	-	-	-	-	-	-	-	-	-
NDVI	-	-	-	-	-	-	-	-	-	-	-	-
RH	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Alt	-	-	✓	✓	-	-	✓	✓	-	-	-	-
Lat	✓	✓	-	-	-	-	-	-	-	-	✓	✓
WD	✓	✓	✓	-	✓	-	✓	-	-	-	✓	-
WS	-	-	-	-	✓	-	✓	-	✓	✓	✓	✓
AP	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Total selected parameter	6	6	6	5	6	4	7	5	5	4	6	5

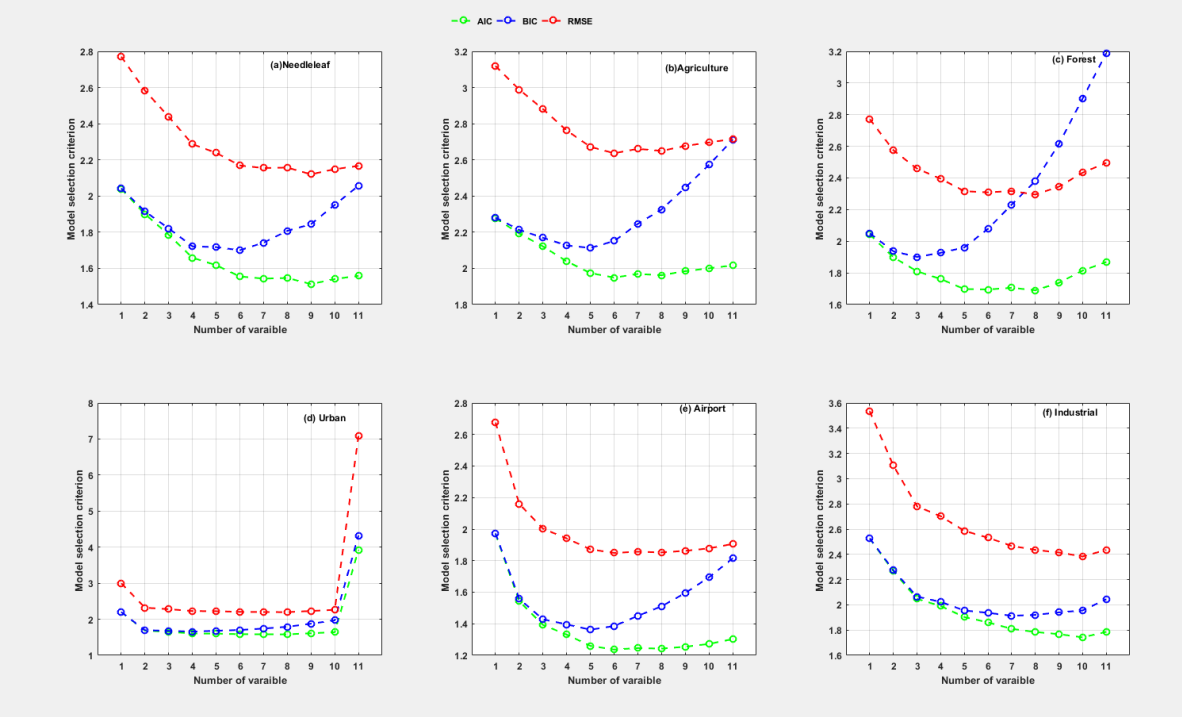

Figure 5.2: The sub-plots (a) to (f) shows the AIC, BIC and RMSE values for estimating T_{2m} for different LC/LU during night-time for the Exhaustive search algorithm.

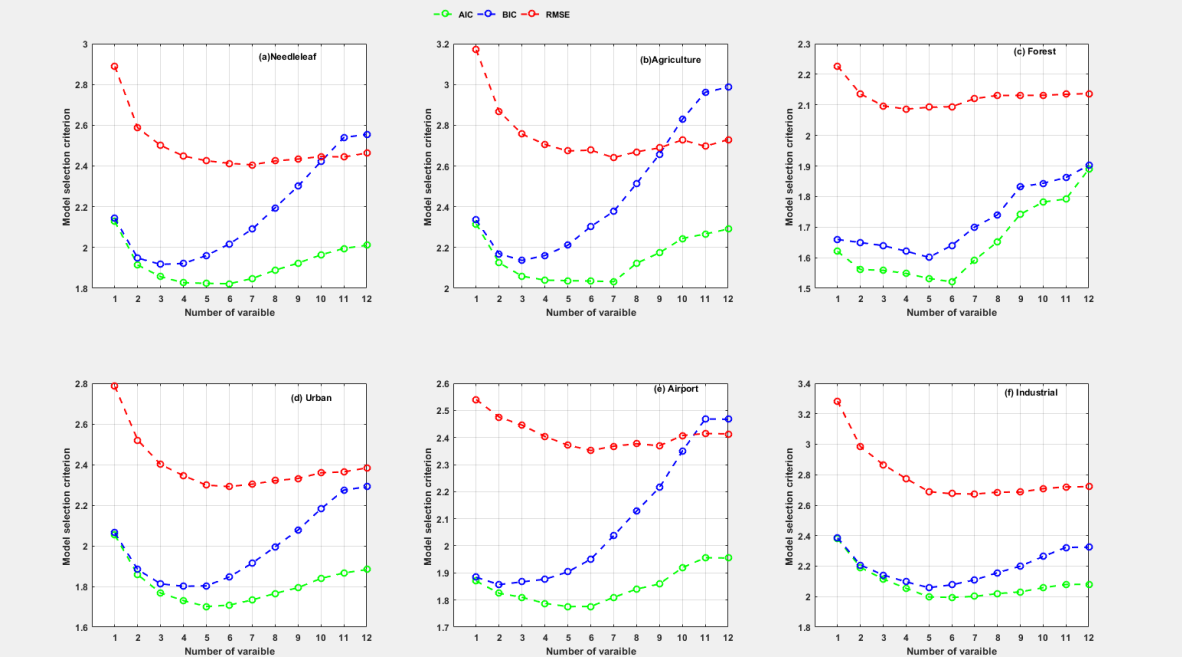

Figure 5.5: The sub-plots (a) to (f) shows the AIC, BIC and RMSE values for estimating T_{air} for different LC/LU during daytime for the greedy best search algorithm.

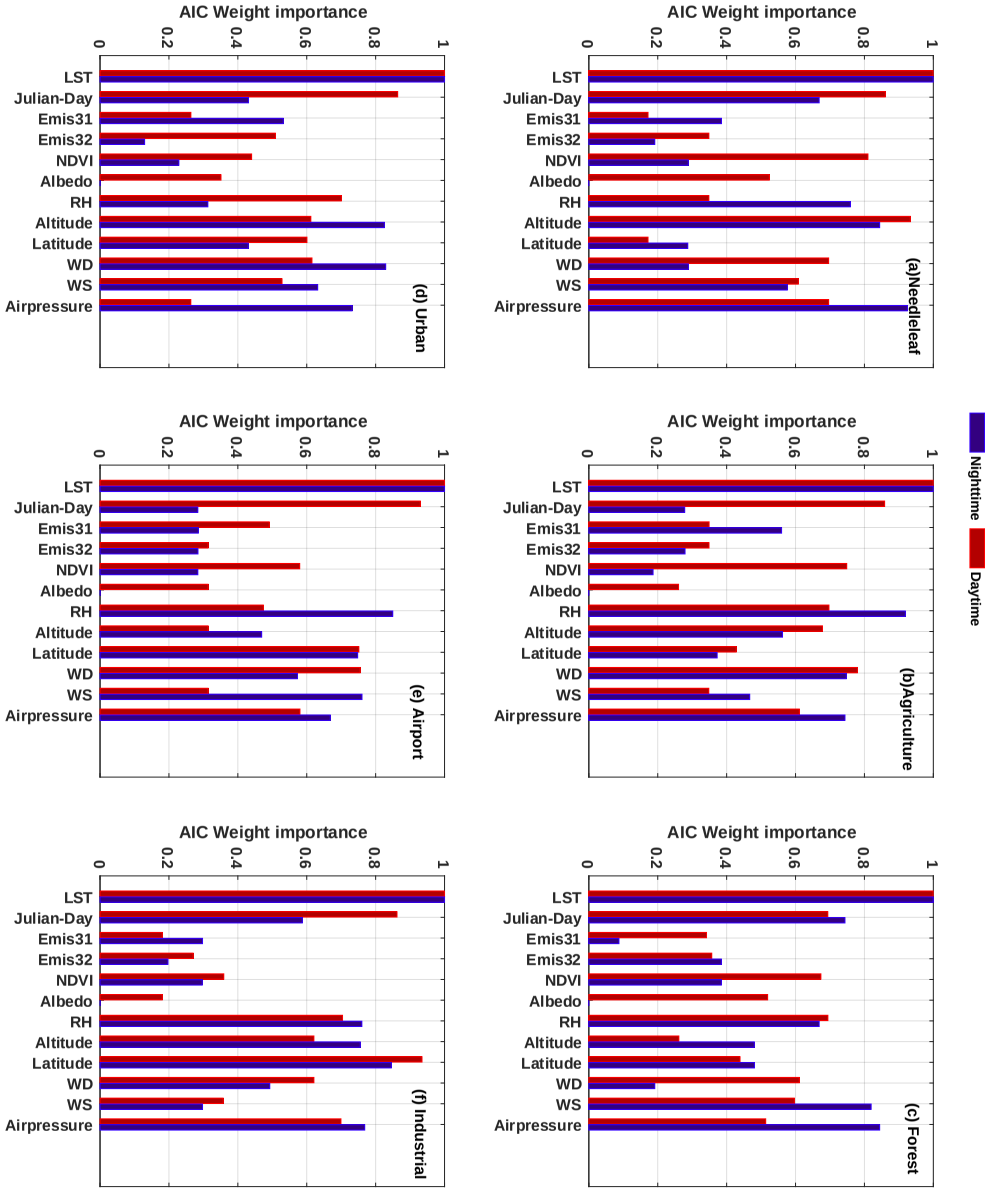

Figure 5.3: Akaike weight importance for estimating T_{air} for different LC/LU are presented in sub-plots (a) to (f) during day and night-time for the Exhaustive search algorithm.

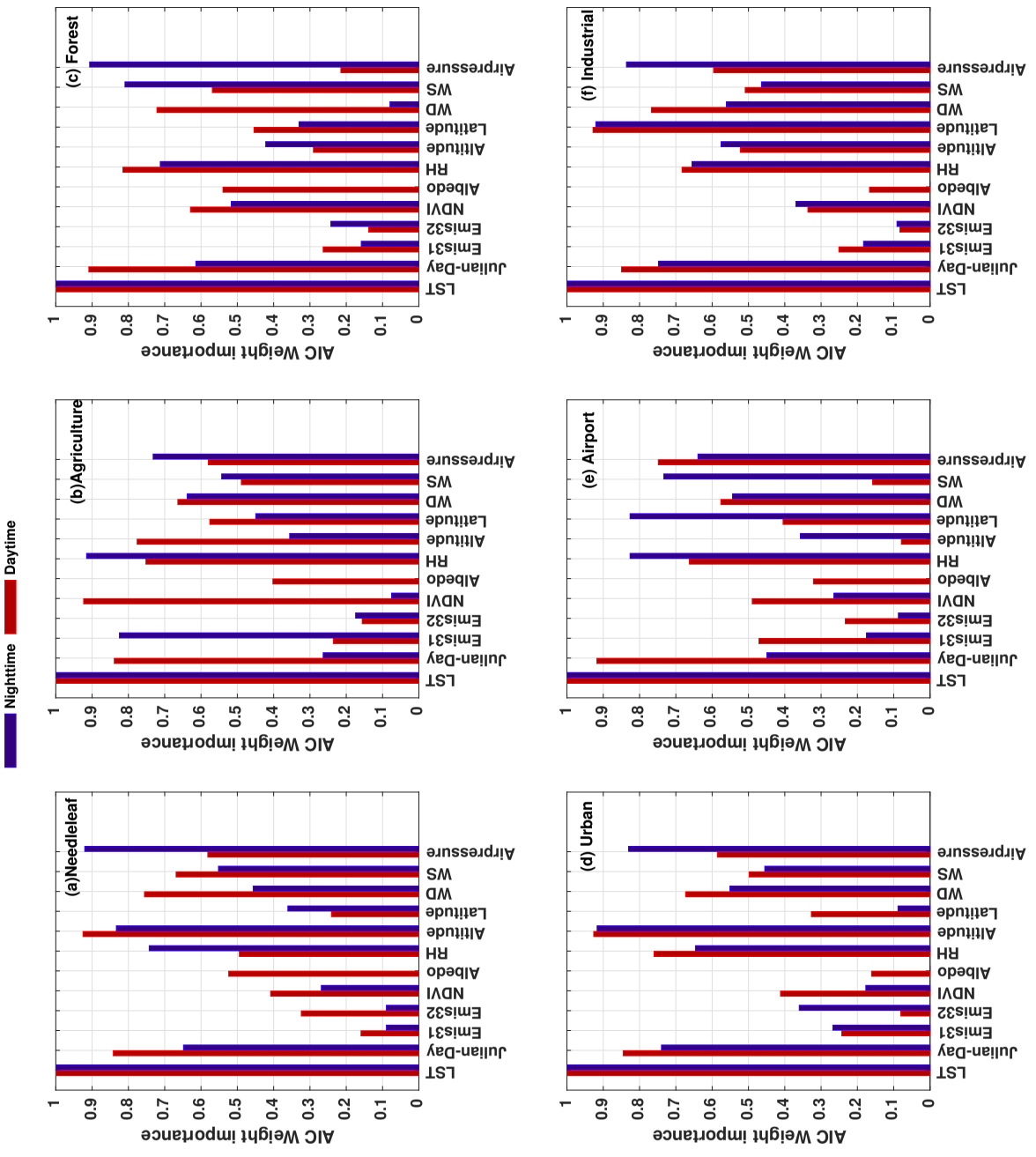

Figure 5.4: Akaike weight importance for estimating T_{air} for different LC/LU are presented in sub plots (a) to (f) during day and night-time for the greedy best first search algorithm.

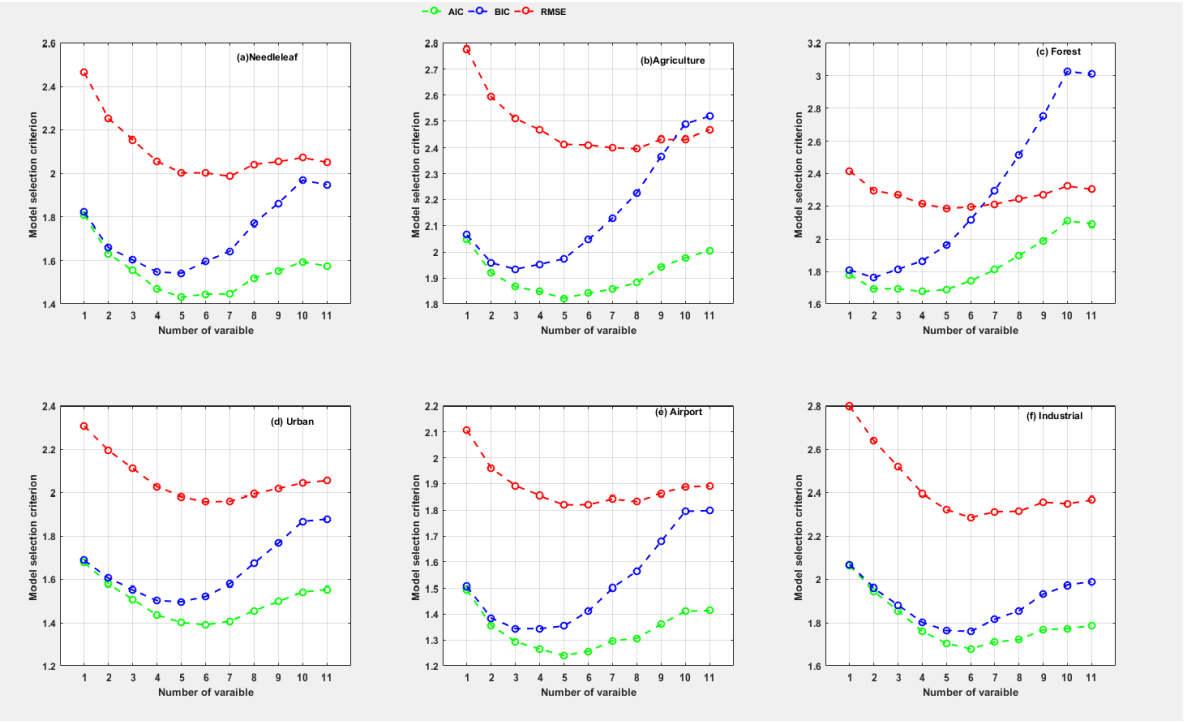

Figure 5.6: The sub-plots (a) to (f) shows the AIC, BIC and RMSE values for estimating T_{air} for different LC/LU during night-time for the greedy best search algorithm.

6 Conclusion

This study presents a comprehensive investigation of the following questions using Satellite data and ground based measurement for six different LC/LU over Berlin for considered year 2007-2013:

Question 1)

Investigation the relationship between T_{air} and Land surface temperature (LST) seasonally during day and night time over six land cover types (LCT), if the land cover type can affect the relationship between land surface temperature and air temperature.

Question 2)

Investigation the relationship between T_{air} and the normalized difference vegetation index (NDVI) seasonally during day and night time over six LCT, if NDVI has different effects on nocturnal and daily air temperatures and how is this effect in different land cover types.

Question 3)

To explore the Spatio-temporal variability of LST and T_{air} relationship: the relationship between remotely sensed LST and T_{air}, is strongly influenced by the local surface heat fluxes, is analysed by overlaying a spatial window of varying size on the MODIS LST grid.

Question 4)

To investigate the relationship between observed T_{air} and the four LST Products over Berlin.

Question 5)

What is the best method to estimate air temperature from remote sensed land surface temperature and how can we find the best estimator?

Question 6)

To rank the features base on using the Greedy Best-First Search and Brute-force search in order to find the optimal subset of feature which are influence the relationship between LST and T_{air} during day and night for different LC/LU and to explore the advantage and disadvantage of using different type of selection criteria such as RMSE, R^2, BIC, and AIC.

Question 7)

Why not just use the global model?

Question 8)

How can we select the best approximate model and evaluate it?

6.1 Result

This section answers the main questions 1 and 2 presented previously in the introduction chapter of the thesis. This study investigates both LST–NDVI and LST–T_{2m} relationships over different LU/LC over Berlin, during different seasons at day and night. In the analysis of different LU/LC types, the results indicate that different LU/LC types have significantly different effects on LST and NDVI as measured by the MODIS in Berlin. The correlation analysis between NDVI and LST for different seasons shows that this relationship depends on the season, time of day, and LU/LC type. The results depict an inverse correlation between LST and NDVI over every LU/LC type.
Moreover, in order to evaluate the influence of LU/LC, a regression analysis between LST and NDVI was utilized and multiple comparisons were made. By considering the slope of regression function for different LU/LC and seasons, it was found that the regression coefficient is dependent on LU/LC type. This means that the NDVI–LST slope is highly dependent on the arrangement of the vegetation type, natural features, rate of ET from surface, and artificial physical features of an area. There is a negative slope for sparse vegetation covers such as industrial, airports, and urban area, but for closed vegetation canopies such as agriculture and needle-leaf trees areas, the slope is positive.
Next to it, the comparisons of mean LST and NDVI values by individual pairings of LU/LC types were made in the research area. The mean LST and NDVI within every LU/LC show a clear negative correlation between LST and NDVI. In certain LU/LC types, it is observed that the relationship between LST and NDVI is varied. This reveals that the LST is affected mainly by the land surface materials and has a close relationship with the abundance of vegetation and the effect to its surrounding.
The strength of relationships between LU/LC variables (e.g. NDVI and vegetation fraction) and LST, however, varies significantly from study to study due to the different measurement of variables and units of analyses (Huang, Guan and Ji 2012; Li et al. 2011).
A statistical analysis was applied to determine the variation of the LST–T_{2m} relationship during day- and night-time, for different seasons over six different LU/LC types. It was found that in most cases, during daytime and in cold seasons, LST is lower than T_{2m}, while in warm seasons, a reverse relationship is observed for all LU/LC types. Generally, for all seasons during night-time, LST is lower than T_{2m}. Moreover, higher differences between $T_{2mNight}$ and LST_{Day} are observed among industrial, urban area, airports, and areas covered by vegetation, bush, and trees (such as forest, agriculture, and needle leaf trees). These results indicate that LU/LC types, seasons, and temporal variation influence this relationship. A linear regression analysis, with the MODIS LST as the independent and T_{2m} as the dependent variable, was applied to analyse this relationship seasonally.
The results reveal that LST_{Day} and LST_{Night} are significantly correlated ($p = 0.0001$) with T_{2mDay} and $T_{2mNight}$. The correlation between LST_{Day} and T_{2mDay} is higher during cold seasons ($0.77 \leq r \leq 0.89$) than in warm seasons ($0.57 \leq r \leq 0.81$). Moreover, for all LU/LC types during cold seasons, a higher agreement was observed during daytime ($0.77 \leq r \leq 0.89$) than night-time ($0.47 \leq r \leq 0.86$), while a reverse relationship

was obtained for warm seasons.
In addition, it would be interesting to examine whether and how the size of the analytical unit may affect the observed relationships between LST–NDVI and LST and T_{air} in future research, considering the inhomogeneity of land surface characteristics within a grid box. In this research, we used the time series of MODIS NDVI between 2007 and 2013. To better understand how well this time series correlates with LST but we recommend to use AVHRR 16-day normalized difference vegetation index composite for compare it with MODIS NDVI. The idea behind this is to take advantage of the maximum temporal overlap between both archives from 2007 to 2013 to test how well is the time series and then we can rely on the underlying time series. The following lines answer the main questions 3, 4 and 5 presented previously. It was shown in Chapter 3 that the MODIS LST product needs a careful pre-processing for detection of cloud-contaminated values and the outliers. Such pre-processing helped to improve the correlation between the MODIS LST and T_{air} measurements. With respect to these results, in chapter 4, at attempts has been done to estimate T_{air} based on the MODIS LST product over six LC/LU class. The main question here is, how the LST product can be used to estimate T_{air} on the immediate proximity of the weather station or over a larger area. To answer this question, this chapter investigates the spatial variability of LST-T_{air} relationship by applying a sliding window of varying size over the MODIS LST grid. But before analyzing the effects of MODIS window size, the daily variability of LST and T_{air} relationship was examined by using separate LST series (over 1x1 window). LST series used in this analysis is a composite time series which includes four daily LST observations (except for cloudy days) from both the MODIS Terra and Aqua day and night overpasses (approximately at 1:30, 10:30, 13:30, 22:30) supplied in the LST L3 product.
The comparison between MODIS LST data and the T_{air} observations shows that LST_{day} and LST_{night} from both Terra and Aqua, with the mean relative bias above and under zero tended to overestimate T_{day} and underestimate T_{night} respectively as Cresswell et al. (1999) found the same result. As shown in the table, a higher relative RMSD and bias values were seen for the Aqua $LST_{daytime}$ than the Terra $LST_{daytime}$ which might be given to the fact that more solar radiation has been received at the time of the Aqua MODIS overpass later in the day. Considering the scatter-plots of LST_{night} and T_{night} from Aqua for the industrial LC type, has higher scattering than daytime observations which are more spread around the 1:1 line (Fig.4.5). This indicates the urban heat island (UHI) phenomena with RMSD=4.21°C.
Both Aqua and Terra LST_{night} underestimated the T_{night} as well except for forest. Moreover, according to RMSD from Tables(4.1-4.3) and MODIS LST from Terra, a higher RMSDs is found for industrial and airport LC types during night time which indicates the UHI phenomena (with RMSD= 4.57°C and 4.32°C respectively). Moreover Tables(4.1-4.3 show that, correlations between the MODIS LST from Terra data are generally stronger from the daytime series compared with those from the night series, except for needle leaf trees. The needle leaf tree type showed more complex correlation patterns from day and night observations. The possible reason for this, is that the values of LST recorded by MODIS observation on this particular LC type is not exactly a representative of the skin temperature of the soil, but rather affected by the

temperature near the top of the trees (canopy temperature). In addition, LST and T_{air} are correlated to a certain degree, with some drawbacks depending on factors, such as land cover type (Jin et al., 2010, Mildrexler et al., 2011).
In general, figures 4.2- 4.7 show that the time-series of the MODIS LST over six LU/LC classes were correlated individually during the day and night-time. They are highly correlated with $R^2 > 0.80$. Moreover, figures 4.2- 4.7 show that, during the warm months the LST_{day} is higher than T_{day} due to strong radiation, while as expected during the cold months LST_{day} is lower than T_{day} for almost all LC/LU. Moreover, almost for all LC/LU, the LST_{night} is close to T_{night}. As due to long wave, radiation from surface LST and T_{air} at night are closer.
Both the Terra and Aqua LST products were compared with the ground-based T_{air} as shown in the figure 4.8 and 4.9, the night-time LST datasets (MOD_{night} and MYD_{night}) and the observed T_{air} are more linearly concentrated along the fitting line than the daytime datasets. Strong correlations were observed between the night-time LST and T_{night} with minimal bias ($0.81< R^2 <0.89$, RMSE< 4.80 and MBE < 2.91°C). Specifically, the MYD_{night} tends to be more accurate for the estimation of T_{air} with lower intercepts, smaller RMSD and MBE than MOD_{night}. For T_{day}, the MYD_{day} had good agreement than MOD_{day} with lower intercept. This is most likely because the Aqua overpass time (1:30 the time when maximum temperature was recorded). However, LST from Aqua and Terra seems to be best for estimation of T_{day} among the LST products.
To sum up, the relationship between LST and T_{air} may vary with time and location because the land surface energy balance is a complex phenomenon that depends on multiple factors (e.g., cloud cover, surface roughness, wind speed and soil moisture (Prince et al., 1998; Zhang et al., 2015). In addition, the LST and T_{air} are different in principle. The satellite remotely sensed LST is a measure of the surface radiation. LST was calculated from the emissivity's surface, which is sensitive to LC, especially during daytime and another reason could be the heat capacity or specific heat of LC. However, the specific heat varies significantly from one LC to another. The variation of the difference between LST_{day} and T_{day} may be due to the different heat capacities or specific heats of LC types. The heat capacity changes with temperature, which may result in different relations at the different times even over the same LC.
Our results showed that the MODIS LST correlates best, with T_{air} measurement during the daytime. To some extent, this outcome was contradictory to the other works in the literature (e.g., Zhang et al., 2011a; Benali et al., 2012) where they have achieved a stronger correlation at night-time compared to daytime. It must be noted, though, that Benali et al. (2012) used MODIS-Terra but not MODIS-Aqua observations. Variations in the MODIS, overpasses time in its 16-day repeated cycle which enabled us to reconstruct the diurnal LST profile over a 7-year period. Although, many studies have shown a higher agreement between LST and T_{air} at night (Zhang et al. 2011a; Benali et al. 2012), this is not the case for all hours of the day or night. During some hours of the night the LST- T_{air} relationship is weaker than some hours during the day. These differences could be because of time of observation and geographical location that affecting the relationship between LST product and T_{air} and therefore, affecting the accuracy estimation of T_{air} based on LST products.
The relationship between the observed T_{night}–T_{day} and LST is not limited to a single

pixel, due to the fact that the temperature of the near-surface air mass in a given area, is influenced by many factors such as energy exchanges with the land surface over a larger area. On the other hand, the T_{air} is impressed by both the local radiation budget and air advection from the surrounding areas. Therefore, for better understanding of the spatial variability in LST- T_{air} relationship, a spatial window with a varying size is examined to discover the optimal spatial extent over which LST agrees best with the T_{air} measurements. To describe the effects of LST window-size on the LST and T_{air} relationship better, firstly, the time-series of LST from a single pixel (1x1 window size) overlapping each weather station were retrieved from the MODIS LST grid and then the LST of 3x3, 5x5, 7x7, 9x9, 11x11, 13x13 and 15x15 pixels were generated, respectively. Secondly, to determine the proper spatial window size for estimating air temperature, correlation coefficient analysis was made for different LC/LU. As shown in figure 4.10, the correlations were improved very slightly when the window size was increased from 1 to 3 pixels for daytime. The highest correlation values were achieved with 3x3 window for all LC/LU during the daytime and at the 1x1 during the night-time. Significance levels of all correlations were found to be at which can be interpreted from p-values (all p-values <0.01). According to these results, the window size was selected for all LC/LU prior to model development for day and night-time data set.
Moreover, three different methods namely SA-SVR, ANN and ANFIS were employed to estimate T_{air} during the day and night-time in Berlin using twelve variables as predictors. The performance of the three models was assessed using cross-validation with k=4 fold over different LC/LU during the day and night-time in order to shows whether the methods generalize well or not (Reunanen, 2003). All samples from each LC/LU were used in turn as the validation data set to test the model, while the remaining samples were used as the training data set to fit the model. RMSE, R^2, MBE and MAE were calculated from the measured and estimated T_{air} values to assess model performance. As shown in Tables(4.4-4.9, ANN model with three layers structure, has higher adjusted R^2 value ranged from 0.93 to 0.97, RMSE ranged from 1.83°C to 2.53°C and MAE ranged from 1.53°C to 1.94°C in test phases for all LC/LU for estimating T_{day}. The results showed that all models have similar capability in the training phase for estimating T_{night} but the ANN has a higher adjusted R^2 which ranged from 0.89 to 0.93, RMSE and ranged from 2.13°C to 2.35°C and also MAE ranged from 1.54°C to 1.84°C values in the test phase in comparison to ANFIS and SA-SVR. The bar plots of RMSE for the three methods on testing data for each LC/LU are shown in Fig 4.11-4.13, respectively. The bar plots depicted the performance of the NN model on the test data which was better than ANFIS and SA-SVR models for all LC/LU during day and night-time. As can be seen from figure 4.14 (b), the SA-SVR and NN models are more robust and stable than ANFIS model regarding their SD values (ranged from 0.03 to 0.08) during the day and night-time and we can say that, these two models are more reliable than the ANFIS model. Moreover, figure 4.15 shows Q-Q diagram of SA-SVR (left), ANFIS (middle) and ANN (right) models. Q-Q diagrams are often used to determine whether the model could extract the behaviour of the observed data (Chambers et al., 1983). As shown in figure 4.15, the models cannot estimate the high temperature for all LC/LU during the day and night-time. The weak performance of all models at high temperature are a consequence of a small

number of data in these temperatures, and this is also highly related to the study area condition (Berlin) which has a short summer and then has only a few numbers of high temperatures. In these cases, the learning algorithm of the three mentioned models have the tendency to underestimate the temperature. Therefore, the generalization of these models for the high temperature is reduced.
In addition, despite moderate to high correlations between LST and T_{air}, LST cannot be directly used for estimating air temperature due to the large difference in MBE (4.1-4.3), while by applying some additional parameters, in three models (Tables (4.4-4.9), It can be seen that the MBE was notably reduced, in all LC/LU during day and night-time.
Furthermore, prediction of long-term monthly air temperature using ANFIS and ANN had been done in the study of Kisi and Shiri (2014). They applied station latitude, longitude and altitude values as input variable to predict the long-term monthly temperature values. They found that the ANN models generally performed better than the ANFIS model in the test period and the determination coefficient of 0.921 and 0.995 were achieved by them for estimating maximum and minimum air temperature using the ANN model in different stations which were higher than ANFIS model. the results of the ANN and ANFIS models in the study of Kisi and Shiri (2014) show the RMSE values range from 1.53 to 4.20°C and 1.18°C to 9.25°C for each station, respectively. Moreover, in the study of Xu et al., (2014), they applied spatially averaged values of LST, NDVI, modified normalized difference water index (MNDWI), latitude, longitude, distance to ocean, altitude, albedo and solar radiation as predictors of T_{air} in linear regression and random forest models for estimating T_{air} in summer periods from 2003 to 2012. In their study, prior to model development, they also investigated the window size effect on the relationship between LST and T_{air}. The cross-validation results of their study show that the random forest model (MAE = 2.02°C, R^2 = 0.74) outperforms the linear regression model (MAE = 2.41°C, R^2 = 0.64) and the distribution of residuals from the random forest model slightly overestimates T_{air}, with a mean residual value of 0.09°C.
Finally, in our study, instead of estimation monthly air temperature and only using the geographical input data, we estimate air temperature during day and night. Moreover, different parameters such as NDVI, Albedo, relative humidity, wind speed, wind direction and Julian day have been taking into consideration, which are representative of seasonal changes. The satisfactory results suggested that this modelling approach is appropriate for estimating air temperature in Berlin over six different LC/LU. In addition, the results indicate that MODIS time series of LST can be successfully combined with ground measurements of temperature to produce accurate and more detailed predications of temperature during day and night time (Hengl et al., 2012). Although the air temperature estimated from satellites tends to be higher than ground-based measurement, the use of satellite remote sensing data can help to overcome the spatial problem of estimating T_{air} particularly in areas with low station density using satellite-based land surface temperature estimation and ground-based relationship between LST and air temperature. To reduce the biases in satellite-estimated air temperature, it can be effective to use retrieval method based on land surface heat budget (e.g. Kato and Yamaguchi, 2005) in future work.

The main questions 6, 7 and 8 are answer as follow: There have been number of studies focusing on the feature selection problem which are done by statistics and machine learning communities for many years. It has received more attention recently because of enthusiastic research in data mining. The aim is to construct a model that predicts/estimate well or explains the relationships in the data. Therefore, the motivation behind this research was to formulate a more efficient means of correctly selecting input variables using ANN models of environmental processes.
In this section, the exhaustive search and forward selection are considered for finding the best candidate model which can estimate the T_{air}, during day and night time for six different LC/LU. Some criteria like RMSE, R^2, AIC and BIC are considered for model selection. The AIC can be termed as a measure of the goodness of fit of any estimated statistical model. The AIC method provides a general measure of the trade-off between information gain and the complexity introduced to the modelling domain by the addition of input variables. This criterion lends itself to clear and simple interpretation and is expected to provide consistent and reliable selection for any data set. The BIC is a type of model selection among a class of parametric models with different numbers of parameters.
we applied ES feature selection method, which is to exhaustively evaluate all possible combinations of the input features, and then find the best subset. Obviously, the computational cos of exhaustive search is prohibitively high, with considerable danger of over-fitting. Hence, people resort to greedy methods, such as forward selection. Our investigation shows that the greediness of the feature selection algorithms greatly improves the efficiency, while does not corrupt the correctness of the selected feature set so that the estimation accuracy using the selected features remains satisfactory.
(According to tables 5.1, 5.2 and figure 5.3, the most important parameters for estimating the T_{2m} for different LC/LU which its AIC weight importance is greater than 0.5 are namely included LST, JD, WD and RH for day time and LST, WD, RH and AP for night time respectively. Moreover, figure 5.1 shows that, BIC always select less variable compare to AIC and RMSE because BIC penalizes larger models more strongly and so will tend to prefer smaller models in comparison to AIC. In addition, the selection results demonstrate LST has more contribution than the other variables to the measurement of Tair during day and night over different LC/LU. Table 5.5 and 5.6 show the final selected model during day and night time for six considered LC/LU based on AIC and BIC criteria. It has been found that AIC and BIC criteria tend to select the variable with Akaike weight importance of more than 0.4 and o.5, respectively. It means that BIC select variables which have strong and moderately strong relationships between a predictor variable and the response variable were associated with Akaike weight importance from 0.5 to 1(Link and Barker 2006; Bolker 2008; Richards 2005).
In addition, by considering tables 5.1-5.4 and figures 5.1, 5.2, 5.5 and 5.6, we found that, those parameters which have impact on the estimation of T_{2m} during day and night time were selected in both algorithms. Both algorithms tend to select the parameters which their sum weight of Akaike were greater than 0.5. Moreover, compared with the exhaustive search, forward selection is much cheaper. However, forward selection may suffer because of its greediness.

Moreover, it has found that the sum weight of wind speed has significant changes during day and night time on airport LC/LU because of impervious surface of this LC/LU in all season (Lin et al., 2012). The wind speed is influenced by many factors, but the presence of tree cover is single factor that distinguishes wind over a forest from wind over more open terrain, such as an airport. After sunset, due to, at night the surface of the earth cools much more rapidly than does the air above the surface, therefore, the wind speed tends to decrease. Because of this difference in cooling ability, it doesn't take long for the ground to become colder than the air above it. The air in close contact with the ground — say in the lowest 300 feet of the atmosphere — then becomes colder than the air above it. This circumstance cause to the development of what is known as a temperature inversion. Inversions dramatically alleviate the amount of mixing that occurs between different vertical layers of the atmosphere. Therefore, after sunset, when the inversion sets up, it is much harder for fast-moving air above the ground to mix down to the surface, where it could appear as a gust of wind.
Furthermore, it has been found the sum weight of RH in all LC/LU during night are higher that night-time the reason is the relative humidity depends not only upon the amount of water vapour present in the air but also on the air temperature. In fact, the relative humidity indicates how much is the percentage of saturation of air (with water vapour) for a given temperature. If it is fully saturated, then the relative humidity is 100 percent. If air holds some amount of water vapour at a particular temperature and is unsaturated, then, at a lower temperature, the same amount of water vapour may be able to saturate it. Hence, for the same amount of moisture content in the air, the relative humidity may be less for a higher temperature and more for a lower temperature. Therefore, naturally the relative humidity is less during daytime and more during night-time. In addition, AIC generally tries to find unknown model that has high dimensional reality, but the Bayesian Information Criteria comes across only true models. It can also be said that Bayesian Information Criteria is consistent whereas AIC. When AIC and BIC will present the danger of outfit under fit, respectively (Aho et al., 2014, Bolker 2008). Compared to AIC and BIC, Theil's adjusted (R^2) is another model selection criterion that is applied to select the number of repressors. Though the determination coefficient (R^2) measures the goodness-of-fit of a model, it almost always increases and never decreases with the number of repressors (according to the table 5.4 which is summary of the model selection procedure applied to the urban LC/LU). Therefore, if we consider R^2 as model selection criteria, then it would always favor larger number of variables. The adjusted R^2 corrects the problem with an adjustment to the degrees of freedom. The two other measurements namely sum of squared errors (SSE) or residual variance can be used for the goodness-of-fit. It should have been considered that minimizing the estimated residual variance is equal to maximizing R^2, therefore residual variance is redundant in cases that R^2 is applied. Cameron (1993) debates that R^2 is not an effective tool for the prevention of data mining because it will rise on the addition of any variable whose t ratio is greater than one when entered the model, although R^2 is used as a model selection criterion.
Overall, it has been found that LC/LU has a key aspect plays on the relationship between T_{air} and LST. The results show that optimal models for estimating T_{air} are

highly depend on LC/LU and time (day or night time). The specific heat capacities of different LC/LU could be a one reason for this variation. The T_{air} is highly depend on the heat transfer process, which was strongly influenced by the local radiation budget (Lin, et al 2016, Yang et al.,2017). Generally, the heat capacity is lower in barren land than forest. Therefore, T_{air} was heated much faster over barren land than forest. Vegetation could also alter latent heat flux, such as enhancing or reducing transpiration (Zeng et al 2015, Kaufmann et al, 2003), and cool the T_{air} in forests (Jeong et al., 2009, Pouteau et al., 2011, Van et al., 2013). In this study, the cooling effect was not explicitly considered because of roughly distribution of meteorological stations across different vegetation types. Therefore, it was difficult to consider the vegetation type in our models. However, land cover also affected land surface albedo, thus, the influence of land cover on estimating T_{air} was conditional and time dependent.
Moreover, another issue that we tried to answer in this study was, what is the pitfall of using the global model? It has been argued that one should make inference from a model with all the factors thought to be important (i.e., a global model). This approach would seem to be simple and avoid the complications of model selection. The serious drawback here is the lack of precision in the estimated parameters (Figure 5.2). A given data set has only a finite amount of information; each time a parameter estimate is made, the information left is reduced. Increasing the number of parameters eventually makes the fitted model unstable and uninformative. The probability of finding factors that are spurious increases. New parameters are estimated but with increasing uncertainty-this phenomenon is an aspect of the Principle of Parsimony. As we can see from Figure 5.2, by adding the irrelevant additional parameter to the model, the accuracy will be decreased.
Finally, feature selection is a process where one tries to identify the useful parameters from among a potentially large set of possible features. The task is obviously hard, and researchers have been tackling it already for decades. Solving the problem properly might today be more important than ever before, because, dataset size seem to grow faster than does the processing power of computers in many applications. In addition, there are several advantages and reasons of performing feature selection which some of them have been enlisted below:

1. Feature selection models with a smaller number of features have higher explain ability and it is easy to interpret.
2. Fewer features lead to enhanced generalization which in turn reduces over-fitting.
3. Feature selection removes data redundancy.
4. Training time (consuming time) of models with fewer features is significantly lower.
5. Models with fewer features are less prone to errors.

The two feature selection techniques have been evaluated to determine their effectiveness for reducing the number of dimensions in a dataset and for improving the accuracy of estimator which uses only these features. In this chapter, we found that, Greedy forward Search is to be a highly effective method for feature selection which has both increased accuracy and reduced the number of dimensions for the dataset

used. An exhaustive search was trailed experimentally, it was found to be very effective for reducing the number of dimensions, and marginally improved the accuracy of the estimator being used regardless the time consuming of the algorithm. In the future work, in order to have better and comprehensive understanding the relationship between LC/LU and T_{air}, applying several feature selection methods would be interesting in order to have deeper investigation. Next to it, we can also apply the deep learning in order to have higher accuracy in the estimation of T_{air}. In this research, the twelve features (as described in chapter 4) were used as important parameters for estimating T_{air}, but in the future work, other static and non-static features such as building fraction, building height, incoming radiation, solar zenith angle, sky view factor, and wall area index should have considered as potential important features which could influence the estimation accuracy as well. Next to it, in order to reduce the over-fitting problem, combinations of different feature selection algorithm can also be examined for the future work as described in the studies of Pohjalainen et al., 2012, and Saeys et al., 2008.

7 Outlook

The T_{air} is an important descriptor of terrestrial environmental conditions across the Earth and plays an important role in multiple biological and physical processes among the hydrosphere, atmosphere, and biosphere. Monitoring and understanding the trends of LST–NDVI and LST–T_{2m} are crucial in the study of regional and global climate change (Yoo et al. 2011). This study investigates both LST–NDVI and LST–T_{2m} relationships over dierent LU/LC over Berlin, during different seasons at day and night with grid size (1 km * 1km). The results indicate that different LU/LC types , seasonal variations, time of day have significantly dierent eects on LST-NDVI and LST–T_{2m} relationships in Berlin. It would be interesting to examine whether and how the size of the analytical unit may aect the observed relationships between LST–NDVI and LST and T_{air} in future research, considering the inhomogeneity of land surface characteristics within a grid box.

In our study, instead of estimation monthly air temperature and using the geographical input data, we estimate air temperature during day and night with different parameters such as NDVI, Albedo, relative humidity, wind speed, wind direction and Julian day which are representative of seasonal changes. The satisfactory results suggested that this modelling approach is appropriate for estimating air temperature in Berlin over six different LC/LU. In addition, the results indicate that MODIS time series of LST can be successfully combined with ground measurements of temperature to produce accurate and more detailed predications of temperature during day and night-time. Although the air temperature estimated from satellites tends to be higher than ground-based measurement, the use of satellite remote sensing data can help to overcome the spatial problem of estimating T_{air} particularly in areas with low station density using satellite-based land surface temperature estimation and ground-based relationship between LST and air temperature but our suggestion, for the future work is to use retrieval method based on land surface heat budget (e.g. Kato and Yamaguchi, 2005) in the future work in order to reduce the biases in satellite-estimated air temperature, which might be a solution. Next to it, we should have considered modified normalized difference water index(MNDWI), distance to river and solar radiation as predictors for estimating T_{air}(Xu. et al. (2014)) in order to have better understanding, view and analyses regarding air estimation during winter and summer periods. In our study, we applied support vector regression model (SVR) which has been widely used to solve non-linear time series problems. For tuning the three parameters of a SVR, we only applied simulated annealing algorithm but for future work, it had better to apply other optimization algorithms such as GA(genetic algorithm), particle swarm optimization (PSO), and Ant Colony Optimization(ACO) and make a comparison between these optimization methods and find the best. In this study we applied two feature selection method for estimating T_{air} during day and night time over different LC/LU using both Terra and Aqua MODIS LST products and auxiliary data from 2007–2013 in order to figure out which parameters, among 12 candidate parameters can described the relationship between LST and T_{air} and has important effect on their relationship. We should have considered other parameters such as BF(x), BH(x), Elv(x), WAI(x), LU(x)

which are referred to building fraction, building height, elevation, wall area index and land use of the given point in order to understand the importance and impact of these parameters in urban and industrial area compared to other LC/LU types. Near surface air, temperature (NSAT) is a primary descriptor of terrestrial environmental conditions. In recent decades, many efforts have been made to develop various methods for obtaining spatially continuous NSAT from gauge or station observations. NSAT has significant role in energy and water exchanges between the land surface and atmosphere (Guan et. al, 2013) and it is the most important component of global climate change which is sensitive to local anthropogenic disturbance (Hansen et. al, 2006). Therefore, the availability of NSAT with a high spatial resolution is necessary for several applications such as hydrology, meteorology, and ecology (Zhu et al.,2013, Yang et al.,2012, Ge et al.,2014, Fu et al.,2011). For decades, many efforts have been made to obtain spatial distributions of various NSAT variables based on the point station measurements, including annual maximum/minimum/mean NSAT(Cristóbal et al.,2008), monthly maximum/minimum/mean NSAT (Ninyerola et al.,2007, El Kenawy et al.,2010, Chen et al.,2015, Evrendilek et al.,2012, Bennie et al.,2010), daily maximum/minimum/mean NSAT (Pape et al., 2009, Gholamnia et al.,2017, Zhou et al.,2017, Good 2015, Peón et al.,2014, Sun et al.,2005), and instantaneous NSAT (Niclos et al., 2014, GLass, 2004). These NSAT retrieval methods can be divided into three groups: (1) spatial interpolation method (Hou et al., 2013), (2) physical-based method (Niclos et al., 2014, Stahl et al.,2006), and (3) regression analysis method (Cristóbal et al., 2008). Considering the high spatial autocorrelation of NSAT, several spatial interpolation methods have been employed to generate spatially continuous NSAT from point station measurements, including inverse distance weighting (IDW), Spline, Kriging, and even more sophisticated methods, such as co-Kriging and elevation-de-trended Kriging techniques (Benavides et al., 2007, Duhan et al., 2013, Vogt et al., 1997). The performance of interpolation methods is highly dependent on the spatial density and distribution of weather stations (Stisen et al., 2007).

Satellite remote sensing provides the ability to extract spatially continuous information of land surface characteristics such as LST and the vegetation index (VI), which are closely relative to NSAT. Many reserachers tried to estimate NSAT using satellite data (Sun et al. 2011, Niclos et al.,2014, Nieto et al., 2011, Kawashima et al., 2000). Cheng et al. (2008), Fu et al. (2011), and Zhu et al. (2013) tried to predict NSAT based on the simple correlation between the NSAT and LST. Multiple linear regression (MLR) analysis using both remote sensing and geographical variables, including LST, VI, latitude, altitude, and so on, as predictors was performed to model NSAT (Cristóbal et al., 2008 , Ninyerola et al., 2007, Peón et al., 2014). However, a global regression analysis may lose local details that can be significant if the relationship is spatially non-stationary. Geographically weighted regression (GWR) is a local modelling technique for analysing spatial analysis, and allows the regression model parameters to vary in space (Foody et al., 2003, Peng et al., 2011). The GWR model was employed by Chen et al. for estimating monthly and eight-days NSAT in China (Chen et al., 2015). Many researches have made contributions to assess the performance of various predicting NSAT models in different regions. Peng et al. interpolated the monthly and annual

NSAT in the Jiangsu province, China, using the IDW, Spline, Kriging, and Co-Kriging models, and the result proved that the Kriging model has a much higher precision than the IDW and Spline models, and that the Co-Kriging model is slightly better than the Kriging model (Zhao et al.,2005). GLASS et al.'s study showed that interpolation models (i.e., the Kriging model), regardless of whether or not satellite data are included, are consistently superior to MLR models, and the Kriging model without satellite data performed similarly to that with satellite data under more general conditions (Hou et al.,2013). Zhao et al. estimated the NSAT in the southern Qilian mountains, China, in which the weather stations are sparse, and the result indicated that the accuracy of the MLR model is higher than that of spatial interpolation models, and the Spline model shows the worst result (Zhao et al.,2005). What can be done next in Berlin area is to estimate near surface air temperature using spatial interpolation model based on mesh free approximation with our proposed method which the aim is to find optimized coefficients for $BF(x)$, $BH(x)$, $Elv(x)$, $WAI(x)$, $LU(x)$ which are referred to building fraction, building height, elevation, wall area index and land use of the given point, respectively.

The reason is that in all estimator models which we applied in this study we did not consider the impact of following factors in formation of UHI micro-climate changes. Moreover, it might be a question why we should consider the mesh free as an approximator? The reason is in many engineering application and science area, computation with high-dimensional data is an important issue. Many traditional numerical methods can either not handle such problems at all, or are restricted to very special situations. Mesh free methods are often better solution to cope with changes in the geometry of the domain of interest than classical discretization techniques. Another advantage is its in-dependency from mesh generation which is the most time-consuming part of any mesh-based numerical simulation.

7.1 The Scattered Data Interpolation Problem (general description)

In many scientific disciplines, one faces the following problem: We are given a set of data (measurements, and locations at which these measurements were obtained), and we want to find a general rule which allows us to deduce information about the process we are studying also at locations different from those at which we obtained our measurements. Thus, we are trying to find a function P_f which is a "good" fit to the given data. There are many ways to decide what we mean by "good", and the only criterion we will consider now is that we want the function P_f to exactly match the given measurements at the corresponding locations. This approach is called interpolation, and if the locations at which the measurements are taken do not lie on a uniform or regular grid, then the process is called scattered data interpolation. To give a precise definition we assume that the measurement locations (or data are labeld $x_j, j = 1, ..., N$, and the corresponding measurements (or data values) are

called y_i We will use X to denote the set of data sites and assume that $x \subset \Omega$ or some region, and Ω in R^s. we will restrict our discussion to scalar-valued data, where $y_j \subset R$. However, much of the following can be generalized easily to problems with vector-valued data. Moreover, we will assume that the data are obtained by sampling some (unknown) function at the data sites $y_i = f(x_i), j = 1, ..., N$. Our notation P_f for the interpolating function emphasizes the connection between the interpolant and the data function. We are now ready for a precise formulation of the scattered data interpolation problem.

Problem 1.1 (Scattered Data Interpolation):Given data $(x_j, y_i), j = 1, ..., N$ with $x_j \in R^s, y_j \in R^s$ find a (continuous) function P_f such that $P_f(x_j) = y_i, j = 1, ..., N$. The fact that we allow x_j to lie in an arbitrary s-dimensional space R^s means that the formulation of Problem 1.1 allows us to cover many different types of applications. If s = 1 the data could, e.g., be a series of measurements taken over a certain time period, therefore the data sites x_j would correspond to given time instances. For $s = 2$ we can think of the data being obtained over a planar region, and so x_j corresponds to the two coordinates in the plane. A convenient and common approach to solving the scattered data problem is to make the assumption that the function P_f is a linear combination of certain basis functions B_k, i.e.:

$$P_f = \sum_{k=1}^{N} C_k B_k, x \in R^s \tag{7.1}$$

Solving the interpolation problem under this assumption leads to a system of linear equations of the form:

$$AC = y \tag{7.2}$$

where the entries of the interpolation matrix A are given by :

$$A_{jk} = B_{xj}, j, k = 1, ..., N, C = [C_1, ..., C_N]^T, y = [y_1, ..., y_N]^T \tag{7.3}$$

Problem 1.1 will be well-posed, i.e., a solution to the problem will exist and be unique, if and only if the matrix A is non-singular. In the univariate setting it is well known that one can interpolate to arbitrary data at N distinct data sites using a polynomial of degree $N - l$.

7.2 Scattered Data Interpolation with more General Polynomial Precision and description of method

Now, we can construct and modify the assumption on the form 1.1 of the solution to the scattered data interpolation Problem 1.1 by adding certain linear polynomials to the expansion. Hereafter P_f is considered as $T(X, t)$ which it is assumed to be:

$$T(X, t) = \sum_{i=1}^{n} \alpha_i W(X, x_i) T(x_i, t) + \sum_{i=1}^{n} \beta_i W(X, x_i) RH(x_i, t) \tag{7.4}$$

Where T is temperature at time t and location $X, i = 1, ..., n$ is the number of observation locations. The T and RH are temperature and relative humidity at time t and location x_i, respectively. The $W(X, x_i)$ and P_x are defined as follow:

$$T(X,t) = \frac{e^{-p(x)\|X-X_i\|^2}}{\sum_{i=1}^{n} e^{-p(x)\|X-X_i\|^2}} \tag{7.5}$$

$$P(x) = \begin{vmatrix} BF(x) \\ BH(x) \\ ELV(x) \\ WAI(x) \\ (LU(x) = LCT_1) \\ (LU(x) = LCT_2) \\ (LU(x) = LCT_3) \\ (LU(x) = LCT_4) \\ (LU(x) = LCT_5) \\ (LU(x) = LCT_6) \end{vmatrix} \begin{vmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \\ p_6 \\ p_7 \\ p_8 \\ p_9 \\ p_{10} \end{vmatrix} \tag{7.6}$$

Where $BF(x)$, $BH(x)$,$ELV(x)$,$WAI(x)$,$LU(x)$ are referred to Building Fraction, Building Height, Elevation, Wall Area Index and Land Use of the given point, respectively. LCT_i is addressed to one of six different land use/class that are considered for this study. The p-coefficient $p_1, ..., p_{10}$ must be optimized in order to get good result in our interpolation problem. The fmin-search, function of MATLAB can be considered for find the minimum of a scalar function of several variables starting at an initial estimate. This is generally referred to as unconstrained non-linear optimization coming back to the scattered data problem. By solving the following linear system, we will be able to find the coefficients (α_i, β_i):

$$P(x) = \begin{vmatrix} \\ W(x_1,x_1),...,W(x_1,x_n),W(x_1,x_1),...,W(x_1,x_n) \\ W(x_2,x_1),...,W(x_2,x_n),W(x_2,x_1),...,W(x_2,x_n) \\ ..... \\ ..... \\ ..... \\ ..... \\ ..... \\ ..... \\ W(x_n,x_1),...,W(x_n,x_n),W(x_n,x_1),...,W(x_n,x_n) \end{vmatrix} \begin{vmatrix} \\ \alpha_1 \\ \alpha_2 \\ . \\ . \\ \alpha_n \\ \beta_1 \\ \beta_2 \\ . \\ . \\ \beta_n \end{vmatrix} = \begin{vmatrix} \\ T(x_1,t_1) \\ T(x_2,t_1) \\ . \\ . \\ . \\ . \\ . \\ . \\ . \\ T(x_n,t_1) \end{vmatrix} \tag{7.7}$$

Where $T(X_i, t_i)$ is referred to the temperature at given time t_i and location x_i. Solving the interpolation problem under this assumption leads to a system of linear equation of the form:

$$AX = y \tag{7.8}$$

Where the entries of the interpolation matrix A is given by equation 4, $X = |\alpha_1, ..., \alpha_n, \beta_1, ..., \beta_n|^T$ and $y = |T(x_1, t_1), ..., T(x_n, t_n)|^T$ and then X is calculated as follow:

$$(A^T A)^+ A^T y = X \tag{7.9}$$

Cross-validation can be used to evaluate the generalizability of a model for estimating the air temperature.

Reference

Asrar, G., Fuchs, M., Kanemasu, E. T., & Hatfield, J. L. 1984. Estimating absorbed photosynthetic radiation and leaf area index from spectral reflectance in wheat. Agronomy Journal, 76, 300–306.

Aarts, E.H.L., Korst, J. 1989. Simulated Annealing and Boltzmann Machines: A Stochastic Approach to Combinatorial Optimization and Neural Computing. Wiley.

Aarts, E.H.L., Lenstra, J.K. 1997. Local Search in Combinatorial Optimization.Wiley, Chichester.

Abramson, D., Krishnamoorthy, M., Dang, H. 1999. Simulated annealing cooling schedules for the school timetabling problem. Asia-Pac. J. Oper. Res. 16, 1–22.

Abdel-Aal RE. 2004. Hourly temperature forecasting using abductive networks. Engineering Applications of Artificial Intelligences 17:543–556.

Abraham, A. 2005. Adaptation of Fuzzy Inference System Using Neural Learning. in Nedjah, Nadia; de Macedo Mourelle, Luiza, Fuzzy Systems Engineering. Theory and Practice, Studies in Fuzziness and Soft Computing, 181, Germany: Springer Verlag, pp. 53–83, doi:10.1007/11339366-3.

Ackerman, S.A.; Holz, R.E.; Frey, R.; Eloranta, E.W.; Maddux, B.C.; McGill, M. 2008. Cloud detection with MODIS. Part II: Validation. J. Atmos. Ocean. Technol, 25, 1073–1086.

A. J. Smola and B. Schölkopf. 2004. A tutorial on support vector regression. Statistics and Computing, vol. 14, no. 3, pp. 199–222.

Anily, S., Federgruen, A. 1987. Simulated annealing methods with general acceptance probabilities. J. Appl. Probab. 24, 657–667.

Allen, D.M., 1974. The relationship between variable selection and data augmentation and a method for prediction. Technometrics 16, 125-127.

Aldersley, A., S. J. Murray, and S. E. Cornell. 2011. Global and Regional Analysis of Climate and Human Drivers of Wildfire. Science of the Total Environment 409:

3472–3481. doi:10.1016/j. scitotenv.2011.05.032.

Allen RG, Pereira L, Raes D, Smith M. 2006. Crop evapotranspiration. FAO Irrigation and Drainage, vol. 56. FAO, Rome.

Atitar, M.; Sobrino, J.A, 2009. A split-window algorithm for estimating LST from Meteosat 9 data: Test and comparison with in situ data and MODIS LSTs. IEEE Geosci. Remote Sens. Lett. 6, 122–126.

Armstrong JS, Collopy F. 2002. Error measures for generalizing about forecasting methods: empirical comparisons. Int J Forecast ; 8(1):69–80.

Al-Alawi, S., Abdul-Wahab, S., Bakheit, C., 2008. Combining principal component regression and artificial neural networks for more accurate predictions of ground-level. Environmental Modelling and Software 23 (4), 396–403.

Akaike, H., 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control AC-19, 716-723.

Alp, M., Cigizoglu, H., 2007. Suspended sediment load simulation by two artificial neural network methods using hydrometeorological data. Environmental Modelling and Software 22 (1), 2–13.

Aho, K., Derryberry, D., & Peterson, T. 2014. Model selection for ecologists: the worldviews of AIC and BIC. Ecology, 95(3), 631-636.

Ben-Ameur, W. 2004.Computing the initial temperature of simulated annealing.Comput.Optim.Appl 29,369385.DOI:10.1023/B:COAP.0000044187.23143.bd.

Boegh, E., H. Soegaard, H. Hanan, P. Kabat, and L. Lesch 1998. A remote sensing study of the NDVI-T_s relationship and the transpiration from sparse vegetation in the Sahel based on high resolution data, Remote Sens. Environ., 69, 224–240.

Berry, M.J.A., and Linoff, G. 1997. Data Mining Techniques, NY: John Wiley & Sons.Blum, A., 1992, Neural Networks in C++, NY: Wiley.

Benali, A., Carvalho, A., Nunes, J., Carvalhais, N., and Santos, A. 2012. Estimating air surface temperature in Portugal using MODIS LST data. Remote Sensing of Environment, 124(0):108–121, doi: 10.1016/j.rse.2012.04.024.

Benavides, R., F. Montes, A. Rubio, and K. Osoro. 2007. Geostatistical Modelling of Air Temperature in a Mountainous Region of Northern Spain. Agricultural and Forest Meteorology 146: 173–188. doi:10.1016/j.agrformet.2007.05.014.

Beal, D. 2005. Selecting the best multiple liner regression model for multivariate data using information criteria.

Beier, C.M.; Signell, S.A.; Luttman, A.; DeGaetano, A.T, 2012. High-Resolution Climate Change Mapping with Gridded Historical Climate Products. Landsc. Ecol., 27, 327–342.

Brunel, J. 1989. Estimation of sensible heat flux from measurements of surface radiative temperature and air temperature at two meters: Application to determine actual evaporation rate. Agricultural and Forest Meteorology, 46(3):179 – 191, doi: 10.1016/0168-1923(89)90063-4.

B. Schölkopf and A. J. Smola, Learning with kernels [Ph.D. Thesis], GMD, Birlinghoven, Germany, 1998.

B. Ustün, W. J. Melssen, M. Oudenhuijzen, and L. M. C. Buydens, 2005. Determination of optimal support vector regression parameters by genetic algorithms and simplex optimization. Analytica Chimica Acta, vol.544, no. 1-2, pp. 292–305.

Bilgili M, Sahin B. 2010. Prediction of long-term monthly temperature and rainfall in Turkey. Energy Sources, Part A. 32: 60–71.

Betsill, M.M, 2001. Mitigating climate change in US cities: Opportunities and obstacles. Local Environ. 6,393–406. https://doi.org/10.1080/13549830120091699.

Bose, N.K. and Liang, P., 1996. Neural Network Fundamentals with Graphs, Algorithms and Applications. McGRaw-Hill Series in Electrical and Computer Engineering. McGraw-Hill, New York.

Bodri, L. and Cermak, V., 2000. Prediction of extreme precipitation using a neural network: application to summer flood occurrence in Moravia. Advances in Engineering Software, 31: 311-321.

Boger, Z., and Guterman, H., 1997, Knowledge extraction from artificial neural network models.” IEEE Systems, Man, and Cybernetics Conference, Orlando, FL, USA.

Burnham, K. P., and D. R. Anderson. 2001. Kullback-Leibler information as a basis for strong inference in ecological studies. Wildlife Research 28 :111-119.

Burnham, K. P., and D. R. Anderson., 2002. Model Selection and Multimodel Inference: a practical information-theoretic approach, 2nd edition. Springer-Verlag, New York.

Burnham, K. P., and D. R. Anderson., 2004. Multimodel Inference: understanding AIC and BIC in model selection. Sociological Methods and Research, 33,261-304.

Burnham, K.P., Anderson, D.R. & Huyvaert, K.P. , 2011. AIC model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons. Behavioral Ecology and Sociobiology, 65, 23–35. doi:10.1007/s00265-010-1029-6.

Bustos, E.; Meza, F.J., 2014. A Method to Estimate Maximum and Minimum Air Temperature Using MODIS Surface Temperature and Vegetation Data: Application to the Maipo Basin, Chile. Theor. Appl. Climatol. Doi: http://dx.doi.org/10.1007/s00704-014-1167-2.

Bowden, G.J., 2003. Forecasting Water Resources Variables using Artificial Neural Techniques. Ph.D., University of Adelaide.
Colombi, A.; De Michele, C.; Pepe, M.; Rampini, A., 2007. Estimation of daily mean air temperature from MODIS LST in Alpine areas. EARSeL eProc. 2007, 6, 38–46.

Bennie, J.; Wiltshire, A.; Joyce, A.; Clark, D.; Lloyd, A.; Adamson, J.; Parr, T.; Baxter, R.; Huntley, B., 2010. Characterising inter-annual variation in the spatial pattern of thermal microclimate in a uk upland using a combined empirical–physical model. Agric. For. Meteorol. 2010, 150, 12–19.

Carlson, T. N., Gillies, R. R., & Perry, E. M. 1994. A method to make use of thermal infrared temperature and NDVI measurements to infer surface soil water content and fractional vegetation cover. Remote Sensing Reviews, 9, 161−173.

Carlson, T. N., Gillies, R. R., & Schmugge, T. J. 1995a. An interpretation of methodologies for indirect measurement of soil water content and fractional vegetation cover. Agricultural and Forest Meteorology, 77, 191−205.

Carlson, T. N., Taconet, O., Vidal, A., Gillies, R. R., Olioso, A., & Humes, K. 1995b. An overview of the workshop on thermal remote sensing held at La Londe les Maures, France, Septe September 20–24, 1993. Agricultural and Forest Meteorology, 77(3),

141−151.

Carlson, T. N.,& Ripley, D. A. 1997. On the relation between NDVI, fractional vegetation cover, and leaf area index. Remote Sensing of Environment, 62, 241−252.

Carlson, T. N., & Arthur, S. T. 2000. The impact of land use/land cover changes due to urbanization on surface microclimate and hydrology: A satellite perspective. Global and Planetary Change, 25, 49−65.

Cassels, V., Sobrino, J. A., & Coll, C. 1992a. On the use of satellite thermal data for determining evapotranspiration in partially vegetated areas. International Journal of Remote Sensing, 13, 2669– 2682.

Carnahan, W. H., & Larson, R. C. 1990. An analysis of an urban heat sink.Remote Sensing of Environment, 33, 65–71.

Campbell, J. B. 2002. Introduction to Remote Sensing. (3rd ed.). New York: The Guilford Press.

Cheng, K.; Su, Y.; Kuo, F.; Hung, W.; Chiang, J., 2008. Assessing the effect of landcover changes on air temperature using remote sensing images—A pilot study in northern Taiwan. Landsc. Urban Plan, 85, 85–96.

Chen, F.; Liu, Y.; Liu, Q.; Qin, F., 2015.A statistical method based on remote sensing for the estimation of air temperature in China. Int. J. Clim, 35, 2131–2143.

Cleveland, W. S. and Loader, C. L. 1996. Smoothing by local regression: Principles and methods, in Statistical Theory and Computational Aspects of Smoothing, W. Haerdle and M. G. Schimek (eds.), Springer (New York), pp. 10-49.

Cristóbal, J.; Ninyerola, M.; Pons, X, 2008. Modeling air temperature through a combination of remote sensing and gis data. J. Geophys. Res. 2008, 113, D13106.

Cristóbal, J.; Ninyerola, M.; Pons, X.; Pla, M., 2006. Improving air temperature modelization by means of remote sensing variables. In Proceedings of the 2006 IEEE International Symposium on Geoscience and Remote Sensing, pp. 2251–2254.

Curtis, P. C, Jr. 1959 n-parameter families and best approximation, Pacific J., pp. 1013-1027.

C. A. Doukim, J. A. Dargham, and A. Chekima. 2010. Finding the number of hidden neurons for an MLP neural network using coarse to fine search technique. in Proceedings of the 10th International Conference on Information Sciences, Signal Processing and their Applications (ISSPA '10), pp. 606–609.

Charles Thurow, 1983. Improving street climate through urban design, American Planning Association.

Cabrera HM., 2002. Respuestas ecofisiológicas de plantas en ecosistemas de zonas con clima mediterráneo y ambientes de altamontaña. Rev Chil Hist Nat 75:625–637.

Carlson TN, CapehartWJ,Gillies RR., 1995. Anewlook at the simplified method for remote sensing of daily evapotranspiration. Remote Sens Environ 54:161–167.

Cameron, S., 1993. Why is the R squared adjusted reported. Journal of Quantitative Economics 9, 183-186.

Caselles, V., Coll, C., Valor, E., & Rubio, E. 1995. Mapping land surface emissivity using AVHRR data: Application to La Mancha, Spain. Remote Sensing Reviews, 12, 311−3330.

Czajkowski, K. P., Goward, S. N., Stadler, S., & Walz, A., 2000. Thermal remote sensing of near surface environmental variables: Application over the Oklahoma Mesonet. The Professional Geographer, 52(2), 345–357.

Cresswell, M. P., Morse, A. P., Thomson, M. C., and Connor, S. J. 1999. Estimating surface air temperatures, from Meteosat land surface temperatures, using an empirical solar zenith angle model.International Journal of Remote Sensing,20(6):1125–1132,doi: 10.1080/014311699212885.

Carrega, P. 1995. A Method for the Reconstruction of Mountain Air Temperatures with Automatic Cartographic Applications. Theoretical and Applied Climatology 52: 69–84. doi:10.1007/BF00865508.

Chen, Y.; Randerson, J.T.; Morton, D.C.; DeFries, R.S.; Collatz, G.J.; Kasibhatla, P.S.; Giglio, L.; Jin, Y.; Marlier, M.E. 2011. Forecasting Fire Season Severity in South America Using Sea Surface Temperature Anomalies. Science, 334, 787–791.

Cigizoglu HK., 2003. Estimation, forecasting and extrapolation of flow data by artificial neural networks. Hydrological Sciences Journal.48(3): 349–361.

Cristóbal, J., M. Ninyerola, and X. Pons., 2008. Modeling Air Temperature through a Combination of Remote Sensing and GIS Data. Journal of Geophysical Research 113: D13106. doi:10.1029/2007JD009318.

Chartzoulakis K, Psarras G, 2005. Global change effects on crop photosynthesis and production in Mediterranean: the case of Crete, Greece. Agric Ecosyst Environ 106:147–157.

Chan, J. C.-W., and D. Paelinckx., 2008. Evaluation of Random Forest and Adaboost Tree-Based Ensemble Classification and Spectral Band Selection for Ecotope Mapping Using Airborne Hyperspectral Imagery. Remote Sensing of Environment 112: 2999–3011. doi:10.1016/j. rse.2008.02.011.

Chai T, Draxler RR., 2014. Root mean square error (RMSE) or mean absolute error (MAE) –Arguments against avoiding RMSE in the literature. Geosci Model Dev, 7(3):1247–50.

Cooch, E., and G. White., 2001. Program MARK: Analysis of data from marked individuals, a gentle introduction, 2nd edition.http://www.cnr.colostate.edu/ gwhite/-mark/mark.html.

C. Liu, D. Jiang, and W. Yang., 2014. Global geometric similarity scheme for feature selection in fault diagnosis," Expert Systems with Applications, vol. 41, issue 8, pp. 3585–3595.

Darsono, S., Labadie, J., 2007. Neural-optimal control algorithm for real-time regulation of in-line storage in combined sewer systems. Environmental Modelling and Software 22 (9), 1349–1361.

Dormann, C.F.; Elith, J.; Bacher, S.; Buchmann, C.; Carl, G.; Carré, G.; Marquéz, J.R.G.; Gruber, B.; Lafourcade, B.; Leitão, P.J. Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. Ecography 2013, 36, 027–046.

De Gooijer, J.G., Abraham, B., Gould, A., Robinson, L., 1985. Methods for determining the order of an autoregressive- moving average process: a survey. International Statistical Review 53, 301-329.

Dawson, C.W., Wilby, R.L., 2001. Hydrological modelling using artificial neural networks. Progress in Physical Geography 25 (1), 80–108.

Dutot, A., Rynkiewicz, J., Steiner, F., Rude, J., 2007. A 24-h forecast of ozone peaks and exceedance levels using neural classifiers and weather predictions. Environmental Modelling and Software 22 (9), 1261–1269.

Doukim, C.A., Dargham, J.A., Chekima, A., 2010. Finding the number of hidden neurons for an MLP neural network using coarse to fine search technique, Information Sciences Signal Processing and their Applications (ISSPA). 10th International Conference on. IEEE, pp. 606–609.

Dodson, R., and D. Marks. 1997. Daily Air Temperature Interpolated at High Spatial Resolution over a Large Mountainous Region. Climate Research 8: 1–20. doi:10.3354/cr008001.

Duhan, D., A. Pandey, K. P. S. Gahalaut, and R. P. Pandey., 2013. Spatial and Temporal Variability in Maximum, Minimum and Mean Air Temperatures at Madhya Pradesh in Central India. Comptes Rendus Geoscience 345: 3–21. doi:10.1016/j.crte.2012.10.016.

De Bruin, H.A.R.; Trigo, I.F.; Jitan, M.A.; TemesgenEnku, N.; van der Tol, C.; Gieske, A.S.M., 2010. Reference crop evapotranspiration derived from geo-stationary satellite imagery: A case study for the Fogera flood plain, NW-Ethiopia and the Jordan Valley, Jordan. Hydrol. Earth Syst. Sci., 14, 2219–2228.

De Forest, E. L., 1873. On some methods of interpolation applicable to the graduation of Irregular series, Annual Report of the Board of Regents of the Smithsonian Institution for 1871, pp. 275-339.

Dyn, N., 1987. Interpolation of scattered data by radial functions, in Topics in Multivariate Approximation, C. K. Chui, L. L. Schumaker, and F. Utreras (eds.), Academic Press (New York), pp. 47-61.

D.B. Crawley, B. Drury., 2008. Estimating the impacts of climate change and urbanization on building performance, Journal of Building Performance Simulation 1 (2) 91–115.

D. Reidel; Kluwer, Dordrecht, Boston, Norwell, MA. Vogt, J., Viau, A., & Paquet, F. 1997. Mapping regional air temperature fields using satellite derived surface skin temperatures. International Journal of Climatology, 17,1559–1579.

E.T. Mansura, R. Mendelsohnb, W. Morrisonc., 2008. Climate change adaptation: a study of fuel choice and consumption in the US energy sector, Journal of Environmental Economics and Management 55 175–193.

Eglese, R.W., 2014. Simulated annealing: A tool for operational research. Eur. J. Oper. Res. 46, 271–281. https://doi.org/10.1016/j.mineng.2014.08.003.

Evrendilek, F., Karakaya, N., Gungor, K., and Aslan, G., 2012. Satellite-based and mesoscale regression modeling of monthly air and soil temperatures over complex terrain in Turkey. Expert Systems with Applications, 39(2):2059–2066, doi: 10.1016/j.eswa.2011.08.023.

Emamifar, S.; Rahimikhoob, A.; Noroozi, A. Daily mean air temperature estimation from MODIS land surface temperature products based on M5 model tree. Int. J. Climatol. 2013, 33, 3174–3181. DOI: 10.1002/joc.3655.

El Kenawy, A.; López-Moreno, J.I.; Vicente-Serrano, S.M.; Morsi, F., 2010. Climatological modeling of monthly air temperature and precipitation in egypt through gis techniques. Clim. Res., 42, 161–176.

Evrendilek, F.; Karakaya, N.; Gungor, K.; Aslan, G., 2012. Satellite-based and mesoscale regression modeling of monthly air and soil temperatures over complex terrain in turkey. Expert Syst. Appl., 39, 2059–2066.

Friedl, M. A., & Davis, F. W. 1994. Sources of variation in radiometric surface temperature over a tall grass prairie. Remote Sensing of Environment, 48, 1–17.

Friedl, M. A. 2002. Forward and inverse modeling of land surface energy balance using surface temperature measurements. Remote Sensing of Environment, 79, 344–354.

Fotheringham, A.S.; Brunsdon, C.; Charlton, M., 2003. Geographically Weighted Regression: The Analysis of Spatially Varying Relationships; John Wiley & Sons: New York, NY, USA, pp. 272–275.

Foody, G., 2003. Geographical weighting as a further refinement to regression modelling: An example focused on the ndvi—Rainfall relationship. Remote Sens. Environ., 88, 283–293.

Florio, E. N., Lele, S. R., Chi Chang, Y., Sterner, R., and Glass, G. E., 2004. Integrating AVHRR satellite data and NOAA ground observations to predict surface air temperature: a statistical approach. International Journal of Remote Sensing, 25(15):2979–2994, doi: 10.1080/01431160310001624593.

Fleischer, M.A., 1995. Simulated annealing: Past, present, and future. In: Alexopoulos, C., Kang, K., Lilegdon, W.R., Goldsman, D., (eds.) Proceedings of the 1995 Winter Simulation Conference, pp. 155–161. IEEE Press, Arlington, Virginia.

Fnaiech, F., Fnaiech, N., Najim, M., 2001. A new feed forward neural network hidden layer neuron pruning algorithm. IEEE International Conference on Acoustics, Speech, and Signal Processing, Proceedings. (ICASSP '01).

Fu, G.; Shen, Z.X.; Zhang, X.Z.; Shi, P.L.; Zhang, Y.J.; Wu, J.S., 2011. Estimating air temperature of an alpine meadow on the Northern Tibetan Plateau using MODIS land surface temperature. Acta Ecol. Sin., 21, 8–13.

Frederick, K.L.; Edward J.T.; Dennis G.T., 2006. The Atmosphere: An Introduction to Meteorology, 10th ed.; Prentice Hall: Upper Saddle River, NJ, USA.

Finardi, S., De Maria, R., D'Allura, A., Cascone, C., Calori, G., Lollobrigida, F., 2008. A deterministic air quality forecasting system for Torino urban area, Italy. Environmental Modelling and Software 23 (3), 344–355.

F. Korn, B. Pagel, and C. Faloutsos, 2011. On the dimensionality curse and the self-similarity blessing. IEEE Trans. Knowl.Data Eng. vol. 13, no. 1, pp. 96–111.

F. Benoît, M. van Heeswijk, Y. Miche, M. Verleysen, and A. Lendasse, 2013. Feature selection for nonlinear models with extreme learning machines, Neurocomputing, vol. 102, pp. 111–124.

Geisser, S., 1975. The predictive sample reuse method with applications. Journal of the American Statistical Association, 70, 320-328.

Goward, S. N., Cruickshanks, G. D., and Hope, A. S., 1985. Observed relation between thermal emission and reflected spectral radiance of a complex vegetated landscape. Remote Sensing of Environment, 18(2):137–146, doi: 10.1016/0034-4257(85)90044-6.

Goward, S.N.; Dye, D., 1997. Global biospheric monitoring with remote sensing. In The Use of Remote Sensing in Modeling Forest Productivity at Scales From the Stand to the Globe; Gholtz, H.L., Nakane, K., Shimoda, H., Eds.; Kluwer Academic: New York, NY, USA, pp. 241–272.

Ghent, D.; Kaduk, J.; Remedios, J.; Ardö, J.; Balzter, H., 2010. Assimilation of land surface temperature into the land surface model JULES with an ensemble Kalman filter. J. Geophys. Res. Atmos.

Gao, H.; Tang, Q.; Shi, X.; Zhu, C.; Bohn, T.J.; Su, F.; Sheffield, J.; Pan, M.; Lettenmaier, D.P.; Wood, E.F., 2014.. Water budget record from Variable Infiltration Capacity (VIC) Model. In Algorithm Theoretical Basis Document for Terrestrial Water Cycle Data Records; Availableonline:http://grid1.cos.gmu.edu:8090/OPeNDAPClient.

Galvão, C., Valença, M., Vieira, V., Diniz, L., Lacerda, E., Carvalho, A., Ludermir, T., 1999. Sistemas inteligentes: Aplicações a recursos hídricos e ciências ambientais. Porto Alegre, UFRGS-ABRH, 246 pgs.

Gardner, M.W. and Dorling, S.R., 1998. Artificial neural networks (the multilayer perceptron) - A review of applications in the atmospheric sciences. Atmospheric Environment, 32(14-15): 2627-2636.

G. Panchal, A. Ganatra, Y. P. Kosta, and D. Panchal., 2011. Behaviour analysis of multilayer perceptrons with multiple hidden neurons and hidden layers. International Journal of Computer Theory and Engineering, vol. 3, no. 2, pp. 332–337.

Gallo, K. P., McNab, A. L., Karl, T. R., Brown, J. F., Hood, J. J., & Tarpley, J. D., 1993. The use of NOAA AVHRR data for assessment of the urban heat island effect. Journal of Applied Meteorology, 32(5), 899−908.

Gallo, K. P., & Owen, T. W., 1999. Satellite based adjustments for the urban heat island temperature bias. Journal of Applied Meteorology, 38, 806−813.

Gillies, R. R., & Carlson, T. N., 1995. Thermal remote sensing of surface soil water content with partial vegetation cover for incorporation into climate models. Journal of Applied Meteorology, 34, 745– 756.

Gillies, R. R., Carlson, T. N., Cui, J., Kustas, W. P., & Humes, K. S., 1997. A verification of the triangle method for obtaining surface soil water content and energy fluxes from remote measurements of the Normalized Difference Vegetation index (NDVI) and surface radiant temperature. International Journal of Remote Sensing, 18, 3145– 3166.

Goward, S. N., Cruickshank, G. D., & Hope, A. S., 1985. Observed relation between thermal emission and reflected spectral radiance of a complex vegetated landscape. Remote Sensing of Environment, 18, 137−146.

Goward, S. N., Xue, Y., & Czajkowski, K. P., 2002. Evaluating land surface moisture conditions from the remotely sensed temperature/vegetation index measurements: an exploration with the simplified simple biosphere model. Remote Sensing of Environment, 79, 225−242.

Goetz, S. J., 1997. Multisensory analysis of NDVI, surface temperature and biophysical variables at a mixed grassland site. International Journal of Remote Sensing, 18, 71–94.

Gurney, R. J., J. p. Ormsby, and D. K. Hall, 1983: Observed relation between thermal emission and reflected spectral radiance of complex vegetated landscape. Permafourth: Fourth Int. Conf., Fairbanks, AK, University of Alaska and National Academy of Science, 401-404.

Gholamnia, M., Alavipanah, S.K., Darvishi Boloorani, A., Hamzeh, S., Kiavarz, M. Diurnal air temperature modeling based on the land surface temperature. Remote Sens. 2017, 9, 915.

Good, E. Daily minimum and maximum surface air temperatures from geostationary satellite data. J. Geophys. Res. Atmos. 2015, 120, 2306–2324.

Guan, H., Zhang, X., Makhnin, O., Sun, Z. Mapping mean monthly temperatures over a coastal hilly area incorporating terrain aspect effects. J. Hydrometeorol. 2013, 14, 233–250.

GLass, G. Integrating AVHR satellite data and noaa ground observations to predict surface air temperature: A statistical approach. Int. J. Remote Sens. 2004, 25, 2979–2994.

G. Hu, L. Hu, H. Li, K. Li, and W. Liu., 2010. Grid resources Prediction with support vector regression and particle swarm optimization. in Proceedings of the third International Joint Conference on Computational Sciences and Optimization (CSO10), vol. 1, pp. 417–422.

Gooijer, J.G., Kumar, K., 1992. Some recent developments in non-linear time series modeling, testing, and forecasting. International Journal of Forecasting 8, 135-156.

Granger, C.W.J., 1993. Strategies for modelling non-linear time- series relationships. The Economic Record 69, 233-238.

Guyon and Elisseeff., 2003. An introduction to variable and feature selection. Journal of Machine Learning Research.

Grueber, C.E., Nakagawa, S., Laws, R.J. & Jamieson, I.G., 2011. Multimodel inference in ecology and evolution: challenges and solutions. Journal of Evolutionary Biology, 24,699–711.

G. C. Cawley, N. L. C. Talbot, and M. Girolami., 2007. Sparse Multinomial Logistic Regression via Bayesian L1 Regularisation. in: B. Schölkopf, J. C. Platt, and T. Hoffmann (eds.), Advances in Neural Information Processing Systems, MIT Press, pp. 209–216.

Gram, J. P., 1883. Uber Entwicklung reeler Functionen in Reihen mittelst der Methode der kleinsten Quadrate, J. Math. 94, pp. 41-73.

G. Bastin and M.Gevers., 1985. Identification and optimal estimation of random fields from scattered pointwise data, Automatica 21 (2), 139-155.

G.Q. Tabios and J.D., 1985. Sales, A comparative analysis of techniques for spatial interpolation of precipitation, Water Resources Bulletin 21 (3), 365-380.

Ge, Q.; Zhang, X.; Zheng, J., 2014. Simulated effects of vegetation increase/decrease on temperature changes from 1982 to 2000 across the eastern china. Int. J. Clim., 34, 187–196.

Gholamnia, M.; Alavipanah, S.K.; Darvishi Boloorani, A.; Hamzeh, S.; Kiavarz, M., 2017. Diurnal air temperature modeling based on the land surface temperature. Remote Sens. 2017, 9, 915.

Good, E., 2015. Daily minimum and maximum surface air temperatures from geostationary satellite data. J. Geophys. Res. Atmos., 120, 2306–2324.

Guan, H.; Zhang, X.; Makhnin, O.; Sun, Z., 2013. Mapping mean monthly temperatures over a coastal hilly area incorporating terrain aspect effects. J. Hydrometeorol., 14, 233–250.

Grimm, N.B.; Faeth, S.H.; Golubiewski, N.E.; Redman, C.L.;Wu, J.; Bai, X.; Briggs, J.M., 2008. Global change and the ecology of cities. Science 2008, 319, 756–760. DOI: 10.1126/science.1150195.

GLass, G., 2004. Integrating avhrr satellite data and noaa ground observations to predict surface air temperature: A statistical approach. Int. J. Remote Sens., 25, 2979–2994.

H.Theil., 1963. Applied Economic Forecasting, North-Holland.

Hansen, J.; Sato, M.; Ruedy, R.; Lo, K.; Lea, D.W.; Medina-Elizade, M. Global temperature change. Proc. Natl. Acad. Sci. USA 2006, 103, 14288–14293.

Hou, P.; Chen, Y.; Qiao, W.; Cao, G.; Jiang, W.; Li, J., 2013. Near-surface air temperature retrieval from satellite images and influence by wetlands in urban region. Theor. Appl. Clim., 111, 109–118.

Hajek, P.; Olej, V., 2012. Ozone prediction on the basis of neural networks, support vector regression and methods with uncertainty. Ecol. Inform., 12, 31–42.

Henderson, D., Jacobson, S.H., Johnson, A.W. 2003. Handbook of Metaheuristics. Kluwer, Boston, MA.

Haykin, S., 1999. Neural Networks: A Comprehensive Foundation. Prentice Hall, Upper Saddle River, NJ.

Haykin, S. 2001. Redes Neurais: princípios e prática. Porto Alegre: Editora Bookman. 900 pgs.

Hastie, T., Tibshirani, R., Friedman, J., 2009. The Elements of Statistical Learning: Data Mining. Inference and Prediction, New York: Springer.

Holden, Z. A., M. A. Crimmins, S. A. Cushman, and J. S. Littell., 2011. Empirical Modeling of Spatial and Temporal Variation in Warm Season Nocturnal Air Temperatures in Two North Idaho Mountain Ranges, USA. Agricultural and Forest Meteorology 151: 261–269. doi:10.1016/j.agrformet.2010.10.006.

Hagan MT, Menhaj MB. 1994. Training feed forward networks with the Marquaradt algorithm. IEEE Transactions on Neural Networks 6: 861–867.

Hachem, S.; Duguay, C.R.; Allard, M., 2012 Comparison of MODIS-derived land surface temperatures with ground surface and air temperature measurements in continuous permafrost terrain. Cryosphere, 6, 51–69.

H.C.Yuan, F. L. Xiong, andX.Y.Huai. 2003. A method for estimating the number of hidden neurons in feed-forward neural networks based on information entropy. Computers and Electronics in Agriculture, vol. 40, no. 1–3, pp. 57–64.

Hunter D, Hao Y, PukishIII MS, Kolbusz J, Wilamowski BM., 2012. Selection of proper neural network sizes and architecture- A comparative study. IEEE Trans Indust Inf 8(2):228–240.

Huang, R.; Zhang, C.; Huang, J.; Zhu, D.; Wang, L.; Liu, J., 2015. Mapping of daily mean air temperature in agricultural regions using daytime and nighttime land surface temperatures derived from Terra and Aqua MODIS data. Remote Sens., 7, 8728–8756. http://dx.doi.org/10.3390/rs70708728.

H. Liu and L. Yu, 2005. Toward integrating feature selection algorithms for classification and clustering. IEEE Trans. Knowl. Data Eng., vol. 17, no. 4, pp. 491–502.

H. Taha. H. Akbari and A. Rosenfeld., 1989. Vegetation microclimate measurements: the Davis project, Lawrence Berkeley Lab. Rep. 24593.

H. Taha, H. Akbari and A. Rosenfeld, 1991. Heat island and oasis effects of vegetative canopies: Micrometeorological field measurements, Theor. Appl Climat., 44. 123.

H. Liu and H. Motoda., 1998. Feature Selection for Knowledge Discovery and Data Mining, London: Kluwer Academic Publishers.

H. Zou and T. Hastie, 2005. Regularization and variable selection via the elastic net. Journal of the Royal Statistical Society: Series B (Statistical Methodology), vol. 67, no. 2, pp. 301–320.

Huang H, Ooka R, Kato S., 2005. Urban thermal environment measurements and numerical simulation for an actual complex urban area covering a large district heating and cooling system in summer.Atmos Environ 39:6362–6375.https://doi.org/10.1016/j.atmosenv.2005.07.018.

Huang Q, Lu Y., 2015. The effect of urban heat island on climate warming in the Yangtze River Delta urban agglomeration in China. Int J Environ Res Public Health

12:8773–8789.

Hansen, J.; Sato, M.; Ruedy, R.; Lo, K.; Lea, D.W.; Medina-Elizade, M. Global temperature change. Proc. Natl. Acad. Sci. USA 2006, 103, 14288–14293.

Hou, P., Chen, Y., Qiao, W., Cao, G., Jiang, W., Li, J. Near-surface air temperature retrieval from satellite images and influence by wetlands in urban region. Theor. Appl. Clim. 2013, 111, 109–118.

Hocking, R.R., 1976. The analysis and selection of variables in multiple regression. Biometrics 32, 1-49.
I. S. Oh, J. S. Lee, and B. R. Moon, 2004. Hybrid genetic algorithms for feature selection. IEEE Trans. Pattern Anal. Mach. Intell., vol. 26, no. 11, pp. 1424–1437.

Ionescu, A., Candau, Y., 2007. Air pollutant emissions prediction by process modelling application in the iron and steel industry in the case of a. Environmental Modelling and Software 22 (9), 1362–1371.

Iglesias, A., Dafonte, C., Arcay, B., Cotos, J., 2007. Integration of remote sensing techniques and connectionist models for decision support in fishing catches. Environmental Modelling and Software 22 (6), 862–870.

Intergov. Panel Clim. Change (IPCC). 2007. The Physical Science Basis. Contribution of working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.

In S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B.Averyt, M. Tignor, & H. L. Miller (Eds.), Cambridge, United Kingdom: Cambridge University Press.

J. S. Sartakhti, M. H. Zangooei, and K. Mozafari., 2011. Hepatitis disease diagnosis using a novel hybrid-method based on support vector machine and simulated annealing (SVM-SA). Computer Methods and Programs in Biomedicine, vol. 108, no. 2, pp. 570–579.

J.Wang, L. Li, D. Niu, and Z. Tan., 2012. An annual load forecasting model based on support vector regression with differential evolution algorithm. ” Applied Energy, vol. 94, pp.65–70.

Jang, Jyh-Shing R. 1991. Fuzzy Modeling Using Generalized Neural Networks and Kalman Filter Algorithm (PDF). Proceedings of the 9th National Conference on Artificial

Intelligence, Anaheim, CA, USA, July 14–19. 2. pp. 762–767.

Jang, J.-S.R., 1993. ANFIS: adaptive-network-based fuzzy inference system. IEEE Transactions on Systems, Man and Cybernetics. 23 (3). doi:10.1109/21.256541.

Jang, Sun, Mizutani., 1997. Neuro-Fuzzy and Soft Computing. Prentice Hall, pp 335–368, ISBN 0-13-261066-3.

Jang, J. D., Viau, A. A., & Anctil, F., 2004. Neural network estimation of air temperatures from AVHRR data. International Journal of Remote Sensing, 25(21), 4541–4554.

Jang JSR. 1993. ANFIS: adaptive-network-based fuzzy inference system. IEEE Transactions on Systems, Man, and Cybernetics 23(3):665–685.

Jang JSR, Sun CT, Mizutani E. 1997. Neurofuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence. Prentice-Hall: New Jersey.

Jang, J.D.; Viau, A.A.; Anctil, F., 2004 Neural network estimation of air temperatures from AVHRR data. Int. J. Remote Sens., 25, 4541–4554.

Jin, M. and Dickinson, R. E. 2010. Land surface skin temperature climatology: benefiting from the strengths of satellite observations. Environmental Research Letters, 5(4):044004, doi: 10.1088/1748-9326/5/4/044004.

Jin, M., 2004.Analysis of land skin temperature using AVHRR observations. Bulletin of American Meteorological Society, BAMS, 85:587–600, doi: 10.1175/BAMS-85- 4-587.

Jain, A., McClendon, R.W., Hoogenboom, G. and Ramyaa, R., 2003. ”Prediction of frost for fruit protection using artificial neural networks. American Society of Agricultural Engineers, St Joseph, MI, ASAE Paper 03-3075.

Jeong, S.J.; Ho, C.H.; Jeong, J.H., 2009. Increase in Vegetation Greenness and Decrease in Springtime Warming over East Asia. Geophys. Res. Lett.,36. http://dx.doi.org/10.1029/20

Janatian, N.; Sadeghi, M.; Sanaeinejad, S.H.; Bakhshian, E.; Farid, A.; Hasheminia, S.M.; Ghazanfari, S., 2016. A statistical framework for estimating air temperature using MODIS land surface temperature data. Int. J. Climatol., http://dx.doi.org/10.1002/joc.4766.

Jang, J.D.; Viau, A.A.; Anctil, F., 2004. Neural network estimation of air temperatures from AVHRR data.Int.J.RemoteSens.,25,4541–4554. http://dx.doi.org/10.1080/01431160310001657533

J. Tang, S. Alelyani, and H. Liu., 2014. Feature Selection for Classification: A Review in: C. Aggarwal (ed.), Data Classification: Algorithms and Applications. CRC Press, 2014.

J. C. Cortizo and I. Giraldez., 2006. Multi Criteria Wrapper Improvements to Naive Bayes Learning," LNCS, vol. 4224, pp. 419–427.

J. M. Cadenas, M. C. Garrido, and R. Martínez., 2013. Feature subset selection Filter–Wrapper based on low quality data, Expert Systems with Applications, vol. 40, pp. 6241–6252.

Jasinski, M. F., 1990. Sensitivity of the Normalized Difference Vegetation Index to subpixel canopy cover, soil albedo, and pixel scale. Remote Sensing of Environment, 32, 169–187.

Kingston, G.B., 2006. Bayesian Artificial Neural Networks in Water Resources Engineering. Ph.D., The University of Adelaide.

Kullback, S., and R. A. Leibler., 1951. On information and sufficiency. Annals of Mathematical Statistics 22 :79-86.

Kawashima, S.; Ishida, T.; Minomura, M.; Miwa, T., 2000. Relations between surface temperature and air temperature on a local scale during winter nights. J. Appl. Meteorol., 39, 1570–1579.

Kaufmann, R.K.; Zhou, L.; Myneni, R.B.; Tucker, C.J.; Slayback, D.; Shabanov, N.V.; Pinzon, J., 2003. The Effect of Vegetation on Surface Temperature: A Statistical Analysis of NDVI and Climate Data. Geophys. Res. Lett., 30. http://dx.doi.org/10.1029/2003GL018251.

Korrmann, M. and Unbehauen, H., 1988. Structure detection in the identification of nonlinear systems. Automatique productive informative Industrielle (APH), 22, 5-25.

K.-Y. Chen and C.-H. Wang. 2007. Support vector regression with Genetic algorithms in forecasting tourism demand. Tourism Management, vol. 28, no. 1, pp. 215–226.

Koulamas, C., Antony, S.R., Jaen, R. 1994. A survey of simulated annealing applications to operations research problems. OMEGA-Int. J. Manage. Sci. 22, 41–56.

K. Ito and R. Nakano., 2005. Optimizing Support Vector regression hyper-parameters based on cross-validation. in Proceedings of the International Joint Conference on Neural Networks, vol. 3, pp.871–876.

Kawashima, S., Ishida, T., Minomura, M., Miwa, T. Relations between surface temperature and air temperature on a local scale during winter nights. J. Appl. Meteorol. 2000, 39, 1570–1579.

Kaufmann, R. K., L. Zhou, R. B. Myneni, C. J. Tucker, D. Slayback, N. V. Shabanov, and J. Pinzon., 2003. The effect of vegetation on surface temperature: A statistical analysis of NDVI and climate data, Geophys. Res. Lett., 30(22), 2147.

Kimes, D. J., 1983. Remote sensing of row crop structure and component temperatures using directional radiometric temperatures and inversion techniques. Remote Sensing of Environment, 13, 33–55.
Kazuhiro Shin-ike., 2010. A Two Phase Method for Determining the Number of Neurons in the Hidden Layer of a 3-Layer Neural Network. SICE Annual Conference, Taipei, Taiwan.

Ke J, Liu X., 2008. Empirical analysis of optimal hidden neurons in neural network modeling for stock prediction. Pac Asia Work Comput Intell Indust Appl, 2:828–832.

Kisi O., 2007. Streamflow forecasting using different artificial neural network algorithms. ASCE Journal of Hydrologic Engineering 12(5): 532–539.

Kurtzman, D., and R. Kadmon., 1999. Mapping of Temperature Variables in Israel: A Comparison of Different Interpolation Methods. Climate Research 13: 33–43. doi:10.3354/cr013033.

Kisi O, Shiri J., 2011. Precipitation forecasting using wavelet-genetic programming and wavelet-neuro-fuzzy conjunction models. Water Resource Management. 25(13): 3135–3152.

Kisi O, Pour Ali Baba A, Shiri J., 2012. Generalized neuro-fuzzy models for estimating daily pan evaporation values from weather data. ASCE Journal of Irrigation and Drainage Engineering 138(4): 349–362.

Kato, Soushi., & Yamaguchi, Yasushi, 2005. Analysis of urban heat-island effects using ASTER and ETM+ Dat: Separation of anthropogenic heat discharge and natural heat radiation from sensible heat flux. Remote Sensing of Environment., 99, 44-45.

Kim, D., and K. Han., 2013. Remotely Sensed Retrieval of Midday Air Temperature Considering Atmospheric and Surface Moisture Conditions. International Journal of Remote Sensing 34: 247–263. doi:10.1080/01431161.2012.712235.

Karnieli, A.; DallOlmo, G., 2003. Remote sensing monitoring of desertification, phenology, and droughts. Manag. Environ. Q. Int. J., 14, 22–38.

Kilbourne EM., 1997.Heat waves and hot environments.In The Public Health Consequences of Disasters, ed. E Noji, pp. 245–69. New York: Oxford Univ. Press.

Kovats RS, Hajat S., 2008. Heat stress and public health: a critical review. Annul Rev Public Health 29:41–55. https://doi.org/10.1146/annurev.publhealth.29.020907.090843.

Kolokotroni M, Ren X, Davies M, Mavrogianni A., 2012. London's urban heat island: impact on current and future energy consumption in office buildings. Energy Build 47:302–311. https://doi.org/10.1016/j.enbuild.2011.12.019.

Kunkel KE, Changnon SA, Reinke BC, Arritt RW., 1996. The July 1995 heatwave in the midwest: a climatic perspective of critical weather factors. Bull. Am. Meteorol. Soc. 77:1507–18.

Luk, K.C., Ball, J.E. and Sharma, A., 2000.A study of optimal model lag and spatial inputs for artificial neural network for rainfall forecasting. Journal of Hydrology, 227: 56-65.

Litschert, S. E., T. C. Brown, and D. M. Theobald., 2012. Historic and Future Extent of Wildfires in the Southern Rockies Ecoregion, USA. Forest Ecology and Management 269: 124–133.doi:10.1016/j.foreco.2011.12.024.

Lofgren, B.M.; Hunter, T.S.; Wilbarger, J., 2011. Effects of using air temperature as a proxy for evapotranspiration in climate change scenarios of Great Lakes basin hydrology. J. Gt. Lakes Res., 37, 744–752.

Lin, S.; Moore, N.J.; Messina, J.P.; de Visser, M.H.; Wu, J., 2012. Evaluation of estimating daily maximum and minimum air temperature with MODIS data in east Africa. Int. J. Appl. Earth Obs. Geoinf., 18, 128–140.

Lindseth, G. The Cities for Climate Protection Campaign (CCPC) and the framing of local climate policy.

Local Environ. 2004, 9, 325–336. https://doi.org/10.1080/1354983042000246252.

Land Surface Temperature Copernicus Global Land Service, 2016. Available online: http://land.copernicus. eu/global/products/lst (accessed on 1 March 2016).

Land Surface Temperature: Global Maps, 2016. Available online: http:/earthobservatory.nasa.gov.GlobalMaps/view.phpd1=MOD11C1-M-LSTDA.

Lambin, E. F., & Ehrlich, D., 1996. The surface temperature–vegetation index space for land cover and land-cover change analysis. International Journal of Remote Sensing, 17, 463–487.

Larson, R. C., & Carnahan, W. H., 1997. The influence of surface characteristics on urban radiant temperatures. Geocarto International, 12,5 – 16.

Lo, C. P., Quattrochi, D. A., & Luvall, J. C., 1997. Application of high-resolution thermal infrared remote sensing and GIS to assess the urban heat island.

Lin, X.; Zhang, W.; Huang, Y.; Sun, W.; Han, P.; Yu, L.; Sun, F., 2016. Empirical Estimation of near-Surface Air Temperature in China from MODIS Lst Data by Considering Physiographic Features". Remote Sens., 8. http://dx.doi.org/10.3390/rs8080629.

L Yu and H. Liu., 2004. Efficient Feature Selection via Analysis of Relevance and Redundancy," J. Mach. Learn. Res., vol. 5, pp.1205–1224.

Link WA, Barker RJ., 2006. Model weights and the foundations of multimodel inference. Ecology 87:2626–2635.

Machon, I., Lopez, H., Rodriguez-Iglesias, J., Maranon, E., Vazquez, I., 2007. Simulation of a coke wastewater nitrification process using a feed-forward neuronal net. Environmental Modelling and Software 22 (9), 1382–1387.

M. Sandri and P. Zuccolotto., 2006. Variable Selection Using Random Forests. in: S. Zani, A. Cerioli, M. Riani, and M. Vichi (eds.), Data Analysis, Classification and the Forward Search, Studies in Classification, Data Analysis, and Knowledge Organization, Springer, pp. 263–270, 2006.

Maier, H.R., 2006. Application of natural computing methods to water resources and environmental modelling. Mathematical and Computer Modelling 44 (5–6), 413–414.

Maier, H.R., Dandy, G.C., 2000. Application of neural networks to forecasting of surface water quality variables: issues, applications and challenges. Environmental Modelling and Software 15, 348.

Maier, H.R., Morgan, N., Chow, C.W.K., 2004. Use of artificial neural networks for predicting optimal alum doses and treated water quality parameters. Environmental Modelling and Software 19 (5), 485–494.

Mostovoy, G.V.; King, R.L.; Reddy, K.R.; Kakani, V.G.; Filippova, M.G., 2006. Statistical estimation of daily maximum and minimum air temperatures from MODIS LST data over the state of Mississippi. GISci. Remote Sens., 43, 78–110. http://dx.doi.org/10.2747/1548-1603.43.1.78.

May, R.J., Dandy, G.C., Maier, H.R., Fernando, T.M.K.G., 2006. Critical values of a kernel-density based mutual information estimator. In: IEEE International Joint Conference on Neural Networks, Vancouver.

Maier, H.R.; Jain, A.; Dandy, G.C.; Sudheer, K.P. 2010. Methods used for the development of neural networks for the prediction of water resource variables in river systems: Current status and future directions. Environ. Model. Softw., 25,891–909.doi: 10.1016/j.envsoft.2010.02.003.

Muller, M., Fill, H. 2003. Redes Neurais aplicadas na propagacao de vazões, in: Simposio Brasileiro de Recursos Hídricos. Curitiba, Brazil, unpaginated CD-Rom Proceedings.

Mostovoy, G. V., King, R. L., Reddy, K. R., Kakani, V. G., and Filippova, M. G., 2006. Statistical Estimation of Daily Maximum and Minimum Air Temperatures from MODIS LST Data over the State of Mississippi. GIScience & Remote Sensing, 43(1):78–110, doi: 10.2747/1548-1603.43.1.78.

Mittal, G.S. and Zhang, J., 2003. Artificial Neural Network-based Psychrometric Predictor.” Biosystems Engineering, 85(3): 283-289.

Maqsood, I., Khan, M.R. and Abraham, A., 2004. An ensemble of neural networks for weather forecasting. Neural Computing & Applications, 13: 112-122.

Maddala GS., 2011. Introduction to econometrics. 3rd ed. John Wiley & Sons.

M. M. Islam and K. Murase, .2001. A new algorithm to design compact two-hidden-layer artificial neural networks. Neural Networks, vol. 14, no. 9, pp. 1265–1278.

Meteotest., 2010. Meteonorm handbook, Part III: Theory Part 2. Accessed online in February 9(2011). doi: http://www.meteonorm.com/media/pdf/mn6-software.pdf.

Manzo-Delgado, L.; Sánchez-Colón, S.; Álvarez, R., 2009. Assessment of Seasonal Forest Fire Risk Using Noaa-Avhrr: A Case Study in Central Mexico. Int. J. Remote Sens., 30, 4991–5013.

Mamdani EH, Assilian S., 1975. An experiment in linguistic synthesis with a fuzzy logic controller. International Journal of Man Machine Studies 7(1): 1–13.

Mildrexler, D.J.; Zhao, M.; Running, S.W., 2011. A global comparison between station air temperatures and MODIS land surface temperatures reveals the cooling role of forests. J. Geophys., 116, 1–15.

Mairhuber, J. C., 1956. On Haar's theorem concerning Chebyshev approximation problems having unique solutions, Proc. Am. Math. Soc. 7, pp. 609-615.

Mccarthy MP, Best MJ, Betts RA., 2010. Climate change in cities due to global warming and urban effects. Geophys Res Lett. https://doi.org/10.1029/2010GL042845.

Nemani, R. R., & Running, S. W., 1989. Estimation of regional surface resistance to evapotranspiration from NDVI and Thermal-IR AVHRR data. Journal of Applied Meteorology, 28, 276−284.

Nemani, R., L. Pierce, S. Running, and S. Goward., 1993. Developing satellite-derived estimates of surface moisture status, J. Appl. Meteorol., 32, 549–557.

Ninyerola, M.; Pons, X.; Roure, J.M. Objective air temperature mapping for the iberian peninsula using spatial interpolation and gis. Int. J. Clim. 2007, 27, 1231–1242.

Niclos, R.; Valiente, J.; Barbera, M.J.; Caselles, V. Land surface air temperature retrieval from eos-modis images. IEEE Geosci. Remote Sens. Lett. 2014, 11, 1380–1384.

Nieto, H.; Sandholt, I.; Aguado, I.; Chuvieco, E.; Stisen, S. Air temperature estimation with msg-seviri data: Calibration and validation of the tvx algorithm for the iberian

peninsula. Remote Sens. Environ. 2011, 115, 107–116.

N.Lu, T.Taylor,W.Jiang,C.Jin,J.Correia,L.R.Leung,P.C.Wong., 2010. Climatechange impacts on residential and commercial loads in the Western U.S. grid, IEEE Transactions on Power Systems 25.

Ninyerola, M.; Pons, X.; Roure, J.M., 2007. Objective air temperature mapping for the iberian peninsula using spatial interpolation and gis. Int. J. Clim., 27, 1231–1242.

Niclos, R.; Valiente, J.; Barbera, M.J.; Caselles, V., 2014. Land surface air temperature retrieval from eos-modis images. IEEE Geosci. Remote Sens., 11, 1380–1384.

Nemani, R.R.; Running, S.W.1989. Estimation of regional surface resistance to evapotranspiration from NDVI and thermal IR AVHRR data. J. Appl. Meteorol. 1989, 28, 276–284.

Nieto, H., I. Sandholt, I. Aguado, E. Chuvieco, and S. Stisen. 2011. Air Temperature Estimation with MSG-SEVIRI Data: Calibration and Validation of the TVX Algorithm for the Iberian Peninsula. Remote Sensing of Environment 115: 107–116. doi:10.1016/j.rse.2010.08.010.

N. Cristianini and J. Taylor., 2000. An Introduction to Support Vector Machines. ” Cambridge University Press, Cambridge, UK.

N. Hoque, D. K. Bhattacharyya, and J. K. Kalita., 2014. MIFS-ND: A mutual information-based feature selection method, Expert Systems with Applications, vol. 41, issue 14, pp. 6371–6385.

Ozkaya, B., Demir, A., Bilgili, M., 2007. Neural network prediction model for the methane fraction in biogas from field-scale landfill bioreactors. Environmental Modelling and Software 22 (6), 815–822.

Oyler, J.W.; Ballantyne, A.; Jencso, K.; Sweet, M.; Running, S.W., 2015. Creating a Topoclimatic Daily Air Temperature Dataset for the Conterminous United States Using Homogenized Station Data and Remotely Sensed Land Skin Temperature. Int. J. Climatol, 35, 2258–2279.

Ozgur Kisia and Jalal Shirib., 2014. Prediction of long-term monthly air temperature using geographical inputs”. Int. J. Climatol. 34: 179–186 (2014) DOI: 10.1002/joc.3676.

Oke, T. R., 1982. The energetic basis of the urban heat island. Quarterly Journal of the Royal Meteorological Society, 108, 1–24.

Owen, T. W., Carlson, T. N., & Gillies, R. R., 1998. An assessment of satellite remotely-sensed land cover parameters in quantitatively describing the climatic effect of urbanization. International Journal of Remote Sensing, 19, 1663−1681.

Panchal G, Ganatra A, Kosta YP, Panchal D., 2011. Behaviour analysis of multilayer perceptrons with multiple hidden neurons and hidden layers. Int J Comput Theory Eng, 3(2):332–337.

Prihodko, L. and Goward, S. N. 1997. Estimation of air temperature from remotely sensed surface observations. "Remote Sensing of Environment, 60(3):335 – 346, doi: 10.1016/S0034-4257(96)00216-7.

Partal T, Kisi O. 2007. Wavelet and neuro fuzzy conjunction model for precipitation forecasting. Journal of Hydrology 342: 199–212.

Prince, S., Goetz, S., Dubayah, R., Czajkowski, K., and Thawley, M., 1998. Inference of surface and air temperature, atmospheric precipitable water and vapor pressure deficit using Advanced Very High-Resolution Radiometer satellite observations: comparison with field observations. Journal of Hydrology, 212-213(0):230–249,doi:10.1016/S0022-169

Pinheiro, A.C.T.; Mahoney, R.; Privette, J.L.; Tucker, C.J., 2006 Development of a daily long-term record of NOAA-14 AVHRR land surface temperature over Africa. Remote Sens. Environ., 103, 153−164.

Purkey DR, Joyce B, Vicuna S et al., 2007. Robust analysis of future climate change impacts on water for agriculture and other sectors: a case study in the Sacramento Valley. Clim Chang 87:109–122.

Peón, J.; Carmen, R.; Javier, F.C., 2014.Improvements in the Estimation of Daily Minimum Air Temperature in Peninsular Spain Using MODIS Land Surface Temperature. Int. J. Remote Sens., 35, 5148–5166.

Pouteau, R.; Rambal, S.; Ratte, J.P.; Gogé, F.; Joffre, R.; Winkel, T., 2011. Downscaling MODIS-Derived Maps Using Gis and Boosted Regression Trees: The Case of Frost Occurrence over the Arid Andean Highlands of Bolivia. Remote Sens. Environ., 115, 117–129. http://dx.doi.org/10.1016/j.rse.2010.08.011.

Pires, J., Martins, F., Sousa, S., Alvim-Ferraz, M., Pereira, M., 2008. Selection and validation of parameters in multiple linear and principal component regressions. Environmental Modelling and Software 23 (1), 50–55.

Pape, R.; Wundram, D.; Löffler, J., 2009. Modelling near-surface temperature conditions in high mountain environments: An appraisal. Clim. Res., 39, 99–109.

Peng, B.; Zhou, Y.; Gao, P.; Ju, W., 2011. Suitability assessment of different interpolation methods in the gridding process of station collected air temperature: A case study in Jiangsu province, China. J. Geo-Inf. Sci.1, 13, 539–548.

Piao, S.; Fang, J.; Ciais, P.; Peylin, P.; Huang, Y.; Sitch, S.; Wang, T. The Carbon Balance of Terrestrial Ecosystems in China. Nature 2009, 458, 1009–1013. http://dx.doi.org/10.1038/nature07944.

P. S. Bradley and O. L. Mangasarian, 1998. Feature selection via concave minimization and support vector machines. in: Proc.15th International Conference on Machine Learning, Madison, Wisconsin, USA, Morgan Kaufmann, pp. 82–90.

Pohjalainen, J., Kadioglu, S., Räsaänen, o., 2012. Feature selection for speaker traits. In:Proc. 13th Annual Conference of the international Speech Communication Association, Portland, OR, USA.

Price, J. C., 1990. Using spatial context in satellite data to infer regional scale evapotranspiration. IEEE Transactions on Geoscience and Remote Sensing, 28, 940−948.

Palmer, W. C. (1968). Keeping track of moisture conditions, nationwide: The new crop moisture index. Weatherwise. 21. 156–161.

Quattrochi, D. A.,& Ridd, M. K., 1998. Analysis of vegetation within a semi-arid urban environment using high spatial resolution airborne thermal infrared remote sensing data. Atmospheric Environment, 32,19–33.

Raynolds, M.; Comiso, J.; Walker, D.; Verbyla, D., 2008. Relationship between Satellite-Derived Land Surface Temperatures, Arctic Vegetation Types, and NDVI. Remote Sens. Environ., 112, 1884–1894. Doi: http://dx.doi.org/10.1016/j.rse.2007.09.008.

Raynolds, M.K.; Walker, D.A.; Maier, H.A., 2006. NDVI Patterns and Phytomass Distribution in the Circumpolar Arctic. Remote Sens. Environ., 102, 271–281.doi:

http://dx.doi.org/10.1016/j.rse.2006.02.016.

Raduly, B., Gernaey, K., Capodaglio, A., Mikkelsen, P., Henze, M., 2007. Artificial neural networks for rapid WWTP performance evaluation: methodology and case study. Environmental Modelling and Software 22 (8), 1208–1216.

Rodriguez, M.J., Serodes, J.-B., 1999. Assessing empirical linear and non-linear modelling of residual chlorine in urban drinking water systems. Environmental Modelling and Software 14, 93–102.

Rissanen, J., 1978. Modelling by shortest data description. Automatic 14, 465-471.

Rissanen, J., 1980. Consistent order-estimates of autoregressive processes by shorts description of data. In: Analysis and Optimization of Stochastic Systems, Jacobs, Ed. O., Davis, M., Dempster, M., Harris, C, Parks, P. (Eds.), Academic Press, New York, pp.451-461.

Razavi, S.; Tolson, B.A. 2011. A New Formulation for Feed-Forward Neural Networks. IEEE Trans. Neural Netw., 22, 1588–1598.Doi:10.1109/TNN.2011.2163169.

Romeo, F., Sangiovanni-Vincentelli, A. 1991. A theoretical framework for simulated annealing. Algorithmic, 6, 302–345.

Rivals I. & Personnaz L., 2000.A statistical procedure for determining the optimal number of hidden neurons of a neural model. Second International Symposium on Neural-Computation, Berlin.

Reich, P.B.; Luo, Y.; Bradford, J.B.; Poorter, H.; Perry, C.H.; Oleksyn, J. 2014. Temperature Drives Global Patterns in Forest Biomass Distribution in Leaves, Stems, and Roots. Proc. Natl. Acad. Sci. USA, 111, 13721–13726.

Ruane, A.C.; McDermid, S.; Rosenzweig, C.; Baigorria, G.A.; Jones, J.W.; Romero, C.C.; Cecil, L.D. 2014. Carbon-Temperature-Water Change Analysis for Peanut Production under Climate Change: A Prototype for the Agmip Coordinated Climate-Crop Modeling Project (C3mp). Glob. Chang. Biol., 20, 394–407.

Rosenzweig, C.; Elliott, J.; Deryng, D.; Ruane, A.C.; Muller, C.; Arneth, A.; Boote, K.J.; Folberth, C.; Glotter, M.;Khabarov, N.; et al., 2014. Assessing Agricultural Risks of Climate Change in the 21st Century in a Global Gridded Crop Model Intercomparison.

Proc. Natl. Acad. Sci. USA, 111, 3268–3273.

Rolland, C., 2003. Spatial and Seasonal Variations of Air Temperature Lapse Rates in Alpine Regions. Journal of Climate 16: 1032–1046.doi:10.1175/1520-0442(2003)016.1032:SASVOA.2.0.

Rizwan AM, Dennis LYC, Liu C., 2008. A review on the generation, determination and mitigation of urban heat island. J Ecol Environ Sci 20:120–128. https://doi.org/10.1016/s1001-0742(08)60019-4.

Roth M., 2007. Review of urban climate research in subtropical regions. Int J Climatol 27:1859–1873. https://doi.org/10.1002/joc.1591.

Reunanen, J., 2003. Over-fitting in making comparison between variable selection methods. Journal of Machine Learning Research. 3,1371-1382.

Saeys, Y., Abeel, T., de Peer, Y.V., 2008. Robust feature selection using ensemble feature selection techniques. In: Proc. European Conference on Machine Learning and Knowledge Discovery in Database, pp.313-325.

Sarrat C, Lemonsu A, Masson V, Guedalia D., 2006 Impact of urban heat island on regional atmospheric pollution. Atmos Environ 40:1743–1758. https://doi.org/10.1016/j.atmosenv.2005.11.037.

Shuxiang X, Chen L. 2008. A novel approach for determining the optimal number of hidden layer neurons for FNN's and its application in data mining. In 5th International Conference on Information Technology and Application (ICITA), 683-686.

Sousa, S.; Martins, F.; Alvimferraz, M.; Pereira, M., 2007. Multiple linear regression and artificial neural networks based on principal components to predict ozone concentrations. Environ. Model. Softw., 22, 97–103.DOI: 10.1016/j.envsoft.2005.12.002.

S. Keerthi, 2002. Efficient tuning of SVM hyper parameters using radius/margin bound and iterative algorithms. "IEEE Transactions on Neural Networks, vol. 13, no. 5, pp. 1225–1229".

Stisen, S., Sandholt, I., NÃ¸rgaard, A., Fensholt, R., and Eklundh, L. 2007. Estimation of diurnal air temperature using MSG SEVIRI data in West Africa. Remote Sensing of Environment, 110(2):262–274, doi: 10.1016/j.rse.2007.02.025.

Steidley, C., Sadovski, A., Tissot, P. and Bachnak, R., 2005. Using an Artificial Neural Network to Improve Predictions of Water Level Where Tide Charts Fail. In M. Ali and F. Esposito (Editors), Innovations in Applied Artificial Intelligence. Springer, Bari, Italy.

Suman, B., Kumar, P., 2006. A survey of simulated annealing as a tool for single and multi-objective optimization. J. Oper. Res. Soc. 57, 1143–1160.

Sugeno M , Tanaka K. 1992. Successive identification of a fuzzy model and its applications to prediction of a complex system. Proc Fuzzy Sets Syst, 42:315–34.

Smith, B.A., McClendon, R.W. and Hoogenboom, G. 2006. Improving Air Temperature Prediction with Artificial Neural Networks. International Journal of Computational Intelligence, 3(3): 179-186.

Smith BA, McClendon RW, Hoogenboom G., 2005. An enhanced artificial neural network for air temperature prediction. Proceedings of World Academy of Science, Engineering and Technology (PWASET)7: 7–12.

Shiri J, Dierickx W, Pour-Ali Baba A, Neamati S, Ghorbani MA., 2011. Estimating daily pan evaporation from climatic data of the State of Illinois, USA using adaptive neuro-fuzzy inference system (ANFIS) and artificial neural networks (ANN). Hydrology Research 42(6):491–502.

Shank DB, Hoogenboom G, McClendon RW., 2008. Dew point temperature prediction using artificial neural networks. Journal of Applied Meteorology and Climatology 47: 1757–1769.

Snyder, W. C., Wan, Z., Zhang, Y., & Feng, Y. Z., 1998. Classification based emissivity for land surface temperature measurement from space. International Journal of Remote Sensing, 19, 2753−2774.

Sobrino, J. A., Raissouni, N., & Li, Z. L. 2001. A comparative study of land surface emissivity retrieval from NOAA data. Remote Sensing of Environment, 75, 256−266.

Stuti Asthana and Rakesh K Bhujade., 2011. Handwritten Multiscript Pin Code Recognition System having Multiple hidden layers using Back Propagation Neural Network. International Journal of Electronics Communication and Computer Engineering, ISSN: 2249-071X, Volume 2, Issue 1.

S.A. Changnon Jr. 1976. Inadvertent weather modifications, WaterResourc. Bull., 12 (1976)695.

S.A. Changnon Jr., 1981. A review and summary, Meteorol. Monogr. 18, No. 40, American Meteorological Society.

Stone, M., 1974. Cross-validation choice and assessment of statistical predictions. Journal of the Royal Statistical Society: Series B 36, 111-147.

Sandholt, I., Rasmussen, K., & Andersen, J., 2002. A simple interpretation of the surface temperature-vegetation index space for assessment of surface moisture status. Remote Sensing of Environment, 79(2–3), 213–224.

Sun, Y. J., Wang, J. F., Zhang, R. H., Gillies, R. R., Xue, Y., & Bo, Y. C., 2005. Air temperature retrieval from remote sensing data based on thermodynamics. Theoretical and Applied Climatology, 80(1), 37–48.

Shamir, E.; Georgakakos, K.P., 2014. MODIS Land Surface Temperature as an Index of Surface Air Temperature for Operational Snowpack Estimation. Remote Sens. Environ., 152, 83–98.
Sousa, S., Martins, F., Alvim-Ferraz, M., Pereira, M., 2007. Multiple linear regression and artificial neural networks based on principle components to predict ozone concentrations. Environmental Modelling and Software 22 (1),97–103.

Serodes, J.-B., Rodriguez, M.J., Ponton, A., 2001. Chlorcast(C): a methodology for developing decision-making tools for chlorine disinfection control. Environmental Modelling and Software 16, 53–62.

S. Alelyani, J. Tang, and H. Liu, 2013. Feature Selection for Clustering: A Review. in: C. Aggarwal and C. Reddy (eds.), Data Clustering: Algorithms and Applications, CRC Press.

S. I. Ali and W. Shahzad., 2012. A Feature Subset Selection Method based on Conditional Mutual Information and Ant Colony Optimization, International Journal of Computer Applications, vol. 60, no. 11, pp. 5–10.

Schwarz, G., 1978. Estimating the dimension of a model. The Annals of Statistics 6, 461-464.

Shanmuganathan, S., Sallis, P., Buckeridge, J., 2006. Self-organizing map methods in integrated modelling of environmental and economic systems. Environmental Modelling and Software 21 (9), 1247–1256.

Shen, S.; Leptoukh, G.G.2011. Estimation of surface air temperature over central and eastern Eurasia from MODIS land surface temperature. Environ. Res., 6, 045206. http://dx.doi.org/10.1088/1748-9326/6/4/045206.

S. Ma and J. Huang., 2008. Penalized feature selection and classification in bioinformatics, Briefings in Bioinformatics, vol. 9, no. 5, pp. 392–403.

S. Das., 2001. Filters, wrappers and a boosting-based hybrid for feature selection, in: Proc. 18th International Conference on Machine Learning, San Francisco, CA, USA, Morgan Kaufmann, pp. 74–81.

Stow, D.A.; Hope, A.; McGuire, D.; Verbyla, D.; Gamon, J.; Huemmrich, F.; Houston, S.; Racine, C.; Sturm, M.;Tape, K.; et al., 2004. Remote Sensing of Vegetation and Land-Cover Change in Arctic Tundra Ecosystems. Remote Sens. Environ., 89, 281–308. Doi:http://dx.doi.org/10.1016/j.rse.2003.10.018.

Sarafrazi, S and Nezamabadi-pour, H., 2013. Facing the classification of binary problems with a GSA-SVM hybrid system, Mathematical and Computer Modelling, vol. 57, issues 1-2, pp. 270–278.

S. Maldonado, R. Weber, and F. Famili., 2014. Feature selection for high-dimensional class-imbalanced data sets using Support Vector Machines, Information Sciences, vol. 286, pp. 228–246.

Symonds,M.R.E. and Moussalli, A., 2011. A brief guide to model selection, multimodel inference and model averaging in behavioural ecology using Akaike's information criterion. Behavioral Ecology and Sociobiology, 65, 13–21.

Sun, Y.-J.; Wang, J.-F.; Zhang, R.-H.; Gillies, R.; Xue, Y.; Bo, Y.-C., 2005. Air temperature retrieval from remote sensing data based on thermodynamics. Theor. Appl. Clim., 80, 37–48.

Stahl, K.; Moore, R.; Floyer, J.; Asplin, M.; McKendry, I., 2006. Comparison of approaches for spatial interpolation of daily air temperature in a large region with complex topography and highly variable station density. Agric. For. Meteorol., 139, 224–236.

Sandholt, I., Rasmussen, K., & Andersen, J., 2002. A simple interpretation of the surface temperature/vegetation index space for assessment of surface moisture status. Remote Sensing of Environment, 79, 213–224.

Sun, D. L., M. Kafatos, 2007: Note on the NDVI-LST relationships and the use of temperature-related drought indices over North America. Geophys. Res. Lett., 34, L24406.

SOBRINO, J.A. and RAISSOUNI, N., 2000, Toward remote sensing methods of land cover dynamic monitoring: Application to Morocco. International Journal of Remote Sensing, 21, pp. 353–366.

Sun, Y.-J.; Wang, J.-F.; Zhang, R.-H.; Gillies, R.; Xue, Y.; Bo, Y.-C. Air temperature retrieval from remote sensing data based on thermodynamics. Theor. Appl. Clim. 2005, 80, 37–48.

Stahl, K.; Moore, R.; Floyer, J.; Asplin, M.; McKendry, I. Comparison of approaches for spatial interpolation of daily air temperature in a large region with complex topography and highly variable station density. Agric. For. Meteorol. 2006, 139, 224–236.

Stisen, S.; Sandholt, I.; Nørgaard, A.; Fensholt, R.; Eklundh, L. Estimation of diurnal air temperature using msg seviri data in west africa. Remote Sens. Environ. 2007, 110, 262–274.

Snyder, W. C., Wan, Z., Zhang, Y., & Feng, Y. Z., 1998. Classificationbased emissivity for land surface temperature measurement from space. International Journal of Remote Sensing, 19, 2753– 2774.

Small, C., 2001. Estimation of urban vegetation abundance by spectral mixture analysis. International Journal of Remote Sensing, 22, 1305– 1334.

T.J. Chandler., 1960. Wind as a factor of urban temperatures: A survey in north-east London, Weather, 25 (1960) 204.

T.R. Oke., 1987. Boundary Layer Climares, Methuen. London,.2nd edn.

T.R. Oke., 1988. The urban energy balance, Prog. Phys. Geogr., 12, 471.

T.R. Oke., 1995. Boundary Layer Climates. London: Methuen.

T.R. Oke., 1973. City size and the urban heat island. Atmos. Environ. 7:769–79.

T.R. Oke., 1982. The energetic basis of the urban heat island. Q J R Meteorol Soc 108:1–24.

T. Karl and P. Jones., 1989. Urban biases in area-averaged surface air temperature trends. Bull. Am. Meteorol. Sot., 70, 265;Environ. Rep. 2, WMO No. 312, Geneva, pp. 101-106.

Takagi T, Sugeno M., 1985. Fuzzy identification of systems and its applications to modeling and control. IEEE Trans Syst Man Cybern, SMC-15(1):116–32.

Tatli H, Sen Z., 1999. A new fuzzy modeling approach for predicting the maximum daily temperature from a time series. Turkish Journal of Engineering and Environmental Science 23: 173–180.

T. Onoda., 1995. Neural network information criterion for the optimal number of hidden units. in Proceedings of the 1995 IEEE International Conference on Neural Networks, vol. 1, pp. 275–280.

Tomislav Hengl, Gerard B. M. Heuvelink, Melita Percec Tadic, Edzer J. Pebesma., 2012. Spatio-temporal prediction of daily temperature using time-series of MODIS LST image. Theor Appl Climatol, 107:265-277.DOI:10.1007//s00704-011-0464-2.

Unger, J.; Gál, T.; Rakonczai, J.; Musci, L.; Szatman, J.; Tobak, Z., 2009. Air temperature versus surface temperature in urban environment. In Proceedings of the 7th International Conference on Urban Climate, Yokohama, Japan.

Vernieuwe H, Georgieva O, De Baets B, Pauwels VRN, Verhoest NEC, De Troch FP., 2005. Comparison of data-driven Takagi-Sugeno models of rainfall-discharge dynamics. Journal of Hydrology 302(1-4): 173–186.

Vogt, J., A. A. Viau, and F. Paquet., 1997. Mapping Regional Air Temperature Fields Using Satellite Derived Surface Skin Temperatures. International Journal of Climatology 17: 1559–1579. doi:10.1002/(SICI)1097-0088(19971130)17.

Van Laarhoven, P.J.M., 1988. Theoretical and Computational Aspects of Simulated Annealing. Centrum voor Wiskunde en Informatica, Amsterdam, Netherlands.

Van De Kerchove, R.; Lhermitte, S.; Veraverbeke, S.; Goossens, R., 2013. Spatio-Temporal Variability in Remotely Sensed Land Surface Temperature, and Its Relationship with Physiographic Variables in the Russian Altay Mountains. Int. J. Appl. Earth Obs. Geoinf., 20, 4–19. http://dx.doi.org/10.1016/j.jag.2011.09.007.

Vancutsem, C., Ceccato, P., Dinku, T., and Connor, S. J., 2010. Evaluation of MODIS land surface temperature data to estimate air temperature in different ecosystems over Africa. Remote Sensing of Environment, 114(2):449 – 465, doi: 10.1016/j.rse.2009.10.002.

Van Laarhoven, P.J.M., Aarts, E.H.L., 1987. Simulated annealing: Theory and applications.

Vazquez, D. P., Reyes, F. J. O., & Arboledas, L. A., 1997. A comparative study of algorithms for estimating land surface temperature from AVHRR data. ” Remote Sensing of Environment, 62, 215–222.

Voogt, J. A.,& Oke, T. R., 1998. Effects of urban surface geometry on remotely-sensed surface temperature. International Journal of Remote Sensing, 19, 895–920.

Wan, Z., P. Wang, and X. Li., 2004, Using MODIS land surface temperature and normalized difference vegetation index products for monitoring drought in the southern Great Plains, USA, Int. J. Remote Sens., 25, 61–72.

Walsh, S. J., Moody, A., Allen, T. R., & Brown, D. G., 1997. Scale dependence of NDVI and its relationship to mountainous terrain. In D. A. Quattrochi, & M. F. Goodchild (Eds.), Scale in Remote Sensing and GIS (pp. 27– 55). Boca Raton, FL: Lewis Publishers.

Weng, Q., 2001. A remote sensing-GIS evaluation of urban expansion and ist impact on surface temperature in the Zhujiang Delta, China. International Journal of Remote Sensing, 22(10), 1999−2014.

Weng, Q., Lu, D., & Schubring, J., 2004. Estimation of land surface temperature–vegetation abundance relationship for urban heat island studies. Remote Sensing of Environment, 89, 467−483.

Wilson, J. S., Clay, M., Martin, E., Stuckey, D., & Vedder-Risch, K.,2003. Evaluating environmental influences of zoning in urban ecosystems with remote sensing. Remote Sensing of Environment, 86, 303– 321.

Wilson.YUE J. XU, W. TAN and L. XU., 2007. The relationship between land surface temperature and NDVI with remote sensing: application to Shanghai Landsat 7 ETM+ data. International Journal of Remote Sensing Vol. 28, No. 15, 10 August 2007, 3205–3226.

Wloczyk, C., Borg, E., Richter, R., and Miegel, K., 2011. Estimation of instantaneous air temperature above vegetation and soil surfaces from Landsat 7 ETM+ data in northern Germany. International Journal of Remote Sensing,32(24):9119–9136, doi: 10.1080/01431161.2010.550332.

Wedge, D., Ingram, D., McLean, D., Mingham, C. and Bandar, Z., 2005. A global-local artificial neural network with application to wave overtopping prediction.

W. Duch, J. Kacprzyk, E. Oja and S. Zadrozny., 2005. Artificial Neural Networks: Formal Models and Their Applications. Springer, Warsaw, Poland.

Willmott, C.J.; Matsuura, K. 2005. Advantages of the Mean Absolute Error (MAE) over the Root Mean Square Error (RMSE) in Assessing Average Model Performance. Clim. Res., 30, 79–82.

Willmott, C.J.; Robeson, S.M. Climatologically Aided Interpolation of terrestrial air temperature. Int. J. Climatol. 1995, 15, 221–229.

Wan, Z.; Zhang, Y.; Zhang, Q.; Li, Z.L. 2004. Quality assessment and validation of the MODIS global land surface temperature. Int. J. Remote Sens., 25, 261–274.

Wan, Z.; Zhang, Y.; Zhang, Q.; Li, Z.L. 2002. Validation of the Land-Surface Temperature Products Retrieved from Terra Moderate Resolution Imaging Spectroradiometer Data. Remote Sens. Environ., 83, 163–180. Doi: http://dx.doi.org/10.1016/S0034-4257(02)00093-7.

Wan, Z.; Li, Z.L., 2008. Radiance-based validation of the V5 MODIS land-surface temperature product. Int. J. Remote Sens., 29, 5373–5395.

Westerling, A.L.; Hidalgo, H.G.; Cayan, D.R.; Swetnam, T.W., 2006. Warming and Earlier Spring Increase Western U.S. Forest Wildfire Activity. Science, 313, 940–943.

Williamson, S.N.; Hik, D.S.; Gamon, J.A.; Kavanaugh, J.L.; Koh, S. Evaluating cloud contamination in clear-sky MODIS Terra daytime land surface temperatures using ground-based meteorology station observations.

J.Clim. 2013, 26, 551–1560. http://dx.doi.org/10.1175/JCLI-D-12-00250.1.

Woolhouse, W. S. B., 1980). Explanation of a new method of adjusting mortality tables, with some observations upon Mr. Makeham's modification of Gompertz's theory, J. Inst. Act. 15, pp. 389-410.

Xu, Y., Qin, Z., and Shen, Y., 2012. Study on the estimation of near-surface air temperature from MODIS data by statistical methods. International Journal of Remote Sensing, 33(24):7629–7643, doi: 10.1080/01431161.2012.701351.

Watkins R, Palmer J, Kolkotroni M, Little fair P., . The London Heat Island—surface and air temperature measurements in a park and street gorges. ASHRAE Trans. 108:419– 27.

Xu, Y.; Knudby, A., and Chak Ho, H., 2014. Estimating daily maximum air temperature from MODIS in British Columbia, Canada. International Journal of Remote Sensing, 35:24, 8108-8121, DOI: 10.1080/01431161.2014.978957.

Yang, J.; Tan, C.; Zhang, T., 2012. Spatial and temporal variations in air temperature and precipitation in the chinese himalayas during the 1971–2007. Int. J. Clim., 33, 2622–2632.

Yates D, Sieber J, Purkey D et al., 2005. WEAP21—a demand-, priority-, and preference-driven water planning model.Water Int 30:487–500.

Yoo, J.-M., Won, Y.-I., Cho, Y.-J., Jeong, M.-J., Shin, D.-B., Lee, S.-J., Lee, Y.-R., Oh, S.-M., and Ban, S.-J. 2011. Temperature trends in the skin/surface, mid-troposphere and low stratosphere near Korea from satellite and ground measurements." Asia-Pacific Journal of Atmospheric Sciences, 47:439–455, doi: 10.1007/s13143-011-0029-4.9.

Y. K.Wu and J. S. Hong. 2007. A literature review of wind forecasting technology in the world. " in Proceedings of the IEEE Lausanne Power Tech, pp. 504–509.

Ying Yang., 2005. Can the Strengths of AIC and BIC Be Shared? Biometrika 92, no. 4 (December 2005), 937–950.

Y. S. Kim, W. N. Street, and F. Menczer., 2002. Evolutionary model selection in unsupervised learning, Intelligent Data Analysis, vol. 6, no. 6, pp. 531–556, 2002.

Yang, Y., 2007b. Consistency of cross validation for comparing regression procedures. The Annals of Statistics, 35, 2450-2473.

Yang, W., Yang, L., & Merchnat, J. W., 1997. An analysis of AVHRR/ NDVI-ecoclimatological relations in Nebraska, USA. International Journal of Remote Sensing, 18, 2161– 2180.

Zhu, W., u, A. L., and Jia, S., 2013. Estimation of daily maximum and minimum air temperature using MODIS land surface temperature products. Remote Sensing of Environment, 130(0):62–73, doi: 10.1016/j.rse.2012.10.034.

Zhang, W., Huang, Y., Yu, Y., and Sun, W., 2011a. Empirical models for estimating daily maximum, minimum and mean air temperatures with MODIS land surface temperatures. International Journal of Remote Sensing, 32(1):1–26, doi: 10.1080/01431161.2011.560622.

Zhang, J.; Gao, S.; Chen, H.; Yu, J.; Tang, Q., 2015. Retrieval of the land surface-air temperature difference from high spatial resolution satellite observations over complex surfaces in the Tibetan Plateau. J. Geophys. Res. Atmos., 120. http://dx.doi.org/10.1002/2015JD0

Zaksek, K., & Schroedter-Homscheidt, M., 2009. Parameterization of air temperature in high temporal and spatial resolution from a combination of the SEVIRI and MODIS instruments. ISPRS Journal of Photogrammetry and Remote Sensing, 64(4), 414–421. doi:10.1016/j.isprsjprs.2009.02.006.

Zheng, X.; Zhu, J.; Yan, Q., 2013. Monthly Air Temperatures Over Northern China Estimated by Integrating MODIS Data with Gis Techniques. J. Appl. Meteorol. Climatol, 52, 1987–2000.

Zeng, L.;Wardlow, B.D.; Tadesse, T.; Shan, J.; Hayes, M.J.; Li, D.; Xiang, D., 2015. Estimation of daily air temperature based on MODIS land surface temperature products over the corn belt in the US. Remote Sens., 7, 951–970. http://dx.doi.org/10.3390/rs70100951.

Zavala M., 2004. Integration of drought tolerance mechanisms in Mediterranean sclerophylls: a functional interpretation of leaf gas exchange simulators. Ecol Model 176:211–226.

Zurr, A.F.; Ieno, E.N.; Elphick, C.S., 2010. A Protocol for Data Exploration to Avoid Common Statistical Problems. Methods Ecol. Evol., 1, 3–14. Doi:http://dx.doi.org/10.1111/j.2041-210X.2009.00001.

Zhao, C.; Nan, Z.; Cheng, G., 2005. Methods for modelling of temporal and spatial distribution of air temperature at landscape scale in the southern Qilian mountains, China. Ecol. Model., 189, 209–220.

Zhou, W.; Peng, B.; Shi, J.; Wang, T.; Dhital, Y.; Yao, R.; Yu, Y.; Lei, Z.; Zhao, R., 2017. Estimating high resolution daily air temperature based on remote sensing products and climate reanalysis datasets over glacierized basins: A case study in the langtang valley, Nepal. Remote Sens., 9, 959.

Zhu, W.; Lu , A.; Jia, S., 2013. Estimation of daily maximum and minimum air temperature using modis land surface temperature products. Remote Sens. Environ., 130, 62–73.

Zhou Y, Shepherd JM., 2010. Atlanta's urban heat island under extreme heat conditions and potential mitigation strategies. Nat Hazards 52:639–668.