Thermal Landscape and its relation to Seasonal Rainfall in Malaga (Spain)

Gemma Muros Esteban

March, 2011

Course Title:	Geo-Information Science and Earth Observation for Environmental Modelling and Management
Level:	Master of Science (MSc)
Course Duration:	September 2009 – March 2011
Consortium partners:	University of Southampton (UK) Lund University (Sweden) University of Warsaw (Poland) University of Twente, Faculty ITC (The Netherlands)

Thermal Landscape and its Relation to Seasonal Rainfall in Andalusia (Spain)

by

Gemma Muros Esteban

Thesis submitted to the University of Twente, faculty ITC, in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation for Environmental Modelling and Management

Thesis Assessment Board

Chair: Prof. Dr. Andrew Skidmore
External Examiner: Dr. Jadu Dash
First Supervisor: MSc. Valentijn Venus
Second Supervisor: Prof. Dr. Ing. Wouter Verhoef

UNIVERSITY OF TWENTE.

ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

Disclaimer

This document describes work undertaken as part of a programme of study at the University of Twente, Faculty ITC. All views and opinions expressed therein remain the sole responsibility of the author, and do not necessarily represent those of the university.

Abstract

Unexpected variations of water availability can have important impacts not only in the natural environment but also in the social environment. Forecasting of seasonal rainfall is a very helpful input for policy planning that allows to improve the adaptive capacity to the oncoming new situations. Moisture and thermal advection are in most of the cases highly related to precipitation events. Advection depends on the temperature gradient and wind vectors. Hence, this study attempts to identify possible moisture and thermal advection sites to predict rainfall in Malaga (Spain) taking into account the variation of land-surface (skin) temperature (form AVHRR-LST and MODIS-LST products) and wind patterns. A time lag of one month was used following the user requirements. The study period was from May to September and the study area covered Northern Africa, the Arabian Peninsula and Southern Europe. The research approach was semi-empirical. Pearson Correlation and Spearman Correlation were used to select candidate sites where land-surface temperature in the study area and rainfall in Malaga were significant correlated whereas Non-parametric Linear Regression was used and validated to evaluate the forecasting skills of those sites. The following sites were chosen for their significance: one site over Spain for the rain in May, one site over Burkina Faso and another over Libya for the rain in June. Among the selected sites, only the site over Burkina Faso was influenced by El Niño-Southern Oscillation (ENSO). The best model was the one for the site in Spain followed by the site in Libya and the site in Burkina Faso.

Acknowledgements

I would like to thanks to the University of Southampton, UK; Lund University, Sweden; University of Warsaw, Poland; and University of Twente-Faculty ITC, The Nederland; for giving me the opportunity to pursue the GEM MSc course. It has given me the chance to widen my knowledge about such as interesting issue as Geoinformation Science and Environmental Modelling.

Especial thanks to my supervisors; MSc Valentijn Venus and Prof. Dr. Ing. Wouter Verhoef for their guidance, advice and help. It has been a very good experience to work with them.

I would also like to thank my course lectures in Southampton, Lund, Warsaw and University of Twente-Faculty-ITC for their enthusiasm and professionalism in their teaching practices.

Especial thanks to Michael A. Bell, International Research Institute for Climate and Society (IRI), for his help at the time of manipulating the IRI/LDEO Climatic Data Library data.

Many, many thanks to my fellow students for their friendship and good company throughout the course. It has been a marvellous experience to have worked and have such as wonderful time together.

Finally, I am very grateful to my family and friends for their support and encouragement in the duration of the course.

Table of contents

List of figures

List of tables

List of abbreviations

AVHRR: Advanced Very High Resolution Radiometer
ENSO: El Niño-Southern Oscillation
ITCZ: Intertropical Convergence Zone
LST: Land-Surface Temperature
LSTA: Land-Surface Temperature Anomalies
MODIS: Moderate Resolution Imaging Spectroradiometer
NAO: North Atlantic Oscillation
NASA: National Aeronautics and Space Administration (USA)
NOAA: National Oceanic and Atmospheric Administration (USA)
NOAA. NCEP: National Oceanic and Atmospheric Administration. National Center of Environmental Prediction (USA)
NOAA.NOMADS: NOAA National Operational Model Archive and Distribution System (USA).
RA: Rainfall Anomalies
SLP: Sea Level Pressure
SST: Sea Surface Temperature
SSTA: Sea Surface Temperature Anomalies
WMO: World Meteorological Organization

1. Introduction

1.1. Background and Significance

Water is a natural resource essential for life on Earth. Unexpected variations of water availability can have important impacts not only in the natural environment but also in the social environment (agricultural yield, hydroelectric power and human water consumption among others). Thus, forecasting of seasonal rainfall is a very helpful input for policy planning that allows to improve the adaptive capacity to the oncoming new situations.

Regional climate change is characterized by a high level of uncertainty. This is due to the complexity of the processes involved not only at different spatial scales - planetary, regional and local- but also at different temporal scales -from sub-daily to multidecadal- (Giorgi et al., 2001). On a continental scale, rainfall distribution patterns are determined by the general circulation of the atmosphere which is driven by the solar energy and the gravitational energy. Hence, latitudinal variations in rainfall are driven by the pressure systems causing rain-bearing fronts. On the other hand, longitudinal variations in rainfall are caused by orography and the distribution of land that determine the potential for convective precipitation, ocean currents and sea-breeze systems. On a micro-scale level, urbanisation can cause highly localized rainfall anomalies (Phillips and McGregor, 2001).

The interest in the field of seasonal rainfall forecasting and the search for seasonal rainfall predictors has increased in the last years. Some authors have investigated the relationship between rainfall variability in the western Mediterranean Region (Iberian Peninsula) and some important known teleconnections such as El Niño-Southern Oscillation (ENSO) and North Atlantic Oscillation (NAO) (Rodo et al., 1997, Muñoz-Díaz and Rodrigo, 2003, Muñoz-Diaz and Rodrigo, 2004a, Muñoz-Díaz, 2004b, Muñoz-Díaz and Rodrigo, 2005, Muñoz-Díaz and Rodrigo, 2006, Frías et al., 2010).

ENSO is a climate pattern that occurs across the equatorial Pacific Ocean causing extreme weather disturbances in many regions of the world. It is characterized by variations in the sea-surface temperatures across the east-central equatorial Pacific Ocean (between 5^oN-5^oS and 170^oW-120^oW) and the associated variations in air surface pressure. El Niño is the extreme warm phase of ENSO while La Niña is the extreme cold phase (NOAA, 2010).

The NAO refers to a redistribution of atmospheric mass between the Arctic and the subtropical Atlantic. The positive phase of the NAO reflects below-normal heights and pressure across the high latitudes of the North Atlantic and above-normal heights and pressure over the central North Atlantic, the eastern United States and Western Europe. The negative phase reflects an opposite pattern. Both phases are associated with changes in the intensity and location of the North Atlantic jet stream, the storm track and the normal patterns of heat and moisture transport (Hurrell, 1995) affecting temperature and precipitation patterns from eastern North America to western and central Europe (Walker and Bliss, 1932, van Loon and Rogers, 1978, Rogers and Van Loon, 1979).

Rodo et al. (1997) studied the relationship between ENSO and NAO by using correlations and cross-spectral analysis (Katz, 1988), singular spectral analysis (Broomhead and King, 1986) and multi-taper method spectral analysis (Thomson, 1982). The authors affirmed that most of the Iberian Peninsula is under NAO influence in winter, with the exception of the eastern part that is positively correlated with ENSO. Furthermore, these authors found that the ENSO influence on the eastern part of Spain has increased over the last part of the 20^{th} century being the percentage of springtime variability even more than 50% on certain areas.

Other authors corroborated the NAO influence over the western part of the Iberian Peninsula in winter and stated that there is a higher probability of abundant rainfall during the negative phase of the NAO. Muñoz-Díaz (2004b) used empirical distribution functions to estimate the changes in the probability of wet and dry winter according to the NAO phases. Moreover, it has been said that the NAO influence is stronger during January in southern Spain (Muñoz-Díaz and Rodrigo, 2003).

Frías et al (2010) used a simple statistical test based on the observed and predicted tercile anomalies and deduced that El Niño causes dry and hot events in spring in the south while La Niña causes dry events in winter in the western part of Spain. However, Muñoz-Díaz and Rodrigo (2005) did not find influence of ENSO during winter but affirmed that in autumn El Niño causes null probability of drought and La Niña causes low probability of wet conditions (except in the north) and in summer La Niña leads to drought in the Southwest of Spain and to a low probability of wet conditions in the next autumn.

At the same time, the relationship between seasonal rainfall in Spain and smaller scale phenomena, such as Sea Level Pressure (SLP) within the region 30N-55N, 25W-20E, has been also studied (Muñoz-Díaz and Rodrigo, 2006). Principal component analysis (Kaufman and Rousseuw, 1990, Ahmed, 1997) and stepwise multiple regression analysis were used to build a model. Rainfall variability in western Iberian Peninsula are explained by variations in the SLP field during winter and spring (around 68%, and 57% of the variability respectively). In spite of all these studies, seasonal rainfall variability in Spain has not been completely explained.

Moisture and thermal advection are, in most of the cases, highly related to precipitation events (University of Illinois, 2010). Thermal advection is the transport of sensible or latent heat by a moving fluid, such as air. Thermal advection is equal to the negative wind vector (–U) multiply by the vector temperature gradient (ΔT). Wind is the flow of air on a large scale caused by the Pressure gradient force together with Coriolis force, Frictional forces and Rotational forces. Three factors, then, make the thermal advection larger: a stronger wind, a larger temperature gradient and a smaller angle between wind direction and temperature gradient (wind blowing normal to the isotherms). Thus, warm advection (warm air transported) refers to winds blowing from warm to cold regions and it is associated with ascending motion as well as cloud or precipitation. Cold advection (cold air transported) refers to winds blowing from cold to warm regions and it is associated with descending motion and clear conditions (Lyndon State College Atmospheric Sciences, 2010). Besides, the probability of heavy precipitation occurrence increases if a cyclone is supplied with an abundance of moisture. Regions of moisture advection are often co-located with regions of warm advection (University of Illinois, 2010).

Malaga is located within the Mediterranean Region with a Mediterranean climate regime with wet winters and dry summers. This region is between the climate conditions of the temperate westerlies (which dominate over central and northern Europe), and the subtropical high pressure belt over North Africa (Figure 1 **(**modified from Barry and Chorly, 1992 in (Harding et al., 2009)). In summer, the subtropical high pressure conditions are displaced from the North of Africa and the Mediterranean comes under the influence of the easterlies. Polar front depressions occasionally may reach the western Mediterranean (Rohling and Hilgen, 1991). During winter, the subtropical conditions are displaced southward, and the Mediterranean is influenced by the temperate westerlies with the associated Atlantic depressions. Most of the rain falls from May to September. However, some rainfall events have taken place during the months of June and August.

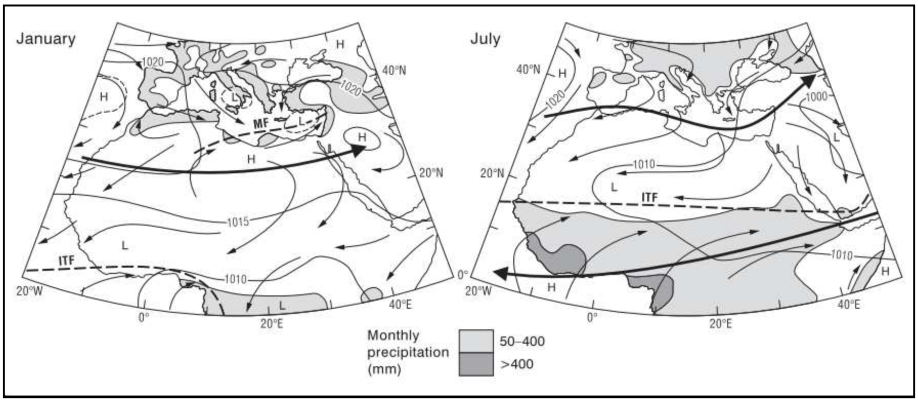

Figure 1. The location of the Mediterranean region in relation to the large scale atmospheric circulation

Because the climate in Malaga is highly influenced by easterlies during the summer period and being the moisture and/or thermal advection caused by the wind and thermal vector, the best period to study the relationship between rainfall variations and land-surface temperature variations over the continents (Northern Africa, the Arabian Peninsula and Southern Europe) would be from May to September.

Accordingly, the significance of this research relies on the following premises:

- the significant relationship between land-surface temperature variations in the study area and rainfall in Malaga
- the influence of wind vector over the area
- the environmental variables which can define the type of advection (moisture and/or thermal advection)
- the influence of ENSO
- the time-lag between land-surface temperature and rainfall
- if this time-lag can be useful for water management
- the identification of predictor sites for seasonal rainfall in Malaga

1.1.1. Conceptual Framework

The climate system is determined by the net radiation. The Earth receives and absorbs energy from the Sun in the form of electromagnetic radiation (mostly light and ultraviolet energy) and re-radiates heat back to the atmosphere and into space (as infrared radiation). Moreover, energy transformations will be different depending on the surface material (ice, water and land) and characteristics (topography, land use, land cover). The difference in the incoming solar radiation between the equator and the poles makes climate to vary with latitude. This phenomenon generates mechanisms of the energy and mass exchange such as evaporation, advection and heat exchange (Figure 2 (Bridgman and Oliver, 2006)).

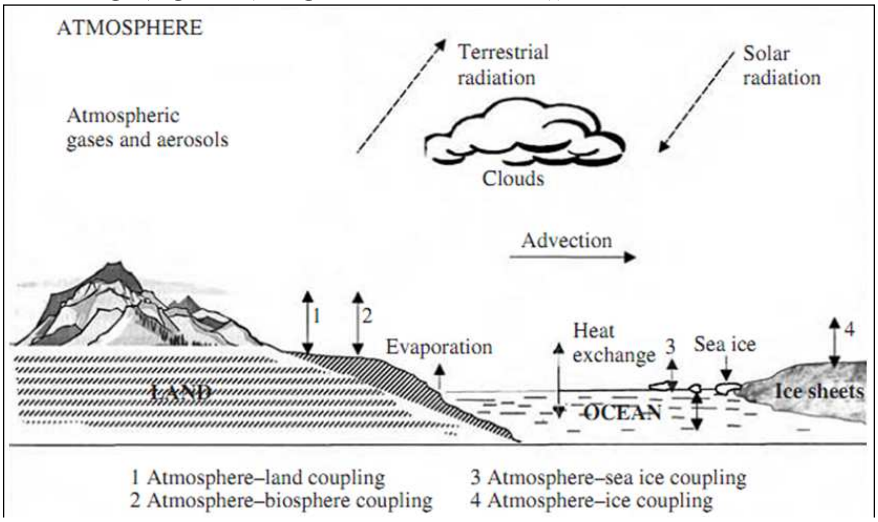

Figure 2. The Climatic System. Energy and Mass Exchange

1.1.1.1. Energy Exchange

The energy balance equation states that the energy arriving at the surface must be equal to the energy leaving the surface for the same time period. Thus:

$$\boldsymbol{R_n = G + LE + H} \quad \textbf{(Equation 1)}$$

Where:

R_n is the net radiation flux density (W/m^2)
G soil heat flux density (W/m^2)
LE is latent heat flux density (W/m^2)
H is sensible heat flux density (W/m^2)

Net radiation (R_n) is defined as the difference between incoming and outgoing long and shortwave radiation on the Earth's surface at a certain moment in time. Soil heat flux density (G) is the rate of flow of heat energy into, from or through the soil. Latent heat flux density (LE) is the flux of heat from the earth's surface to the atmosphere that is associated with evaporation or condensation of water vapour at the surface (soil and vegetation). Sensible heat flux (H) is the transference of heat from the surface to the atmosphere that is not associated with phase changes of water but is associated to the change of temperature of the air.
The sensible heat flux (H) equation is as follow:

$$H = \rho \, C_p \frac{[T_{aero} - T_{air}]}{r_{aero}} \quad \textbf{(Equation 2)}$$

Where:

ρ is the air density (kg/m^3)
C_p is the air specific heat at constant pressure (J/kg K)
r_{aero} is the aerodynamic resistance to heat transport between the surface and the reference level (s/m)
T_{aero} is the aerodynamic (land-surface) temperature (K)
T_{air} is the air temperature at the measurement height (K)

Sensible heat flux (H) is proportional to the difference between aerodynamic temperature and absolute temperature of the air at a measurement height. Aerodynamic resistance to heat transport (r_{aero}) is determined by wind speed, surface roughness, displacement height, and the thermal instability of the atmosphere. Reference heights for temperature and aerodynamic resistance must be identical to express sensible heat flux (H) (Norman and Becker, 1995).

This study attempts to model semi-empirically the energy exchange and water cycle by isolating only one parameter from the energy balance equation. This parameter is the aerodynamic temperature (land-surface temperature) which is an essential factor that influences moisture/air motion. This influence will be highlighted through the wind vectors. Considering that the present study has been carried out at a regional scale and taking into account that the atmosphere is the best model in itself, the statistical analysis will allow to see the relationship between land-surface temperature variations in the candidate predictor sites and seasonal rainfall in Malaga.

1.1.2. Land-Surface Temperature

1.1.2.1. Aerodynamic Temperature versus Land-Surface (Skin) Temperature

The term Aerodynamic Temperature relates to the efficiency of heat exchange between the land surface and overlying atmosphere within the Energy Balance Equation (Kustas et al., 2007). On the other hand, Land-Surface Temperature or Skin Temperature refers to the weighted soil and canopy radiation emitted and reflected into the sensor. This temperature is captured by a narrow wavelength band from the Instantaneous Field of View (IFOV) and from a specific angle.

Aerodynamic temperature may fall between air surface temperature and skin temperature. However, skin daily temperature tends to be higher than aerodynamic temperature at midday and lower than aerodynamic temperature at night (Sun and Mahrt, 1995). Huband and Monteith (1986) and Chehbouni et al. (1996) modelled aerodynamic temperature from skin temperature while other authors assumed a thin boundary layer over the leaves or soil where molecular diffusion generates the difference between aerodynamic temperature and skin temperature (Oleson et al., 2008).

Norman and Becker (1995) stated that land-surface (skin) temperature is equivalent to aerodynamic temperature when land surface is homogeneous and is in thermal equilibrium within the Instantaneous Point of View of the sensor.

1.1.2.2. Land-Surface (Skin) Temperature retrieval

The main advantage of using land-surface (skin) temperature from remotely sensed data is that it provides a better spatial footprint of environmental variables by reducing point observational biases and providing new estimates in areas which had not been observed before (Legates, 2000). The main disadvantage is that land-surface (skin) temperature is not comparable to aerodynamic temperature for heterogeneous surfaces which are not in equilibrium within the Instantaneous Point of View of the sensor (Norman and Becker, 1995).

Accurate retrieval of land surface temperature from satellite images is challenging due to the atmospheric attenuation (absorption and emission) of thermal radiation, and the nonblack-body property of the observed land-surface. The atmospheric attenuation, especially the attenuation caused by the presence of water vapour, affects the transmission of the emitted radiation from the Earth to the satellite sensor (Bastiaanssen, 1995, Qin and Karnieli, 1999). To correct the absorption of atmospheric water vapour, a so-called 'split-window' technique is commonly applied by using split-data in the far-thermal infrared range (10-13 μm) (Caselles et al., 1997, Qin and Karnieli, 1999, Parodi, 2000).

The data used in the present study, AVHRR LST for Africa (Pinheiro et al., 2006) and MODIS LST (Salisbury et al., 2002), are based in the following algorithms: Ulivieri et al. split-window algorithm (Ulivieri et al., 1994) for band 4 and 5 for AVHRR and Wan and Dozier split-window algorithm (Wan and Dozier, 1996, Wan, 2008) for bands 31 and 32 for MODIS LST.

1.2. Research Problem

On a regional level, variations in seasonal rainfall are still difficult to forecast due to the complexity of the atmospheric phenomena. Although it has been demonstrated that seasonal rainfall variability is linked to global atmospheric processes (NAO, ENSO) and smaller scale processes (SLP) within the Iberian Peninsula, there is still a part of this variability that has not been explained yet. Since moisture and thermal advection (caused by wind and temperature gradient), may be related to this seasonal rainfall variability, to explore the relationship between land surface temperature variations within the study area and seasonal rainfall variability in Malaga is a challenge.

1.3. Assumptions

This research is based on the following assumptions:

- the surfaces under study are homogeneous and in thermal equilibrium
- the air density (ρ), the air specific heat at constant pressure (C_p), the aerodynamic resistance to heat transport (r_{aero}), the air temperature at the measurement height (T_{air},) the soil heat flux density (G), and the latent heat flux density (LE) are constant.

1.4. Research Questions and Objectives

Research terminology:

- Candidate sites: possible moisture and/or thermal advection sites where LSTA´s are significant correlated with RA´s in Malaga and where the wind direction points to Malaga.
- Predictor sites: candidate sites which are suitable to forecast RA´s and therefore suitable predictor sites.

The research questions and objectives are shown in the following table:

Table 1. Research objectives and research questions

Research Objectives	Research Questions
To study if there is a significant relationship between LSTA´s over candidate sites and RA´s in Malaga.	Is there a significant relationship between LSTA´s over candidate sites and RA´s in Malaga?
To study if the candidate sites are not influenced by El Niño-Southern Oscillation (ENSO) and can be used as a independent predictor of rainfall in Malaga.	Is there a significant relationship between rainfall in candidate sites and Sea Surface Temperature Anomalies in El Niño Regions?
To evaluate the forecasting skills of the candidate sites by building and validating a bivariate model LSTA-RA.	Are the candidate sites suitable enough to forecast seasonal rainfall in Malaga?
To find out if the candidate sites are sources of moisture and/or thermal advection by analysing other environmental variables (monthly evapotranspiration, monthly moisture over the surface, monthly atmospheric moisture column and monthly rainfall).	Are the candidate sites sources of moisture and/or thermal advection?
To find out if the time lag between LSTA´s over the candidate sites and RA´s in Malaga is sufficient to improve water management in Andalusia.	Is the time lag between the LSTA´s in the candidate sites and RA´s sufficient to satisfy user-requirements from water managers?

1.5. Hypothesis

Hypothesis 1:
This Hypothesis is stated to test if there is a significant correlation between LSTA´s in candidate sites and RA´s in Malaga.

H_0= There is no significant correlation between LSTA´s over the candidate sites and RA´s in Malaga.
H_a= There is a significant correlation between LSTA´s over the candidate sites and RA´s in Malaga.

Hypothesis 2:
This Hypothesis is stated to test if there is no significant correlation between rainfall in the candidate sites and Sea Surface Temperature Anomalies (SSTA) in El Niño Regions. If it is so, these sites can be used to forecast RA´s in Malaga independently from the ENSO phenomena.

H_0= There is a significant correlation between rainfall in the candidate sites and Sea Surface Temperature Anomalies in El Niño Regions.
H_a= There is no significant correlation between rainfall in the candidate sites and Sea Surface Temperature Anomalies in El Niño Regions.

Hypothesis 3:
This Hypothesis is stated to test if the time lag between the LSTA´s in the candidate sites and RA´s is sufficient for water management purposes.

H_0= The time lag between the LSTA´s in the candidate sites and RA´s is not sufficient for water management purposes.
H_a= The time lag between the LSTA´s in the candidate sites and RA´s is sufficient for water management purposes.

2. Materials and Methods

2.1. Research Approach

First of all, in order to choose the best time lag a statistical analysis was carried out. There was not any significant time lag, thus the maximum time lag required by the users was chosen. Secondly, the pre-selection of candidate sites was carried out after computing a correlation analysis (LSTA´s-RA´s) in IRI/LDEO Climate Data Library (Columbia University, 2010). The computer language used was the Ingrid Language (IRI/LDEO Climatic Data Library, 2010). Correlation maps were obtained with a correlation coefficient value for every pixel. Only data available in this library could be used to compute the correlation, so AVHRR from 1995-1999 and MODIS from 2003-2010 were selected. These two data sets had to be analysed separately.

Although the data were not normally distributed (Apendix 1), Pearson correlation (Equation 3) was chosen to pre-select the candidate sites because it showed to be more effective than Spearman's. Since there was not previous knowledge of the system, the way the data should be grouped was tested. The correlation for Seasonal LSTA´s and RA´s averages (May-Sep) showed less significant areas and lower correlation coefficients than the results for every month. Thus, every month was analyzed separately.

Only those sites showing significant correlation with the same sign in both data sets were pre-selected for further analysis. Those pre-selected sites showing a significant Spearman Correlation Coefficient for the whole study period (1995-1999 and 2003-2010) were considered the final candidate sites. Spearman Correlation Coefficient was calculated from LSTA´s and RA´s ranking scores. First, the spatial LST average was obtained for every year under study. These data were used to get the LST average. LSTA's (Equation 4) and RA's (Equation 5) were calculated for every site.

To study the independence of the final candidate sites from ENSO, Pearson Correlation between SSTA (between 5^oN-5^oS and 170^oW-120^oW) and Rainfall was computed in such sites. Furthermore, to obtain a better understanding of the possible causality under the correlation some environmental variables were considered. Finally, a nonparametric model consisting in linear regression of the scores of the two variables was built and validated for every candidate sites (Sheskin, 2000).The analysis and map preparations were done in ArcGIS.10 and the statistic analysis was completed using XLSTATS add-in for Excel 2007 available at http://www.xlstat.com/.

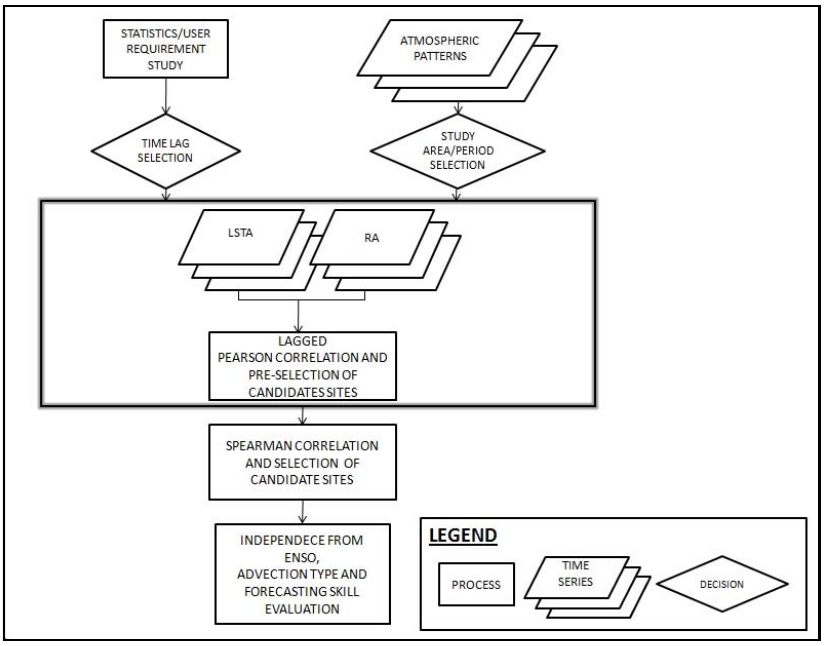

Figure 3. Research Approach. General overview

The next figure illustrates the process and data used in detail.

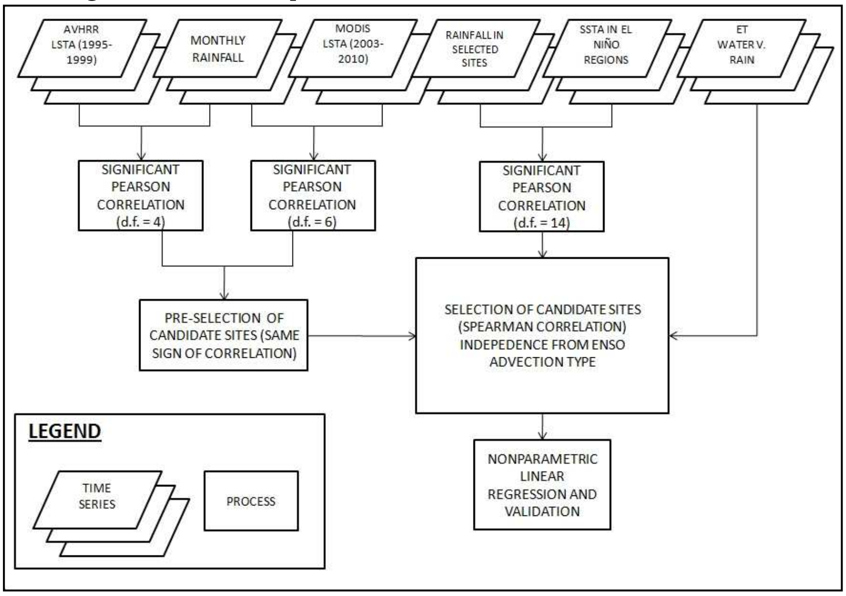

Figure 4. Methodology. Detail of the analysis

(note: ET = Evapotranspiration, Water V. = Water Vapour).

2.1.1. Statistical tests and assumptions

2.1.1.1. Pearson Correlation

The correlation is defined as the measure of linear association between two variables. The correlation coefficient is bounded by -1 and 1. If the correlation is exactly -1, there is a perfect, negative linear association between the two variables. Conversely, if the correlation is exactly 1, there is a perfect, positive linear correlation. Secondly, the square of the correlation describes the amount of variability in one variable that is described by the other variable. Correlation does not imply causation or a physical relationship of any kind, correlations are only associated with observed instances of events.

Pearson-Product Moment Correlation coefficient (r) (Pearson, 1896, 1900) is calculated as follows:

$$r = \frac{1}{n-1} \sum_{i=1}^{n} \left(\frac{x_i - \bar{x}}{s_x}\right)\left(\frac{y_i - \bar{y}}{s_y}\right) \quad \textbf{(Equation 3)}$$

Anomalies are those values above or below average. Then:

$$LSTAi = (LSTi - \overline{LST}) \quad \textbf{(Equation 4)}$$

$$RAi = (Ri - \bar{R}) \quad \textbf{(Equation 5)}$$

Lagged Correlation LSTA-RA would then be:

$$\mathrm{r} = \frac{1}{\mathrm{n}-1} \sum_{\mathrm{i}=1}^{\mathrm{n}} \left(\frac{\mathrm{LSTA}_{(\mathrm{i}-\mathrm{lag})} - \overline{\mathrm{LSTA}}}{\mathrm{S}_{\mathrm{LSTA}}}\right)\left(\frac{\mathrm{RA}_{\mathrm{i}} - \overline{\mathrm{RA}}}{\mathrm{S}_{\mathrm{RA}}}\right) \quad \textbf{(Equation 6)}$$

where:

i is the time

$$\overline{\mathrm{LSTA}} = \frac{1}{\mathrm{n}} \sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{LSTAi} = 0 \quad \textbf{(Equation 7)}$$

$$\overline{\mathrm{RA}} = \frac{1}{\mathrm{n}} \sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{RAi} = 0 \quad \textbf{(Equation 8)}$$

Note that the mean of the anomalies is always zero and the Standard Deviation of the anomalies has the same value than the Standard Deviation of the sample. Then, when the correlation is computed for only one month, the result of the sample is exactly the same than the correlation of its anomalies. Nevertheless, the use of anomalies is more suitable to visualize and interpret the statistical results.

The next table shows the minimum threshold for the Pearson Correlation Coefficient (r) at a given significance level and degree of freedom (Snedecor GW and Cochran W.G., 1980).

Table 2. Significant levels for Pearson Correlation Coefficient

Significant Level	*r (df = 4)*	*r (df = 6)*	*r (df = 12)*
90%	0.729	0.622	0.458
95%	0.811	0.707	0.532
98%	0.882	0.789	0.612
99%	0.917	0.834	0.661

The Pearson product-moment correlation coefficient is based on the following assumptions: a) The sample of *n* subjects for which the value *r* is computed, it is randomly selected from the population it represents; b) The level of measurement upon which every of the variables is based is interval or ratio; c) The two variables have a bivariate normal distribution (every of the variables and the linear combination of the two variables are normally distributed); d) Existence of homoscedasticity. Homoscedasticity exists in a set of data if the relationship between the X and Y variables is of equal strength across the whole range of both variables.

2.1.1.2. Spearman Correlation

Spearman Correlation Coefficient (Spearman, 1904) is calculated using the same equation than in Pearson but instead of the LSTA's and RA's values, the ranking score of those are used.

The next table shows the minimum threshold for the Spearman Correlation Coefficient (rho) at a given significance level and degree of freedom.

Table 3. Significant levels for Spearman Correlation Coefficient

Significant Level	*rho (df = 4)*	*rho (df = 6)*	*rho (df = 12)*
90%	0.829	0.643	0.464
95%	0.886	0.738	0.538
98%	0.943	0.833	0.622
99%	1.000	0.881	0.675

Spearman's rank-order correlation coefficient assumes that the ratio data are rank-ordered. It is used when one or more of the assumptions of the Pearson product-moment correlation coefficient have been saliently violated.

2.1.2. Linear regression, Analysis of Variance and Goodness of fit statistics

2.1.2.1. Linear regression, Analysis of Variance (ANOVA)

Linear regression is used to model the relationship between a scalar variable and one or more variables denoted x. One of its uses is prediction or forecasting. In linear regression, data are modelled using linear functions, and unknown model parameters are estimated from the data. Linear regression focuses on the conditional probability distribution of y given x. The method used is the least square method (Montgomery and Peck, 1992). It is shown is the following equations:

$$\hat{y} = a + bx + e \quad \textbf{(Equation 9)}$$

$$b = r\,\frac{S_x}{S_y} \quad \textbf{(Equation 10)}$$

$$a = \bar{y} - b\bar{x} \quad \textbf{(Equation 11)}$$

Where:

a	is the intercept
b	is the slope
e	is the error
S_x	is the variance of x
S_x	is the variance of y
$\bar{y}$	is y average
$\bar{x}$	is x average

The analysis of variance (ANOVA) is a method to test the significance of the regression. This approach uses the variance of the observed data to determine if a regression model can be applied to the observed data. The observed variance is partitioned into components (Table 4) that are then used in the test for the significance of the regression. Thus, being β the slope of the regression, the null hypothesis is H_0: $\beta = 0$ and the alternative hypothesis is H_a: $\beta \neq 0$. If the *p-value* resulting in the ANOVA is lower than the significance level a, the null hypothesis is rejected and the regression results to be significant. The next table is the ANOVA table where y_i is the observe value, $\bar{y}$ is the average of the observed values and $\hat{y}_i$ is the predicted value.

Table 4. ANOVA table

Source	Df	Sum of Squares (SS)	Mean Squares (MS)	F	Pr> F
Model/Regression (R)	*P*	$\sum_{i=1}^{n}(\hat{y}_i - \overline{y})^2$	$\frac{\sum_{i=1}^{n}(\hat{y}_i - \overline{y})^2}{p}$	$\frac{MS_{Model}}{MS_{Error}}$	*p-value*
Error (E)	*n*-*p*-1	$\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$	$\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n-p-1}$		
Corrected Total (T)	*n*-1	$\sum_{i=1}^{n}(y_i - \overline{y})^2$	$\frac{\sum_{i=1}^{n}(y_i - \overline{y})^2}{n-1}$		

The next figure illustrates the Sum of Squares for the Model (SSR), the Sum of Squares for the Error (SSE) and the Sum of Squares Total (SST).

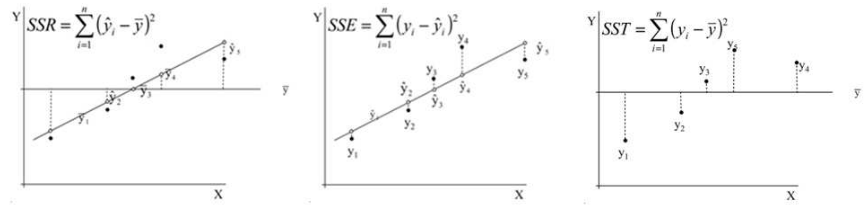

Figure 5. Sum of Squares Model (SSR), Sum of Squares Error (SSE) and Sum of Squared Total (SST)

The principal assumptions for linear regression and analysis of variance in linear regression are: a) linearity of the relationship between dependent and independent variables; b) independence of the errors (no serial correlation); c) homoscedasticity (constant variance) of the errors (versus time, versus the predictions (or versus any independent variable)); d) normality of the error distribution (Apendix 2). If any of these assumptions is violated, then the forecasts, confidence intervals, and insights yielded by the regression model may be inefficient or seriously biased or misleading.

2.1.2.2. Goodness of fit statistics

a) Coefficient of determination (R^2) and adjusted R^2:
The coefficient of determination is a statistical measure of how well the regression line approximates the real data points. An R^2 of 1 indicates that the regression line perfectly fits the data. Adjusted R^2 is a modification of R^2 for the number of explanatory terms in a model. Unlike R^2, the adjusted R^2 increases only if a new term improves the model more than it would be expected by chance. Adjusted R^2 not only varies depending on the size of the R^2 and the number of independent variables,

but also on the sample size (n). So this corrects for the fact that standard regression overestimates population parameters.

$$R^2 = 1 - \frac{SSE}{SST} \quad \textbf{(Equation 12)}$$

Where:
SSE is the Sum of Squares Error
SST is the Sum of Squares Total

$$R^2{}_{Adjusted} = 1 - (1 - R^2)\frac{n-1}{n—k-1} \quad \textbf{(Equation 13)}$$

Where:
n is the sample size
k is the number of independent variables

b) Mean Square Error (MSE):
Mean Square Error assesses the quality of an estimator in terms of its variation and unbiasedness. An MSE of zero, meaning that the estimator predicts observations of the parameter with perfect accuracy, would be the ideal result, but in fact it is never occurs. Values of MSE may be used for comparative purposes (Equation in Table 4).

c) Root Mean Square Error (RMSE):
The root mean square error is defined as the square root of the mean square error. It is a frequently-used measure of the differences between values predicted by a model or an estimator and the values actually observed from the variable being modelled or estimated. RMSE is a good measure of precision. These individual differences are also called residuals, and the RMSE serves to aggregate them into a single measure of predictive power.

d) Prediction Residual Sum of Squares (PRESS) and Root Mean Squared Prediction Error (PRESS RMES):
As in Allen (1971) with *n* as the sample size, the model equation is fitted to *n-1* and a prediction taken for the remaining one. The difference between the recorded data value and the value given by this model is called a prediction residual. PRESS is the sum of squares of the prediction residuals. The square root of PRESS is PRESS RMSE (root mean square prediction error). The PRESS statistic gives a good indication of the predictive power of the model. Minimizing PRESS is desirable. Overfitting can be evaluated by comparing PRESS RMSE with RMSE. The PRESS statistic is a surrogate measure of crossvalidation of small sample sizes and a measure for internal validity. Small values indicate that the model is not overly sensitive to any single data point.

d) Bias
Bias of an estimator is the difference between this estimator's expected value and the true value of the parameter being estimated.

$$Bias = \frac{\sum_{i=1}^{n}(\widehat{y_i} - y_i)^2}{n} \quad \textbf{(Equation 14)}$$

Where:

$\hat{y}_i$ is the predicted value

y_i is the observe value

n is the sample

2.2. Study Area

The study area was selected according to the atmospheric pattern. As it has been mentioned before, due to its latitude Malaga can be affected by the westerlies (mostly in winter) and by the easterlies (mostly in summer). When it is affected by the westerlies, because of its position in the extreme of Western Europe, the air masses have maritime origin (Atlantic Ocean). Sometimes, the polar continental air mass coming from Siberia can reach Spain, but rarely reaches Malaga. When Malaga is affected by the easterlies, the main air mass comes from African continent (Sahara) (University of Valencia, 2010). Since the objective of this research is to study LSTA´s over the continents and its influence over RA´s in Malaga, the study area selected was mainly northern Africa. Larger extension was considered since atmospheric processes can have further influences. Thus, the study area is encompassed within 0ºN- 40ºN latitude and 20ºW-60ºE longitude.

2.2.1. Physical Geography

The study area extent and physical characteristics are shown in the following figure:

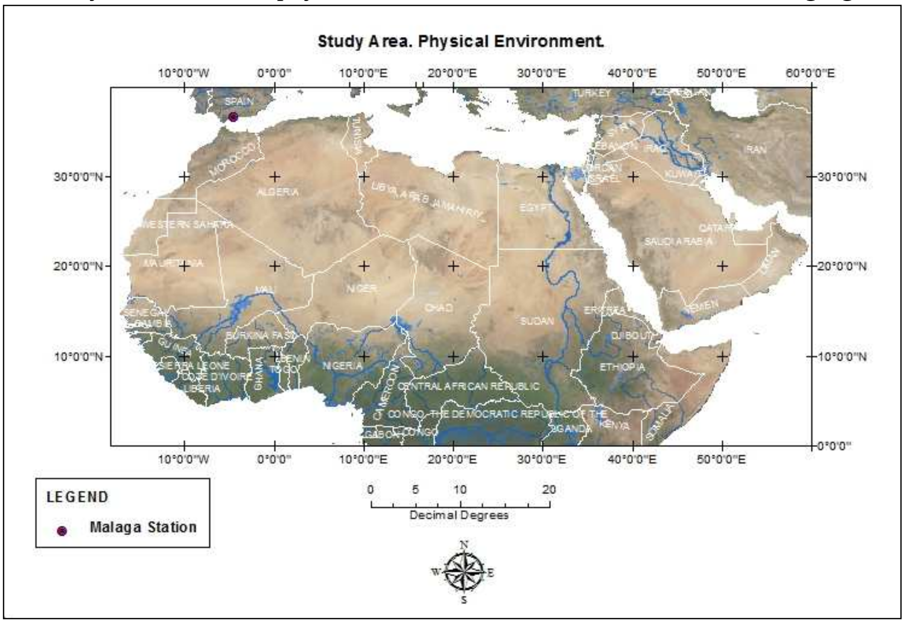

Figure 6. Study Area. Physical Environment.

2.2.1.1. Climatic Zones and Permanent Water Sources

There are four different climatic zones within the study area which correspond with vegetation density: Mediterranean (Northern Coast of Africa and South of Europe), Arid (Sahara and Arabian Desert), Semiarid (Northern Africa and Sahel) and Tropical (Forest of the equatorial region). The classical Mediterranean climate, where Andalusia is included, is characterized by warm and dry summers, and mild and wet winters. It is opposite to the tropical monsoon climates, which comprises a pluvial maximum in the warm months. The main permanent water sources are in the northern and southern part of the study area. In the middle part, the Sahara desert and Arabian Desert do not present any permanent source of water except for the Nile.

2.3. Data Available and Collected Data

2.3.1. Data Available

The next table summarizes the data available. Most of the data were available in IRI/LDEO Climate Data Library (Columbia University, 2010) and in MODIS Atmosphere websites (NASA, 2010b).

Table 5. Data available: Source and Characteristics

Variable	**Rainfall**	**Land Surface Temp. (LST)**		**Water Vapour Column**	**Wind (U/V)**	**SSTA**	**Evaporation and Water Vapour at the Surface**
Source	Climate Anomaly Monitoring System (NOAA)	NOAA	NASA	NASA	NOAA.NOMADS	NOAA. NCEP	NASA
Product Name	Station Precipitation from Station	AVHRR-LST (Pinheiro et al., 2006)	MYD11A2 v-005 (Aqua) (Wan and Zhao-Liang, 1997)	MYD08_M3 (Aqua)	20^{th} Century Reanalysis (Compo et al., 2006)	Reyn_Smith v2 .monthly .ssta (Reynolds et al., 2002)	NOAH025_M (NASA, 2010a)
Available in	IRI/LDEO	IRI/LDEO	IRI/LDEO	MODIS ATM.	IRI/LDEO	IRI/LDEO	IRI/LDEO
Available from	1981	1995-2000*	2002-2010	2002	1870-2008	1981	Feb 2000
Spatial Res.	-	8 km**	1 km**	5 km	2 degrees	1 degree	0.25 degree
Temporal Res.	Monthly	Daily	8 Days	2/day	Monthly	Monthly	Monthly

*In the visualisation only the years from 1995 to 1999 are considered due to computational problems.

** Due to computational problems the data resolution had to be decreased to 0.10 degrees.

2.3.1.1. Land-Surface Temperature (Predictor)

2.3.1.1.1. Data Specifications

Two different products were used to carry out the analysis: AVHRR day-time Land-Surface Temperature and MODIS Aqua day-time Land-Surface Temperature. Both data sets were available in the IRI data library (MODIS Terra was not).

AVHRR Land-Surface Temperature dataset corresponds to the daily, daytime NOAA-14 AVHRR land surface temperature (LST) over continental Africa for the period 1995-2000 (Pinheiro et al., 2006). The local equatorial crossing time is from 12:00 to 14:00 (approximately).

MODIS Aqua Land-Surface Temperature MYD11A2 version_005 (LST) 8-day data are composed from the daily 1-kilometer LST product (MYD11A1) and stored on a 1-km grid as the average values of clear-sky LSTs during an 8-day period. The local equatorial crossing time is approximately 13:30 in an ascending node with a sun-synchronous, near-polar, circular orbit. The algorithm used in this product was developed by Wan, Dozier and Zhao-Liand (Wan and Dozier, 1996, Wan and Zhao-Liang, 1997, Wan, 2008). Although the first image available is in July 2002, the period used was from 2003 to 2010 in order to have the same degree of freedom for every month (14 years).

2.3.1.1.2. Data Quality and Data Comparison

Pinheiro et al. (2006) validated AVHRR LST product over a savanna field site. An uncertainty below 1.5 K for daytime retrievals was found. Nevertheless, the authors suggested a more robust validation for further evaluation.

According to Wan (2008) MODIS LST V5 was validated in 47 clear-sky cases being the accuracy of the MODIS LST product better than 1 K in most cases and the root of mean squares of differences less than 0.7 K for all cases. They stated that the quantity and quality of MODIS LST products depend on clear-sky conditions due to the limitation of the thermal infrared remote sensing. Nevertheless, later on the author (2010) included the possibility of errors in desert regions due to the uncertainties in the classification-based emissivity values. Other examples of validation can be found (Coll et al., 2009).

AVHRR and MODIS have similar local passing time. Comparison analysis could not be carried out since the dates when the tow dataset were available did not match. Nevertheless, Zhong et al. (2010) estimated LST over the Tibetan Plateau by using two split-window algorithms, one for AVHR, and the other for MODIS simultaneously. In the validation process they obtained an average percentage error (PE) of 10.5% for AVHRR and 8.3% for MODIS. The results from AVHRR agreed with MODIS, but the latter displayed a higher level of accuracy.

2.3.1.2. Rainfall (Predictand)

The next table shows Malaga station characteristics:

Table 6. Malaga Station Characteristics

Station (IWMO) Code	*Name*	*Longitude*	*Latitude*	*Elevation (m)*
8482	MALAGA/AEROPUERTO	4.48W	36.67N	16

Due to the LST availability, the years under study were divided into two periods: 1995-2000 and 2003-2010. Monthly rainfall (mm) from May to September is illustrated in the following figure. July was excluded because the value in those years was zero except for July 2003 which was just 3 mm. The months with higher rain values are May and September. The years 2000, 2003, 2004, 2005 and 2006 shown less amount of rainfall compared with the rest of the study period.

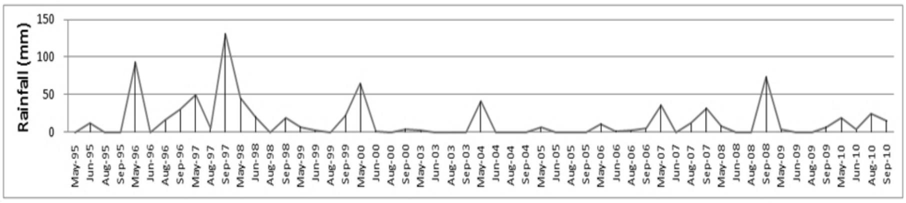

Figure 7. Monthly rainfall in Malaga

The next table shows the summary statistics of rainfall for the study period. All the data series was completed. Although the data were obtained from an official source, no information about the quality was provided. The maximum value was registered in September 1997 (131 mm).

Table 7. Summary Statistics for Rainfall in Malaga

Variable	N	Missing data	Min	Max	Mean	Std. Deviation
Rainfall (mm)	69	0	0.0	131.0	11.971	23.754

2.3.1.3. Wind

The wind data were obtained from the 20th Century Reanalysis (Compo et al., 2006). This data set contains the U and V wind components for different pressure levels (from 1000mb to 10mb) at 2 degrees spatial resolution. The 850mb pressure level is generally used to diagnose thermal advection forcing precipitation systems. This level is generally above the boundary layer, so that winds are unaffected by surface friction, yet low enough to reflect the stronger thermal gradients near the surface (McGill University, 2003).

2.3.2. User-requirement Study

In order to know which was the best time lag to be used in this research, a user requirement study was carried out. Stakeholders from agriculture, reservoirs and natural areas sector, were contacted by telephone. Then, the survey was sent to them by e-mail. Among other questions, the most important was: “how long in advance would you need the forecasting information in order to improve water management in Andalusia?” Several options were given: at least 1 week, at least 2 weeks, at least 3 weeks, at least 1month or others.

3. Results and discussion

3.1. Time-Lag Selection: Analysis and user-requirement study

For the selection of the time lag between LSTA and RA in Malaga a preliminary statistical analysis was carried out. The analysis showed that there was not any significant time lag. Thus, it was decided to choose the maximum time lag so all user demands could be satisfied.

The survey was sent to 35 users, answers from 24 were obtained. The next figure shows the user-requirement study results. The maximum time in advance the users would need was 1 month. Thus, this was the chosen time lag to carry out the analysis.

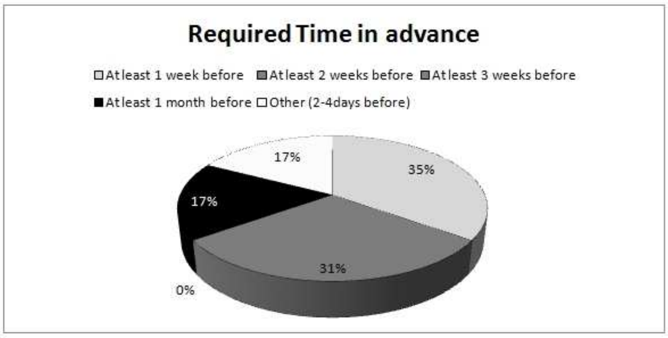

Figure 8. Time lag required by users

According to the information above, if the candidate sites shown as suitable predictor of rainfall in Malaga, then the Null Hypothesis (H_0) of the *Hypothesis 3* would be rejected and the time lag between LSTA´s in the candidate sites and RA´s would be sufficient for water management purposes.

3.2. Predictor Site Selection and Evaluation

As it has been mentioned before, the correlation for Seasonal LSTA´s and RA´s averages (May-Sep) showed less significant areas and lower correlation coefficients than the results for every month. These differences could be due to the dilution of the information after data averaging. The main advantage of considering the results month by month is that the users could have the forecasting information updated every month.

The used of Pearson correlation to identify the sites resulted to be more effective than Spearman's, because when transforming interval/ratio data into a rank-order some information is lost. Pearson Correlation is generally a more powerful test than Spearman's (Sheskin, 2000).

The proposed method to pre-select candidate site (merging areas with the same correlation sign for every of the data set) allowed good results for rainfall in May and June but not in August and September. This is because the correlation was computed for every data set independently. The fact that a high correlation is obtained for every data set in the same site, does not imply a high correlation after merging both data. A high difference between the mean and standard deviation, seen together and individually, causes also a high difference in their correlation values too.

On the other hand, the two data sets after correlation showed different outputs. In most of the cases, AVHRR highlighted larger areas than MODIS. Moreover, after the rejection of the areas with different correlation signs only a few areas remained. These differences in the output could be due to the data quality and differences. MODIS quality is slightly higher that AVHRR (around 0.5 k) (Pinheiro et al., 2006, Wan, 2008, Zhong et al., 2010). This fact could influence the sign of the anomaly and consequently the correlation value.

In spite of all, one candidate site was identified for the RA in May (Spain) and two candidate sites for the RA in June (in Burkina Faso and Libya).

3.2.1. Selection and evaluation of candidate sites for RA´s in May

First of all, the pre-selection of candidate sites for RA´s in May was carried out. The method proposed here showed to be suitable for this purpose allowing to select one final candidate site over Spain.

Secondly, the influence of ENSO over the candidate site was analyzed. Also, some environmental variables were studied to determine the type of advection occurring over the site.

Finally, the candidate site forecasting skills where evaluated.

3.2.1.1. Pre-selection of Candidate Sites for RA´s May

The next figure shows the map from the Pearson Product Moment Correlation between LSTA-April and RA-May once the results of the two data sets were merged. There are some significant negatively correlated areas in Spain and Northern Africa as well as some significant positively correlated areas below 20ºN longitude. Furthermore, the wind direction vectors over the selected sites points to Malaga. Among all the different sites, the site over Spain was selected because of its significant size.

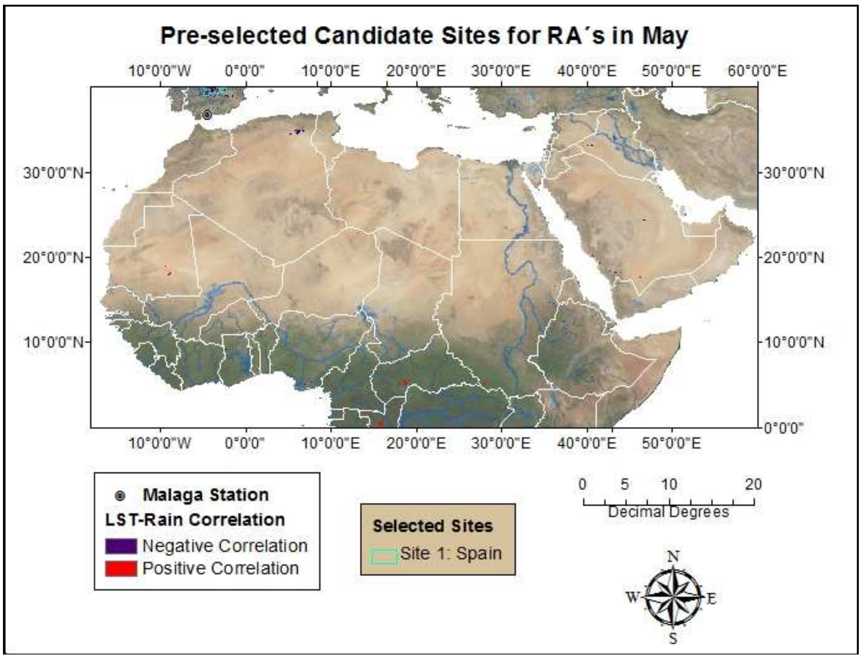

Figure 9. Pre-selected Candidate Sites for RA´s in May

The next figures (10 and 11) show the correlation maps for each data set and the wind patterns.

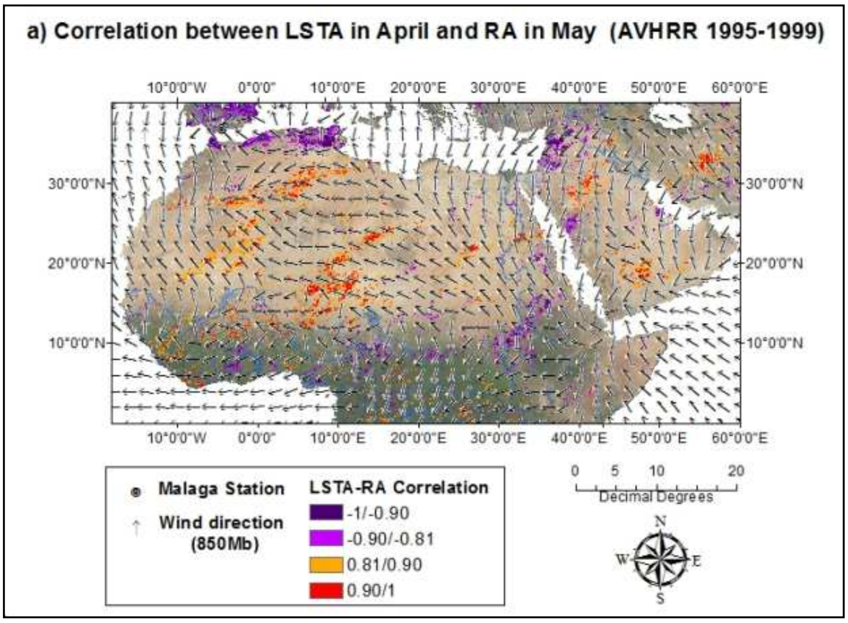

Figure 10. AVHRR. Significant Correlation RA´s in May

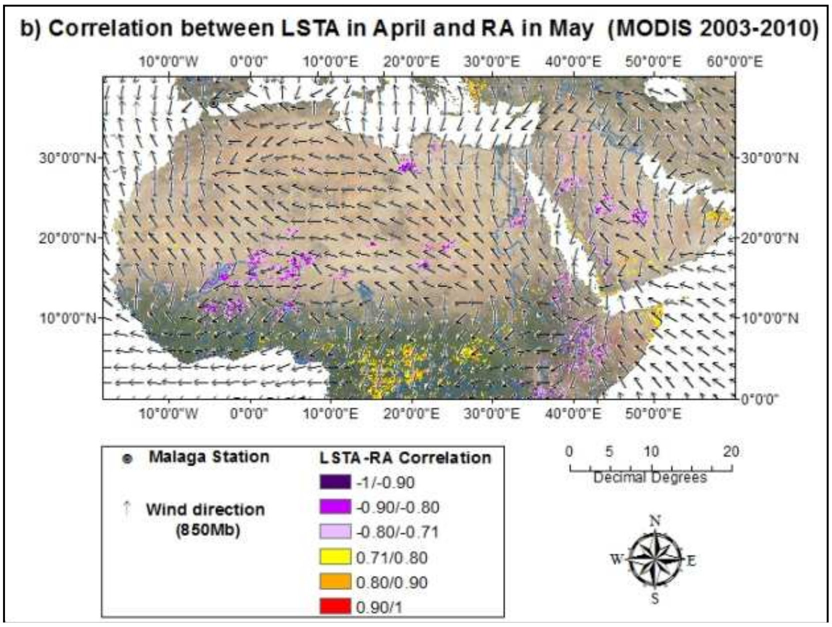

Figure 11. MODIS. Significant Correlation RA´s in May

3.2.1.2. Selection of Candidate Sites for RA´s May

The following table illustrates the values of the Pearson and Spearman Correlation Coefficients for the study period (1995-2000/2003-2010) and for every of the dataset periods (AVHRR 1995-1999 and MODIS 2003-2010). Spearman Correlation coefficient within the study period is significant. For that reason, the Null Hypothesis (H_0) from the *Hypothesis 1* can be rejected and there is a significant correlation between LSTA-April over the site in Spain and RA-May in Malaga for the study period. According to the *p-value*, the probability that these two variables are correlated is 99.99%. This result suggests that when the ranking score of LSTA-April over the site, which has been identified in Spain is above average the ranking score of RA-May over Malaga is below average and vice-versa. On the other hand, the values in the table can be used to evaluate the effectiveness of the proposed method. In this case, the use of Pearson Correlation for the pre-selection of the candidate sites allowed to obtain a significant Spearman correlation after averaging the LSTA in space and time.

Table 8. Correlation Coefficients. Candidate Site for RA´s in May. Spain

Variables	Study period	1995-1999	2003-2010
Pearson	**-0.887**	**-0.902**	**-0.795**
Spearman	**-0.807***	**-0.900**	-0.452

Values in bold are different from 0 with a significance level alpha=0.05
** The p-value for Spearman within the study period is 0.001*

The next figure shows the scatter plot of the data. A negative correlation between the LSTA's and RA's can be noticed.

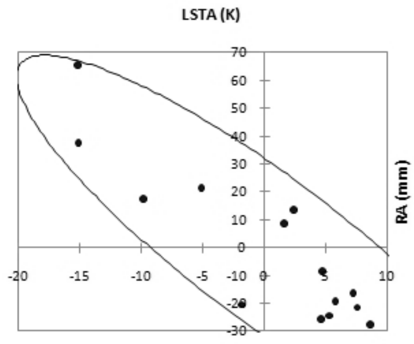

Figure 12. Scatter Plot. Candidate Sites for RA´s in May. Spain

3.2.1.3. Selected sites and relationship with ENSO

In the next map, the correlation between SSTA and Rainfall in the candidate site in Spain at a significance level of a = 0.05 is shown for Apr-May 1995-2010. There are not many significant correlated areas located on El Niño Regions (between 5^{o}N-5^{o}S and 170^{o}W-120^{o}W). According to this, then the Null Hypothesis (H_0) of the *Hypothesis 2* is rejected, which means that there is no significant correlation between rainfall in the candidate site in Spain and Sea Surface Temperature Anomalies in El Niño Regions.

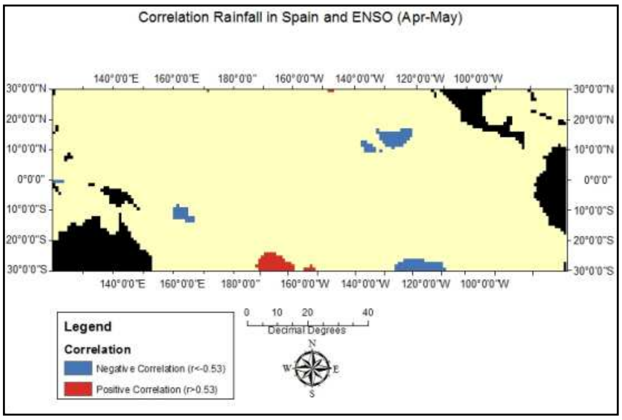

Figure 13. ENSO and Rainfall in the Candidate Site: Spain

3.2.1.4. Environmental Variables in the selected sites

The next table contains some of the environmental variables that could help to explain the causes of the relation between LSTA´s and RA´s in Malaga. These variables together, especially the average rainfall, suggests that this site could be a source of moisture and thermal advection.

Table 9. Environmental variables. Candidate Sites for RA´s in May.

Site	Area (sq km)	Average Evapotranspiration (Kg/m2/s)	Water Vapour Near Surface (kg/kg)	Water Vapour Column (cm)	Average Rainfall Aug-Sep (mm)
Site 1: Spain	16,142.71	1.230E-05	0.0063.	1.123	43.32

3.2.1.5. Forecasting skills of selected sites

Model: RA (Ranking Score) = 13.55-0.806*LSTA (Ranking Score)

Next table summarize the Analysis of Variance. The probability value (Pr>F) is bellow the significance level a = 0.05, the null hypothesis is rejected and the regression significant.

Table 10. ANOVA. Candidate Site for RA´s in May. Spain

Source	DF	Sum of squares	Mean squares	F	Pr > F
Model	1	148.010	148.010	22.344	0.000
Error	12	79.490	6.624		
Corrected Total	13	227.500			

Computed against model Y=Mean(Y)

The quality of the model is summarized in the next table. The coefficient of determination, R^2, indicates that 65.1 % of the variance in the ranking score of RA-May occurring in Malaga is explained by LSTA-April over the candidate site in Spain. The *Adjusted* R^2 is slightly slower due to the sample size. The MSE is lower than for the other candidate sites. The comparison RMSE and Press RMSE indicates that the model could not be overfitted.

Table 11. Goodness of fit. Candidate Site for RA´s in May. Spain

DF	R²	Adjusted R²	MSE	RMSE	Press RMSE	Bias
12	0.651	0.621	6.624	2.574	2.892	0

Next figure represents the regression line together with the confidence interval at a significant level a = 0.05 for the mean and the predictions.

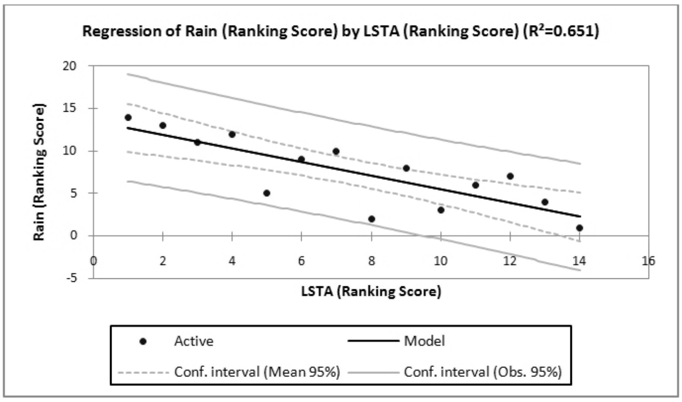

Figure 14. Model. Candidate Site in Spain

To sum up, LSTA-April over the site in Spain was negatively correlated with RA-May in Malaga giving a Spearman correlation value of -0.807 at a *p-value* of 0.001. This suggests that a negative LSTA-April over the identified site could cause a positive RA-May in Malaga and vice-versa. From the environmental variables it can be inferred that the site could be a source of moisture. In addition, the wind direction from this site is pointing to Malaga. On the other hand, this site also showed a significant linear regression and a good fit of the model. Furthermore, it does not show totally dependency from ENSO.

Romero et al. (1999) classified the atmospheric patterns that affect rainfall in Spain. They recognized 19 circulation patterns which vary with the seasons. Two of the patterns identified (AP14 and AP15) were characterized by troughs over Spain from the south-easterly flows from the Mediterranean (warm and humid) induced by low pressure over the south of Spain. AP15 was often found during spring and autumn while AP14 was found mainly in spring and summer. These patterns could be related to the high negative correlation found over the identified site in Spain, meaning that the increase of moisture over the site decreases its temperature and the moisture advection causes rainfall in Malaga.

3.2.2. Selection and evaluation of candidate sites for RA´s in June

First of all, the pre-selection of candidate sites for RA´s in June was carried out. The method proposed here showed to be suitable for this purpose allowing to select two final candidate sites: one over Burkina Faso and another one over Libya.

Secondly, the influence of ENSO over the candidate sites was analyzed. Also, some environmental variables were studied to determine the type of advection occurring over these sites.

Finally, the candidate site forecasting skills where evaluated.

3.2.2.1. Pre-selection of Candidate Sites for RA´s June

The next figure shows the map from the Pearson Product Moment Correlation between LSTA-May and RA-June once the results of the two data sets were merged. Most of the resulted areas are negatively correlated with RA-June in Malaga. They are mainly located over the Sahara desert and over the Sahel region (10ºN-20ºN longitude). The areas showing a positive correlation LSTA-RA are mainly located in the Eastern part of the study area. Due to their significant size two sites with negative correlation were chosen for further analysis: one over Libya and another over Burkina Faso. Besides Malaga is influenced by the easterlies coming from the selected sites.

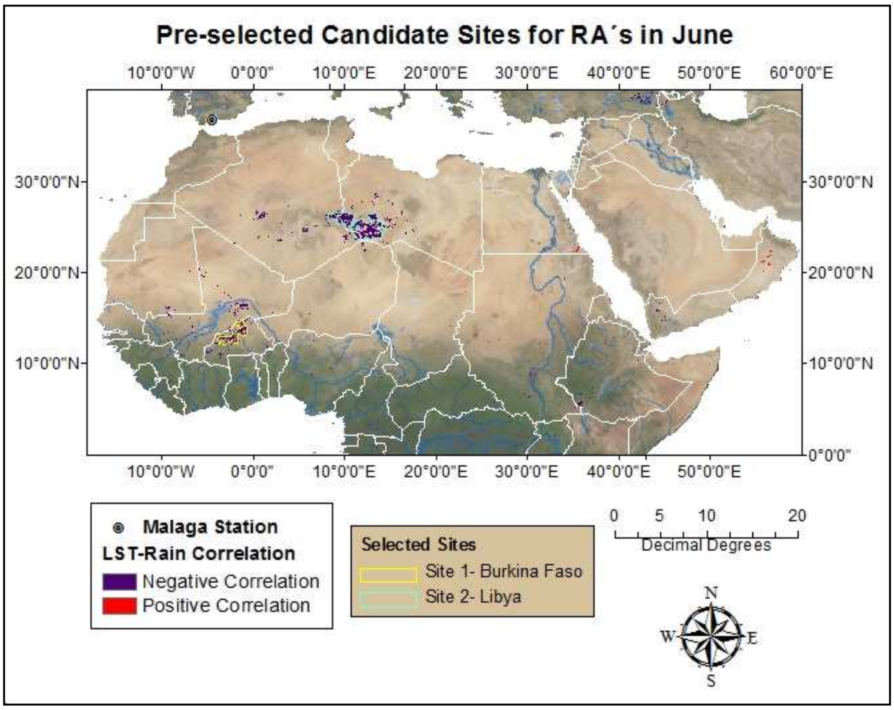

Figure 15. Pre-selected Candidate Sites for RA´s in June

The next figures (16 and 17) show the correlation maps for each data set and the wind patterns.

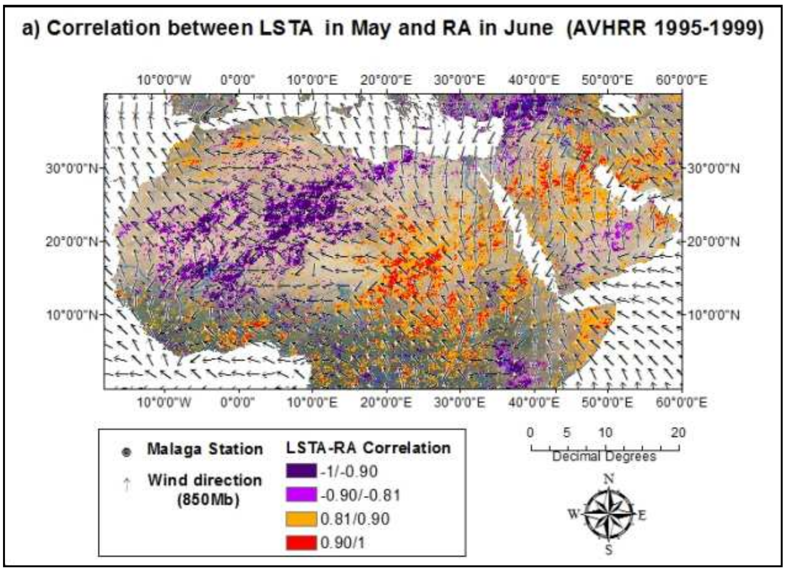

Figure 16. AVHRR. Significant Correlation RA´s in June

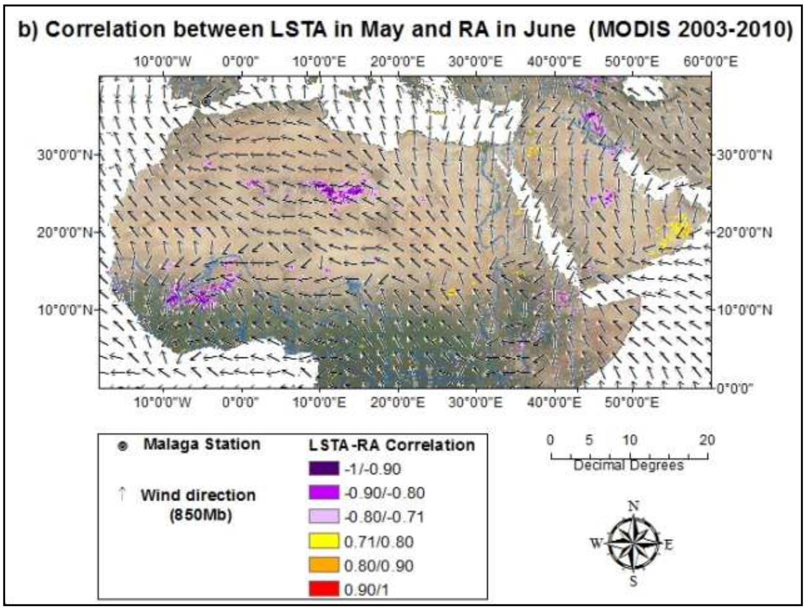

Figure 17. MODIS. Significant Correlation RA´s in June

3.2.2.2. Selection of Candidate Sites for RA´s June

a) Site 1: Burkina Faso

The following table illustrates the values of the Pearson and Spearman Correlation Coefficients as it can be seen in table 8. The Spearman Correlation coefficient within the study period is shown to be significant. The Null Hypothesis (H_0) from the *Hypothesis 1* can be rejected and there is a significant correlation between LSTA-May over the site in Burkina Faso and RA-June in Malaga for the study period. According to the *p-value,* the probability that these two variables are correlated is 97.6%. This result suggests that when the ranking score of LSTA-May over Burkina Faso is above average, bellow average ranking score of RA-June is generally observed over Malaga and vice-versa. On the other hand, the values in the table can be used to evaluate the effectiveness of the proposed method. In this case, the use of Pearson Correlation for the identification of the sites allowed to obtain a significant Spearman correlation after averaging the LSTA in space and time.

Table 12. Correlation Coefficients. Candidate Site for RA´s in June. Burkina Faso

Variables	Study period	1995-1999	2003-2010
Pearson	-0.697	**-0.784**	**-0.843**
Spearman	**-0.609***	-0.673	-0.299

Values in bold are different from 0 with a significance level alpha=0.05
** The p-value for Spearman within the study period is 0.024*

The next figure shows the scatter plot of the data. A negative correlation between the LSTA's and RA's can be noticed.

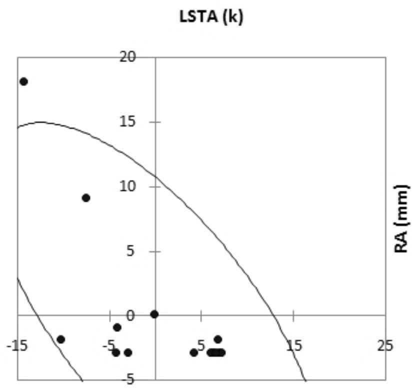

Figure 18. Scatter plot. Candidate Site for RA´s in June. Burkina Faso

b) Site 2: Libya

The following table illustrates the values of the Pearson and Spearman Correlation Coefficients as it can be seen in table 8. The Spearman Correlation coefficient within the study period is shown to be significant. The Null Hypothesis (H_0) from the *Hypothesis 1* can be rejected and there is a significant correlation between LSTA-May over the site in Libya and RA-June in Malaga for the study period. According to the *p-value*, the probability these two variables are correlated is 98.9%. This result suggests that when the ranking score of LSTA-May over the site in Libya is above average, bellow average ranking score of RA-June is generally observed over Malaga and vice-versa. On the other hand, the values in the table can be used to evaluate the effectiveness of the proposed method. In this case, the use of Pearson Correlation for the identification of the sites allowed obtaining a significant Spearman correlation after averaging the LSTA in space and time.

Table 13. Correlation Coefficients. Candidate Site for RA´s in June. Libya

Variables	Study period	1995-1999	2003-2010
Pearson	**-0.758**	**-0.811**	**-0.929**
Spearman	**-0.668**	-0.518	-0.763

Values in bold are different from 0 with a significance level alpha=0.05
** The p-value for Spearman within the study period is 0.011*

The next figure shows the scatter plot of the data. A negative correlation between the LSTA's and RA's can be noticed.

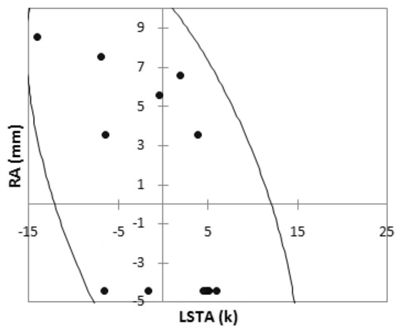

Figure 19. Scatter plot. Candidate Site for RA´s in June. Libya

3.2.2.3. Candidates Sites and relationship with ENSO

a) Site 1: Burkina Faso

In the next map, the correlation between SSTA and Rainfall in Burkina Faso at a significance level of a=0.05 is shown for May-June 1995-2010. There are some blue areas located on El Niño Regions (between 5^{o}N-5^{o}S and 170^{o}W-120^{o}W) indicating that there is a significant negative correlation between SSTA in May and June and Rainfall in Burkina Faso during the same months. This means that when SST's are above normal, during El Niño conditions, the rainfall in May and June are below the average in this site. According to this, the Null Hypothesis (H_0) of the *Hypothesis 2* is not rejected which means that there is a significant correlation between rainfall in the candidate site in Burkina Faso and Sea Surface Temperature Anomalies in El Niño Regions.

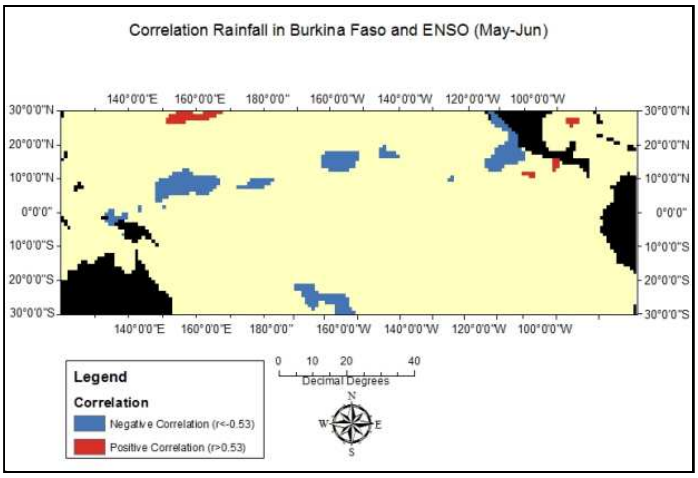

Figure 20. ENSO and Rainfall in the Candidate Site: Burkina Faso

b) Site 2: Libya

Contrary to the situation in Burkina Faso, no significant correlation can be noticed between rainfall in May and June and El Niño. Accordingly, the Null Hypothesis (H_0) of the *Hypothesis 2* is rejected, which means that there is no significant correlation between rainfall in the candidate site in Libya and Sea Surface Temperature Anomalies in El Niño Regions.

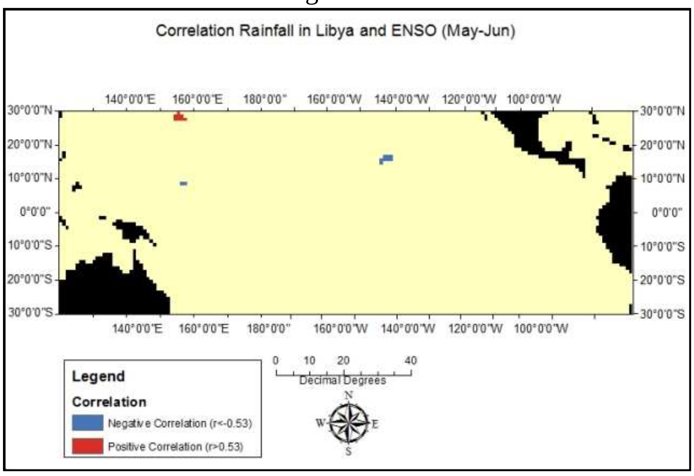

Figure 21. ENSO and Rainfall in the Candidate Site: Libya

3.2.2.4. Environmental Variables in the selected sites

The next table contains some of the environmental variables that could help to explain the causes of the relation between land-surface temperature anomalies and rainfall anomalies in Malaga. It can be noticed that the site in Burkina Faso contains higher amount of water vapour near the surface and also in the water vapour column which indicates it could be a source of moisture. Furthermore, the average rainfall for May and June is much higher than for Libya. Consequently, since there is more water to evapotranspirate, the evapotranspiration value is also higher. Burkina Faso could be a source of moisture and thermal advection while Libya could just be a source of thermal advection.

Table 14. Environmental variables. Candidate Sites for RA´s in June.

Site	Area (sq km)	Average Evapotranspiration (Kg/m2/s)	Water Vapour Near Surface (kg/kg)	Water Vapour Column (cm)	Average Rainfall May-Jun (mm)
Site 1: Burkina Faso	39,761.07	1.099E-05	0.0090	3.53	79.88
Site 2: Libya	105,801.44	1.177E-06	0.0044	1.57	2.35

3.2.2.5. Forecasting skills of candidate sites

a) Site 1: Burkina Faso

Model: RA (mm) = 11.264-0.778*LSTA (k)

Next table summarize the Analysis of Variance. The probability value (Pr>F) is bellow the significance level a = 0.05. The null hypothesis is consequently rejected and the regression significant.

Table 15. ANOVA. Candidate Site for RA´s in June. Burkina Faso

Source	DF	Sum of squares	Mean squares	F	Pr > F
Model	1	137.710	137.710	6.618	0.024
Error	12	249.719	20.810		
Corrected Total	13	387.429			

Computed against model Y=Mean(Y)

The quality of the model is summarized in the next table. The coefficient of determination, R^2, indicates that 35.5 % of the variance in the ranking score of RA-June occurring in Malaga is explained by LSTA in May over the candidate site in Burkina Faso. The *Adjusted* R^2 is slightly slower due to the sample size. The MSE is higher than for the models of the other candidate sites. The comparison RMSE and Press RMSE indicates that the model could be overffited.

Table 16. Goodness of fit. Candidate Site in Burkina Faso

DF	R^2	Adjusted R^2	MSE	RMSE	Press RMSE	Bias
12	0.355	0.302	20.81	4.562	5.212	0

Next figure represents the regression line together with the confidence interval at a significant level a = 0.05 for the mean and the predictions.

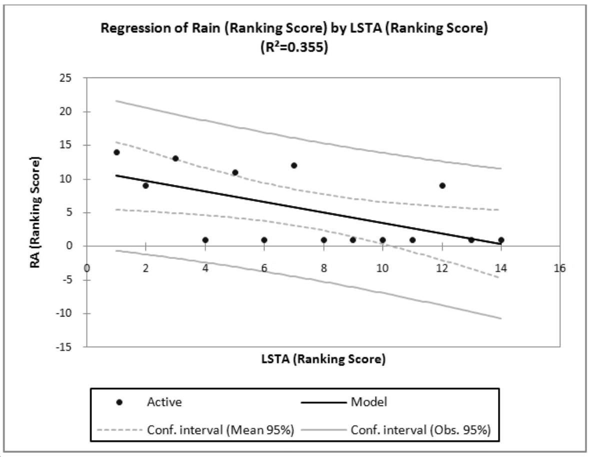

Figure 22. Model. Candidate Site in Burkina Faso

b) Site 2: Libya

Model: Rainfall Ranking Score = 11.857-0.857*LSTA Ranking Score

Next table summarize the Analysis of Variance. The probability value (Pr>F) is bellow the significance level a = 0.05. The null hypothesis is therefore rejected and the regression significant.

Table 17. ANOVA. Candidate Site for RA´s in June. Libya

Source	DF	Sum of squares	Mean squares	F	Pr > F
Model	1	167.143	167.143	9.105	0.011
Error	12	220.286	18.357		
Corrected Total	13	387.429			

Computed against model Y=Mean(Y)

The quality of the model is summarized in the next table. The coefficient of determination, R^2, indicates that 43.1 % of the variance in the ranking score of RA-June occurring in Malaga is explained by LSTA in May over the candidate site in Libya. The *Adjusted* R^2 is slightly slower due to the sample size. The MSE is lower than for the candidate site in Burkina Faso. The comparison RMSE and Press RMSE indicates that the model could be overffited.

Table 18. Goodness of fit. Candidate site in Libya

DF	R²	Adjusted R²	MSE	RMSE	Press RMSE	Bias
12	0.431	0.384	18.357	4.285	4.901	0

Next figure represents the regression line together with the confidence interval at a significant level a = 0.05 for the mean and for the predictions.

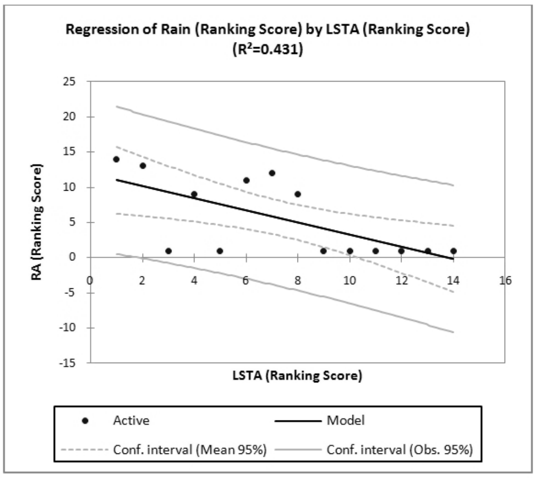

Figure 23. Model. Candidate Site in Libya

To sum up, LSTA-May over the site in Burkina Faso was negatively correlated with RA-June. Spearman correlation value was -0.609 at a *p-value* of 0.024. This suggests that a negative LSTA-April over the identified site could cause a positive RA-May in Malaga and vice-versa. From the environmental variables it can be inferred that the site could be a source of moisture. Furthermore, the wind direction from this site points to Malaga. Then, this site showed a significant linear regression and the goodness of fit of the model was not as better as for the site in Spain. However, with regards to Spain, it did not show totally independency from ENSO. This ENSO influence over Burkina Faso and the Sahel region has also been observed by Janicot et al. (2001).

On the other hand, LSTA-May over the site in Libya also showed a significant negative correlation with RA-May in Malaga with a Spearman correlation value of -0.668 at a *p-value* of 0.011. This suggests again that a negative LSTA-April over the identified site could cause a positive RA-May in Malaga and vice-versa. As in the cases before, the wind direction is also pointing to Malaga. The environmental variables suggest that this site is not a source of moisture but a source of warm advection. The significant linear regression showed slightly better results than the site in Burkina Faso. Contrarily to the other sites, the site over Libya showed significant independence from ENSO. This means that the site over Libya could improve the forecasting of rainfall in Malaga by itself.

These candidate sites identified over Africa could be related to the tropical intrusion occurring during the monsoon of western Africa. This monsoon is considered to be caused by the seasonal shifts of the Intertropical Convergence Zone (ITCZ) together with the great seasonal temperature and humidity differences between the Sahara and the equatorial Atlantic Ocean. The site over Libya is within the Sahara desert. The monsoon normally reaches western Africa in June (where the site over Burkina Faso is located) and moves to the south in October (AMMA, 2010). During the transition seasons, when the circulation patterns change, and in June and August the West Africa monsoon can reach its northernmost location crossing the Sahara desert (Andrew et al., 2009) and in some cases can reach Spain. Examples are shown in Knippertz et al. (2003) and EUMESAT (2007). Grams et al. (2010) also observed the influence of the Atlantic inflow in the Saharan heat low.

Furthermore, the site over Libya could be related to the Southern Advection over Spain. It is known that during the summer period, the dry masses of air coming from northern Africa (Southern Advection) cross the Mediterranean Sea and when charged of water vapour they produce rainfall in Southern Spain. The tropical continental air mass from Africa can also meet a maritime polar air mass in the western part of Spain an produce red rain (Saharan dust) in the South and East of Spain (University of Valencia, 2010). As cited in Rodriguez et al. (2001) the injection of particles to high atmospheric levels is due to the thermal convective activity over the Sahara desert (Carlson and Prospero, 1972; Prospero and Carlson, 1972; Westphal et al., 1988). The Saharan dust reaches the Iberian Peninsula when the North Atlantic anticyclone (Azores high) is displaced westward and the North African high is centred over Algeria which is beside Libya. This phenomenon mainly happens during summer. The summer dust events may be caused also by South-western depressions or by the introduction of Atlantic air masses.

3.2.3. Selection and evaluation of candidate sites for RA´s in August

The pre-selection of candidate sites for RA´s in August was carried out. The method proposed here showed to be not suitable for this purpose. Thus, no final candidate sites were selected.

3.2.3.1. Pre-selection of Candidate Sites for RA´s August

The next figure shows the map from the Pearson Product Moment Correlation between LSTA-July and RA-August once the results of the two data sets were merged. The resulting common areas sharing the same sign of the correlation are shown. Most of the resulted areas are negatively correlated with RA-July in Malaga. They are mainly located over the Arabian Peninsula. Due to its significant size a site over Saudi Arabia which is negatively correlated with RA-July in Malaga was chosen. The wind pattern illustrated in figures 25 and 26 does not show direct wind direction pointing to Malaga, though indirectly it does.

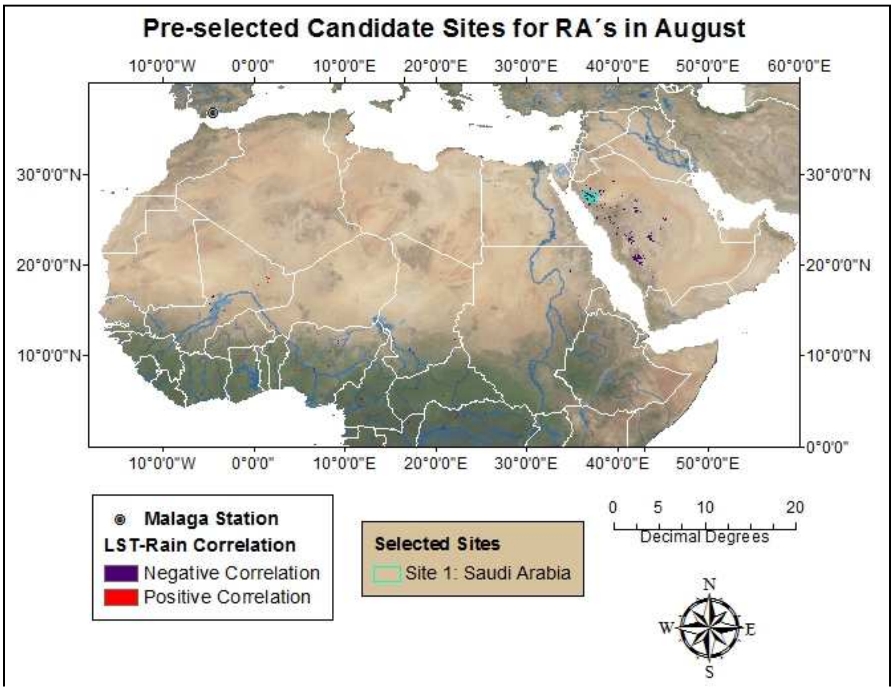

Figure 24. Pre-selected Candidate Sites for RA´s in August

The next figures (25 and 26) show the correlation maps for each data set and the wind patterns.

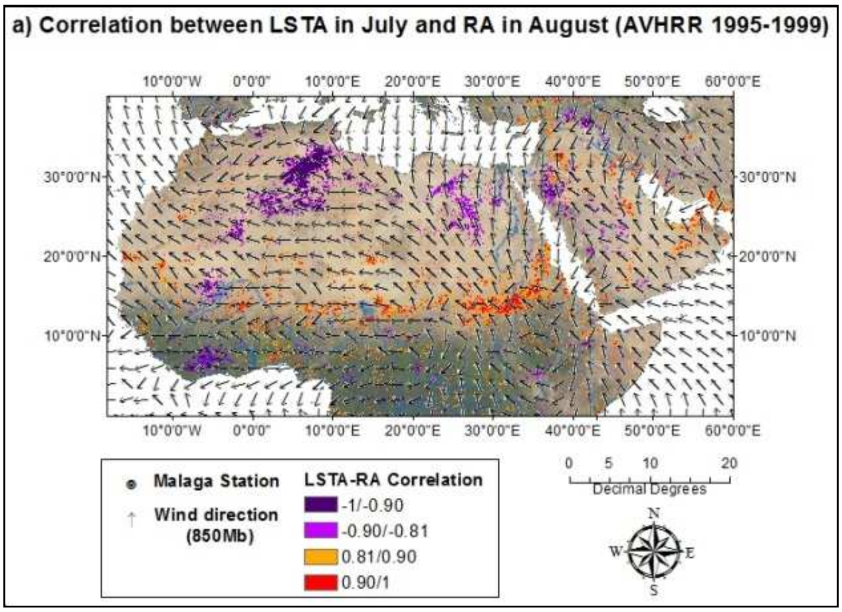

Figure 25. AVHRR. Significant Correlation RA´s in August

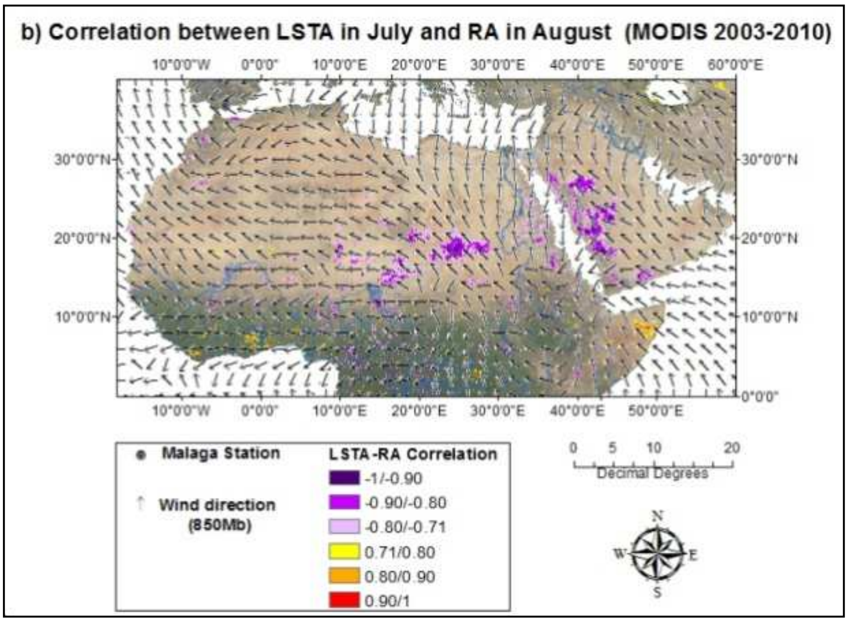

Figure 26. MODIS. Significant Correlation RA´s in August

3.2.3.2. Selection of Candidate Sites for RA´s August

The following table illustrates the values of Pearson and Spearman Correlation Coefficients as it can be seen in table 8. Spearman Correlation coefficient within the study period is significant. Then, the Null Hypothesis (H_0) from the *Hypothesis 1* cannot be rejected and there is no significant correlation for the study period between LSTA-July over the site in Saudi Arabia and RA-August in Malaga. According to the *p-value*, the probability that these two variables are not correlated is 53.5%. On the other hand, the values in the table can be used to evaluate the effectiveness of the proposed method. In this case, the use of Pearson Correlation for the identification of the sites did not allow to obtain a significant Spearman correlation after averaging LSTA in space and time. Furthermore, the Pearson correlation coefficient obtained for the whole study period also drops significantly.

Table 19. Correlation Coefficients. Candidate Site for RA´s in August. Saudi Arabia

Variables	Study period	1995-1999	2003-2010
Pearson	-0.133	**-0.890**	**-0.862**
Spearman	-0.213	-0.162	-0777

Values in bold are different from 0 with a significance level alpha=0.05
** The p-value for Spearman within the study period is 0.465*

The next figure shows the scatter plot of the data. The absence of correlation between the LSTA's and RA's can be noticed.

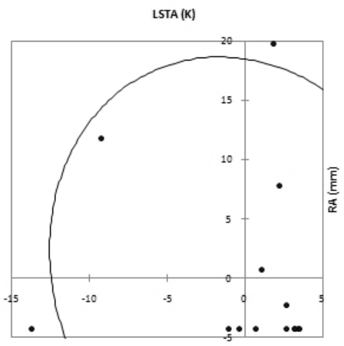

Figure 27. Scatter Plot. Candidate Site: Saudi Arabia

The decrease of the Pearson Correlation coefficient occurs when merging the two data sets. It is due to the variation of the mean and the standard deviation values when considering all the data under the study period, especially LST values (Table 20). There is a difference of 5ºK between the data 1995-1999 and 2003-2010. Furthermore, the Standard Deviation varies also in almost four units.

Table 20. Candidate Site in Saudi Arabia. Statistics

Time Period	mean R	Std. Dev R	mean LST	Std. Dev. LST
1995-2000/2003-2010	4.214	7.627	322.180	5.137
1995-1999	4.200	6.942	320.400	4.263
2003-2010	4.750	8.812	325.004	0.608

The change in the mean causes differences in the anomaly values. Figure 28 shows the LSTA-August and RA-September for Saudi Arabia for the entire period. Figure 29 and Figure 30 are similar to Figure 28 but for the time period of every dataset. The main difference is found between LSTA for the entire study period and LSTA for the period 2003-2010. In the entire period all LSTA are positive while in the period 2003-2010 there are some negative values. On the other hand, from figure 28 it can be inferred that the period 1995-1999 was cooler that 2003-2010 in the site over Saudi Arabia, especially in 1996 and 2000.

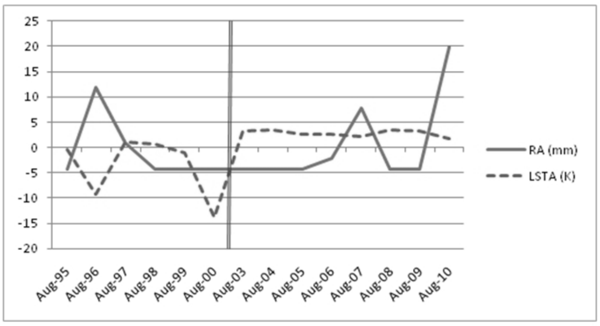

Figure 28. LSTA´s in Saudi Arabia and RA´s in Malaga

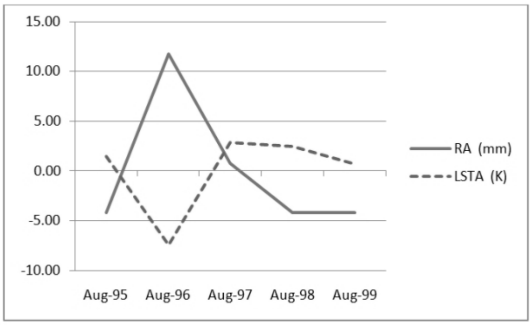

Figure 29. LSTA´s in Saudi Arabia and RA´s in Malaga (1995-1999)

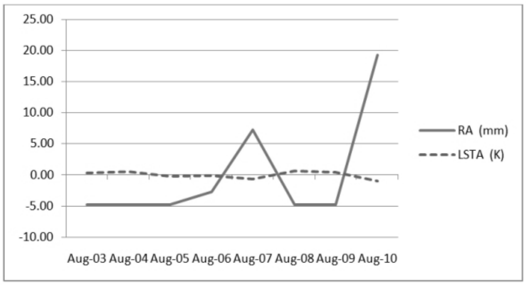

Figure 30. LSTA´s in Saudi Arabia and RA´s in Malaga (2003-2010)

3.2.4. Selection and evaluation of candidate sites for RA´s in September

The pre-selection of candidate sites for RA´s in August was carried out. The method proposed here showed to be not suitable for this purpose. Thus, no final candidate sites were selected.

3.2.4.1. Pre-selection of Candidate Sites for RA´s September

The next figure shows the map from the Pearson Product Moment Correlation between LSTA-August and RA-September once the results of the two data sets were merged. Some of the resulted areas located over the Arabian Peninsula are negatively correlated with RA-September in Malaga. Some others located over Sahel region (10ºN-20ºN longitude) are positively correlated. Due to their significant size the area positively correlated over Mali was selected for further analysis. Furthermore, Malaga is influenced by the easterlies coming from the selected site.

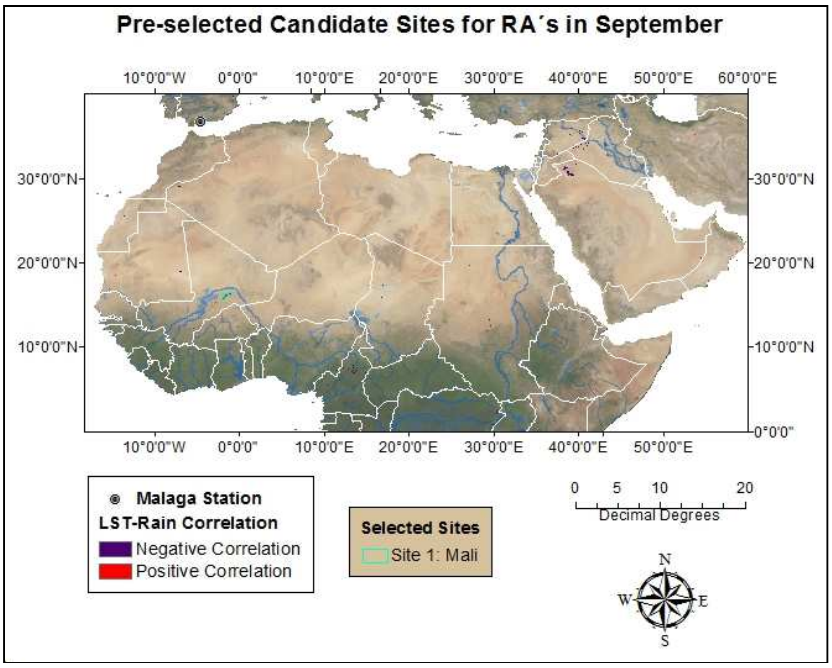

Figure 31. Pre-selected Candidate Sites for RA´s in September

The next figures (32 and 33) show the correlation maps for each data set and the wind patterns.

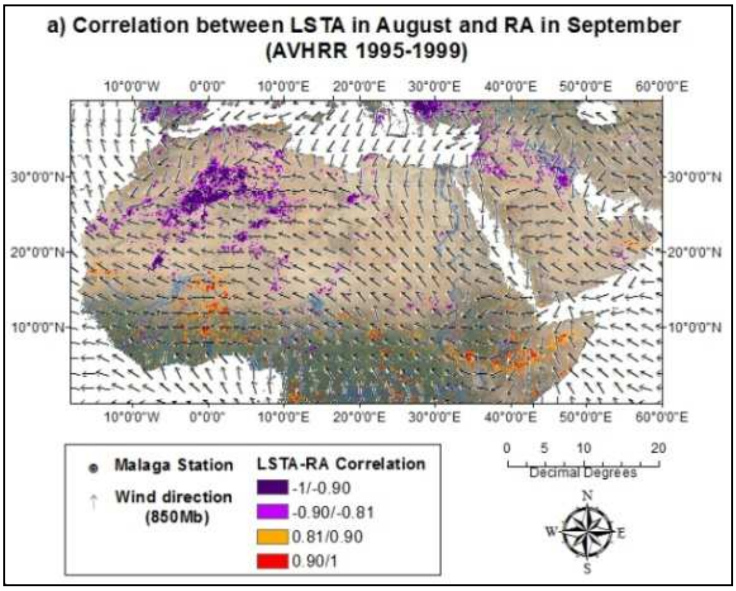

Figure 32. AVHRR. Significant Correlation RA´s in September

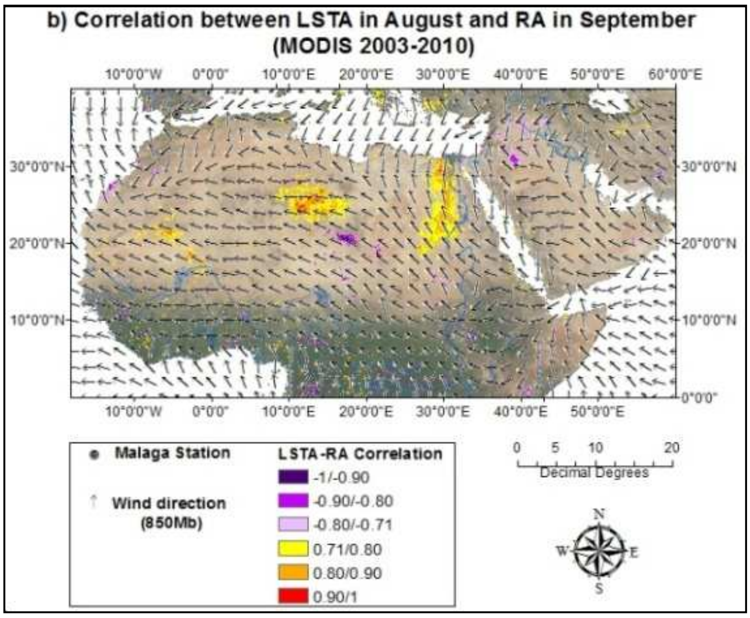

Figure 33. MODIS. Significant Correlation RA´s in September

3.2.4.2. Selection of Candidate Sites for RA´s September

The following table illustrates the values of the Pearson and Spearman Correlation Coefficients as it can be seen in table 8. The Spearman Correlation coefficient within the study period is shown to be significant. Then, the Null Hypothesis (H_0) from the *Hypothesis 1* cannot be rejected and it can be said that there is no significant correlation between LSTA in May over the site in Saudi Arabia and RA in June in Malaga for the study period. According to the *p-value*, the probability these two variables are not correlated is 96.9%. On the other hand, the values in the table can be used to evaluate the effectiveness of the proposed method. In this case, the use of Pearson Correlation for the identification of the sites did not allow obtaining a significant Spearman correlation after averaging the LSTA by space and time. Furthermore, the Pearson correlation coefficient obtained for the whole study period also drops significantly.

Table 21. Correlation Coefficients. Candidate Site for RA´s in September. Mali

Variables	Study period	1995-1999	2003-2010
Pearson	0.206	**0.946**	**0.784**
Spearman	0.011	0.7000	0.256

Values in bold are different from 0 with a significance level alpha=0.05
** The p-value for Spearman within the study period is 0.969*

The next figure shows the scatter plot of the data. The absence of correlation between the LSTA's and RA's can be noticed.

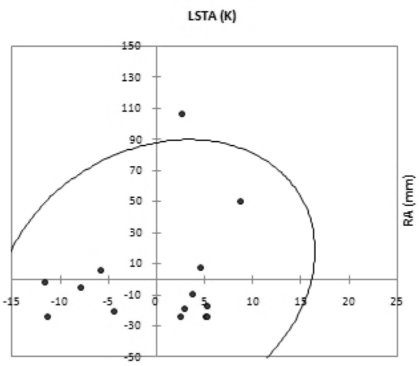

Figure 34. Scatter Plot. Candidate Site: Mali

As in RA-August, the decrease of the Pearson Correlation coefficient occurs when merging the two data sets. It is due to the variation of the mean and in the standard deviation values when considering all the data under the study period. In this case both rainfall and LST show very different mean and standard deviation values (Table 22).

Table 22. Candidate Site in Mali. Statistics

Time Period	mean R	Std. Dev R	mean LST	Std. Dev. LST
1995-2000/2003-2010	24.214	36.703	307.438	6.725
1995-1999	40.400	51.830	300.700	5.775
2003-2010	16.625	25.601	312.203	1.939

The change in the mean causes differences in the anomaly values. This is illustrated in the next figures. Figure 35 shows the LSTA-August and RA-September for Mali for the entire period. Figure 36 and Figure 37 are similar to Figure 35 but for the time period of every dataset. The main difference occurs between the LSTA for the entire study period and LSTA for the period 2003-2010. In the entire period all the LSTA are positive while in the period 2003-2010 there are some negative values. On the other hand, from figure 35 it can be inferred that the period 1995-1999 was cooler that 2003-2010 in the site over Mali, especially in 1995 and 1999.

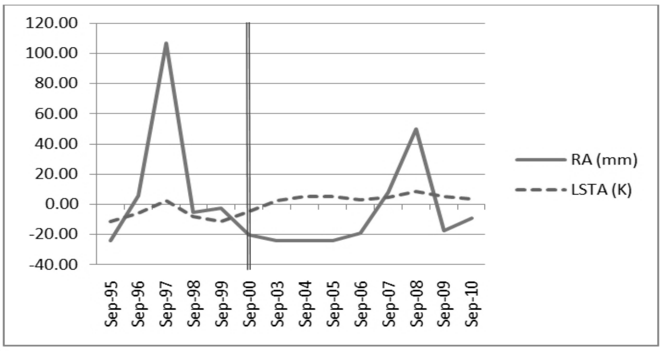

Figure 35. LSTA´s in Mali and RA´s in Malaga

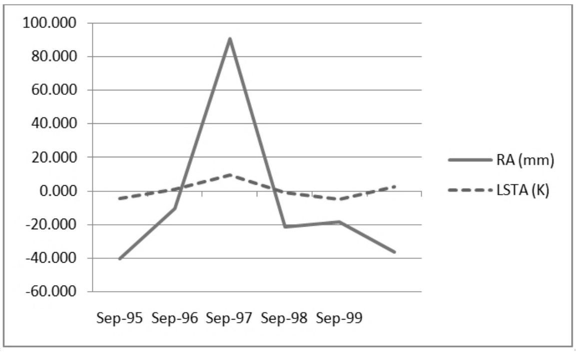

Figure 36. LSTA´s in Mali and RA´s in Malaga (1995-1999)

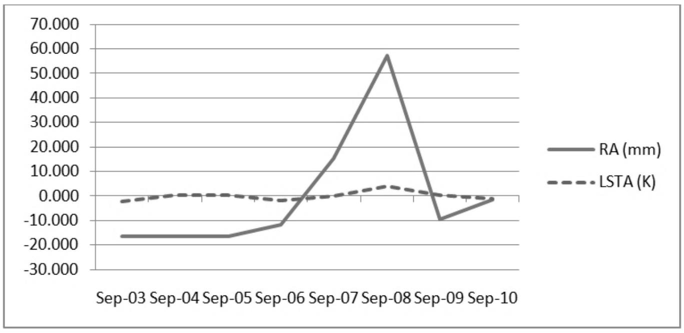

Figure 37. LSTA´s in Mali and RA´s in Malaga (2003-2010)

3.3. Sources of error and uncertainities

The first uncertainty in this research comes from the assumptions in which it is based: first, the simplification of all the terms of the Energy Balance Equation; second, the treatment of radiative temperature (LST) as thermodynamic temperature, even over heterogeneous areas, and in areas under no thermal equilibrium. It is known the fact that radiative temperature from satellite images tends to overestimate the aerodynamic temperature value (Friedl, 2002). However, since this research does not pretend to obtain numerical results as numerical models do, but to give an approximation of the real world phenomena, this fact may not be an important issue.

The decrease in the resolution of the data in order to carry out the correlations to select the areas may introduce a source of error was well as AVHRR and MODIS data may give way to error especially over deserted areas (Wan, 2010). Another consideration is that the radiative temperature from the images is the one obtained at the passing time of the satellite (noon) so it is just again a simplification of the reality. The rainfall data comes from the meteorological station located in Malaga airport. All the data series were completed. Although the data were obtained from an official source, no information about the quality was provided.

4. Conclusion and Recommendation

The main objective of this research was to identify predictor sites of rainfall in Malaga within the study area. Since moisture and thermal advection are in most of the cases highly related to precipitation events and being advection dependent on the temperature gradient and wind vectors, the variables taken into account were land-surface (skin) temperature and wind patterns. All the statistics tests were taken at a significance level a = 0.05 and for a degree of freedom of 12. The following conclusions have been obtained:

With regards to the first objective: To study if there is a significant relationship between LSTA´s over candidate sites and RA´s in Malaga, it has been found that there exits some specific sites correlated with RA´s in Malaga. The following sites were chosen for their significant: one site over Spain for the rain in May (rh_o= -0.807), one site over Burkina Faso (rh_o= -0.609) and another over Libya (rh_o= -0.668) for the rain in June.

With regards to the second objective: To study if the candidate sites are not influenced by El Niño-Southern Oscillation (ENSO) and can be used as a independent predictor of rainfall in Malaga, it has been found that, among the selected sites, only the site over Burkina Faso was totally influenced by ENSO.

With regards to the third objective: To evaluate the forecasting skills of the candidate sites by building and validating a bivariate model LSTA-RA, although all the model were significant, the best model was the one for the site in Spain (R^2= 0.651; RMSE = 2.574) followed by the site in Libya (R^2 = 0.431; RMSE = 4.285) and the site in Burkina Faso (R^2 = 0.355; RMSE = 4.562).

With regards to the forth objective: To find out if the candidate sites are sources of moisture and/or thermal advection by analysing other environmental variables (monthly evapotranspiration, monthly moisture over the surface, monthly atmospheric moisture column and monthly rainfall), the sites in Spain and Burkina Faso could be a source of moisture and thermal advection and the site in Libya could only be a source of thermal advection.

With regards to the final objective: To find out if the time lag between LSTA´s over the candidate sites and RA´s in Malaga is sufficient to improve water management in Andalusia, as a significant time lag was not found after the statistical analysis of the data, to meet this objective it was chosen the maximum time lag required by the users.

The recommendations for further studies are the following:

- To enlarge the number of years to be studied
- To enlarge the study area
- To include more meteorological stations
- To analyze other time lags
- To include more variable in the models
- To improve the thermodynamic temperature estimation by using NDVI or LAI
- To carry out a more detailed analysis of the environmental variables that affect advection.

5. References

AHMED, B. Y. M. 1997. Climatic classification of Saudi Arabia: an application of factor-cluster analysis. *GeoJournal, 41.1, 69–84, 1997.*

ALLEN, D. M. 1971. Mean square error of prediction as a criterion for a selection variables. *Technometrics,* 13**,** 469-475.

AMMA. 2010. *The African Monsoon* [Online]. Available: http:/www.amma-international.org/article.php3?id_article=10 [Accessed 2010, November 14].

ANDREW, H., JEAN, P. & TOM, H. 2009. Chapter 3: The climate system. *In:* WOODWARD, J. (ed.) *The Physical Geography of the Mediterranean.* Oxford University Press.

BASTIAANSSEN, W. G. M. 1995. *Regionalization of surface flux densities and moisture indicators in composite terrain : a remote sensing approach under clear skies in Mediterranean climates* Report 109, Agricultural Research Department, Wageningen, The Netherlands.

BRIDGMAN, H. A. & OLIVER, J. E. (eds.) 2006. *The Global Climate System. Patterns, Process and Teleconnections*: Cambridge University Press.

BROOMHEAD, D. S. & KING, G. P. 1986. Extracting qualitative dynamics from experimental data. *Physica D20 : 217-236.*

CASELLES, V., COLL, C. & VALOR, E. 1997. Land surface emissivity and temperature determination in the whole HAPEX-Sahel area from AVHRR data. *International Journal of Remote Sensing,* 18**,** 1009-1027.

CHEHBOUNI, A., LO SEEN, D., NJOKU, E. G. & MONTENY, B. M. 1996. Examination of the difference between radiative and aerodynamic surface temperatures over sparsely vegetated surfaces. *Remote Sensing of Environment,* 58**,** 177-186.

COLL, C., WAN, Z. & GALVE, J. M. 2009. Temperature-based and radiance-based validations of the V5 MODIS land surface temperature product. *J. Geophys. Res.,* 114**,** D20102.

COLUMBIA UNIVERSITY. 2010. *IRI/LDEO Climate Data Library* [Online]. The International Research Institute for Climate and Society (IRI). Earth Institute. Columbia University. Available: http://iridl.ldeo.columbia.edu/index.html [Accessed 2010, August 1].

COMPO, G. P., WHITAKER, J. S. & SARDESHMUKH, P. D. 2006. Feasibility of a 100-year reanalysis using only surface pressure data. *Bulletin of the American Meteorological Society,* 87**,** 175-190.

EUMESAT. 2007. *West African Monsoon crosses the Sahara desert (August 2007)* [Online]. Available: http://oiswww.eumetsat.org/WEBOPS/iotm/iotm/20070808_monsoon/20070808_monsoon.html [Accessed 2010, December 1].

FRÍAS, M. D., HERRERA, S., COFIÑO, A. S. & GUTIÉRREZ, J. M. 2010. Assessing the Skill of Precipitation and Temperature Seasonal Forecasts in Spain: Windows of Opportunity Related to ENSO Events. *Journal of Climate,* 23**,** 209-220.

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

GIORGI, F., CHRISTENSEN, J., HULME, M., STORCH, H. V., WHETTON, P., JONES, R., MEARNS, L., FU, C., ARRITT, R., BATES, B., BENESTAD, R., BOER, G., BUISHAND, A., CASTRO, M., CHEN, D., CRAMER, W., CRANE, R., CROSSLY, J., DEHN, M., DETHLOFF, K., DIPPNER, J., EMORI, S., FRANCISCO, R., FYFE, J., GERSTENGARBE, F., GUTOWSKI, W., GYALISTRAS, D., HANSSEN-BAUER, I., HANTEL, M., HASSELL, D., HEIMANN, D., JACK, C., JACOBEIT, J., KATO, H., KATZ, R., KAUKER, F., KNUTSON, T., LAL, M., LANDSEA, C., LAPRISE, R., LEUNG, L., LYNCH, A., MAY, W., MCGREGOR, J., MILLER, N., MURPHY, J., RIBALAYGUA, J., RINKE, A., RUMMUKAINEN, M., SEMAZZI, F., WALSH, K., WERNER, P., WIDMANN, M., WILBY, R. & WILD, M., XUE, Y., 2001. Regional Climate Information- Evaluation and Projections, Climate Change 2001: The Scientific Basis. *Contribution of Working Group to the Third Assessment Report of the Intergouvernmental Panel on Climate Change. [Houghton, J.T. et al. (eds)]. Cambridge University Press: Cambridge. 583-638*

GRAMS, C. M., JONES, S. C., MARSHAM, J. H., PARKER, D. J., HAYWOOD, J. M. & HEUVELINE, V. 2010. The Atlantic inflow to the Saharan heat low: observations and modelling. *Quarterly Journal of the Royal Meteorological Society,* 136**,** 125-140.

HARDING, A. E., PALUTIKOF, J. & HOLT, T. 2009. The climatic system. *In:* WOODWARD, J. C. (ed.) *The physical geography of the Mediterranean.* Oxford University Press, Oxford, 69-88.

HUBAND, N. D. S. & MONTEITH, J. L. 1986. Radiative surface temperature and energy balance of a wheat canopy. *Boundary-Layer Meteorology,* 36**,** 1-17.

HURRELL, J. W. 1995. Decadal trends in the North Atlantic Oscillation: Regional temperatures and precipitation. *Science*, 269, 676-679.

IRI/LDEO CLIMATIC DATA LIBRARY. 2010. *Ingrid Function Documentation* [Online]. Available: http://iridl.ldeo.columbia.edu/dochelp/Documentation/ [Accessed 2010, September 1].

JANICOT, S., TRZASKA, S. & POCCARD, I. 2001. Summer Sahel-ENSO teleconnection and decadal time scale SST variations. *Climate Dynamics,* 18**,** 303-320.

KATZ, R. 1988. Use of cross-correlations in the search for teleconnections. *International Journal of Climatology,* 8**,** 241-253.

KAUFMAN, L. & ROUSSEUW, P. J. 1990. Finding Groups in Data: An Introduction to Cluster Analysis. *Wiley, New York, 1990.*

KNIPPERTZ, P., FINK, A. H., REINER, A. & SPETH, P. 2003. Three Late Summer/Early Autumn Cases of Tropical–Extratropical Interactions Causing Precipitation in Northwest Africa. *Monthly Weather Review,* 131**,** 116-135.

KUSTAS, W., ANDERSON, M., NORMAN, J. & LI, F. 2007. Utility of Radiometric–aerodynamic Temperature Relations for Heat Flux Estimation. *Boundary-Layer Meteorology,* 122**,** 167-187.

LEGATES, D. R. 2000. An introduction. *The Professional Geographers,* 52**,** 233-234.

LYNDON STATE COLLEGE ATMOSPHERIC SCIENCES. 2010. *Thermal Advection* [Online]. Available: http://apollo.lsc.vsc.edu/~wintelsw/MET2110/notes/lesson13.baroclinic/ [Accessed 2010, November 2].

MCGILL UNIVERSITY. 2003. *Thermal Advection.* [Online]. Montreal. Available: http://www.meteo.mcgill.ca/wxlab/ATOC-546/notes/lesson09.thermal_advection/ [Accessed 2010, December 3].

MONTGOMERY, D. C. & PECK, E. A. (eds.) 1992. *Introduction to linear regression analysis (2nd ed.)*: New York: John Wiley and Sons, Inc.

MUÑOZ-DÍAZ, D. 2004b. Impacts of the North Atlantic Oscillation on the probability of dry and wet winters in Spain. *Climate Research,* 27**,** 33.

MUÑOZ-DIAZ, D. & RODRIGO, F. S. 2004a. Spatio-temporal patterns of seasonal rainfall in Spain (1912–2000) using cluster and principal component analysis: comparison. *Annales Geophysicae,* 22**,** 1435-1448.

MUÑOZ-DÍAZ, D. & RODRIGO, F. S. 2003. Effects of the North Atlantic oscillation on the probability for climatic categories of local monthly

rainfall in southern Spain. *International Journal of Climatology,* 23**,** 381-397.
MUÑOZ-DÍAZ, D. & RODRIGO, F. S. 2005. Influence of the El Niño-Southern Oscillation on the probability of dry and wet seasons in Spain. *Climate Research,* 30**,** 1-12.
MUÑOZ-DÍAZ, D. & RODRIGO, F. S. 2006. Seasonal rainfall variations in Spain (1912-2000) and their links to atmospheric circulation. *Atmospheric Research,* 81**,** 94-110.
NASA. 2010a. *GLDAS_NOAH025_M. PRODUCT DESCRIPTION AND DOCUMENTATION* [Online]. Available: http://mirador.gsfc.nasa.gov/collections/GLDAS_NOAH025_M__001.shtml [Accessed 2010, January 1].
NASA. 2010b. *MODIS Atmosphere* [Online]. Available: http://modis-atmos.gsfc.nasa.gov/ [Accessed 2010, August 12].
NOAA. 2010. *ENSO. National Oceanic and Atmospheric Administration's. National Weather Services.* [Online]. Available: http://www.cpc.ncep.noaa.gov/products/analysis_monitoring/lanina/index.html [Accessed 21 August 2010].
NORMAN, J. M. & BECKER, F. 1995. Terminology in thermal infrared remote sensing of natural surfaces. *Agricultural and Forest Meteorology,* 77**,** 153-166.
OLESON, K. W., NIU, G. Y., YANG, Z. L., LAWRENCE, D. M., THORNTON, P. E., LAWRENCE, P. J., STÖCKLI, R., DICKINSON, R. E., BONAN, G. B., LEVIS, S., DAI, A. & QIAN, T. 2008. Improvements to the Community Land Model and their impact on the hydrological cycle. *Journal of Geophysical Research,* 113 G01021.
PARODI, G. N. 2000. AVHRR Hydrological Analysis System (AHAS). ITC Water Resources Division. Enschede, The Netherlands.
PEARSON, K. 1896. Mathematical contributions to the theory of evolution — III. Regression, heredity and panmixia. *Philosophical Transactions of the Royal Society of London***,** Series A 187, 253–318.
PEARSON, K. 1900. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can reasonably be supposed to have arisen in a random sampling. *Philosophical Magazine***,** 5, 157–175.
PHILLIPS, I. D. & MCGREGOR, G. R. 2001. The relationship between synoptic scale airflow direction and daily rainfall: a methodology applied to Devon and Cornwall, South West England. *Theoretical and Applied Climatology,* 69**,** 179-198.

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 Sensing of Environment,* 103**,** 153-164.

QIN, Z. & KARNIELI, A. 1999. Progress in the remote sensing of land surface temperature and ground emissivity using NOAA-AVHRR data. *International Journal of Remote Sensing,* 20**,** 2367 - 2393.

REYNOLDS, R. W., RAYNER, N. A., SMITH, T. M., STOKES, D. C. & WANG, W. 2002. An improved in situ and satellite SST analysis for climate. *J. Climate,* 15**,** 1609-1625.

RODO, X., BAERT, E. & COMIN, F. A. 1997. Variations in seasonal rainfall in southern Europe during the present century: Relationships with the North Atlantic Oscillation and the El Nino Southern Oscillation. *Climate Dynamics,* 13**,** 275-284.

RODRÍGUEZ, S., QUEROL, X., ALASTUEY, A., KALLOS, G. & KAKALIAGOU, O. 2001. Saharan dust contributions to PM10 and TSP levels in Southern and Eastern Spain. *Atmospheric Environment,* 35**,** 2433-2447.

ROGERS, J. C. & VAN LOON, H. 1979. The Seesaw in Winter Temperatures between Greenland and Northern Europe. Part II: Some Oceanic and Atmospheric Effects in Middle and High Latitudes. *Monthly Weather Review,* 107**,** 509-519.

ROHLING, E. J. & HILGEN, F. J. 1991. The eastern Mediterranean climate at times of sapropel formation: a review. *Geologie en Mijnbouw* 70**,** 253-264.

ROMERO, R., SUMNER, G., RAMIS, C. & GENOVÉS, A. 1999. A classification of the atmospheric circulation patterns producing significant daily rainfall in the Spanish Mediterranean area. *International Journal of Climatology,* 19**,** 765-785.

SALISBURY, C., ZHANG, Y. & WAN, Z. 2002. MODIS Land Surface Temperature *In:* ICESS (ed.). University of California, Santa Barbara.

SHESKIN, D. J. 2000. *Handbook of parametric and nonparametric statistical procedures,* Boca Raton, London, New York, Washington, D.C., Chapman & Hall/CRC Press

SNEDECOR GW & COCHRAN W.G. 1980. *Statistical Methods,* Iowa State University Press, Ames, Iowa

SPEARMAN, C. 1904. The proof and measurement of association between two things. *The American Journal of Psychology,* 15**,** 72-101.

SUN, J. & MAHRT, L. 1995. Determination of surface fluxes from the surface radiative temperature. *Journal of Atmospheric Sciences,* 52 1096–1106

THOMSON, D. J. 1982. Spectrum estimation and harmonic analysis. *Proc IEEE 70, 1055-1096.*

ULIVIERI, C., CASTRONUOVO, M. M., FRANCIONI, R. & CARDILLO, A. 1994. A split window algorithm for estimating land surface temperature from satellites. *Advances in Space Research,* 14**,** 59-65.

UNIVERSITY OF ILLINOIS. 2010. *Moisture Advection along the 850 mb surface* [Online]. Available: http://ww2010.atmos.uiuc.edu/(Gh)/guides/mtr/af/adv/madv.rxml [Accessed 2010, December 1].

UNIVERSITY OF VALENCIA. 2010. MASAS DE AIRE, SITUACIONES ATMOSFÉRICAS Y TIPOS DE TIEMPO EN ESPAÑA. Available: http://www.valencia.edu/portasec/materiales_geo/geografia_fisica/1/curs/cont/nf3/pdf/c3.pdf [Accessed 2010, November 3].

VAN LOON, H. & ROGERS, J. C. 1978. The Seesaw in Winter Temperatures between Greenland and Northern Europe. Part I: General Description. *Monthly Weather Review,* 106**,** 296-310.

WALKER, G. T. & BLISS, E. W. 1932. World Weather V. *Memoirs of the Royal Meteorological Society,* 4**,** 53-84.

WAN, Z. 2008. New refinements and validation of the MODIS Land-Surface Temperature/Emissivity products. *Remote Sensing of Environment,* 112**,** 59-74.

WAN, Z. 2010. *Validation of the V5 MODIS Land-Surface Temperature Product Worldwide* [Online]. NASA. Available: http://modis.gsfc.nasa.gov/sci_team/meetings/201001/presentations/land/wan.pdf [Accessed 2010, October 22].

WAN, Z. & DOZIER, J. 1996. A generalized split-window algorithm for retrieving land-surface temperature from space. *Geoscience and Remote Sensing, IEEE Transactions,* 34**,** 892-905.

WAN, Z. & ZHAO-LIANG, L. 1997. A physics-based algorithm for retrieving land-surface emissivity and temperature from EOS/MODIS data. *Geoscience and Remote Sensing, IEEE Transactions,* 35**,** 980-996.

ZHONG, L., MA, Y., SU, Z. & SALAMA, M. 2010. Estimation of land surface temperature over the Tibetan Plateau using AVHRR and MODIS data. *Advances in Atmospheric Sciences,* 27**,** 1110-1118.

6. Apendices

Apendix 1. Normality Tests

Rainfall in May

Rainfall in May. Summary statistics:

Variable	N	Missing data	Obs. without missing data	Min.	Maxi.	Mean	Std deviation
Rainfall MAY	14	0	14	0.00	93.00	27.500	28.194

Shapiro-Wilk test (Rainfall MAY):	
W	0.865
p-value	0.035
Alpha	0.05

Test interpretation:

H0: The sample follows a Normal distribution.

Ha: The sample does not follow a Normal distribution.
As the computed *p-value* is lower than the significance level alpha=0.05, one should reject the null hypothesis H0, and accept the alternative hypothesis Ha.

The risk to reject the null hypothesis H0 while it is true is lower than 3.54%.

Anderson-Darling test (Rainfall MAY):	
A^2	0.731
p-value	0.043
Alpha	0.05

Test interpretation:

H0: The sample follows a Normal distribution.

Ha: The sample does not follow a Normal distribution.
As the computed *p-value* is lower than the significance level alpha=0.05, one should reject the null hypothesis H0, and accept the alternative hypothesis Ha.

The risk to reject the null hypothesis H0 while it is true is lower than 4.33%.

Rainfall in June

Summary statistics:

Variable	Obs.	Obs. with missing data	Obs. without missing data	Min.	Max.	Mean	Std. deviation
Rainfall JUNE	14	0	14	0.000	21	2.857	6.112

Shapiro-Wilk test (Rainfall JUNE):

W	0.547
p-value	< 0.0001
Alpha	0.05

Test interpretation:

H0: The sample follows a Normal distribution.

Ha: The sample does not follow a Normal distribution.
As the computed *p-value* is lower than the significance level alpha=0.05, one should reject the null hypothesis H0, and accept the alternative hypothesis Ha.

The risk to reject the null hypothesis H0 while it is true is lower than 0.01%.

Anderson-Darling test (Rainfall JUNE):

A^2	2.828
p-value	< 0.0001
Alpha	0.05

Test interpretation:

H0: The sample follows a Normal distribution.

Ha: The sample does not follow a Normal distribution.
As the computed *p-value* is lower than the significance level alpha=0.05, one should reject the null hypothesis H0, and accept the alternative hypothesis Ha.

The risk to reject the null hypothesis H0 while it is true is lower than 0.01%.

Rainfall in August

Summary statistics:

Variable	Obs.	Obs. With missing data	Obs. without missing data	Min.	Max.	Mean	Std. deviation
Rainfall AUG	14	0	14	0.000	24.000	4.214	7.628

Shapiro-Wilk test (Rainfall AUG):

W	0.640
p-value	< 0.0001
Alpha	0.05

Test interpretation:

H0: The sample follows a Normal distribution.

Ha: The sample does not follow a Normal distribution.
As the computed *p-value* is lower than the significance level alpha=0.05, one should reject the null hypothesis H0, and accept the alternative hypothesis Ha.

The risk to reject the null hypothesis H0 while it is true is lower than 0.01%.

Anderson-Darling test (Rainfall AUG):

A^2	2.307
p-value	<0.0001
Alpha	0.05

Test interpretation:

H0: The sample follows a Normal distribution.

Ha: The sample does not follow a Normal distribution.
As the computed *p-value* is lower than the significance level alpha=0.05, one should reject the null hypothesis H0, and accept the alternative hypothesis Ha.

The risk to reject the null hypothesis H0 while it is true is lower than 0.01%.

Rainfall in September

Summary statistics:

Variable	Obs.	Obs. with missing data	Obs. without missing data	Min.	Max.	Mean	Std. deviation
Rainfall SEP	14	0	14	0.000	131.000	24.214	36.703

Shapiro-Wilk test (Rainfall SEP):

W	0.692
p-value	0.000
Alpha	0.05

Test interpretation:

H0: The sample follows a Normal distribution.

Ha: The sample does not follow a Normal distribution.
As the computed *p-value* is lower than the significance level alpha=0.05, one should reject the null hypothesis H0, and accept the alternative hypothesis Ha.

The risk to reject the null hypothesis H0 while it is true is lower than 0.03%.

Anderson-Darling test (Rainfall SEP):

A^2	1.629
p-value	0.000
Alpha	0.05

Test interpretation:

H0: The sample follows a Normal distribution.

Ha: The sample does not follow a Normal distribution.
As the computed *p-value* is lower than the significance level alpha=0.05, one should reject the null hypothesis H0, and accept the alternative hypothesis Ha.

The risk to reject the null hypothesis H0 while it is true is lower than 0.02%.

Apendix 2. Normality Test for residuals

Rainfall in May
Spain

Summary statistics:

Variable	Obs.	Obs. with missing data	Obs. without missing data	Min.	Max.	Mean	Std. deviation
Residual	14	0	14	-5.097	3.13	0.000	2.473

Jarque-Bera test (Residual):

JB (Observed value)	2.267
JB (Critical value)	5.991
DF	2
p-value	0.322
Alpha	0.05

Test interpretation:

H0: The sample follows a Normal distribution.

Ha: The sample does not follow a Normal distribution.
As the computed *p-value* is greater than the significance level alpha=0.05, one cannot reject the null hypothesis H0.

The risk to reject the null hypothesis H0 while it is true is 32.19%.

Rainfall in June
Burkina Faso

Summary statistics:

Variable	Obs	Obs. with missing data	Min.	Max.	Me an	Std. deviati on
Residual	14	0	.152	7.073	0.0	4.383

Jarque-Bera test (Residual):

JB (Observed value)	0.687
JB (Critical value)	5.991
DF	2
p-value	0.709
Alpha	0.05

Test interpretation:

H0: The sample follows a Normal distribution.

Ha: The sample does not follow a Normal distribution.
As the computed *p-value* is greater than the significance level alpha=0.05, one cannot reject the null hypothesis H0.

The risk to reject the null hypothesis H0 while it is true is 70.94%.

Libya

Summary statistics:

Variable	Observations	Obs. with missing data	Min.	Max.	Mean	Std. Dev.
Residual	14	0	-8.286	6.143	0.000	4.116

Jarque-Bera test (Residual):

JB (Observed value)	0.794
JB (Critical value)	5.991
DF	2
p-value	0.672
Alpha	0.05

Test interpretation:

H0: The sample follows a Normal distribution.

Ha: The sample does not follow a Normal distribution.
As the computed *p-value* is greater than the significance level alpha=0.05, one cannot reject the null hypothesis H0.

The risk to reject the null hypothesis H0 while it is true is 67.23%.