UNIVERSIDADE FEDERAL DE JUIZ DE FORA

INSTITUTO DE CIENCIAS BIOLÓGICAS (ICB)

PROGRAMA DE PÓS-GRADUAÇÃO EM ECOLOGIA (PGECOL)

Alberto Benito Mirambell

Título:

APPLICATION OF SPATIOTEMPORAL TECHNIQUES TO ESTIMATE EVAPOTRANSPIRATION IN THE PARAÍBA DO SUL RIVER WATERSHED

Juiz de Fora

2018

Alberto Benito Mirambell

Título:

APPLICATION OF SPATIOTEMPORAL TECHNIQUES TO ESTIMATE EVAPOTRANSPIRATION IN THE PARAÍBA DO SUL RIVER WATERSHED

Dissertação apresentada ao Programa de Pós-graduação em Ecologia (PGECOL), da Universidade Federal de Juiz de Fora como requisito parcial a obtenção do grau de Mestre em Manejo de Ecossistemas e Conservação da Biodiversidade. Área de concentração: Ecologia

Orientador: Dr. Celso Bandeira de Melo Ribeiro, Doutor em Engenharia Agrícola, Universidade Federal de Viçosa

Juiz de Fora

2018

Ficha catalográfica elaborada através do programa de geração
automática da Biblioteca Universitária da UFJF,
com os dados fornecidos pelo(a) autor(a)

Benito Mirambell, Alberto.
APPLICATION OF SPATIOTEMPORAL TECHNIQUES TO ESTIMATE EVAPOTRANSPIRATION IN THE PARAÍBA DO SUL RIVER WATERSHED/ Alberto Benito Mirambell. -- 2018.
203 f. : il.

Orientador: Celso Bandeira de Melo Ribeiro
Coorientador: Ricardo Guimarães Andrade
Dissertação (mestrado acadêmico) - Universidade Federal de Juiz de Fora, Instituto de Ciências Biológicas. Programa de Pós Graduação em Ecologia, 2018.

1. Evapotranspiration. 2. Penman-Monteith. 3. Artificial Neural Networks. 4. METRIC algorithm. 5. Remote Sensing . I. Bandeira de Melo Ribeiro , Celso, orient. II. Guimarães Andrade, Ricardo, coorient. III. Título.

Alberto Benito Mirambell

Título:

APPLICATION OF SPATIOTEMPORAL TECHNIQUES TO ESTIMATE EVAPOTRANSPIRATION IN THE PARAÍBA DO SUL RIVER WATERSHED

Dissertação apresentada ao Programa de Pós-graduação em Ecologia, da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do grau de Mestre em Manejo de Ecossistemas e Conservação da Biodiversidade. Área de concentração: Ecologia

Aprovada em 23 de fevereiro de 2018

BANCA EXAMINADORA

_____

Dr. Celso Bandeira de Melo Ribeiro

Universidade Federal de Juiz de Fora

_____

Dr. Ricardo Guimarães Andrade

Embrapa, Gado de Leite, Juiz de Fora

_____

PhD. Otto Corrêa Rotunno Filho
Universidade Federal do Rio de Janeiro

_____

Dr. Jonathas Batista Gonçalves Silva

Universidade Federal de Juiz de Fora

Dedico este trabalho ao meu pai, pois devo muito do que eu sou a ele, àminha mãe que me viu partir na procura de um futuro melhor há sete anos já, e sobre tudo dedico este trabalho à minha esposa, Claudia, que foi meu apoio e minha companheira durante os quatro anos que morei no Brasil e compartilhou momentos ruins e bons, que acompanhou todo esse crescimento e transformação, até chegar ao que eu sou hoje.

AGRADECIMENTOS

Ao chegar neste ponto e olhando tudo o que foi realizado e todas as pessoas que encontrei durante estes últimos dois anos, quero agradecer especialmente aqueles que me trouxeram alguma coisa de bom. A Marcel Pereira, por ter sempre sua mão estendida e um oferecimento de ajuda. A Clayton Felício e ao Professor Flávio Barbosa por fazer parte da minha primeira publicação internacional. Ao Professor Roberto da Gama, pelas aclarações necessárias na hora certa. À Rodrigo Santis e Jonata Jefferson Andrade, pela inestimável ajuda com à aplicação de uma das metodologias empregadas ao longo deste trabalho. A Ricardo Guimarães, pela paciência e um excelente tratamento. A Anne Carvalho pelas palavras de suporte e ânimo. À Nishan Battarai, por compartilhar comigo alguns recursos que foram muito úteis. Ao Brasil por me mostrar à importância das coisas que eu achava “óbvias”. O agradecimento mais especial e sincero é para meu orientador, Celso, que não só me orientou academicamente durante este trabalho, mas também em um momento da minha vida onde precisava de uma “luz”, ele foi essa “luz” que me ajudou a descobrir uma capacidade e um talento que nem eu sabia que tinha. GRACIAS de corazón.

RESUMO

No dia de hoje, qualquer estudo relacionado aos recursos hídricos e seus usos, tais como irrigação, abastecimento de água e geração de energia, é de suma importância em função dos cenários que vivemos atualmente face às variabilidades climáticas. O uso eficiente desses recursos faz-se cada vez mais necessário, envolvendo fatores como a estimativa de algumas variáveis relacionadas ao ciclo hidrológico, notadamente a evapotranspiração. Há quase trinta anos a FAO recomendou o uso da equação de Penman-Monteith para a estimativa da evapotranspiração de referência. Desde então tem sido aplicada com sucesso em diferentes regiões e sob diferentes climas. No entanto, esta abordagem tem algumas desvantagens, entre elas, o fato de depender de medições de campo de parâmetros climáticos, tais como temperatura, humidade do ar, velocidade do vento e radiação solar. Além disso, essas medições são pontuais em referência ao local de operação da estação meteorológica e podem não representar de forma fidedigna as condições climáticas dos ambientes circundantes. Nos últimos tempos, com o avanço da tecnologia, o desenvolvimento de potentes linguagens de programação orientados à análise de dados, o surgimento das técnicas na área de inteligência artificial e do tratamento de grandes volumes de dados ("*Big Data"*), surgiram ferramentas com grande potencial para melhorar a forma como se tratam os eventos naturais ou antrópicos, permitindo maior eficiência e produtividade. Nessa linha, o objetivo principal do presente estudo é o uso desse conjunto de tecnologias para uma estimação confiável e robusta da evapotranspiração, na medida que constitui uma variável fundamental no fechamento do balanço hídrico no nível de uma bacia hidrográfica. Complementarmente, essa estimativa poderia ser também empregada como indicativo da perda água em uma cultura pelo agricultor. Em especial, dois procedimentos foram aplicados ao longo deste trabalho: redes neurais artificiais (RNA's) e o algoritmo METRIC. O primeiro está associado a uma ferramenta com base em inteligência artificial, capaz de reproduzir o comportamento de certas variáveis com um alto nível de semelhança abrindo a possibilidade de gerar predições a curto-maio prazo que ajude no gerenciamento dos recursos hídricos por parte dos comités de bacia e outros entes responsáveis por eles. Por outro lado, METRIC permite usar imagens de satélite para estimar evapotranspiração em escala horária, capturando as abruptas mudanças que sofrem algumas variáveis climáticas ao longo do dia, sendo esta uma informação de vital importância para os agricultores determinarem a irrigação com maior confiabilidade. Os resultados obtidos após a aplicação de ambos os procedimentos, que compõem a abordagem metodológica deste trabalho, foram muito satisfatórios e com uma alta correlação com aqueles gerados pela metodologia considerada como referência. Assim sendo, pode-se concluir que ambos procedimentos formam um referencial apropriado na estimativa de valores de evapotranspiração que podem ser transferidos à prática agrícola com a certeza de uma melhora constante a tenor da rápida e imparável evolução da tecnologia na área da agricultura de precisão.

Palavras-chave: Evapotranspiração, Penman-Monteith, Redes Neurais Artificiais, METRIC, Sensoriamento Remoto

ABSTRACT

Nowadays, any study related to water resources and its usage, such as irrigation, water consumption and energy production is a central issue due to the climate change scenario we are currently living. The efficient use of such resources is a must and involves several factors, among them, the estimate of some hydrologic cycle-related variables, highlighting evapotranspiration, among them. About thirty years ago, the FAO recommended Penman-Monteith equation as the most trustworthy and representative methodology to estimate reference crop evapotranspiration. Since then, it has been applied successfully over different regions and under diverse weather conditions. However, this approach has some cons, such as its dependency on ground measurements of most common climatological parameters: temperature, relative air humidity, wind speed or solar radiation. In addition, these measurements are punctual on the weather station's location and may not fully represent surrounding environments' conditions. Lately, thanks to technological advances, the development of powerful programming data analysis-oriented languages, the rising of artificial intelligence, as well as big data, we have a wide variety of tools to improve the way we analyse natural phenomena, making it more efficient and productive. Therefore, the main objective of the present study is the use of such technologies aiming to estimate reliable evapotranspiration values, as a central parameter on water resources management at watershed basis, or even as an indicator of crop water loss, by farmers. Mainly, two different technology-based approaches have been applied along this dissertation pursuing the objective previously mentioned: artificial neural networks (ANN's) and METRIC algorithm. The former is an artificial intelligence-based tool, capable of "recording" specific variables behaviour and succeed in "mimicking" them at a high resemblance level, favouring the possibility of short-half term forecasts to help watershed committees and other responsible bodies manage water resources. On the other hand, METRIC algorithm uses satellite imagery in order to estimate evapotranspiration hourly and able thus to catch the disrupting changes some of the climatological variables suffer along the day, turning into a vital piece of information for farmers, since they can design irrigation schedule more precisely. Results obtained after both procedures' application, which compose the methodological approach throughout this study, fully satisfied our expectations and showed a high correlation to those results estimated by the methodology of reference. To sum up, we conclude that both approaches are reliable at estimating reference crop evapotranspiration and can be transferred to the agricultural management assuring a steady improvement due to the quick and unstoppable evolution in technology on the "agriculture of precision" field.

Keywords: Evapotranspiration, Penman-Monteith, Artificial Neural Networks, METRIC, Remote Sensing

SUMÁRIO

CHAPTER II

ESTIMATION OF MISSING VAUES VIA ARTIFICIAL NEURAL NETWORKS

CHAPTER III

INFLUENCE OF CLIMATOLOGICAL PARAMETERS ON THE EVAPOTRANSPIRATION PROCESS

CHAPTER VI

REMOTE SENSING: THE CUTTING-EDGE TOOL ON EVAPOTRANSPIRATION ESTIMATES

INTRODUCTION

Throughout the present study, the reader will find the word “evapotranspiration” more than one hundred times, which points out who is the main character in this story. To start from, evapotranspiration is defined as a complex physical parameter composed by two procedures that happen simultaneously in the nature and which mainly depends on some climatological parameters. This is only the starting point of a story trying to highlight the importance of water, as well as the necessity to use it as efficiently as possible. It is here where we find the connection point with the author and the reason that led him to spend all the effort invested in this research.

Born in the bosom of a farming family in the Southeast of Spain, the author learnt soon how to deal with scarcity of water, the regular condition in the region. A region responsible for the 73% of table grapes consumed in Spain and up to 65.000 Tn of grapes exported to more than 15 countries over the world. Figures highlighting that a vast part of the regional economy strongly depends on vineyards. Not many harvesting procedures or transport chain elements have changed in the last decades, the productivity has raised and the production costs has gone down. However, producers, cooperatives and insurance companies started to worry about water lately and its accessibility in the near future, since drought turned into the regular condition. It was obvious something had to be done, since the gravitational irrigation system, in use until then, was recognised as a low efficient one. It required great volumes of water, since a large part of it was lost by deep percolation. It existed even an irrigation schedule where the farmer had to “reserve” its turn as the rest of farmers who depended on the same irrigation community, so there was a sort of waiting list. This irrigation procedure required some previous and posterior tasks, apart from the irrigation event itself, which used to take several hours. Definitely, it was time to change something.

All public bodies and private investors agreed on installing a much more efficient irrigation system, which had brought excellent results on other areas where was already in operation: drip irrigation. With it, nowadays farmers save up to 65% of the water volume they used to spend with the gravitational system. Besides from that, they had on their hands the decision on when, where and how much to irrigate; irrigators community were able to manage

water and users more easily and with less work; infrastructure was fully modernised, allowing a better and more accurately monitoring of water consumption and a great improvement in efficiency. Even though, drip irrigation could lose its efficiency and advantages when managed inadequately.

Until today, traditional farming is based on experience passed from parents to children, with no sign of scientific base. In the last two decades, new-generation farmers have introduced some agronomy concepts, devices and procedures. More and more fields are under agronomy engineering companies aiming to raise productivity at the same time as to reduce costs and environmental impacts. High-tech is all around and farming could not be set apart. From this union, Precision Agriculture was born, defined as a farming management concept based on observing, measuring and responding to inter and intra-field variability in crops. Its goal is to become a decision support system for whole farm management, allowing to optimize returns on inputs while preserving resources. It is within this framework where the author of the present work looked for a high-tech procedure able to be applicable on the field with results good enough to represent a significant improvement in water resources management in the region described above.

The main objective of the present study is thus to apply cutting-edge procedures to estimate evapotranspiration at a high resemblance level to the reference approach, Penman-Monteith equation. Among these high-tech procedures, we would highlight the application of METRIC algorithm, developed by Rick Allen between 2000 and 2005, as a head of the Idaho Department of Water Resources. METRIC allows to estimate evapotranspiration instantaneously, which permits to capture the great changes some climatological parameters suffer along the day. These changes also make evapotranspiration to vary widely, whose fluctuations do not go unnoticed for METRIC. Such high frequency in monitoring evapotranspiration gives to farmers the opportunity of when, where and how much water is lost by evapotranspiration and thus, when, where and how much water is needed to be replaced by drip irrigation.

What this approach permits is to take the decision of irrigating off the farmer’s hands and leave it to the climatological conditions. Values of evapotranspiration estimated by METRIC will determine whether it irrigates or not, how much water must be replaced, as well as the right timing and the right place to irrigate. The other high-tech procedure applied throughout this study is the artificial neural network, a tool based on artificial intelligence

capable of estimating new evapotranspiration values from previous ones at high resemblance level. It implies that artificial neural networks are able to produce reliable evapotranspiration estimates at low cost.

However, the path to get there has been long and many questions have risen along it; questions related to the diverse ways of measuring and estimating evapotranspiration, how the different climatological parameters influence evapotranspiration or if it is possible to use other factors, besides evapotranspiration, as indicators of soil moisture and thus, of water availability. The reader may consider these ones as secondary objectives, but in any case less essential than the main ones, as they can help the reader get a global view of the evapotranspiration as a whole, as well as the individual factors that rule such process.

The chapters' order in this document has been designed to finally reach a reliable evapotranspiration value estimated by METRIC, which is the main target of this study and follow the order of happening of the different experimental procedures that compose the present study. As a high-tech approach and part of a rising study field, known as "precision agriculture", remote sensing is currently considered as a high-potential tool due to the fierce spatial race involving the greatest economies in the world, such as USA, Europe, China or Russia. This competition will surely bring some short-half term benefits, such as better spatiotemporal resolution products, an easier access to earth observation imagery, sensors able to detect more and more parameters, development of new and more capable algorithms and so on. The future perspectives are massive so as to amplify the fields of study, the analysing tools, the information sources and the diversity of results. But it is essential to keep in mind that any study, analysis, result, interpretation, conclusion or a whole research must focus on its applicability in the real world, bringing to it the benefits that are supposed to it.

To get to an end and as the reader can realise, the main objective of this work is a clear reflection of the author's personal reason to carry out the present research: develop an unmanned remote sensing-based approach able to estimate crop water demand with enough assurance so as to irrigate by drip irrigation only what has been previously evapotranspirated.

CHAPTER I
EVAPOTRANSPIRATION: THE KEY CONCEPT TO MANAGE SUCCESSFULLY THE SOIL WATER DEMAND

1. EVAPOTRANSPIRATION

Evapotranspiration (*ET*) is an important component of the hydrologic cycle. Its estimation plays a central role in different fields related to hydrology such as hydrologic water balance, impact of land uses assessment, water resources planning and management and irrigation system design (ALLEN *et al.*, 1994, p. 92).

First of all, it would be necessary to have a deep insight into the essentials of evapotranspiration as a physical process influenced by different climatic parameters. Basically, evapotranspiration is the physical process through which water is transferred to the atmosphere both by evaporation from land and other surfaces and via transpiration through plants, so it is a combined process whereby soil moisture returns to the atmosphere (ALLEN *et al.*, 1998, p. 236). As it is a physical mechanism compound by two processes running simultaneously it gets unable to separate to each other when it comes to measuring its quantities or estimating its values. Conversely, we can expose both individually in order to better comprehend for what accounts each of them or what their driving climatic parameters are.

1.1 Evaporation

Evaporation is the process through which liquid water is converted to water vapour and removed from the evaporating surface. Evaporation takes place in every surface containing any trace of water. Therefore, water evaporates from a variety of surfaces, such as lakes, rivers, pavements, soils and wet vegetation (SLATYER; MCILROY, 1961, p. 395; PERRIER, 1982, p. 448).

Energy is required in order to vaporize the liquid water while solar radiation and air temperature represent the main sources of energy. Once the molecules of water have been transformed into vapour, the difference between the water vapour pressure at the evaporating surface and that of the surrounding atmosphere acts as the driving force to remove water vapour from the evaporating surface. This difference slowly decreases as the process continues since the surrounding air becomes gradually saturated until getting to a point where the difference is zero; at this point the process would stop if the wet air would not be transferred to the atmosphere. It is here where wind speed acts replacing the saturated air with drier air. Thus, solar radiation, air temperature, air humidity and wind speed represent the main climatic parameters to consider when assessing the evaporation process.

Minor parameters should be considered when we take into account the evaporation process taking place on any soil surface. Here, the degree of shading of the crop canopy and the availability of soil moisture are other factors that affect the evaporation process (MONTEITH, 1965, p. 266; SMITH *et al.*, 1991, p. 218). There are surfaces capable of constantly suppling enough water so as to satisfy the evaporation demand; in these places the evaporation process is ruled uniquely by the meteorological conditions. However, these cases are rarely on the nature, where usually the intervals between rains and irrigation are widely variable and sometimes even the ability of the soil to conduct moisture to reach the surface is small (JENSEN *et al.*, 1990, p. 970; SMITH *et al.*, 1991, p. 214) Under these limiting conditions in terms of water availability, evaporation decreases rapidly and may cease almost completely within a few days.

The evaporation rate even changes along with the growing period of the crop. Indeed, Evaporation rate decreases as the crop develops, since the degree of shading of the canopy increases cutting down, at the same time, the portion of the direct solar radiation reaching the evaporating surface. As an example, for crops on initial-growing stages, water is predominantly lost by evaporation due to the absence of canopy, while late-growing stages are characterized by high transpiration rates (>90%), as a well-developed canopy covers completely the soil (RIJTEMA, 1965, p. 172; PERRIER, 1985, p. 887).

1.2 Transpiration

Transpiration consists of the vaporization of liquid water contained in plant tissues and the vapour removal to the atmosphere by means of leaf stomata. The water, extracted from the soil by the roots, is transported through the plant stems along with some nutrients until the leaves, where the vaporization occurs and it is controlled by the stomatal aperture (ALLEN *et al.*, 1989, p. 662; ALLEN *et al.*, 1994, p. 92; ALLEN *et al.*, 1998, p. 240). Transpiration also depends on the energy supply being radiation, air temperature, air humidity and wind the driving factors which influence the process. Water availability, salinity conditions of the soil, the ability of it to conduct water to the roots also determine the transpiration rate, as well as crop characteristics, environmental aspects and cultivation practices should be considered when assessing transpiration (JENSEN *et al.*, 1990, p. 970; SMITH *et al.*, 1991, p. 263).

It is obvious that evapotranspiration measures an amount of water, although, most times evapotranspiration is understood as a rate, so it is necessary to have an "amount of time or piece of surface" so as to reference that measure (PENMAN, 1948, p. 145; MONTEITH, 1965, p. 421). Evapotranspiration is usually measured in terms of millimetres (mm) of water depth per unit time, such as hour, day, month or year and per hectare (ha). So that, being 1 hectare 10,000 m^2 and 1 mm corresponds to 0.001 m, a *ET* rate of 1 mm day^{-1} would correspond to a loss of 10 m^3 ha^{-1} day^{-1}. Recalling that *ET* is a physical process which requires energy to vaporize the liquid water into vapour, *ET* can also be considered as a part of an energy balance (RIJTEMA, 1965, p. 211). This approach is as reasonable as the one it has been proposed here. Accordingly, *ET* can be expressed in terms of energy or heat to vaporize the water received per unit area. This energy, the so-called latent heat of vaporization (Equation 1), depends on the water temperature. It is taken as standard that 2.45 MJ are needed to vaporize 1 kg or 0.001 m^3 of water. Thus, 2.45 MJ per m^2 are able to vaporize 0.001 m or 1 mm of water, and therefore 1 mm of water is equivalent to 2.45 MJ m^{-2}. The evapotranspiration rate expressed in units of MJ m^{-2} day is represented by $\boldsymbol{\lambda}ET$, widely known as the latent heat flux and being part of the energy balance equation as follows:

$$R_n = \lambda ET + G + H \qquad \textbf{(1)}$$

Where R_n is the net radiation, the only entrance of energy to the system, as mentioned above. *G* is known as the soil heat flux, which accounts for the amount of energy gained and lost by the soil along the period under study. It tends to be zero on daily measures, since the energy income tends to be cancelled out by the energy outgoings. *H* is the sensible heat flux and accounts for the convection and conduction mechanisms taking place right above of the soil. It allows the heat emitted by the soil to dislocate somewhere else and avoid over-saturated conditions for the surrounding atmosphere .

2. TERMINOLOGY RELATED TO EVAPOTRANSPIRATION

Evapotranspiration has always been a concept widely used as far as water resources management is concerned, but there have always been some ambiguity in the use of such terms as potential *ET* or reference crop *ET*. In order to solve this issue, the Food and Agricultural Organization of the United Nations (FAO) and through the publication of FAO Irrigation and Drainage Paper No. 56 in 1990 (ALLEN *et al.*, 1998, p. 156) shed some light on the problem, helping users with uniformity on the use of such terminology. Indeed, evapotranspiration is the term used to refer to the physical process described in the first section. When it comes to the terminology created along the years of research by hundreds of scientists who worked on the field, it gets certainly confusing to distinguish ones from others. At this point, it makes necessary to clarify the differences between them as we will use them several times along this dissertation.

Basically, we can distinct between reference crop evapotranspiration (ET_o), crop evapotranspiration under standard conditions (ET_c) and crop evapotranspiration under non-standard conditions ($ET_{c\ adj}$).

2.1 Reference crop evapotranspiration (ET_o)

ET rate from a well-watered reference surface is called as "reference crop *ET*" or "reference evapotranspiration" and denoted as *ETo* (ALLEN *et al.*, 1998, p. 245). For the FAO, a reference surface must be a hypothetical surface "closely resembling an extensive surface of green grass of uniform height, actively growing, completely shading the ground and with adequate water".

DOORENBOS & PRUITT (1977, p. 398) defined reference crop evapotranspiration rate as "the rate of evapotranspiration from an extensive surface of 0.8-1.5 m tall, green grass cover of uniform height, actively growing, completely shading the ground and no short of water". As we can notice, the reference surface equates to a hypothetical grass reference crop with specific features. By definition, *ETo* is a measure independent of the type of crop, its

growing stage, management practices and soil conditions, as the reference surface is “not short of water”. In its definition only the effects produced by its driving climatological parameters are taken into account. In other words, *ETo* is considered as a way of assessing the evaporative demand of the atmosphere.

This fact entails that other *ET* measures taken in different places, over different type crops, under different management practices can be compared to one another since it exists a common reference surface for both surfaces. Once in this point, the focus moves towards the climatological parameters that rule the behavior of the *ETo*. They will be different depending on which method we would be using to measure or estimate *ETo*. We will be back on this subject later on.

2.2 Crop evapotranspiration under standard conditions (ET_c)

It refers to the Evapotranspiration rate of any crop free of any disease, well-fertilized and grown in large areas under ideal soil and water conditions and reaching its optimal development and full production under the given climatic conditions (PENMAN, 1948, p. 145; MONTEITH, 1965, p. 256; ALLEN *et al.*, 1998, p. 109). The type of crop, the leaves anatomy, the stomatal conductance, the growing stage or even the albedo gain importance now and the compound of these factors represents the main difference with respect to *ETo*, which does not take these factors into account, as mentioned previously.

All of these parameters are somewhat summarize and integrated into coefficients that account for the differences between different crops. This Crop Coefficients, as they are properly called and denoted as *Kc* (Equation 2) can be directly calculated by the following equation and determined experimentally straight away from several crops:

$$K_c = \frac{ET_c}{ET_o} \quad \textbf{(2)}$$

2.3 Crop evapotranspiration under non-standard conditions ($ET_{c\ adj}$)

It corresponds to the *ET* rate of crops under climatological conditions and management practices different from the ideal ones which define the crop evapotranspiration under standard conditions (ALLEN *et al.*, 1998, p. 156). Most times these are the actual circumstances found in the Nature and even along cultivated fields as it is common for these crops to suffer any disease or plague, register non-optimal salinity measurements, experience water shortage or waterlogging periods or low soil fertility.

Any of these non-desired conditions, pretty usual on the other hand, make the *ET* rate deviate from *ETc* values, generally slowing down its values and, consequently, leading to a reduced crop growth or low plant density. To calculate the $ET_{c\ adj}$ (Equation 3) is necessary to include into the equation a hydric stress factor or coefficient (K_s) which accounts for the effect of all of these limitations on the crop in the following way:

$$ET_{cadj} = K_s * K_c * ET_o \quad \textbf{(3)}$$

3 FACTORS INFLUENCING EVAPOTRANSPIRATION

As it was presented in the previous paragraphs, each "type" of *ET* is influenced by different factors. Firstly, *ETo* is influenced only by climatological parameters, namely radiation, air temperature, air relative humidity and wind speed. In this case and as mentioned above, *ETo* can be considered as a measurement of the evaporative power of the atmosphere (SMITH *et al.*, 1991, p. 123; ALLEN *et al.*, 1998, p. 249).

Radiation is the main entrance of energy to the system composed by the plant, the soil underneath and the surrounding atmosphere. The amount of radiation reaching an evaporative surface depends on its location on the Earth and the period of the year, which greatly influences the incidence angle of the solar rays. Cloudiness is other important factor that determine the final amount of radiation passing through the atmosphere as it behaves as filter. Furthermore, from the radiation filtered by the clouds only a little portion is invested in transpiration, since it is also used to heat up the soil, the atmosphere, the water bodies, to run the photosynthesis and so many others mechanisms which also need a source of energy to work.

Air temperature is a direct consequence of the inbound radiation since it heats up both the soil and the atmosphere. The sensible heat from the surrounding air transfers energy to the crop exerting, at the same time, such a controlling effect on the *ET* rate. It is obvious then, that in a sunny, warm day the *ET* rate would be higher than in a cloudy, cool day (RIJTEMA, 1965, p. 485).

As mentioned along the first paragraphs of this item 3, radiation and air temperature act as the main driving factors for water evaporation, although the difference between the water vapor pressure right on the evaporating surface and the surrounding air is the key process in charge of removing the vapor just emitted (SLATYER; MCILROY, 1961, p. 562). Air relative humidity can be as determinant as radiation or air temperature since a dry atmosphere acts with a heavy removal power and cause high water consumptions, provided that there is water available. On the other hand, saturated environments exert a deterrent effect on evapotranspiration, despite the high radiation levels reaching the evaporative surface.

As the evapotranspirative process moves on the surrounding environment saturates gradually. It is necessary to remove this humid air and replace it for dry air in order to continue the process. Wind speed is the one in charge to do so. Indeed, the removing process of vapor depends greatly on the wind speed and on the air turbulence which are responsible for transferring large quantities of air over the evaporative surface (JENSEN *et al.*, 1990, p. 970; SMITH *et al.*, 1991, p. 254).

Secondly, crop type, crop variety and development stage are others factors to be taken into account when assessing the *ET* of crops growing intensively on large and well-managed fields, apart from the meteorological parameters described above. Differences in resistance to transpiration, crop height, crop roughness, reflection, ground cover and crop rooting characteristics result in different *ET* levels in different types of crops under identical environmental conditions (IRMAK *et al.*, 2003, p. 347). All these differences are summarized and integrated somewhat into the so-called crop coefficients and denoted as *Kc*. Once the *ET* rate count on these crop coefficients we are no longer referring to *ETo*, but crop evapotranspiration under standard conditions (ET_c). By recalling its definition, *ETc* refers to the *ET* rate of any crop free of any disease, well-fertilized and grown in large areas under ideal soil and water conditions and reaching its optimal development and full production under the given climatic conditions.

Finally, a third group of parameters that exert an effect on *ET* rate are related not to the crop itself, but to the environmental conditions and managements practices dominant on the crops. Factors such soil salinity, low soil fertility, use of fertilizers, existence of hard soil layers, presence of plagues or parasites and a wrong soil management can limit the optimal development of the crop, as well as reduce the *ET* rate. A central factor to be taken into account seriously is the soil moisture availability. Its effect is determined both by the magnitude of the hydric deficit and by the soil type. Low water content in the soil would restrict its extraction by the rooting system slowing down the *ET* rate as there is no enough water to transpirate (KOTSOPOULOS; BABAJIMOPOULOS, 1997, p. 256).

Conversely, waterlogging conditions can damage the rooting system and reduce its capability of extraction. The maximum water quantity a soil is able to retain varies with the soil texture and structure, the organic matter content and the rooting system depth. Organic matter content of the soil is especially important as it can stock water up to 20 times its weight.

Organic soils with a medium texture and high content of very thin sand and sediments are characterized by higher disposable capacity of water, while clay soils present medium values.

4 DETERMINATION OF EVAPOTRANSPIRATION BY TRADITIONAL INSTRUMENTS

Despite evapotranspiration is not a physical process easy to determine, it can be measured experimentally. To do so, very specific equipment is required aiming to obtain high accurate registers of the climatological parameters under study. Field experimentally procedures are generally costly requiring very precise procedures and measurements being fully applied and appropriately analysed exclusively by researchers. For the sake of these reasons, these methods have become restricted to the research field and used for calibrating the measurements taken and validating the later results.

5.1 Lysimeters

Lysimeters, also called soil columns, are receptacles containing a sample of the surrounding soil destined to measure the quantity of water lost by the soil sample within either by evaporation or/and transpiration (HOWELL *et al.*, 1991, p. 9). It is also used to determine the content in soluble salts and nutrients the soil loses by leaching. The optimal surface area and size of a lysimeter depend mainly on the research aim, the filling procedure, the lower boundary, and the location of installation. Lysimeters are usually filled with the natural crop and the maximal root penetration depth should be taken into account. They can be filled with either non-disturbed or disturbed soil or non-reactive materials such as sand. Disturbed lysimeters can be filled with the disturbed horizons of natural soil or artificial material. In this case, the natural texture and the spatial heterogeneity will change and, consequently will result in changes of water and solute flow (HOWELL *et al.*, 1991, p. 9).

Despite the lysimeter installation can be done in laboratories, within greenhouse boundaries or at special lysimeter facilities, its installation is broadly carried out on the field as one of the objectives of using this equipment is to replicate the natural conditions within the lysimeter (MAKKINK, 1957, p. 288; GREBET E CUENCA, 1991, p. 18; NEALE *et al.*, 1991, p. 169). In this case the upper surface of the lysimeters should be equal to the ground surface

to minimize microclimatic changes. The space between the lysimeter and the surrounding soil should also be minimized to reduce artificial temperature gradients within the soil block. For acquisition of the surface run-off, lysimeters can also be fitted with run-off / overflow tubes. Its applications are varied since they depend in a large extent to the objective of the research. The main applications are mentioned just below:

1. To obtain direct measurements of both evaporation and transpiration assisting in the study of the parameters affecting such physical mechanisms.
2. To use these direct measurements so as to compare them with *ET* values estimated by indirect determining methods.
3. To carry out leaching process studies in order to better understand how the demineralizing process occurs.
4. To determine the magnitude and to analyse the percolating process within a demarcated space.
5. As a way of knowing the quantities of water, nutrients and fertilizers are needed to be replaced after the run-off and percolating take place.

5.1.1 Types of lysimeters

There are two basic types of lysimeter:

A) weighing lysimeters (WL) for agricultural purposes: in this case, potential evaporation is obtained as the difference between added and drained water quantity. Weighing lysimeters have been used for many years to measure and study water use, to calibrate reference *ET* methods for a local area, and to develop crop-coefficient functions for specific crops (MEYER *et al.*, 1990, p. 75). A weighing lysimeter provides a direct measure of the amount of water used in evaporation and transpiration by isolating and continuously monitoring a vegetated area in a field. A number of researchers around the world have reported on recent studies using lysimeters to develop crop-coefficient functions for a variety of crops, such as pulse crops in India (PANDEY, 2012, p. 58), corn in Spain (MARTÍNEZ-COBO, 2008, p. 345), rice and sunflower in India (TYAGI *et al.*, 2000, p. 54), wheat and maize in China (LIU; LUO, 2010, p. 40), and cotton and wheat in the USA (KO *et al.*, 2009, p. 1697). Early weighing lysimeters

often contain mechanical mechanisms and electrical circuitry to obtain relatively high-resolution measurements, and required regular maintenance. This could result in high costs in terms of purchase, operation, and labor. Advances in electronics, datalogging equipment, and strain-gage loadcells have allowed for the design of simpler lysimeters, which are less expensive, accurate and reliable, and require less maintenance.
With respect to the way of measuring the variation on the parameters we aim to determine, another classification can be done, discerning two types of weighing lysimeters:

A1) mechanical weighing lysimeter, where the soil sample is placed directly on the balance. The sensitivity will be high, if friction can be reduced using an advanced support construction.

A2) hydraulic weighing lysimeter, where the soil sample is placed in a tank floating on a fluid. Changes in level reflect weight changes in the sample. Extremely small weight changes can be detected using this method.

B) non-weighing lysimeters (NWL) for chemical analysis on drainage water. Also called Drainage lysimeters, they operate on the principle that *ET* is equal to the amount of rainfall and irrigation water added to the system, minus percolation, runoff and soil moisture changes. Since the percolation is a slow process, drainage lysimeters are accurate only for a long periods for which the water content at the beginning exactly equals that at end (MEYER; MATEOS, 1990, p. 237). Therefore, records of drainage lysimeters should be presented only in terms of a long-period more than one day (CHANG, 1968, p. 410) and they are not useful for estimating hourly *ET*. ALLEN et al. (1990, p. 51) discusses two types of non-weighing lysimeters:

B1) non- weighing constant water-table type, which provides reliable data in areas where a high water table normally exists and where the water table level is maintained essentially at the same level inside as outside the lysimeters.

B2) non-weighing percolation type, in which changes in water stored in the soil are determined by sampling or neutron methods or other soil humidity sensors like *TDR*, and the rainfall and percolation are measured.

For more specifications and other features, a collection of lysimeters of varying design was published by the World Meteorological Organization in 1983.

To sum up, in order to obtain reliable hourly evapotranspiration measurements, the ideal instrument is a high precision weighing lysimeter with an accuracy better that 0.1 mm of water (ABOUKHALED *et al.*, 1982, p. 68). The dimension of lysimeters used for estimation of water consumption of herbage and trees is usually about 3 – 10 m^3, in order to avoid rim effect.

MiniLysimeters (ML), characterized by reduced soil volume (less than 1 m^3), have been recently adopted due to reduced installation and managements costs and good accuracy of measurement (OKE, 2004, p. 26). Therefore, due to the limited weight of the soil block, ML can adopt the affordable and easily available load cell technology that cannot be easily applied to classic lysimeters (ERVIN E KOSKI, 2001, p. 250; DIAZ-ESPEJO; VERHOEF, 2003, p. 623). Water need of plants in lysimetric experiments is supplied by natural rain, irrigation or by an artificially maintained water table ("constant water table" lysimeters). Gravitational drainage systems are cheaper, easy to install and require practically no maintenance.

5.2 Evaporation pan estimation

Evaporation from an open water surface provides an index of the integrated effect of radiation, air temperature, air humidity and wind on evapotranspiration (SHUTTLEWORTH; WALLACE, 1985, p. 855; SHUTTLEWORTH, 1993, p. 562). However, differences in the water and cropped surface produce significant differences in the water loss from an open water surface and the crop. The pan has proved its practical value and has been used successfully to estimate reference evapotranspiration by observing the evaporation loss from a water surface and applying empirical coefficients to relate pan evaporation to ET_o.

The evaporation rate from pans filled with water is easily obtained. In the absence of rain, the amount of water evaporated during a period (mm/day) corresponds with the decrease in water depth in that period. Although the pan responds in a similar fashion to the same climatic factors affecting crop transpiration, several factors produce significant differences in loss of water from a water surface and from a cropped surface (THOM *ET AL.*, 1981, p. 736; BRUTSAERT, 1982, p. 230). These are some examples:

- Reflection of solar radiation from water in the shallow pan might be different from the assumed 23% for the grass reference surface.
- Storage of heat within the pan can be appreciable and may cause significant evaporation during the night while most crops transpire only during the daytime.
- Heat transfer through the sides of the pan occurs and affects the energy balance.
- There are also differences in turbulence, temperature and humidity of the air immediately above the respective surfaces.

Despite that the difference between pan-evaporation and evapotranspiration of cropped surfaces, the use of pans to predict *ETo* for periods of 10 days or longer may be warranted. According to SHUTTLEWORTH and WALLACE (1985, p. 501), the pan evaporation is related to the reference evapotranspiration by an empirically derived pan coefficient:

$$ETo = K_P * E_{pan} \quad \textbf{(4)}$$

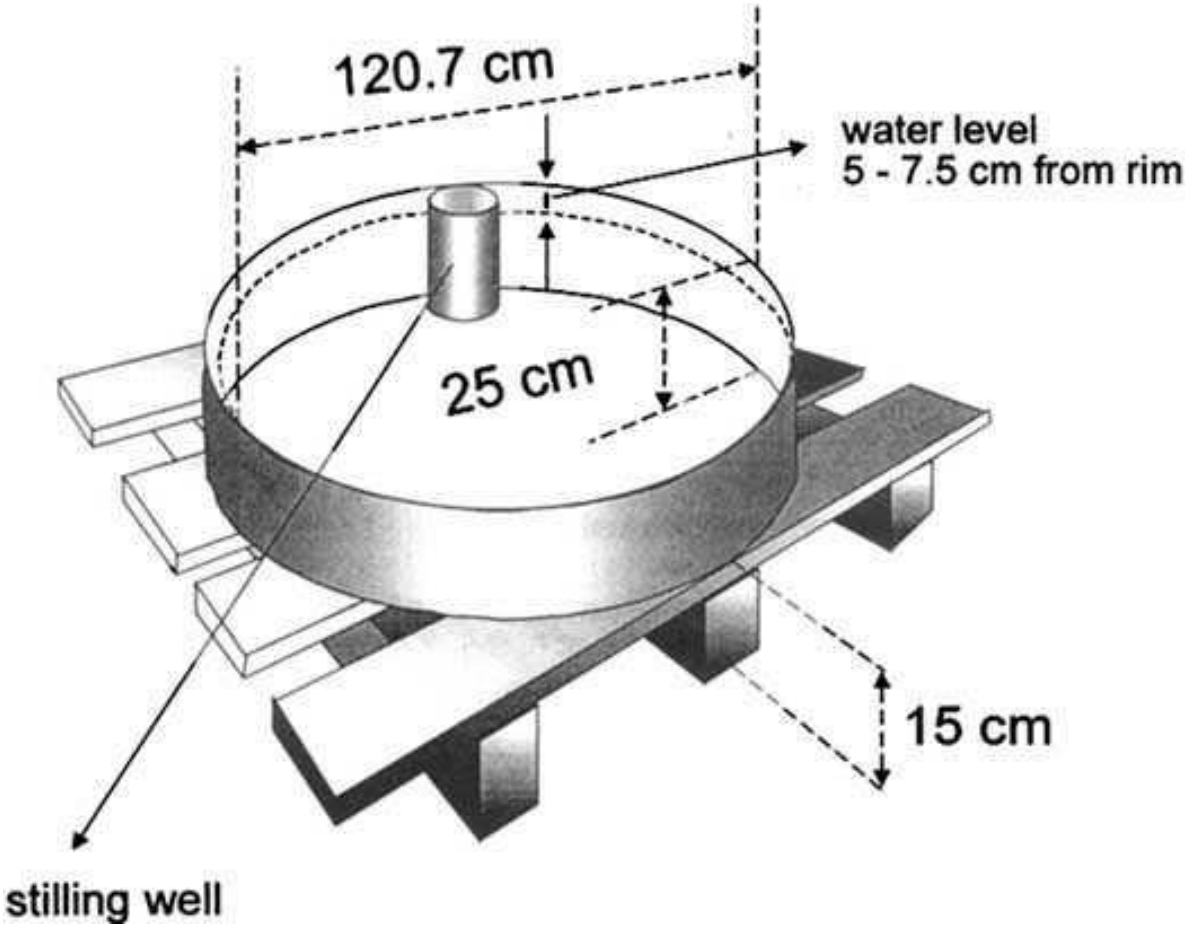
where K_p = pan coefficient (dimensionless) and E_{pan} = pan evaporation (mm day^{-1}).

Basic features as colour, size or position have a noticeable significance on the measurements taken resulting in specific pan coefficients for each type of evaporation pan. Surface coverage where the pan is located, its surroundings as well as the wind and the general moisture conditions are others determinant factors to have under consideration when selecting an appropriate pan coefficient (SHUTTLEWORTH; WALLACE, 1985, p. 855; SHUTTLEWORTH, 1993, p. 567). There are basically two main types of evaporation pans:

A) Class A pan: The Class A Evaporation pan is circular, 120.7 cm in diameter and 25 cm deep (Figure 1). It is made of galvanized iron (22 gauge) or Monel metal (0.8 mm). The pan is placed on a wooden open frame platform which is 15 cm above ground level. The pan must be level. It is filled with water to 5 cm below the rim, and the water level should not be allowed to drop to more than 7.5 cm below the rim. The water should be regularly renewed, at least weekly, to eliminate extreme turbidity. The pan, if galvanized, is painted annually with aluminium paint. Screens over the pan are

not a standard requirement and should preferably not be used. Pans should be protected by fences to keep animals away. The site should preferably be under grass, 20 by 20 m, open on all sides to permit free circulation of the air. It is preferable that stations to be located in the centre or on the leeward side of large cropped fields. Pan readings are taken daily in the early morning at the same time that precipitation is measured. Measurements are made in a stilling well that is situated in the pan near one edge. The stilling well is a metal cylinder of about 10 cm in diameter and some 20 cm deep with a small hole at the bottom (World Meteorological Organization, 1983; ALLEN *et al.*, 1991, p. 346).

Figure 1. Illustration of a standard design for a class A evaporation pan



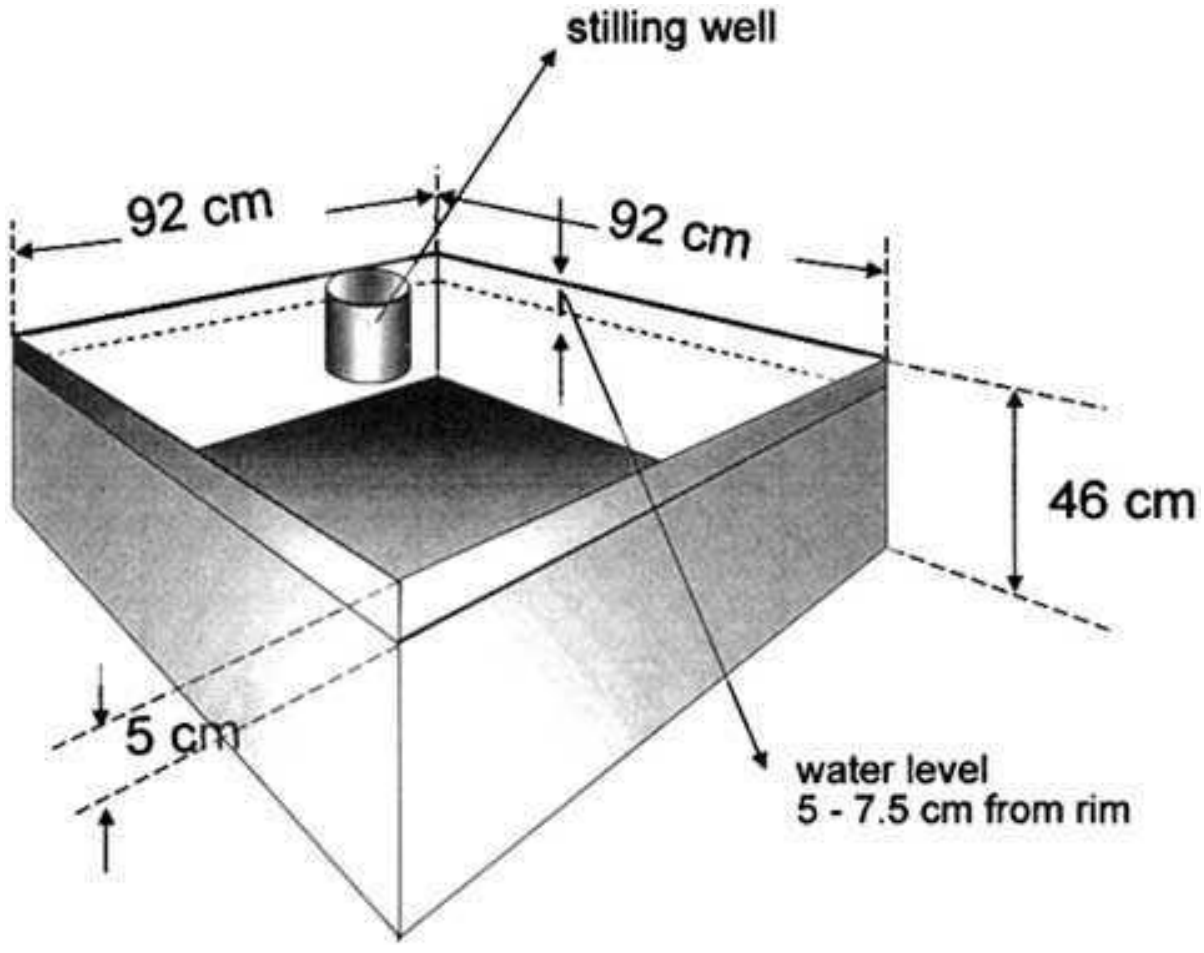
Source: FAO Irrigation and Drainage Paper. Crop Evapotranspiration (guidelines for computing crop water requirements

B) Colorado Sunken pan: The Colorado sunken pan is 92 cm (3 ft) square and 46 cm (18 in) deep, made of 3 mm thick iron, placed in the ground with the rim 5 cm (2 in) above the soil level (Figure 2). Also, the dimensions 1 m square and 0.5 m deep are

frequently used. The pan is painted with black tar paint. The water level is maintained at or slightly below ground level, i.e., 5-7.5 cm below the rim.

Measurements are taken similarly to those for the Class A pan. Siting and environment requirements are also similar to those for the Class A pan. Colorado Sunken pans are sometimes preferred in crop water requirements studies, as these pans give a better direct estimation of the reference evapotranspiration than does the Class A pan. The disadvantage is that maintenance is more difficult and leaks are not visible.

Figure 2. Illustration of a standard design for a Colorado Sunken pan



Source: FAO Irrigation and Drainage Paper. Crop Evapotranspiration (guidelines for computing crop water requirements)

6. ENERGY BALANCE APPROACH AND MICRO-CLIMATIC METHODS

Evapotranspiration is a combined process result of two processes, which run simultaneously and are activated by a common energy source: solar radiation. Indeed, *ET* is controlled by the interchange of energy on the crop surface and, at the same time, restricted by the amount of energy available (MONTEITH, 1965, p. 265; ALLEN *et al.*, 1989, p. 662; ALLEN *et al.*, 1998, p. 655). It is possible then, to think about *ET* in terms of energy and predict it by applying the conservation of energy principle. In other words, the quantity of income energy reaching the evaporative surface must be equal to the quantity of inbound energy leaving the system composed by the plant, the soil surface underneath and the surrounding environment in a given period of time. This circumstance can be illustrated by the following equation:

$$R_n = G + \lambda ET + H \quad \textbf{(5)}$$

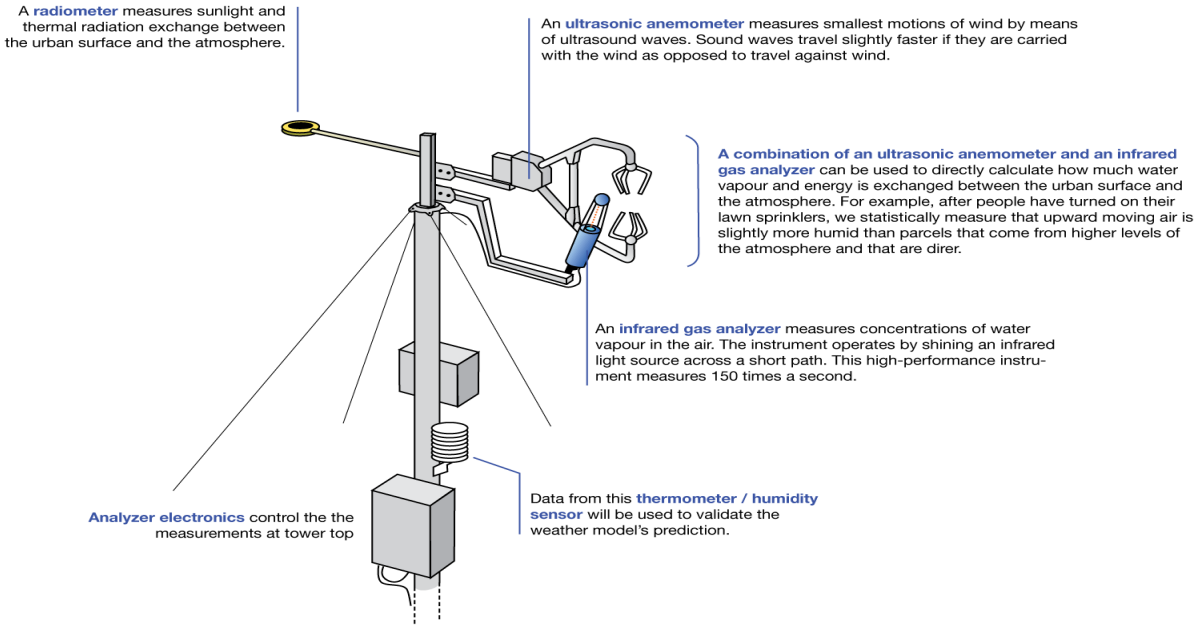
Where *Rn* represents the solar net radiation reaching the evaporative surface and appears isolated on the left side as showing it is the unique income energy source. On the right hand side the different exit routes of the system: *G* is the soil heat flux which accounts for the quantity of energy invested in warming up the soil surface underneath. λET, also called the latent heat flux, measures the portion of energy allocated in the *ET* process.

Finally, the sensible heat flux denoted as *H* accounts for the energy that heat up the surrounding air and it is doubtlessly the most complex parameter to be calculated since high accurate measurements of temperature gradients above the evaporative surface are essential (ALLEN, 1986, p. 368; ALLEN *et al.*, 1989, p. 662; ALLEN *et al.*, 1998, p. 156). Other processes run meanwhile and spend such an insignificant portions of energy that their "contributions" are negligible. Some considerations to be taken into account include that this approach does not take into consideration the energy rate transferred horizontally by advection, thus the equation is applied uniquely over large areas of homogeneous vegetation.

Another method of estimating evapotranspiration is the mass transfer method. This approach considers the vertical movement of small parcels of air (eddies) above a large

homogeneous surface. The eddies transport material (water vapour) and energy (heat, momentum) from and towards the evaporating surface (KIZER *et al.*, 1990, p. 181). By assuming steady state conditions and that the eddy transfer coefficients for water vapour are proportional to those for heat and momentum, the evapotranspiration rate can be computed from the vertical gradients of air temperature and water vapour via the Bowen ratio (BLAD; ROSENBERG, 1974, p. 236). Other direct measurement methods use gradients of wind speed and water vapour. These methods and other methods such as eddy covariance (Figure 3), require accurate measurement of vapour pressure, and air temperature or wind speed at different levels above the surface (KIZER *et al.*, 1990, p. 181). Therefore, their application is restricted to primarily research situations.

Figure 3. Illustration of standard flux tower 'Vancouver-Oakridge', equipped with an eddy-covariance system to directly measure the energy and water exchange.< http://www.geog.ubc.ca>



Source. Image taken from < http://www.geog.ubc.ca>

7. EVAPOTRANSPIRATION DETERMINED VIA METEOROLOGICAL DATASET

After the presentation of the different ways of measuring crop reference evapotranspiration (*ETo*) directly on the field, the great investment both in equipment, material and personnel, the difficulty in getting to the locations or the attention to be paid when applying the measuring protocols are some of the reasons why these field procedures have become restricted to primarily research area (ALLEN *et al.*, 1998, p. 245).

A more affordable alternative to these methods is the application of mathematical approaches based on several climate parameters, which are divided into empirical and physical models. Some of the methods are only valid under specific climatic and agronomic conditions and cannot be applied under conditions different from those under which they were originally developed (ALLEN, 2003, p. 163).

The former are based on statistical functions of approximation between meteorological parameters and *ETo* values (THORNTHWAITE, 1948, p. 94; BLANEY; CRIDDLE, 1950; JENSEN; HAISE, 1963, p. 41; HARGREAVES; SAMANI, 1985, p. 99). On the other hand, physical models are based on physical principles associated with the three most influential factors for *ET*: the amount of energy, the water vapor flux and the supply of moisture (CHOW *et al.*, 1988, p. 233). The most representative of these methods is the Penman combined process (PENMAN, 1948, p. 145). This approach relates evaporation dynamics to net radiation flux and aerodynamic transport characteristics of a natural surface. Observing that latent heat transference through plants is not only influenced by abiotic factors, Monteith introduced a surface conductance term that accounted for the response of leaf stomata to its hydrologic environment. This modified form of the Penman equation is widely known as the Penman-Monteith evapotranspiration model (MONTEITH, 1965, p. 366).

Aiming to assess the performance of these and other estimation procedures, a major study was undertaken within the framework of the Committee on Irrigation Water Requirements of the American Society of Civil Engineers (ASCE). The performance of 20 different methods was analyzed by comparing their results lysimeter data from 11 locations under variable climatic conditions. The results revealed a greatly uneven performance of the

methods mentioned above. The European Community, through a parallel study assessed the performance of various ETo methods using data from different lysimeter studies in Europe.

The studies confirm the relatively accurate and consistent performance of the Penman-Monteith approach in both arid and humid climates in both the ASCE and European studies. Despite that, these results revealed the need for formulating a standard method for the computation of *ETo*. Therefore, the final agreement of the Committee was that the Penman-Monteith equation, since then renamed Penman-Monteith FAO 56 method, became recommended as the sole standard method.

Many scientists proved the reliability of the Penman-Monteith FAO 56 (*PM56*) method for estimating *ETo* (MCNAUGHTON; JARVIS, 1984, p. 278; ALLEN, 1986, p. 368; ALLEN *et al.*, 1989, p. 662; SOUZA; YODER, 1994; CHIEW *et al.*, 1995, p. 21; KASHYAP; PANDA, 2001, p. 25; IRMAK *et al.*, 2003, p. 347; ITENFISU *et al.*, 2003, p. 448). Several studies demonstrated that *ETo* estimates obtained by *PM*56 model significantly resembles to observed *ETo* values (ALLEN *ET AL.*, 1994, P. 92; VENTURA *et al.*, 1999, p. 170; HOWELL *ET al.*, 2000, p. 242; WRIGHT *et al.*, 2000, p. 232).

PM 56 equation has basically two advantages: first, it is applicable to a wide range of climates and local conditions with no need to be calibrated; second, it is a method previously validated using lysimeters. It also has a drawback, though; *ET* is a complex and nonlinear phenomenon due to its dependency on several interconnected climatological parameters, such as air temperature, mean relative humidity, wind speed and insolation as sunshine hours. A very common problem among researchers on this field is that only temperatures have been broadly registered via weather stations since the fifties. In other words, historical meteorological data is usually incomplete. The *PM56* equation is defined as follows:

$$ET_o = \frac{0.408(R_n - G) + \gamma \frac{900}{T+273} U_2 (e_s - e_a)}{\Delta + \gamma(1 + 0.34U_2)} \quad \textbf{(6)}$$

where:

ETo : Reference Crop Evapotranspiration estimated by *PM* method (mm day^{-1}).

Δ: slope vapour pressure curve (kPa $^oC^{-1}$)

Rn: net radiation at the crop surface (MJ m^{-2} day^{-1})

G: soil heat lux (MJ m^{-2} day^{-1})

γ = psychrometric constant (kPa $^oC^{-1}$)

T: mean daily air temperature (^oC)

U_2: wind speed at 2 meters height (m s^{-1})

e_s: saturation vapour pressure (kPa) and

e_a: actual vapour pressure (kPa).

ETo provides a standard comparative method whereby it is possible to compare *ET* values gathered in different places, crop types, growing periods or weather conditions. The equation is fed with climatological data of solar radiation, air temperature, air relative humidity and wind speed. The compilation of the dataset must be accomplished at 2 meters above the soil surface, over "an extensive surface of green grass of uniform height, actively growing, completely shading the ground and with adequate water" (DOORENBOS; PRUITT, 1977, p. 423). Even so, any empirical equation based on climatic parameters is able to produce a 100% reliable result due to the equation simplicity itself or to errors (discrete or systematic) on the gathering of the climate measurements. Despite that, the panel of experts agreed on using the definition of hypothetical reference crop surface, over which Penman-Monteith FAO 56 equation is based on, as an homogenized comparison value so as to compare estimated values from different locations.

The Penman-Monteith FAO 56 equation is a simple representation of the physical and physiological factors governing the evapotranspiration process. By using the FAO Penman-Monteith definition for *ETo* it is possible to calculate crop coefficients at research sites by relating the measured crop evapotranspiration (ET_c) with the calculated ET_o, i.e., $K_c = ET_c/ET_o$. In the crop coefficient approach, differences in the crop canopy and aerodynamic resistance relative to the hypothetical reference crop are accounted for within the crop coefficient.

8. SOIL WATER BALANCE APPROACH

Back to the energy balance, *ET* is roughly defined as the result of subtracting the "energy costs" from the unique income, radiation. Similarly, *ET* can be also thought in terms of water; in this case, the inbound and outbound channels are not about energy, but matter; water to be precise (PERRIER *et al.*, 1976, p. 84; PERRIER; TUZET, 1991, p. 155; BASTIAANSSEN, 1995B, p. 512). The method consists in assessing the entering and leaving water fluxes from and towards the rooting system in a given period of time. In the following, the equation that illustrates this process:

$$Pp + I + C = R_{off} + Pc + ET + \Delta FS \qquad \textbf{(7)}$$

Where precipitation (*Pp*), irrigation (*I*) and capillarity (*C*) represent the water inputs to the system, while run-off (R_{off}), percolation (*Pc*), evapotranspiration (*ET*) and the difference between the inbound and outbound sub-superficial water flux (*ΔFS*). Precipitation and irrigation provide the required amount of water to the root zone, although not all the water is used, a meaningful portion is lost by means of run-off and percolation which will eventually recharge the table water. Water also can be conveyed from the phreatic layer upwards to the surface via capillarity or even horizontally transferred through sub-superficial flux inwards or outwards the root zone (DIAZ-ESPEJO; VERHOEF, 2003, p. 623).

The net quantity gain or lost is illustrated on the equation as *ΔFS*. Nevertheless, apart from greatly steep slope circumstances, *ΔFS* is usually not taken under consideration. Measure sub-superficial flux, deep percolation and capillarity are complex to determine and to measure them it is generally necessary to invest in costly equipment and high-trained researchers (TANNER; PELTON, 1960, p. 413). Moreover, this processes take long time to finish off, so their measurements have become restricted to medium and long-term studies in the las decade.

9. COMPUTING OF METEOROLOGICAL PARAMETERS

To apply Penman-Monteith FAO 56 equation, it is necessary to arrange some measures of the climatic parameters described in early sections. Those parameters are responsible for suppling the amount of energy required to vaporize the soil water, maintaining such a temperature and relative humidity on the surrounding environment that would favour the stomatal opening and the transpiration process and a removing power of the vapour accumulated right above the evaporative surface (ALLEN *et al.*, 1998, p. 248). As a reminder, we are referring to radiation, air temperature, air relative humidity and wind speed. Let us present them individually so as to better understand how they can be computed or to what extent they can influence on the reference crop evapotranspiration estimation.

9.1 Radiation

As mentioned above, radiation is the only inbound source of energy on the Earth; everything depends on it, so it represents the driving force of the rest of the physical processes. Radiation can be measured by using piranometers, radiometers, solarimeters or even heliographs.

Radiation supplies the energy required to hold entire trophic chains, starting from the photosynthetic producers which transform this solar energy into biochemical energy available for the rest of the trophic levels. Radiation is also the responsible for heating up the oceans in such irregular way that allow to distinguish streams which convey nutrients and heat to different locations in a non-stop circulation known as the “Global Ocean Conveyor Belt”. In spite of exerting as the driving force for thousands of physical processes which are not possible to describe here, not all the radiation generated by the Sun reaches the Earth surface. Indeed,

the amount of radiation reaching the top of the atmosphere depends on the angle of incidence between the solar rays and the atmosphere surface, it is called as solar constant and has an approximate value of 0.082 MJ m^{-2} min^{-1} (DOORENBOS; PRUITT, 1977, p. 561). This value varies along the year, varies depending on the latitude and even along the day and is called as extraterrestrial radiation denoted as *Ra* and calculated as follows:

$$R_a = \frac{24*60}{\pi} G_{sc} d_r [\omega_s \sin(\varphi) \sin(\delta) + \cos(\varphi) \cos(\delta) \sin(\omega_s)] \quad \textbf{(8)}$$

where:

Ra: extraterrestrial radiation (MJ m^{-2} min^{-1})

G_{sc}: Solar constant = 0.082 MJ m^{-2} min^{-1}

d_r: inverse relative distance Earth-Sun (dimensionless)

ω_s: solar incidence angle at dawn (radians)

ϕ: latitude (radians)

δ: solar declination (radians)

The inverse relative distance Earth-Sun (*dr*) and the solar declination are defined as:

$$d_r = 1 + 0.033 * \cos\left(\frac{2\pi}{365} J\right) \quad \textbf{(9)}$$

$$\delta = 0.409 * sen\left(\frac{2\pi}{365} J - 1.39\right) \quad \textbf{(10)}$$

where *J* is the Julian day of the year

The solar incidence angle at dawn, ω_s, is given by:

$$\omega_s = ar\cos[-\tan(\varphi)\tan(\delta)] \quad \textbf{(11)}$$

Having the quantity of radiation reaching the top of the atmosphere, it is possible to calculate the maximum daylight hours (*N*) as:

$$N = \frac{24}{\pi}\omega_s \quad \textbf{(12)}$$

Now we are in conditions of calculating the solar radiation (*Rs*), which accounts for the portion of extraterrestrial radiation passing through the atmosphere and reaching an horizontal plane. As the radiation penetrates the atmosphere, some of the radiation is scattered, reflected or absorbed by the atmospheric gases, clouds and dust. As the sun emits energy, by means of electromagnetic waves characterized by short wavelengths, *Rs* is also referred to as shortwave radiation. R_s may range from 75% of *Ra* in a clear day to 25% of *Ra* in a cloudy day (JENSEN *et al.*, 1990, p. 970; IRMAK *et al.*, 2003, p. 370). *Rs* is also known as global radiation as pointing out the sum of direct shortwave solar radiation diffuse radiation from other reflective surfaces. Here, two new factors appear in the equation so as to represent that portion of extraterrestrial radiation reaching the ground:

$$R_S = \left(a_S + b_S \frac{n}{N}\right) R_a \quad \textbf{(13)}$$

where:

Rs: Solar radiation (MJ m^{-2} dia^{-1})

n : actual sunshine hours (hours)

N: maximum sunshine hours (hours)

a_s: regression coefficient. Accounts for the portion of the extraterrestrial radiation reaching the ground in cloudy days ($n = 0$)

b_s: portion of the extraterrestrial radiation reaching the ground in clear days ($n = N$)

The so-called term relative sunshine duration, denoted as *n/N*, expresses the atmospheric cloudiness and also depends on the latitude and the date. Not being in power of measurements of solar radiation, a_s and b_s are recommended to take values of 0.25 and 0.50, namely. In these cases, clear-sky solar radiation, R_{so} (when $n = N$) can be computed as follows:

$$R_{SO} = \left(0.75 + 2 * 10^{-5} z\right) R_a \quad \textbf{(14)}$$

where z is the altitude of the meteorological station above sea (m).

From the solar radiation (*Rs*) reaching the ground, part of it is reflected as effect of the albedo (α) and the rest is computed as Net Income Shortwave Radiation (*Rns*) as in the following:

$$R_{nS} = (1 - \alpha) R_S \quad \textbf{(15)}$$

Where *Rns* and *Rs* are calculated as MJ m^{-2} dia^{-1}. Albedo (α) is dimensionless and widely varies with respect to the type of vegetal cover, incidence angle of solar rays or surface slope. Its value may range from 0.95 for a fresh snowed surface to 0.05 for a humid and bare piece of ground. A regular green cover usually presents a value comprehend between 0.20 and 0.25. In fact, the hypothetical reference crop surface, α has a value of 0.23 (ALLEN *et al.*, 1989, p. 662; ALLEN *et al.*, 1998, p. 249).

The radiation absorbed by the Earth is transformed into thermal energy and emitted by it as longwave emission. This longwave radiation (*Rnl*) is responsible for heating the air temperature up and hence maintaining the habitability conditions on the Earth. Part of it is reflected by the atmosphere, so we also have longwave radiation reaching the ground. That is to say that the Earth both emits and absorbers longwave radiation. *Rnl* is the net difference between these inbound and outbound fluxes and is expressed quantitatively by the *Stefan-Boltzmann law* as follows:

$$R_{nl} = \sigma \left[\frac{T_{\max,K}^4 + T_{\min,K}^4}{2} \right] \left(0.34 - 0.14\sqrt{e_a}\right) \left(1.35 \frac{R_S}{R_{SO}} - 0.35\right) \quad \textbf{(16)}$$

where:

R_{nl}: net outgoing longwave radiation (MJ m^{-2} day^{-1}),

σ: *Stefan-Boltzmann* constant = 4.903*10-9 (MJ K^{-1} m^{-2} day^{-1}),

$T_{max, K}$: maximum absolute temperature during the 24-hour period (K = °C + 273.16),

$T_{min, K}$: minimum absolute temperature during the 24-hour period (K = °C + 273.16),

e_a: actual vapour pressure (kPa),

R_s: measured or calculated solar radiation (MJ m^{-2} day^{-1}),

R_{so}: calculated clear-sky radiation (MJ m^{-2} day^{-1}).

The term (0.34-0.14$\sqrt{\ e_a}$) expresses the correction for air humidity, and will be smaller as the humidity increases. The effect of cloudiness is expressed by (1.35 R_s/R_{so} - 0.35). The term becomes smaller if the cloudiness increases and hence R_s decreases. The smaller the correction terms, the smaller the net outgoing flux of longwave radiation (ALLEN *et al.*, 1989, p. 662; JENSEN *et al.*, 1990, p. 970; ALLEN *et al.*, 1998, p. 249). Note that the R_s/R_{so} term will always be within 0 and 1. Indeed, *Rs/Rso* commonly varies between 0.33 and 1. Eventually, Net Radiation (*Rn*) is computed as the difference between short and long wave radiation represented by such a simple equation:

$$R_n = R_{nS} - R_{nl} \quad \textbf{(17)}$$

Rn is usually positive along daytime, negative during night and positive for 24-hour periods, excepting conditions of extraordinary high altitudes. Part of the inbound radiation

reaching the ground is invested in heating the soil up in a process which spend little portion of radiation. Heat soil flux, as is called and denoted as *G*, tend to be zero if it is computed daily due to the gains (during daylight hours) and losses (at night) along the day which cancel out to each other, so $Gday \approx 0$ (BASTIAANSSEN, 1995B, p. 653).

If *G* has to be calculated monthly, the equation completely changes to:

$$G_m = 0.07(T_{m+1} - T_{m-1}) \text{ or } G_m = 0.14(T_m - T_{m-1}) \quad \textbf{(18a, 18b)}$$

where:

Tm: mean air temperature on month m,

Tm+1: mean air temperature on the following month,

Tm-1: mean air temperature on the previous month

If *G* has to be calculated hourly,

$$G_{hr} = 0.1R_n \text{ or } G_{hr} = 0.5R_n \quad \textbf{(19a, 19b)}$$

Depending on whether G_{hr} is computed for daylight or nocturnal periods, respectively.

9.2 Air temperature

Radiation is absorbed both by the atmosphere and by the ground increasing the air temperature. This sensible heat of the surrounding environment transfers energy to the crop exerting such a controlling effect on the evapotranspiration rate. For instance, the water loss due to evapotranspiration would be higher in a sunny, warm day than in a cloudy, cool day. Temperature is measured inside shelters (Stevenson screens or ventilated radiation shields) placed in line with World Meteorological Organization standards at 2 m above the ground (World Meteorological Organization, 1983).

Temperature is commonly measured by means of thermometers, thermocouples fitted inside the shelters that register maximum and minimum values of temperature along a 24-hours period. Termographs plot the instantaneous temperatures during the sampling period. Automatic weather stations commonly sample the temperature every minute and report hourly averages in addition to maximum and minimum values over the 24-hours period. Temperature can be expressed both as Celsius (ºC) and Fahrenheit degrees (ºF).

9.1 Relative Air Humidity

It refers to the water content in the air. It can be expressed in different ways, such as vapor pressure, dew point temperature or the most common, air relative humidity. Air relative humidity (*RH*) expresses the air saturation degree as a function of the vapor actual pressure (e_a) in a given temperature and the vapor saturation pressure ($e^o(T)$) under such temperature and can be defined as follows:

$$HR = \frac{e_a}{e^o(T)} * 100 \qquad \textbf{(20)}$$

where:

HR: relative humidity (%),

e_a: vapor actual pressure (kPa),

$e^o(T)$: vapor saturation pressure (kPa).

In other words, relative humidity is a quotient of the water content the air actually contains in a certain temperature and the content it could be contain if the air would be saturated at such temperature (MAKKINK, 1957, p. 288; HARGREAVES E SAMANI, 1985, p. 99) . Although the vapor actual pressure (e_a) may be relative constant along the day, the relative humidity varies from a maximum at dawn and a minimum when the incidence angle of the solar rays is near zero.

As the vapor saturation pressure depends on the air temperature, which changes along the day exerting the same effect on the former. It is not possible to measure directly the vapour actual pressure, but we can measure it indirectly by using psycrometers. In that case, the vapour pressure can be computed from the difference between two thermometers, the so-called dry and wet bulb thermometers. Relative humidity measurements often carry errors from diverse origins. If we are not 100% sure of the reliability of the registered values, it is advisable to compare them with dew point temperatures or minimum air temperature. It is also possible to compute *RH* values from dew point temperatures, whereas the equipment required and its maintenance discourage users to do so.

Due to the non-linear relation between temperature and air humidity, both included within the Penman-Monteith equation, vapor pressure for some period must be calculated as the mean between the vapor pressure for the maximum air temperature and the vapor pressure for the minimum air temperature for the given period as follows:

$$e_s = \frac{e^o(T\max) + e^o(T\min)}{2} \qquad \textbf{(21)}$$

where:

e_s: mean saturation vapor pressure (kPa),

$e^o(Tmax)$: saturation vapor pressure at the maximum air temperature (kPa)

$e^o(Tmin)$: saturation vapor pressure at the minimum air temperature (kPa)

Saturation vapor pressure can be computed in function of the air temperature in the following way:

$$e^o(T_{max}) = 0.6108 * \exp\left(\frac{17.27 * T_{max}}{T_{max} + 237.3}\right) \quad \textbf{(22)}$$

As saturation vapor pressure at maximum temperature and

$$e^o(T_{min}) = 0.6108 * \exp\left(\frac{17.27 * T_{min}}{T_{min} + 237.3}\right) \quad \textbf{(23)}$$

As saturation vapor pressure at minimum temperature.

The use of the mean temperature (*Tmean*) instead of maximum and minimum temperatures lead us to underestimations of the mean saturation vapor pressure (MONTEITH, 1965, p. 654; MCNAUGHTON; JARVIS, 1984, p. 278). The corresponding vapor pressure coefficient, a parameter which expresses the atmosphere evaporating energy, will also be lesser and the result will be some underestimation of the evapotranspiration rate of the reference crop.

Conversely, to calculate the *Slope of saturation vapor* pressure curve (*Δ*) the FAO recommends to use the *Tmean* of the air instead of calculating a *ΔTmax* for the *Tmax* and a *ΔTmin* for the *Tmin (*SMITH *et al.,* 1991*;* ALLEN *et al.,* 1998, p. 156*).* Thus, the equation of the *Slope of saturation vapor* pressure curve (*Δ*) is described as in the following:

$$\Delta = \frac{4098 * \left[0.6108 * \exp\left(\frac{17.27 * T_{mean}}{T_{mean} + 237.3} \right) \right]}{(T_{mean} + 237.3)^2} \quad \textbf{(24)}$$

where:

Δ: slope of saturation vapor pressure curve at air temperature *Tmean* (kPa $^{o}C^{-1}$), *exp[..]:* 2.7183 (base of natural logarithm) raised to a given power and

$$T_{mean} = \frac{T_{\max} + T_{\min}}{2} \quad \textbf{(25)}$$

Actual vapor pressure (e_a) can be computed in different ways depending on what parameters we have. In our case, we did have neither dew point temperature data nor psycrometric data so we used the third way: using mean air relative humidity (*RHm*) in the following equation:

$$e_a = \frac{RH_m}{100} e_S \quad (26)$$

where:

$$e_S = e^{o}(T_{mean}) = 0.6108 * \exp\left(\frac{17.27 * T_{mean}}{T_{mean} + 237.3} \right) \quad \textbf{(27)}$$

The difference between e_a and e_s is known as vapor pressure deficit and is determined simply by subtracting e_a from e_s: (e_s-e_a).

9.2 Wind Speed

For computing the evapotranspiration rate, wind speed plays an important role since acts as the driving force which removes the water vapor from the evaporative surface avoiding the over-saturation of the surrounding atmosphere. It is obvious that the key features for wind are its direction and speed. The later must be referred both to a length and time unit, mostly meters divided by seconds (m s^{-1}).

Wind speed can be measured through anemometers. Its measurement requires special attention since it varies with the height whereby measurements are taken. Near the surface tends to be low and increases with the height due, in part, to the existence of a vegetal cover right above the surface. For the sake of it the standard height to install anemometers is 2 meters above the ground.

9.3 Other essential atmospheric parameters for the estimation of *ETo*

Along this study we make clear that the driving parameters governing the process of evapotranspiration are basically four: radiation, air temperature, air relative humidity and wind speed. Having said that, the Penman-Monteith equation (ALLEN *et al.*, 1989, p. 662; JENSEN *et al.*, 1990, p. 970; ALLEN *et al.*, 1998, p. 249) is fed with some other parameters apart from the ones cited above. They are the atmospheric pressure and the psychrometric constant. Atmospheric pressure is the pressure exerted by the atmosphere weight. The evaporation at high locations is in part due to the low pressure on those locations. So, the higher the location is, the lower the atmospheric pressure. Its dependency is expressed in the following equation:

$$P = 101.3\left(\frac{293 - 0.0065z}{293}\right)^{5.26} \tag{28}$$

where:

P: atmospheric pressure (kPa),

Z: altitude (m).

On the other hand, the psychrometric constant is computed as follows:

$$\gamma = \frac{c_p * P}{\varepsilon \lambda} = 0.665 * 10^{-3} * P \quad \textbf{(29)}$$

where:

γ: psychrometric constant (kPa °C^{-1})

λ: latent heat of vaporization = 2.45 (MJ kg^{-1}),

$C_{p:}$ specific heat at constant pressure = 1.013 10^{-3} (MJ kg^{-1} °C^{-1}),

ε: ratio molecular weight of water vapour / dry air = 0.622.

The procedure here described corresponds to the computing of the *ETo* on a daily basis. That is to say that the estimation process starts from gathering daily measurements of the driving parameters from the weather stations comprehend within the study area. Next, we apply of all the previous set of equations aiming to compute the corresponding daily *ETo* within a given period. Nevertheless, *ETo* can be also computed hourly, ten-day, monthly or annually.

FAO recommends estimate *ETo* daily or monthly, as hourly estimates could present systematic errors and entail overestimation of the *ETo* (ALLEN *et al.*, 1989, p. 662; ALLEN *et al.*, 1998, p. 249; WRIGHT *et al.*, 2000, p. 165). In the case of the estimation of annual *ETo* values, other parameters should be taking into account, such as dawn timing, which varies with the location and the season.

Soil heat flux is another parameter that changes the way it is considered within an hourly/monthly approach; in fact, (*G*) no longer tends to zero since the considered period is not anymore a whole day, but one hour or one complete month so it requires to be computed.

Having said that, when the soil is warming up (spring) or cooling down (autumn), the soil heat flux for monthly periods may become significantly relative to the mean monthly R_n (ITENFISU *et al.*, 2003, p. 448). In these cases, *G* cannot be ignored and its value should be determined from the mean monthly air temperatures of the previous and next month. Similarly, soil heat flux present very different values on late/early hours; when the sun sets the temperature drops and the incidence angle of the solar rays is such small that the radiation reduces its value in a big way. On the other hand, midday hours, when the solar rays present an incidence angle of approximately 90º, the soil heat flux reaches the maximum daily values. Hourly estimations of *ETo* must include these soil heat flux variations along a 24-hours period and make a clear difference between daylight hours and darkness hours.

Conversely, the computation of hourly *ETo* estimates would be more appropriate in locations where meaningful variations may occur along a 24-hours period. Such variations comprehend changes in wind speed (coastal areas, for instance), dew point temperature (microclimatical areas) or cloudiness (mountain areas). Noticeable changes in any of these climatological factors can carry erroneous 24-hours-period averages which do not represent the true evaporative power of the environment (CAPRIO, 1974, p. 364). Hourly estimation of *ETo* using the Penman-Monteith FAO 56 equation requires to introduce some adjustments, although thorough this study we estimate *ETo* on a daily basis, which equation has already been described in previous paragraphs.

10. ALTERNATIVES TO PENMAN-MONTEITH FAO 56 EQUATION TO ESTIMATE *ETo*

10.1 Temperature-based equations

Hargreaves & Samani equation (HARGREAVES E SAMANI, 1985, p. 99; ALLEN *et al.*, 1998, p. 249): $PET_{Har} = 0.0023(Ra/\lambda)\sqrt{(T_{\max} - T_{\min})} * (T_{mean} + 17.8)$ **(30)**

Thornthwaite's equation (THORNTHWAITE, 1948, p. 94):

$$PET_{Th} = 1.6(N_{di} * N_i/360) * (10T_{mean}/I)^a \quad \textbf{(31)}$$

McGuinness-Bordne's equation (OUDIN *et al.*, 2005, p. 306):

$$PET_{MGB} = (Ra/\lambda * \rho) * (T_{mean} + 5/68) \quad \textbf{(32)}$$

Romanenko's equation (OUDIN *et al.*, 2005, p. 306)

$$PET_{Rom} = 4.5 * [1 + (T_{mean}/25)]^2 * (1 - e_a/e_S) \quad \textbf{(33)}$$

Hammon's equation (LU *et al.*, 2005, p. 421):

$$PET_{Ham} = 0.1651(\mathrm{DL}/12) * 1.2[216.7e_S/(T_{mean} + 273.3)] \quad \textbf{(34)}$$

Hammon's equation (OUDIN *et al.*, 2005, p. 306):

$$PET_{Ham} = (DL/12)^2 * \exp(T_{mean}/16) \quad \textbf{(35)}$$

Hammon's equation (ROSENBERRY *et al.*, 2004, p. 497):

$$PET_{Ham} = 0.55\left(DL/12\right)^2 * 25.4\left[SVD/100\right] \quad \textbf{(36)}$$

McCloud's equation (JACOBS *et al.*, 2001, p. 245; IRMAK *et al.*, 2003, p. 370):

$$PET_{McC} = 1.07 + 0.01^{(T_{mean}-32)} \quad \textbf{(37)}$$

where:

Table 1.Description of the variables appeared on the empirical temperature-based equations

Symbol	Definition	Unit	Calculation
R_a	extraterrestrial radiation	MJ m^{-2}	$R_a = \frac{24*60}{\pi} G_{SC} * d_r \left[\begin{matrix}\omega_s \sin(\varphi)\sin(\delta) + \\ \cos(\varphi)\cos(\delta)\sin(\omega_s)\end{matrix}\right]$
λ	latent heat of evaporization	MJ kg^{-1}	$\lambda = 2.501 - \left(2.361*10^{-3}\right)*T_{mean}$
T_{max}	maximum temperature	°C	Measured values
T_{min}	minimum temperature	°C	Measured values
T_{mean}	mean temperature	°C	Measured values
I	Annual calorific efficiency index	dimensionless	$I = \left(T_{month}/5\right)^{1.51}$; $I = \sum_{j=1}^{12} i_j$
a	constant	dimensionless	$a = 67.5*10^{-8} I^3 - 77.1*10^{-6} I^2 + 0.0179 I + 0.492$
N_{di}	number of the considered days	days	Measured values
N_i	astronomical duration of the day	hours	Measured values
ρ	water density	1000(kg m^{-3})	Measured values
e_a	actual vapour pressure	KPa	$e_a = 0.6108 * \exp\left(\frac{17.27 * T_{dew}}{T_{dew} + 237.3}\right)$
e_s	mean saturation vapour pressure	KPa	$e_S = \frac{e^o(T\max) + e^o(T\min)}{2}$
DL	day length	hours	Measured values
SVD	saturated vapour density at mean temperature	g m^{-3}	$SVD = 216.7\left[e_S/\left(T_{mean} + 273.3\right)\right]$

Source: A wide variety of articles related to the evolution of empirical equations to determine ETo. Gathered by the author form diverse sources

10.2 Radiation-based equations

Penman FAO24 equation (DOORENBOS; PRUITT, 1977, p. 564):

$$PET_{FAO24} = C\left[\left(\Delta/\Delta+\gamma\right)\left(R_n/G\right)+\left(\gamma/\Delta+\gamma\right)*2.7\left(1+0.86u_2\right)\left(e_S-e_a\right)\right] \quad \textbf{(38)}$$

Blaney-Criddle's equation (DOORENBOS; PRUITT, 1977, p. 564):

$$PET_{B-C} = \left(0.0043RH_{\min}-n/N-1.41\right)+b_b\left[p\left(0.46T_{mean}+8.13\right)\right] \quad \textbf{(39)}$$

Radiation equation:

$$PET_{Rad} = b_R\left(R_S\left(\Delta/\Delta+\gamma\right)\right)-0.3 \quad \textbf{(40)}$$

Prestley-Taylor's equation:

$$PET_{P-T} = 1.26\left(\Delta/\Delta+\gamma\right)R_n \quad \textbf{(41)}$$

Hansen's equation (HANSEN, 1984, p. 212):

$$PET_{Han} = 0.7\left(\Delta/\Delta+\gamma\right)*R_s/\lambda \quad \textbf{(42)}$$

Caprio's equation (CAPRIO, 1974, p. 364):

$$PET_{Cap} = 6.1*10^{-6}*R_S*(1.8*T_{mean}+1) \quad \textbf{(43)}$$

Jensen and Haise equation (JENSEN; HAISE, 1963, p. 41):

$$PET_{J\&H} = \left(Rs/\lambda\right)*0.025*T_{mean}+0.08 \quad \textbf{(44)}$$

Turc's equation (TURC, 1961, p. 49):

For RH < 50%:

$$PET_{Turc} = 0.013 * \left(T_{mean} / T_{mean} + 15\right) * \left(R_S * 23.8846 + 50\right) * \left(1 + \left(50 - RH\right)/70\right) \quad \textbf{(45)}$$

For RH > 50%:

$$PET_{Turc} = 0.013 * \left(T_{mean} / T_{mean} + 15\right) * \left(R_S * 23.8846 + 50\right) \quad \textbf{(46)}$$

Makkink's equation:

$$PET_{Mak} = \left[0.061\Delta/\Delta + \lambda * R_s / \lambda - 0.012\right] * 10 \quad \textbf{(47)}$$

De Bruin's equation:

$$PET_{DBr} = \left\{R_s / \lambda \left[\Delta/\left(0.85\Delta + 0.63\gamma\right)\right]\right\} * 10 \quad \textbf{(48)}$$

Table 2.Description of the variables appeared on the empirical radiation-based equations

Symbol	Definition	Unit	Calculation
C	adjustment factor	dimensionless	Review KOTSOPOULOS & BABAJIMOPOULOS, (1997, p. 451)
Δ	Slope of saturation vapour pressure curve	KPa ºC^{-1}	$\Delta = \frac{4098 * \left[0.6108 * \exp\left(\frac{17.27 * T_{mean}}{T_{mean} + 237.3}\right)\right]}{(T_{mean} + 237.3)^2}$
Rn	Net radiation	MJ m^{-2}	$R_n = R_{nS} - R_{nl}$
G	Soil heat flux	MJ m^{-2}	Depending on the time lapse under study
γ	Psychrometric constant	KPa ºC^{-1}	Review ALLEN et al., 1998, p. 249 $\gamma = 0.665 * 10^{-3} * P_{atm}$
U_2	Wind speed	M seg^{-1}	Measured values
RHmin	Minimum relative air humidity	%	Measured values
n/N	Relative sunlight duration	dimensionless	n: Measured values $N = 24/\pi * w_s$

b_b	Adjustment facrtor		$b_b = 0.81917 - 0.0040922\,RH_{min} + 1.0705\,n/N + 0.065649u_2 - 0.0059684\,RH_{min} * n/N - 0.0005967\,RH * n/N$
p	mean daily percentage of daytime hours a year	%	Measured values
br	Adjustment factor		$b_r = 1.0656 - 0.0012795RH_{mean} + 0.044953U_2 - 0.0002003RH_{mean} * U_2 - 0.00031508RH_{mean}^2 - 0.001102U_2^2$
Rs	Solar radiation	MJ m^{-2}	$R_S = \left(a_S + b_S \frac{n}{N}\right)R$ Measured values
RH	Relative air humidity	%	$RH = \frac{e_a}{e^o(T)} * 100$

Source: A wide variety of articles related to the evolution of empirical equations to determine ETo. Gathered by the author form diverse sources

11. STUDY AREA AND DATA SET

The data set used throughout this research was composed by a total of 26 meteorological stations belonging to the Brazilian Meteorological National Institute (INMET) and sparsed within and around Paraíba do Sul watershed <http://sigaceivap.org.br/siga-ceivap/map>. The complete set of weather stations was split into 4 groups: data clusters 1 and 2, composed by conventional weather stations, encompassed a 31-year period, ranging from January 1st, 1985 to December 31st, 2015 totalizing 11322 measures of daily meteorological data. Data clusters 3 and 4 included automated weather stations, whose instalments occurred from 2000 onward. That is to say that stations belonging to clusters 3 and 4 show different periods of study, ranging from May 5th, 2000 to September 9th, 2016, as shown in Table 3. Aside from that, it is important to pinpoint that for clusters 3 and 4 the measures were registered hourly; further computations turn them into daily measures.

The historical registers embraced 5 central climatological parameters to compute *ETo: maximum temperature (Tmax), minimum temperature (Tmin),* both temperatures measured in Celsius degrees (°C), *Global Solar Radiation (Rs)* measured in MJ m^{-2} dia^{-1}, *relative air humidity (RH)*, represented as a % and *wind speed (U_2)*, measured in m s^{-1} at a 10 meters high. However, U_2 would require further computations applied following ALLEN *et al.*, (1998, p. 249) as it needs to be expressed in m s^{-1} at a 2 meters high.

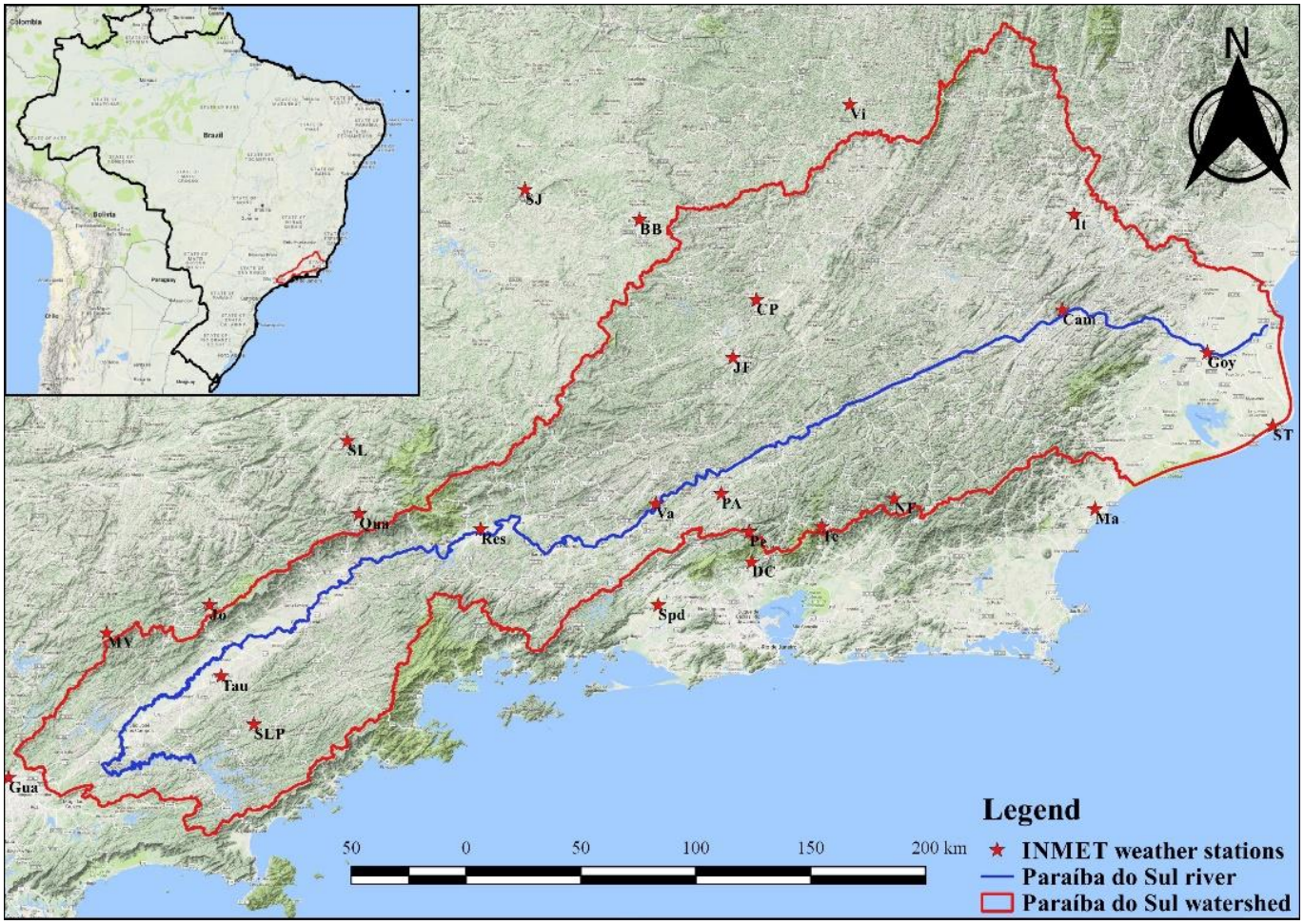
Located in the South-East of Brazil, Paraíba do Sul watershed is one of the most significant river basins in Brazil in terms of urban water supply. Water is mostly used for domestic and industrial purposes but also for irrigation and hydropower generation. It is transboundary as it involves several states, such as Minas Gerais, Rio de Janeiro and São Paulo. Climate of the study area has been classified as Humid Subtropical (Cwa) according to Köppen climate classification (GEIGER, 1961; KÖPPEN, 1884, 1918, 1936, 2011) and depicted in Figure 4.

Table 3. Distinctive features of every INMET weather station that take part into the present study, clustered into four clusters

Data		INMET weather stations		altitude (m)	Latitude (degrees)	Longitude (degrees)	WMO code	Start date	Finish date	% of available measurements
Conventional stations	**Data cluster 1**	Barbacena	BB	1126	-21.22	-43.76	83689	01/01/1985	31/12/2015	7894/11322
		Coronel Pacheco	CP	435	-21.54	-43.25	86822	01/01/1985	31/12/2015	7294/11322
		Juiz de Fora	JF	939.96	-21.76	-43.35	83692	01/01/1985	31/12/2015	8977/11322
		Paty de Alferes	PA	507	-22.35	-43.41	83049	01/01/1985	31/12/2015	6703/11322
		Resende	Res	439.89	-22.45	-43.43	83738	01/01/1985	31/12/2015	6638/11322
		Lavras	Lv	918.84	-21.75	-45	83687	01/01/1985	31/12/2015	9118/11322
		São Lourenço	SL	953.2	-22.10	-45.01	83736	01/01/1985	31/12/2015	9491/11322
		Viçosa	Vi	712.20	-20.77	-42.85	83642	01/01/1985	31/12/2015	8589/11322
	Data Cluster 2	Campos Goytacazes	Goy	17	-21.75	-41.33	83698	01/01/1985	31/12/2015	2498/11322
		Itaperuna	Ita	123.59	-21.2	-41.9	83695	01/01/1985	31/12/2015	6491/11322
		Campos Jordão	Jor	1672	-22.75	-45.6	83714	01/01/1985	31/12/2015	3046/11322
		Guarulhos	Gua	735	-23.43	-46.46	83075	01/01/1985	31/12/2015	10273/11322
		Taubaté	Tau	577	-23.03	-45.55	83784	01/01/1985	31/12/2015	4009/11322
Automated stations	**Data cluster 3**	Duque de Caxias	DC	33	-22.58	-43.27	86877	21/10/2002	30/09/2016	4273/5094
		Petrópolis	Pe	1777	-22.47	-43.29	86876	22/10/2006	30/09/2016	3008/3632
		São João del Rei	SJ	991	-21.11	-44.24	86849	09/06/2006	30/09/2016	3633/3767
		Teresópolis	Te	980	-22.45	-42.97	86888	01/11/2006	30/09/2016	3598/3622
		Nova Friburgo	NF	1046	-22.33	-42.68	86889	17/09/2010	30/09/2016	2101/2206
		Valença	Va	367	-22.36	-43.70	86875	27/09/2006	30/09/2016	3567/3657
	Data cluster 4	Monte Verde	MV	1545	-22.86	-46.04	86870	19/12/2004	30/09/2016	3925/4304
		São Luis Paratinga	SLP	862	-23.22	-45.41	86912	01/11/2007	30/09/2016	3135/3412
		Cambuci	Ca	60	-21.58	-41.95	86854	20/11/2002	30/09/2016	5604/5974
		Passa-Quatro	PQ	1017	-22.39	-44.96	86873	30/05/2007	30/09/2016	3412/5064
		São Tomê	ST	7	-22.04	-41.05	86890	13/06/2008	30/09/2016	3511/3662
		Macae	Ma	25	-22.37	-41.81	86891	22/09/2006	30/09/2016	3006/3032
		Seropédica	Spd	35	-22.76	-43.02	86878	24/05/2000	30/09/2016	3176/32/57

Source. Data collected by the author from INMET website http://www.inmet.gov.br/

Figure 4. Study area: Paraíba do Sul river watershed location with the INMET stations object of study highlighted by a red star and their corresponding initials

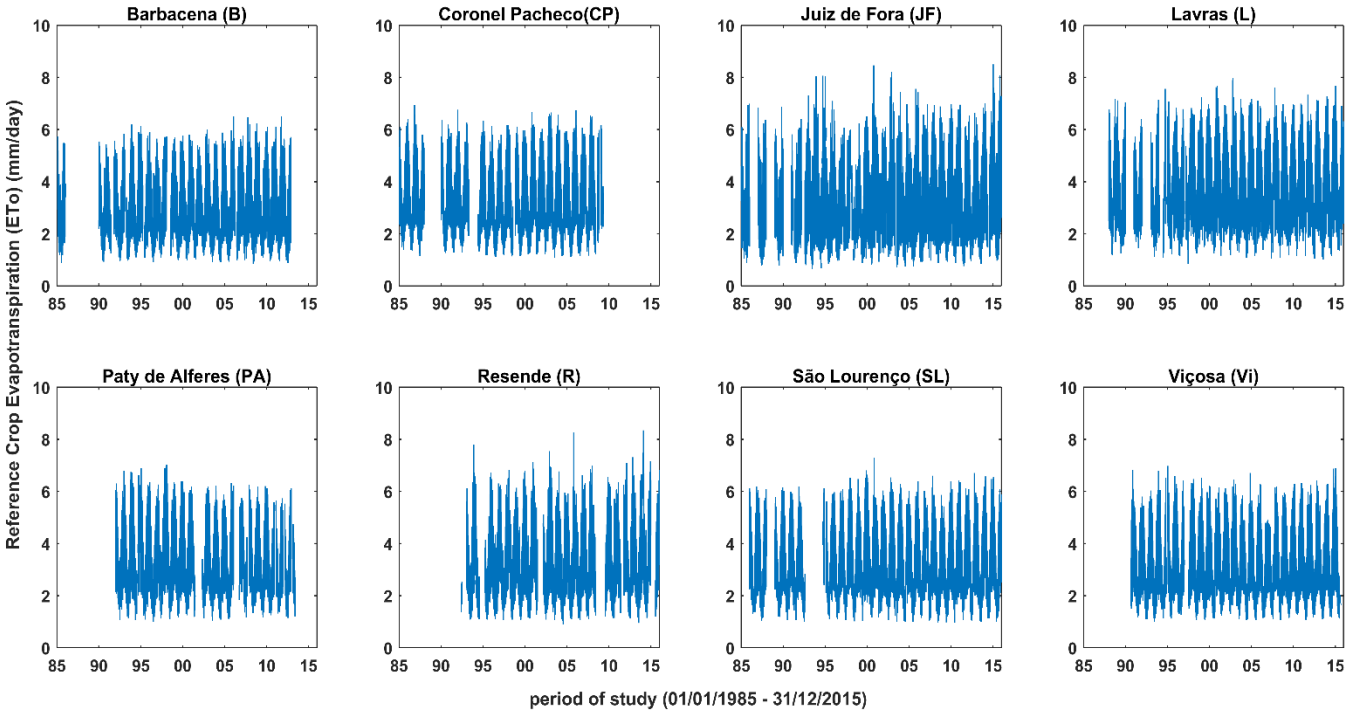


Source. Google Physical Map processed by the author by means of ArcGIS software. Shapefiles of Paraíba do Sul river watershed and riverflow taken from <http://sigaceivap.org.br/siga-ceivap/map>

12. COMPUTATION OF *ETo* FROM INMET DATA SET

In this section, we proceeded to compute *ETo* by using Penman-Monteith FAO 56 equation (Equation 6)

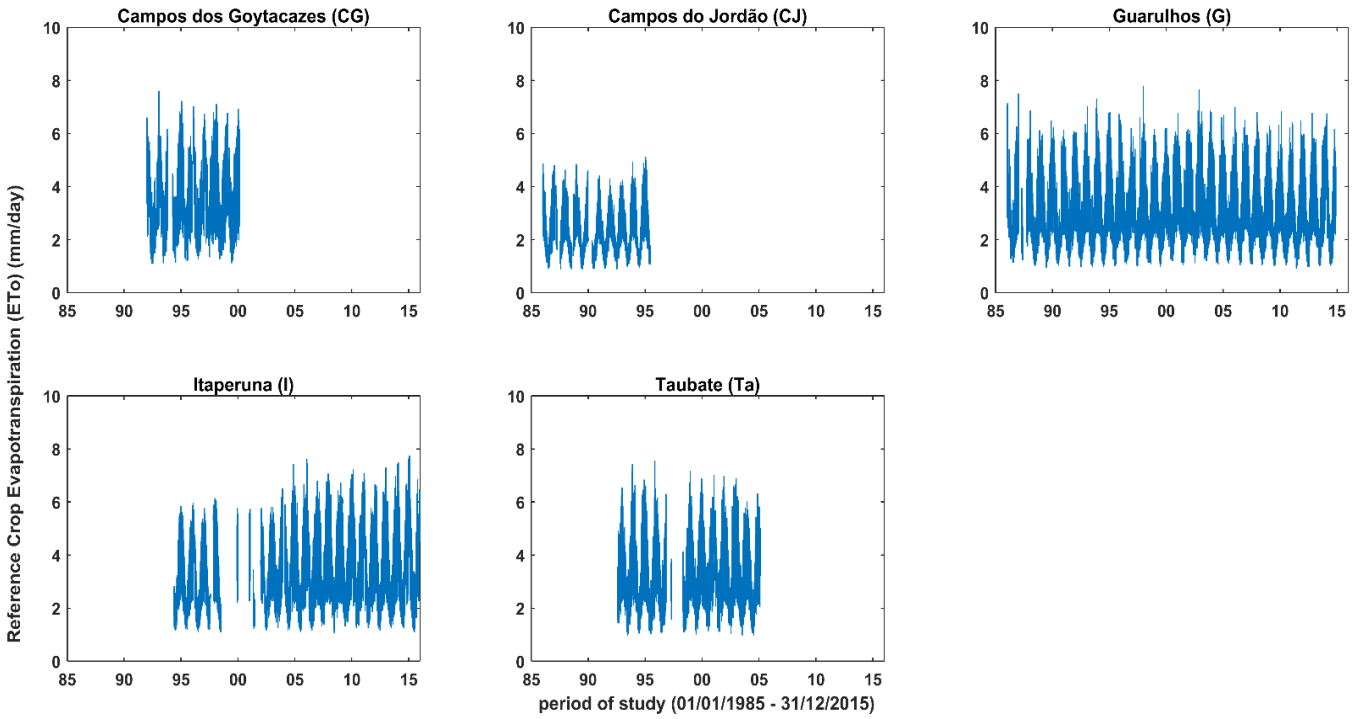
Figure 5. ETo evolution throughout the study period of 31 years of conventional stations belonging to Cluster 1



Source. Result of plotting INMET data through MatLab. Plots performed by the author

The only event deserved to be pinpointed is the higher-than-usual *ETo* values reached in JF, L and R, which some of them went over 8 mm/day.

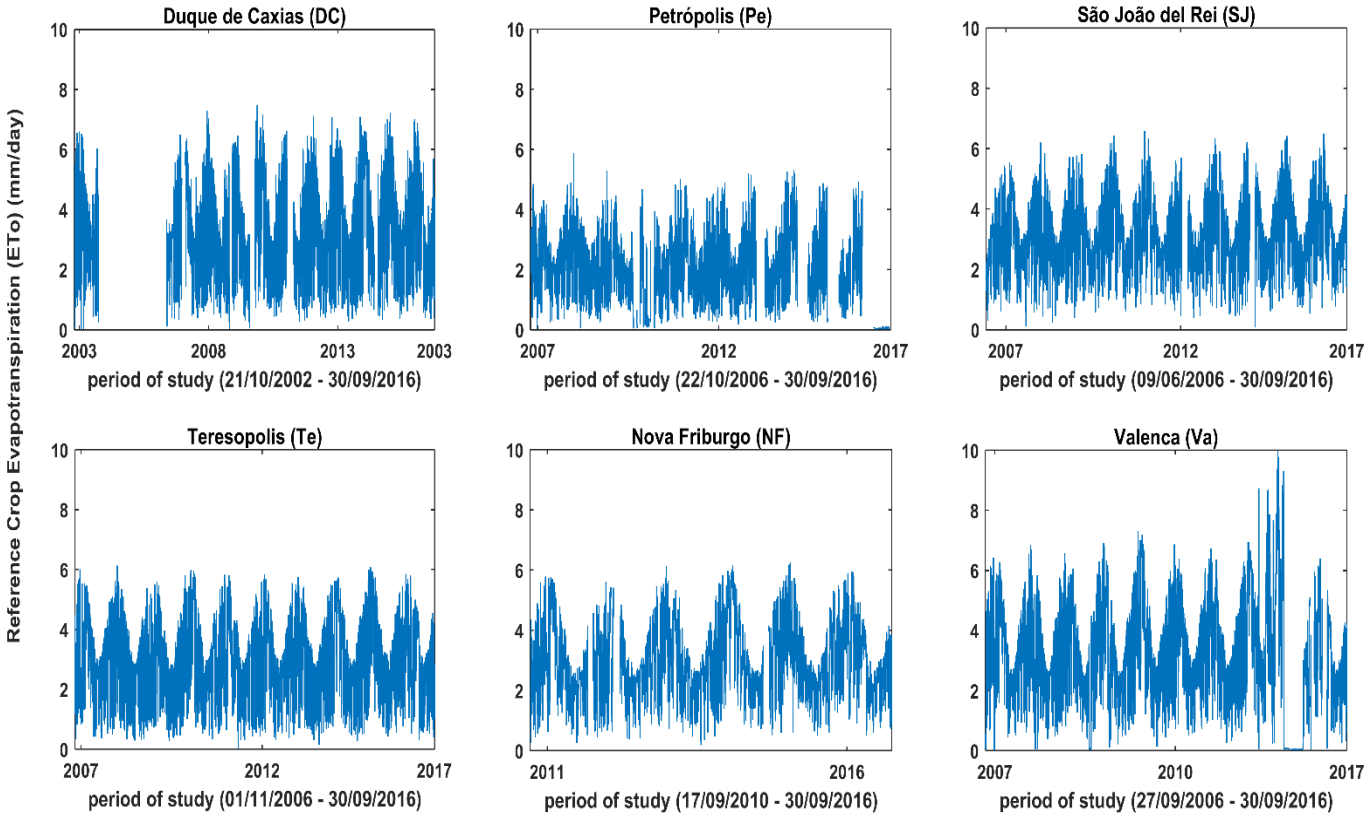
Figure 6. ETo evolution throughout the study period of 31 years of conventional stations belonging to Cluster 2



Source. Result of plotting INMET data through MatLab. Plots performed by the author

In this cluster, the volume of missing values is really high. For instance, in CG, CJ and Ta we have only 8, 9 and 11 years of climatological measures, respectively, out of the 31-year period of study. That may affect the robustness and the reliability of further analyses and conclusions.

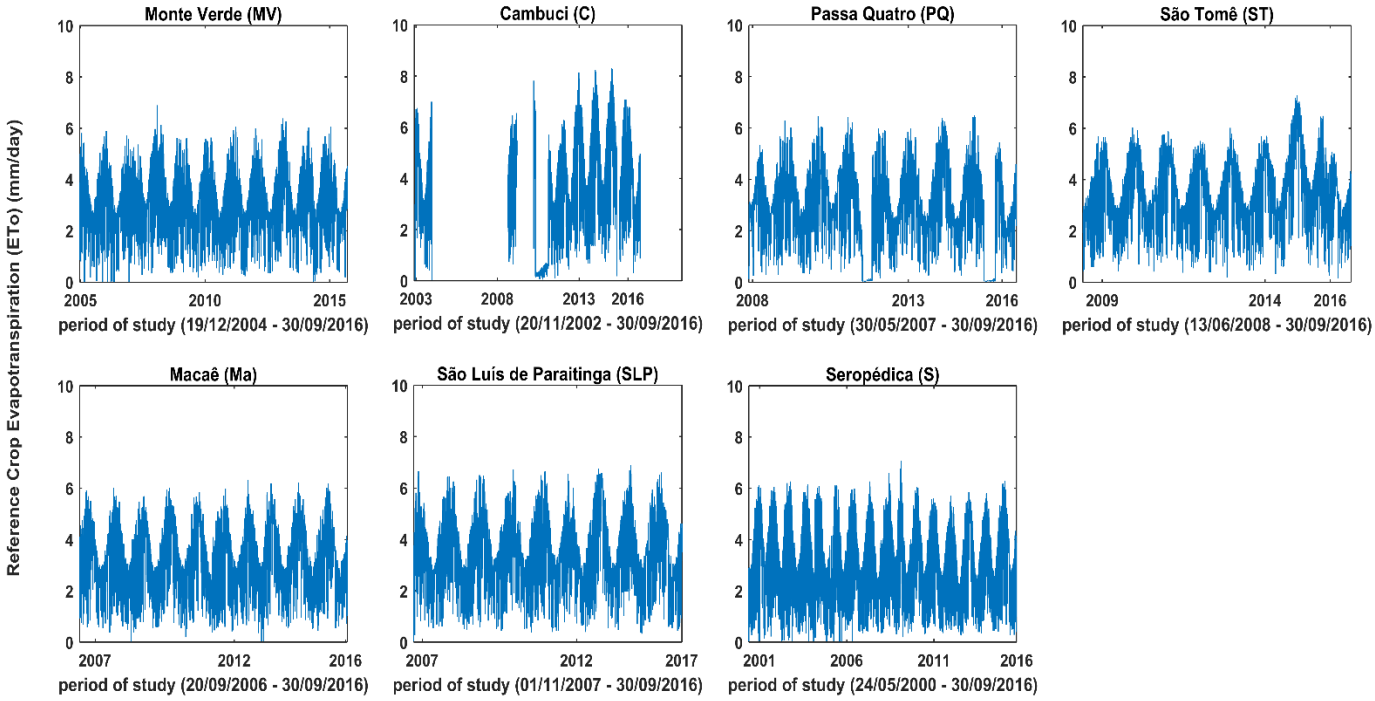
Figure 7. ETo evolution throughout the corresponding study periods of automated stations belonging to Cluster 3



Source. Result of plotting INMET data through MatLab. Plots performed by the author

The most remarkable from cluster 3 is the low average *ETo* values in Pe, mostly below 2.5 mm/day, when the average remains around 4 mm/day. In addition, the extraordinary *ETo* values in Va reached values over 9 mm/day at the beginning of 2003.

Figure 8. ETo evolution throughout the corresponding study periods of automated stations belonging to Cluster 4



Source. Result of plotting INMET data through MatLab. Plots performed by the author

Within cluster 4 it is striking the great gap in C, resting only 6 years running of available *ETo* registers. Apart from that, it is remarkable the fact that we obtained negative values of evapotranspiration, which can be interpreted as water retention or absorption of water, but in any case neither of these events would bring negative values for evapotranspiration, so it must be due to systematic computation errors.

PM56 equation has been applied to each daily measure that compose our dataset. Along the previous section we could visualize the behavior of *ETo* along the study periods and some extraordinary events were highlighted. But this is only the starting point, since plotting *ETo* values along some years is not enough to get any meaningful insight.

So, once we have *ETo* estimates via *PM56* equation and discarding the usage of lysimeters as a direct method of measuring *ETo* due to its high cost both financially and technically, the issue now is to find a procedure to these gaps in. This issue, which is recurrent among both researchers and students, will be treated over the next chapter.

Aiming to show a bit more about our data set and its behaviour over the years that compose the study period without being too extensive, we opted for applying an exploratory

analysis of yearly *ETo* values estimated by *PM56* (Equation 6). As it has been done so far, the present data set is presented split into 4 clusters based on their measurement frequencies: conventional stations (daily measurements): cluster 1 and cluster 2; and automated weather stations (hourly measurements): cluster 3 and cluster 4. In Tables 2, 3 ,4 and 5, m, A and M values account for minimum, mean and maximum annual *ETo* values for each of the 31 years which compose our data set.

CLUSTER 1

Table 4. Annual minimum, average and maximum ETo values for weather stations belonging to Cluster 1

Year	B			CP			JF			L			PA			R			SL			Vi		
	m	A	M	m	A	M	m	A	M	m	A	M	m	A	M	m	A	M	m	A	M	m	A	M
85	0.87	2.94	5.72	1.39	3.32	6.28	0.88	3.10	6.99	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
86	0.00	0.00	0.00	1.34	3.45	6.94	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.30	3.44	6.19	0.00	0.00	0.00
87	0.00	0.00	0.00	1.16	3.30	6.19	0.77	3.10	6.85	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.05	3.32	6.09	0.00	0.00	0.00
88	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.19	3.68	7.16	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
89	0.00	0.00	0.00	0.00	0.00	0.00	0.77	3.01	6.88	1.24	3.52	7.12	0.00	0.00	0.00	0.00	0.00	0.00	1.18	3.27	6.15	0.00	0.00	0.00
90	0.93	3.10	5.53	1.27	3.35	6.35	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.06	3.43	6.01	1.51	3.96	6.82
91	1.25	3.10	5.55	1.29	3.35	6.76	0.87	3.00	6.07	1.45	3.60	6.40	0.00	0.00	0.00	0.00	0.00	0.00	1.07	3.35	6.19	1.18	3.08	5.80
92	0.94	2.71	5.80	1.29	3.15	5.93	0.79	2.96	7.04	0.00	0.00	0.00	1.06	3.44	6.42	1.36	2.24	5.90	1.00	2.84	5.62	1.09	2.86	5.80
93	0.96	2.85	6.19	1.91	3.95	6.27	0.65	3.20	8.03	1.12	3.52	7.14	1.13	3.54	6.78	1.08	3.14	7.79	0.00	0.00	0.00	1.01	3.33	6.80
94	1.10	2.97	5.89	1.33	3.16	5.90	0.67	3.21	8.06	1.09	3.87	7.56	1.05	3.60	6.61	1.10	3.04	6.57	2.40	4.25	6.10	1.14	3.07	6.52
95	1.07	3.10	6.14	1.18	3.23	6.34	0.84	3.19	8.04	1.21	3.70	7.08	1.08	3.46	6.90	1.08	2.96	6.72	1.10	3.26	5.92	1.12	3.37	6.98
96	0.92	2.75	5.89	1.11	3.12	6.14	0.97	2.85	6.77	1.15	3.61	6.85	0.99	3.30	6.36	1.11	3.21	6.47	1.09	3.31	6.01	1.28	3.27	6.42
97	1.08	2.92	5.71	1.17	3.28	6.18	1.34	3.24	6.05	0.84	3.74	6.82	1.12	3.64	6.95	1.06	3.26	6.49	1.06	3.35	6.51	1.51	3.57	6.33
98	1.15	3.08	5.74	1.14	3.16	6.00	0.98	2.74	4.87	1.26	3.83	6.78	1.10	3.50	7.01	1.08	3.11	6.82	1.02	3.32	6.47	1.08	3.21	6.47
99	0.99	3.00	5.69	1.14	3.32	6.11	0.88	2.93	6.50	1.48	3.97	7.24	1.16	3.47	6.40	1.06	3.13	6.80	1.02	3.51	6.81	1.08	3.26	6.53
00	0.98	2.94	5.79	1.22	3.33	6.67	0.81	3.14	8.46	1.24	3.87	7.66	1.25	3.44	6.21	1.12	3.19	7.11	1.16	3.37	7.30	1.12	3.12	6.21
01	1.00	2.99	5.56	1.22	3.53	6.20	0.93	3.35	6.47	1.16	4.03	7.20	1.19	3.62	6.18	1.27	3.56	6.87	1.16	3.47	6.03	1.22	3.46	6.15
02	1.16	2.92	5.99	1.21	3.53	6.60	1.11	3.48	8.22	1.17	4.09	7.99	1.14	3.14	6.11	1.08	3.28	7.54	1.06	3.46	6.54	1.31	3.46	6.44
03	0.91	2.81	5.64	1.06	3.37	6.66	0.72	3.43	7.44	1.28	3.85	7.16	1.04	3.35	6.23	1.07	3.32	6.92	1.00	3.36	6.27	1.10	3.26	6.28
04	0.91	2.65	5.88	1.15	3.08	5.98	0.74	2.97	6.97	1.07	3.72	7.04	1.08	3.17	5.87	0.89	3.08	6.77	1.00	3.21	6.14	1.06	2.97	6.71
05	0.93	2.59	5.96	1.12	3.24	6.07	0.75	2.92	7.56	1.14	3.72	7.17	1.10	3.17	6.31	1.11	3.13	8.27	1.00	3.17	6.02	1.06	3.03	5.86
06	0.99	3.03	6.50	1.14	3.28	6.74	0.91	3.28	7.42	1.37	3.83	6.90	1.84	4.06	6.21	1.07	3.14	6.61	1.15	3.31	6.38	1.15	2.98	6.29
07	1.00	3.14	6.48	1.17	3.59	6.54	0.97	3.48	6.98	1.06	4.13	7.61	1.17	3.41	6.25	1.11	3.51	7.00	1.04	3.37	6.60	1.10	2.97	6.20
08	0.83	2.83	6.24	1.21	3.43	6.48	0.88	3.08	6.89	1.07	3.74	6.97	1.34	3.35	6.19	1.21	3.13	5.54	0.96	3.13	5.86	1.07	3.02	5.95
09	1.09	2.81	5.93	2.19	4.18	6.17	0.93	3.03	6.43	1.18	3.72	6.78	1.17	3.46	6.20	1.19	3.25	6.55	0.97	3.08	6.39	1.15	3.14	5.88
10	0.95	3.11	6.11	0.00	0.00	0.00	0.88	3.36	6.96	1.14	4.03	6.99	1.24	3.31	6.12	1.12	3.35	6.28	0.97	3.44	6.41	1.28	3.39	6.31
11	0.96	3.00	6.51	0.00	0.00	0.00	0.76	3.07	6.95	1.14	3.90	7.35	1.31	3.15	5.73	1.34	3.51	6.46	1.07	3.36	6.36	1.05	3.15	6.12
12	0.89	2.95	5.71	0.00	0.00	0.00	0.95	3.56	6.63	1.05	4.02	7.15	1.14	3.20	6.12	1.12	3.38	7.31	1.04	3.40	6.70	1.16	3.37	6.42
13	0.00	0.00	0.00	0.00	0.00	0.00	0.85	3.12	6.84	1.02	3.70	6.91	1.21	2.80	4.75	0.95	3.06	6.64	1.00	3.23	6.19	1.11	3.01	6.01
14	0.00	0.00	0.00	0.00	0.00	0.00	1.11	4.01	7.21	1.15	4.32	7.28	0.00	0.00	0.00	1.14	3.53	8.34	1.12	3.72	6.58	1.17	3.69	6.87
15	0.00	0.00	0.00	0.00	0.00	0.00	0.96	3.48	8.50	1.35	3.86	7.67	0.00	0.00	0.00	1.19	3.51	6.83	1.00	3.35	6.54	1.12	3.23	6.89

Source: data computed by the author

CLUSTER 2

Table 5. Annual minimum, average and maximum ETo values for weather stations belonging to cluster 2

Year	CG			CJ			G			I			Ta		
	m	A	M	m	A	M	m	A	M	m	A	M	m	A	M
85	0.00	0.00	0.00	0.00	0.00	0.00	4.50	4.50	4.50	0.00	0.00	0.00	0.00	0.00	0.00
86	0.00	0.00	0.00	0.92	2.42	4.87	1.12	3.04	7.16	0.00	0.00	0.00	0.00	0.00	0.00
87	0.00	0.00	0.00	0.88	2.66	4.81	1.23	3.54	7.51	0.00	0.00	0.00	0.00	0.00	0.00
88	0.00	0.00	0.00	0.87	2.35	4.84	1.01	2.98	6.87	0.00	0.00	0.00	0.00	0.00	0.00
89	0.00	0.00	0.00	0.88	2.24	4.62	0.94	3.00	6.49	0.00	0.00	0.00	0.00	0.00	0.00
90	0.00	0.00	0.00	0.90	2.12	4.43	1.02	3.26	6.22	0.00	0.00	0.00	0.00	0.00	0.00
91	0.00	0.00	0.00	0.92	2.21	4.28	0.99	3.12	6.10	0.00	0.00	0.00	0.00	0.00	0.00
92	1.08	3.27	6.60	0.87	2.23	4.18	0.97	3.09	6.37	2.20	2.92	3.94	1.43	3.58	6.54
93	1.12	3.52	7.60	0.98	2.29	4.94	0.96	3.09	7.31	0.00	0.00	0.00	0.96	3.26	7.44
94	1.14	3.46	6.84	0.92	2.43	4.90	0.96	3.12	6.78	1.16	3.00	5.84	1.04	3.40	6.85
95	1.20	3.67	7.23	1.06	2.49	5.11	1.00	3.08	6.80	1.19	3.11	5.64	1.02	3.22	7.56
96	1.26	3.29	7.02	0.00	0.00	0.00	1.07	2.99	6.35	1.08	3.22	5.97	1.05	3.05	6.31
97	1.34	3.92	6.80	0.00	0.00	0.00	0.99	3.14	7.79	1.19	3.33	6.14	1.56	3.08	4.12
98	1.26	3.72	7.13	0.00	0.00	0.00	1.07	3.01	6.40	1.09	3.35	6.12	1.11	3.04	7.18
99	1.12	3.47	6.77	0.00	0.00	0.00	0.97	3.12	6.21	2.20	3.99	5.79	1.00	3.39	6.88
00	1.97	4.69	6.94	0.00	0.00	0.00	1.19	3.19	6.21	0.00	0.00	0.00	1.14	3.35	6.32
01	0.00	0.00	0.00	0.00	0.00	0.00	1.04	3.39	6.78	1.23	3.03	5.73	1.11	3.45	7.03
02	0.00	0.00	0.00	0.00	0.00	0.00	1.12	3.49	7.66	1.15	3.35	5.81	1.02	3.45	6.88
03	0.00	0.00	0.00	0.00	0.00	0.00	1.20	3.33	6.88	1.24	3.52	6.52	0.98	3.31	6.21

04	0.00	0.00	0.00	0.00	0.00	0.00	1.06	3.12	6.61	1.16	3.39	7.42	0.97	3.10	6.33
05	0.00	0.00	0.00	0.00	0.00	0.00	1.08	3.13	6.71	1.10	3.48	6.67	2.05	3.89	5.86
06	0.00	0.00	0.00	0.00	0.00	0.00	1.04	3.24	6.98	1.16	3.62	7.64	0.00	0.00	0.00
07	0.00	0.00	0.00	0.00	0.00	0.00	1.06	3.20	6.81	1.31	3.84	7.07	0.00	0.00	0.00
08	2.07	2.07	2.07	0.00	0.00	0.00	1.03	2.99	6.35	1.08	3.47	6.82	0.00	0.00	0.00
09	0.00	0.00	0.00	0.00	0.00	0.00	1.01	2.92	6.31	1.31	3.64	6.85	2.19	2.20	2.22
10	0.00	0.00	0.00	0.00	0.00	0.00	1.04	2.95	6.83	1.25	3.82	7.23	0.00	0.00	0.00
11	0.00	0.00	0.00	0.00	0.00	0.00	0.92	2.96	6.48	1.24	3.53	7.09	0.00	0.00	0.00
12	0.00	0.00	0.00	0.00	0.00	0.00	0.96	3.23	6.77	1.15	3.75	7.13	0.00	0.00	0.00
13	0.00	0.00	0.00	0.00	0.00	0.00	1.01	3.03	6.23	1.25	3.49	7.30	0.00	0.00	0.00
14	0.00	0.00	0.00	0.00	0.00	0.00	1.02	3.29	6.75	1.25	3.92	7.49	0.00	0.00	0.00
15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.11	3.79	7.74	0.00	0.00	0.00

Source: data computed by the author

CLUSTER 3

Table 6. Annual minimum, average and maximum ETo values for weather stations belonging to cluster 3

Year	DC			Pe			SJ			Te			NF			Va		
	m	**A**	**M**	**m**	**A**	**M**	**m**	**A**	**M**	**m**	**A**	**M**	**m**	**A**	**M**	**m**	**A**	**M**
2008	0.40	3.44	7.28	0.24	2.29	5.88	0.12	3.13	6.20	0.36	2.99	6.14	0.00	0.00	0.00	0.43	3.36	6.85
2009	0.03	2.98	6.53	0.05	2.08	5.30	0.23	2.98	6.07	0.29	2.65	5.60	0.00	0.00	0.00	0.05	2.94	6.66
2010	0.04	3.30	7.48	0.04	1.99	4.68	0.58	3.09	6.18	0.27	2.73	5.81	0.00	0.00	0.00	0.03	3.03	6.39
2011	0.43	3.37	7.15	0.07	1.91	5.02	0.61	3.31	6.58	0.34	2.99	6.00	0.26	2.97	5.81	0.10	3.27	7.31
2012	0.27	3.05	6.62	0.26	2.12	4.81	0.64	3.16	6.33	0.05	2.75	5.81	0.26	2.75	5.77	0.14	3.38	7.19
2013	0.39	3.50	7.10	0.19	2.31	5.19	0.39	3.11	6.34	0.31	3.03	5.84	0.38	2.89	6.14	0.50	3.53	6.73
2014	0.29	3.18	7.09	0.25	2.03	5.25	0.46	3.05	6.22	0.15	2.68	5.61	0.18	2.76	5.83	0.25	3.19	6.37
2015	0.34	3.92	7.00	0.26	2.59	5.31	0.09	3.24	6.31	0.28	3.15	6.09	0.52	3.28	6.15	0.06	4.58	10.48
2016	0.37	3.40	7.23	0.25	2.05	4.93	0.53	3.29	6.43	0.44	3.00	5.93	0.50	2.97	6.26	0.02	1.32	6.16

Source: data computed by the author

CLUSTER 4

Table 7. Annual minimum, average and maximum ETo values for weather stations belonging to cluster 4

Year	MV			C			PQ			ST			Ma			SLP			S		
	m	**A**	**M**	**m**	**A**	**M**	**m**	**A**	**M**	**m**	**A**	**M**	**m**	**A**	**M**	**m**	**A**	**M**	**m**	**A**	**M**
2008	0.0	0.8	5.4	0.0	0.0	0.0	0.0	1.9	5.8	0.0	0.0	0.0	0.0	3.5	7.0	0.0	3.2	6.4	0.0	3.2	7.1
2009	0.0	1.9	5.3	0.0	1.2	6.4	0.0	3.0	5.6	0.0	3.4	6.2	0.0	2.9	6.4	0.0	2.7	6.2	0.0	2.9	6.9
2010	0.0	2.3	5.3	0.0	0.7	5.7	0.0	3.1	5.9	0.0	3.8	6.6	0.0	3.4	7.0	0.0	2.9	6.0	0.0	2.6	6.3
2011	0.0	2.7	5.5	0.0	1.1	7.2	0.0	3.1	6.2	0.0	4.0	7.2	0.0	3.6	7.3	0.0	2.9	6.2	0.0	2.3	6.8
2012	0.0	2.5	5.5	0.0	2.5	6.4	0.0	2.4	5.8	0.0	3.8	6.9	0.0	3.0	6.8	0.0	2.9	6.4	0.0	3.3	7.2
2013	0.0	2.6	6.0	0.0	3.2	7.4	0.0	3.0	6.2	0.0	4.0	7.3	0.0	3.1	6.9	0.0	2.5	6.3	0.0	3.2	7.3
2014	0.0	2.4	5.7	0.0	3.4	7.3	0.0	3.5	6.5	0.0	3.8	7.2	0.0	2.8	7.0	0.0	2.9	6.2	0.0	2.9	6.7
2015	0.0	2.8	5.6	0.0	3.9	7.6	0.0	3.2	6.6	0.0	3.8	6.8	0.0	3.9	7.4	0.0	3.2	6.5	0.0	3.5	7.5
2016	0.0	2.0	5.7	0.0	3.0	7.0	0.0	1.1	5.7	0.0	3.3	6.1	0.0	3.7	8.1	0.0	2.6	6.6	0.0	3.1	7.8

Source: data computed by the author

CHAPTER II
ESTIMATION OF MISSING VALUES VIA ARTIFICIAL NEURAL NETWORKS

1. MISSING VALUES WITHIN HISTORICAL CLIMATOLOGICAL REGISTERS: A RECURRENT ISSUE AMONG RESEARCHERS

Climate has always been matter of study for mankind since it has dealt with nature first and agriculture later. Very primitive gadgets were used for some centuries along with religious beliefs to somehow predict what would be next station weather like, harvest abundance or likely plagues. After Second World War, technology made possible climate measurement, analysis and continuous monitoring. Indeed, climate monitoring turned into a priority for worldwide governments as the first global warming signs were revealed. Taking advantage of the ultimate technology developed during the warlike conflict, thousands of weather stations were installed in order to know a bit more about the climate behaviour and somehow, predict natural phenomena such as hurricanes, earthquakes and, by then, the rising global warming. This net has been broaden and managed globally since then by National Meteorological Services as a part of the World Meteorological Organization World Weather Watch, which networks the observant stations to national, regional and global weather along with climate prediction centres 24 hours a day in real-time.

Nowadays, weather stations are high-technology devices built taking into account the exposure they would suffer once they are installed outdoors. Despite that, they usually suffer with inclement weather: wind gasps, gales, rainstorms, thunderstorms, flood, intense run-off or soil erosion as the most common events. Their actions can result in a bad operating or even the complete ruin of the equipment, which obviously lead us to errors in the measurement of the climatological parameters we are interested in. In worst cases, data registers show great gaps with no inputs at all. This is a recurrent problem according to scientist, researchers or students: incomplete registers that make their analyses discontinuous and their corresponding conclusions less robust and meaningful.

The reliability of the automated weather stations has greatly improved the gathering of meteorological parameters in last decades. Even though, failures in weather stations still

happen and techniques to fill gaps in data sequences turn it necessary in order to use these "estimated" values as model inputs. In those cases, which are rather usual, FAO has a protocol to follow aiming to estimate these missing values and encourages weather-data users not to use alternative approach for the estimation of *ETo* if they count only with limited climatological parameters.

That protocol recommends the importation of these parameters from nearby weather stations on the corresponding dates where the original measurements are missing. Doubtless, the procedure of importing values from nearby weather stations are restrained by some conditions which must be met and are usually different depending on the climatological parameter to be imported. We have been no an exception and gaps have been found in all parameters that take part in the Penman-Monteith FAO 56 equation. Thus, the main objective of this chapter is to show the alternative path we followed aiming to estimate *ETo* values even when we faced incomplete climatological datasets. With it, we hope to aid scientist, researchers and students who face this recurrent problem in their day-by-day work.

2. THE FAO PROTOCOL TO REFILL MISSING VALUES

When importing a set of values from a given register to a different one we can refer to them as importing (*Xi*) and receiving register (*Yi*), namely. In order to use portions of data set X_i to replace data set Y_i, both data sets must be homogeneous. In other words, they need to represent the same conditions. The procedure of completing incomplete data sets is applied only after proving the existence of homogeneity within importing and receiving registers or carrying out the corrections required in case of non-homogeneity. The substitution procedure proposed herein consists of using an appropriate regression analysis. The procedure for importing nearby data into an incomplete data set can be summarized as follows:

1. Select a nearby weather station whose *Xi* data set length covers all periods for which Y_i data is missing.

2. Characterize the data sets from the nearby importing station X_i and of the station having missing data Y_i by computing the mean ($\bar{x}$) and the standard deviation (S_x) for both data sets X_i and *Yi*:

$$\bar{x} = \sum_{i=1}^{n} \frac{x_i}{n} \quad \text{and} \quad \bar{y} = \sum_{i=1}^{n} \frac{y_i}{n} \qquad \textbf{(49a, 49b)}$$

$$S_x = \sqrt{\sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n-1}} \quad \text{and} \quad S_y = \sqrt{\sum_{i=1}^{n} \frac{(y_i - \bar{y})^2}{n-1}} \qquad \textbf{(50a, 50b)}$$

For periods where measures are present in both data sets, where x_i and y_i are individual observations from data sets X_i and Y_i, and n is the number of observations in each set.

3. Perform a regression of *y* on *x* within these periods:

$$\hat{y}_i = a + bx_i \quad \textbf{(51)}$$

with

$$b = \frac{\mathrm{cov}_{xy}}{s_x^2} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})*(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \quad \textbf{(52)}$$

$$a = \bar{y} - b\bar{x} \quad \textbf{(53)}$$

where *a* and *b* are empirical regression constants, and cov_{xy} is the covariance between X_i and Y_i. Plot all points x_i and y_i and the regression line for the range of observed values. If deviations from the regression line increase with *y*, then substitution is not recommended because this indicates that the two datasets have a different behaviour relative to the particular weather variable, and they may not be homogeneous. Another nearby station should be selected in those cases. If correlation is high enough we can carry on.

4. Compute the correlation coefficient r:

$$r = \frac{\mathrm{cov}_{xy}}{s_x^2 s_y} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})*(y_i - \bar{y})}{\sqrt{\left(\sum_{i=1}^{n}(x_i - \bar{x})^2 * \sum_{i=1}^{n}(y_i - \bar{y})^2\right)}} \quad \textbf{(54)}$$

A high r^2 ($r^2 \geq 0.7$) indicates good conditions and perhaps sufficient homogeneity for replacing missing data in the incomplete data series. Apart from r^2, *b* can also be used as criteria for selecting the best nearby station, although we preferred only to use r^2.

5. Compute the data for the missing periods $k = n+1, n+2...$, m using the regression equation (Equation 51) characterized by the parameters *a* and *b*. The complete receiving register *Yi* with dimension m will now be:

$$Y_j = y_i \text{ and } Y_j = \hat{y}_k \qquad \textbf{(55a, 55b)}$$

3. RESULTS OF CORRELATION BETWEEN NEARBY WEATHER STATIONS

In order to carry out this procedure in an organized way, a total of 26 weather stations, which comprises the present research, were clustered into 2 classes: conventional and automated weather stations, which in turn were subdivided into two clusters each, aiming to better visualize the results: cluster 1 and 2 and clusters 3 and 4, respectively.

Here, each climatological parameter was treated individually; in other words, correlations with respect to individual parameters were performed between nearby datasets. Thus, we can check the results obtained in terms of indexes r^2 and *b* in Table 8, Table 9, Table 10, Table 11 and Table 12:

CONVENTIONAL STATIONS CLASS

Maximum Temperature (*Tmax*)

Table 8. Correlation results in terms of r and b among conventional stations as far as Maximum Temperature (Tmax) is concerned

		B	CG	CJ	CP	Gu	It	JF	Lv	PA	Re	SL	Ta	Vi
B	r^2	1.00	0.60	0.52	**0.77**	0.48	**0.75**	**0.79**	**0.75**	**0.75**	0.65	0.59	0.54	**0.74**
	b	0.00	10.22	3.67	7.33	5.16	7.46	0.60	7.96	5.31	5.15	9.59	7.25	6.77
CG	r^2	0.60	1.00	0.36	0.64	0.40	**0.80**	0.69	0.44	0.69	0.55	0.37	0.44	0.61
	b	2.03	0.00	4.22	6.29	3.30	2.52	-2.11	9.86	2.43	2.97	10.92	5.48	5.31
CJ	r^2	0.52	0.36	1.00	0.49	**0.70**	0.42	0.53	0.64	0.62	0.63	**0.72**	**0.78**	0.45
	b	9.14	16.49	0.00	13.46	3.50	15.30	7.10	11.64	9.60	7.77	10.02	6.08	13.11
CP	r^2	**0.77**	0.64	0.49	1.00	0.46	**0.78**	**0.81**	0.64	**0.80**	0.67	0.58	0.57	**0.74**
	b	-1.05	5.36	1.51	0.00	1.66	3.57	-4.09	6.66	0.33	0.79	6.52	3.67	3.34
Gu	r^2	0.48	0.40	**0.70**	0.46	1.00	0.45	0.53	0.48	0.65	**0.72**	0.55	**0.84**	0.38
	b	10.21	16.11	3.87	14.45	0.00	15.25	7.67	14.02	9.78	7.22	12.74	5.45	14.56
It	r^2	**0.75**	**0.80**	0.42	**0.78**	0.45	1.00	**0.78**	0.54	**0.75**	0.63	0.45	0.48	**0.74**
	b	0.12	3.81	3.05	3.46	2.63	0.00	-2.86	8.29	1.91	1.95	9.50	5.05	3.75
JF	r^2	**0.79**	0.69	0.53	**0.81**	0.53	**0.78**	1.00	0.65	**0.84**	**0.71**	0.55	0.60	**0.73**
	b	4.79	10.86	5.33	8.61	6.12	9.19	0.00	11.00	6.06	6.33	11.92	7.98	8.81
Lv	r^2	**0.75**	0.44	0.64	0.64	0.48	0.54	0.65	1.00	0.65	0.61	**0.79**	0.57	0.64
	b	-1.52	9.47	-2.44	4.36	0.23	6.80	-2.17	0.00	1.99	0.92	2.44	1.62	4.17
PA	r^2	**0.75**	0.69	0.62	**0.80**	0.65	**0.75**	**0.84**	0.65	1.00	**0.80**	0.64	**0.72**	0.66
	b	1.90	7.33	1.08	5.48	-0.05	6.00	-1.69	8.28	0.00	0.84	7.79	2.42	6.65
Re	r^2	0.65	0.55	0.63	0.67	**0.72**	0.63	**0.71**	0.61	**0.80**	1.00	0.64	**0.79**	0.53
	b	5.19	11.32	2.54	8.88	0.84	9.68	2.05	10.16	4.78	0.00	9.20	2.89	10.15
SL	r^2	0.59	0.37	**0.72**	0.58	0.55	0.45	0.55	**0.79**	0.64	0.64	1.00	0.67	0.53
	b	2.08	11.69	-2.95	6.94	-0.76	9.43	0.61	3.66	3.09	0.81	0.00	0.21	6.73
Ta	r^2	0.54	0.44	**0.78**	0.57	**0.84**	0.48	0.60	0.57	**0.72**	**0.79**	0.67	1.00	0.44
	b	7.01	13.50	0.23	9.93	-1.21	12.54	4.22	10.73	5.99	3.33	8.88	0.00	11.75
Vi	r^2	**0.74**	0.61	0.45	**0.74**	0.38	**0.74**	**0.73**	0.64	0.66	0.53	0.53	0.44	1.00
	b	0.30	7.26	2.53	4.42	4.64	4.28	-1.89	6.78	3.54	4.02	8.11	6.38	0.00

Source. data computed by the author

Minimum Temperature (*Tmin*)

Table 9. Correlation results in terms of r² and b among conventional stations as far as Minimum Temperature (Tmin) is concerned

		B	CG	CJ	CP	Gu	It	JF	Lv	PA	Re	SL	Ta	Vi
B	***r²***	1.00	**0.74**	0.69	**0.77**	0.63	**0.75**	0.68	**0.84**	**0.78**	**0.75**	**0.77**	**0.76**	**0.77**
	b	0.00	10.99	-5.38	1.47	4.74	7.78	5.65	2.79	-0.28	4.07	-3.44	0.99	1.86
CG	***r²***	**0.74**	1.00	0.63	0.65	0.67	**0.80**	**0.73**	**0.74**	0.69	**0.72**	0.66	**0.72**	0.68
	b	-7.74	0.00	-16.19	-8.24	-5.03	-2.46	-2.77	-6.00	-11.14	-6.09	-14.65	-10.14	-7.81
CJ	***r²***	0.69	0.63	1.00	0.64	0.66	0.60	0.59	**0.75**	**0.72**	**0.70**	**0.82**	**0.76**	0.62
	b	7.73	15.82	0.00	9.95	9.70	13.60	10.58	8.91	7.46	10.21	4.00	7.26	8.64
CP	***r²***	**0.77**	0.65	0.64	1.00	0.53	**0.79**	0.56	**0.80**	**0.80**	**0.75**	**0.77**	**0.74**	**0.82**
	b	2.19	11.80	-5.81	0.00	5.82	8.12	6.78	3.60	0.17	4.29	-2.87	1.92	1.92
Gu	***r²***	0.63	0.67	0.66	0.53	1.00	0.56	**0.74**	0.67	0.58	0.66	0.61	**0.73**	0.51
	b	1.41	10.18	-6.61	2.58	0.00	8.06	3.83	2.58	0.49	3.19	-3.43	-0.87	2.92
It	***r²***	**0.75**	**0.80**	0.60	**0.79**	0.56	1.00	0.61	**0.77**	**0.77**	**0.76**	**0.72**	**0.72**	**0.81**
	b	-3.44	5.98	-9.84	-5.75	0.66	0.00	2.16	-1.99	-6.82	-1.86	-10.01	-4.92	-4.64
JF	***r²***	0.68	**0.73**	0.59	0.56	**0.74**	0.61	1.00	0.68	0.61	0.62	0.59	0.65	0.53
	b	-1.06	7.96	-8.35	-0.05	0.13	5.74	0.00	0.63	-2.70	1.51	-5.68	-2.07	0.90
Lv	***r²***	**0.84**	**0.74**	**0.75**	**0.80**	0.67	**0.77**	0.68	1.00	**0.81**	**0.82**	**0.87**	**0.83**	**0.79**
	b	-0.36	9.79	-7.39	-0.81	3.11	6.14	4.54	0.00	-2.30	1.64	-6.47	-1.57	0.07
PA	***r²***	**0.78**	0.69	**0.72**	**0.80**	0.58	**0.77**	0.61	**0.81**	1.00	**0.79**	**0.82**	**0.80**	**0.80**
	b	3.20	12.41	-3.70	3.19	6.41	9.12	7.38	4.53	0.00	5.24	-1.82	2.35	3.39
Re	***r²***	**0.75**	**0.72**	**0.70**	**0.75**	0.66	**0.76**	0.62	**0.82**	**0.79**	1.00	**0.80**	**0.84**	**0.73**
	b	0.18	9.72	-7.18	-0.11	3.04	6.11	4.85	1.46	-2.24	0.00	-5.85	-1.79	0.45
SL	***r²***	**0.77**	0.66	**0.82**	**0.77**	0.61	**0.72**	0.59	**0.87**	**0.82**	**0.80**	1.00	**0.84**	**0.76**
	b	5.44	14.24	-1.66	5.93	8.23	11.48	9.05	6.45	4.28	7.61	0.00	4.81	6.13
Ta	***r²***	**0.76**	**0.72**	**0.76**	**0.74**	**0.73**	**0.72**	0.65	**0.83**	**0.80**	**0.84**	**0.84**	**1.00**	**0.71**
	b	2.57	11.85	-4.31	2.49	4.95	8.85	6.54	3.78	0.88	4.22	-2.87	0.00	3.42
Vi	***r²***	**0.77**	0.68	0.62	**0.82**	0.51	**0.81**	0.53	**0.79**	**0.80**	**0.73**	**0.76**	**0.71**	1.00
	b	1.78	11.40	-3.89	1.01	6.00	7.31	6.75	3.19	-0.31	4.28	-3.33	1.45	0.00

Source. data computed by the author

Solar Radiation (*Rs*)

Table 10. Correlation results in terms of r² and b among conventional stations as far as Solar Radiation (Rs) is concerned

		B	CG	CJ	CP	Gu	It	JF	Lv	PA	Re	SL	Ta	Vi
B	**r²**	1.00	0.49	0.29	**0.72**	0.33	0.60	**0.73**	0.63	0.67	0.59	0.51	0.39	0.62
	b	0.00	4.72	7.28	2.93	6.45	4.54	2.12	5.93	3.90	3.83	6.73	6.32	4.30
CG	**r²**	0.49	1.00	0.12	0.54	0.24	**0.70**	0.52	0.32	0.53	0.38	0.25	0.26	0.51
	b	4.41	0.00	9.82	5.02	7.63	3.62	4.30	8.94	5.13	5.68	9.74	8.16	5.64
CJ	**r²**	0.29	0.12	1.00	0.24	0.50	0.16	0.23	0.45	0.38	0.46	0.56	0.61	0.19
	b	6.28	9.89	0.00	7.64	3.10	9.98	7.12	6.47	5.91	3.90	4.41	2.99	9.23
CP	**r²**	**0.72**	0.54	0.24	1.00	0.28	0.65	**0.77**	0.56	0.69	0.56	0.45	0.36	0.65
	b	1.92	3.69	7.88	0.00	6.93	3.71	1.50	6.10	3.12	3.62	6.82	6.62	3.70
Gu	**r²**	0.33	0.24	0.50	0.28	1.00	0.24	0.31	0.38	0.44	0.51	0.49	**0.74**	0.20
	b	6.46	8.35	5.31	8.13	0.00	9.39	6.89	8.45	6.38	4.46	6.85	2.80	9.53
It	**r²**	0.60	**0.70**	0.16	0.65	0.24	1.00	0.60	0.41	0.56	0.45	0.31	0.28	0.65
	b	2.82	1.78	8.77	2.82	7.39	0.00	2.78	7.60	4.28	4.68	8.55	7.31	3.32
JF	**r²**	**0.73**	0.52	0.23	**0.77**	0.31	0.60	1.00	0.54	**0.71**	0.58	0.44	0.38	0.60
	b	2.35	4.39	8.29	2.43	6.81	4.68	0.00	6.80	3.39	3.86	7.54	6.54	4.61
Lv	**r²**	0.63	0.32	0.45	0.56	0.38	0.41	0.54	1.00	0.61	0.57	**0.73**	0.48	0.47
	b	0.87	5.27	4.07	2.57	4.22	5.00	2.11	0.00	2.40	1.84	2.27	3.43	4.00
PA	**r²**	0.67	0.53	0.38	0.69	0.44	0.56	**0.71**	0.61	1.00	**0.71**	0.60	0.55	0.51
	b	2.05	3.88	5.96	2.52	4.53	4.24	1.65	5.15	0.00	1.70	4.95	3.87	4.76
Re	**r²**	0.59	0.38	0.46	0.56	0.51	0.45	0.58	0.57	**0.71**	1.00	0.59	0.66	0.41
	b	3.39	6.26	5.59	4.62	4.28	6.34	3.55	6.31	3.46	0.00	5.76	3.39	6.59
SL	**r²**	0.51	0.25	0.56	0.45	0.49	0.31	0.44	**0.73**	0.60	0.59	1.00	0.60	0.36
	b	2.79	7.17	3.71	4.45	3.04	6.83	3.61	2.98	2.84	1.89	0.00	2.34	5.87
Ta	**r²**	0.39	0.26	0.61	0.36	**0.74**	0.28	0.38	0.48	0.55	0.66	0.60	1.00	0.25
	b	5.12	7.31	3.69	6.43	1.58	8.24	5.50	6.96	4.63	2.40	5.20	0.00	8.40
Vi	**r²**	0.62	0.51	0.19	0.65	0.20	0.65	0.60	0.47	0.51	0.41	0.36	0.25	1.00
	b	2.48	3.56	8.13	2.75	7.85	3.20	2.68	6.84	4.77	5.12	7.95	7.59	0.00

Source. data computed by the author

Relative Humidity (*RH*)

Table 11. Correlation results in terms of r² and b among conventional stations as far as Relative Humidity (RH) is concerned

		B	CG	CJ	CP	Gu	It	JF	Lv	PA	Re	SL	Ta	Vi
B	r^2	1.00	0.23	0.34	0.21	0.25	0.40	0.68	0.61	0.39	0.43	0.37	0.31	0.45
	b	0.00	52.40	47.81	52.73	38.70	30.38	8.54	1.77	41.00	22.63	27.79	40.78	35.80
CG	r^2	0.23	1.00	0.07	0.14	0.08	0.49	0.29	0.17	0.23	0.20	0.12	0.10	0.26
	b	21.93	0.00	57.34	45.01	43.35	2.63	9.83	16.39	35.93	21.26	36.02	45.94	29.72
CJ	r^2	0.34	0.07	1.00	0.04	0.38	0.11	0.30	0.47	0.24	0.43	0.42	0.52	0.21
	b	16.20	59.74	0.00	62.58	10.64	44.20	18.31	-8.04	38.23	6.37	5.65	8.26	40.79
CP	r^2	0.21	0.14	0.04	1.00	0.08	0.21	0.28	0.28	0.23	0.20	0.20	0.13	0.22
	b	30.97	52.25	67.85	0.00	47.23	34.81	18.54	9.86	40.49	29.13	29.93	46.51	41.19
Gu	r^2	0.25	0.08	0.38	0.08	1.00	0.11	0.31	0.27	0.24	0.45	0.29	0.56	0.11
	b	37.87	61.18	47.30	61.85	0.00	50.22	29.43	22.16	47.48	18.14	31.66	26.74	57.09
It	r^2	0.40	0.49	0.11	0.21	0.11	1.00	0.46	0.34	0.32	0.27	0.22	0.14	0.49
	b	24.72	37.74	61.66	49.11	48.75	0.00	15.76	14.28	41.87	29.33	36.81	49.54	29.97
JF	r^2	0.68	0.29	0.30	0.28	0.31	0.46	1.00	0.52	0.43	0.48	0.32	0.33	0.41
	b	18.93	51.48	51.85	50.69	38.25	31.29	0.00	12.02	42.49	24.46	34.84	41.79	41.03
Lv	r^2	0.61	0.17	0.47	0.28	0.27	0.34	0.52	1.00	0.40	0.51	0.61	0.41	0.41
	b	29.54	59.89	49.18	54.58	44.64	41.75	29.34	0.00	48.32	29.59	26.60	43.77	46.19
PA	r^2	0.39	0.23	0.24	0.23	0.24	0.32	0.43	0.40	1.00	0.46	0.41	0.38	0.30
	b	14.12	44.00	45.24	42.79	28.96	22.55	4.04	-3.76	0.00	3.77	9.82	25.16	32.96
Re	r^2	0.43	0.20	0.43	0.20	0.45	0.27	0.48	0.51	0.46	1.00	0.50	0.56	0.27
	b	30.38	55.54	44.15	55.17	29.67	40.49	22.55	11.98	40.63	0.00	24.26	31.18	48.12
SL	r^2	0.37	0.12	0.42	0.20	0.29	0.22	0.32	0.61	0.41	0.50	1.00	0.48	0.29
	b	32.82	59.14	46.47	53.25	37.54	42.33	32.03	3.54	41.39	19.56	0.00	34.00	45.31
Ta	r^2	0.31	0.10	0.52	0.13	0.56	0.14	0.33	0.41	0.38	0.56	0.48	1.00	0.17
	b	26.55	58.95	34.93	55.79	11.03	45.31	22.55	4.34	34.40	3.98	10.78	0.00	49.19
Vi	r^2	0.45	0.26	0.21	0.22	0.11	0.49	0.41	0.41	0.30	0.27	0.29	0.17	1.00
	b	14.03	44.95	47.56	43.28	45.11	15.84	11.73	1.54	38.34	24.59	24.14	42.92	0.00

Source. data computed by the author

Wind speed (U_2)

Table 12. Correlation results in terms of r² and b among conventional stations as far as Wind Speed (U2) is concerned

		B	CG	CJ	CP	Gu	It	JF	Lv	PA	Re	SL	Ta	Vi
B	r^2	1.00	0.02	0.00	0.01	0.02	0.01	0.02	0.09	0.01	0.04	0.06	0.08	0.04
	b	0.00	1.57	1.03	1.09	1.25	0.56	2.44	1.81	1.38	0.90	0.79	1.24	1.04
CG	r^2	0.02	1.00	0.00	0.04	0.03	0.00	0.07	0.00	0.10	0.01	0.03	0.03	0.16
	b	1.26	0.00	1.22	0.83	1.20	0.74	2.14	2.25	1.20	1.08	0.89	1.39	0.76
CJ	r^2	0.00	0.00	1.00	0.00	0.01	0.26	0.05	0.00	0.07	0.00	0.01	0.01	0.00
	b	1.32	2.12	0.00	1.24	1.36	-0.01	2.51	2.36	1.82	1.13	0.90	1.86	1.41
CP	r^2	0.01	0.04	0.00	1.00	0.01	0.03	0.08	0.03	0.01	0.05	0.01	0.04	0.03
	b	1.23	1.74	1.05	0.00	1.35	0.32	2.15	2.05	1.47	0.81	0.92	1.39	1.12
Gu	r^2	0.02	0.03	0.01	0.01	1.00	0.01	0.05	0.08	0.00	0.10	0.10	0.07	0.00
	b	1.24	1.59	0.84	1.09	0.00	0.54	2.32	1.92	1.46	0.72	0.73	1.35	1.27
It	r^2	0.01	0.00	0.26	0.03	0.01	1.00	0.02	0.03	0.06	0.05	0.02	0.00	0.00
	b	1.38	1.74	1.05	0.99	1.39	0.00	2.61	2.18	1.50	1.06	1.01	1.61	1.25
JF	r^2	0.02	0.07	0.05	0.08	0.05	0.02	1.00	0.01	0.02	0.04	0.04	0.01	0.05
	b	1.16	1.22	0.62	0.82	1.02	0.44	0.00	2.17	1.35	0.83	0.80	1.60	0.94
Lv	r^2	0.09	0.00	0.00	0.03	0.08	0.03	0.01	1.00	0.01	0.09	0.08	0.09	0.00
	b	0.87	1.68	1.00	0.87	0.85	0.36	2.51	0.00	1.34	0.51	0.63	1.14	1.34
PA	r^2	0.01	0.10	0.07	0.01	0.00	0.06	0.02	0.01	1.00	0.01	0.00	0.07	0.05
	b	1.15	0.82	1.90	0.89	1.29	1.35	2.17	2.02	0.00	0.91	0.98	1.01	0.79
Re	r^2	0.04	0.01	0.00	0.05	0.10	0.05	0.04	0.09	0.01	1.00	0.11	0.09	0.00
	b	1.25	1.65	0.97	0.93	1.13	0.47	2.39	1.98	1.41	0.00	0.83	1.41	1.24
SL	r^2	0.06	0.03	0.01	0.01	0.10	0.02	0.04	0.08	0.00	0.11	1.00	0.08	0.01
	b	1.10	1.50	0.83	1.07	0.98	0.46	2.37	1.89	1.45	0.68	0.00	1.36	1.16
Ta	r^2	0.08	0.03	0.01	0.04	0.07	0.00	0.01	0.09	0.07	0.09	0.08	1.00	0.02
	b	0.95	1.85	1.27	0.87	1.04	0.38	2.58	1.78	1.37	0.64	0.70	0.00	1.14
Vi	r^2	0.04	0.16	0.00	0.03	0.00	0.00	0.05	0.00	0.05	0.00	0.01	0.02	1.00
	b	1.16	1.07	1.13	0.97	1.41	0.79	2.23	2.41	1.28	1.18	0.94	1.50	0.00

Source. data computed by the author

As we can check in Tables 8 and 9, gaps found in *Tmax* and *Tmin* datasets can easily be refilled with data from nearby stations, since the correlation index r^2 goes over 0.70 in all comparisons, pointing out a high assembly between the confronted datasets. That is to say that every gap in *Tmax* or *Tmin* datasets will be completed with values from another station. In both tables, stations whose datasets can be refilled have been distinguished with background gray color filling in the cell and the numeric values in bold. The refilling procedure followed the steps described in section 2 of this same chapter, where, the FAO protocol is described and where we would use Equation 3.

The scenario for *Rs* is slightly different, where CJ and Vi are the only stations that do not match with any other station where to import *Rs* values from, as shown in Table 10. There are no restrictions on this issue for the rest of stations.

As far as *RH* and U_2, no matches at all were found, so there is no way to refill *RH* and U_2 gaps following this procedure.

It is in the current scenario, when other protocols do not meet the requirements, where we propose an alternative path to reach the objective of estimating *ETo* values from incomplete data sets.

Correlations among conventional stations have been possible due to the complete agreement on the extensions of their datasets (01/01/1985 – 31/12/2015). Something that does not happen within automated stations, where each of them has a different installment date, so the datasets comprise different extensions, making their correlation not possible. This is another reason why we propose an alternative way of estimating *ETo* values, which will be fully exposed in the following sections.

4. OUR PROPOSAL: ARTIFICIAL NEURAL NETWORKS (*ANN*)

Artificial Neural Networks (*ANNs*) are mathematical models whose architecture has been inspired by biological neural networks and are considered as very appropriate tools for modelling nonlinear processes. This is the case of evaportanspiration, a process influenced by several climatological parameters which does not behave linearly and that justifies the methodology proposed here and its suitability to it. *ANNs* are capable of learning from examples, recognising repeated patterns and generalising complex relationships among a large amount of data to produce eventually meaningful results, even when input data contains errors or is incomplete, which is the problem this study expects to solve. The main advantage of the *ANNs* is its ability of solving problems which are difficult to formalize (SUDHEER *et al.*, 2005, p. 220). *ANNs* allow us to capture deep complex characteristics of data (GALVÃO *et al.*, 2004, p. 232).

In terms of internal structure and operations, an *ANN* basically consists in three layers: an input, a hidden and an output layer, each of them composed by an array of processing elements (*PE*). A *PE* is a model whose components are analogous to the elements of an actual neuron. An *ANN* is a network where all its components in a given layer are interconnected to all components in the next layer, but not between elements within the same layer. Input data is stored in the input layer, in fact, each input parameter gets into and it is stored in the *ANN* through a neuron or node. That is to say that each parameter is represented by a neuron.

The function of this first layer is to provide information to the network. At the entrance of the hidden layer, all input parameters randomly receive a weight, which ranks them in terms of importance according to the model the *ANN* is trying to simulate. This represents the first part of a processing element; the second part consist in a nonlinear filter, usually called “transfer function”. Its aim is to limit the output values between two asymptotes. The most common transfer function is the sigmoidal function. It is a function that varies gradually between two asymptotic values, typically 0 and 1 or -1 and +1.

The hidden layer is actually which allows the network to model complex functions. The number of nodes compounding the hidden layer is determined by trial and error. There could be more than one hidden layer but the use of a single hidden layer along with a sigmoidal

transfer function is widely recognised as the most frequent network topology (CYBENKO, 1989, p. 314; HORNIK *et al.*, 1989, p. 366)

Finally, the output layer can be understood as "the exit door" for values predicted by the network and it is composed only by one node, the output. The type of network topology described here is known as Multilayer Perceptron (*MLP*) (fausset, 1994, p. 226). Within the *ANN*, as the sign spreads forward layer-by-layer (feed-forward) the error values propagate backwards (backpropagation) for a better adjustment of the weights (synaptic weight adjustments). Three different algorithms can be applied: Quasi-Newton (*Q-N*) (HAYKIN, 1994, p,. 321); Levenberg-Marquardt (*L-M*), (HAGAN; MENHAJ, 1994, p. 993); Backpropagation with variable learning rate (*BPVL*) (SOARES; NADAL, 1999, p. 213).

An extensive review about several *ANN* applications in hydrology field is gathered in ASCE Task Committee (2000) and (GOVINDARAJU; RAO, 2000, p. 165). Furthermore, recent studies on *ANNs* applications in this area include rainfall–runoff modelling (WILBY *et al.*, 2003, p. 181; LIN; CHEN, 2004, p. 8); river stage forecasting (IMRIE *et al.*, 2000, 183; LEKKAS *et al.*, 2001, p. 164; CAMPOLO *et al.*, 2003, p. 398); reservoir operation (JAIN *et al.*, 1999, p. 271); land drainage design (LINGIREDDY, 1998, p. 144); describing soil water retention curve (SUDHEER; JAIN, 2004, p. 844) and optimization or control problems (WEN; LEE, 1998, p. 436; BHATTACHARYA *et al.*, 2003, p. 245). Some of those studies have also shown that *ANNs* are even more accurate than conventional methods in flow forecasting and drainage design (YANG *et al.*, 1996, p. 533; ZEALAND *et al.*, 1999, p. 48).

ANNs can be also applied for estimating *ETo*. Several studies have been carried out during the past decade aiming the estimation of evapotranspiration through the application of *ANNs* (JOTHIPRAKASH *et al.*, 2002, p. 20; TRAJKOVIC *et al.*, 2003, p. 457; KISI, 2006, p. 646; 2006, p. 260; 2006, p. 1105; 2007, p. 1934; ZANETTI *et al..*, 2007, p. 89; JAIN *et al.*, 2008, p. 234; KISI, 2008, p. 260; CHAUHAN; SHRIVASTAVA, 2009, p. 837; DAI *et al.*, 2009, p. 442; IZADIFAR; ELSHORBAGY, 2010, p. 425)

KUMAR *et al.* (2002, p. 512) concluded that *ANNs* trained with lysimetric data performed better than *PM* equation in terms of *ETo* estimates. SUDHEER *et al.* (2005, p. 23) used a radial basis function neural network for estimating actual *ET* from limited data and compared with lysimeter *ETo* values. KISI (2006, p. 161) used feed-forward neural networks with Levenberg-Marquardt training algorithm for estimating *ETo*. Results were compared with

those obtained through linear regression. ZANETTI *et al.* (2007, p. 323) applied Levenberg-Marquardt training algorithm in *ANNs* for the estimation of *ETo* with minimum and maximum temperatures. *ANNs* have also been used for forecasting monthly *ETo* values (TAHIR, 1998, p. 546; TRAJKOVIC *et al.*, 2003, p. 457).

All these studies are somewhat related to our proposal since they point out *ANNs* as useful tools at estimating *ETo*. Consequently, we also expect to get accurate *ETo* values when applying *ANNs* onto our incomplete datasets. Thus, the scope of this chapter is to demonstrate that *ANNs* are an efficient methodology to estimate *ETo* values at high performance resulting in accurate results even when the dataset is incomplete, as it is in our case.

4.1 Estimating *ETo* through *ANNs*

The same weather datasets described in chapter I, section 10 (Study area and dataset) were used to apply different *ANNs* built in order to test its potentiality for estimating *ETo*. Each daily measure, which comes into the network through the input layer, is considered as a pattern of the climatic behaviour within the study area and it is from this dataset where the network learns from.

As mentioned above, it is usual to find missing values within large datasets and we were not an exception. That is to say that the amount of values available to run the networks are diverse in number and were specified back in chapter I, Table 4 (Study area and dataset section) This means that networks for different stations were built, trained, validated and tested by using different amounts of values. Moreover, all datasets were treated individually, which means that the procedures described in the following were applied to each dataset.

- Gaps were previously removed from each dataset, so we used only recordings where the 5 climatological parameters (used afterwards as inputs) were available in a way that we also could compute *ETo* by means of Penman-Monteith FAO56 equation.
- 70% of it was destined for training; during this phase, the network is trained to associate outputs (*ETo*(*PM*)) with input patterns (*Tmax*, *Tmin*, *RH*, *Rs*, U_2,). This process is usually known as Learning Process, as it requires a memorization process of the wide variety of input patterns and its associated outputs.

- The 30% spare was divided into two equal parts and used for validation and testing, respectively; during validation, the network emulates what it learned before and tries to perform the best on new patterns. It is here where the most important application of neural networks takes place: Pattern Recognision; during this step, the network identifies the input pattern and tries to produce the associated output. The power of neural networks makes sense when a pattern that has no output associated with it, is given as an input (testing phase). In this case, the network gives the output that corresponds to a taught input pattern that is least different from the given pattern. That is to say that during the testing phase the network estimates *ETo* values basing on what it learnt previously during the Learning Process.

4.2 Building and Running an *ANN*

Our whole dataset is composed by 26 weather stations as mentioned in chapter I, section 10. However, for workload and efficiency reasons we chose JF as a representative station to perform different *ANNs* aiming to discern which architecture is sufficient or most suitable to estimate *ETo* values with enough accuracy. Those architectures were adopted and run afterwards on the rest of the stations.

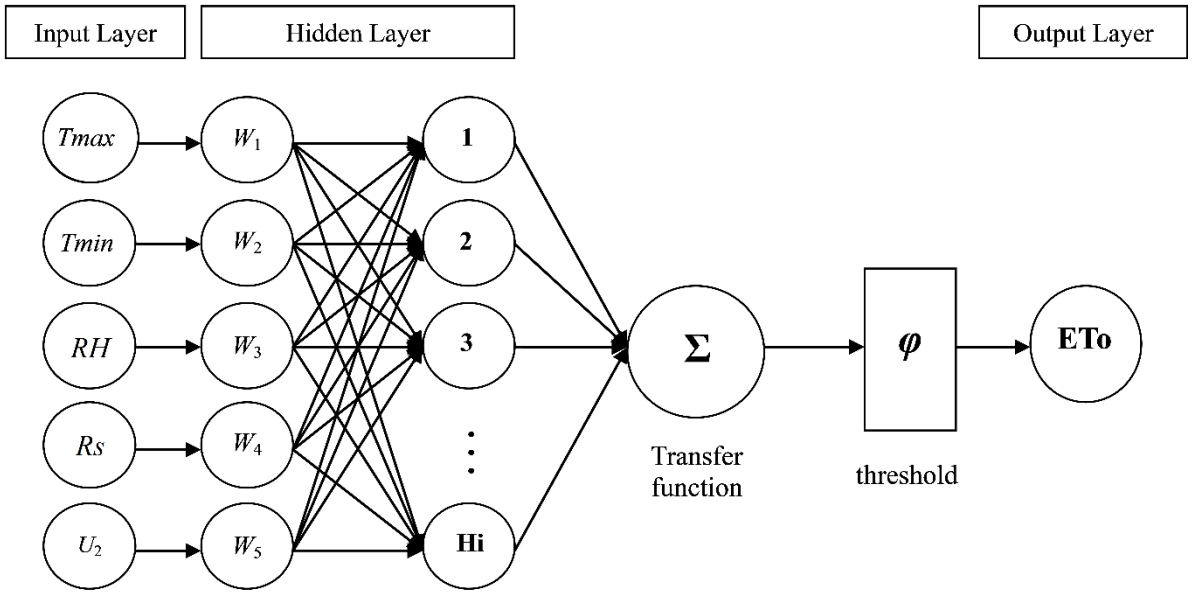
So, firstly, we built *ANNs* with an input layer composed by 5 input parameters: maximum air temperature (*Tmax*), minimum air temperature (*Tmin*), solar radiation (*Rs*), mean relative humidity (*RH*) and mean wind speed (U_2). Latitude (decimal degrees) and altitude (meters above sea) were excluded from the study as they were considered as constants. The function of this first layer is to provide information to the network.

Secondly, as far as the nodes that compound the hidden layer (Hi in Figure 2) were concerned, networks with 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45 and 50 nodes were trained. The number of nodes in the hidden layer is a recurrent issue as there is no an “absolute truth” about it (COULIBALY *et al.*, 2000, p. 213). Indeed, some studies on the issue revealed that large-than-necessary networks tend to over-fit the training samples and bring poor performances. Some researches demonstrated that one hidden layer is enough to represent the non-linear relationship between the climatic parameters and *ETo* (ARCA *et al.*, 2001, p. 97; KUMAR *et al.*, 2002, p. 233). At the entrance of the hidden layer, all input parameters

randomly received a weight (Wij in Figure 9), which ranked them in terms of importance accordingly to the model the *ANN* is trying to simulate.

Finally, the output layer was only composed by one node instead: daily *ETo* values estimated by PM method *(ETo(PM))*, considered as the target values. The complete architecture just described is depicted in Figure 9.

Figure 9. Basic architecture of an Artificial Neural Network adapted to the problematic treated along this second chapter



Source. Adaptation of a standard architecture of an Artificial Neural Network taken from <https://www.kdnuggets.com/>

To accomplish this task, the software Matlab R2015a was used, both for applying Penman-Monteith method to estimate *ETo* values and for building and running of *ANNs*. Aiming to assess the *ANNs* performance when estimating *ETo* values (*ETo*(*ANN*)), the following statistical indices were applied:

$$MSE = \frac{\left(\sum_{i=1}^{n}\left(x_i - y_i\right)^2\right)}{n} \quad \textbf{(56)}$$

$$Y = mX + b \quad \textbf{(57)}$$

$$r^2 = \left(\frac{\frac{1}{n}\sum_{i=1}^{n}(x-\bar{x})(y-\bar{y})}{\sigma_x \sigma_y} \right) \quad \textbf{(58)}$$

where:

MSE = Mean Squared Error (mm day^{-1});

n = number of observations;

x_i = *ETo* (mm day^{-1}) estimated by *PM* (*ETo*(*PM*));

y_i = *ETo* (mm day^{-1}) estimated by the *ANNs* (*ETo*(*ANN*));

r^2 = *determination coefficient*;

x = mean of x_i;

y = mean of y_i;

σ_x = standard deviation of x_i ;

σ_y = standard deviation of y_i;

Y = estimatted values (*ETo*(*ANN*));

X = target values (*ETo(PM)*);

m and *b* = *slope* and *intercept* of the line of best fit between *ETo*(*PM*) and *ETo*(*ANN*).

Just recalling, this second chapter began with the target of estimating *ETo* by means of an alternative methodology instead of using the standard method, Penman-Monteith FAO 56 equation as there were climatic parameters missing that did not allow the calculation of *ETo*. However, up to now, *ANN's* were built putting in place all parameters required to estimate *ETo*, but this is far from the reality as we just mentioned. So, another challenge comes up: estimate *ETo* values only with part of the parameters needed to do so.

After a prospective analysis of our datasets, we confirmed that *RH* and U_2 were constantly lacking in the historical register used along this study. The lack of these two climatic parameters throughout our dataset was actually the reason why we considered *ANNs* as an

approach to estimate *ETo* since Penman-Monteith method cannot be applied when there are missing inputs.

So, in order to imitate what we have found throughout previous prospective analyses, *ANNs* were built removing *RH* and U_2 from the input layer, as a one-at-a-time process in order to check whether it was still possible to estimate *ETo* at a high resemblance level. Thus, once *RH* and U_2 inputs were removed, *ANNs* performances were compared with *ETo*(*PM*) by means of some statistical indices which allowed us to quantify the adjustment level between target values (*ETo*(*PM*)) and forecasting values (*ETo*(*ANN*)) as it follows**:**

$$c = dr \tag{59}$$

$$d = 1 - \frac{\sum_{i=1}^{n}\left(y_i - x_i\right)^2}{\sum_{i=1}^{n}\left(\left|y_i - \bar{x}\right| + \left|x_i - \bar{x}\right|\right)^2} \tag{60}$$

$$r = \frac{\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)}{\sigma_x \sigma_y} \tag{61}$$

where:

c = performance index (CAMARGO; SENTELHAS, 1997, p. 321);

d = adjustment coefficient (WILLMOTT, 1981, p. 546);

r = correlation coefficient.

5. RESULTS OF TESTING SEVERAL ARCHITECTURES ON JF DATASET

Different networks were built with a variety of nodes in the hidden layer (*ANN* architecture*, Table 13) and the statistical indices described above are shown in Table 13 for each of them. *MSE* values account for the average difference between the target and the predicted values expressing the quality of an estimator, in this case, the *ANN*. The closer to zero value, the better is the estimator and its performance. Learning cycles (also called number of epochs) refers to the number of trials the network performs to reduce the differences previously mentioned to a minimum.

Table 13. ANN's performance indices r^2, m, b and comparison between ETo(PM) and ETo(ANN) estimated values

ANN* architecture	***MSE* validation**	**Learning cycles**	**r^2**	***m***	***b***	***ETo*(*PM*)-*ETo*(*ANN*)**
5_1_1	0.112	128	0.880	0.88	0.37	-0.052
5_2_1	0.089	162	0.899	0.89	0.33	0.010
5_3_1	0.074	34	0.906	0.91	0.28	0.074
5_4_1	0.083	13	0.906	0.91	0.29	0.057
5_5_1	0.079	20	0.910	0.91	0.28	0.051
5_10_1	0.146	10	0.935	0.93	0.24	-0.055
5_15_1	0.157	7	0.937	0.93	0.23	-0.068
5_20_1	0.156	20	0.937	0.92	0.24	0.028
5_25_1	0.159	17	0.935	0.91	0.26	0.066
5_30_1	0.145	6	0.929	0.95	0.19	0.037
5_35_1	0.158	13	0.931	0.92	0.30	-0.035
5_40_1	0.160	5	0.937	0.93	0.25	0.053
5_45_1	0.148	6	0.933	0.93	0.27	0.023
5_50_1	0.158	7	0.935	0.92	0.26	-0.073

Source: data computed by the author

Note. * ANN architecture: They all were built based on this topology: In_Hi_Op, where In = parameters in the input layer; Hi = nodes in the hidden layer; Op = output. *ETo*(*PM*)-*ETo*(*ANN*): difference between *ETo* values estimated by *PM* method and *ETo* values predicted by the *ANN*.

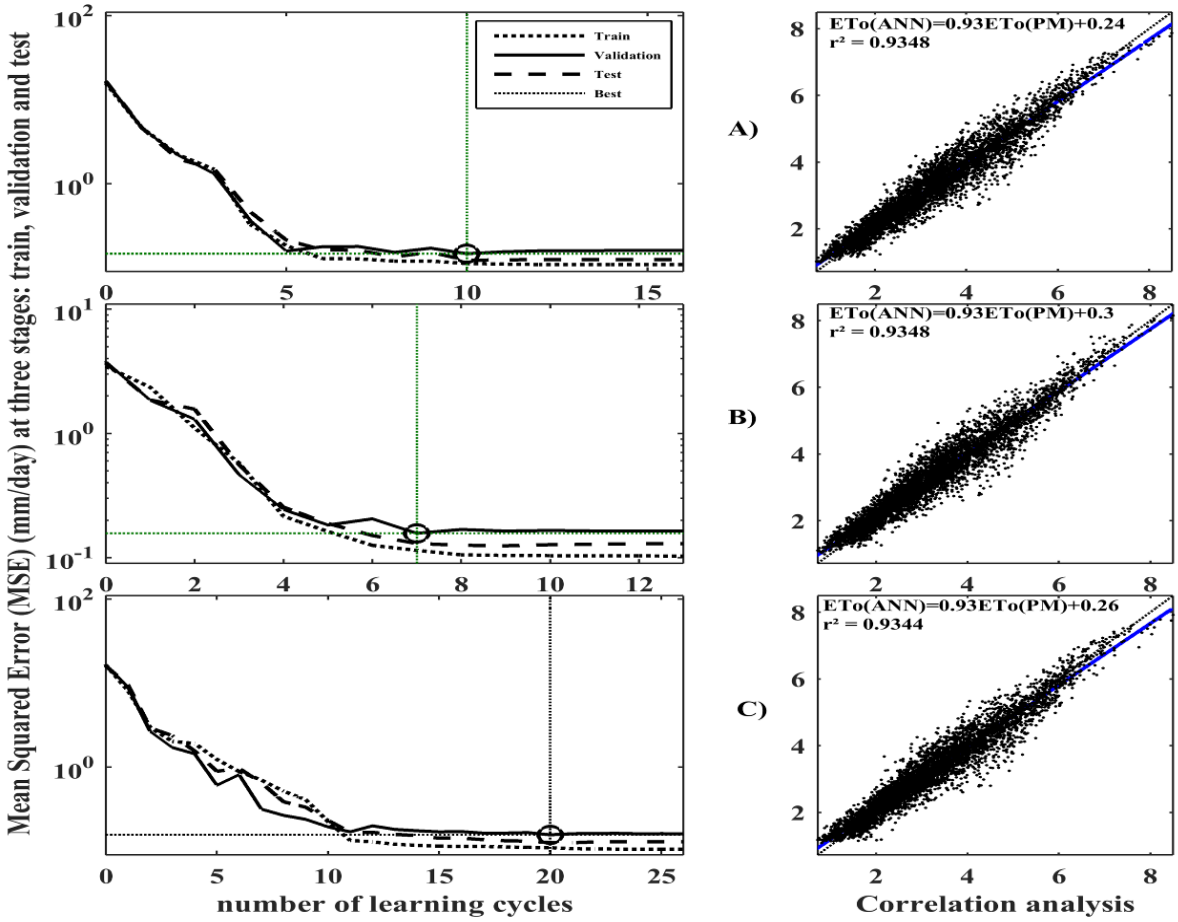
Architectures with 15 and 20 nodes in the hidden layer presented the highest r^2 values, 0.937 accompanied with a value of 0.935 for networks with 10 and 25 nodes in the hidden layer

as is shown in Table 6. In other words, networks containing 10, 15, 20 and 25 hidden nodes could be considered as those reaching the minimum *MSE* values along with the highest r^2 values. In fact, except networks with 1 and 2 nodes in the hidden layer, all r^2 values obtained were above of 0.90 which is considered by most authors as an excellent *ANN* performance.

According to *MSE* values, networks with 2, 3, 4 and 5 neurons would be apparently the most appropriate for *ETo* estimates, as they present the lowest *MSE* values: 0.089, 0.074, 0.083 and 0.079, respectively. Furthermore, several authors used *MSE* as the criterion to assess *ANNs* topology (KUMAR *et al.*, 2002, p. 233; ZANETTI *et al.*, 2007, p. 89; JAIN *et al.*, 2008, p. 234).

However, if we focus only on *MSE* values, the results obtained in this study and shown in Table 6 can lead us to a misunderstanding since the same networks showing the minimum *MSE* values also needed greater number of learning cycles and larger r^2 values. Although, 10, 15 and 20 hidden-nodes networks brought *MSE* values of 0.146, 0.157, 0.156, respectively along with less learning cycles required and much higher r^2 values, so these architectures were considered as the most suitable to estimate *ETo* for JF station and their performances are shown in Figure 10.

Figure 10. Performances of 10, 15 and 20-hidden-nodes architectures applied on JF station. Performance in terms of MSE, numbers of learning cycles needed to reach the minimum MSE value and determination coefficient (r^2) from correlation analyses for (A)



Source: plots computed by the author

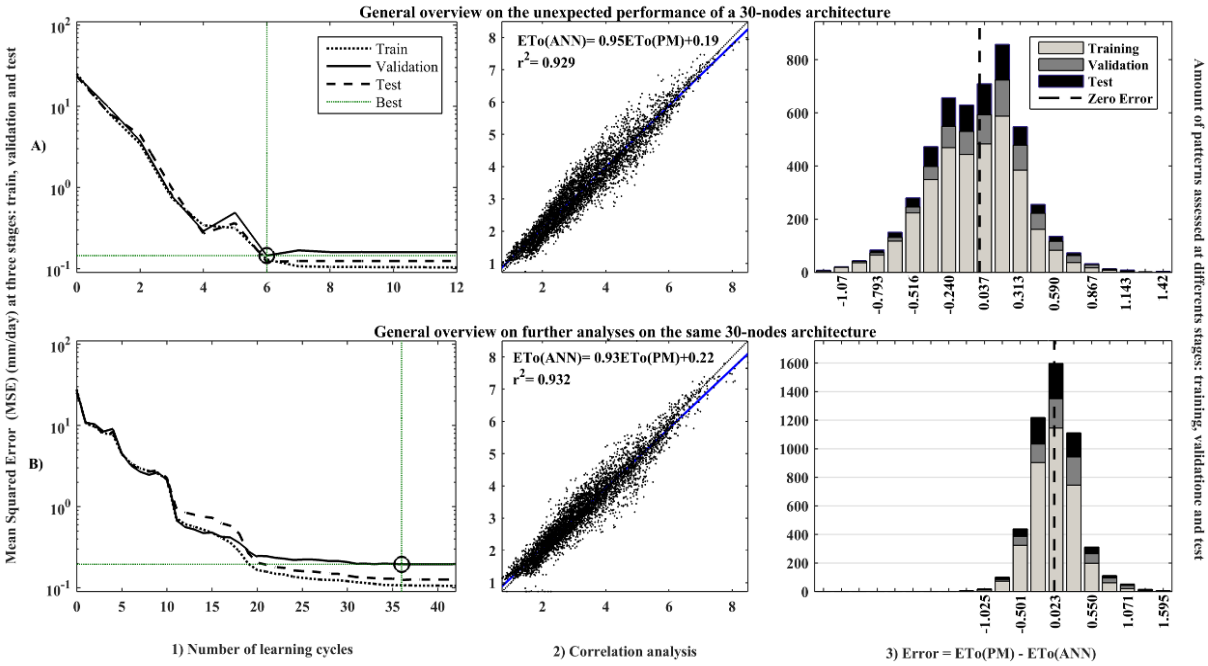
Still focused on *MSE* values, we can check in Table 6 that an extraordinary value of 0.145 was obtained for a 30-hidden-nodes network. In concordance with it, the number of learning cycles required to reach the minimum error at validation stage gave us a raw idea of which *ANN* architectures neither over-fit the data nor underestimate the results.

As far as number of learning cycles is concerned, networks with 10, 15, 30, 40, 45 and 50 were ranked as the ones with less completed learning cycles: 10, 7, 6, 5, 6 and 7, namely. Again, 30-hidden-nodes network brought an unexpected result; 6 learning cycles were required to reach such a low *MSE* value as it is shown in Table 13 and clearly visible in Figure 11, A1. In addition, *slope* (*m*) and *intercept* (*b*) were also displayed in Table 6 so as to confirm the outstanding behaviour of 30-nodes network, which delivered remarkable values of 0.95 and 0.19, corresponding to the maximum *slope* and the minimum *intercept* among the architectures assessed. On the other hand, the r^2 value for this network (0.929) was actually the poorest

among the architectures under study. Further analyses were carried out in order to determine wether 30-hidden-nodes network was, indeed, the best architecture or just an outlier.

After this unexpected performance of a 5_30_1 network, we carried out some further analyses. Training, validating and testing again the same 5_30_1 topology brought different results: now the network needed 36 learning cycles (6 times more) to reach the minimum *MSE*, which changed from 0.145 (Figure 11, A1) to a value of 0.195 (Figure 11, B1), the highest among the architectures under study. The difference between the *ETo*(*PM*) and *ETo*(*ANN*) for JF station can be checked in Figure 11, A3 and B3. It is presented as a histogram where all patterns were clustered in 20 classes or bins around the line of zero error. It gave us an idea on how accurate *ANNs* can be, since most bins yielded a difference below 0.1 mm day^{-1} for JF station.

Figure 11. Unexpected behaviour of a 30-hidden-nodes architecture and results of further analyses on JF dataset. Comparison between (A) the first set of results of a 5_30_1 network and (B) further analyses. 1) Number of learning cycles required to reach the minimum MSE



Source: plots computed by the author

Slope and *intercept* moved to completely opposite extreme values: 0.93 and 0.22, namely, the lowest *slope* and the highest *intercept* (Figure 11, B2). Surprisingly, r^2 slightly increased until 0.932 (Figure 11, B2) and *ETo*(*PM*) - *ETo*(*ANN*) decreased meaningfully from 0.037 (Figure 11, A3) to 0.023 (Figure 11, B3).

Taking into account these results and probably requiring deeper analyses, we cannot consider a 30-hidden-nodes architecture as a trustworthy network, due to the poor *slope* and *intercept* values, the high *MSE* values and the large amount of learning cycles to reach it.

5.1 Estimating *ETo* with missing values of *RH* and U_2 on JF dataset

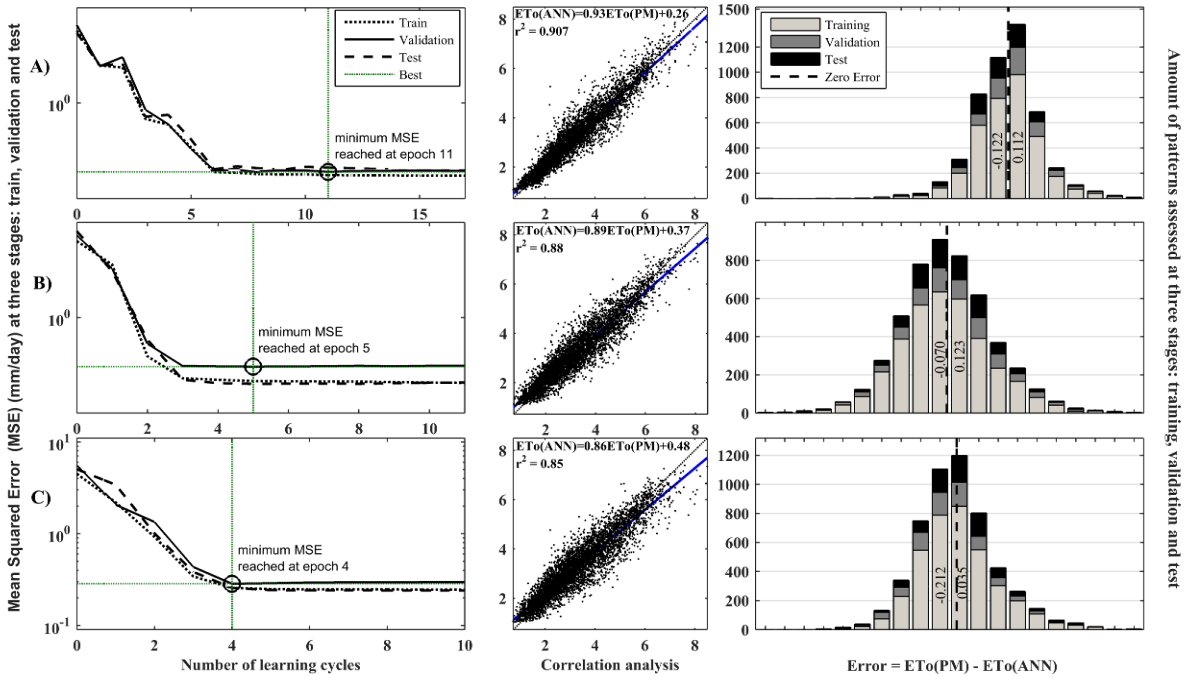
Here, we built *ANNs*, still on JF dataset as a test, with 10, 15, 20, 30 nodes in the hidden layer, 1 output only (*ETo*(*PM*)) and with no U_2, as input (4U_2_10/15/20/30_1); with no *RH* as input (4RH_10/15/20/30_1); with neither U_2, nor *RH* as inputs (3_10/15/20/30_1); and finally with only *Rs* as input (1_10/15/20/30_1). Their performances were assessed by using some statistical indices as performance index (*c*), adjustment coefficient (*d*) and regression coefficient (*r*), previously presented in section 4.2 (Equations 59, 60 and 61).

Table 14. Performance indices from architectures with missing inputs still on JF dataset

Missing inputs	Architectures	c	d	r
None	5_10_1	0.95	0.983	0.967
	5_15_1	0.952	0.983	0.968
	5_20_1	0.952	0.983	0.968
	5_30_1	0.948	0.982	0.965
U_2	4_10_1	0.928	0.975	0.952
	4_15_1	0.929	0.976	0.952
	4_20_1	0.928	0.975	0.952
	4_30_1	0.921	0.972	0.948
RH	4_10_1	0.907	0.967	0.938
	4_15_1	0.908	0.968	0.939
	4_20_1	0.909	0.968	0.939
	4_30_1	0.909	0.968	0.939
U_2, RH	3_10_1	0.888	0.960	0.925
	3_15_1	0.890	0.961	0.927
	3_20_1	0.890	0.960	0.926
	3_30_1	0.890	0.961	0.927
Tmax, Tmin, U_2, RH	1_10_1	0.664	0.859	0.773
	1_15_1	0.665	0.860	0.774
	1_20_1	0.663	0.858	0.773
	1_30_1	0.666	0.86	0.774

Source: data computed by the author

Figure 12. General overview on the performance of networks with missing inputs on JF dataset. Effect of removing some input parameters in the performance of (A) 4U2_15_1; (B) 4HR_30_1; (C) 3_30_1 topologies in terms of MSE (left column), correlation analyses between *ETo(PM)* and *ETo(ANN)* (central column) and Error (right column) measured in mm day^{-1}, which accounts for the difference between the target *ETo*(*PM*) values and the predicted values *ETo*(*ANN*) values



Source: plots computed by the author

The comparisons brought outstanding *c* and *d* values as depicted in Figure 12: 0.929 and 0.976, respectively for a 4U$_2$_15_1 architecture; 0.909 and 0.968 for a 4RH_30_1 architecture; 0.890 and 0.961 for a 3_30_1 architecture. 1_10/15/20/30_1 network was rejected due to its low *r*, *c* and *d* values.

That means that an *ANNs* with only three input parameters (*Tmax*, *Tmin* and *Rs*) can offer excellent adjustments (CAMARGO; SENTELHAS, 1997, p. 327) as it is visible in Figure 12C. According to what it can be inferred from Figure 4 and confirmed by *c* and *d* values in Table 2, *ETo* values estimated by *ANNs* are highly resembled to those estimated by *PM* method, even when both U_2 and *RH* are not available.

Once we have tested different architectures on JF dataset and proved the convenience of some architectures over others, even with *RH* and U_2 as missing inputs, we are now in condition to spread the application of such *ANN* approach to the rest of datasets in the following section.

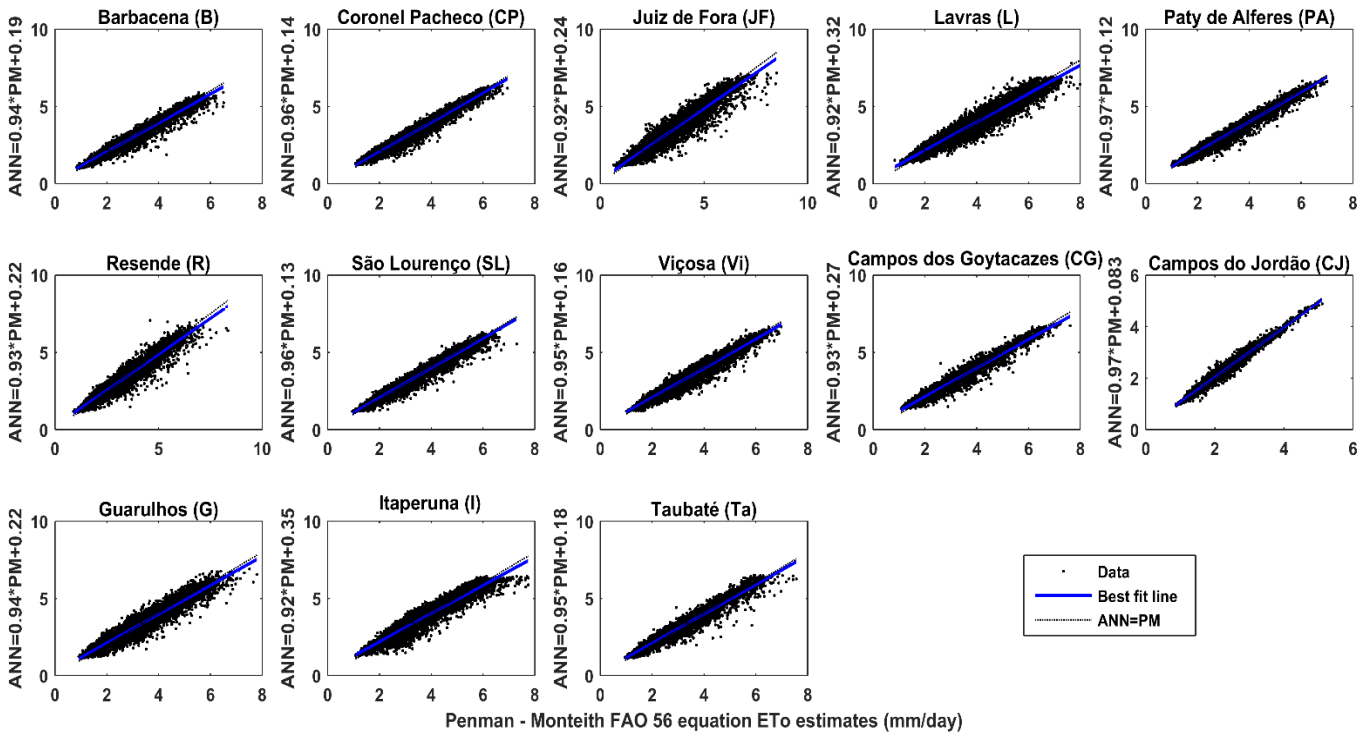
6. RESULT OF WIDENING *ANNS* APPROACH TO ALL STATIONS UNDER STUDY

An *ANN* with an architecture fitted only by *Tmax*, *Tmin* and *Rs* as inputs and 15 nodes in the hidden layer (3_15_1 architecture) was built, trained, validated and tested for each station individually afterwards. Again, the 26 stations and their corresponding data sets were clustered into 2 classes: conventional and automated weather stations, which in turn were subdivided into two clusters each, aiming to better visualize the results: cluster 1 and 2 and clusters 3 and 4, respectively. The results of applying the *ANN* just mentioned above for each dataset is shown in terms of correlation analyses, performances and error histograms in the following.

6.1 CONVENTIONAL STATIONS CLASS

6.1.1 Correlation analysis results

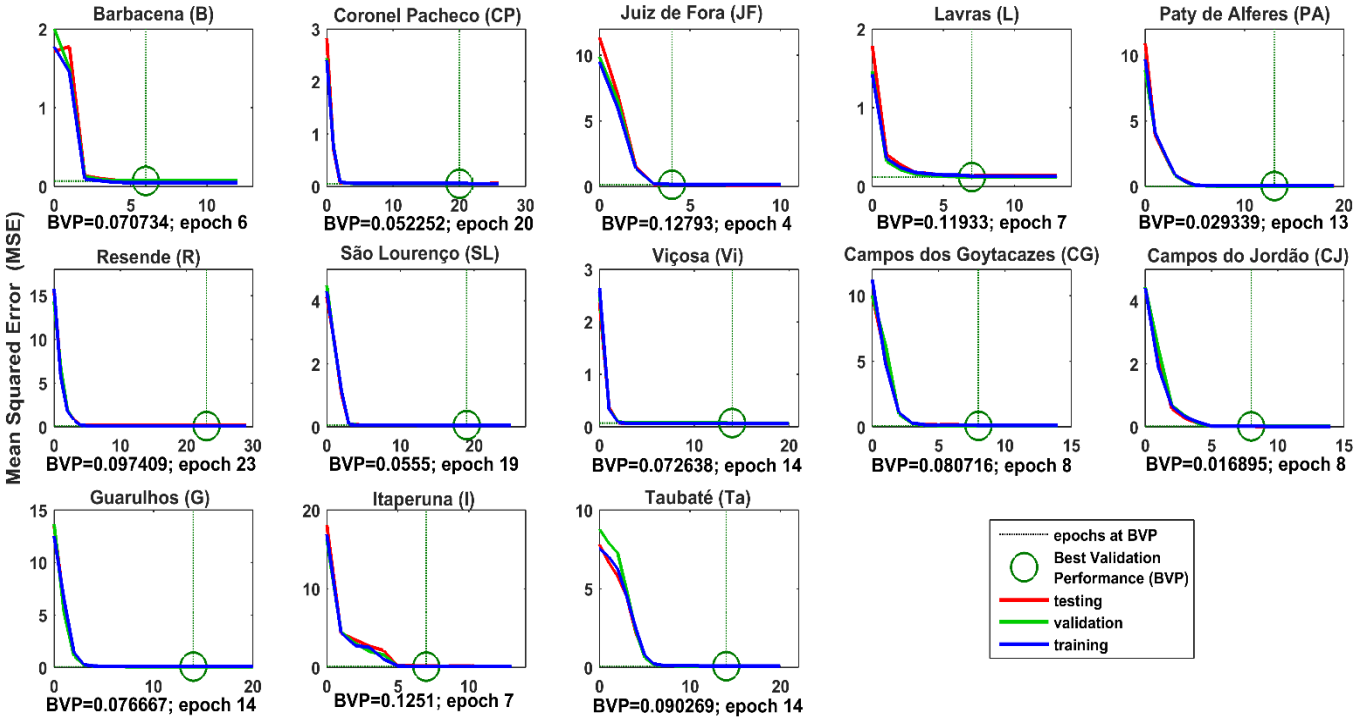
Figure 13. Correlation analysis applied on all the stations belonging clusters 1 and 2 showing a high degree of similarity between ETo(PM) and ETo(ANN)



Source: plots computed by the author

6.1.2 *ANN's* performance

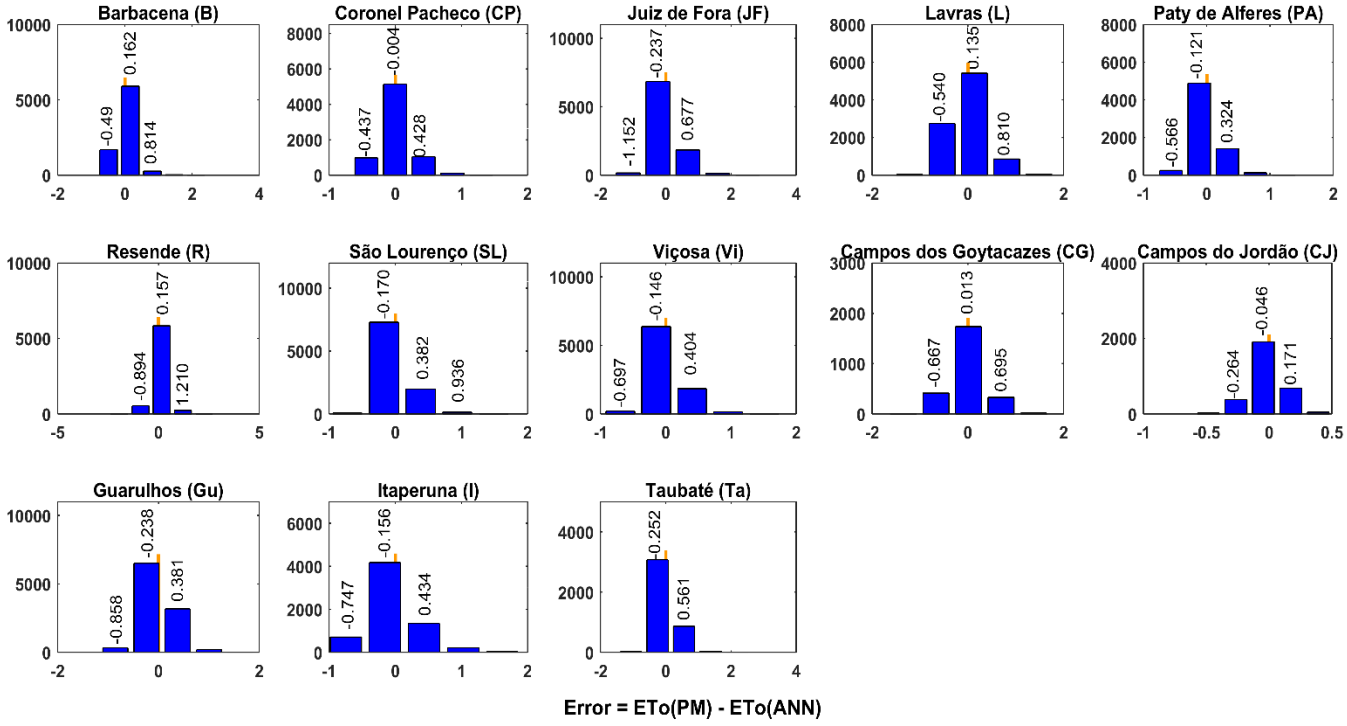
Figure 14. ANN's performance and number of epochs needed to reach the minimum MSE between ETo(PM) and ETo(ANN) for all the stations belonging to clusters 1 and 2



Source: plots computed by the author

6.1.3 Error histograms

Figure 15. Error histogram measured in mm day^{-1} comparing ETo(PM) and ETo(ANN) values for all the stations belonging to clusters 1 and 2

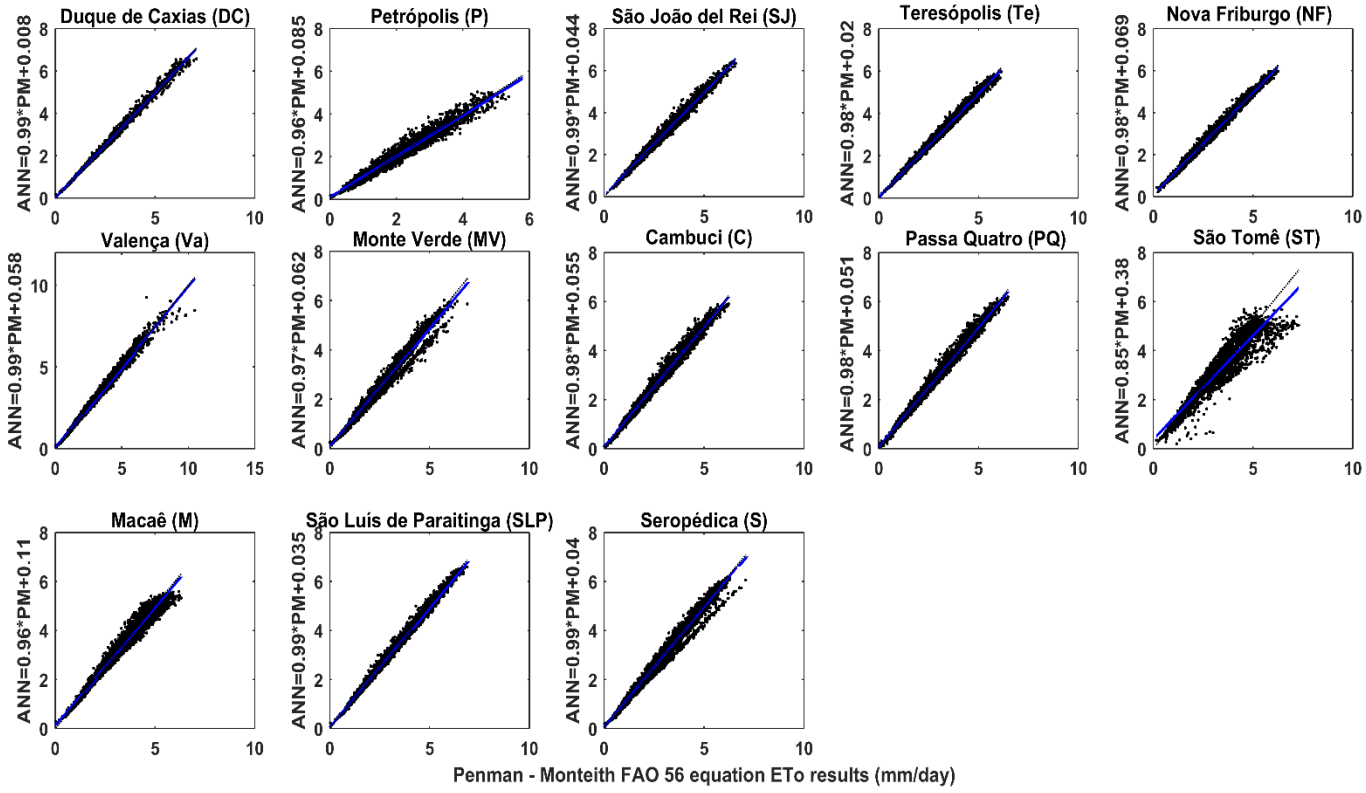


Source: plots computed by the author

6.2 AUTOMATED STATIONS CLASS

6.2.1 Correlation analysis results

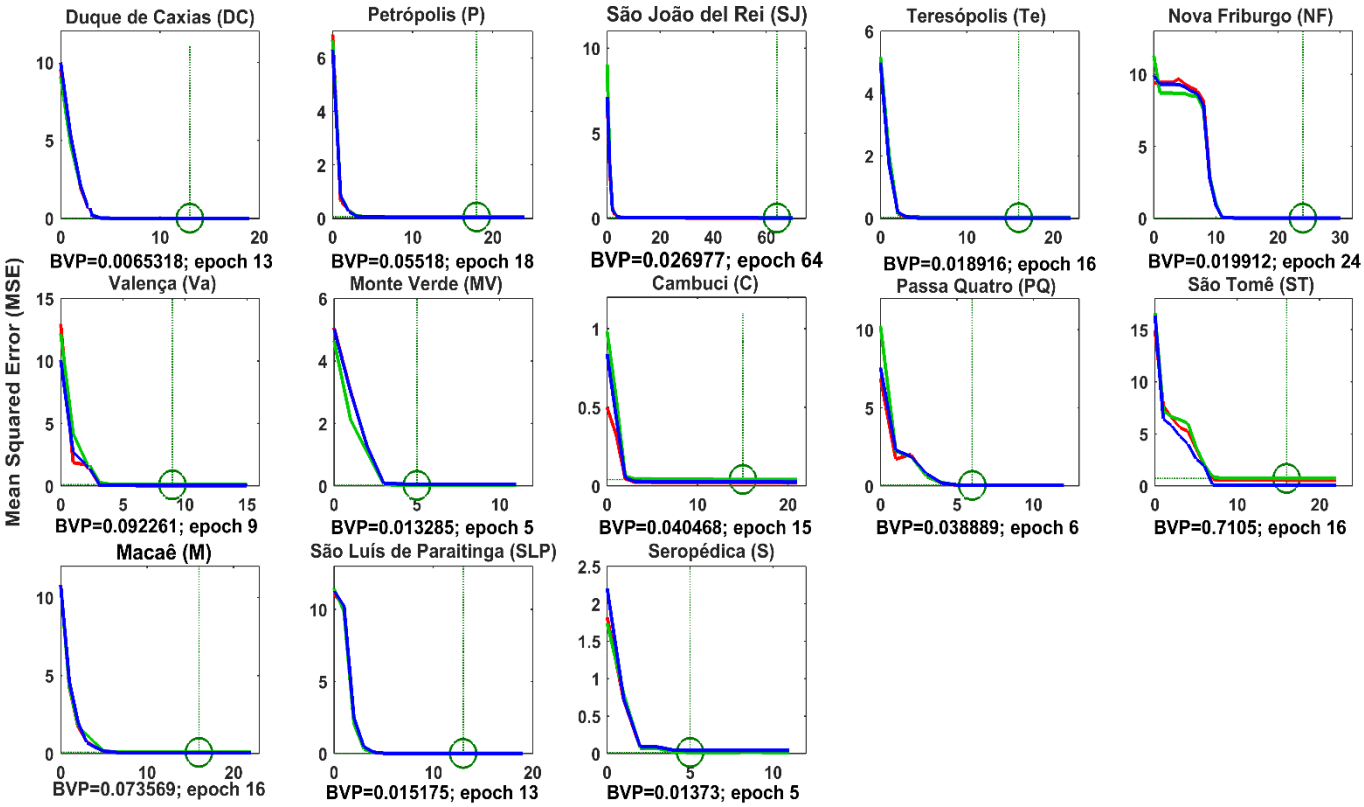
Figure 16. Correlation analysis applied on all the stations belonging clusters 3 and 4 showing a high degree of similarity between ETo(PM) and ETo(ANN)



Source: plots computed by the author

6.2.2 *ANN's* performance

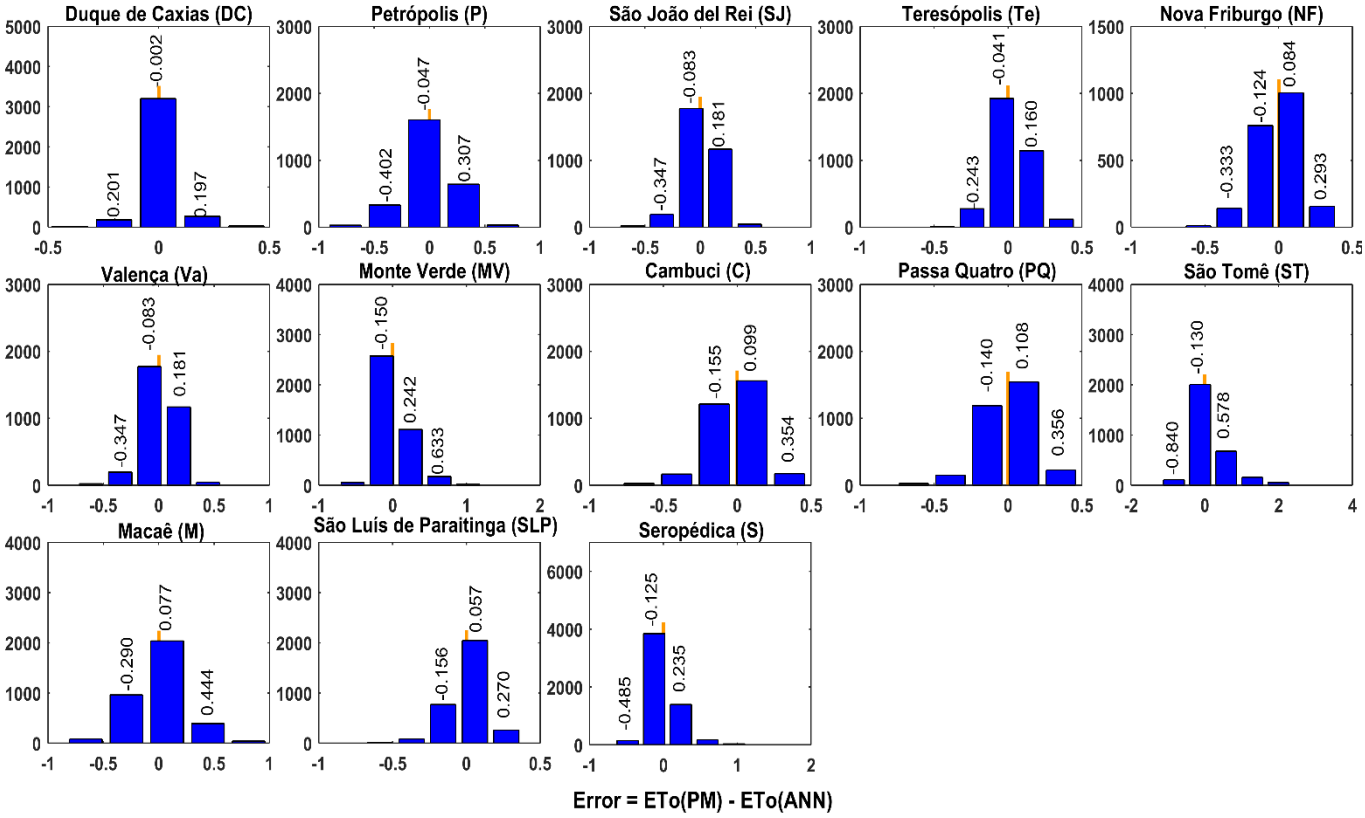
Figure 17. ANN's performance and number of epochs needed to reach the minimum MSE between ETo(PM) and ETo(ANN) for all the stations belonging to clusters 3 and 4



Source: plots computed by the author

6.2.3 Error histograms

Figure 18. Error histogram measured in mm day^{-1} comparing ETo(PM) and ETo(ANN) values for all the stations belonging to clusters 3 and 4



Source: plots computed by the author

The importance of these results and its proper reading have demonstrated that *ANNs* approach can be an efficient tool to estimate *ETo* which can also be applied for water resources management, irrigation schedule or crop irrigation systems design.

As we can check in Figure 13, regression r^2 and *intercept* coefficient *b* among conventional stations remained steadily over 0.92 (except stations L, JF and I) and below 0.35 (except station L), respectively. In regard to performance, Best Validation Performance (BVP) stays continuously below 0.128, a negligible difference when compared with our reference, *ETo(PM)* as shown in Figure 14. Values considered outstanding in literature by CAMARGO and SENTELHAS (1997, p. 327).

In Figure 16, regression r^2 and *intercept* coefficient *b* among automated stations keep their values over 0.96 (except stations P and M) and below 0.11 (except station M) respectively, with the exception of ST whose values reached 0.85 and 0.38 for r² and b, namely. As far as

performance is concerned, Best Validation Performance (BVP) stays continuously below 0.09 as depicted in Figure 17.

7. SPEAKING IN TERMS OF WATER COLUMN HEIGHT

As far as error histograms are concerned, its proper interpretation does not reduce to a plain results reading, but a water-related interpretation. Since this study has had *ETo* as the leading role of the story and it can even be thought, described and expressed as a water volume, we thought error histograms results could be more clearly expressed and interpreted in terms of difference in water column height. With it, we meant to demonstrate that the amount of evapotranspirated water predicted by *ANN's* method was outstandingly similar to those estimated by *PM* method. Those differences can be checked in Table 15.

Table 15. Differences between target (ETo(PM)) and predicted values (ETo(ANN)) interms of water column height. Mean daily target height: mean daily milimeters of water estimated by PM method for the whole study period. Mean daily predicted height: mean

Stations		Mean daily target height (mm day^{-1})	Mean daily predicted height (mm day^{-1})	Mean daily difference (mm day^{-1})	Total difference (mm)
Conventional	**B**	2.92	2.93	-0.01	-45.65
	CP	3.33	3.33	0.00	9.80
	JF	3.19	3.19	0.00	34.50
	L	3.83	3.83	0.00	-27.81
	PA	3.39	3.39	0.00	-14.84
	R	3.24	3.24	0.00	4.05
	SL	3.34	3.34	0.00	-12.53
	Vi	3.22	3.22	0.00	4.71
	CG	3.57	3.57	0.00	-2.82
	CJ	2.33	2.33	0.00	-7.07
	G	3.13	3.13	0.00	-28.09
	I	3.52	3.52	0.00	8.08
	Ta	3.30	3.30	0.00	12.99
Automated	**DC**	3.07	3.13	0.01	26.39
	P	2.05	2.16	-0.01	-20.38
	SJ	3.13	3.15	0.01	16.26
	Te	2.88	2.88	0.02	84.51
	NF	2.92	2.93	0.01	16.45
	Va	3.17	3.22	-0.04	-148.59
	MV	2.98	2.98	0.00	0.28
	C	2.75	2.75	0.00	-0.72
	PQ	2.89	2.89	0.00	-0.27
	ST	3.40	3.40	0.00	3.12
	Ma	3.18	3.18	0.00	-0.56
	SLP	3.35	3.35	0.00	0.51
	S	3.01	3.01	0.00	-3.48

As we can check in Table 15, mean daily predicted height barely does not vary at all from *ETo*(*PM*) values, which has been a surprising result, even for us. Furthermore, the total difference is still low, with a maximum mean accumulated difference of 148.60 mm. for the whole period under study, which barely encompassed 10 years (27/09/2006 - 30/09/2016); in other words, an annual mean difference of 14.86 mm. Now it is clearer to perceive the high accuracy of the method proposed throughout this chapter.

8. CONCLUSION

A neural network with only three inputs (*Tmax*, *Tmin* and *Rs*) in the input layer, no more than 20 neurons in the hidden layer, fitted with a hyperbolic tangent sigmoid transfer function and guided by the Levenberg-Marquardt algorithm was sufficient to reach outstanding results. The resemblance level was, at all times, above 92% with respect to the reference approach (*PM56*), considered as an excellent performance according to the classification presented by CAMARGO and SENTELHAS (1997, p. 327).

9. DISCUSSION

The scope of this study was to assess how accurate Artificial Neural Networks could be at estimating reference crop evapotranspiration (*ETo*). The issue that motivated us to conduct this study was the recurrent problem among researchers of facing incomplete date registers. To accomplish this task, different *ANNs* were built as it was previously explained and fed with the same number of inputs as those required by Penman-Monteith method, which is considered as the standard method to estimate *ETo*. After training, validating and testing the *ANNs*, the results delivered by them were compared with those estimated by *PM* method and the diferrent statistical indices pointed out *ANN* approach as a high accurate tool to estimate *ETo*.

The neural networks performing the most accurate results were built presenting some parameters missing in the input layer (*RH* and U_2), a single hidden layer with no more than 20 neurons in it, fitted with a hyperbolic tangent sigmoid transfer function and guided by the Levenberg-Marquardt algorithm. This architecture was sufficient to reach outstanding results, with a resemblance above 92%, considered as an excellent performance according to the classification presented by CAMARGO and SENTELHAS (1997, p. 327) and a maximum MSE of 0.128 mm day^{-1}, taking off couple of exceptions. It confirmed the unpredicted capability of an *ANN* to reproduce at a high resemblance to the target model, in our case, the *PM* model, even with incomplete datasets.

These results confirmed our main hypothesis: *ANNs* were able to estimate *ETo* values with a high accuracy. Previous studies on the field reached the same conclusion (KUMAR *et al.*, 2002, p. 233; ZANETTI *et al.*, 2007, p. 89; JAIN *et al..*, 2008, 2225-2234; KHOOB, 2008, p. 39; LANDERAS *et al.*, 2008, p. 565; CHAUHAN; SHRIVASTAVA, 2012, p. 30). Translating these results into water column height, the difference between predicted *ETo* and target values was negligible in both daily measures and extraordinary low in accumulated volume for the whole study period.

Despite the excellent results obtained throughout this research some other questions arose, such as:

- The possibility of training, validating and testing different *ANNs* with a particular dataset and then estimating new patterns for a different dataset. It can be special useful in those cases where a weather station has been out of order for long periods and its dataset presents

really wide blanks.

- Testing to what extent climatic conditions could limit *ANNs* application. Different locations would probably point out different climatic parameters, like those more influent on *ETo*, due to the difference in altitude, radiation, wind speed or humidity conditions. In those cases, local sensitivity analysis would be a useful approach in order to confirm such influence. The approach would be the following: clustering the whole bunch of stations with respect to climatological conditions. Similar climatological conditions may represent similar behaviours of temperatures, radiation, relative humidity or wind speed, which would lead to a similar *ETo* range of values. It is known that in mountainous areas, relative air humidity, wind speed and temperatures differ so much from those in valley or coastal areas. The question is whether an artificial neural network can be trained, validated and tested in different locations, as long as we stay within an area with constant climatological conditions. That would be another interesting perspective to test the importance of some weather factors over others.

- As *ANNs* approach demonstrated to be an efficient tool to estimate *ETo* accurately, further analyses may be focused on future forecasting as irrigation schedule planning and management tool for those public institutions in charge of dealing with water demand, licencing, infrastructure or services.

- The chance of validating *ANNs* approach by using forecasted *ETo* values in other methodologies, in turn based on a water balance. This sort of procedures may allow us to model other physical processes where *ETo* plays a significant role.

Thus, deeper analyses and different approaches may be carried out in further studies.

CHAPTER III
INFLUENCE OF THE CLIMATOLOGICAL PARAMETERS ON THE EVAPORATION PROCESS

1. INTRODUCTION

As mentioned in previous chapters, *ET* has been distinguished as an important factor within the hydrologic cycle and accounts for the water lost both from the soil by evaporation and through the plants by transpiration. *ETo* is usually the term more widely utilized so as to refer analyze, compute or simply, to refer to the evapotranspirative process.

It is well known *ETo* is a complex process determined by several climatological parameters such as temperatures, solar radiation, relative air humidity and wind speed, all of which suffer from spatial and temporal variations. Such complexity lies on the existence of non-linear relations to the climatic parameters just mentioned. For instance, the influence of a single meteorological factor can differ from one place to another or from one time to a different one. Hence, several parameters varying simultaneously but in different scales could also lead to upward or downward trends in *ETo* (GOYAL, 2004, p. 11).

The extent in variation of *ETo* induced by a single factor is also unknown. Therefore, the dominant meteorological factors related to *ETo* changes need to be identified with the aim to better comprehend the effects of those on *ET* as a part of the hydrologic cycle. Its importance moved to a central position among scientific issues in the last decade due to the need of identifying which meteorological factors and to what extent those affect or can be influenced by the global climate change.

Literature is full of examples showing this relation: the reporting of meaningful decreasing trends in pan evaporation and *ETo* in recent decade, within a range from -1 to -5 mm/year over worldwide regions including European Russia, Siberia, United States, India, Israel, Australia and New Zealand (PETERSON, 1995, p. 688; CHATTOPADHYAY; HULME, 1997, p. 73; BRUTSAERT; PARLANGE, 1998, p. 30; LAWRIMORE; PETERSON, 2000, p. 546; GOLUBEV *et al.*, 2001, p. 668; COHEN *et al.*, 2002, p. 91; OHMURA; WILD, 2002, p. 346; HOBBINS *et al.*, 2004, p. 432; RODERICK; FARQUHAR, 2004, p. 190; 2005, p. 239). Variations in pan evaporation and *ETo* in China have been reported in recent years (THOMAS, 2000, p. 396; LIU *et al.*, 2004, p. 213; GAO *et al.*, 2006, p. 12;

WANG *et al.*, 2007, p. 23; ZHANG *et al.*, 2007, p. 546; YIN *et al.*, 2010, p. 28). The decline in *ETo* and its possible driving factors in response to climate change have stimulated extensive research and debate.

Despite this, scientific community have reached no agreement about the underlying causes of *ETo* changes. Some of them have attributed decreasing *ETo* to a decreasing in solar radiation or sunshine hours in Russia and United States (PETERSON, 1995, p. 688), China (THOMAS, 2000, p. 396; LIU *et al.*, 2004, p. 213; GAO *et al.*, 2006, p. 12), the Yangtze River Basin (WANG *et al.*, 2007, p. 23) and Israel (COHEN *et al.*, 2002, p. 91). Conversely, others pointed out wind speed as the main cause in Australia (RAYNER, 2007, p. 394; RODERICK *et al.*, 2007, p. 451), the Tibetan Plateau (CHEN *et al.*, 2006, p. 319) and Canadian Prairies (BURN; HESCH, 2007, p. 73), to relative humidity in India (CHATTOPADHYAY; HULME, 1997, p. 73) and to maximum temperature in China (CONG; YANG, 2008, p. 131). All these studies suggested mechanisms of *ETo* changes, whereas there is still much controversy around it.

In brief, the determining factors in *ET*o trends are yet to be distinguished. It is here, where we strongly believe in the necessity of further and deeper studies on the sensibility of *ETo*, generally estimated by the most broadly used method, Penman-Monteith FAO 56 equation, to its driving factors: *temperatures (Tmax, Tmin), solar radiation (Rs), relative humidity (RH)* and *wind speed (U_2).* Rainfall has not been taken into account throughout this chapter as it does not take part of Penman-Monteith FAO 56 equation. However, its importance is broadly known as it represents the main source of water income to the soil. Indeed, it is expected to provoke other climatic factors to behave differently depending on the presence or absence of rainfall. It is the case of air temperature, whose increase can result in a rise of evapotranspiration rate until reaching a threshold where only the presence of rainfall would allow the evaporative process to continue, otherwise the crop may enter a hydrological stress state.

Having not much literature addressing this issue in Brazil, the main objectives of this paper are: 1) to contribute to a more thorough understanding of the impacts of single climatic parameters in *ETo*; 2) determine whether exists any correlation between the climatic factors that influence *ETo* in the South East region of Brazil, Paraíba do Sul watershed specifically, during the periods analyzed.

2. MATERIALS AND METHODS

Three different methodologies were applied throughout the present study aiming to unveil those non-linear relations between *ETo* and the parameters that rule the process that we already exposed back in the Introduction section. By applying these three methodologies we intend to confirm the influence of the 5 climatological parameters under study on *ETo*. It is true that they would treat the issue from different perspectives, use different analyses tools and probably bring out different results. It is within this variety of perspectives where we expect to find points in common or similar insights about several behaviours. A particular methodology may be limited to a point the other one can reach or hidden correlations may be reveal if applying the appropriate method.

These three different methodologies were Principal Components Analysis, Global Sensitivity Analysis and Local Sensitivity Analisys and are being fully descrived in the following.

2.1 Principal Components Analysis

Principle component analysis (PCA) is a non-parametric statistical technique used to examine the interrelations among a set of variables in order to identify the underlying structure of those variables and to reduce a complex data set to a lower dimension allowing to reveal likely hidden dynamics. When computing a PCA, each data sample is considered as a vector of m-dimensional space, where m is the number of variables under study. Equivalently, every sample can be treated as a vector that lies in an m-dimensional space spanned by an orthonormal basis. As most of the statistical analyses, the data set object of study should meet some requirements prior to its analysis. Some of the assumptions any data set must fulfil are the following:

i. Linearity: this first obstacle is overcome by changing to an orthonormal basis.
ii. Gaussian distribution where variance and mean, which is zero, are sufficient statistics.
iii. Large variances have important dynamics as the data itself is highly redundant.

iv. Principal components are orthogonal to each other.
For each dataset PCA was applied individually following the steps described in the following:

Let *A* be an *n x m* matrix containing raw data composed by *n* measurements and *m* variables, where each measurement is considered as a *m*-dimensional vector where its mean ($\vec{\mu}$) is:

$$\vec{\mu} = \frac{1}{n}(x_1 + x_2 + \cdots + x_n) \quad \textbf{(62)}$$

It is common to reproject the data by subtracting the mean $\vec{\mu}$ from each sample measurement $\vec{X}_{ij}$ turning the mean to zero. A new matrix is created: *B*, an *n x m* matrix whose *ith* column is determined as $\vec{X}_{ij} - \vec{\mu}$ and meet both requirements of linearity and orthogonality of the new axis.

$$B = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1j} \\ x_{21} & x_{22} & \cdots & x_{2j} \\ \vdots & \vdots & \cdots & \vdots \\ x_{i1} & x_{i2} & \cdots & x_{ij} \end{bmatrix} \quad \textbf{(63)}$$

from which we are able now to compute the covariance matrix *S*, defined as follows:

$$S = \frac{1}{n-1} BB^T \quad \textbf{(64)}$$

where *S* is symmetric and it is here where we can easily find the *eigenvalues* $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m \geq 0$ and its corresponding orthonormal eigenvectors $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_m$. These eigenvectors are actually the new principal components that account for the total variance of the data, *T*.

Since *S* is a symmetric matrix, Spectral Theorem can be applied, which states that for any symmetric matrix *A*, $\lambda\vec{v} = A\vec{v}$, where $\vec{v}$ is the collection of eigenvectors of *A* and λ represents the correlation coefficients or eigenvalues of *A.* Aiming to compute the eigenvalues, let us analyze this same expression for the covariance matrix *S* instead:

$\lambda\vec{v} = S\vec{v}$, where $\lambda = \lambda I$ and *I* is an identity matrix that allows us to diagonalise *S*, a way of maximizing the variances, located on the diagonal entries and minimizing the covariances. We can rewrite the former equation as $(\lambda I)\vec{v} = S\vec{v}$ and finally reaching to a conclusion where $(\lambda I - S)\vec{v} = 0$.

Its result is an m-degree equation from where the eigenvalues λ_m are computed. The direction in the space R^m given by $\vec{u}_1$ (the first principal direction) explains or accounts for an amount λ_1 of the total variance *T*, which is actually λ_1/T. Similarly, the second principal direction $\vec{u}_2$ accounts for the fraction $\vec{u}_2$ of the total variance T and so on. From the same equation, it is straightforward to find the eigenvectors $\vec{v}_m$ by replacing the eigenvalues just obtained. Thus, $\vec{u}_1$ points towards the most significant direction of the data set. Among directions that are orthogonal to $\vec{u}_1$, $\vec{u}_2$ being orthogonal to $\vec{u}_1$ points in the second most significant direction of the data set and so on.

One of the most meaningful results are the so-called “loadings”. These refer to the existing correlation between the individual variables and the new PC’s. Loadings point out which variables are contained in which PC as well as how much each variable is correlated to each PC. This proceeding was programmed and applied by using MATLAB R2015a.

In this study, since the variables under study presented different variances and were recorded under their corresponding measurement units, it was necessary to standardize the whole data set. As a result, each variable was transformed into a new variable with zero mean and unit standard deviation. This task pursues to void the overweight of variables measured at higher scales cancelling out the differences among them.

A recurrent issue among researchers is about the number of PC’s that significantly represents the whole data. In this study we used the Broken Stick model which was first presented by MACARTHUR (1957, p. 564) during his studies on bird communities structure. In this specific case, the apportioned resource is represented by the total variance of our data set. Taking into account that each eigenvalue is considered as a measurement of each

components' variance, a given principal component is retained only if its corresponding eigenvalue is larger than the value given by the broken stick model. The *broken stick model* is an easy methodology to implement. Considering an interval *J* between 0 and 1 the model splits *J* into *n* subintervals by randomly selecting *n* - 1 points from a uniform distribution within the same interval. Superior subintervals will represent larger eigenvalues; in other words, the components retaining larger variances would be ordered in the first positions.

2.2 Global Sensitivity Analysis

Global sensitivity analysis focuses on the variance of a given model output Y and more precisely on how the input variability influences the variance output (SALTELLI, 2002, p. 297). It enables to determine which parts of the output variance are due to the different inputs by means of the estimation of *Sobol* indices. They are a central tool in sensitivity analysis since they give a quantitative and a rigorous overview on how different inputs influence the output. To apply global sensitivity analysis methods, it is necessary to assume in the following that the given variables are defined on the interval [0, 1]. Sobol sensitivity indices at first order *Si* for the variable *X(i)* are then defined with:

$$S_i = \frac{V_i}{V} = \frac{Var\left(E\left(Y \middle| X^{(i)}\right)\right)}{Var\left(Y\right)} \quad \textbf{(65)}$$

If *Si* is close to 1, then the variable *X(i)* has a great influence on *Y*. When the input dimension *p* increases, the number of *Sobol* indices rises exponentially and its estimation turns more complex. For that purpose, total sensitivity indices S_{Ti} are introduced for each variable *X(i)*:

$S_{T_i} = \sum_{k \# i} S_k$, where S_{Ti} represents all sets of indices that contain *i*. **(66)**

Sobol indices are often estimated by means of *Monte-Carlo* methods. Searching for a better understanding, let us consider two *N* sample-size realizations of *X*:

$$\tilde{X}_k^i = \left(x_{k1}^i, \mathrm{K}\ , x_{kp}^i\right), \text{ where k=1,…,N and i=1,…,p} \quad \textbf{(67)}$$

Each sample corresponds to a single point within the input space with *p* dimensions. From these two *N* sample-size realizations of *X*, first order sensitivity index $\hat{S}_i$ for the input *X(i)* is estimated in the following way:

$$\hat{S}_i = \frac{\hat{V}_i}{V} = \frac{\hat{U}_i - \hat{\phi}_0^2}{\hat{V}} \text{ where } \hat{\phi}_0 = \frac{1}{N}\sum_{k=1}^{N}\phi\left(x_{k1}^1, \mathrm{K}\ , x_{kp}^1\right) \quad \textbf{(68)}$$

The variance is estimated by definition with $\hat{V} = \frac{1}{N}\sum_{k=1}^{N}\phi^2\left(x_{k1}^1, \mathrm{K}\ , x_{kp}^1\right) - \hat{\phi}_0^2$ **(69)**

Second order sensitivity indices *Sij* are estimated in the following manner:

$$S_{ij} = \frac{\hat{U}_{ij} - \phi_0^2 - \hat{V}_i - \hat{V}_j}{\hat{V}} \quad \textbf{(70)}$$

Sensitivity indices of superior orders are then derived from theses relations. Total sensitivity indices can be estimated with the following relations:

$$\hat{S}_{T_i} = 1 - \frac{\hat{U}_{\tilde{i}} - \hat{\phi}_0^2}{\hat{V}} \quad \textbf{(71)}$$

A first attempt aiming to investigate the sensitivity of various evaporation models and their rates to changes on their corresponding driving weather parameters was conducted by MC CUEN (1973, p. 612). After him, several studies were carried out on sensitivity of *ETo* to weather parameters, mostly conducted on a single station dataset. Since then, only fewer studies have encompassed regional or seasonal behaviour of the sensitivity of *ETo* to its driving climatic factors.

BEVEN (1979, p. 502) investigated the sensitivity of actual crop evapotranspiration estimates (*ETa*) to random changes in input data across England and Wales, concluding that *ETa* estimates from Penman-Monteith FAO 56 equation were more sensitive to aerodynamic and canopy resistance parameters than to vegetation type. (SINGH; XU, 1997, p. 473) investigated the influence of evaporation changes computed by mass transfer-based equation in Switzerland. Evaporation resulted to be more sensitive to vapour pressure gradient, less sensitive to wind speed and most insensitive to air temperature data. GOYAL (2004, p. 123), whose research was conducted across India, found that an increase in average air temperature caused a proportional increasing on *ETo*, being less sensitive to net radiation. GONG *et al.* (2006, p. 641) found relative humidity as the most sensitive variable followed by radiation and air temperature across Yangtze River basin. According to LIU *et al.* (2010, p. 251) and YU *et al.* (2002, 244), solar radiation resulted to be the most sensitive variable after estimating *ETo* by means of Penman Monteith equation, followed by relative humidity in Yellow River basin.

All these studies suggested mechanisms of *ET*o variation, but as we could see there is yet so much controversy on this issue. It is obvious that *ET*o is determined by several meteorological parameters, all of which show complex spatial and temporal variations. Indeed, the influence of the same meteorological variables on *ET*o differs between climatic regions. Moreover, different combinations of changes in meteorological variables, such as temperature, relative humidity and radiation, also lead to upward or downward trends in *ET*o. The range of *ET*o variation induced by a single factor is also uncertain.

To perform such analysis for each and every dataset gathered from the 26 INMET weather stations that compose this research, it needed to be programmed and applied by using MATLAB R2015a, which has already this function among its built-in functions. Nevertheless, Penman-Monteith FAO56 equation had to be modelled by using Simulink Design Optimization software. *Tmax*, *Tmin*, *RH*, *Rs* and U_2 were set as independent variables within the model, whereas everything else was considered as constants: altitude, latitude and Julian date as we were not interested in the influence these factor might have on *ETo*. *Monte-Carlo* techniques were utilized aiming to generate samples so as to explore the design space. The workflow proceeded as follows:

1. Meeting the requirements specified by the Simulink Design Optimization software in order to make Penman-Monteith FAO56 model work properly. For more information, visit <https://es.mathworks.com/help/sldo/ug/what-is-sensitivity-analysis.html>

2. Sampling the model parameters using experimental design principles. That is, for each variable, generate multiple values the parameter can assume. The number of samples recommended are *10Np*, where *Np* is the number of variables to be analysed, in this case, *50*. However, *100* samples were built, in our case.

3. Defining the parameter sample space by specifying probability distributions for each parameter. Simulink even allows us to specify parameter correlations.

4. Evaluate the function under study (Penman-Monteith FAO56) at each sample point. Simulink enters the model, one by one, all the samples previously generated. It is advisable to plot the results so as to visually analyse any trends on the output.

5. Analyse the relation between the function under study and the samples. Analysis methods in MATLAB R2015a include Correlation, Partial Correlation, Kendall Correlation, Rank Correlation, Partial Rank Correlation, Standardized Regression and Rank Standardized Regression.

2.3 Local Sensitivity Analysis: Morris method

Local sensitivity analysis is able to determine to what extent a small variation on an input value $X^0 = X_1^0, \dots, X_p^0$ influences the output value $y = \emptyset(X^0)$. To accomplish this, it gets necessary to estimate:

$$A_i = \frac{\partial y}{\partial x_i}\left(x_1^0, \text{K}\ , x_p^0\right) \quad \textbf{(72)}$$

That defines the effect on the output *Y* generated by a change on input $X^{(i)}$ near a nominal value X_i^0. A classical approach to compute this quantity of variation is to consider a "*One At a Time*" method, as the Morris Method.

The Morris method (MORRIS, 1991, p. 174) is a specialized randomised *One-At-a-Time SA* design. It is an efficient and reliable technique to identify and rank important variables (MORRIS, 1991, p. 174; CAMPOLONGO *et al.*, 2007, p. 518). The method relies on a sensitivity measure, *"the elementary effect"* (*EE*), which is a local measure, but the final sensitivity measurements are obtained by averaging these elementary effects computed at different points to evenly spaced values of each parameter over its entire range (NORTON, 2009, p. 674). Traditional "*One At a Time*" sensitivity analysis design, where each variable is tested individually, assumes that as long as all variables are changed by the same relative amount, the variable that causes the largest variation on the output is the most influent.

The Morris method is simple to implement, and it gives easily interpreted results. However, it also has a drawback, it only provides qualitative sensitivity measures by ranking the input factors in order of importance. Thus, the standardised effect of a positive or negative change (Δ) of an input variable can be computed using the corresponding *EEs*, define it as follows:

$$EE_i(x) = \left[y\left(x_1, x_2, \text{K}\ x_{i-1}, x_i + \Delta, x_{i+1}, \text{K}\ , x_k\right) - y(x)\right] / \Delta \quad \textbf{(73)}$$

where Δ is the magnitude of the step.

Morris method uses the same principle as Global Sensitivity Analyses when changing one variable at a time by the same amount of variation Δ within a range of intervals, which is known as an experiment. While Morris method requires only *k + 1* experiments to calculate one *EE* for each of the *k* input variables, traditional "*One At A Time*" SA designs require *2k* experiments. Although, the main difference lies on the fact that here the experiments are arranged in order to create a trajectory throughout the variable space.

MORRIS (1991, p. 174) proposed two simple measures of the combination of *EEs* pertaining to the same input variable: the mean (μ), which represents the strength between the input variable and the output response (CAMPOLONGO; BRADDOCK, 1999, p. 12); the standard deviation (σ), interpreted as an indicator of possible existence of interactions between variables and the non-linear effect on the output. And a third term, μ^* introduced by CAMPOLONGO *et al.* (2007, p. 518) which is considered as the average of the distribution of absolute values of EE_i and provides the overall sensitivity of the *i-th* input variable and is define as follows:

$$\mu^* = \frac{\sum_{n=1}^{r} \left| EE_n \right|}{r} \qquad \textbf{(74)}$$

Some of the studies carried out by using Morris method had hydrology as a general study field. For instance, SUN *et al.* (2012, p. 24) applied Morris method to assess model sensitivities to input parameter values in a water quality model in Australia. SHEN *et al.* (2008, p. 269) conducted a research aiming to better understand which factors most influenced non-point source pollution by using the software *SWAT*. FRANCOS *et al.* (2003, p. 320) applied Morris method as a first approach of a two-step procedure aiming to provide quantitative estimations of sensitivity in terms of variance decomposition procedures for both hydrological and water quality determinants. PAPPENBERGER *et al.* (2008, p. 201) investigated how a one-dimensional flood inundation model (*HEC-RAS*) could be influenced by different factors on the River Alzette, Luxembourg by using five procedures, Morris method among them.

Along this research, Morris method was applied by using the open-source programming language Python and a specific Sensitivity Analysis Library, SALib (for download and further information, please visit <https://github.com/SALib/SALib>) where Morris method can be found and implemented. This time the workflow proceeded as follows:

1. It is determined the range each variable is within.
2. Each variable registers are turned into values comprised between 0 and 1 so as to cancel out any influence from the different scales parameters were registered.
3. The own set of registered values of *Tmax, Tmin, Rs, RH* and U_2 would determine the number of samples generated. Recalling that we only need $k+1$ experiments to compute the corresponding *EE*.

It proceeds to structure the “problem”, to which Morris method would propose some forecasts.

3. RESULTS AND DISCUSSION

3.1 PCA results

Based on the results showed in Table 16, among conventional weather stations gathered in clusters 1 and 2, *Tmax* is correlated to PC1 between a minimum of 0.55 (Ta) and a maximum of 0.61 (B, L). Its correlation to PC2 ranges from a minimum of 0.17 (SL) to a maximum of 0.39 (Ta). *Rs* is correlated to PC1 from a minimum of 0.52 (CG, CJ) and a maximum of 0.62 (L) and between -0.34 (I, G) and a positive value of 0.05 (CP) to PC2. *RH* is correlated to PC1 from a minimum of a -0.42 (CG, CJ) to a maximum of -0.56 (Ta) and from 0.34 (Ta) and 0.63 (JF) to PC2. These are actually the three variables contained in PC1 as they all brought loadings over 0.42 (CG, CJ).

Table 16. Maximum and minimum loadings pertaining to clusters 1 and 2 (conventional weather stations) representing the level of correlation between raw variables (Tmax, Tmin, Rs, RH, U2) and the new Principal Components

Conventional Stations		CLUSTER 1								CLUSTER 2				
		B	**CP**	**JF**	**PA**	**R**	**L**	**SL**	**Vi**	**CG**	**CJ**	**I**	**G**	**Ta**
	Tmax	**0.61**	0.60	0.59	0.59	0.59	**0.61**	0.57	0.56	0.56	0.56	0.59	0.59	**0.55**
	Tmin	0.23	0.22	0.41	0.23	0.20	0.19	0.16	0.18	0.35	0.35	0.31	0.31	0.08
	Rs	0.59	0.60	0.55	0.60	0.61	**0.62**	0.61	0.56	**0.52**	**0.52**	0.54	0.54	0.60
PC1	RH	-0.47	-0.43	-0.43	-0.48	-0.48	-0.46	-0.50	-0.48	**-0.42**	**-0.42**	-0.50	-0.50	**-0.56**
	U2	-0.05	0.21	0.00	0.12	0.04	-0.01	0.19	0.33	0.33	0.33	-0.11	-0.11	0.14
	Var.	2.23	2.11	2.49	2.26	2.31	2.19	2.33	2.45	2.47	2.48	2.39	2.39	2.24
	%	44.63	42.26	**49.83**	45.27	46.25	43.75	46.69	48.99	49.48	49.56	47.75	47.87	44.85
	Tmax	0.21	0.28	0.20	0.26	0.22	0.26	**0.17**	0.18	0.26	0.25	0.29	0.29	**0.39**
	Tmin	0.79	0.62	0.69	0.71	0.81	0.73	0.81	0.83	0.74	0.74	0.77	0.77	0.84
	Rs	-0.09	**0.05**	-0.24	-0.04	-0.09	-0.08	-0.08	-0.16	-0.21	-0.20	**-0.34**	**-0.34**	-0.16
PC2	RH	0.55	0.53	**0.63**	0.50	0.51	0.55	0.47	0.47	0.56	0.57	0.45	0.46	**0.34**
	U2	-0.17	-0.51	0.19	-0.43	0.16	-0.31	0.30	0.20	-0.17	-0.17	0.00	-0.01	0.03
	Var.	1.28	1.36	1.09	1.25	1.30	1.47	1.31	1.23	1.11	1.11	1.22	1.21	1.34
	%	25.61	27.13	21.71	24.92	25.91	**29.41**	26.21	24.54	22.27	22.24	24.41	24.25	26.78
	Tmax	-0.04	-0.09	0.07	-0.07	0.19	-0.06	0.34	0.34	0.25	0.25	-0.02	-0.02	-0.19
	Tmin	0.21	0.57	0.12	0.44	0.05	0.32	0.17	0.15	-0.12	-0.12	0.10	0.10	0.06
	Rs	0.07	-0.28	-0.08	-0.16	-0.02	0.09	0.05	0.19	0.29	0.28	0.08	0.08	-0.12
PC3	RH	0.02	0.14	0.12	0.14	0.14	0.15	0.16	0.06	-0.13	-0.13	-0.09	-0.09	-0.07
	U2	0.98	0.76	-0.98	0.87	-0.97	0.93	-0.91	-0.91	-0.91	-0.91	0.99	0.99	0.97
	Var.	1.00	0.88	1.00	0.96	1.02	0.98	0.96	0.82	0.83	0.82	0.99	0.99	1.01
	%	20.09	17.57	19.97	19.28	**20.31**	19.69	19.23	16.49	16.52	16.49	19.88	19.90	20.18

Source: data computed by the author

Based on the results showed in Table 17, among automated weather stations gathered in clusters 3 and 4, *Tmax* is correlated to PC1 between a minimum of 0.51 (DC, Ma) and a

maximum of 0.64 (MV). Its correlation to PC2 ranges from a minimum of 0.20 (P) to a maximum of 0.41 (Ma). *Rs* is correlated to PC1 from a minimum of 0.4 (MV) and a maximum of 0.51 (SJ) and between -0.57 (MV) and a positive value of 0.05 (Spd) to PC2. *RH* is correlated to PC1 from a minimum of a -0.24 (MV) to a maximum of -0.46 (Te) and from positive values of 0.28 (P) and 0.72 (SJ) to PC2. Similarly to what was found within clusters 1 and 2, *Tmax, Rs* and *RH* are the three variables contained in PC1 as they all brought loadings over 0.4 (MV), except MV in the case of *RH*, which showed only a 0.24, quite low value if comparing to the rest.

Most times, only the first two PC's are enough to explain a great portion of the total variance of our whole data set, represented by the variable *Var*. Based on our results, the variance explained into the first two PC's for all variables under study appears also in Tables 16 and 17 represented as percentage and the symbol %.

Table17. Loadings pertaining to clusters 3 and 4 (automated weather stations) representing the level of correlation between raw variables (Tmax, Tmin, Rs, RH, U2) and the new Principal Components

Automated stations		CLUSTER 3						CLUSTER 4						
		DC	**P**	**SJ**	**Te**	**NF**	**Va**	**MV**	**SLP**	**C**	**PQ**	**ST**	**Ma**	**Spd**
	Tmax	**0.51**	0.58	0.58	0.52	0.61	0.56	**0.64**	0.56	0.53	0.61	0.56	**0.51**	0.52
	Tmin	0.48	0.56	0.54	0.50	0.59	0.54	0.61	0.54	0.52	0.59	0.56	0.50	0.50
	Rs	0.44	0.43	**0.51**	0.45	0.41	0.41	**0.40**	0.47	0.44	0.46	0.45	0.48	0.46
PC1	RH	-0.45	-0.38	-0.33	**-0.46**	-0.30	-0.41	**-0.24**	-0.37	-0.42	-0.26	-0.30	-0.36	-0.43
	U2	0.33	-0.12	-0.10	0.28	-0.09	0.25	0.04	0.20	0.29	-0.06	0.29	0.36	0.29
	Var.	3.14	2.55	2.55	3.01	2.37	2.70	2.15	2.70	2.89	2.41	2.70	2.98	3.12
	%	**62.83**	51.01	51.10	60.25	47.33	54.00	42.98	54.06	57.84	48.13	53.97	59.54	62.42
	Tmax	0.40	**0.20**	0.34	0.40	0.24	0.35	0.27	0.39	0.38	0.21	0.33	**0.41**	0.34
	Tmin	0.46	0.26	0.43	0.43	0.33	0.38	0.36	0.46	0.39	0.32	0.35	0.46	0.38
	Rs	-0.05	-0.10	-0.40	-0.09	-0.41	-0.08	**-0.57**	-0.39	-0.01	-0.24	-0.04	-0.09	**0.05**
PC2	RH	0.40	**0.28**	**0.72**	0.41	0.34	0.44	0.69	0.69	0.47	0.64	0.65	0.56	0.38
	U2	-0.68	0.90	-0.14	-0.69	0.75	-0.73	-0.03	-0.11	-0.69	0.62	-0.59	-0.55	-0.77
	Var.	0.97	0.99	1.20	1.03	1.10	1.05	1.52	0.99	1.03	1.15	0.97	1.12	0.96
	%	19.31	19.79	23.91	20.57	21.90	21.00	**30.35**	19.76	20.65	23.00	19.45	22.32	19.26
	Tmax	-0.21	0.33	0.09	-0.23	0.19	-0.23	-0.04	-0.08	-0.23	0.17	0.04	-0.04	0.28
	Tmin	-0.30	0.35	0.15	-0.30	0.21	-0.30	0.00	-0.04	-0.31	0.19	0.08	0.03	0.36
	Rs	0.70	-0.49	-0.04	0.64	-0.17	0.70	0.01	-0.17	0.66	-0.30	-0.23	0.00	-0.59
PC3	RH	-0.27	0.60	0.05	-0.32	0.74	-0.31	0.06	0.14	-0.35	0.49	0.61	0.70	0.46
	U2	-0.55	-0.42	0.98	-0.59	-0.58	-0.52	1.00	0.97	-0.54	-0.78	0.75	0.71	0.49
	Var.	0.62	0.94	0.98	0.70	0.92	0.73	1.00	0.93	0.69	0.90	0.77	0.49	0.55
	%	12.38	18.78	19.66	14.09	18.39	14.69	**20.01**	18.51	13.81	18.00	15.37	9.82	10.97

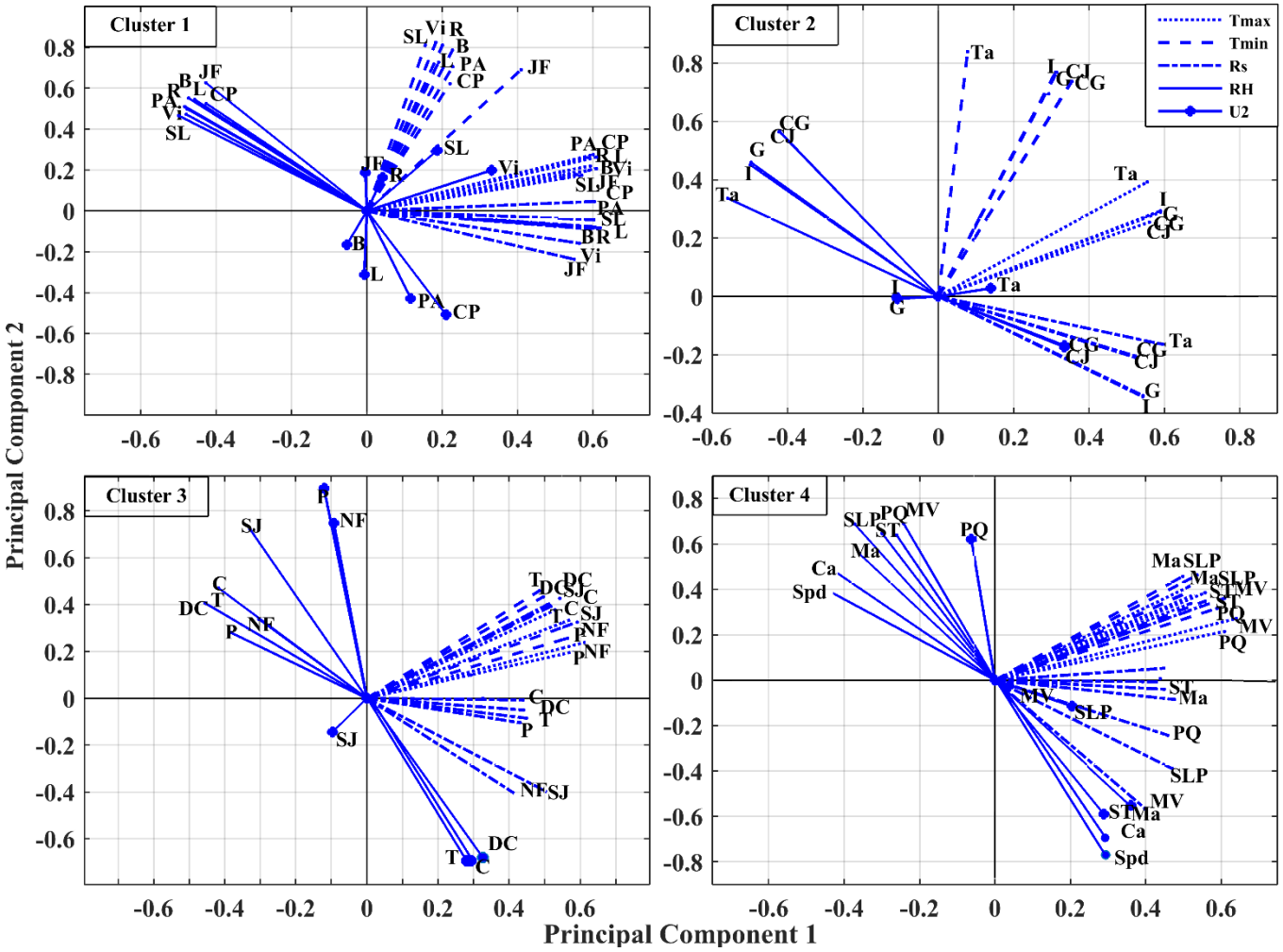
Source: data computed by the author

By combining loadings and eigenvectors (PC's variance) into the same plot, we get the so-called biplot. It represents a new re-projected coordinates system based on the orthogonal

directions of the first two PC's. Loadings are reprojected now onto this new coordinates system showing the way they correlate to any of the PC's.

A biplot of loadings and eigenvectors is shown in Figure 19. It is important to draw the reader's attention to the fact that the data represented on this plot is not the actual data, but an orthogonal transformation of it. The new orthogonal data allows us to re-project the whole data set on the direction of the greatest variance (abscissa axis), corresponding to PC1 and on the direction of the second greatest variance (ordinate axis), corresponding to PC2 as depicted in Figure 19.

Figure 19. .Biplot showing loadings of the variables under study and its corresponding eigenvectors on the new orthogonal basis for the four clusters of weather stations



Source: biplots computed by the author

The second PC is orthogonal to the first one and indeed represents the second greatest variance contained into our original data set. It can be easily checked by paying attention to the

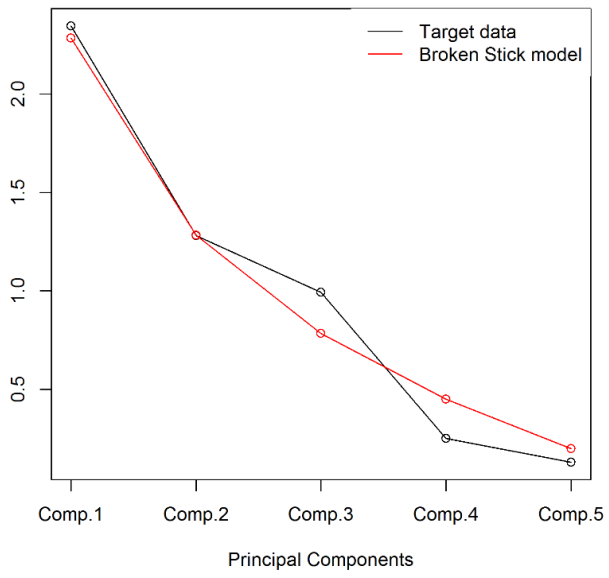
PC's variance item (*Var*), placed both in Tables 16 and 17. For those stations gathered into clusters 1 and 2, the variance carried by PC1 with respect to the whole data set reaches an average value of 46.70% with a maximum of 49.83% (JF), PC2 carries an average of 25.03% and a maximum of 29.41% (L) and PC3 represents only an average of 18.89% with a maximum of 20.30% (R).

For those stations gathered into clusters 3 and 4, the variance carried by PC1 with respect to the whole data set reaches an average value of 54.26% with a maximum of 62.83% (DC), PC2 carries an average of 21.63% and a maximum of 30.35% (MV) and PC3 represents only an average of 15.72% with a maximum of 20% (MV).

For both conventional and automated INMET stations these three first PC's represent over 90% of the total variance of the original data set and, as the *broken stick model* forecasted, are enough to explain the total variance of our data set.

In Figure 20, we can see the *broken stick model* represented by a red line and the target data behaviour and its variance split into the different PC's and represented by a black line. As explained above, only the number of PC's located above the line drawn by the *broken stick model* are significant enough at representing such variance. So, in our case and as we can see in Figure 20, the variance in our dataset in represented meaningfully mostly in the first three PC's.

Figure 20. Overlap between Broken Stick Model and our data set showing the most representative variables



Source: plot computed by the author through R statistics programming language

Now, it is time to transfer those general statements onto the results obtained over this study:

From the results gathered both in Table 16 and 17 and taking some exceptions off, *Rs*, *Tmax*, *Tmin* and, most of the times, U_2 have positive sign within PC1 for all data clusters which means that the higher these parameters are, the higher *ETo*. Indeed, among all variables, *Tmax* accounts for the highest portion of PC1, with an average value of 0.58 among conventional stations and 0.56 among automated stations, which means that *Tmax* is the parameter whose variability reflects the most on *ETo*. On the other hand, *RH* appeared with negative sign within PC1 showing an inversely proportional relation to *ETo*, in other words, the lower *RH* is, the higher *ETo* can reach. These results are fully aligned to those we expected to reach.

Going further, we can note that PC1 clustered the five climatological parameters under study into two groups. One group apparently represents the sources of energy: *Rs*, *Tmax* and *Tmin* as the driving forces that switch on the evapotranspiration process. The other group may be considered as the parameters which facilitate the process to continue under appropriate conditions; *RH* and U_2 are in this second group. *RH* loadings have a negative sign within the PC1, which means that the lowest *RH* gets, the highest *ETo* reaches. U_2 sometimes acquires negative values in contrast to what is expected.

Conversely, it is clear that U_2 has little influence on evapotranspiration due to its low magnitude within PC1, averaging 0.10 over conventional stations and 0.15 over automated stations.

Within PC2, *Tmin* appears as the most correlated variable to PC2 computing 0.75 as an average among conventional stations and 0.38 among automated stations. PC2 apparently appears to speak about the existence of some kind of correlation between the variables under study. Indeed, climatological parameters that influence the surrounding atmosphere such as temperatures *(Tmax and Tmin)* are well-known to be strongly related to *RH* and their combination acts as the driving force which rules the dynamics of the air and its movement. Another hypothesis would speak about the influence of the intensity of a given parameter. *Rs* represents the main source of energy and it is responsible for switching the evapotranspiration process on, independently of the amount of MJ m^{-2} day^{-1}, that is to say that as long as there is light, there is evapotranspiration. A great amount of *Rs* would lead the crop to an excess of evapotranspirated water and consequently to potential hydrological stress conditions.

As far as temperatures are concerned, since the temperature gets higher and higher, evapotranspiration also rises until reaching a critical level where the leaves stomata close aiming to keep a minimum content of water inside the plant and avoid dehydration. *Tmax* represents this case, where high temperatures can take the crops to those unsustainable conditions. It occurs the same with the *RH;* as the humidity level in the air rises, the air humidity gradient gets smaller meanwhile the surrounding atmosphere gets saturated until equalling the water vapour content with the leaves.

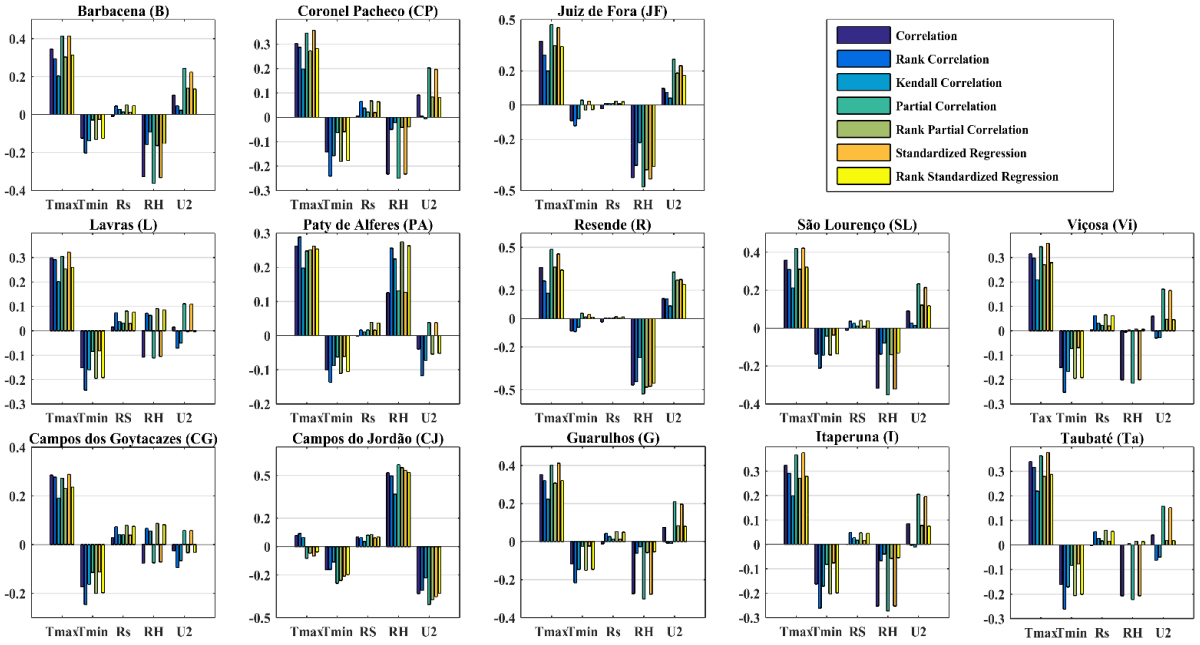
Wind speed *(U_2)* appears to be set apart from the rest of the parameters since their values are negligible within PC1, 0.10 among conventional stations and 0.15 among automated stations. Oppositely, U_2 turns into the far highest parameter in PC3 averaging an absolute value of 0.92 for data clusters1 and 2 and 0.68 for data clusters 3 and 4. In addition, the parameter-reduction process brought two PC's that explained the 70.24 % of the data variation, according to the *broken stick model*.

3.2 Global sensitivity analysis results

Most analysis options applied as a part of the global sensitivity analysis over conventional stations (clusters 1 and 2) pointed out *Tmax* as the most sensitive parameter on the Penman-Monteith FAO 56 equation/ETo, specifically in 10 out of 13 stations and always with a positive sign, as depicted in Figure 21. Within those 10 stations, *RH* appeared as the second most influential parameter in 7 of them. In addition, *RH* was ranked in first position in 3 stations: JF, Re and CJ. In this third station *RH* surprisingly appeared with a positive sign, an event also repeated in PA if we check in Figure 21. *Tmin* was, most of the times, the third parameter in order of influence.

On the other hand, *Rs* and U_2 appeared as the least sensitive parameters on *ETo* in 7 stations. It is worthwhile to mention the contrasting results brought by CJ, where *RH* and U_2 turned into the two most influential parameters, completely opposite to what we found so far.

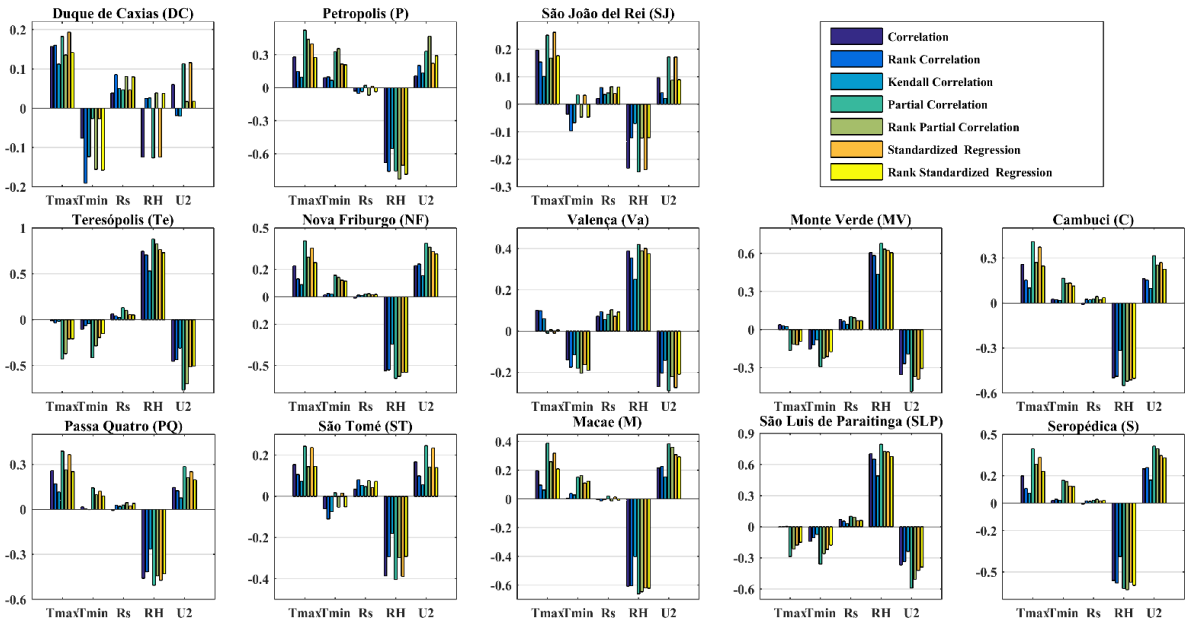
Figure 21. Climatological parameters sensitivity results after applying a Global Sensitivity Analysis upon conventional stations (clusters 1 and 2) and analysed by different analysis methodologies



Source: plots computed by the author

With respect to automated stations, *RH* was the far most sensitive parameter on Penman-Monteith FAO 56 equation/ETo, specifically in 11 out of 13 stations, as it is shown in Figure 22. Its sign was negative in 7 stations and positive in 4 out of the former 11 stations. *Tmax* was ranked as the second most influential in 8 out of the former 11 stations, besides being ranked in first position in DC and SJ. Always appearing with the opposite sign to *RH*. Within automated stations, U_2 appeared in third position, really close to the values brought by *Tmax* and also appearing with the opposite sign to *RH*. Again, *Rs* appeared as the least sensitive parameter on *ETo*, this time along with *Tmin*.

Figure 22. Climatological parameters sensitivity results after applying a Global Sensitivity Analysis upon automated stations (clusters 3 and 4) and analysed by different analysis methodologies



Source: plots computed by the author

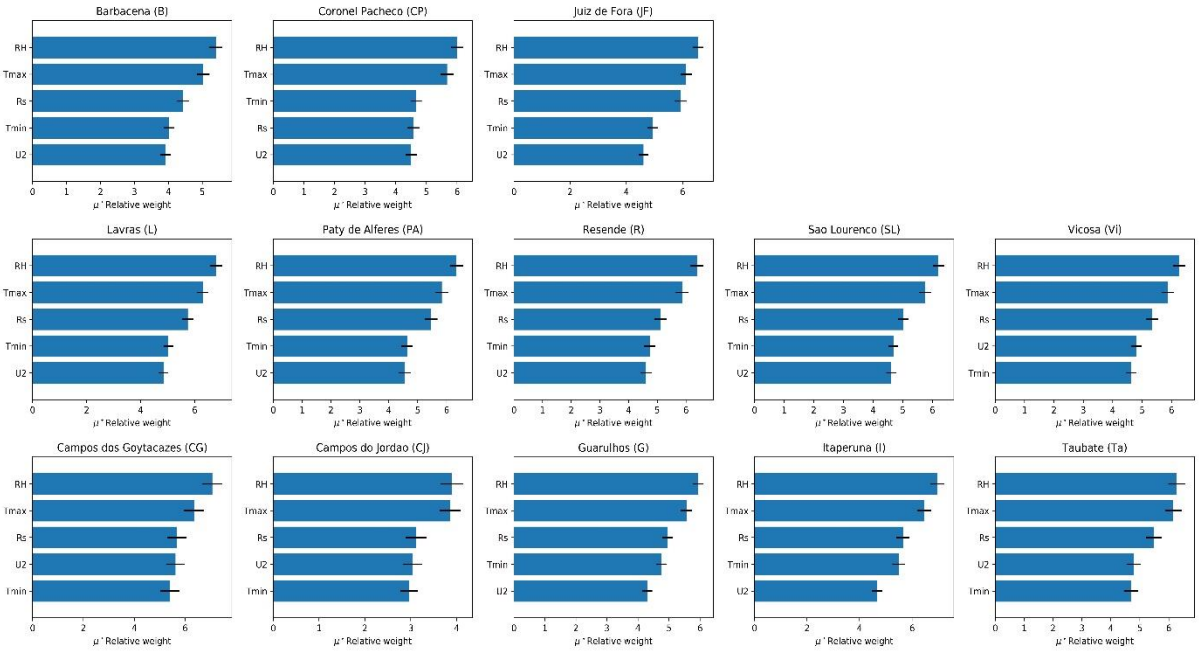
At first, there seems to be an incoherence when contrasting results from both sensitivity analyses, as *Tmax* and *RH* are pointed out as the most sensitive parameter within conventional (Figure 21) and automated (Figure 22) stations, respectively. It is well-known that *RH* changes greatly along the day reaching its highest values early in the morning and its lowest values at midday, a behaviour completely opposite to the one shown by *temperatures*. Is in this shifting behaviour and in the fact that automated stations take measures hourly where we may find the answer to this incoherence. Hourly measures are able to better reflect broad changes in *RH* while daily measures tend to reduce this feature to an average value, losing its strength. *RH* fluctuations are even higher than those registered by *temperature,* but this feature usually fades away when we operate with daily measures.

3.3 Morris method results

In figure 23, we can see the different variables under study ranked in order of importance, as it was mentioned previously. The index used to rank them as μ^* which indicates

RH as the most sensitive parameter on the evapotranspiration process, followed by *Tmax-Rs* for all the conventional stations, except CP whose third parameter in importance was *Tmin*. This order of sensitivity has remained steady for the whole bunch of conventional stations. After those three, *Tmin-U_2* appeared as the less sensitive, interchanging sometimes their positions.

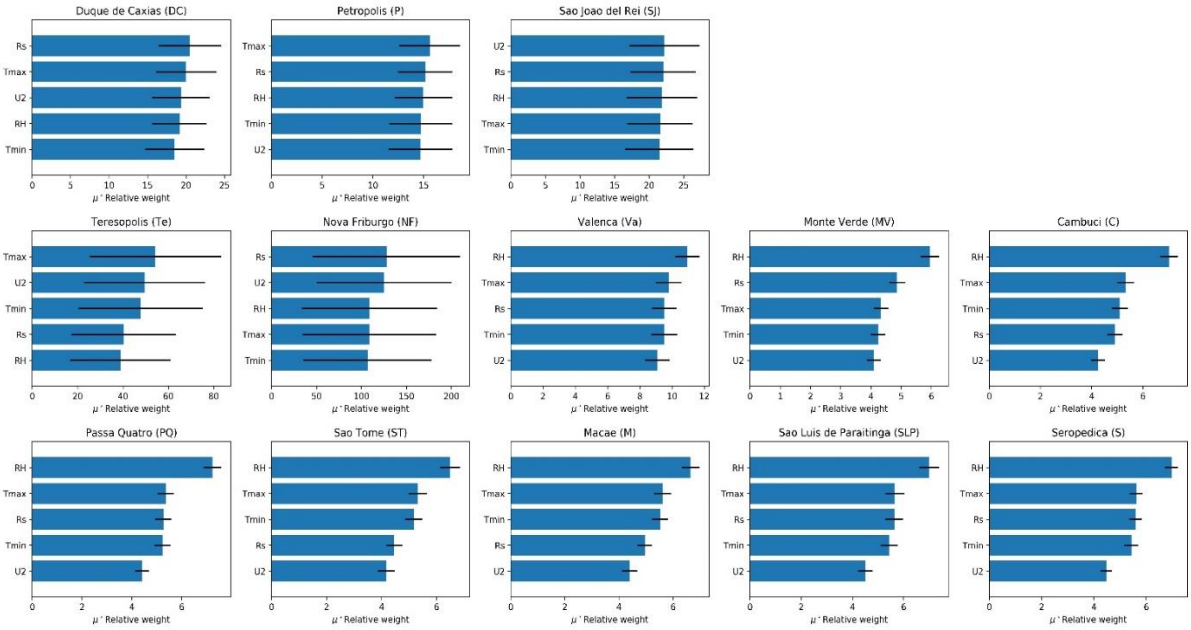
Figure 23. Climatological parameters ranked in order of influence after applying Morris method upon conventional stations (clusters 1 and 2)



Source: plots computed by the author by using an open-source scientific library (SALib) on Python programming language

In Figure 24, results were more disparate and different parameters were recognised as the most sensitive. Still, *RH* appeared in first position as far as sensitivity is concerned in 8 out of 13 stations; from those, in 3 of them, *Tmax*-T*min* followed *RH*, in 4 of them *Rs*-*Tmax* were located in second and third position. Aside from that, *Rs* was ranked in first position in DC, and NF, followed by *Tmax-U2* and *U_2-RH*, respectively. *Tmax* was also ranked as the most sensitive among all variables under study in P and Te.

Figure 24. Climatological parameters ranked in order of influence after applying Morris method upon automated stations (clusters 3 and 4)



Source: plots computed by the author by using an open-source scientific library (SALib) on Python programming language

4. CONCLUSION

All three approaches described in this third chapter, PCA, GSA and Morris method, pointed out *Tmax* and *RH* as the most influential parameters to evapotranspiration, as they have always been ranked within the three first positions.

5. DISCUSSION

Regarding PCA, first principal component clustered the parameters into two subclasses. One class with positive sign and composed by *solar radiation*, *maximum temperature* and *minimum temperature* apparently represents the sources of energy switching on the evaporative process; the other group, with negative sign and formed by *relative humidity* and *wind speed* may be considered as "helpers", keeping the process under appropriate conditions showing a low variance on evapotranspiration. These results agree with those reached by IKUDAYISI; ADEYEMO (2016, p. 210) where these three parameters also appeared with a positive sign and even in a very similar location within the biplot pointing out their directly proportional relation to *ETo* and thus backing our own results.

Taking all stations under study as a unique cluster, PCA results highlighted *Tmax* and *Rs* as the most sensitive parameters to *ETo*, since an average of 57% and 51% of their respective variances are represented in PC1. Next, *RH* reached a 42% of sensitivity to *ETo* model, with a negative sign instead.

Due to their negative values, we can claim that a low *relative humidity* is the most suitable condition to allow evapotranspiration proceed as the atmosphere around the evaporating surface is far from being saturated and the humidity gradient is still high. A result in concordance with those from CHATTOPADHYAY; HULME (1997, p. 451), which concluded that *relative humidity* appeared as the variable most strongly associated with the changes in potential evapotransporation, particularly in the pre-monsoon and monsoon sea. BAKHTIARI; LIAGHAT (2011, p. 320) also concluded that daily estimates of *ETo* were sensitive to *vapour pressure deficit*, which depends on *relative humidity*.

As long as there is light, *evapotranspiration* keeps working. This statement have been deducted from the results for *minimum temperature* within the second principal component, which shows the convenience of mild temperatures as the most suitable condition to maintain a high level of *evapotranspiration* without compromising the crop health.

Wind speed was pointed out as a non-correlated variable to the rest due to their high values within the third principal component. Its interpretation may correspond to an attempt of explaining its shifting dynamics, explicitly illustrated by the eigenvectors in Figure 20. Wind

speed is, due to their low values within the first principal component, the least influential parameter on *ETo* estimation. This partial result regarding *wind speed* is difficult to contrast as most researches have been developed either on arid or semi-arid environments, where, conversely to our results, *wind speed* resulted one of the most influential parameters.

From the comparison between both sensitivity analyses, it is remarkable the agreement between the results brought by both methodologies, pointing out *RH* and *Tmax* as the most influent parameters on Penman-Monteith FAO56/*ETo* in most of the stations under study. *Tmax* and *RH* were placed in first position by global sensitivity analysis and Morris method, namely. This conclusion is shared by ESLAMIAN *et al.* (2011, p. 459), who reached the same result, recognising *RH* and temperature as the most sensitive parameters after a study carried out in Iran after analysing 54 years of meteorological data from 5 weather stations. They added that the effect of *RH* on evapotranspiration is stronger than the one from U_2 and even that humidity with closing 100% makes evapotranspiration decrease and therefore wind speed effect decreases too. Lower *RH* values provoke high evapotranspiration rates making U_2 changes affect evapotranspiration with more severity.

YIN *et al.* (2010, p. 511) also got to a similar conclusion after a large study evolving monthly data from 1971 and 2008 gathered from a total of 603 meteorological stations around China: *RH* was the most sensitive variable for annual potential evapotranspiration. On the other hand, they determined that *RH* was not the determining factor on the study due to its changing trend which was statistical insignificant. These researchers also focused on the importance of studying spatial-temporal differences influencing changes in *ETo* and its driving mechanisms as in the only way of putting some light on the complex nonlinear relationships among the parameters that rule the evapotranspiration. After attribution analyses for changes in potential evapotranspiration during the past 48 years in China estimated by radiation modified Penman-Monteith FAO 56 equation, the same three authors agreed on the same conclusion: *RH* was revealed as the most sensitive variable followed by *sunshine duration*, while *minimum temperature* was pointed out as the least sensitive variable. They noticed that decrements of -10% in *RH* would result in 6.2 – 7.5% increase of *ETo* apart from registering decreasing trends in *RH* in most station in China.

ESTÉVEZ *et al.* (2009, p. 78) carried out a study using daily data from 87 automated weather stations located in Andalusia (southern Spain). They introduced both systematic and random errors in order to test the sensitivity of those errors on Penman-Monteith FAO 56

equation. Sensitivity coefficients exhibited great fluctuations throughout the year, especially for *RH* and *temperature*, with their largest values found in August and December respectively. Estimated sensitivity coefficients for *RH* showed the highest spatial and temporal variability among locations.

GONG *et al.* (2006, p. 912) used a historical register composed by a data set of 150 meteorological stations with daily observations of *maximum, minimum* and *average air temperature* at 2 m height, U_2 measured at 10 m height, *RH* (2 m height) and *daily sunshine duration* for the period 1960 – 2000. Sensitivities of reference evapotranspiration to these major climatic variables were studied and large spatial and temporal variations of sensitivity were at first noticed. The study showed that *RH* was the most sensitive variable in general for the basin, followed by *shortwave radiation* and *air temperature*, which had similar sensitivities. U_2 has the least impact.

IRMAK *et al.* (2006, p. 320), after their research pointed out wind speed as a minor influent parameter on *ETo* in humid climates than under arid conditions, where small variations in wind speed may result in larger variations in *ETo* rate. Results in accordance with the ones we obtained in the current study.

Back to our results an according to what was revealed by global sensitivity analysis results, *Rs* was not considered so influential since was ranked in last position in 19 out of the whole set of 26 stations. This event may be due to the fact that global sensitivity analysis assumes all variables under analysis to de independent from one another, although this independence does not exist among the variables here mentioned. As *Rs* increases, *temperatures* make the air capable of storing larger and larger volumes of water vapour and *RH* ends up rising.

The combination of temperature and humidity rules the formation of air masses and its movements turning into wind. Another hypothesis regards to sensitivity analysis approaches based on linear regression, which can inaccurately measure sensitivity when the model response is nonlinear, with respect to its inputs, as is the case of *ETo.* In such cases, variance-based measures are more appropriate, as Morris. May be some interactions can occur when the perturbation of two or more inputs simultaneously causes variation in the output greater than that of varying each of the inputs alone. Such interactions are present in any model that is non-additive, but would be neglected by one-at-a-time perturbations, like GSA is.

The effect of interactions can be measured by the total-order sensitivity index, a kind of index Morris method computes. Aside from these arguments to the detriment of GSA, its results are valuable as restrain the most sensitive variables to two: *Tmax* and *RH*.

CHAPTER IV

REMOTE SENSING: THE CUTTING-EDGE TOOL ON EVAPOTRANSPIRATION ESTIMATES

1. A REVIEW OF SURFACE ENERGY BALANCE MODELS

Along the present study, we have pointed out several times the importance of evapotranspiration as a key concept within hydrological cycle. We also have exposed the reasons why its estimation, computation or measurement can offer a valuable information to manage water supply to those bodies in charge of such task. Water is supplied naturally in the form of rainfall; part of it is evapotranspirated and returns to the atmosphere so as to keep running with the never-ending hydrological cycle. An important part is incorporated into crops and other commodities and can be tracked nowadays thanks to an innovative concept that aims to estimate water consumption on every product or service: Water Footprint (MAUPIN *et al.*, 2014, p. 151).

Another part is ingested (and most times wasted) by livestock and humans, so consequently is unavailable for other demands on a water supply (MAUPIN *et al.*, 2014, p. 151). Most water consumption is through actual evapotranspiration (*ETa*), which is an important component of the water cycle. For instance, in the United States, it has been estimated that about 70 % of precipitation on land returns to the atmosphere through *ETa* (CARR *et al.*, 1990, p. 261). In addition, also in the United States, more than 80 % of water consumption is for agriculture (CARR *et al.*, 1990, p. 261), most of which stems from *ETa*. Therefore, water resource users and managers have a vested interest in accurately determining consumptive water use, especially when considering the effect of population growth and climate change on water demand and supply. Distribution of water resources depends on knowing the volume of water that initially is available for use and knowing how much of that water is consumed, thus making it unavailable for additional uses.

As we could confirm throughout the present study, many approaches have been developed for measuring or estimating *ETo in last decades*, which constitutes a large portion of consumptive water use (ALLEN *et al.*, 2011, p. 920). In chapter I, we reviewed some direct methods of measuring *ETo*: lysimeters. However, using these instruments usually involves

substantial expense and effort, aside from requiring well-trained personnel. Mostly, lysimeters are exclusively used by academic institutions for research purposes.

ETo can also be measured indirectly by using evaporation pans (EP), Bowen ratio (BR) stations, eddy covariance (EC) techniques, and scintillometers. Local, state, and federal water management regulatory agencies need good-quality water-use estimates for different land surfaces to assess short and long-term water management, planning, and allocations on a watershed scale. The ability to accurately estimate the magnitude of *ETo*, which at some locations represent the main water flux, would go a long way toward computing the water balance and planning the use of available water resources. However, and according to PEACOCK and HESS (2004, p. 289) is the most complex flux to quantify. The reason lies on the fact that evapotranspiration is highly dynamic in space and time because of the complex interaction of soil, vegetation, and climate, as we already highlighted on previous chapters.

For a review of the usage of evaporation pans and lysimeters, we encourage the reader to be back to chapter I, where this issue was widely described, among others.

Bowen ratio and eddy correlation techniques offer alternatives for measuring surface energy fluxes, including evapotranspiration, at a footprint scale. Despite the high accuracy of these techniques, they may not be practical when quantifying water use at regional scales due to the number of measurement sites needed and the operational expense of such a dense network. With respect to BR method, it has been applied both in agricultural crops and natural vegetation (FRITSCHEN, 1965, p. 48; INMAN-BAMBER E MCGLINCHEY, 2003, p. 138; SCOTT *et al.*, 2003, p. 335; LEE *et al.*, 2004, p. 124; TEIXEIRA *et al.*, 2007, p. 42). Examples of EC measurements can also be found (CLEVERLYE *et al.*, 2002, p. 197; HUMPHREYS *et al.*, 2003, p. 125; LUND; SOEGAARD, 2003, p. 144; PRUEGER *et al.*, 2004, p. 100; Teixeira *et al.* 2008, p. 127). However, the spatial variability of these energy fluxes is significant and extrapolation of energy balance data directly from flux tower to a surrounding landscape environment can lead to inaccurate regional estimates.

On the other hand, aiming to increase the spatial scales, *ETo* was obtained by means of scintillometers across several kilometres paths (MEIJNINGER *et al.*, 2002, p. 83), but these field methods cannot provide a fair sample of a whole biome (WYLIE *et al.*, 2003, p. 255). The use of these instruments, although requiring less expense and training than directly

measuring *ETo*, still implies considerable workload. Additionally, these direct and indirect measurements of *ETo* are limited to the sites and times at which they are taken.

A simple technique for estimating *ETo* over larger extents and longer time periods involves the use of crop coefficients (*Kc*) (ALLEN *et al.*, 1998, p. 109). A crop coefficient (*Kc*) is a factor that relates *ETo* of a plant to that of a reference state by parameterizing several characteristics of the plant and the soil. Crop coefficients have been developed for numerous plant species. This technique can be scaled to larger extents or longer time periods and transferred among sites since crop coefficients are fixed parameters, although a crop type may have several factors depending on the number of growth stages (for example, initial and development). Applying this technique more broadly, however, is difficult due to complications with determining crop types or growth stages from aerial photography or satellite imagery. Furthermore, this technique makes a questionable assumption that local conditions affecting parameters are spatially homogenous. Despite these limitations, the crop coefficient technique is still used worldwide because of its simplicity (ALLEN, 2000, p. 41; ALLEN *et al.*, 2005, p. 13).

Surface Energy Balance Algorithm for Land (SEBAL) procedure consists of a suite of algorithms able to solve the complete energy balance equation:

$$\boldsymbol{\lambda}ET = Rn - G - H \qquad \textbf{(75)}$$

where:
$\boldsymbol{\lambda}ET$ = latent heat flux (the energy used to evaporate water);
Rn = net radiation at the surface;
G = soil heat flux;
H = sensible heat flux to the air.

Rn is computed for each pixel using albedo and transmittances computed from short wave bands and using long wave emission computed from the thermal band. *G* is predicted using vegetation indices computed from combinations of bands and net radiation. *H* is calculated from several factors: surface temperature and a wind speed measurement from

ground data, and estimated surface roughness and surface-to-air temperature differences predicted from vegetation indices.

All computations are made specific to each pixel in the image. Iterative predictions of sensible heat are improved using atmospheric stability corrections based on Monin-Obukhov. Endpoints for H within a satellite image are bounded by known evaporative conditions at key reference-points. These reference points include pixels having little or no evaporation where $H \simeq Rn - G$ (burned or desert areas) and pixels representing well-irrigated areas, where $H \simeq 0$ so that $\boldsymbol{\lambda}ET \simeq Rn - G$. More details about the energy balance approach and its application are fully explained in section 5 (Materials and Methods).

2. THE ROLE OF REMOTE SENSING ON HYDROLOGY

The demand for quantification of evapotranspiration over large areas is growing. Estimates based on satellite remote sensing offer reasonable means for meeting this demand. The advantage of applying remote sensing-based evapotranspiration procedures is that the water used by the soil-water-vegetation system can be derived directly without the need to quantify other hydrological processes. Therefore, the remote sensing-based estimation of evapotranspiration has potential in quantification of large-scale water balances; one of the main reasons why this field is growingly being exploited. Satellite remote sensing overcomes these issues with a broad spatial coverage and the potential exists for indirect evapotranspiration measurement (VÖRÖSMARTY *et al.*, 2000, p. 288).

Remote sensing has been a useful tool in developing methods for estimating consumptive water use from *ETo* that are scalable and transferable. This fact turns essential as the apportionment of water resources is highly affected by environmental and economic circumstances differing in extent and spatiotemporal resolution. Research over many years has led to the development of remote sensing methods that are reliably reproducible and effective in estimating *ETo.* Since satellites first began collecting data on natural resources in the 1970s, researchers have been developing models to process these data for estimating *ETo* (IDSO et al., 1975, p. 349*;* JACKSON *et al.*, 1977, p. 656). Some remote sensing methods for estimating *ETo* are focused at very local scales, whereas others are focused at scales ranging from regional or continental (SENAY *et al.*, 2013, p. 591; SINGH; SENAY, 2015, p. 9) to global (MU *et al.*, 2007, p. 536). These methods also range from simple (JACKSON *et al.*, 1977, p. 656) to complex (BASTIAANSSEN *et al.*, 1998, p. 212; ALLEN *et al.*, 2007, p. 406). Several remote sensing methods can be used to estimate *ETo* at the high spatial resolution of agricultural fields and the large extent of river basins, a scale that is useful to water resource managers.

Remote sensing is, thus, an excellent mean to determine and map the spatial and temporal structure of energy balance components, including evapotranpiration, at a basin scale. Vegetative indexes, remotely obtained, along with agro-meteorological data, can be used to extrapolate evapotranspiration values from local to regional scales. Hydrological models are in general too complex and costly for this extrapolation because of the difficulty in obtaining data

sets in different hydrological uniform sub-areas and due to this deficiency, these models can yield unreliable results (NAGLER *et al.*, 2007, p. 110).

For decades scientists have been piecing together clues about Earth's climate from ancient rocks, ice cores, patterns in the growth rings of trees, historical weather records, and satellite data. Now scientists are accumulating evidence that human industrial and agricultural activities may be accelerating naturally-occurring changes in our climate, and that we are contributing to such hazards as global warming, rising sea level, ozone depletion, acid rain, and loss of biodiversity.

People all over the world, not just scientists, but also farmers, fishermen, and policymakers are asking questions about global change and the consequences for life on Earth. To answer those questions, scientists must study the Earth as an interacting system because the relationships among life, land, oceans and atmosphere are tightly interwoven. Clearly, the Earth is a complex, changing system and its global dynamics are being subject of study and deep analysis for scientists to better understand natural climate changes and human responsibility for those changes (CARR *et al.*, 1990, p. 254).

To improve global models, scientists need information, or data, for every region of the Earth, every day for many years. The most powerful way to collect such comprehensive data is through the use of satellite sensors that measure different types of energy coming from the Earth. The way the Earth reflects or emits electromagnetic energy into space gives scientists valuable information about its ecological, hydrological, and meteorological conditions.

Understanding the importance of studying global dynamics, the United States Congress instituted the U.S. Global Change Research Program in 1990. NASA's Earth Science Enterprise is making substantial contributions to the program through its Earth Observing System (EOS). Through EOS, NASA is working with the national and international scientific community to design, develop, and launch advanced satellite sensors that collect data across a wide spectrum of energy: ultraviolet, visible, infrared, and microwave. With these sensors, NASA will collect multi-year data sets that will help to answer questions about global change.

3. MODIS SENSOR

The first EOS satellite, called Terra, was launched on December 18, 1999, carrying five remote sensors. The most comprehensive EOS sensor is MODIS, the Moderate-resolution Imaging Spectroradiometer. MODIS offers a unique combination of features:

a) It detects a wide spectral range of electromagnetic energy.
b) It takes measurements at three spatial resolutions (levels of detail).
c) It takes measurements all day, every day and
d) It has a wide field of view.

This continual, comprehensive coverage allows MODIS to complete an electromagnetic picture of the globe every two days. MODIS's frequent coverage complements other imaging systems such as Landsat's Enhanced Thematic Mapper Plus, which reveals the Earth in finer spatial detail, but can only image a given area once every 16 days, although too infrequently to capture many of the rapid biological and meteorological changes that MODIS observes.

Terra is the first large, multi-instrument EOS satellite, and its orbit around the Earth is timed so that it passes from north to south across the equator in the morning. A second EOS satellite, Aqua, will also carry a MODIS instrument. Aqua will pass south to north over the equator in the afternoon. Working in tandem to see the same area of the Earth in the morning and the afternoon, the two satellites will help scientists ensure MODIS's and other instruments' measurement accuracy by optimizing cloud-free remote sensing of the surface and minimizing any optical effects, like shadows or glare, that are unique to morning or afternoon sunlight.

Having morning and afternoon sensors also permits investigation of changes that occur over the course of the day, such as the build-up or dissipation of clouds and changes in sea or land surface temperature.

MODIS instruments are designed to take measurements in spectral regions that have been used in previous satellite sensors. MODIS is adding to existing knowledge by extending data sets collected by heritage sensors such as:

1. The National Oceanic and Atmospheric Administration's (NOAA);
2. The Advanced Very High Resolution Radiometer (AVHRR), used for meteorology and monitoring sea surface temperature, sea ice, and vegetation;
3. The Coastal Zone Color Scanner (CZCS);
4. The Sea-viewing Wide Field of View Sensor (SeaWiFS), used to monitor ocean biological activity; Landsat, used to monitor terrestrial conditions and
5. NOAA's High Resolution Infrared Radiation Sounder (HIRS), used to observe atmospheric conditions.

(Source: <https://modis.gsfc.nasa.gov/about/media/modis_brochure.pdf>)

3.1 What can MODIS offer us?

By extending these data sets, MODIS promotes the continuity of data collection, a essential feature for understanding both long- and short-term change in the global environment. Some of the most easily recognizable changes are occurring on land. Human-induced changes such as deforestation, urbanization, and hydroelectric and irrigation projects combine with the Earth's existing cycles of fire, erosion, and floods to change our landscape. While we can often see these changes happening on a local scale, it is impossible to assess global effects through fieldwork alone. With its near-daily coverage of the Earth's surface, MODIS provides comprehensive measurements that scientists and land managers need to make informed decisions about managing Earth's natural resources from season to season, year to year, and decade to decade.

Monitoring and assessing conditions on the Earth's surface is critical to understanding the impacts of weather and climate change and human activities. MODIS provides global maps

of several land surface characteristics, including surface reflectance, albedo (the percent of total solar energy that is reflected back from the surface), land surface temperature, and vegetation indices. Vegetation indices tell scientists how densely or sparsely vegetated a region is and help them to determine how much of the sunlight that could be used for photosynthesis is being absorbed by the vegetation. The maps will provide the basis for MODIS's real-time global monitoring of subtle changes in vegetation that may signal biospheric stress, such as pollution, drought, or temperature extremes, which in turn could be used to predict and prevent wildfire danger or crop failure. MODIS provides data for land cover maps that tell scientists not only whether an area is vegetated, but also what kind of vegetation is growing there, separating coniferous forests from deciduous forests, or cropland from grassland. In addition to 11 categories of vegetation, the maps recognize various non-vegetated surfaces, including bare soil, water, and urban areas; 17 land cover types in all.

(Source: <https://modis.gsfc.nasa.gov/about/media/modis_brochure.pdf >)

3.2 MODIS, an essential tool for a multidisciplinary perspective

MODIS's high-quality, daily measurements allow scientists to track changes in land cover types and land use, to determine where forested land is becoming deforested, where grassland is becoming cropland, or where burned land is returning to natural vegetation. MODIS collects data on land surface characteristics that are crucial for modeling the Earth's carbon cycle, or the exchange of carbon between the Earth's life, land, oceans, and atmosphere.

Modeling the carbon cycle is especially important given increases in atmospheric carbon dioxide, a known greenhouse gas. Land surface variables such as amount and type of vegetation, soil composition, and temperature play a significant role in calculating mass and energy exchange between the Earth's surface and the atmosphere. Since plants remove carbon dioxide from the atmosphere as they make food, scientists call them a carbon sink; when vegetation is burned, it becomes a carbon source. Estimates of carbon storage in plants vary widely, but MODIS data will help reduce that uncertainty by allowing scientists to make more accurate estimates of global photosynthesis and productivity.

MODIS measures the photosynthetic activity of land and marine plants (phytoplankton) to yield better estimates of how much of the greenhouse gas is being absorbed and used in plant productivity. Coupled with the sensor's surface temperature measurements, MODIS' measurements of the biosphere are helping scientists track the sources and sinks of carbon dioxide in response to climate changes.

MODIS measures the properties of clouds such as the distribution and size of cloud droplets in both liquid water and ice clouds. MODIS also measures the properties of aerosols; these enter the atmosphere from manmade sources like pollution and biomass burning and natural sources like dust storms, volcanic eruptions, and forest fires. MODIS helps scientists determine the amount of water vapour in a column of the atmosphere and the vertical distribution of temperature and water vapour, measurements crucial to understanding Earth's climate system.

MODIS maps the areal extent of snow and ice brought by winter storms and frigid temperatures. The sensor observes the "green wave" that sweeps across continents as winter gives way to spring and vegetation blooms in response. It sees where and when disasters strike, such as volcanic eruptions, floods, severe storms, droughts or wildfires and will hopefully help people get out of harm's way.

MODIS' bands are particularly sensitive to fires; they can distinguish flaming from smoldering burns and provide better estimates of the amounts of aerosols and gases fires release into the atmosphere.

MODIS sees changes in the Pacific phytoplankton populations that may signal the onset of the famous El Niño/La Niña climatic siblings well ahead of their arrival. In turn, by coupling its sea surface temperature and ocean colour measurements, MODIS has observed the impacts El Niño and La Niña have on the microscopic marine plant. MODIS also has a unique channel for measuring chlorophyll fluorescence. All plants bombarded with light begin to glow, or fluoresce, but in wavelengths that our eyes cannot see. The more plants fluoresce, the less energy they are using for photosynthesis. Thus, MODIS not only maps the distribution of phytoplankton, it also helps us gauge its health.

(Source: <https://modis.gsfc.nasa.gov/about/media/modis_brochure.pdf>)

3.3 Technical specifications

MODIS has a viewing swath width of 2,330 km, and it images the entire Earth each 1 to 2 days in 36 spectral bands, or groups of wavelengths, ranging from 0.405 to 14.385 µm. It collects data at three spatial resolutions: 250, 500, and 1,000 meters. The average rate of data collection is 6.1 megabits each second. Using physically and empirically based algorithms, high-speed computers process MODIS data to yield 44 global data products describing many of Earth's vital signs, from ocean and land surface temperatures to the physical properties of clouds. Consequently, MODIS tracks a wider array of the earth's vital signs than any other Terra sensor. For instance, the sensor measures the percent of the planet's surface that is covered by clouds almost every day. This wide spatial coverage enables MODIS, together with Multi-angle Imagery SpectroRadiometer (MISR) and Clouds and Earth's Radiant Energy System (CERES), to help scientists determine the impact of clouds and aerosols on the Earth's energy budget.

(Source: <https://modis.gsfc.nasa.gov/>)

4. ACCESS TO DATA

A) Data Ordering Tools & Resources

Listed below are a variety of tools and resources that can be utilized with MODIS Data. Each listing provides a link to the tool or resource and a short summary of its use.

AppEEARS - https://lpdaacsvc.cr.usgs.gov/appeears/

Built on top of middleware services, the AppEEARS interface enables users to input precise sample locations - such as a field study site or a flux tower - and access analysis-ready data from land MODIS products held by NASA's LP DAAC. AppEEARS provides interactive time series and scatter plots, allowing users to preview and interact with their query results before downloading them. Users are also able to view quality information and pixel values in table format in addition to the interactive plots.

Earth Data - http://earthdata.nasa.gov/

Earth Data is a website containing a vast amount of information on use and access of all NASA Earth Observing System Data and Information System (EOSDIS) data products. All MODIS data products can be accessed from the website. The Earth Data website provides many user resources, including tutorials, webinars, as well as data search, discovery and processing information. Discipline specific information (e.g. Atmosphere, Cryosphere, Land, Ocean, Human Dimensions) is also featured on the website.

EarthExplorer (EE) - http://earthexplorer.usgs.gov/

EarthExplorer is an U.S. Geological Survey (USGS) data search and order website that provides access to multi-sensor satellite and airborne data sets in the long-term earth science USGS data archive. The MODIS Land data products available from this websites include Land Surface Reflectance (M*D09), Land Surface Temperature and Emissivity (M*D11), Land Cover (MCD12), Vegetation Indices (M*D13, MOD44A,B), LAI/FPAR (M*D15), Water Mask (MOD44W), Thermal Anomalies and Fire (M*D14, MCD45), Gross Primary Productivity (M*D17) and BRDF and Albedo (MCD43).

EOS Clearing House (ECHO) / Reverb - http://reverb.echo.nasa.gov/reverb

The Reverb | ECHO website is a metadata and data service search tool. The Reverb | ECHO website contains a data catalog of NASA Earth Observing System data and a registry for related data services (e.g. reformatting, pattern recognition). There are 200 MODIS related data sets available through Reverb. These MODIS data sets include standard products as well as in situ and satellite comparison data sets (e.g. SAFARI 2000 campaign, with ASTER and MODIS fire data).

Giovanni - http://giovanni.gsfc.nasa.gov/giovanni/

Giovanni and Giovanni 4 are Web applications developed by the GES DISC to provide a simple, intuitive way to visualize, analyze, and access Earth science remote sensing data, including large data volumes, particularly from satellites, without having to download the data. The MODIS Atmosphere data products are available to analyze with this tool. Giovanni 4 is the next generation of Giovanni, designed to be faster, more interactive and easier to learn than its predecessor.

Global Change Master Directory (GCMD) - http://gcmd.gsfc.nasa.gov/

NASA's Global Change Master Directory (GCMD) provides a directory listing of Earth science data sets and service descriptions, including all MODIS data products. The GCMD is one of the largest public earth and environmental science metadata inventories currently known.

HDF-EOS to GeoTIFF converter (HEG) - http://hdfeos.org/software/heg.php

The HDF-EOS to GeoTIFF converter (HEG) tool converts MODIS hdf formatted data files into GeoTIFF, native binary or HDF-EOS Grid. The HEG tool also has reprojection, resampling, subsetting, mosaicking and metadata creation capabilities.

MODIS Land Global Browse Images - http://landweb.nascom.nasa.gov/cgi-bin/browse/browse.cgi

MODIS Level 1 and select Land data products are available as quick browse images at global resolution. These images allow for a rapid synoptic quality assessment of data. Collections 4, 5, and 6 are included for browsing as available per product.

MODIS Level 1 Atmosphere Archive and Distribution System (LAADS Web) - https://ladsweb.nascom.nasa.gov

LAADS Web is the web interface to the Level 1, Atmosphere Archive and Distribution System. LAADS Web provides quick and easy access to all MODIS Level 1, 2 and 3 Atmosphere and Land data products with a number of post processing options. Post processing options include subset by parameter, area or band, mosaicked, reprojected or masked. The website also provides quick look true colour RGB and false colour images of selected data sets.

MODIS Swath-to-Grid Toolbox (MS2GT) - https://nsidc.org/data/modis/ms2gt/index.html

The MODIS Swath-to-Grid Toolbox (MS2GT) reads MODIS swath HDF-EOS data files and produces gridded data binary files. MS2GT can output data in a variety of map projections. MS2GT can also produce a seamless output grid from multiple MODIS swath input files.

MODIS Swath Reprojection Tool (MRT Swath) and MODIS Reprojection Tool (MRT)

https://lpdaac.usgs.gov/tools/modis_reprojection_tool_swath

https://lpdaac.usgs.gov/tools/modis_reprojection_tool

The MODIS Swath Reprojection Tool reads MODIS swath HDF-EOS files and produces binary HDF-EOS Grid or GeoTiff files of gridded data in different map projections. The MODIS Reprojection Tool reprojects MODIS Level-2G, Level 3 and Level 4 land data products to a user specified output map projection type and output resolution. The MODIS Reprojection Tool is developed and maintained by the LP-DAAC and is available for free to all registered users.

MODIS Web Service - https://daac.ornl.gov/MODIS/MODIS-menu/modis_webservice.html

The MODIS Web Service is a tool that provides users with a means to order subsets of MODIS Land Products through standards based SOAP (Simple Object Access Protocol) established open source formats. Examples of currently supported open source formats: (1) OPeNDAP DAP, (2) OpenSearch and (3) Web Map Tile Service. The Open-source Project for a Network Data Access Protocol (OPeNDAP) Data Access Protocol (DAP) is a method for requesting and moving data on the web. Several of the DAACs including Physical Oceanography DAAC, GES DISC and LaRC offer the capability to use OPeNDAP to retrieve MODIS data products. OpenSearch is a protocol or format structure used to communicate

search requests and search results. OpenSearch is used at LANCE-MODIS and LAADS Web. Web Map Tile Service (WMTS) is a protocol or format to host georferenced map tiles. WMTS is used with NASA EOSDIS Global Imagery Browse Services (GIBS) for serving MODIS data.

ORNL DAAC MODIS Land Products Subsetting and Visualization Tools

http://daac.ornl.gov/modisglobal

The ORNL DAAC MODIS Land Products Subsetting and Visualization Tools allow users to request subsets of specific MODIS Land Products synoptically or programmatically. The subsetter tool outputs data in ASCII format, and provides the option to visualize data in temporal resolution with graphs and statistic outputs.

(Source: <https://modis.gsfc.nasa.gov/>)

B) MODIS ready-to-use Evapotranspiration product MOD16

This project is part of NASA/EOS and aims to estimate global terrestrial evapotranspiration from earth land surface by using satellite remote sensing data. MOD16 global evapotranspiration product can be used to calculate regional water and energy balance, soil water status; hence, it provides key information for water resource management. With long-term ET data, the effects of changes in climate, land use, and ecosystems disturbances (e.g. wildfires and insect outbreaks) on regional water resources and land surface energy change can be quantified.

The MOD16 global evapotranspiration (ET)/latent heat flux (LE)/potential ET (PET)/potential LE (PLE) datasets are regular 1 km^2 land surface ET datasets for the 109.03 Million km^2 global vegetated land areas at 8-day, monthly and annual intervals. The dataset covers the time period 2000 through 2010. Future years will be produced and posted periodically, but not in near-real time.

(Source: <https://modis.gsfc.nasa.gov/>)

5. MATERIALS AND METHODS

This section provides an overview of a remote sensing method used for estimating water consumed through *ETa*, Mapping Evapotranspiration at High Resolution with Internalized Calibration (METRIC) and is discussed in detail, including the theory behind the continued improvement of these method and some of its applications in *ETa* estimation. Various qualities of this method, including the extent and spatiotemporal resolution of model estimates and its accuracy, the cost, and the ease of implementation, are also discussed along the present section. With it, we intended to provide a synopsis of a recently developed technique that currently seem most applicable to *ETa* estimation at scales appropriate for water resource management, along with a discussion of the potential for cloud computing to enable the operability of these techniques over large extents at high spatiotemporal resolution.

The pioneer methodology in computing *ETa* by applying an energy balance was Surface Energy-Balance Algorithm for Land (SEBAL), from which METRIC was developed. A critical issue with SEBAL is, however, defining anchor pixels. When anchor pixels are not present on the imagery, assumptions may not hold, and then the SEBAL approach has been less than satisfactory (SINGH *et al.*, 2008, p. 122). To solve this dilemma, ALLEN *et al.* (2007a, p. 406, 2007b, p. 394) introduced the METRIC (Mapping Evapotranspiration at High Resolution using Internalized Calibration) model using similar principles to those of SEBAL, but with refinements to anchor pixels and various energy balance components. They introduced a better proxy for scaling satellite overpass time evaporation (*ETrF*) to 24-hour periods for the well-watered and fully vegetated areas of the image which represent the alfalfa-based reference evapotranspiration surface (*ETr*).

In METRIC, *LE* at the cold pixel with full canopy cover ($LAI > 4$) is taken as 1.05 *ETr*, where *ETr* is the alfalfa-based reference evapotranspiration. For the hot pixel, the *LE* estimate is based on a bare soil water balance using the FAO 56 approach (Allen *et al.*, 1998, D05109). METRIC estimates a so-called "fraction of reference *ET* (*ETrF*)" defined as the "ratio of instantaneous *ET* to the reference crop evapotranspiration (*ETr*)" (ALLEN *et al.*, 2007a, p. 406; ALLEN *et al.*, 2007b, p. 394). The *ETrF* estimates from METRIC using remotely sensed data are instantaneous; even so, it allows estimating *LE* on a 24-hour basis. The daily reference crop evapotranspiration (*ETr24*) is multiplied with *ETrF* to estimate daily *ET* at each pixel of the

satellite scene. *ETrF* is similar to evaporative fraction used by BASTIAANSSEN *et al.* (1998, p. 212).

For comparison purposes, theoretical and computational basis of SEBAL is described in (BASTIAANSSEN *et al.*, 1998, p. 212; BASTIAANSSEN, 2000, p. 100; BASTIAANSSEN *et al.*, 2005, p. 93).

5.1 METRIC algorithm steps and considerations

As none of the satellites measure vapour content near surface, *ET* is generally determined from satellite imagery by applying an energy balance at the surface, where *ET* is calculated as a residual of the surface energy equation (Equation 1) which has already been presented back in section 1:

$$LE = Rn - G - H \quad \textbf{(75)}$$

where:
LE = latent energy consumed by *ET* (W m^{-2});
Rn = net radiation (sum of all incoming and outgoing short-wave and long-wave radiation at the surface) (W m^{-2});
G = sensible heat flux conducted into the ground (W m^{-2});
H = sensible heat flux convected to the air (W m^{-2}).

It is important to draw readers' attention to the fact that METRIC is fed with hourly measures of the parameters needed. This, along with the application of an Energy Balance method allow the algorithm to produce *actual/instantaneous ET* rather than *potential ET*. That would allow us to highlight water shortage periods. Accuracy on *Rn*, *G*, and *H* computing is essential to reach accurate *LE* values, as it depends on the formers. Avoiding this dependency on *LE*, METRIC calibration is carried out on *H* instead of *LE* so as to cancel out all intermediate estimation errors and biases.

In the following, we will see how the vast number of variables which compound the model are computed and which other inputs we would need. To start from, most of the algorithms used in METRIC for *Rn* and *G* computation stem from those used in early SEBAL applications by (BASTIAANSSEN *et al.*, 1998, p. 212). However, albedo in METRIC was updated following (TASUMI *et al.*, 2007, p. 89) to improve accuracy over a wide range of surface conditions.

Net Radiation (Rn)

Actual net radiation flux at the surface (Rn) represents the radiant energy at the surface that is partitioned into H, G, and LE. Rn is computed by subtracting all outgoing radiant fluxes from all incoming radiant fluxes, including solar and thermal radiation.

$$Rn = RS\downarrow - \_RS\downarrow + RL\downarrow - RL\uparrow - (1 - \mathcal{E}o)\, RL\downarrow \quad \textbf{(76)}$$

where:

$RS\downarrow$ = incoming short-wave radiation (W m^{-2});

α = surface albedo (dimensionless);

$RL\downarrow$=incoming long-wave radiation (W m^{-2});

$RL\uparrow$=outgoing long-wave radiation (W m^{-2});

$\mathcal{E}o$=broad-band surface thermal emissivity (dimensionless).

Incoming Solar Radiation (Rs↓)

Incoming broad-band short-wave radiation ($Rs\downarrow$) represents the main source of energy responsible for turning the evapotranspirative process on. For images comprising areas not larger than 25,000 km² and with smooth slope, $Rs\downarrow$ can be calculated as a constant as it follows:

$$R_s \downarrow = \frac{G_{sc} * \cos\theta_{rel} * \tau_{sw}}{d^2} \quad \textbf{(77)}$$

where:

G_{sc} = solar constant (1367 W m^{-2});

θ_{rel} = solar incidence angle;

d^2 = square of the relative Earth–Sun distance;

τ_{sw} = broad-band atmospheric transmissivity.

Computed *Rs*↓ for clear sky conditions (*Rso*) is considered to have the same or better accuracy as measured *Rs*↓ from an automated weather station (ALLEN, 1996, p. 106).

Broad-band atmospheric transmissivity (τ_{sw})

Broad-band atmospheric transmissivity τ_{sw} is calculated using a simple equation:

$$\tau_{sw} = 0.75 + 2*10^{5} * z \quad \textbf{(78)}$$

where z = elevation above sea level (m).

That means that there is need of each pixel altitude composing the image. The best option in this case is to arrange a Digital Elevation Model (DEM) where z-values (altitudes) are estimated based on a Projected Reference System. In our case, the DEM has been used throughout the present study was downloaded from Shuttle Radar Topographic Mission (SRTM) website, where it is easy to find DEM's for vast areas over the entire world.

As far as the incidence angle (θ_{rel}) is concerned, it refers the angle between the solar beam and a vertical line perpendicular to the land surface for uneven surfaces. For horizontal flat surfaces, θ_{rel} is simply equivalent to solar zenith angle. However, for sloping surfaces, θ_{rel} must be calculated pixel by pixel, using the surface slope and aspect information derived again from a DEM and the following equation from (DUFFIE; BECKMAN, 2013, p. 1324):

$$\cos\theta_{rel} = \sin(\delta)\sin(\phi)\cos(s) - \sin(\delta)\cos(\phi)\sin(s)\cos(\gamma) + \cos(\delta)\cos(\phi)\cos(s)\cos(\omega) + \cos(\delta)\sin(\phi)\sin(s)\cos(\gamma)\cos(\omega) + \cos(\delta)\sin(\gamma)\sin(s)\sin(\omega) \quad \textbf{(79)}$$

where:

δ = declination of the Earth (radians);

ϕ = latitude of the pixel (radians);

s = surface slope, always positive and representing the downward slope in any direction (s = 0 for flat surfaces, $s = \pi/2$ radians for vertical downward slope) (radians);

γ = surface aspect angle (γ = 0 for slopes oriented to south, $\gamma = -\pi/2$ radians for slopes oriented to east, $\gamma = \pi/2$ radians for slopes oriented to west, and $\gamma = \pm\pi$ radians for slopes oriented to north (radians);

ω = hour angle (ω = 0 at solar noon, ω is negative in the morning, and positive in the afternoon (radians).

Surface Albedo (α)

Surface albedo (α) is the ratio of incoming short-wave solar radiation that is reflected by the surface and represents the integrated reflectance across the short-wave spectrum (0.2 to 3.2 µm.). Albedo is calculated as integration of surface reflectivities with weighting functions of the corresponding bands (bands 1–5 and 7 of Landsat, bands 1–9 of ASTER and bands 1-5 and 7 of MODIS). Therefore, albedo is calculated as:

$$\alpha = \sum_{b=1}^{n}\left(\rho_{s,b} w_b\right) \quad \textbf{(80)}$$

where w_b represents the weights/coefficients of each band and fully described in Liang (2001). For MODIS applications, Equation 80 turns into the following:

$$\alpha = 0.16*b1 + 0.291*b2 + 0.243*b3 + 0.116*b4 + 0.112*b5 + 0.08*b7 - 0.0015\pi \quad \textbf{(81)}$$

where b1, b2, b3, b4, b5 and b7 are the different MODIS bands.

Outgoing long-wave radiation ($RL\uparrow$)

Outgoing long-wave radiation (*RL*↑) emitted from the surface is driven by surface temperature and surface emissivity as it follows:

$$R_L \uparrow = \varepsilon_o \sigma T_s^4 \quad \textbf{(82)}$$

where $\mathcal{E}_o$ = broad-band surface emissivity (dimensionless);
σ = Stefan-Boltzman constant (5.67×10^{-8} W m^{-2} K^{-4});
Ts = surface temperature (K°)

Broad-band Surface emissivity ($\mathcal{E}_o$)

Broad-band Surface emissivity ($\mathcal{E}_o$) is computed using an empirical equation by TASUMI (2003, p. 412) based on soil and vegetative thermal spectral emissivities housed in the MODIS UCSB Emissivity Library (2004) where:

$\mathcal{E}_o = 0.95 + 0.01 * LAI$, for those pixels with $LAI \leq 3$ and **(83a)**
$\mathcal{E}_o = 0.98$, for those pixels with $LAI > 3$ **(83b)**

where *LAI* = leaf area index, which is the ratio of the total leaf area per unit of ground area (m^2 m^{-2}).

LAI is an indicator of biomass and canopy resistance to vapour flux, and is computed using an empirical equation stemming from BASTIAANSSEN *et al.* (1998, p. 212)

$$LAI = -\frac{\ln\left[{}^{(0.69-SAVI)}/_{0.59}\right]}{0.91} \quad \textbf{(84)}$$

$$SAVI = \frac{(1+L)(NIR-R)}{L+(NIR+R)} \quad \textbf{(85)}$$

where *SAVI* = soil adjusted vegetation index (dimensionless);
L = constant, often set to *L* = 0.5 as it was here.

LAI computation brings out maximum values of 6.0 and 0.69, for *LAI* and *SAVI*, respectively. Beyond this, *LAI* remains steady at a maximum value and does not change significantly. The accuracy of *LAI* estimation depends on site and soil characteristics and crop types (BASTIAANSSEN *et al.*, 1998, p. 212). In other words, *LAI* and *SAVI* calculations may require local or regional calibration. Even though, the impact of errors in *LAI* on the METRIC energy balance computation is negligible.

Surface Temperature (T_s)

As regards to Surface Temperature (T_s), it is important to draw the reader's attention to the fact that MODIS sensor is capable of measuring this parameter and it is accessible towards its digital platform offering a ready-to-use surface temperature product: MOD11A1/A2. For more details, please, find the information about this product here: (https://modis.gsfc.nasa.gov/data/dataprod/mod11.php).

Incoming long-wave radiation ($RL\downarrow$)

Incoming long-wave radiation is the downward thermal radiation flux originating from the atmosphere and is traditionally computed using the Stefan–Boltzmann equation, as it follows:

$$RL\downarrow = \mathcal{E}a\ \sigma\ T_s^4 \quad \textbf{(86)}$$

where $\mathcal{E}a$ = effective atmospheric emissivity (dimensionless);
T_s = near-surface air temperature (K°).

Atmospheric emissivity ($\mathcal{E}a$)

An empirical equation for $\mathcal{E}a$ developed by BASTIAANSSEN *et al.* (1998, p. 212) is applied by using coefficients developed by ALLEN (2000, p. 587) and data collected over alfalfa crops in Idaho.

$$\mathcal{E}a = 0.85(-\ln \tau sw)^{0.09} \quad \textbf{(87)}$$

where τsw = broad-band atmospheric transmissivity for shortwave radiation calculated from Equation 78.

Soil heat flux (*G*)

Soil heat flux (*G*) is the rate of heat storage in the soil and vegetation due to conduction. General METRIC applications compute *G* as a ratio *G*/*Rn* using an empirical equation by BASTIAANSSEN (2000, p. 564) representing values near midday as it follows:

$$G/Rn = (Ts - 273.15)*(0.0038 + 0.0074\alpha)*(1 - 0.98\ NDVI^4) \quad \textbf{(88)}$$

where:
Ts = effective surface temperature (Kº) gathered from MOD11A1/A2 MODIS product;
α = surface albedo, computed as in Equation 81.

An alternative option to compute *G* that has been successfully applied in METRIC was developed by TASUMI (2003, p. 745) as it follows:

$$G/Rn = 0.05 + 0.18e^{-0.521\ LAI} \quad \text{when } LAI \geq 0.5 \quad \textbf{(89a)}$$

$$G/Rn = 1.80(Ts - 273.15)/Rn + 0.084 \quad \text{when } LAI < 0.5 \quad \textbf{(89b)}$$

These equations have been used in their origin and predicted relatively accurate *G* values for irrigated crops (TASUMI, 2003, p. 745). It suggests that *G*/*Rn* increases with increasing albedo (indicative of bare soils that often have high reflection) and decreases with increasing vegetation (due to shading by the canopy). It also suggests that *G*/*Rn* decreases with increasing leaf area, for the same reason, and that for bare soil, *G* increases in proportion to surface temperature.

Sensible heat flux (*H*)

Sensible heat flux is the portion of energy used in heating up the air and it is calculated by using the following equation:

$$H = \frac{\rho_{air} C_p dT}{r_{ah}} \quad \textbf{(90)}$$

where:

ρ_{air} = air density (kg m^{-3});

C_p = specific heat of air at constant pressure (1004 J kg^{-1} K^{-1});

r_{ah} = aerodynamic resistance (m s^{-1}) between two near surface heights, *z1* and *z2* (generally 0.1 and 2 m) computed as a function of estimated aerodynamic roughness of the particular pixel.

Air density (ρ_{air})

Air density (ρ_{air}) in the aerodynamic equation is calculated using standard equations for mean atmospheric pressure (*P*) and the universal gas law and simplifying for the effect of vapour pressure

$$\rho_{air} = \frac{101325 * \left(1 - 1.2184 * 10^{-5} * z\right)^{5.26}}{T_S * R} \quad \textbf{(91)}$$

where:

ρ_{air} = air density (kg m^{-3});

z = pixel altitude (m)

R = specific gas constant (287 J kg^{-1} K^{-1});

Ts = surface air temperature at the pixel (K^{o})

Aerodynamic transport (r_{ah})

In METRIC, aerodynamic transport (r_{ah}) calculation uses wind speed values extrapolated from some blending height above the ground surface (typically 100 to 200 m) and an iterative stability correction scheme based on the Monin–Obukhov functions (ALLEN *et al.*, 2007, p. 394).

$$r_{ah} = \frac{\ln\left(z_2/z_1\right)}{k u_*} \quad \textbf{(92)}$$

where:

z1 and *z2* = heights above the zero-plane displacement of the vegetation where the endpoints of *dT* are defined;

u_* = friction velocity (m s^{-1});

k = von Karman's constant (0.41).

Friction velocity (u_*)

Friction velocity (u_*) is computed by using the logarithmic wind law assuming neutral atmospheric conditions at blending height.

$$u_* = \frac{k u_{200}}{\ln\left(200/z_{0m}\right)} \quad \textbf{(93)}$$

where:

u_{200} = wind speed (m s^{-1}) at a blending height, usually at 200 m.

z_{om} = momentum roughness length (m)

Wind speed at blending height (u_{200})

The wind speed at an assumed blending height, 200 m. above the weather station (u_{200}) is calculated as it follows:

$$u_{200} = \frac{u_w \ln\left(200/z_{0m}\right)}{\ln\left(z_x/z_{0m}\right)} \quad \textbf{(94)}$$

where:

u_w = wind speed (m s^{-1}) measured at a weather station at z_x height above the surface;

z_{0m} = momentum roughness length for the weather station surface (m).

The value for u_{200} is assumed constant for the satellite image. This assumption allows for the use of a constant relation between *dT* and *Ts*.

Momentum roughness length (Z_{0m})

In METRIC, momentum roughness length (z_{om}), measured in meters, is estimated for each pixel according to land use type or amount of vegetation. In agricultural areas, z_{om} and, in turn, crop height is generally proportional to the *LAI*, and is therefore calculated as a function of *LAI* for "short" agricultural crops (less than about 1 m in height) using the following relationship from TASUMI (2003, p. 742):

$$z_{0m} = 0.018LAI \quad \textbf{(95)}$$

where *LAI* = leaf area index (dimensionless) and computed as in Equation 84.

There is an alternative way of computing Z_{0m} due to its complexity when computing it among a large wide of crop types within the same image. This alternative way is based on NDVI index, previously computed and can be reviewed in ALLEN (2002, p. 198).

Near-surface temperature gradient (*dT*)

The *dT* parameter (K^o) represents the near-surface temperature difference between *z1* and *z2* heights. *dT* is used in Equation 90 because of the difficulty in estimating surface temperature (T_s) accurately from satellite due to uncertainty in atmospheric attenuation or contamination and radiometric calibration of the sensor. The temperature gradient *dT* is designed to "float" above the surface, beyond the height for sensible heat roughness (z_{oh}) and zero plane displacement, and can be approximated as a relatively simple linear function of T_s as pioneered by bastiaanssen (1995a, p. 895) as in the following:

$$dT = a + bTs \quad \textbf{(96)}$$

where:

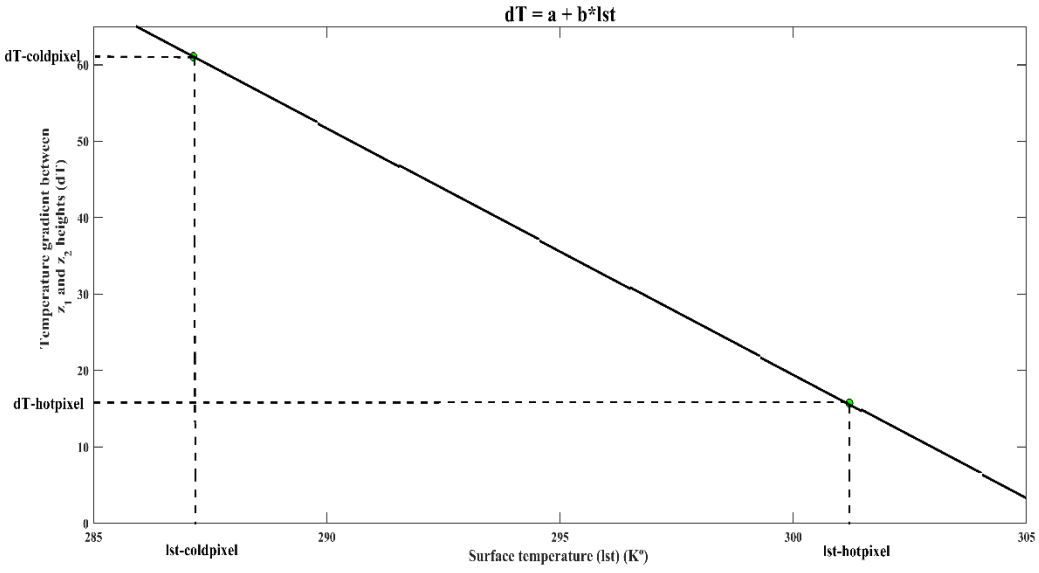
a and *b* = empirically determined constants for a given satellite image.

Ts = surface temperature (K°) of the corresponding hot/cold pixel.

The determination of *a* and *b* involves locating a hot/dry pixel (in a fallow agricultural field) with high *Ts* value and a cold/wet pixel with a low *Ts* value (typically on an irrigated agricultural setting) in the remote sensing image. These is an arduous task as it is usually done manually by the user. Failing in selecting the proper pixel may entail systematic errors and inaccurate *ETo* values. So, it is always better to be assisted by experienced users or follow a sort of procedure guide.

Both *a* and *b* can be also thought as the resultant *slope* and *intercept*, namely, when applying a linear regression between the tandems (*dT*, *Ts*) for both pixels. Here, Figure 25 illustrates the relation between *dT* and *Ts* for a better comprehension.

Figure 25. Correlation between surface temperature (lst) and temperature gradient (dT) between z1 and z2 for both reference pixels (hot and cold pixel)



Source. Plot computed bay the author from a random dT, computed previuosly

Once these pixels have been identified, the energy balance equation (Equation 75) can be solved independently for both pixels as:

$$H_{hot} = (R_n - G)_{hot} - LE_{hot} \quad \textbf{(97a)}$$

$$H_{cold} = (R_n - G)_{cold} - LE_{cold} \quad \textbf{(97b)}$$

where:

$H_{cold/hot}$ = sensible heat flux at cold/hot pixel (W m^{-2});

Rn = net radiation flux at the surface at cold/hot pixel (W m^{-2});

G = soil heat flux at cold/hot pixel (W m^{-2});

$LE_{cold/hot}$ = latent heat flux at cold/hot pixel (W m^{-2})

As far as the cold pixel is concerned, it is assumed to have latent heat flux (*LE*) equal to 1.05 times that expected for a tall reference crop, commonly assuming alfalfa as a reference crop. Thus,

$$LE_{cold} = 1.05\ ETr\ \lambda \quad \textbf{(98)}$$

where *ETr* is the hourly (or shorter time interval) alfalfa reference *ET* values, calculated by using the standardized ASCE Penman-Monteith equation. So, Equation 97b turns into the following:

$$H_{cold} = (R_n - G)_{cold} - (1.05 ET_r \lambda) \quad \textbf{(99)}$$

The reason of using a coefficient of 1.05 to estimate LE_{cold} lies on the fact that the cold pixel typically has an *ET* rate of 5% larger than that for the reference *ET(ETr)* (WRIGHT, 1982, p. 74) due to wet soil surface beneath a full vegetation canopy that will tend to increase the total *ET* rate (ALLEN *et al.*, 2007, p. 406). Conversely, the dry conditions that characterize

the hot pixel entail no evapotranspiration in such site, as there is no significant water supply on the bare soil. This condition assumes *LE* term to be 0 and, thus, Equation 97a turns into the following:

$$H_{hot} = (R_n - G)_{hot} \quad \textbf{(100)}$$

Rn_{cold}, G_{cold} and rah_{cold} are collected straight away from the cold pixel as well as Rn_{hot}, G_{hot} and rah_{hot} are collected from the hot pixel. At this point, we are in condition of computing both dT_{cold} and dT_{hot} as follows:

$$dT_{cold} = \frac{H_{cold} r_{ah_{cold}}}{\rho_{air_{cold}} C_p} \quad \textbf{(101a)}$$

$$dT_{hot} = \frac{H_{hot} r_{ah_{hot}}}{\rho_{air_{hot}} C_p} \quad \textbf{(101b)}$$

And *a* and *b* are then computed as it follows:

$$a = \frac{dT_{hot} - dT_{cold}}{Ts_{hot} - Ts_{cold}} \quad \textbf{(102)}$$

$$b = \frac{dT_{hot} - a}{Ts_{hot}} \quad \textbf{(103)}$$

So as to eventually calculate *dT* as it was defined in Equation 96.

Stability Correction Functions

Equations 99 and 100 actually represent a first approximation where neutrality conditions are assumed. The scope is to have a first set of *H* values from where to begin an iterative process that take us to subsequent calculations where atmospheric stability conditions are, effectively, considered.

The *Monin–Obukhov length* (*L*) defines the stability conditions of the atmosphere in the iterative process. *L* is the height at which forces of buoyancy (or stability) and mechanical mixing are equal, and is calculated as a function of heat and momentum fluxes as it follows:

$$L = -\frac{\rho_{air} C_p u_*^3 T_s}{kgH} \quad \textbf{(104)}$$

where:

g = gravitational acceleration (9.807 m s^{-2});

k = Von Karman's constant (dimensionless) (0.41)

Cp = specific heat of air at constant pressure (1004 J kg^{-1} K^{-1});

u_* = friction velocity (m s^{-1}) computed trough Equation 93;

Ts = surface air temperature (K°), obtained from the product MOD11A2;

ρ_{air} = air density (kg m^{-3}), computed trough Equation 91;

H = sensible heat flux (W m^{-2}) ,computed trough Equation 90.

Depending on the sign of *L*, values of the integrated stability corrections for momentum (Ψ_m) and heat transport (Ψ_h) are computed using formulations by PAULSON (1970, p. 371) and Webb (1970, p. 842).

When $L<0$, the lower atmospheric boundary layer is unstable so the following corrections are required:

$$\Psi_{m(200m)} = 2\ln\left(\frac{1+x_{(200m)}}{2}\right) + \ln\left(\frac{x_{(200m)}^2}{2}\right) - 2arctan\left(x_{(200)}\right) + 0.5\pi \quad \textbf{(105)}$$

$$\Psi_{h(2m)} = 2\ln\left(\frac{1+x_{(2m)}^2}{2}\right) \quad \textbf{(106)}$$

$$\Psi_{h(0.1m)} = 2\ln\left(\frac{1+x_{(0.1m)}^2}{2}\right) \quad \textbf{(107)}$$

where:

$$x_{(200m)} = \left(1 - 16\frac{200}{L}\right)^{0.25} \quad \textbf{(108)}$$

$$x_{(2m)} = \left(1 - 16\frac{2}{L}\right)^{0.25} \quad \textbf{(109)}$$

$$x_{(0.1m)} = \left(1 - 16\frac{0.1}{L}\right)^{0.25} \quad \textbf{(110)}$$

When L>0, the boundary layer is stable and the following correction are needed:

$$\Psi_{m_{(200m)}} = -5\left(\frac{2}{L}\right) \quad \textbf{(111)}$$

$$\Psi_{h_{(2m)}} = -5\left(\frac{2}{L}\right) \quad \textbf{(112)}$$

$$\Psi_{h_{(0.1m)}} = -5\left(\frac{0.1}{L}\right) \quad \textbf{(113)}$$

When L>0, values for $x_{(200\ m)}$, $x_{(2\ m)}$ and $x_{(0.1\ m)}$ have no meaning and their values are set to 1. These atmospheric stability corrections would affect also u_* and r_{ah}, whose new corrected values are defined as it follows:

$$u_* = \frac{ku_{200}}{\ln\left({}^{200}/_{z_{0m}}\right) - \Psi_{m_{(200m)}}} \quad \textbf{(114)}$$

$$r_{ah} = \frac{\ln\left({}^{z_2}/_{z_1}\right) - \Psi_{h(z_2)} + \Psi_{h(z_1)}}{ku_*} \quad \textbf{(115)}$$

where:

z_2 = commonly 2 m;

z_1 = commonly 0.1 m;

k = Von Karman's constant (dimensionless) (0.41);

u_* = friction velocity (m), obtained through Equation 93;

$\Psi m_{(200m)}$ = stability correction for momentum transport at 200 m. high, obtained through Equations 105 or 11 depending on the L value (dimensionless).

u_{200} = wind speed at a blending height (commonly 200 m), obtained through Equation 94.

z_{0m} = momentum roughness length (m), obtained through Equation 95.

$\Psi h(z_2)$ = stability correction for heat transport at z_2 height (dimensionless), obtained through Equation 106 or 112 depending on the L value.

$\Psi h(z_1)$ = stability correction for heat transport at z_1 height (dimensionless), obtained through Equation 107 or 113 depending on the *L* value.

Now that the corresponding corrections have been applied both on u_* and *rah*, with their new corrected values we will be able to compute a new Monin-Obukhov length (*L)*. *L* values are again evaluated with respect to their atmospheric conditions and another iteration is run if such conditions are not met. The iterative process keeps running until *rah* value stabilizes.

One utility of using energy balance approach is that actual/instantaneous evapotranspiration (*Eta*) rather than potential evapotranspiration (*ETp*) is computed so that reductions in evapotranspiration caused by a shortage of soil moisture are captured.

Taking into account that *Rn* represents the net radiation available for the system and that part of this energy serves to warm up the soil (*G*) and the atmosphere (*H*), *LE* can be thought as the amount of energy consumed by *ETa*. Subsequently, *LE* is a concept that includes the evapotranpirative mechanism and the required energy to run it and it is used to estimate *ETa* for each pixel of the satellite image as in the following:

$$ET_{inst} = 3600 \frac{LE}{\lambda} \quad \textbf{(116)}$$

where:

ET_{inst} = actual/instantaneous *ET*, called *ETa* throughout this study (mm hour^{-1}).

3600 = a factor for converting from seconds to hours;

LE = latent heat flux, (W m^{-2});

λ = latent heat of vaporization, (J kg^{-1});

$$\lambda = [2.501 - 0.00236(T_s - 273.15)] * 10^6 \quad \textbf{(117)}$$

ETa is now extrapolated to a daily time scale by calculating *ETr* fraction (*ETrF*), which is equivalent to the crop coefficient (*Kc*) for the cold pixel (ALLEN *et al.*, 2007, p. 406).

$$ET_rF = \frac{ET_{inst}}{ET_r} \quad \textbf{(118)}$$

where ET_r = reference *ET* at the time of the satellite image, (mm h^{-1});

It is assumed that *ETrF* is constant throughout the day, an assumption that (ALLEN *et al.* (2007a, p. 406); ALLEN *et al.* (2007b, p. 745)) corroborated by means of observational data. Eventually, *ETrF* is used to calculate daily *ETa (ET24):*

$$ET_{24} = ET_rF * ET_{r24} \quad \textbf{(119)}$$

where *ETr24* = *ETr* accumulated over 24 hours on the date of the satellite image, (mm).

5.2 Vegetation indices agreement on rainfall measurements

Several vegetation indices are obtained as secondary products as METRIC algorithm runs. This condition must not make us underestimate its importance and the meaningful insights we can obtain from them. At least, three of them are derived from METRIC application: Normalized Difference Vegetation Index (NDVI), LAI (Leaf Area Index) and SAVI (Soil-adjusted Vegetation Index).

NDVI has evolved as a primary tool for monitoring vegetation changes and the interpretation of the impact of short- to long-term climatic events on the biosphere (DAVENPORT; NICHOLSON, 1993, p. 389; GOWARD; PRINCE, 1995, p. 564; SRIVASTAVA *et al.*, 1997, p. 952; EKLUNDH, 1998, p .570).

The ability of remote sensing to monitor intra-annual and inter-annual spatial variability of vegetation provides a basis for spatio-temporal investigations of all types. Accurate in situ soil moisture information for large geographic areas is difficult to obtain; therefore, the relationship between precipitation and NDVI has been used to assess soil moisture (PETTORELLI *et al.*, 2005, p. 510). Translation of rainfall into soil moisture is a complex process, as is the translation of soil moisture into vegetation productivity. Consequently, the

importance of rainfall in controlling soil moisture varies both temporally and spatially, as a function of rainfall amount and distribution, as well as the ability of respective soils to store moisture (STEPHENS; LYONS, 1998, p. 224). Descriptive statistical relationships between rainfall, NDVI and soil moisture provide reliable means to identify drought severity (WYLIE *et al.*, 2003, p. 255). The prediction of the onset of drought conditions is an important tool for use in crop management within agriculture. The NVDI has also been used to improve our predictions and impact assessments of disturbances such as drought (SINGH *et al.*, 2003, p. 402) and flood (WANG *et al.*, 2003, p. 148). The use of the NVDI in the monitoring of drought relies on the sensitivity of the index to vegetation dryness.

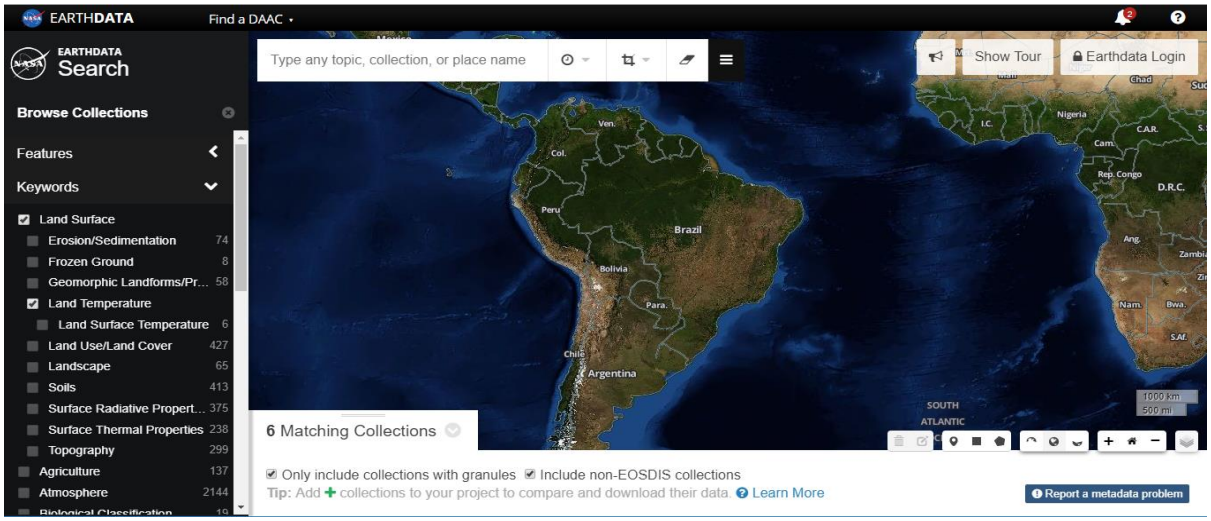
The relation between precipitation and vegetation varies between plants and during different growth stages (ROSENTHAL *et al.*, 1985, p. 90) the seasonality has to be taken into account when addressing the correlation between the two parameters (JI; PETERS 2003, p. 98). Vegetation furthermore responds to precipitation with a time lag, and precipitation should be accumulated over several months (FABRICANTE *et al.*, 2009, p. 753). Based on that strong and proved relation, we can predict short-period rainfall values or refill incomplete data registers.

6. WORKFLOW

This section illustrates all the steps made so as to reach the final results, from the imagery gathering to the *ET24* computing. The inputs needed are also detailed as well as the sites where they were taken from. The aim of this flowchart is to make the reader understand the whole process form the outset and, at the same time, permit other students/researchers to apply METRIC algorithm and how to get all the required inputs to run it, as a contribution to the scientific community.

6.1 Imagery extraction from EARTHDATA <https://search.earthdata.nasa.gov/>

Figure 26. Screenshot of EarthData website showing the user interface to allow any user to download data easily

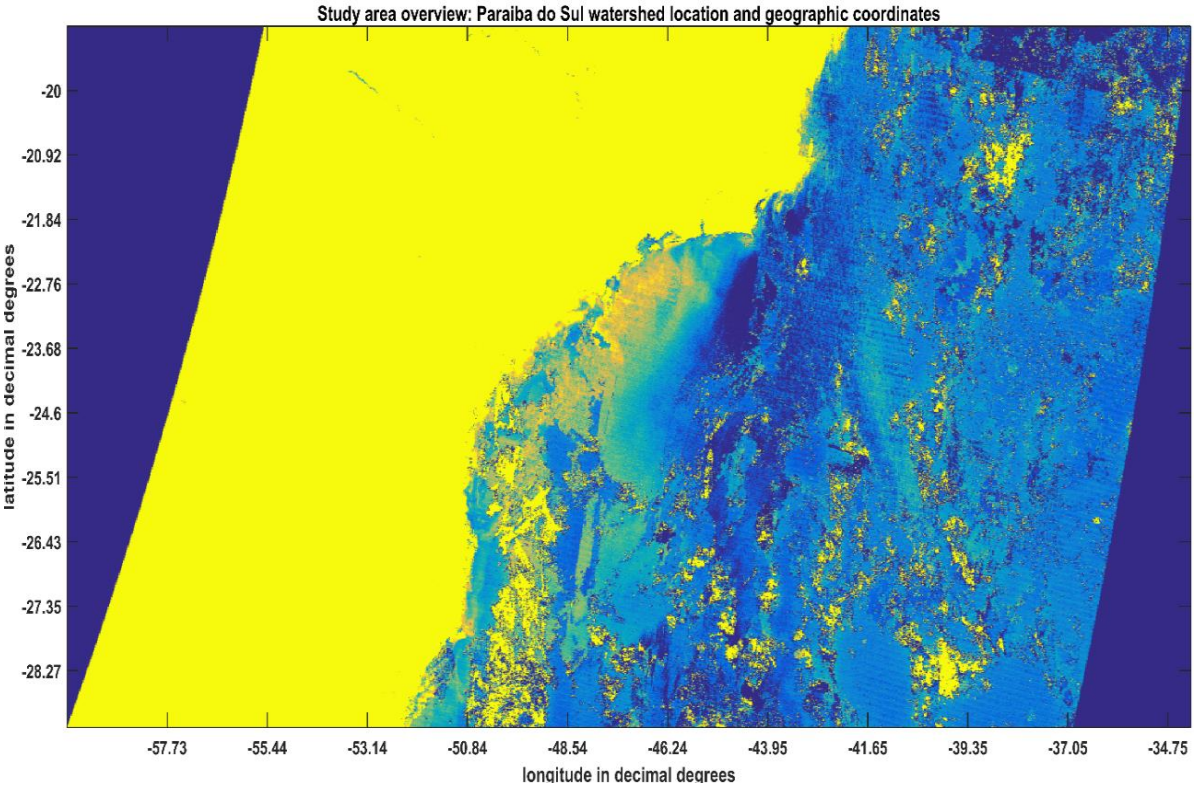


Source: <https://search.earthdata.nasa.gov/>

- from Keywords: select Land Surface / Surface Radiative Properties and Land Temperature; from platform: select Terra; from instruments: select MODIS
- from Temporal (clock symbol): enter the period of time of your interest; from Spatial (clipping symbol): select Grid coordinates, select MODIS sinusoidal as Grid and enter h and v values of your area of interest. If needed, visit <https://modis-land.gsfc.nasa.gov/MODLAND_grid.html>
- select the product of your interest among collections found and download data by choosing Direct download and click on View Download Links
- DownThemAll! <https://addons.mozilla.org/en-US/firefox/addon/downthemall/> can download them all data at once quickly. Imagery is eventually downloaded in format .hdf.

6.2 .hdf files conversion to .tiff using MODIS Reprojection Tool <https://lpdaac.usgs.gov/tools/modis_reprojection_tool>

Figure 27. .tiff file yielded by Modis Reprojection Tool after resampling two MODIS images and trandforming from sinusoidal to geographic reference system

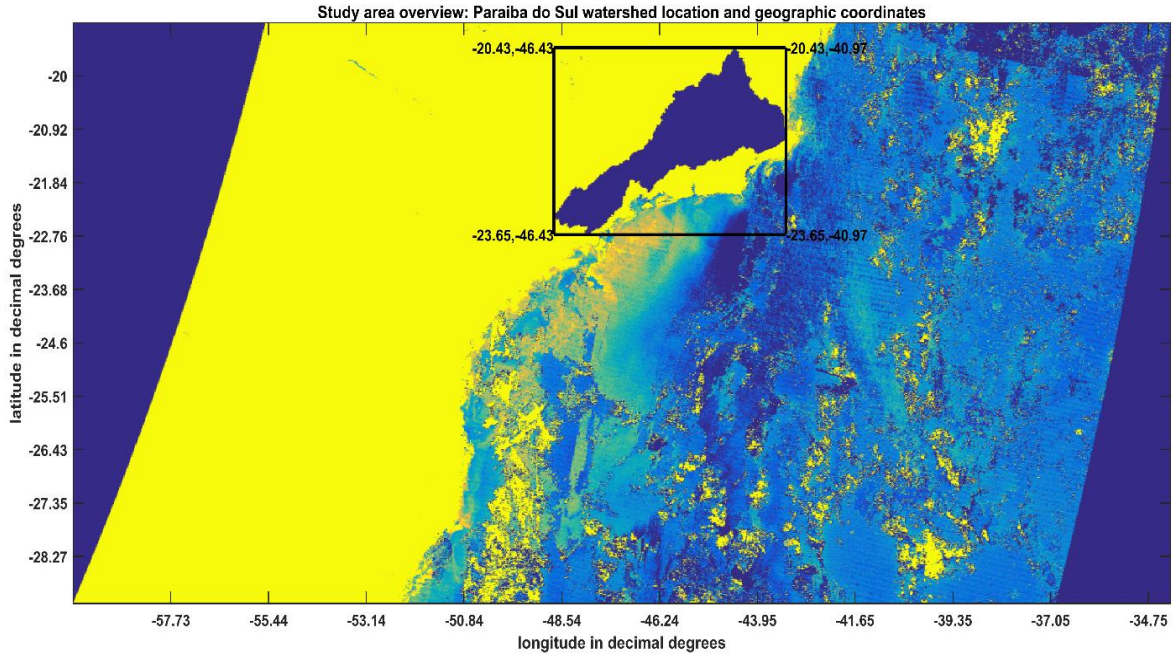


Source: https://search.earthdata.nasa.gov/

- download the suitable MODIS Reprojection Tool for your computer. <https://lpdaac.usgs.gov/tools/modis_reprojection_tool>
- download MRT user guide (recommended). Follow the recommended steps for installation
- This useful tool is essential if you work with MODIS imagery.
- It converts .hdf to .tiff files into the desire projection and datum. It allows to select only the required bands. It allows to resample into a different pixel size. It allows to piece several images together
- It contains a tool (MRTBatch.jar) that allow us to do all of the above in batch processing, very useful when working with a great volume of images.

6.3 Clipping our study area out the image

Figure 28. Overlapping of Paraíba do Sul watershed on the entire resampled area and detail about its specific location on decimal degrees



Source: https://search.earthdata.nasa.gov/

- Our study area is divided by two h tiles: h13 an h14, so we needed to piece both images together to get the entire study area, and this is what we see in this figure.

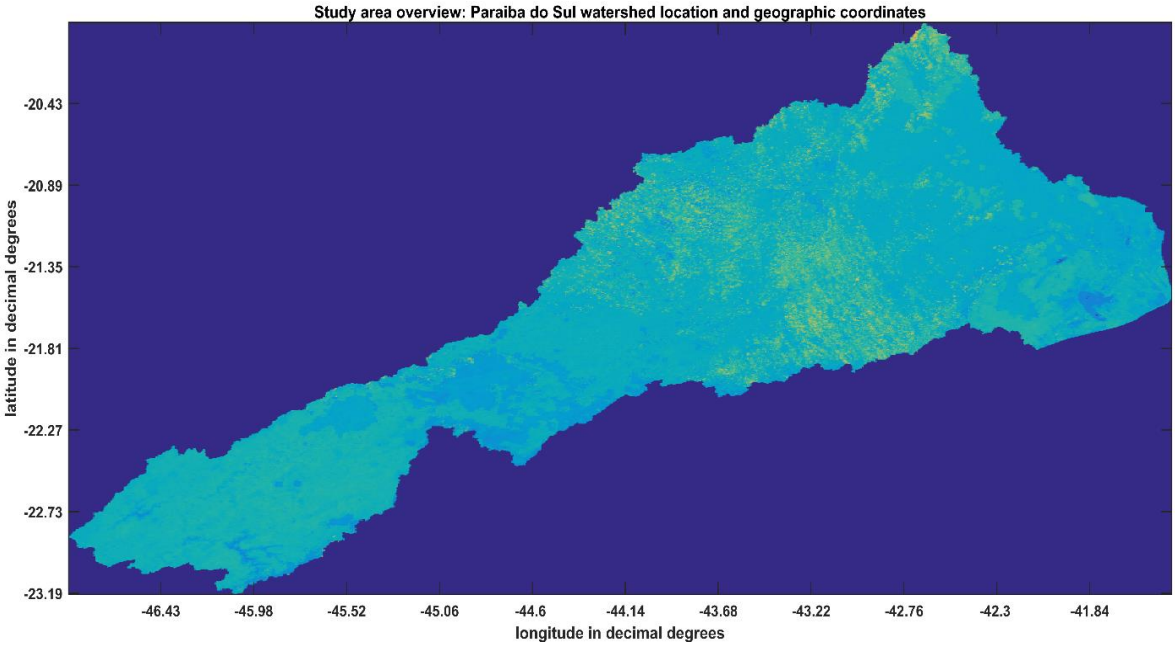
- It is advisable to arrange a shapefile archive of your area of interest. Our study area shapefile was obtained from <http://sigaceivap.org.br/siga-ceivap/map>.

- Bounding box information from both files is essential to precisely locate the shapefile into the image, as the area of study is usually smaller than a MODIS image. In the figure above, we can appreciate the specific location of our study area within the resample images.

- It is recommended to clip the actual area of interest and on work on what we are interested in. In the above figure, we can see where the cut will be done.

- Extract the area of interest and make a binary mask of it, where pixels with a value of 1 will allow any other values to be expressed and where pixels with a value of 0 will remain in 0 and compose the image background. The figure below is the result of clipping, the extension we will be working with all along the present research.

Figure 29. Result of clipping our study area out from Figure 3

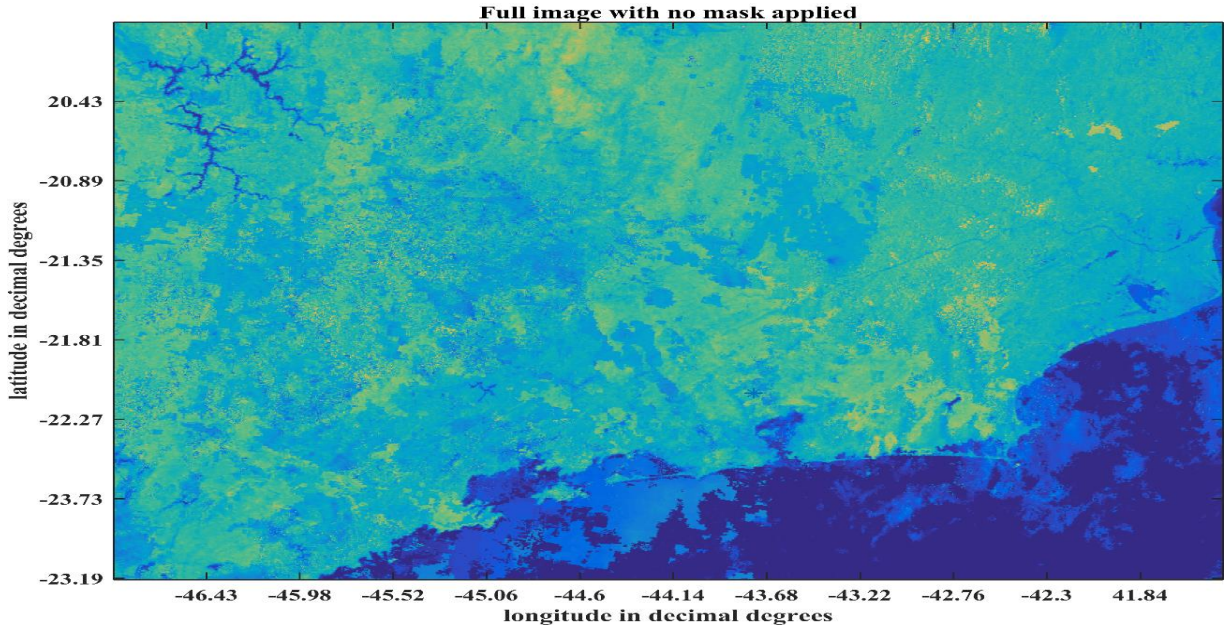


Source: <https://search.earthdata.nasa.gov/>

From this point on, it is advisable to work only with the pixels that take part in our study area as we can see in Figure 29. Here, pixels outside the study are converted to NaN (not a number value), a non-operating value, which means that any computation will bring no results.

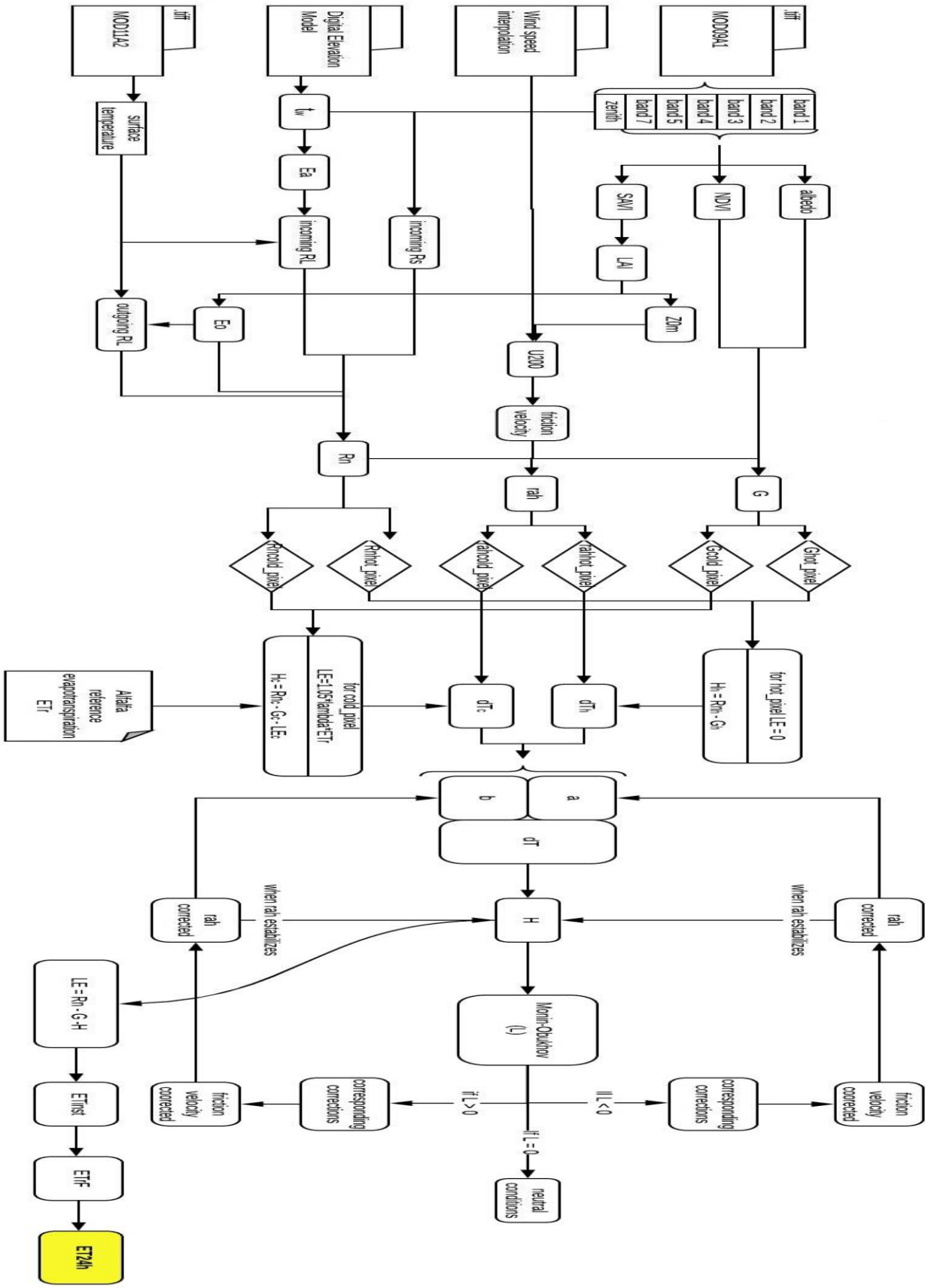
In our case, as mentioned back in chapter I, section 10 (Study area and data set), our INMET weather stations are located within and around Paraíba do Sul watershed, which means that we had some stations outside the contour drawn in Figure 29. However, pixels corresponding to the these stations' location are part of the study. In such case, in order to count on the largest number of stations, no mask was applied on any image and every computation composing METRIC algorithm was applied in full images as the one shown in Figure 30.

Figure 30. Same area as Figure 5 but with no mask overlapping so as to permit weather stations outside Paraíba do Sul watershed to be taken into account in further analyses



Source: <https://search.earthdata.nasa.gov/>

Figure 31. METRIC algorithm application



Source: workflow made by the author

7. DATASET SPECIFICATIONS

7.1 Regarding data availability

Before showing the results reached after applying some analysis on the images gathered for this study, it is important to draw reader's attention to the fact that METRIC algorithm is a methodology still quite dependable on ground measurements, represented by the data collected from the INMET weather station, in our case. Indeed, METRIC generates, at first, instantaneous evapotranspiration values (*ETinst* in equations 42 and 44), well-known as actual evapotranspiration (*ETa*). It means that METRIC algorithm brings evapotranspiration values on the satellite overpass time, and that is why these values are to be converted to *ET24* (24h-period accumulated evapotranspiration) through equation 45.

This essential requirement forced us to discard Clusters 1 and 2 conventional stations as they only produce daily measurements, making impossible apply METRIC method on them. So, in this section, our dataset is only composed by 13 INMET automated weather stations (Clusters 3 and 4), which do measure the required climatological parameters hourly.

7.2 Regarding study period

Another issue it is needed to be warned is related to the study period; it has been reduced too. Automated weather stations were installed by INMET only from the 2000's and not all at the same time. It means that among our dataset we have stations installed in 2000 and stations installed in 2010. This fact lead us to reduce the study period to a 9-year period ranging from 01/01/2007 to 31/12/2015, totalizing 3287 available measurements. Besides that, we have to take into account the missing values in each station data set.

7.3 Regarding analysis frequency

Main inputs needed for applying METRIC are remote sensing imagery as mentioned above. This imagery is produced by the different operating satellites with different frequencies. In the case of MODIS, it takes an image of the entire Earth twice a day, if we count on both Terra and Aqua platforms. As we are working with Terra platform imagery, we have only one image per day. It is a high temporal resolution if compared with other platforms imagery-production frequency. Albeit, it entails some "risks" represented by the cloud presence that can make impossible the analysis of part of the image.

To overcome this situation, MODIS platform offers a product, which is the combination of a 8-day period. For instances, images taken on julian days between 1 and 8 are combined into a composition of 8 days. It is important to warn the reader that this combination is not a product derived from average values of every day secene encompassed within the boundaries of the "combined period". It is actually a kind of puzzle composed by cloud-free regions/pixels of each image taken within those 8 days. For instance, first composition of every images collection (composition 001) consists of an unknown number of pixels belonging to day 1, day 2, day 3 and so on, until day 8.

Along this study, these compositions were used as inputs for METRIC algorithm aiming not to have problem relating to cloud-coverage.

7.4 Regarding *wind speed* and *ETr* interpolations

Using image compositions entailed another issue the reader must know and understand. The fact that these compositions encompass 8 days measurements entails that one year of images is not anymore composed by 365 images, but only by 46 images. This event turns an important point to be taken into account as wind speed measurements, along with the *ETr* were, respectively, measured and computed from daily measurements. It means that, in both cases, 365 values are available. As mentioned previously, there is no way of knowing which pixel in

the composition comes from which day image, so we opted for computing the mean both of wind speed and *ETr* values so as to reduce the original dataset from 365 to 46.

7.5 Regarding rainfall dataset

In vegetation indices evolution section, we discussed the role some vegetation indices, such as NDVI, LAI and SAVI can play in monitoring primary production or simply as vegetative condition indicators by contrasting them against rainfall measurements also made by the INMET. To do so, we needed a ground-measured rainfall dataset. It was obtained, as the rest of ground measurements via INMET weather station.

Again, we needed to take into account that one year of MODIS 8-day compositions yields 46 image compositions. In order to contrast vegetation indices (46 values of each index per year), with daily rainfall values (365 values per year), the latter had to be "cut down" by computing 8-day periods mean. Later on, we decided to reduce them even more and remain with only 12 values, one per month, not to contaminate the plot and be able to reflect any similar dynamics.

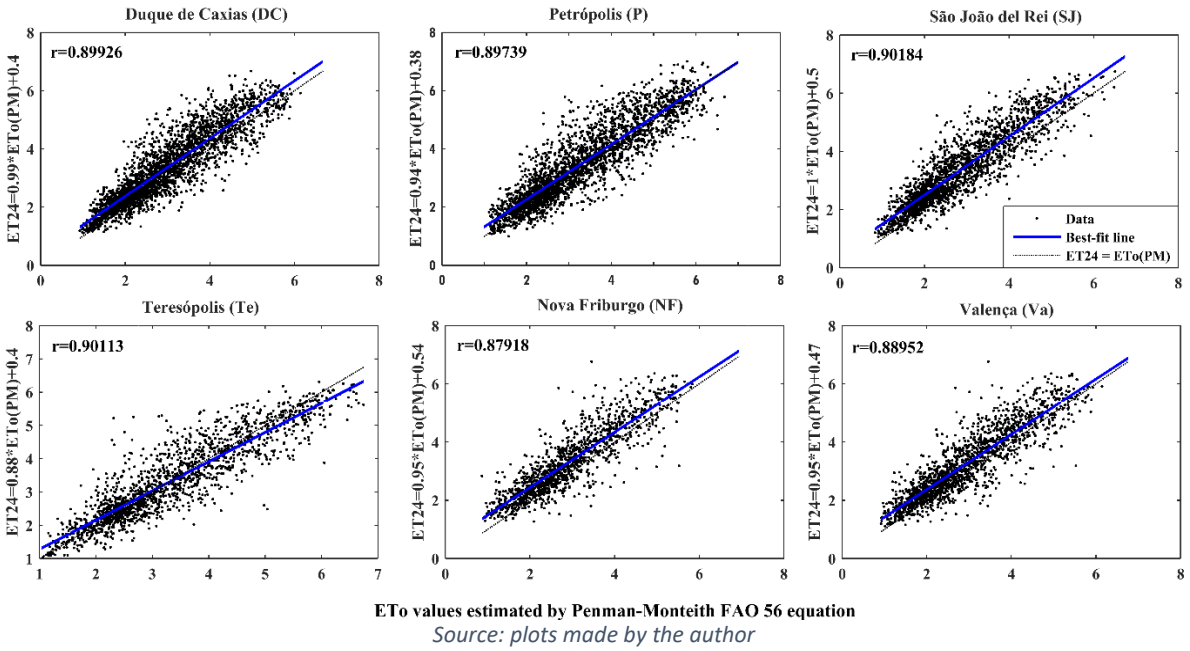
8 RESULTS

8.1 Correlations analysis between *ET24(METRIC)* and *ETo(PM)*

In this section, we wanted mainly to collate *ETo* values estimated by Penman-Montetih FAO 56 (*ETo(PM)*) equation (as the reference method) and *ET24* values yielded by METRIC algorithm (*ET24(METRIC)*). As the image is composed by many pixels and encompassed by a huge area where our INMET stations are located, the recommended way of comparing both sets of *ETo* values is to gather the geographical coordinates of the stations, look for them on the *ET24* raster map, get the values compressed in the corresponding pixels for the study period and plot them against the ETo values obtained via Penman-Monteith equation.

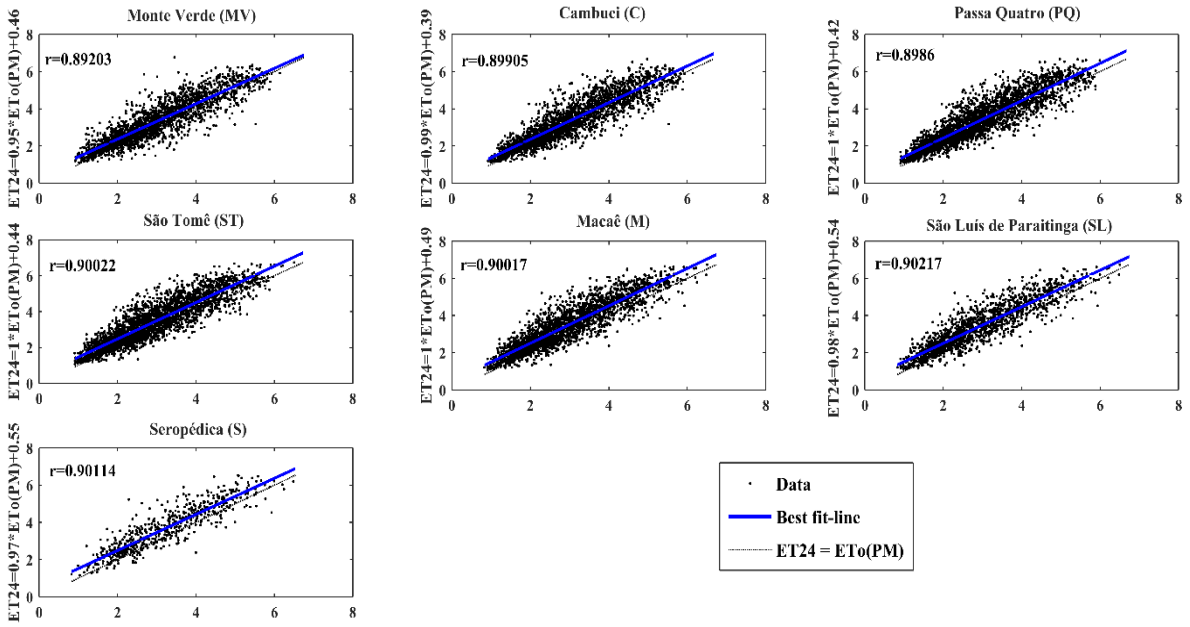
Further objectives of this fourth chapter is to check similarity between them and determine to what extent METRIC method (as a representative of other Energy Balance methods) and remote sensing imagery can represent an actual alternative to conventional methodologies or measuring devices. To do so, as a primary analysis we compared both sets of values by applying a correlation analysis. The results brought by this analysis can be checked in Figures 32 and 33:

Figure 32. Correlation analysis between ETo(PM) as independent variable and ET24(METRIC) as dependant variable conducted on Cluster 3 stations. High r values around 0.90 demonstrate the grate similarity between both estimates



Source: plots made by the author

Figure 33.Correlation analysis between ETo(PM) as independent variable and ET24(METRIC) as dependant variable conducted on Cluster 4 stations. High r values around 0.90 demonstrate the grate similarity between both estimates and the potential of METRIC to



Source: plots made by the author

As we can see in Figure 32 and Figure 33, correlation coefficient (*r*) reached values around 0.90, which speaks loud about the great similarity between *ETo(PM*) and *ET24(METRIC)* both for Cluster 3 and 4. It is easy to notice some plots having a lower point density; the reason lies on the fact that not all stations present the same amount of available measures. Although this procedure has been applied to values within the same period of study (01/01/2007 – 31/12/2015), the number of missing values in each station is variable. Despite this disagreement among stations, *r* values keep over 0.90, pointing out the reliability of METRIC and its independency of the amount of available data.

On the other hand, it is also true that some *ET24* products had to be discard for not reaching the expected results or because after a reasonable number of iterations, r_{ah} dis not remain steady. Such event shows still not enough robustness of the model and may need deeper calibration procedures (IDSO *et al.*, 1975, p. 349; MU *et al.*, 2007, p. 536; SENAY *et al.*, 2013, p. 591).

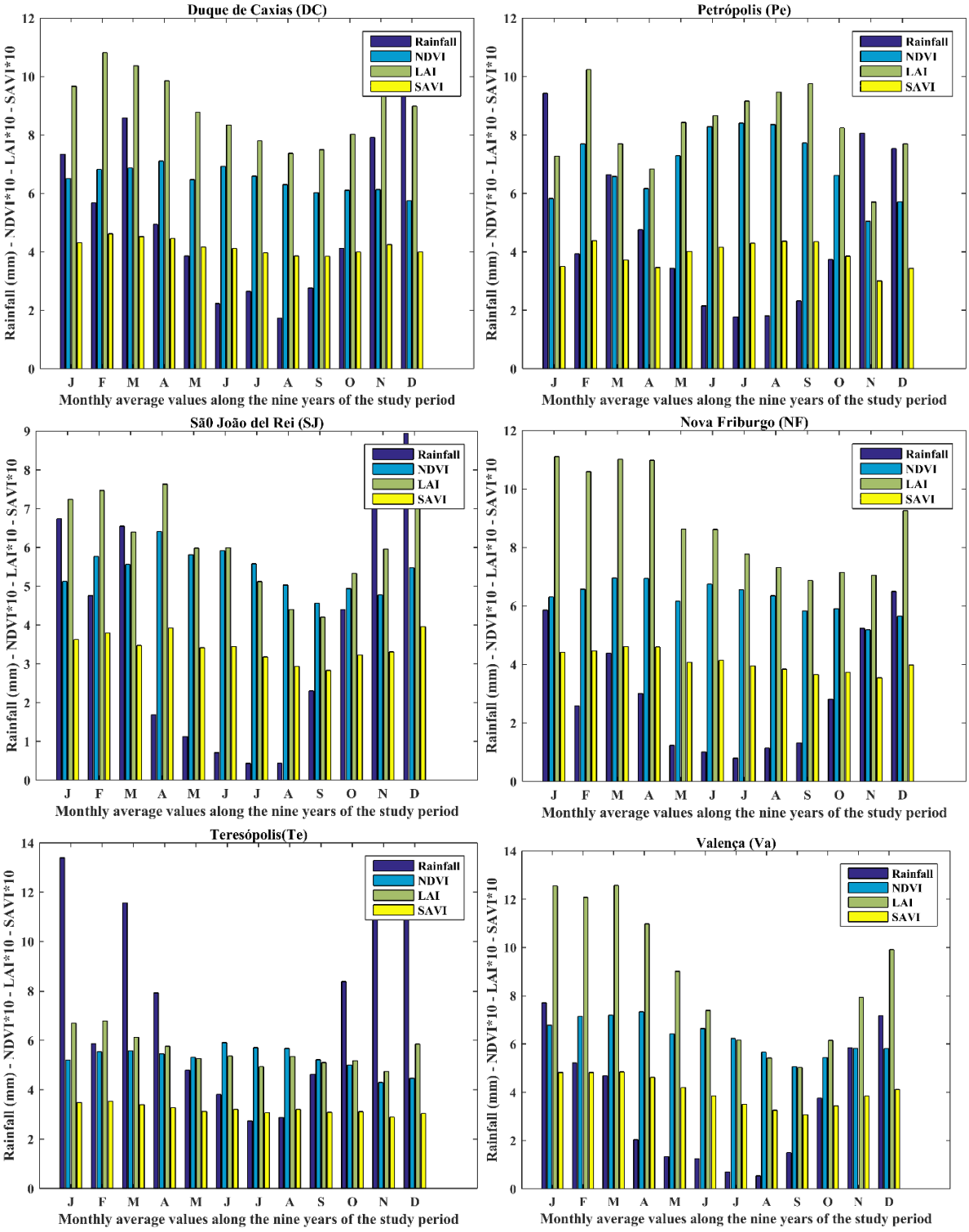
8.2 Relation between vegetation indices and rainfall

In Figure 34 and Figure 35, we plotted rainfall against the three vegetation indices mentioned above so as to determine the nature of those existing relations. For plotting and visibility purposes, all vegetation indices were multiplied by 10 so as to reach similar values to rainfall and be easier to read, interpretate and, overall, to make little changes more noticeable. As we can appreciate in both Figures 34 and 35, all three vegetation indices progressed in harmony among them and with rainfall. Specific stations, such as Pe, NF and Ma showed a particular behaviour, where NDVI and LAI rose whilst rainfall went down.

Another point that deserves to be highlighted regards to hard rain periods. During hard rain episodes, NDVI, LAI and SAVI responses are smoother though. Probably, the well-known delay between rainfall and vegetative production is, in part, responsible for this behaviour. SAVI and LAI are the indices where this delay is more easily detected, whilst NDVI seems to have a rapid response to water availability.

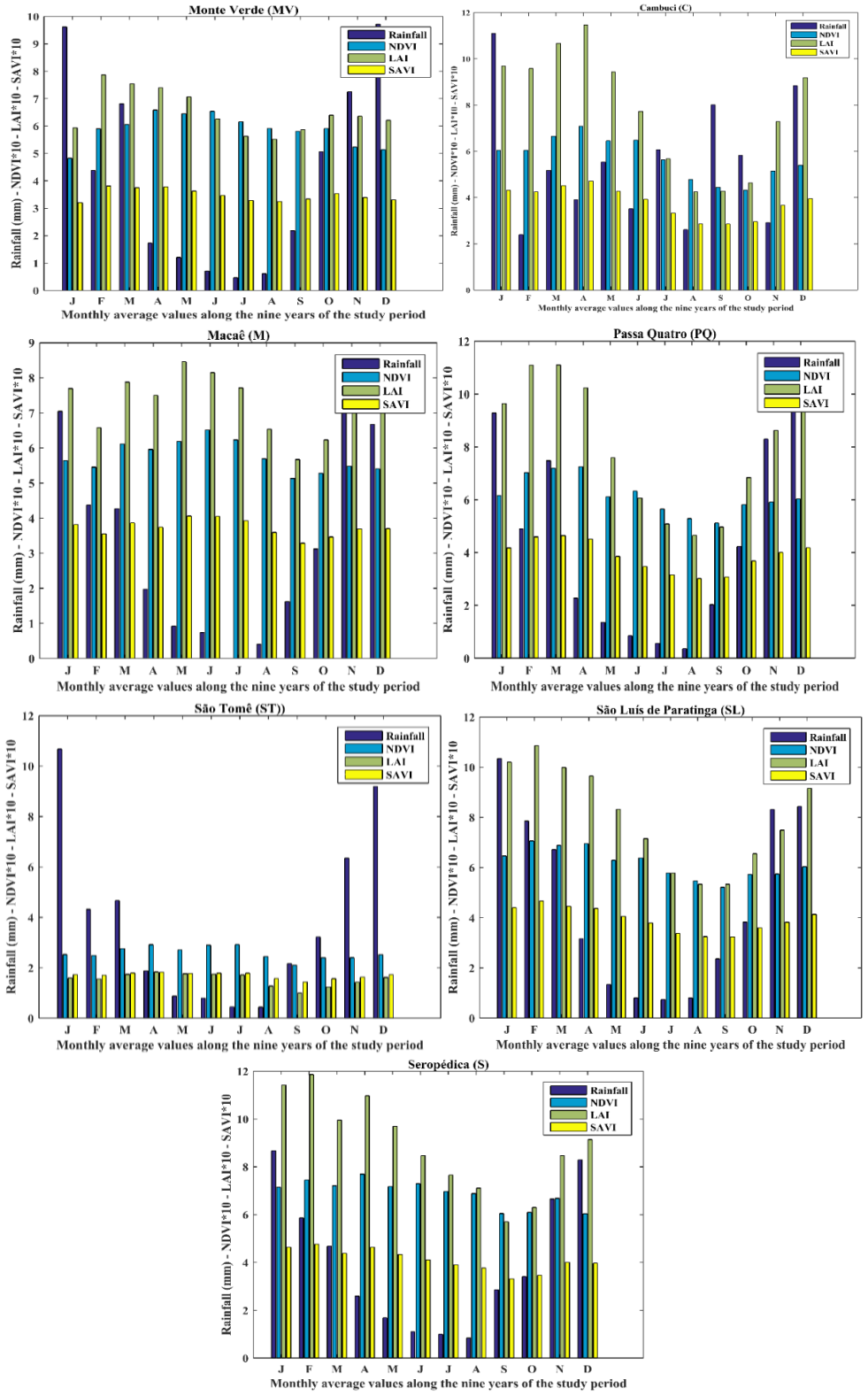
It is thought that the other part of responsibility lies on the soil and its natural capacity to store and deploy water. It is in this point where we find the connection between vegetation indices and evapotranspiration. In other words, high NDVI, LAI or SAVI values are, somehow, a reflect of a well-watered condition, an essential requirement to make evapotranspiration work.

Figure 34 . Contrasting Rainfall mesured values from Cluster 3 weather stations against NDVI, LAI and SAVI yielded during METRIC application procedure as vegetative and primary production indicators.



Source: plots made by the author

Figure 35. .Contrasting Rainfall mesured values from Cluster 4 weather stations against NDVI, LAI and SAVI yielded during METRIC application procedure as vegetative and primary production indicators.



Source: plots made by the author

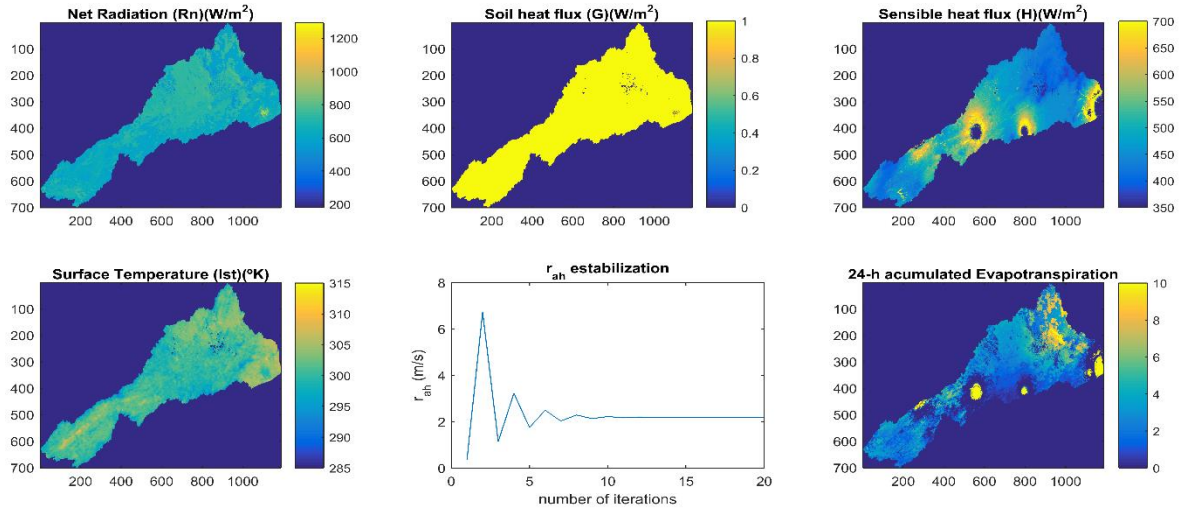
8.3 Exploratory analysis of some images

Along this last chapter we warned the reader about the reduction on our data set and on the study period due to reasons already exposed. In spite of it, the amount of images analysed has been pretty large, as we analysed 46 compositions of 8 days every year, for 9 years, which totalizes 414 images. For obvious reasons, we do not intend to show the results of all of them but a few examples where we could appreciate the range of expected values of some variables, such as Net radiation (Rn), Soil heat flux (G), Sensible heat flux (H), Surface temperature (lst), rah estabilization and, the final product, 24-hour accumulated evapotranspiration. Rn, G and H account for the amount of energy available and the losses of such energy, respectively. From them we are able to compute Latent heat flux (LE) through Equation 75, and eventually, *ETinst* as a residual product.

We were specially interested in showing the stabilization process of *rah* as it determines the number of iterations where it pursues the atmospheric stability condition. As we can see in Figures 36-39, *rah* estabilizes after 6-8 iterations, in agreement with ALLEN et al. (2013, p. 612).

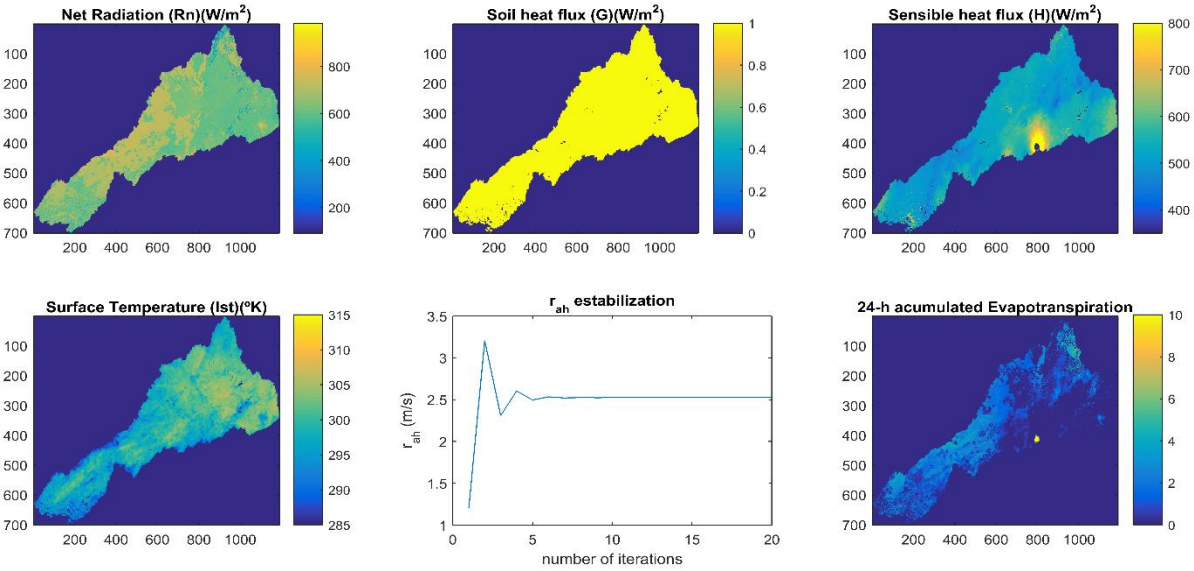
With respect to the main product of this last chapter (and somehow of the present work), *ET24* over Paraíba do Sul river watershed, it obtained a wide variety of values. For instance, it is recognisible the high values of *ET24* reached on Teresópolis-Petrópolis metropolitan area. Not much surprising if we take into account that it represents an important population spot and the probable heat island effect could have an impact. More surprising are the high values of *ET24* computed on Valença (Va), which is one of our automated stations.

Figure36. Composition 041(09/feb-16/feb) 2007. Illustration of some energetic features, such as Rn, G, H, lst, r_{ah} estabiliation and teh final product, 24-h accumulated eavapotranspiration, ET24



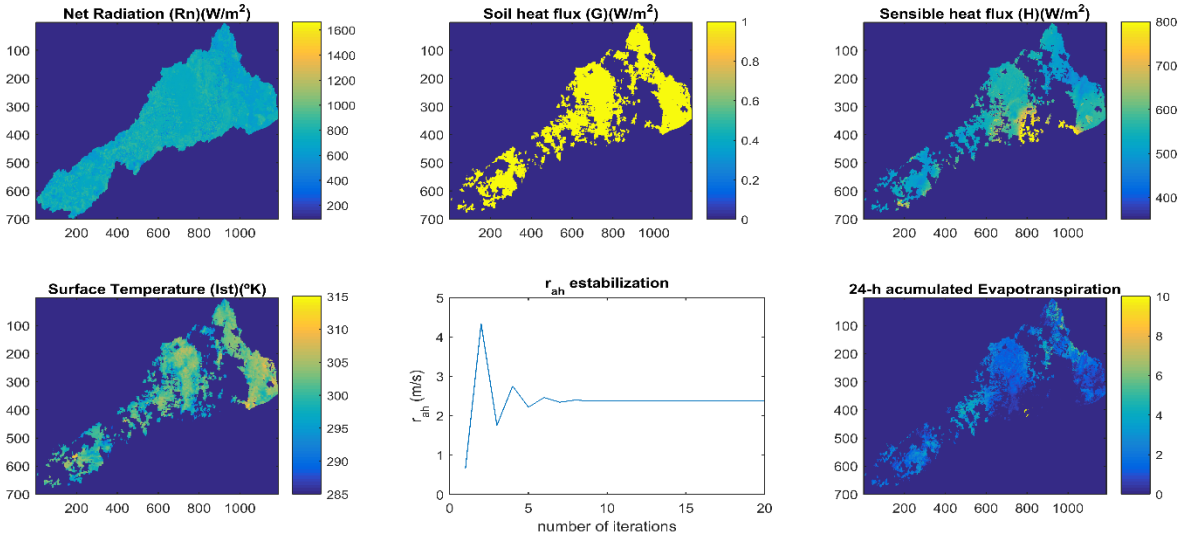
Source: plots made by the author

Figure 37. Composition 265(22/sep – 30/sep) 2008. Illustration of some energetic features, such as Rn, G, H, lst, r_{ah} estabiliation and teh final product, 24-h accumulated eavapotranspiration, ET24



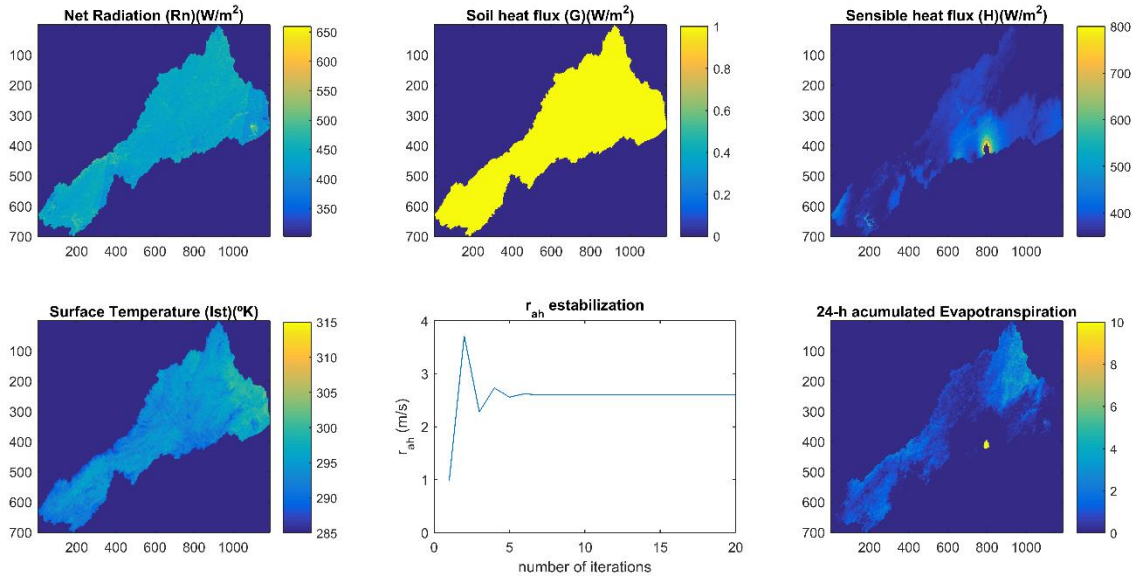
Source: plots made by the author

Figure 38. Composition 345 (11/12 – 19/12) 2009. Illustration of some energetic features, such as Rn, G, H, lst, r_{ah} estabiliation and teh final product, 24-h accumulated eavapotranspiration, ET24



Source: plots made by the author

Figure 39. Composition 161 (10/06 – 18/06)) 2010. Illustration of some energetic features, such as Rn, G, H, lst, r_{ah} estabiliation and teh final product, 24-h accumulated eavapotranspiration, ET24



Source: plots made by the author

9. CONCLUSION

The main analysis, made back in section 8.1, were a set of correlation analysis contrasting these to sources of ETo values. High r values, around 0.90 confirm our hypothesis and agree with results reached by other researchers in previous works, such as MORTON et al. (2013, p. 342) or REYES-GONZÁLEZ et al. (2017, p. 213). Thus, we can conclude that METRIC algorithm application on MODIS images over our study area generates daily evapotranspiration estimates highly similar to those values estimated by the reference method (*PM56*).

10. DISCUSSION

METRIC uses digital image data collected by Landsat and other remote-sensing satellites that record thermal infrared, visible, and near-infrared radiation from the earth's surface. ET is computed on a pixel-by-pixel basis for the instantaneous time of the satellite image. The process is based on a complete energy balance for each pixel, where ET is predicted from the residual amount of energy remaining from the classical energy balance, where ET = net radiation - heat to the soil - heat to the air.

Reasons why METRIC is attractive to western water resources management are:

- METRIC calculates actual *ET* rather than potential *ET* and does not require knowledge of crop type. (No satellite-based crop classification is needed);

- METRIC relies heavily on theoretical and physical relationships, but provides for the introduction and automated calibration of empirical coefficients and relationships to make the process operational and accurate;

- The use of reference *ETr* in calibration of METRIC and the use of *ETrF* in extrapolation to *ET24* provides general equivalency and congruency with *ET* as estimated using the traditional *Kc-ETr* approach, where *ETr* = alfalfa reference *ET* calculated using the ASCE-EWRI standardized PenmanMonteith equation (ASCE-EWRI 2005). This congruency is valuable for using *ET* maps generated by METRIC water rights management where water rights are based on previous *Kc-ETr* calculations;

- METRIC is autocalibrated for each image using ground-based calculations of *ETr* (made using weather data) where accuracy of the *ETr* estimate has been established by lysimetric and other studies in which we have high confidence.

ET maps created using METRIC, SEBAL, and similar remote sensing based processing systems may some day be routinely used as input to daily and monthly operational and planning models for reservoir operations, ground-water management, irrigation water supply planning, water rights regulation, and hydrologic studies.

The main objective of this fourth chapter was to confirm the capability of METRIC algorithm of producing reliable ETo values when compared to those yielded by the reference method (Penman-Monteith FAO 56 equation). The main analysis, made back in section 8.1, were a set of correlation analysis contrasting these to sources of ETo values. High r values, around 0.90 confirm our hypothesis and agree with results reached by other researchers in previous works, such as MORTON et al. (2013, p. 342) or REYES-GONZÁLEZ et al. (2017, p. 213).

Another reason for using an Energy Balance approach to implement a remotely sensed-based algorithm, such as METRIC, is to rapidly and operationally estimate evapotranspiration over large areas. The calibration of Energy Balance-based models is typically a time-consuming process, and as existing literature shows, there can be considerable variation in Evapotranspiration estimates from different users (MORTON et al., 2013, p. 342). Automated algorithms also allow for the ability to quantify the impact of model uncertainty on evapotranspiration.

Energy Balance-based models, such as METRIC still depends in a big way on ground climatological parameters measurements, such as temperatures, relative air humidity or wind speed. Its internal calibration is based on ETr, which most of the times is presented as an advantage, but makes the model even more dependable on measured inputs, which was one of the reason why these remote sensing-based algorithms were developed.

Continued development of this or similar automated algorithms may make it possible to provide operational evapotranspiration estimates that can be produced quickly and at minimal cost. Further analysis needs to be conducted to test how well this automated algorithm will perform in other study areas with different distributions of crop types and growing cycles.

METRIC model (also SEBAL) represents a maturing technology for deriving a satellite-driven surface energy balance for estimating ET from the earth's surface and have the potential to become widely adopted by water resources communities. METRIC is designed to produce high quality and accurate maps of ET for focused regions smaller than a few hundred kilometres in scale and at high resolution. Thus, use of reference ET for calibration and for extrapolating from instantaneous to ET24 helps to account for regional advection effects on ET and improves congruency with the traditional reference ETKc approach.

Requirements for applying METRIC are trained experts having good background in energy balance and radiation physics and adequate knowledge of vegetation characteristics as well as high quality hourly (or shorter) weather data. The requirement for trained experts having good physics background and high quality weather data is a distinct disadvantage of METRIC, and in some regard of SEBAL, in that highly specialized personnel are required to operate the model with credibility. It is evident that METRIC is highly reliant on the operator's ability to select appropriate cold and hot pixels. Current development work has tested the use of a statistical procedure to automate the selection of the hot and cold pixels with relatively good success (BURNETT et al., 2008, p. 79, KJAERSGAARD et al., 2009, p. 561).

However, METRIC is currently not a complete hands-off model. Continued comparisons with other remote sensing-based approaches are needed to understand the advantages/disadvantages of the different techniques and to gain a better understanding of the uncertainties in various model components and their impacts on final ET estimates.

METRIC is a hybrid model that combines remotely sensed energy balance via satellite data and ground-based evapotranspiration via in situ meteorological measurements in order to determine evapotranspiration. If calibrated properly, the model could be a viable tool to estimate water use in managed ecosystems in subhumid climates on a large scale, and particularly to assess short- and long-term water management, planning, and allocations.

However, there are a few constraints with application of the model to create monthly and seasonal ET maps. Currently, a number of efforts are being made at the University of Nebraska-Lincoln to use METRIC in agricultural water management, which is also the dream the author of the present work pursues: a real-world application for farmers. Some of those efforts include: (1) calibrating model algorithms against measurements over different vegetation, climate, and water regimes; (2) testing submodels to estimate H, G, LAI, and other variables under various land surfaces and developing improved algorithms (or localized calibration of the model) and if needed; (3) validating the model in the advective conditions; (4) developing a GIS-based soil water model to account for background evaporation; (5) automating hot and cold pixel selection; and (6) comparing pixel-by pixel values with other remote sensing-based models.

FINAL CONCLUSION

As it was the objective, throughout the present study we have applied different approaches to estimate evapotranspiration on the Paraíba do Sul river watershed; beginning from the traditional ones, which represent the past, characterized by inefficient and high-costly procedures, but widely utilized. On the other hand, we also applied cutting-edge procedures, based on high technology and ultimate scientific approaches, as artificial neural networks and remote sensing. After different analyses and interpretation of the results, we can conclude that both procedures are capable of estimating evapotranspiration at a high resemblance level, due to the results obtained and showed along the present work.

Even though, there are some issues we must discuss. Along chapter II, we built, trained, validated and tested artificial neural networks aiming to estimate evapotranspiration. The results were highly satisfactory, providing that any dataset was considered individually and the target was to refill incomplete historical registers. It is in this point where we must broaden the options available and find solutions to day-by-day issues which aids decision-making bodies or entities.

In chapter II, it was already mentioned the option to cluster weather stations whose climatological conditions are similar. Assuming similar conditions, validation stage may be reduced to a minimum, but can rise as essential when we change region. In any case and after the all the work carry out along the present research, something is clear: the power of the neural networks so as to produce reliable values of evapotranspiration from previous ones. From that point on, artificial neural networks disclose as tools able to forecast at short and half term; a very helpful approach to irrigation schedule, water rights license, hydrologic infrastructure planning or aquifer depletion management. Apart from that, artificial neural networks do not need a great amount of infrastructure or logistic, but a well-trained team with robust mathematical and physic background.

In the case of remote sensing-based approaches and models, like METRIC or SEBAL as the most widely used, most cases users do not hesitate to apply them assuming climatological conditions (from its origin) completely different from the conditions where the approach is being applied by the user. It is the case of METRIC, developed by Richard Allen under the climatological conditions of Idaho. Obviously, Idaho climatological conditions do not have

anything to do with Southeast of Brazil weather. As a scientist, we must doubt at all times and hesitate whether it is appropriate to apply algorithms with no constraints. Of course, it is much easier to apply something that has already done and tested by many others, but it is not the reproductibility what really matters, but the adaptation of such methodology or approach. Only by doing this way, we may get more precise, get more robust results and be more secure when making descisions based on such results, and, of course, assumptions.

In any case, the perspectives for both approaches are inspiring, that is why we focused on them along the present work. As a high-tech approach and part of a rising study field, known as “precision agriculture”, remote sensing is currently considered as a high-potential tool due to the fierce spatial race involving the greatest economies in the world, such as USA, Europe, China or Russia. This competition will surely bring some short-half term benefits, such as better spatiotemporal resolution products, an easier access to earth observation imagery, sensors able to detect more and more parameters, development of new and more capable algorithms and so on. The future perspectives are massive so as to amplify the fields of study, the analysing tools, the information sources and the diversity of results. But it is essential to keep in mind that any study, analysis, result, interpretation, conclusion or a whole research must focus on its applicability in the real world, bringing to it the benefits that are supposed to it.

To get to an end, and as a personal thought, there is no point in study any issue deeply if we cannot get any benefit from that, as the same manner that farmers will not save any water if we do not give them the right elements.

REFERENCES

ABOUKHALED, A.; ALFARO, A.; SMITH, M. Lysimeters, FAO Irrigation and Drainage Paper No. 39. **FAO, Rome, Italy**, p. 68, 1982.

ALLEN, R. G. Evapotranspiration for southwest Florida from satellite-based energy balance. **Rep. Prepared for Tampa Bay Water**, 2002.

ALLEN, R. G. A Penman for all seasons. **Journal of Irrigation and Drainage Engineering,** v. 112, n. 4, p. 348-368, 1986. ISSN 0733-9437.

ALLEN, R. G. Assessing integrity of weather data for reference evapotranspiration estimation. **Journal of Irrigation and Drainage Engineering,** v. 122, n. 2, p. 97-106, 1996. ISSN 0733-9437.

ALLEN, R. G.. Using the FAO-56 dual crop coefficient method over an irrigated region as part of an evapotranspiration intercomparison study. **Journal of hydrology,** v. 229, n. 1, p. 27-41, 2000. ISSN 0022-1694.

ALLEN, R. G. et al. Automated calibration of the metric-landsat evapotranspiration process. **JAWRA Journal of the American Water Resources Association,** v. 49, n. 3, p. 563-576, 2013. ISSN 1752-1688.

ALLEN, R. G. et al. Lysimeters for evapotranspiration and environmental measurements. 1991, ASCE.

ALLEN, R. G. et al. Operational estimates of reference evapotranspiration. **Agronomy journal,** v. 81, n. 4, p. 650-662, 1989. ISSN 0002-1962.

ALLEN, R. G. et al. Evapotranspiration information reporting: I. Factors governing measurement accuracy. **Agricultural Water Management,** v. 98, n. 6, p. 899-920, 2011. ISSN 0378-3774.

ALLEN, R. G. et al. Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. **FAO, Rome,** v. 300, n. 9, p. D05109, 1998.

ALLEN, R. G. et al. **Crop evapotranspiration : guidelines for computing crop water requirements**. Rome: Food and Agriculture Organization of the United Nations, 1998. xxvi, 300 p. ISBN 9251042195

0254-5284 ;.

ALLEN, R. G. et al. FAO Irrigation and drainage paper No. 56. **Rome: Food and Agriculture Organization of the United Nations,** v. 56, n. 97, p. e156, 1998.

ALLEN, R. G. et al. FAO-56 dual crop coefficient method for estimating evaporation from soil and application extensions. **Journal of Irrigation and Drainage Engineering,** v. 131, n. 1, p. 2-13, 2005. ISSN 0733-9437.

ALLEN, R. G. et al. An update for the calculation of reference evapotranspiration. **ICID bulletin,** v. 43, n. 2, p. 35-92, 1994.

ALLEN, R. G. et al. Satellite-based energy balance for mapping evapotranspiration with internalized calibration (METRIC)—Applications. **Journal of Irrigation and Drainage Engineering,** v. 133, n. 4, p. 395-406, 2007. ISSN 0733-9437.

ALLEN, R. G.; TASUMI, M.; TREZZA, R. Satellite-based energy balance for mapping evapotranspiration with internalized calibration (METRIC)—Model. **Journal of Irrigation and Drainage Engineering,** v. 133, n. 4, p. 380-394, 2007. ISSN 0733-9437.

ARCA, B.; BENISCASA, F.; VINCENZI, M. Evaluation of neural network techniques for estimating evapotranspiration. **Evolving Solution with Neural Networks, Italia: Baratti, R. y De Canete**, p. 62-97, 2001.

BAKHTIARI, B.; LIAGHAT, A. Seasonal sensitivity analysis for climatic variables of ASCE-Penman-Monteith model in a semi-arid climate. **Journal of Agricultural Science and Technology,** v. 13, p. 1135-1145, 2011. ISSN 1680-7073.

BASTIAANSSEN, W. SEBAL-based sensible and latent heat fluxes in the irrigated Gediz Basin, Turkey. **Journal of hydrology,** v. 229, n. 1, p. 87-100, 2000. ISSN 0022-1694.

BASTIAANSSEN, W. et al. SEBAL model with remotely sensed data to improve water-resources management under actual field conditions. **Journal of Irrigation and Drainage Engineering,** v. 131, n. 1, p. 85-93, 2005. ISSN 0733-9437.

BASTIAANSSEN, W. G. et al. A remote sensing surface energy balance algorithm for land (SEBAL). 1. Formulation. **Journal of hydrology,** v. 212, p. 198-212, 1998. ISSN 0022-1694.

BASTIAANSSEN, W. G. **Regionalization of surface flux densities and moisture indicators in composite terrain. A remote sensing approach under clear skies in Mediterranean climates**. SC-DLO, 1995a. ISBN 9054854650.

BEVEN, K. A sensitivity analysis of the Penman-Monteith actual evapotranspiration estimates. **Journal of hydrology,** v. 44, n. 3-4, p. 169-190, 1979. ISSN 0022-1694.

BHATTACHARYA, B.; LOBBRECHT, A.; SOLOMATINE, D. Neural Networks and Reinforcement Learning in Control of Water System. 2003.

BLAD, B. L.; ROSENBERG, N. J. Lysimetric calibration of the Bowen ratio-energy balance method for evapotranspiration estimation in the central Great Plains. **Journal of Applied Meteorology,** v. 13, n. 2, p. 227-236, 1974. ISSN 0021-8952.

BLANEY, H. F.; CRIDDLE, W. D. Determining water requirements in irrigated areas from climatological and irrigation data. **US Soil Conservation Service**, 1950.

BRUTSAERT, W. Energy Budget and Related Methods. In: (Ed.). **Evaporation into the Atmosphere**: Springer, 1982. p.209-230.

BRUTSAERT, W.; PARLANGE, M. Hydrologic cycle explains the evaporation paradox. **Nature,** v. 396, n. 6706, p. 30-30, 1998. ISSN 0028-0836.

BURN, D. H.; HESCH, N. M. Trends in evaporation for the Canadian Prairies. **Journal of hydrology,** v. 336, n. 1, p. 61-73, 2007. ISSN 0022-1694.

CAMARGO, A. D.; SENTELHAS, P. C. Avaliação do desempenho de diferentes métodos de estimativa da evapotranspiração potencial no Estado de São Paulo, Brasil. **Revista Brasileira de agrometeorologia,** v. 5, n. 1, p. 89-97, 1997.

CAMPOLO, M.; SOLDATI, A.; ANDREUSSI, P. Artificial neural network approach to flood forecasting in the River Arno. **Hydrological Sciences Journal,** v. 48, n. 3, p. 381-398, 2003. ISSN 0262-6667.

CAMPOLONGO, F.; BRADDOCK, R. The use of graph theory in the sensitivity analysis of the model output: a second order screening method. **Reliability Engineering & System Safety,** v. 64, n. 1, p. 1-12, 1999. ISSN 0951-8320.

CAMPOLONGO, F.; CARIBONI, J.; SALTELLI, A. An effective screening design for sensitivity analysis of large models. **Environmental Modelling & Software,** v. 22, n. 10, p. 1509-1518, 2007. ISSN 1364-8152.

CAPRIO, J. M. The solar thermal unit concept in problems related to plant development and potential evapotranspiration. In: (Ed.). **Phenology and seasonality modeling**: Springer, 1974. p.353-364.

CARR, J. et al. **National Water Summary 1987: Hydrologic events and water supply and use. Annual report**. Geological Survey, Reston, VA (USA). Water Resources Div. 1990

CHANG, J. H. **Climate and agriculture: an ecological survey**. Transaction Publishers, 1968. ISBN 0202364895.

CHATTOPADHYAY, N.; HULME, M. Evaporation and potential evapotranspiration in India under conditions of recent and future climate change. **Agricultural and Forest Meteorology,** v. 87, n. 1, p. 55-73, 1997. ISSN 0168-1923.

CHAUHAN, S.; SHRIVASTAVA, R. Performance evaluation of reference evapotranspiration estimation using climate based methods and artificial neural networks. **Water Resources Management,** v. 23, n. 5, p. 825-837, 2009. ISSN 0920-4741.

CHAUHAN, S.; SHRIVASTAVA, R. Estimating Reference Evapo-transpiration Using neural computing Technique. **Journal of Indian Water resources Society,** v. 32, p. 22-30, 2012.

CHEN, S.; LIU, Y.; AXEL, T. Climatic change on the Tibetan Plateau: potential evapotranspiration trends from 1961–2000. **Climatic Change,** v. 76, n. 3-4, p. 291-319, 2006. ISSN 0165-0009.

CHIEW, F. et al. Penman-Monteith, FAO-24 reference crop evapotranspiration and class-A pan data in Australia. **Agricultural water management,** v. 28, n. 1, p. 9-21, 1995. ISSN 0378-3774.

CHOW, V.; MAIDMENT, D.; MAYS, L. **Applied hydrology**. 1988. ISBN 0070108102.

CLEVERLY, J. R. et al. Seasonal estimates of actual evapo-transpiration from Tamarix ramosissima stands using three-dimensional eddy covariance. **Journal of Arid Environments,** v. 52, n. 2, p. 181-197, 2002. ISSN 0140-1963.

COHEN, S.; IANETZ, A.; STANHILL, G. Evaporative climate changes at Bet Dagan, Israel, 1964–1998. **Agricultural and Forest Meteorology,** v. 111, n. 2, p. 83-91, 2002. ISSN 0168-1923.

CONG, Z.; YANG, D. Does evaporation paradox exist in China. **Hydrol. Earth Syst. Sci. Discuss,** v. 5, n. 4, p. 2111-2131, 2008.

CYBENKO, G. Approximation by superpositions of a sigmoidal function. **Mathematics of control, signals and systems,** v. 2, n. 4, p. 303-314, 1989. ISSN 0932-4194.

DAI, X. et al. Artificial neural network models for estimating regional reference evapotranspiration based on climate factors. **Hydrological processes,** v. 23, n. 3, p. 442, 2009. ISSN 0885-6087.

DAVENPORT, M. L.; NICHOLSON, S. E. On the relation between rainfall and the Normalized Difference Vegetation Index for diverse vegetation types in East Africa. **International Journal of Remote Sensing,** v. 14, n. 12, p. 2369-2389, 1993. ISSN 0143-1161.

DIAZ-ESPEJO, A.; VERHOEF, A. The effect of varying soil moisture on the energy balance and co2 exchange of sunflowers: a study using a mini weighing lysimeter system. EGS-AGU-EUG Joint Assembly, 2003. p.9623.

DOORENBOS, J.; PRUITT, W. O. Guidelines for predicting crop water requirements. **FAO Irrigation and Darinage,** v. Paper No. 24, 1977.

DUFFIE, J. A.; BECKMAN, W. A. **Solar engineering of thermal processes**. John Wiley & Sons, 2013. ISBN 0470873663.

EKLUNDH, L. Estimating relations between AVHRR NDVI and rainfall in East Africa at 10-day and monthly time scales. **International Journal of Remote Sensing,** v. 19, n. 3, p. 563-570, 1998. ISSN 0143-1161.

ERVIN, E.; KOSKI, A. Trinexapac-ethyl effects on Kentucky bluegrass evapotranspiration. **Crop science,** v. 41, n. 1, p. 247-250, 2001. ISSN 1435-0653.

ESLAMIAN, S.; KHORDADI, M. J.; ABEDI-KOUPAI, J. Effects of variations in climatic parameters on evapotranspiration in the arid and semi-arid regions. **Global and Planetary Change,** v. 78, n. 3, p. 188-194, 2011. ISSN 0921-8181.

ESTÉVEZ, J.; GAVILÁN, P.; BERENGENA, J. Sensitivity analysis of a Penman–Monteith type equation to estimate reference evapotranspiration in southern Spain. **Hydrological Processes,** v. 23, n. 23, p. 3342-3353, 2009. ISSN 1099-1085.

FABRICANTE, I.; OESTERHELD, M.; PARUELO, J. Annual and seasonal variation of NDVI explained by current and previous precipitation across Northern Patagonia. **Journal of Arid Environments,** v. 73, n. 8, p. 745-753, 2009. ISSN 0140-1963.

FAUSSET, L. Fundamentals of neural networks. **Architecture, Algorithm and Application. Prentice Hall**, 1994.

FRANCOS, A. et al. Sensitivity analysis of distributed environmental simulation models: understanding the model behaviour in hydrological studies at the catchment scale. **Reliability Engineering and System Safety,** v. 79, p. 205–218, 2003.

FRITSCHEN, L. J. Accuracy of evapotranspiration determinations by the Bowen ratio method. **Hydrological Sciences Journal,** v. 10, n. 2, p. 38-48, 1965. ISSN 0020-6024.

GALVÃO, R. K. et al. Linear-wavelet networks. **International journal of applied mathematics and computer science,** v. 14, n. 2, p. 221-232, 2004. ISSN 1641-876X.

GAO, G. et al. Spatial and temporal variations and controlling factors of potential evapotranspiration in China: 1956–2000. **Journal of Geographical Sciences,** v. 16, n. 1, p. 3-12, 2006. ISSN 1009-637X.

GOLUBEV, V. S. et al. Evaporation changes over the contiguous United States and the former USSR: A reassessment. **Geophysical Research Letters,** v. 28, n. 13, p. 2665-2668, 2001. ISSN 0094-8276.

GONG, L. et al. Sensitivity of the Penman–Monteith reference evapotranspiration to key climatic variables in the Changjiang (Yangtze River) basin. **Journal of hydrology,** v. 329, n. 3, p. 620-629, 2006. ISSN 0022-1694.

GOVINDARAJU, R. S.; RAO, A. R. Artificial neural networks in hydrology: ii, hydrologic applications. 2000.

GOWARD, S. N.; PRINCE, S. D. Transient effects of climate on vegetation dynamics: satellite observations. **Journal of Biogeography**, p. 549-564, 1995. ISSN 0305-0270.

GOYAL, R. Sensitivity of evapotranspiration to global warming: a case study of arid zone of Rajasthan (India). **Agricultural Water Management,** v. 69, n. 1, p. 1-11, 2004. ISSN 0378-3774.

GREBET, P.; CUENCA, R. H. History of lysimeter design and effects of environmental disturbances. Lysimeters for Evapotranspiration and Environmental Measurements, 1991, ASCE. p.10-18.

HAGAN, M. T.; MENHAJ, M. B. Training feedforward networks with the Marquardt algorithm. **IEEE transactions on Neural Networks,** v. 5, n. 6, p. 989-993, 1994. ISSN 1045-9227.

HANSEN, S. Estimation of potential and actual evapotranspiration. **Hydrology Research,** v. 15, n. 4-5, p. 205-212, 1984. ISSN 0029-1277.

HARGREAVES, G. H.; ALLEN, R. G. History and evaluation of Hargreaves evapotranspiration equation. **Journal of Irrigation and Drainage Engineering,** v. 129, n. 1, p. 53-63, 2003. ISSN 0733-9437.

HARGREAVES, G. H.; SAMANI, Z. A. Reference crop evapotranspiration from temperature. **Applied engineering in agriculture,** v. 1, n. 2, p. 96-99, 1985.

HAYKIN, S. Neural Networks: A Comprehensive Foundation: Macmillan College Publishing Company. **New York**, 1994.

HOBBINS, M. T.; RAMÍREZ, J. A.; BROWN, T. C. Trends in pan evaporation and actual evapotranspiration across the conterminous US: Paradoxical or complementary? **Geophysical Research Letters,** v. 31, n. 13, 2004. ISSN 1944-8007.

HORNIK, K.; STINCHCOMBE, M.; WHITE, H. Multilayer feedforward networks are universal approximators. **Neural networks,** v. 2, n. 5, p. 359-366, 1989. ISSN 0893-6080.

HOWELL, T. et al. Irrigated fescue grass ET compared with calculated reference grass ET. National irrigation symposium. Proceedings of the 4th Decennial Symposium, Phoenix, Arizona, USA, November 14-16, 2000., 2000, American Society of Agricultural Engineers. p.228-242.

HOWELL, T. A.; SCHNEIDER, A. D.; JENSEN, M. E. History of lysimeter design and use for evapotranspiration measurements. Lysimeters for Evapotranspiration and Environmental Measurements:, 1991, ASCE. p.1-9.

HUMPHREYS, E. et al. Annual and seasonal variability of sensible and latent heat fluxes above a coastal Douglas-fir forest, British Columbia, Canada. **Agricultural and Forest Meteorology,** v. 115, n. 1, p. 109-125, 2003. ISSN 0168-1923.

IDSO, S. et al. The utility of surface temperature measurements for the remote sensing of surface soil water status. **Journal of geophysical research,** v. 80, n. 21, p. 3044-3049, 1975. ISSN 2156-2202.

IKUDAYISI, A.; ADEYEMO, J. Effects of Different Meteorological Variables on Reference Evapotranspiration Modeling: Application of Principal Component Analysis. **World Academy of Science, Engineering and Technology, International Journal of Environmental, Chemical, Ecological, Geological and Geophysical Engineering,** v. 10, n. 6, p. 632-636, 2016.

IMRIE, C.; DURUCAN, S.; KORRE, A. River flow prediction using artificial neural networks: generalisation beyond the calibration range. **Journal of hydrology,** v. 233, n. 1, p. 138-153, 2000. ISSN 0022-1694.

INMAN-BAMBER, N.; MCGLINCHEY, M. Crop coefficients and water-use estimates for sugarcane based on long-term Bowen ratio energy balance measurements. **Field Crops Research,** v. 83, n. 2, p. 125-138, 2003. ISSN 0378-4290.

IRMAK, S.; ALLEN, R.; WHITTY, E. Daily grass and alfalfa-reference evapotranspiration estimates and alfalfa-to-grass evapotranspiration ratios in Florida. **Journal of Irrigation and Drainage Engineering,** v. 129, n. 5, p. 360-370, 2003. ISSN 0733-9437.

IRMAK, S. et al. Solar and net radiation-based equations to estimate reference evapotranspiration in humid climates. **Journal of irrigation and drainage engineering,** v. 129, n. 5, p. 336-347, 2003. ISSN 0733-9437.

IRMAK, S. et al. Sensitivity analyses and sensitivity coefficients of standardized daily ASCE-Penman-Monteith equation. **Journal of Irrigation and Drainage Engineering,** v. 132, n. 6, p. 564-578, 2006. ISSN 0733-9437.

ITENFISU, D. et al. Comparison of reference evapotranspiration calculations as part of the ASCE standardization effort. **Journal of Irrigation and Drainage Engineering,** v. 129, n. 6, p. 440-448, 2003. ISSN 0733-9437.

IZADIFAR, Z.; ELSHORBAGY, A. Prediction of hourly actual evapotranspiration using neural networks, genetic programming, and statistical models. **Hydrological processes,** v. 24, n. 23, p. 3413-3425, 2010. ISSN 1099-1085.

JACKSON, R.; REGINATO, R.; IDSO, S. Wheat canopy temperature: a practical tool for evaluating water requirements. **Water Resources Research,** v. 13, n. 3, p. 651-656, 1977. ISSN 1944-7973.

JACOBS, J. M.; SATTI, S. R.; FITZGERALD, J. M. Evaluation of reference evapotranspiration methodologies and AFSIRS crop water use simulation model. **Publ. SJ2001-SP8. Available online at http://www. sjrwmd. com/programs/outreach/pubs/techpubs/sj2001-sp8. pdf (Verified 20 October, 2003)**, 2001.

JAIN, S.; DAS, A.; SRIVASTAVA, D. Application of ANN for reservoir inflow prediction and operation. **Journal of Water Resources Planning and Management,** v. 125, n. 5, p. 263-271, 1999. ISSN 0733-9496.

JAIN, S.; NAYAK, P.; SUDHEER, K. Models for estimating evapotranspiration using artificial neural networks, and their physical interpretation. **Hydrological Processes,** v. 22, n. 13, p. 2225-2234, 2008. ISSN 1099-1085.

JENSEN, M.; BURMAN, R.; ALLEN, R. Evaporation and irrigation water requirements. ASCE Manuals and Reports on Eng. Practices No. 70. **Am. Soc. Civil Eng., New York, NY**, p. 978-970, 1990.

JENSEN, M. E.; HAISE, H. R. Estimating evapotranspiration from solar radiation. **Proceedings of the American Society of Civil Engineers, Journal of the Irrigation and Drainage Division,** v. 89, p. 15-41, 1963.

JI, L.; PETERS, A. J. Assessing vegetation response to drought in the northern Great Plains using vegetation and drought indices. **Remote Sensing of Environment,** v. 87, n. 1, p. 85-98, 2003. ISSN 0034-4257.

JOTHIPRAKASH, V.; RAMACHANDRAN, M.; SHANMUGANATHAN, P. Artificial neural network model for estimation of REF-ET. **Journal of the Institution of Engineers(India), Part CV, Civil Engineering Division,** v. 83, n. 1, p. 17-20, 2002. ISSN 0373-1995.

KASHYAP, P.; PANDA, R. Evaluation of evapotranspiration estimation methods and development of crop-coefficients for potato crop in a sub-humid region. **Agricultural water management,** v. 50, n. 1, p. 9-25, 2001. ISSN 0378-3774.

KHOOB, A. R. Artificial neural network estimation of reference evapotranspiration from pan evaporation in a semi-arid environment. **Irrigation Science,** v. 27, n. 1, p. 35-39, 2008. ISSN 0342-7188.

KISI, O. Daily pan evaporation modelling using a neuro-fuzzy computing technique. **Journal of Hydrology,** v. 329, n. 3, p. 636-646, 2006. ISSN 0022-1694.

KISI, O. Evapotranspiration estimation using feed-forward neural networks. **Hydrology Research,** v. 37, n. 3, p. 247-260, 2006. ISSN 0029-1277.

KISI, O. Generalized regression neural networks for evapotranspiration modelling. **Hydrological sciences journal,** v. 51, n. 6, p. 1092-1105, 2006. ISSN 0262-6667.

KISI, O. Evapotranspiration modelling from climatic data using a neural computing technique. **Hydrological Processes,** v. 21, n. 14, p. 1925-1934, 2007. ISSN 1099-1085.

KISI, O. The potential of different ANN techniques in evapotranspiration modelling. **Hydrological Processes,** v. 22, n. 14, p. 2449-2460, 2008. ISSN 1099-1085.

KIZER, M. A.; ELLIOTT, R. L.; STONE, J. F. Hourly ET model calibration with eddy flux and energy balance data. **Journal of Irrigation and Drainage Engineering,** v. 116, n. 2, p. 172-181, 1990. ISSN 0733-9437.

KO, J. et al. Determination of growth-stage-specific crop coefficients (Kc) of cotton and wheat. **Agricultural Water Management,** v. 96, n. 12, p. 1691-1697, 2009. ISSN 0378-3774.

KOTSOPOULOS, S.; BABAJIMOPOULOS, C. Analytical estimation of modified Penman equation parameters. **Journal of Irrigation and Drainage Engineering,** v. 123, n. 4, p. 253-256, 1997. ISSN 0733-9437.

KUMAR, M. et al. Estimating evapotranspiration using artificial neural network. **Journal of Irrigation and Drainage Engineering,** v. 128, n. 4, p. 224-233, 2002. ISSN 0733-9437.

LANDERAS, G.; ORTIZ-BARREDO, A.; LÓPEZ, J. J. Comparison of artificial neural network models and empirical and semi-empirical equations for daily reference evapotranspiration estimation in the Basque Country (Northern Spain). **Agricultural water management,** v. 95, n. 5, p. 553-565, 2008. ISSN 0378-3774.

LAWRIMORE, J. H.; PETERSON, T. C. Pan evaporation trends in dry and humid regions of the United States. **Journal of Hydrometeorology,** v. 1, n. 6, p. 543-546, 2000. ISSN 1525-7541.

LEE, X. et al. Micrometeorological fluxes under the influence of regional and local advection: a revisit. **Agricultural and Forest Meteorology,** v. 122, n. 1, p. 111-124, 2004. ISSN 0168-1923.

LEKKAS, D.; IMRIE, C.; LEES, M. Improved non-linear transfer function and neural network methods of flow routing for real-time forecasting. **Journal of Hydroinformatics,** v. 3, n. 3, p. 153-164, 2001. ISSN 1464-7141.

LIANG, S. Narrowband to broadband conversions of land surface albedo I: Algorithms. **Remote Sensing of Environment,** v. 76, n. 2, p. 213-238, 2001. ISSN 0034-4257.

LIN, G.-F.; CHEN, L.-H. A non-linear rainfall-runoff model using radial basis function network. **Journal of Hydrology,** v. 289, n. 1, p. 1-8, 2004. ISSN 0022-1694.

LINGIREDDY, S. Aquifer parameter estimation using genetic algorithms and neural networks. **Civil Engineering Systems,** v. 15, n. 2, p. 125-144, 1998. ISSN 1028-6608.

LIU, B. et al. A spatial analysis of pan evaporation trends in China, 1955–2000. **Journal of Geophysical Research: Atmospheres,** v. 109, n. D15, 2004. ISSN 2156-2202.

LIU, Q. et al. The temporal trends of reference evapotranspiration and its sensitivity to key meteorological variables in the Yellow River Basin, China. **Hydrological Processes,** v. 24, n. 15, p. 2171-2181, 2010. ISSN 1099-1085.

LIU, Y.; LUO, Y. A consolidated evaluation of the FAO-56 dual crop coefficient approach using the lysimeter data in the North China Plain. **Agricultural Water Management,** v. 97, n. 1, p. 31-40, 2010. ISSN 0378-3774.

LU, J. et al. **A comparison of six potential evapotranspiration methods for regional use in the southeastern United States1**: Wiley Online Library 2005.

LUND, M.; SOEGAARD, H. Modelling of evaporation in a sparse millet crop using a two-source model including sensible heat advection within the canopy. **Journal of hydrology,** v. 280, n. 1, p. 124-144, 2003. ISSN 0022-1694.

MACARTHUR, R. H. On the relative abundance of bird species. **Proceedings of the National Academy of Sciences,** v. 43, n. 3, p. 293-295, 1957. ISSN 0027-8424.

MAKKINK, G. Testing the Penman formula by means of lysimeters. **J. Inst. of Water Eng.,** v. 11, p. 277-288, 1957.

MARTÍNEZ-COB, A. Use of thermal units to estimate corn crop coefficients under semiarid climatic conditions. **Irrigation Science,** v. 26, n. 4, p. 335-345, 2008. ISSN 0342-7188.

MAUPIN, M. A. et al. **Estimated use of water in the United States in 2010**. US Geological Survey. 2014. (2330-5703)

MC CUEN, R. H. The role of sensitivity analysis in hydrologic modeling. **Journal of hydrology,** v. 18, n. 1, p. 37-53, 1973. ISSN 0022-1694.

MCNAUGHTON, K.; JARVIS, P. Using the Penman-Monteith equation predictively. **Agricultural Water Management,** v. 8, n. 1-3, p. 263-278, 1984. ISSN 0378-3774.

MEIJNINGER, W. et al. Determination of area-averaged water vapour fluxes with large aperture and radio wave scintillometers over a heterogeneous surface–Flevoland field experiment. **Boundary-Layer Meteorology,** v. 105, n. 1, p. 63-83, 2002. ISSN 0006-8314.

MEYER, W. et al. Effects of soil type on soybean crop water use in weighing lysimeters. **Irrigation Science,** v. 11, n. 2, p. 69-75, 1990. ISSN 0342-7188.

MEYER, W.; MATEOS, L. Soil type effects on soybean crop water use in weighing lysimeters. **Irrigation Science,** v. 11, n. 4, p. 233-237, 1990. ISSN 0342-7188.

MONTEITH, J. **Evaporation and environment. The state and movement of water in living organisms. Symposium of the society of experimental biology, Vol. 19 (pp. 205-234)**: Cambridge: Cambridge University Press 1965.

MORRIS, M. D. Factorial sampling plans for preliminary computational experiments. **Technometrics,** v. 33, n. 2, p. 161-174, 1991. ISSN 0040-1706.

MORTON, C. G. et al. Assessing calibration uncertainty and automation for estimating evapotranspiration from agricultural areas using METRIC. **JAWRA Journal of the American Water Resources Association,** v. 49, n. 3, p. 549-562, 2013. ISSN 1752-1688.

MU, Q. et al. Development of a global evapotranspiration algorithm based on MODIS and global meteorology data. **Remote sensing of Environment,** v. 111, n. 4, p. 519-536, 2007. ISSN 0034-4257.

NAGLER, P. et al. Evapotranspiration in a cottonwood (Populus fremontii) restoration plantation estimated by sap flow and remote sensing methods. **Agricultural and Forest Meteorology,** v. 144, n. 1, p. 95-110, 2007. ISSN 0168-1923.

NEALE, C. M.; KRUSE, E. G.; YODER, R. E. Field experience with hydraulic weighing lysimeters. Lysimeters for evapotranspiration and environmental measurements, 1991, ASCE. p.160-169.

NORTON, J. P. Selection of Morris trajectories for initial sensitivity analysis. **IFAC Proceedings Volumes,** v. 42, n. 10, p. 670-674, 2009. ISSN 1474-6670.

OHMURA, A.; WILD, M. Is the hydrological cycle accelerating? **Science,** v. 298, n. 5597, p. 1345-1346, 2002. ISSN 0036-8075.

OKE, T. R. Initial guidance to obtain representative meteorological observations at urban sites. 2004.

OUDIN, L. et al. Which potential evapotranspiration input for a lumped rainfall–runoff model?: Part 2—Towards a simple and efficient potential evapotranspiration model for rainfall–runoff modelling. **Journal of hydrology,** v. 303, n. 1, p. 290-306, 2005. ISSN 0022-1694.

PANDEY, P. K.; PANDEY, V. Lysimeter based crop coefficients for estimation of crop evapotranspiration of black gram (Vigna Mungo L.) in sub-humid region. **International Journal of Agricultural and Biological Engineering,** v. 4, n. 4, p. 50-58, 2012. ISSN 1934-6352.

PAPPENBERGER, F. et al. Multi-method global sensitivity analysis of flood inundation models. **Advances in water resources,** v. 31, n. 1, p. 1-14, 2008. ISSN 0309-1708.

PAULSON, C. A. The mathematical representation of wind speed and temperature profiles in the unstable atmospheric surface layer. **Journal of Applied Meteorology,** v. 9, n. 6, p. 857-861, 1970. ISSN 0021-8952.

PEACOCK, C.; HESS, T. Estimating evapotranspiration from a reed bed using the Bowen ratio energy balance method. **Hydrological Processes,** v. 18, n. 2, p. 247-260, 2004. ISSN 1099-1085.

PENMAN, H. L. Natural evaporation from open water, bare soil and grass. Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 1948, The Royal Society. p.120-145.

PERRIER, A. Land surface processes: vegetation. **Land surface processes in atmospheric general circulation models**, p. 395-448, 1982.

PERRIER, A. Updated evapotranspiration and crop water requirement definitions. **Crop Water Requirements**, p. 885-887, 1985.

PERRIER, A. et al. A new device for continuous recording of the energy balance of natural surfaces. **Agricultural Meteorology,** v. 16, n. 1, p. 71-84, 1976. ISSN 0002-1571.

PERRIER, A.; TUZET, A. Land surface processes: Description, theoretical approaches, and physical laws underlying their measurements. In: (Ed.). **Land Surface Evaporation**: Springer, 1991. p.145-155.

PETERSON, T. Evaporation losing its strength. **Nature,** v. 377, p. 687-688, 1995.

PETTORELLI, N. et al. Using the satellite-derived NDVI to assess ecological responses to environmental change. **Trends in ecology & evolution,** v. 20, n. 9, p. 503-510, 2005. ISSN 0169-5347.

PRUEGER, J. H. et al. Aerodynamic parameters and sensible heat flux estimates for a semi-arid ecosystem. **Journal of Arid Environments,** v. 57, n. 1, p. 87-100, 2004. ISSN 0140-1963.

RAYNER, D. Wind run changes: the dominant factor affecting pan evaporation trends in Australia. **Journal of Climate,** v. 20, n. 14, p. 3379-3394, 2007. ISSN 1520-0442.

REYES-GONZÁLEZ, A. et al. Comparative Analysis of METRIC Model and Atmometer Methods for Estimating Actual Evapotranspiration. **International Journal of Agronomy,** v. 2017, 2017. ISSN 1687-8159.

RIJTEMA, P. E. **An analysis of actual evapotranspiration**. 1965. Pudoc

RODERICK, M. L.; FARQUHAR, G. D. Changes in Australian pan evaporation from 1970 to 2002. **International Journal of Climatology,** v. 24, n. 9, p. 1077-1090, 2004. ISSN 1097-0088.

RODERICK, M. L. et al.. Changes in New Zealand pan evaporation since the 1970s. **International Journal of Climatology,** v. 25, n. 15, p. 2031-2039, 2005. ISSN 1097-0088.

RODERICK, M. L. et al. On the attribution of changing pan evaporation. **Geophysical Research Letters,** v. 34, n. 17, 2007. ISSN 1944-8007.

ROSENBERRY, D. O. et al. Comparison of 13 equations for determining evapotranspiration from a prairie wetland, Cottonwood Lake area, North Dakota, USA. **Wetlands,** v. 24, n. 3, p. 483-497, 2004. ISSN 0277-5212.

ROSENTHAL, W. D.; BLANCHARD, B. J.; BLANCHARD, A. J. Visible/infrared/microwave agriculture classification, biomass, and plant height algorithms. **IEEE Transactions on Geoscience and Remote Sensing**, n. 2, p. 84-90, 1985. ISSN 0196-2892.

SALTELLI, A. Making best use of model evaluations to compute sensitivity indices. **Computer Physics Communications,** v. 145, n. 2, p. 280-297, 2002. ISSN 0010-4655.

SCOTT, C. A.; BASTIAANSSEN, W. G.; AHMAD, M.-U.-D. Mapping root zone soil moisture using remotely sensed optical imagery. **Journal of Irrigation and Drainage Engineering,** v. 129, n. 5, p. 326-335, 2003. ISSN 0733-9437.

SENAY, G. B. et al. Operational evapotranspiration mapping using remote sensing and weather datasets: A new parameterization for the SSEB approach. **JAWRA Journal of the American Water Resources Association,** v. 49, n. 3, p. 577-591, 2013. ISSN 1752-1688.

SHEN, Z. et al. Parameter uncertainty analysis of the non-point source pollution in the Daning River watershed of the Three Gorges Reservoir Region, China. **Science of the total environment,** v. 405, n. 1, p. 195-205, 2008. ISSN 0048-9697.

SHUTTLEWORTH, W. J. **Evaporation. Chapter 4 in Handbook of Hydrology**: McGraw-Hill, USA 1993.

SHUTTLEWORTH, W. J.; WALLACE, J. Evaporation from sparse crops-an energy combination theory. **Quarterly Journal of the Royal Meteorological Society,** v. 111, n. 469, p. 839-855, 1985. ISSN 1477-870X.

SINGH, R. K.; SENAY, G. B. Comparison of four different energy balance models for estimating evapotranspiration in the Midwestern United States. **Water,** v. 8, n. 1, p. 9, 2015.

SINGH, R. P.; ROY, S.; KOGAN, F. Vegetation and temperature condition indices from NOAA AVHRR data for drought monitoring over India. **International Journal of Remote Sensing,** v. 24, n. 22, p. 4393-4402, 2003. ISSN 0143-1161.

SINGH, V.; XU, C. Sensitivity of mass transfer-based evaporation equations to errors in daily and monthly input data. **Hydrological Processes,** v. 11, n. 11, p. 1465-1473, 1997. ISSN 0885-6087.

SLATYER, R.; MCILROY, I. Evaporation and the principle of its measurement. **Practical Meteorology**, 1961.

SMITH, M. et al. Report on the expert consultation on procedures for revision of FAO guidelines for prediction of crop water requirements. **Rome: FAO**, 1991.

SOARES, D. S.; NADAL, J. Aplicação de uma Rede Neural Feedforward com Algoritmo de Levenberg-Marquardt para Classificação de Alterações do Segmento ST do Eletrocardiograma. 1999.

SOUZA, D.; YODER, R. Et estimation in the northeast of Brazil: Hargreaves or Penman-Monteith equation? **Paper**, 1994. ISSN 0149-9890.

SRIVASTAVA, S. et al. Interlinkages of NOAA/AVHRR derived integrated NDVI to seasonal precipitation and transpiration in dryland tropics. **International Journal of Remote Sensing,** v. 18, n. 14, p. 2931-2952, 1997. ISSN 0143-1161.

STEPHENS, D.; LYONS, T. Rainfall-yield relationships across the Australian wheatbelt. **Australian Journal of Agricultural Research,** v. 49, n. 2, p. 211-224, 1998. ISSN 1444-9838.

SUDHEER, K.; GOSAIN, A.; RAMASASTRI, K. Closure to "Estimating Actual Evapotranspiration from Limited Climatic Data Using Neural Computing Technique" by KP Sudheer, AK Gosain, and KS Ramasastri. **Journal of Irrigation and Drainage Engineering,** v. 131, n. 2, p. 220-220, 2005. ISSN 0733-9437.

SUDHEER, K.; JAIN, A. Explaining the internal behaviour of artificial neural network river flow models. **Hydrological Processes,** v. 18, n. 4, p. 833-844, 2004. ISSN 1099-1085.

SUN, X. et al. Three complementary methods for sensitivity analysis of a water quality model. **Environmental Modelling & Software,** v. 37, p. 19-29, 2012. ISSN 1364-8152.

TAHIR, S. Estimating potential evaporation using artificial neural network. In: (Ed.): Jakarta: INACID, 1998.

TANNER, C.; PELTON, W. Potential evapotranspiration estimates by the approximate energy balance method of Penman. **Journal of geophysical research,** v. 65, n. 10, p. 3391-3413, 1960. ISSN 2156-2202.

TASUMI, M. **Progress in operational estimation of regional evapotranspiration using satellite imagery**. 2003. ISBN 0496281399.

TASUMI, M.; ALLEN, R.; TREZZA, R. Estimation of at surface reflectance and albedo from satellite for routine, operational calculation of land surface energy balance. journal of hydrologic engineering. 2007.

TEIXEIRA, A. D. C.; BASTIAANSSEN, W.; BASSOI, L. Crop water parameters of irrigated wine and table grapes to support water productivity analysis in the Sao Francisco river basin, Brazil. **Agricultural Water Management,** v. 94, n. 1, p. 31-42, 2007. ISSN 0378-3774.

THOM, A.; THONY, J. L.; VAUCLIN, M. On the proper employment of evaporation pans and atmometers in estimating potential transpiration. **Quarterly Journal of the Royal Meteorological Society,** v. 107, n. 453, p. 711-736, 1981. ISSN 1477-870X.

THOMAS, A. Spatial and temporal characteristics of potential evapotranspiration trends over China. **International Journal of Climatology,** v. 20, n. 4, p. 381-396, 2000. ISSN 0899-8418.

THORNTHWAITE, C. W. An approach toward a rational classification of climate. **Geographical review,** v. 38, n. 1, p. 55-94, 1948. ISSN 0016-7428.

TRAJKOVIC, S.; TODOROVIC, B.; STANKOVIC, M. Forecasting of reference evapotranspiration by artificial neural networks. **Journal of Irrigation and Drainage Engineering,** v. 129, n. 6, p. 454-457, 2003. ISSN 0733-9437.

TURC, L. Estimation of irrigation water requirements, potential evapotranspiration: a simple climatic formula evolved up to date. **Ann Agron,** v. 12, n. 1, p. 13-49, 1961.

TYAGI, N.; SHARMA, D.; LUTHRA, S. Determination of evapotranspiration and crop coefficients of rice and sunflower with lysimeter. **Agricultural Water Management,** v. 45, n. 1, p. 41-54, 2000. ISSN 0378-3774.

VENTURA, F. et al. An evaluation of common evapotranspiration equations. **Irrigation Science,** v. 18, n. 4, p. 163-170, 1999. ISSN 0342-7188.

VÖRÖSMARTY, C. J. et al. Global water resources: vulnerability from climate change and population growth. **science,** v. 289, n. 5477, p. 284-288, 2000. ISSN 0036-8075.

W.M.O. Guide to Meteorological Instruments and Observing Practices. v. nº 8 (fifth edition), 1983.

WANG, Q. et al. Using NOAA AVHRR data to assess flood damage in China. **Environmental monitoring and assessment,** v. 82, n. 2, p. 119-148, 2003. ISSN 0167-6369.

WANG, Y. et al. Changes of pan evaporation and reference evapotranspiration in the Yangtze River basin. **Theoretical and Applied Climatology,** v. 90, n. 1-2, p. 13-23, 2007. ISSN 0177-798X.

WEBB, E. K. Profile relationships: The log-linear range, and extension to strong stability. **Quarterly Journal of the Royal Meteorological Society,** v. 96, n. 407, p. 67-90, 1970. ISSN 1477-870X.

WEN, C. G.; LEE, C. S. A neural network approach to multiobjective optimization for water quality management in a river basin. **Water Resources Research,** v. 34, n. 3, p. 427-436, 1998. ISSN 1944-7973.

WILBY, R.; ABRAHART, R.; DAWSON, C. Detection of conceptual model rainfall—runoff processes inside an artificial neural network. **Hydrological Sciences Journal,** v. 48, n. 2, p. 163-181, 2003. ISSN 0262-6667.

WRIGHT, J. L. New evapotranspiration crop coefficients. **Proceedings of the American Society of Civil Engineers, Journal of the Irrigation and Drainage Division,** v. 108, n. IR2, p. 57-74, 1982.

WRIGHT, J. L.; ALLEN, R. G.; HOWELL, T. A. Conversion between evapotranspiration references and methods. 2000.

WYLIE, B. K. et al. Calibration of remotely sensed, coarse resolution NDVI to CO 2 fluxes in a sagebrush–steppe ecosystem. **Remote Sensing of Environment,** v. 85, n. 2, p. 243-255, 2003. ISSN 0034-4257.

YANG, C.-C.; PRASHER, S.; LACROIX, R. Applications of artificial neural networks to land drainage engineering. **Transactions of the ASAE,** v. 39, n. 2, p. 525-533, 1996.

YIN, Y. et al. Attribution analyses of potential evapotranspiration changes in China since the 1960s. **Theoretical and Applied Climatology,** v. 101, n. 1-2, p. 19-28, 2010. ISSN 0177-798X.

YIN, Y.; WU, S.; DAI, E. Determining factors in potential evapotranspiration changes over China in the period 1971–2008. **Chinese Science Bulletin,** v. 55, n. 29, p. 3329-3337, 2010. ISSN 1001-6538.

YU, P.-S.; YANG, T.-C.; CHOU, C.-C. Effects of climate change on evapotranspiration from paddy fields in southern Taiwan. **Climatic Change,** v. 54, n. 1, p. 165-179, 2002. ISSN 0165-0009.

ZANETTI, S. et al. Estimating evapotranspiration using artificial neural network and minimum climatological data. **Journal of Irrigation and Drainage Engineering,** v. 133, n. 2, p. 83-89, 2007. ISSN 0733-9437.

ZEALAND, C. M.; BURN, D. H.; SIMONOVIC, S. P. Short term streamflow forecasting using artificial neural networks. **Journal of hydrology,** v. 214, n. 1, p. 32-48, 1999. ISSN 0022-1694.

ZHANG, Y. et al. Trends in pan evaporation and reference and actual evapotranspiration across the Tibetan Plateau. **Journal of Geophysical Research: Atmospheres,** v. 112, n. D12, 2007. ISSN 2156-2202.