Revista Brasileira de Computação Aplicada, November, 2019

DOI: 10.5335/rbca.v11i3.10026
Vol. 11, Nº 3, pp. 146–154
Homepage: seer.upf.br/index.php/rbca/index

ISSN 2176-6649

ORIGINAL PAPER

Availability analysis of the Brazilian's national weather measurement system

Diego de Freitas Bezerra[1], José Clodoalves da Silva Júnior[1], Glauco Estácio Gonçalves[1], Victor Wanderley Costa de Medeiros[1]

[1]Department of Statistics and Informatics - Federal Rural University of Pernambuco, Recife, PE – Brazil

*diego.freitas@ufrpe.br, jose.clodoalves@ufrpe.br, glauco.goncalves@ufrpe.br, victor.medeiros@ufrpe.br

Received: 2019-10-01. **Revised:** 2019-10-15. **Accepted:** 2019-10-23.

Abstract

Weather measurement systems became an important tool for the efficient operation of various economic activities. Automated irrigation systems, that improve agricultural productivity and reduce the consumption of water resources, relies on data collected by these systems, for example. Due to the inherent complexity of these systems (i.e. stations with multiple sensors communicating through multiple communication channels to cloud services), it is very important to have measures that clarify how faults behave allowing better planning of maintenance and establish a degree of systems' reliability. This work presents a study of the availability of all meteorological stations of the National Institute of Meteorology - INMET installed in the Brazilian territory in the year 2017. The results present the first analysis of this parameter and serve both for academic and commercial users, as a form of measurement of these systems' reliability, as well as for weather measurement infrastructure providers as a tool for improving the effectiveness of their maintenance policy and as a support for the strategic planning of new investments.

Keywords: reliability analysis; systems availability; evapotranspiration

Resumo

Os sistemas de medição do tempo tornaram-se uma ferramenta importante para o funcionamento eficiente de diversas atividades econômicas. Sistemas automatizados de irrigação, que melhoram a produtividade agrícola e reduzem o consumo de recursos hídricos, dependem de dados coletados por esses sistemas, por exemplo. Devido à complexidade inerente desses sistemas (i.e., estações com múltiplos sensores se comunicando através de múltiplos canais de comunicação para serviços em nuvem), é muito importante ter medidas que esclarecem como as falhas se comportam, permitindo um melhor planejamento de manutenção e um grau de confiabilidade de sistemas. Este trabalho apresenta um estudo da disponibilidade de todas as estações meteorológicas do Instituto Nacional de Meteorologia-INMET instaladas no território brasileiro no ano de 2017. Os resultados apresentam a primeira análise deste parâmetro e servem tanto para usuários acadêmicos como comerciais, como forma de mensuração da confiabilidade desses sistemas, bem como para os provedores de infraestrutura de medição meteorológica como ferramenta para melhorar a eficácia de sua política de manutenção e como suporte para o planejamento estratégico de novos investimentos.

Palavras-Chave: análise de confiabilidade; disponibilidade de sistemas; evapotranspiração

1 Introduction

Today weather measurement systems provide continuous, updated, and autonomous monitoring

of climate variables, which are necessary for alerts of harmful incidents and weather forecast. The weather measurement systems are also fundamental for decision making in a broad range of economic activities as farms – for irrigation –, industries – for monitoring air and bodies of water –, data centers – for free cooling –, and others.

An weather measurement system is composed of ground stations, radar stations, satellites, and other instruments that send hourly or minute-based measurements over data communication networks to servers in data centers for treatment, processing, and analysis. This complex chain for data acquisition and processing involves a number of IT components that can fail. Starting from a sensor at the ground station to the network card at a database server, all components of the weather measurement system can fail.

A system is under failure when the service that it delivers deviates from its original purpose (Arjannikov et al., 2017). Considering a broad view, there are three main possible deviations: complete failures, partial failures, or bad measurements. A system is in a complete failure when an essential component fails and the station becomes inoperative. A partial failure can be detected when a non-essential component is damaged and some of the data is lost. The bad measurements occur when a sensor provide inconsistent data.

The bad measurements can be considered as a failure because it deviates the system from its original purpose, which is founded on delivering trustworthy data. This type of deviation, if not treated, put people and business that depends on the collected data at serious risk. Moreover, treating this problem is hard because there is no established general method for detecting drifts and coping it (von Arx et al., 2013).

Partial and complete failures can be studied using methods from the reliability theory, which offers statistical methods for the analysis of how a system behaves under failures. This theory formally defines properties as reliability, availability, risk, etc, and it also provides mathematical models for assessing those properties (Rausand and Arnljot, 2004). Availability provides a simple measure to capture the amount of time a system delivers its service during a stated period. Despite simple, this metric is versatile, since it can be recursively employed to assess subsystems (or the components) of a system.

Under the perspective of the weather measurement systems maintainer, the cautious study of the availability of the systems' components is important to drive its maintenance policies and forthcoming investments. On the other hand, considering under the perspective of the system's final users, the station availability is important to allow the choice of the right source of data when some alternative stations are available.

This paper analyzes the availability, during the year of 2017, of the network of automatic weather stations of the National Institute of Meteorology (Instituto Nacional de Meteorologia - INMET), which collects environmental data all over Brazil. These data are an important source for academic studies and commercial applications in Brazil. The results found in our study intends to give a first look concerning availability of this important service, which can be important to its users and to the institute.

This paper is organized as follows: Section 2 describes the process employed into our evaluation; Section 3 presents the obtained results; Section 4 presents a discusses of the obtained results; At last, Section 5 discusses some conclusions and future work.

2 Data and method

This section details each step of the availability assessment study conducted by us over the INMET's weather measurement system. This study evaluates only the automatic station network and it do not covers the conventional (read by humans) stations and the radiosondes belonging to INMET. We choose to study only the automatic ground stations due to its relevance to a broad community and because the automatic nature of the telemetry process tends to reduce the sources of failure.

Moreover, we do not consider other important components of the system as database servers, routers, web servers, applications, and so on. In other words, our availability analysis focuses on the indispensable part of the system, whose failure produces irrecoverable gaps at the historical data, which is the most valuable product of the INMET's weather measurement system. It is important to highlight also that we assess the system from outside, i.e. in the point of view of the final users of the data. This way, we intend to provide to the academic community results that can be useful to critically evaluate the most important weather measurement service of Brazil.

2.1 Data

The historical series of the weather data analyzed in this work present hourly measurements ranging from jan/01/2017 to dec/31/2017. These series contains data about 490 stations distributed among the Brazilian states. For each automatic weather station, one can retrieve from INMET's website[1] data as latitude and longitude, and hourly wheater data from 17 environmental measures.

The names of these environmental measurements and the variable names were defined to reference it in this paper: average temperature of the air (T_{mean}), maximum temperature of the air (T_{max}), minimum temperature of the air (T_{min}), average relative humidity of the air (RH_{mean}), maximum relative humidity of the air (RH_{max}), minimum relative humidity of the air (RH_{min}), average dew point ($Tdew_{mean}$), maximum dew point ($Tdew_{max}$), minimum dew point ($Tdew_{min}$), average air pressure (P_{mean}), maximum air pressure (P_{max}), minimum air pressure (P_{min}), wind direction (u_{dir}), wind speed (u_2), wind burst (u_{burst}), solar

[1]Data from those stations can be found at http://www.inmet.gov.br/portal/index.php?r=estacoes/estacoesAutomaticas

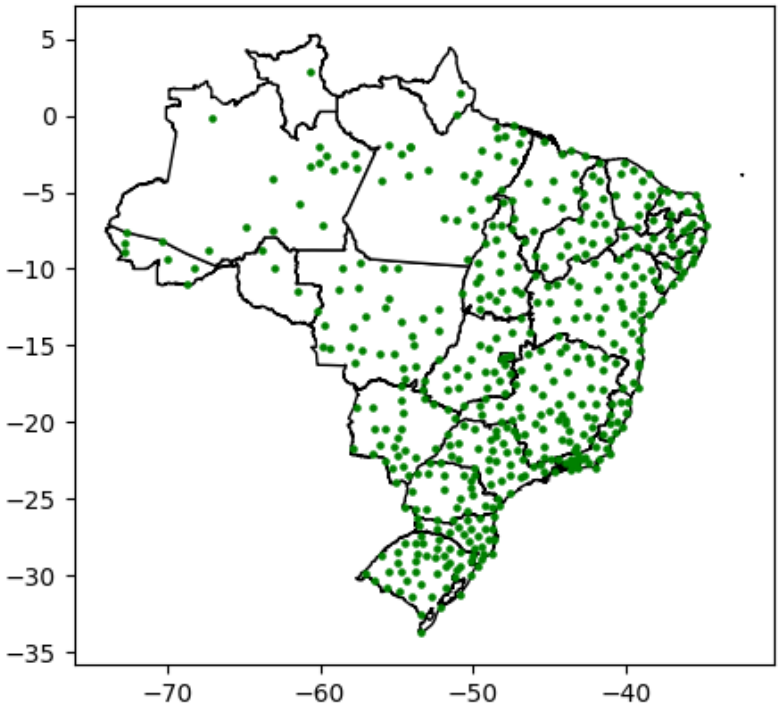

Figure 1: Spatial distribution of weather stations

radiation (R_s), and precipitation (pr).

Please note that, although the environmental measures can be derived from the same physical sensor (e.g., the hourly maximum and minimum temperatures can be obtained from the time series of instantaneous measurements of the temperature sensor), this paper calls each environmental measure as a sensor. We chose this name firstly because the technical note that describes INMET's automatic weather stations network (INMET, 2011) does not explains how each of the 17 environmental measurement are captured, and secondly because our analysis is from the point of view of a final user, that tend to use each environmental measure without considering its relationships.

The data set, available in CSV format (comma separated values), presents 77,885,160 measurements distributed among the 490 automatic weather stations during the period 8,760 hours. It is important to highlight that were initially obtained data from 523 stations but 33 stations did not present a minimum of 8,760 hours of operation. These stations start to operate during 2017 and did not complete the one-year cycle. Thus, based on this criterion, they were not considered in this analysis.

In order to illustrate the coverage of the INMET's network of ground stations, Fig. 1 diplays the 490 stations in the Brazillian territory based on its respective latitude and longitude data.

The figure shows that the INMET network was present at all Brazilian states in 2017, but it also depicts a high concentration of stations close to the coastline of Brazil (encompassing several states at the Northeast, Southeast, and South regions) that scatter as one moves to the west of the country. In order to better illustrate this skewed coverage, the Table 1 shows the total number of stations that operated throughout the year of 2017 in the respective territory[2]. The table orders the Brazilian states by the territory area covered by an station in the state (increasing order).

Although the state of Minas Gerais (MG) presents the largest number of stations in its territory, the dimension of this state makes each station covers a considerable area. The opposite happens with the state of Rio de Janeiro (RJ) that presents 20 stations and an area approximately 13 times smaller, and with the Federal District (DF) that has only 2 stations in an area almost 8 times smaller than that of the RJ. The table also shows that Roraima (RR) and Amazonas (AM) have extremely large areas for each station to serve. In the case of Roraima, there is a single station for an area of 224,300.805 km^2, which is almost 103 times larger

[2]Area measurements presented in Table 1 were obtained from the Brazilian Institute of Geography and Statistics (IBGE) at https://www.ibge.gov.br/cidades-e-estado

Table 1: Spatial distribution of weather stations by state

State	Stations	Area per station (km^2)	State	Stations	Area per station (km^2)
RJ	20	2,189.079	MS	27	13,227.612
DF	2	2,889.999	PI	19	13,242.733
SC	21	4,558.950	BA	42	13,446.011
AL	6	4,641.357	TO	20	13,886.021
ES	9	5,120.767	GO	23	14,787.239
SE	4	5,479.611	AC	7	23,446.248
RS	42	6,708.045	MA	15	22,129.129
PB	8	7,058.554	MT	30	30,106.748
RN	7	7,544.444	PA	27	46,220.564
SP	31	8,007.085	RO	4	59,441.323
PE	12	8,173.000	AP	2	71,414.261
PR	23	8,665.563	AM	15	103,943.125
MG	60	9,775.345	RR	1	224,300.805
CE	13	11,452.895			

than the area covered by each station in RJ.

2.2 Method

In our analysis, a station is considered to be the aggregation of several components as the datalogger, its firmware, its programming software, battery, solar panel, and the communication link, but it excludes the station sensors because these were treated individually. This way, we assume that the availability estimate of an station is impacted by failures of each one of these components. But, the root cause of a station failure is not discussed in this paper due to the limits of our analysis.

To analyze the availability of a system, subsystem, or component, one must observe the frequency and duration of failures during the time interval at which the system should operate. Thus, the availability is a function of the mean time to failure (MTTF) and the mean time to repair (MTTR) of a system. Formally, it is defined as:

$$availability = \frac{MTTF}{MTTF + MTTR} \quad (1)$$

Our main assumption in this paper is that the availability of each station, as well as the availability of each sensor at each station, can be inferred from the existing gaps in the correspondent measurement. In other words, the records of the historical series that presented missing measurements were considered as a failure in the respective sensor during the respective hour.

In turn, in order to measure station availability, we assume that when no data is collected (i.e., there is no data from all sensors), the station is under failure. This assumption does not implies that the sensor is in failure. Actually, in those cases, we assume that there is no sensor failure. In other words, we assume that the probability of simultaneous failures of the station and the sensor are negligible.

Note that the analysis of the station and sensor availability allows estimating the availability of environmental indexes that require data from one sensor (e.g., precipitation) or data from a set of sensors (e.g., evapotranspiration). In order to illustrate such a concept, we estimated in this work the reference evapotranspiration (ET_o), which is an important component of the hydrological cycle defined as the combination of the processes of water loss by evaporation from the soil and transpiration from vegetation (Xavier and Brochado, 2017).

The low availability of measurements can difficult or even make unfeasible the hourly estimation of evapotranspiration (Moura et al., 2010), which is considered an important information for water management in agriculture (Jensen and Allen, 2016). This way, our analysis seeks to verify how the availability of weather data can affect the availability of the hourly estimates of evapotranspiration.

There are several methods based on meteorological data to estimate ET_o, however the FAO-56 Penman-Monteith, Hargreaves and Turc methods are commonly used to estimate ET_o. This set of methods are interesting to demonstrate our point about the service availability of a sensor because they demand a different set of parameters to be calculated. The FAO-56 Penman-Monteith method, for example, demands a high number of environmental parameters, whereas the Hargreaves and Turc methods are most recommended when there is low availability of environmental data (Fisher and Pringle III, 2013). Each method requires different parameters to calculate ET_o, therefore to estimate the availability of a service for calculating ET_o we employ a set of specific sensors for each method.

The FAO-56 Penman-Monteith method for daily ET_o estimation (Fisher and Pringle III, 2013) may be written as:

$$ET_o = \frac{0.408\Delta(R_n - G) + \gamma\left(\frac{900}{T_{mean}+273}\right)u_2(e_s - e_a)}{\Delta + \gamma(1 + 0.34u_2)} \quad (2)$$

where R_n is the net solar radiation, G is the soil heat flux, γ is the psychometric constant, e_s is the saturation pressure, e_a is the actual vapor pressure, and Δ is the

slope of vapor curve. This way, it is possible to estimate ET_o from data of the INMET service combining the following sensors: the average air temperature, wind speed, solar radiation, and the dew point temperature.

For the Hargreaves method the evaporation estimate is based on T_{mean}, T_{max}, and T_{min} (Fisher and Pringle III, 2013) and it may be written as:

$$ET_o = 0.023(T_{mean} + 17.8)(T_{max} - T_{min})^{0.5} R_a \quad (3)$$

where R_a is the extraterrestrial radiation. R_a is estimated based on a specific location and day of the year, which is independent of the INMET service.

The Turc method depends on the maximum and minimum air temperature and daily solar radiation (Fisher and Pringle III, 2013). Thus, the ET_o can be obtained from:

$$ET_o = 0.40\left(\frac{T_{mean}}{T_{mean} + 15}\right)(R_s + 50) \quad (4)$$

where R_s is the solar radiation.

Note that each method requires a specific set of sensors from the INMET weather measurement service to estimate the ET_o. Therefore, following the systems reliability theory and assuming independence of the sensors and station, the service availability for the reference evapotranspiration of a specific method is the product among the availability of the station and the respective availability of each sensor needed to calculate ET_o with the considered method. Eq. (5) presents the calculation of the ET_o availability:

$$A_{ET_o} = A_{station} \prod_{i=1}^{n} A_{sensor_i}, \quad (5)$$

where $A_{station}$ is the station availability and A_{sensor_i} is the availability of each sensor of the set of sensors required by the method.

Please note that it is possible that some missing variables can be estimated from other variables, for example, R_s can be estimated from T_{max} and T_{min} (Fisher and Pringle III, 2013). However, it is not the purpose of this paper to analyze alternatives for estimating missing environmental variables from other most available variables. The focus is just to illustrate how the availability of the service for calculating evapotranspiration in its standard form can be found.

3 Results

This section presents the availability analysis of the INMET's weather measurement system. Section 3.1 gives a broad view of the system, focusing on the availability of the stations, whereas Section 3.2 focuses on the sensors (see Table 5). Section 3.3 discusses the availability of environmental indexes that depends on multiple sensors.

3.1 Station availability

A general analysis of the weather stations was performed using the dataset obtained from INMET. The MTTF and MTTR of each station were calculated in order to assess its availability. The Table 2 summarizes the MTTF and MTTR data. One can note that at least 75% of the stations operate during the year with at least one failure. It is also noteworthy that at least 25% of the stations operate less than a week (168h) before a failure occurs. This percentage represents 122 stations that fail at least once a week.

When a failure occurs its service is quickly recovered, since 75% of the stations are repaired in less than a day (20.9h). Despite this, there is a considerable number of stations whose time to repair is high, with the mean reaching more than 5 days (120.0h) and the maximum recovery time being 7.5 months (5,425.0h). It is also possible to verify that the 25% stations with the highest MTTR value presents very highly variable MTTR. This variation occurs in the interval between 21 hours and 7.5 months.

From the values obtained for the MTTF and MTTR of each station, the summary of the availability of the stations was calculated and presented at Table 3.

From these results, one can note that in average the INMET station were down during 33 days (about 9%) in 2017, but it must be considered that the top 25% most available stations were down less than 2 hours in a year, actually, a total of 104 (about 21%) stations were fully available during 2017. Moreover, 99 stations (about 20%) presented at least two 9's of availability and 287 stations (59%) have less than two 9's availability.

Fig. 2 presents the spatial proportion of stations with availability lower than two 9's. The sidebar defines the proportion of stations (considering only the stations at the respective Brazilian state) and the more intense color, the greater percentage of stations with low availability.

In general, states from the North region of Brazil present a low availability. Considering Roraima (RR) and Amapá (AP), for example, all stations are poorly

Table 2: Weather stations MTTF and MTTR

	Min	1st Qu.	Median	Mean	3rd Qu.	Max	SD
MTTF	6.0h	127.7h	792.4h	2822.0h	4379.5h	8760.0h	3401.1h
MTTR	0.0h	1.0h	7.9h	128.3h	20.9h	5425.0h	490.6h

Table 3: Weather stations availability

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	SD
0.2351	0.8713	0.9760	0.9090	0.9998	1.0000	0.1385

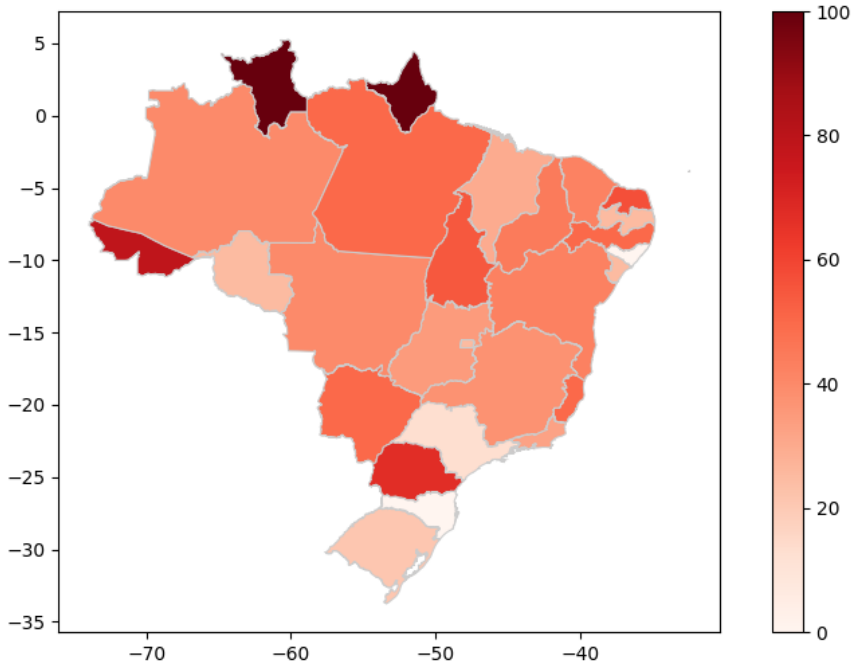

Figure 2: Spatial distribution of stations (%) with less than two 9's of availability

available, and the state of Acre has 86% of its stations in this case. However, sucho characteristic can be found in other states from different regions of Brazil, Paraná (PR) (78%) and Rio Grande do Norte (RN) (71%) are examples of states with a high number of poorly available stations. On the other hand, the states of Santa Catarina (SC) and Alagoas (AL) presented the smaller percentage of stations with low availability (only 33%). The top 5 is completed by the states of São Paulo (SP), Rio Grande do Sul (RS) and Sergipe (SE) with 42%, 48% and 50%, respectively.

Particularly, it can be highlighted 12 stations that were available for less than 50% of the time of the analysis. Four of these stations are in the state of Acre. In other words, those stations were unavailable for more than one semester.

Table 4 names those stations and informs the Brazilian state where they are located and the correspondent availability in 2017.

Table 4: Twelve weather stations with the lowest availability rates

City	State	Availability
Feijó	AC	0.2351
Rio Branco	AC	0.2471
Apui	AM	0.3036
Paranatinga	MG	0.3104
Porto Walter	AC	0.3412
Marechal Thaumaturgo	AC	0.3446
Buritis	MG	0.3692
Ariquemes	RO	0.4178
Salto do Céu	MG	0.4390
Campos Lindos	TO	0.4477
Barbalha	CE	0.4549
Recife	PE	0.4919

3.2 Sensor availability

Based on its MTTF and MTTR, the availability of each sensor at each station was calculated. Table 5 displays an aggregated view of the availability of sensors from different stations.

Despite presenting a high standard deviation value, with confidence interval of 95%, the T_{mean} sensors have the highest average value in terms of availability and are fully available in more than 90% of the cases analyzed (91.84%). On the other hand, the precipitation sensors present the lowest average value of availability among all the other sensors, being fully available in 60%. However, note that all the sensors, 50% of the cases present availability at least three 9's. This means that they have been unavailable for only 9 hours throughout the year 2017.

From Table 5 we can verify that the variation of the availability between sensors that provide data of T_{min}, T_{mean}, and T_{max} is subtle. The same occurs at the relative air humidity, dew point temperature, and pressure sensors. This characteristic can imply in the existence of a correlation between the failures of these sensors. Thus, the Fig. 3 presents a correlation matrix to analyze the correlation of failures between the sensors.

Note that there is a positive correlation between the sensors that offer data of maximum, mean, and minimum of the environmental variables and this correlation is considerably high for these sensors. Thus,

Table 5: Sensors availability

Sensor	Min	1st Qu.	Median	Mean	3rd Qu.	Max	SD
T_{min}	0.5451	0.9983	0.9998	0.9931	1.0000	1.0000	0.0360
T_{mean}	0.5440	1.0000	1.0000	0.9967	1.0000	1.0000	0.0296
T_{max}	0.5451	0.9984	0.9998	0.9931	1.0000	1.0000	0.0210
RH_{min}	0.0138	0.9946	0.9994	0.9724	1.0000	1.0000	0.1037
RH_{mean}	0.0138	0.9998	1.0000	0.9767	1.0000	1.0000	0.0992
RH_{max}	0.0138	0.9952	0.9995	0.9749	1.0000	1.0000	0.0954
$Tdew_{min}$	0.0138	0.9946	0.9995	0.9738	1.0000	1.0000	0.0984
$Tdew_{mean}$	0.0138	0.9998	1.0000	0.9792	1.0000	1.0000	0.0915
$Tdew_{max}$	0.0138	0.9949	0.9995	0.9745	1.0000	1.0000	0.0965
P_{min}	0.2634	0.9985	0.9998	0.9907	1.0000	1.0000	0.0563
P_{mean}	0.2634	0.9984	0.9998	0.9902	1.0000	1.0000	0.0600
P_{max}	0.2634	0.9985	0.9998	0.9908	1.0000	1.0000	0.0557
u_{dir}	0.0000	0.9997	1.0000	0.9790	1.0000	1.0000	0.1002
u_2	0.0000	0.9995	1.0000	0.9786	1.0000	1.0000	0.1001
u_{burst}	0.0000	0.9950	0.9995	0.9747	1.0000	1.0000	0.1035
R_s	0.1468	1.0000	1.0000	0.9832	1.0000	1.0000	0.0891
pr	0.1566	0.8891	1.0000	0.9260	1.0000	1.0000	0.1473

it can be said that generally when the P_{mean} sensor fails, the P_{min} sensor also fails, for example. Note that the same is true for temperature, wind, dew point, and humidity sensors.

3.3 Availability of environmental indexes

This section evaluates the availability of the evapotranspiration index (A_{ET_o}) using the availability data from each station and sensor. For this analysis were considered the FAO-56 Penman-Monteith, Hargreaves, and Turc methods. Based on Eqs. (2) to (5), the Table 6 presents a summary of the availability obtained in this analysis.

The index estimation is highly impacted by the unavailability of data. In average, the less impacted method (Turc's method) was unavailable during about 62 days in 2017, and the most impacted (FAO-56 Penman-Monteith method) cannot be calculated during about 103 days. Such unavailability could cause strong losses to the different actors that depend on these indexes like farmers, watershed managers, researches, and so on. For a crop depending on evapotranspiration data for daily irrigation, for example, the continuous unavailability of the index during 80 days could promote losses on water and yields and even huge losses in the case of short-cycle crops such as small vegetables, roots, and *leguminosae* (Allen et al., 1998).

4 Discussion

The analysis of the availability of the weather stations carried out in this work presented a considerable amount of stations with failures during the period of one year. Considering the set of 12 stations with the lowest availability, one can observe that 4 stations are in the state of Acre and 2 of them were the less available stations during 2017: Feijó and Rio Branco. Other 3 stations in this ranking are at the Brazilian's North region (Apui, Ariquemes, and Campos Lindos), which confirms our previous observation about this region. Most of these stations are located out of the capital cities of each state, only 2 of them are at capitals (Rio Branco e Recife). In addition, from Fig. 2 we noticed that the distribution of stations with low availability by states is concentrated in the states of Roraima and Amapá due to the small number of stations present in those states.

The analysis of the sensors availability shows that at least 25% sensors are fully available, but there are occurrences of sensors that do not operate for a whole year. For sensors related to wind data it is possible to verify this occurrence. Sensors of relative air humidity and temperature of dew point also presented low availability. In these cases, given the availability, the sensors were available for only five days (121 hours). The solar radiation sensors in 81% of the cases present fully availability. But there is an occurrence where the availability is 54 days, representing the lowest availability among the radiation sensors. In general, it is also possible to verify that the availability of the sensors is strongly correlated. Measurements of temperature, relative humidity and dew point temperature, for example, showed a strong correlation with each other.

The analysis performed to evaluate the availability of environmental indexes verified that the higher the number of parameters the method presents lower

Table 6: Availability of the evapotranspiration methods

Method	Min	1st Qu.	Median	Mean	3rd Qu.	Max	SD
FAO-56	0.0000	0.5054	0.8310	0.7164	0.9936	1.0000	0.3029
Hargreaves	0.0130	0.6490	0.9149	0.7887	0.9990	1.0000	0.2647
Turc	0.0518	0.7341	0.9391	0.8303	0.9995	1.0000	0.2250

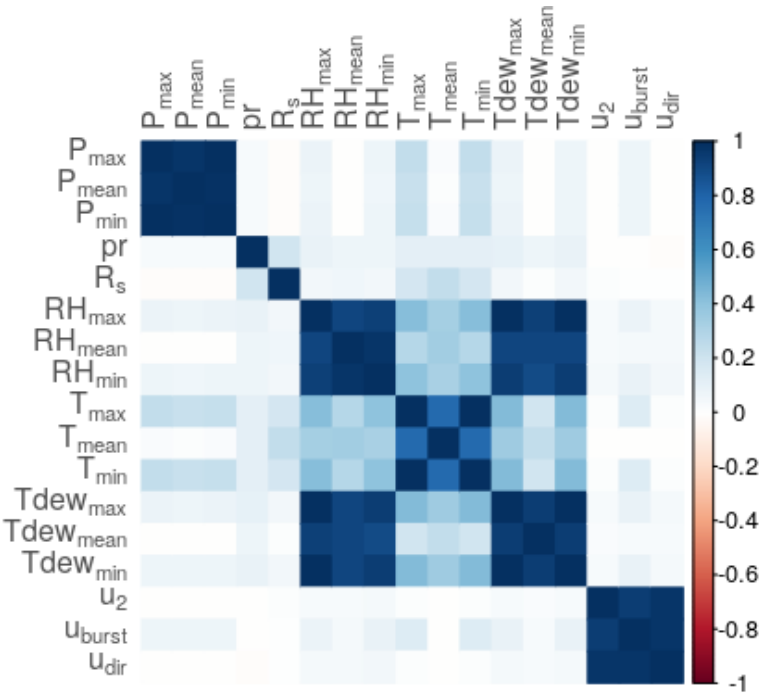

Figure 3: Correlation matrix of the availability of all sensors studied

the availability in its standard form. The FAO-56 Penman-Monteith method, because it requires many parameters, was the least available among the other methods. Moreover, the method requires data from the wind speed sensor which presented a low mean availability (see Table 5) and, for some stations, this sensor was unavailable throughout the year which prevented the estimation of ET_o through the FAO-56 Penman-Monteith method in those stations.

Analyzing the overall yearly availability of Turc's and Hargreaves' methods, the better availability can be attributed mainly to the smaller set of sensor that these methods depends on, but also to the availability of the sensors in this set. Whereas Hargreaves' method depends on the sensors related to temperature only (T_{min}, T_{mean}, and T_{max}), the Turc's method is dependent on T_{mean} and R_s sensors, and all those sensors are highly available (less than 25% of the cited sensors have an availability below two 9's).

5 Conclusion and further work

This work analyzes the availability of the network of automatic weather stations of the INMET. The analysis was performed considering the 490 weather stations distributed throughout the states of Brazil, verifying the availability of these stations, their respective sensors, service availability, and an analysis of the correlation between failures.

We observed that 59% of the network of automatic weather stations presents availability less than two 9's and only 21% is totally available. This analysis also presented the spatial distribution of failures and a notorious concentration of lower available stations in the North region of Brazil. We also observed that the T_{mean} sensors are those that present higher average availability, whereas the *pr* sensors present the lowest average. This same conclusion can be extended to the analysis of the weather data service and from this analysis, the availability of evapotranspiration as an environmental indices service was verified, presenting greater availability from the Hargreaves method.

As future work, we intend to expand the availability analysis considering previous years. In addition, we intend to carry out a study of the problem of bad measurements in the network of automatic weather stations of the INMET, trying to identify when a sensor drifts from its original calibration and starts to provide inconsistent data.

Acknowledgements

The authors would like to thank to Foundation for the Support of Science and Technology of the State of Pernambuco (FACEPE) (IBPG-0188-1.03/18) and National Council for Scientific and Technological Development (CNPq) (Process 424979/2016-0) for their financial support.

References

Allen, R. G., Pereira, L. S., Raes, D., Smith, M. and others (1998). Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56, *FAO, Rome*

300(9): D05109.

Arjannikov, T., Diemert, S., Ganti, S., Lampman, C. and Wiebe, E. C. (2017). Using markov chains to model sensor network reliability, *Proceedings of the 12th International Conference on Availability, Reliability and Security*, ARES '17, ACM, New York, NY, USA, pp. 6:1–6:10. http://doi.acm.org/10.1145/3098954.3098979.

Fisher, D. K. and Pringle III, H. (2013). Evaluation of alternative methods for estimating reference evapotranspiration, *Agricultural Sciences* **4**(08): 51. http://www.scirp.org/journal/PaperInformation.aspx?PaperID=36456.

INMET (2011). Rede de estações meteorológicas automáticas do inmet, *Technical report*, Instituto Nacional de Meteorologia. Available at https://bit.ly/2Z5jny9.

Jensen, M. E. and Allen, R. G. (2016). Evaporation, evapotranspiration, and irrigation water requirements, American Society of Civil Engineers.

Moura, C. R. W., Zolnier, S. A., Ribeiro, A. and Oliveira, R. A. D. (2010). Evaluation of methods of estimation of hourly evapotranspiration for hydroponic lettuce under protected environment, *Engenharia Agrícola* **30**: 646 – 656. https://doi.org/10.1590/S0100-69162010000400009.

Rausand, M. and Arnljot, H. (2004). *System reliability theory: models, statistical methods, and applications*, Vol. 396, John Wiley & Sons.

von Arx, G., Dobbertin, M. and Rebetez, M. (2013). Detecting and correcting sensor drifts in long-term weather data, *Environmental Monitoring and Assessment* **185**(6): 4483–4489. https://doi.org/10.1007/s10661-012-2831-6.

Xavier, F. and Brochado, M. L. C. (2017). Use of spatial visualization for pattern discovery in evapotranspiration estimation., *GEOINFO*, pp. 346–356.