

Article

## Análise Comparativa Entre Medidas Meteorológicas da Estação Convencional e Automática da Fazenda Lageado no Município de Botucatu, São Paulo, Brasil

José Rafael Franco<sup>1</sup> , Matheus Rodrigues Ranieiro<sup>1</sup>, Marcus Vinícius Contes Calça<sup>1</sup>, Sérgio Augusto Rodrigues<sup>1</sup>, Alexandre Dal Pai<sup>1</sup>, Enzo Dal Pai<sup>2</sup>

<sup>1</sup>*Departamento de Bioprocessos e Biotecnologia, Faculdade de Ciências Agrônomicas de Botucatu, Universidade Estadual Paulista “Júlio de Mesquita Filho”, Botucatu, SP, Brasil.*

<sup>2</sup>*Departamento de Engenharia Rural e Socioeconomia, Faculdade de Ciências Agrônomicas de Botucatu, Universidade Estadual Paulista “Júlio de Mesquita Filho”, Botucatu, SP, Brasil.*

Recebido em: 31 de Março de 2022 - Aceito em: 13 de Maio de 2022

### Resumo

O processo de desativação de uma estação meteorológica convencional (EMC) e sua substituição por uma automática (EMA) deve ser feito após a comparação de medições concomitantes entre ambas. Portanto, o objetivo deste estudo foi comparar medidas de 2015 e 2016 da EMC e EMA da Faculdade de Ciências Agrônomicas (UNESP) de Botucatu (SP) - Brasil, no intuito de avaliar se existe diferença nas medidas e criar modelos de regressão linear para construção de uma série de medidas única entre as estações. Foram comparadas medidas de temperatura máxima, mínima e média do ar, umidade relativa média do ar, velocidade do vento e precipitação, utilizando o método de reamostragem (Bootstrap) e de validação cruzada (K-Fold), bem como os indicadores estatísticos de índice de concordância, erro médio de viés, erro médio de viés relativo, raiz do erro médio quadrático, raiz do erro médio quadrático relativo, coeficiente de correlação e de determinação ( $R^2$ ). A maior diferença foi encontrada na velocidade do vento ( $R^2 = 0,746$ ) e a menor na precipitação ( $R^2 = 0,984$ ). Os resultados demonstram que apesar da diferença na composição dos instrumentos existe concordância entre os elementos, tornando segura a desativação da EMC e permitindo estimar as medidas da EMA para reconstrução de uma série histórica única.

**Palavras-chave:** comparação, validação, homogeneização.

## Comparative Analysis Between Meteorological Measurements of the Conventional and Automatic Station of Fazenda Lageado in the Municipality of Botucatu, São Paulo, Brazil

### Abstract

The disable process of a conventional meteorological station (EMC) and replacing it with an automatic one (EMA) must be done after comparing concomitant measurements between both. Therefore, the objective of this study was to compare 2015 and 2016 measurements of the EMC and EMA of the School of Agricultural Sciences (UNESP) of Botucatu (SP) - Brazil to evaluate if there is a difference in measurements and create a linear regression models for the construction of a single measurement serie between stations. Measures of maximum, minimum and average air temperature, average relative humidity, wind speed and precipitation were compared using the resampling (Bootstrap) and cross-validation (K-Fold) method, as well as the statistical indicators of index of concordance, mean bias error, relative mean bias error, root mean square error, relative root mean square error, correlation and determination coefficient ( $R^2$ ). The biggest difference was found in wind speed ( $R^2 = 0.746$ ) and the smallest in precipitation ( $R^2 = 0.984$ ). The results demonstrate that despite the difference in the composition of the instruments, there is agreement between the elements, making it safe to disable the EMC and allowing the estimation of the EMA measurements for the reconstruction of a single historical series.

**Keywords:** comparison, validation, homogenization.

## 1. Introdução

O estudo sobre o clima é de extrema importância para a sociedade. Para tanto, torna-se necessário medidas meteorológicas procedentes de fontes confiáveis. Esse tipo de medição, sobre os elementos do clima como temperatura do ar, umidade relativa do ar, velocidade e direção do vento, precipitação entre outros, é realizada em superfície terrestre através de estações meteorológicas, que podem ser classificadas como de origem convencional/mecânica (EMC) ou automática/eletrônica (EMA). Uma EMC exige a presença diária de um observador meteorológico para coletar as medidas dos instrumentos, já a EMA opera com sensores eletrônicos que permitem o registro das medições automaticamente, por meio de um sistema de aquisição de dados (datalogger), possibilitando a programação, armazenamento, processamento e transmissão de informações (Galina e Verona, 2004).

Com o avanço da tecnologia, muitos instrumentos meteorológicos convencionais estão entrando em desuso, devido a sua complexidade de operação e a necessidade de se ter um observador meteorológico atuando diariamente para registrar as medidas. A falta de equipamentos para substituição é outro problema, muitos instrumentos convencionais já não são mais fabricados, forçando a migração para sensores eletrônicos automáticos. A utilização de sensores automáticos apresenta diversas vantagens, entre elas: registro de medidas de forma automática em áreas remotas de difícil acesso, sensores com maior exatidão nas medidas, maior número de registros de medidas em um determinado período e menor tempo de resposta entre os registros, assim como a transmissão de dados em tempo real (Strassburger *et al.*, 2010; Almeida e Hermenegidio, 2013).

O acesso à informação meteorológica em tempo real possibilita uma variedade de aplicações que facilita e agiliza tomadas de decisões em diferentes campos da agricultura como: manejo de irrigação, aplicação de defensivos agrícolas, risco de incêndio, previsão de geadas, ocorrência de pragas e de doenças, classificação climática e zoneamento agrícola (Ribeiro *et al.*, 2017; Carvalho *et al.*, 2019).

Os sensores eletrônicos estão cada vez menores (compactos) e apresentando custos mais acessíveis, atualmente, de forma que torna ainda mais viável sua aquisição e implantação em novas áreas para coleta de medidas meteorológicas e substituição de instrumentos convencionais (Strassburger *et al.*, 2010; Almeida e Hermenegidio, 2013).

Segundo Sentelhas *et al.* (1997) surge uma questão importante ao substituir uma EMC por uma EMA, referente à homogeneização da série histórica de dados medidos de forma convencional com a nova série medida de forma automática. Portanto, esforços devem ser feitos para transferir a confiança de instrumentos convencionais para

sensores eletrônicos, com diferentes composições e números de amostragem. Além do processo de substituição dos instrumentos as duas séries devem ser unificadas em uma única série histórica, com início nos dados convencionais (EMC) e concatenada com os dados automáticos (EMA) a partir da data de desativação da EMC.

O processo de substituição de uma EMC por uma EMA não é um ato simples de instalar os novos equipamentos automáticos e desativar os instrumentos convencionais (Almeida e Hermenegidio, 2013).

Recomenda-se que o processo de substituição de um sistema por outro deva ser realizado após coleta de medidas simultâneas entre as estações meteorológicas e verificar se existe mudança significativa entre as medidas, após esse estudo é identificado se é seguro ou não desativar a EMC. O processo de análise e comparação das medidas deve ser realizado para cada localidade, no período mínimo de 1 ano de coleta de dados (Sentelhas *et al.*, 1997, Strassburger *et al.*, 2010).

Portanto, o objetivo deste estudo foi realizar uma análise comparativa entre as medidas diárias de precipitação (Prec), temperatura máxima (Tmax), média (Tmed) e mínima do ar (Tmin), umidade relativa média do ar (UR) e velocidade média do vento a 2 m de altura ( $U_2$ ) a fim de verificar se é seguro desativar a EMC e quando possível criar equações para estimar as medidas da EMA a partir de medições realizadas pela EMC para reconstruir a série histórica entre as estações meteorológicas.

## 2. Material e Métodos

### 2.1. Localização e clima

O estudo foi realizado na EMC e EMA (22°54' S, 48°27' O e 786 m) do Departamento de Engenharia Rural e Socioeconomia da Faculdade de Ciências Agrônomicas (UNESP) de Botucatu - São Paulo - Brasil. Botucatu é uma cidade semi-rural localizada na região centro-sul do estado de São Paulo e a aproximadamente 221 km do oceano Atlântico. Com bioma Cerrado e Mata Atlântica, o município possui área territorial de 1.482,682 km<sup>2</sup> e população de 148 mil habitantes, bem como é localizado entre duas extensas bacias (Barra Bonita e Jurumirim), com sua principal atividade sendo a agropecuária, o cultivo de cana-de-açúcar e eucalipto (IBGE, 2021; Dal Pai *et al.*, 2016; Rossi *et al.*, 2018).

Seguindo os critérios de Köppen, o clima de Botucatu apresenta classificação Cwa, mesotérmico (subtropical e temperado) com verões quentes e chuvosos e invernos frios e secos (Rossi *et al.*, 2018; SMAABC, 2002). A série histórica registrada pela EMC da Faculdade de Ciências Agrônomicas (UNESP) de Botucatu no período de 1971 a 2016 apresentou os seguintes valores: temperatura média do ar de 21,04 °C, umidade relativa média de 69,8% e precipitação acumulada anual média de

1521,76 mm. O mês mais quente ocorre em fevereiro, com temperatura máxima de 28,59 °C, o mês mais chuvoso ocorre em janeiro, com precipitação acumulada média de 286,60 mm e umidade relativa média de 75,79%. O período mais frio ocorre entre junho e julho, com temperatura média de 17,79 °C. O mês mais seco ocorre em agosto, com precipitação acumulada média de 38,02 mm e umidade relativa do ar média de 60,79%.

### 2.2. Instrumentos e medidas

Na **Tabela 1** são apresentados os instrumentos utilizados pela EMC e EMA, seus elementos sensíveis, fabricante, resolução e unidade de medida para mensurar os elementos do clima.

Dos instrumentos presentes na EMC apenas o termohigrógrafo realiza medida em escala horária, os demais instrumentos realizam medições em escala diária, por esta razão foi adotada a escala diária para comparação de todos os elementos.

As medidas coletadas na EMC foram registradas por meio de um observador meteorológico diariamente. Os dados da EMA foram coletados de forma automática, utilizando um datalogger CR1000 da Campbell Scientific, com coleta de dados a cada 5 min.

As temperaturas máxima e mínima do ar na EMC foram registradas pelos termômetros de máxima e mínima, a temperatura média do ar foi calculada utilizando as temperaturas extremas do dia pelo método da FAO (Medeiros *et al.*, 2012).

Para realizar a comparação dos dados de precipitação foram utilizadas as medidas diárias da EMC e horárias da EMA das 9:00 h até as 9:00 h (valor acumulado) do outro dia (Pereira *et al.*, 2008; Oliveira *et al.*, 2010).

A **Tabela 2** apresenta o período da aquisição de dados de cada instrumento e a composição das medidas.

**Tabela 2** - Período amostral e cálculo para obtenção dos elementos do clima pela EMC e EMA.

Elemento	EMC	EMA
Tmax	> valor em 24 h	> valor em 24 h
Tmin	< valor em 24 h	< valor em 24 h
Tmed	$T_{med} = (T_{max} + T_{min})/2$	$\sum T_{Med_j}/288^{**}$
UR	$\sum UR_i/24^*$	$\sum UR_j/288^{**}$
Prec	Leitura às 9:00 h	Somatória das medidas coletadas das 9:00 h do dia n às 9:00 h do dia n + 1
U <sub>2</sub>	Média diária	$\sum U_{2j}/288^{**}$

\*Medida obtida a cada 1 h (i = 01:00, 02:00, ... 24:00)

\*\*Medida obtida a cada 5 min (j = 0:00, 00:05, ... 23:55).

### 2.3. Dados analisados

A EMC teve início em 1971 até 2016, sua interrupção na coleta de dados foi causada por defeitos apresentados nos instrumentos, não havendo equipamentos para substituição, pois alguns modelos não são mais fabricados. O funcionamento da EMA teve início no final de 2014 e se mantém em operação até os dias atuais.

Para realização do estudo foi necessário medidas registradas simultaneamente pela EMC e EMA no mesmo período, devido a interrupção da EMC o período de dados analisado foi de 01/01/2015 a 31/12/2016 na resolução temporal diária. Os elementos meteorológicos analisados foram: temperatura máxima do ar (°C); temperatura mínima do ar (°C); temperatura média do ar (°C); umidade relativa média do ar (%); precipitação (mm) e velocidade média do vento a 2 m de altura (m s<sup>-1</sup>). Para garantia da qualidade das medidas os dados passaram por um processo de filtragem.

O filtro foi utilizado para eliminar medidas incorretas presentes na série de dados, preservando a qualidade e confiabilidade das informações. As medidas eliminadas pelo filtro não foram utilizadas neste estudo.

**Tabela 1** - Instrumentos utilizados para mensurar os elementos do clima na EMC e EMA: fabricante, sensibilidade/resolução e unidade de medida.

Estação	Variável	Elemento/ sensor	Fabricante	Sensibilidade/resolução	Unidade de medida
EMC	Tmax	Mercúrio	R. FUESS	0,2 °C	°C
	Tmin	Álcool	R. FUESS	0,2 °C	°C
	Tmed	—	—	—	°C
	UR	Harpa de Cabelo	Salmoiraghi	5%	%
	Prec	Pluviômetro Ville de Paris	IR	0,1 mm	mm
	U <sub>2</sub>	Conjunto de 3 Canecas	Lambrecht	0,5 m s <sup>-1</sup>	km h <sup>-1</sup>
EMA	Tmax	Termistor	Campbell Scientific	0,1 °C	°C
	Tmin	Termistor	Campbell Scientific	0,1 °C	°C
	Tmed	Termistor	Campbell Scientific	0,1 °C	°C
	UR	Capacitor	Campbell Scientific	0,8%	%
	Prec	Báscula	Campbell Scientific	2% no intervalo de operação	mm
	U <sub>2</sub>	Conjunto de 3 Canecas	Campbell Scientific	0,12 m s <sup>-1</sup>	m s <sup>-1</sup>

O método utilizado para filtragem foi o desvio relativo percentual Eq. (1), que permitiu identificar a grandeza da diferença entre os dados da amostra primária (medidas convencionais) com relação aos dados da amostra secundária (medidas automáticas) (Arenales e Darezzo, 2008).

$$DRi = \sum_{i=1}^n \left( \frac{|Y' - Y_i|}{|Y_i|} \right) 100 \quad (1)$$

onde  $Y'$  representa as medidas da EMA,  $Y_i$  as medidas da EMC e  $i$  representa o dia que a medida foi registrada.

Foram definidos critérios de cortes para eliminar as medidas que ultrapassaram os limites superior e inferior para cada variável. Os valores limites são definidos em porcentagem, indicando o intervalo das medidas válidas. Para identificação do valor de corte foi gerado um gráfico de dispersão das medidas utilizando o Microsoft Excel 365, onde foi feita uma observação dos pontos pelo método de análise gráfica exploratória, ajustando a dispersão entre os pontos de medidas com a eliminação de valores extremos de DR.

Para realizar a reconstrução das bases de dados foi utilizado o método de equação linear simples Eq. (2), permitindo estimar os dados da EMA em um período anterior a sua instalação por meio das medidas da EMC. Criando desta forma uma série de dados com início em 1971 mensurada com instrumentos convencionais e adicionada a série automática em 2015.

$$y = bx + a \quad (2)$$

onde  $y$  são as medidas da EMA estimadas e  $x$  as medidas registradas pelos instrumentos da EMC.

Devido à falta de mais anos de medidas para garantir uma maior confiança na análise estatística foi utilizado o método de reamostragem (Bootstrap), que consiste na reutilização de um mesmo conjunto de dados por várias vezes sorteados de forma aleatória. A reamostragem permite criar diferentes alternativas para testes, obtendo o desvio padrão e o intervalo de confiança em percentil dos dados analisados. Neste método é recomendado a reamostragem com utilização de 1000 repetições. A utilização da reamostragem apresenta uma boa opção para realizar vários testes de comparação em cenários diferentes, utilizando a mesma série de dados (Hesterberg *et al.*, 2003).

A validação cruzada foi o método utilizado em combinação com o método de reamostragem para avaliar o desempenho do modelo em diferentes situações amostral, dividindo os dados em dois grupos, uma parte para criação do modelo e outra parte para validação. No método K-Fold de validação cruzada, os dados foram separados em blocos, e em cada bloco foram divididos aleatoriamente entre a criação do modelo e os dados de treinamento (Rohani *et al.*, 2018).

A combinação dos métodos de reamostragem e validação cruzada permitiram a criação de um laço com 1000 repetições, que em cada passagens do laço os 2 anos de medidas registradas são sorteadas aleatoriamente e separados entre 70% das medidas para realizar a comparação entre as estações e gerar os coeficientes interceptor, angular e coeficiente de determinação ( $R^2$ ) e os 30% das medidas restantes foram aplicados na equação gerada (utilizando o valor do interceptor e angular encontrado) para estimar os dados da EMA e comparar com os dados medidos da EMC, avaliando a capacidade do modelo em fazer previsões.

Os valores de saída do laço de repetição foram a média e desvio padrão dos indicativos e o intervalo de confiança em percentil dos coeficientes da equação.

## 2.4. Indicativos estatísticos

Os indicativos estatísticos foram utilizados para avaliar duas situações. A primeira foi para comparar os dados obtidos de forma convencional (EMC) e automática (EMA). A segunda foi analisar as medidas registradas pela EMA com as medidas estimadas pela equação utilizando as medidas da EMC.

Para a comparar os dados foi realizado a análise da regressão e visualizações gráficas. Foram utilizados os coeficientes de determinação ( $R^2$ ) e correlação ( $r$ ) e os seguintes indicativos estatísticos: índice de concordância de Willmott ( $d$ ); Mean Bias Error ( $MBE$ ) e relativo ( $rMBE$ ); Root Mean Square Error ( $RMSE$ ) e relativo ( $rRMSE$ ).

O índice de concordância  $d$  Eq. (3), descrito por Willmott *et al.* (1985), tem seus valores que variam entre 0 (nenhuma concordância) e 1 (concordância perfeita). O  $MBE$  Eq. (4) quantifica a diferença entre os pares de dados, comparando-os a fim de identificar uma tendência significativa em superestimar ou subestimar sistematicamente os valores. O  $RMSE$  Eq. (6) estima o erro absoluto indicado pelas diferenças entre os pares de dados avaliados, dando um maior peso para grandes erros.

Para os indicativos  $MBE$  e  $RMSE$  valores próximos de zero são desejáveis. Os valores relativos  $rMBE$  Eq. (5) e  $rRMSE$  Eq. (7) fornecem o grau de correspondência percentual proporcional à média da observação. O  $r$  Eq. (8) fornece o grau de correlação entre os pares de dados comparados e  $R^2$  Eq. (9) informa quanto a variável preditora (EMC) consegue explicar a variabilidade da variável resposta (EMA) (Raniero *et al.*, 2020; Rincón *et al.*, 2018).

$$d = 1 - \frac{\sum_{i=1}^n (Y'_i - Y_i)^2}{\sum_{i=1}^n (|Y' - Y| + |Y_i - Y|)^2} \quad (3)$$

$$MBE = \frac{\sum_{i=1}^n (Y'_i - Y_i)}{n} \quad (4)$$

$$rMBE(\%) = \frac{MBE}{\bar{Y}} \times 100 \tag{5}$$

$$RMSE = \left[ \frac{\sum_{i=1}^n (Y'_i - Y_i)^2}{n} \right]^{\frac{1}{2}} \tag{6}$$

$$rRMSE(\%) = \frac{RMSE}{\bar{Y}} \times 100 \tag{7}$$

$$r = \frac{\sum_{i=1}^n (Y'_i - \bar{Y}') (Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (Y'_i - \bar{Y}')^2 \sum_{i=1}^n (Y_i - \bar{Y})^2}} \tag{8}$$

$$R^2 = \frac{\left[ \sum_{i=1}^n (Y'_i - \bar{Y}') (Y_i - \bar{Y}) \right]^2}{\sum_{i=1}^n (Y'_i - \bar{Y}')^2 \sum_{i=1}^n (Y_i - \bar{Y})^2} \tag{9}$$

Na etapa de comparação das medidas entre a EMC e EMA,  $Y'$  representa as medidas da EMC,  $Y$  as medidas de referência (EMA), e  $n$  o número de observações,  $\bar{Y}'$  é a média das medidas da EMC e  $\bar{Y}$  a média das medidas de referência (EMA).

Na avaliação de desempenho dos modelos de estimativa,  $Y'$  representa as medidas estimadas pela equação utilizando as medidas da EMC,  $Y$  os dados de referência (EMA), e  $n$  o número de observações,  $\bar{Y}'$  é a média das medidas dos valores estimados pela equação e  $\bar{Y}$  a média das medidas de referência (EMA).

Todo o processamento de dados foi realizado utilizando a linguagem de programação R. Uma linguagem de

programação estatística, disponível como software livre (R Core Team, 2021).

### 3. Resultados e Discussão

#### 3.1. Filtro de dados

A aplicação do filtro de  $DR_i$  permitiu eliminar as medidas incorretas registradas, não sendo utilizadas na comparação. A Tabela 3 apresenta o valor total de medidas de cada elemento e a quantidade de medidas eliminadas pelo filtro. O elemento com maior número de registros eliminados foi a Tmin com 28,01%, os erros ocorreram devido a falhas no termômetro de mínima, que utiliza álcool como elemento sensível, diferente do termômetro da Tmax, que teve a menor quantidade de dados eliminados, com 5,15% das medidas, que utiliza o mercúrio como elemento sensível. Os erros registrados na Tmin interferiram no cálculo da Tmed, que foi a segunda variável que teve mais dados eliminados pelo filtro.

#### 3.2. Comparação da EMC x EMA

A Tabela 4 apresenta os resultados do laço de 1000 repetições onde foram gerados os valores da equação para

**Tabela 3** - Aplicação do filtro Desvio Relativo Percentual (DR) nos elementos analisados no período de 01/01/2015 a 31/12/2016 registrados pela EMC e EMA.

Variável	Total	Ponto de corte	Dados eliminados	Restante
Tmin	707	18%	198	509
Tmed	703	13%	142	561
U <sub>2</sub>	706	66%	121	585
UR	587	18%	50	537
Prec	208	74%	11	197
Tmax	718	14%	37	681

**Tabela 4** - Média e desvio padrão dos coeficientes das equações.

Elemento	Coefficiente	Estimativa Bootstrap Média (desvio padrão)	IC (95%) percentil	R <sup>2</sup>
Prec	Intercepto (a)	0,539 (0,132)	[0,2998;0,8184]	0,984
	Inclinação (b)	0,977 (0,008)	[0,9577;0,9904]	
Tmax	Intercepto (a)	0,578 (0,132)	[0,3137; 0,8279]	0,961
	Inclinação (b)	0,979 (0,005)	[0,9706; 0,9890]	
Tmin	Intercepto (a)	2,509 (0,178)	[2,1755; 2,8480]	0,844
	Inclinação (b)	0,904 (0,012)	[0,8811; 0,9277]	
Tmed	Intercepto (a)	0,641 (0,153)	[0,3451;0,9363]	0,917
	Inclinação (b)	0,970 (0,007)	[0,9557; 0,9850]	
UR	Intercepto (a)	5,376 (0,945)	[3,4685;7,1137]	0,811
	Inclinação (b)	0,928 (0,013)	[0,9046; 0,9537]	
U2	Intercepto (a)	0,670 (0,016)	[0,6392;0,7002]	0,746
	Inclinação (b)	1,161 (0,017)	[1,1298;1,1946]	

estimar as medidas da EMC, o intervalo em percentil e o valor do coeficiente de determinação  $R^2$  da comparação entre a EMC e EMA.

### 3.2.1. Precipitação

Para efetuar a comparação, os dias que não ocorreram precipitação foram desconsiderados. As medidas coletadas pelas estações apresentaram valores próximos em precipitações menores que 40 mm. Os valores também foram similares para dias em que ocorreu precipitação intensa, próxima a 140 mm. O valor acumulado na série completa foi de 3593,15 mm registrados na EMC e 3618,79 mm pela EMA. A Fig. 1a indica a associação entre precipitações (Prec) obtidas pela EMC e EMA. O resultado apresentou forte correlação positiva, com pouca dispersão dos pontos. A forte correlação também pode estar associada a pouca quantidade de dias chuvosos no período (197 dias).

### 3.2.2. Temperatura máxima

A temperatura máxima (Tmax) apresentou valores semelhantes entre as estações meteorológicas em temperaturas elevadas (35 °C) e temperaturas amenas (15 °C). Os instrumentos apresentaram grande estabilidade nas medidas entre as estações. A Fig. 1b mostra a correlação entre as medidas diárias de temperatura máxima do ar obtidas pela EMC e EMA, apresentando forte correlação positiva com pouca dispersão.

### 3.2.3. Temperatura mínima

A temperatura mínima do ar (Tmin) apresentou variação durante o período analisado na comparação. No intervalo de janeiro a dezembro de 2015 as medidas da EMA superestimaram as medidas da EMC. No final de dezembro de 2015 até abril de 2016 o termômetro de mínima da EMC apresentou problemas, gerando medidas inconsistentes, que foram eliminadas no processo de filtragem. Após a substituição do termômetro em abril de 2016 as medidas apresentaram resultados próximos entre as estações. O valor médio da temperatura mínima do ar foi de 15,24 °C na EMC e 16,29 °C na EMA. A Fig. 1c exibe a correlação entre as medidas diárias da temperatura mínima do ar obtida pela EMA e EMC. O gráfico apresentou correlação positiva, com grande dispersão dos pontos. A temperatura mínima foi a medida entre as temperaturas que apresentou maior espalhamento dos pontos, causado pelo mal funcionamento do termômetro.

### 3.2.4. Temperatura média

A temperatura média (Tmed) foi estimada pela EMC utilizando apenas duas medidas extremas (temperatura máxima e mínima) do dia, diferente da EMA que realiza a média de todas as 288 medidas de temperatura registradas. O mal funcionamento do termômetro de mínima e a limitação do número de medidas pela EMC não apresentaram bons resultado no intervalo de dezembro de 2015 a abril de 2016. Nos demais intervalos os valores apresentaram

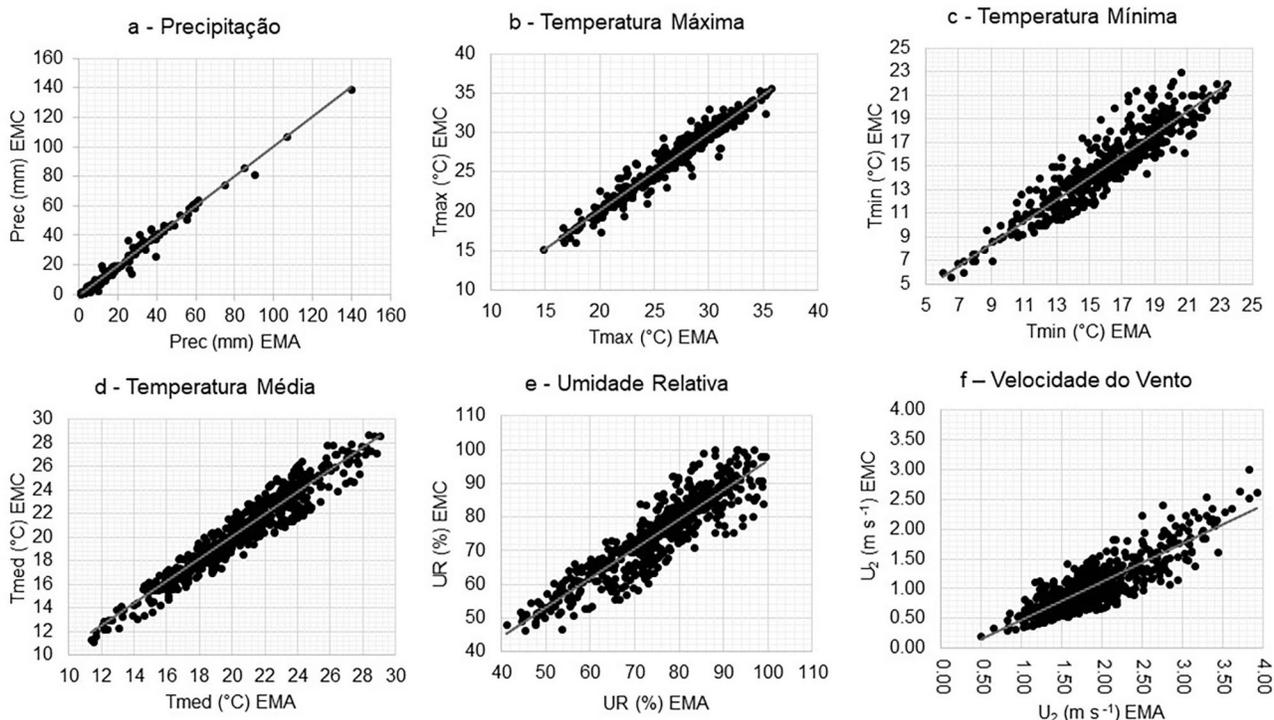


Figura 1 - Correlação entre a EMC e EMA para o Município de Botucatu, São Paulo, Brasil.

resultados próximos entre a EMC e EMA. A Temperatura média apresentou valor médio de 20,82 °C medido na EMC e 20,83 °C na EMA. A Fig. 1d exibe a correlação entre as medidas obtidas pela EMC e EMA. O gráfico apresentou uma correlação positiva, com maior espalhamento em comparação a temperatura máxima. Este espalhamento ocorreu por causa das falhas registradas na temperatura mínima da EMC.

### 3.2.5. Umidade relativa

A umidade relativa do ar (UR) apresentou maior diferença entre as estações no início de janeiro de 2015 a junho de 2015. Após esse período as medidas apresentaram maior exatidão. A EMC e EMA apresentaram o mesmo valor médio na série de 74,87%. A Fig. 1e exibe a correlação entre as medidas diárias da umidade relativa média do ar obtida pela EMA e EMC. O gráfico apresenta correlação positiva com grande espalhamento dos pontos. Esse espalhamento é causado pela diferente sensibilidade dos instrumentos, como especificado pelo fabricante. O higrógrafo da EMC possui uma sensibilidade (precisão nominal) de 5% e da EMA de 0,8%. Como o higrógrafo da EMC utiliza cabelo humano como elemento sensível, com o passar dos anos esse material pode apresentar problemas em sua capacidade de expandir-se e contrair-se com a variação da umidade relativa do ar, registrando medidas incorretas.

### 3.2.6. Velocidade do vento

A velocidade do vento a 2 m de altura ( $U_2$ ) obtidas pela EMA apresentaram valores maiores que as medidas registradas na EMC. O valor médio da velocidade do vento foi de 1,028 m s<sup>-1</sup> na EMC e 1,864 m s<sup>-1</sup> na EMA. A Fig. 1f exibe a correlação entre as medidas de velocidade média do vento a 2 m de altura obtidas pela EMA e EMC. O gráfico apresentou uma correlação positiva com agrupamento dos pontos para velocidades próximas a 1,5 m s<sup>-1</sup> e maior espalhamento com o aumento da velocidade do vento. Essa diferença ocorre pela sensibilidade na composição dos sensores. Como especificado pelo fabricante, o anemômetro da EMC possui uma sensibilidade (precisão nominal) de 0,5 m s<sup>-1</sup> e da EMA de 0,12 m s<sup>-1</sup>. Outro fator a ser considerado é a manutenção do instrumento convencional ao passar do tempo, referente a lubrificação do eixo que pode afetar a rotação das conchas e interferir nas medidas.

## 3.3. Análise dos indicadores estatísticos

Os indicadores estatísticos (Tabela 5) foram aplicados em toda a série de dados no processo de comparação, assim como foi realizado em outros trabalhos presentes na literatura e seus resultados foi comparado com este trabalho (Sentelhas *et al.*, 1997; Pereira *et al.*, 2008; Oliveira *et al.*, 2010; Ribeiro *et al.*, 2017; Carvalho *et al.*, 2019).

**Tabela 5** - Indicativos estatísticos da comparação de EMC e EMA.

Elemento	MBE	rMBE	RMSE	rRMSE	r	d
Prec	-0,130	-0,708	2,602	14,166	0,992	0,995
Tmax	-0,024	-0,091	0,780	2,909	0,980	0,990
Tmin	-1,049	-6,437	1,650	10,120	0,919	0,932
Tmed	-0,017	-0,080	0,971	4,659	0,958	0,978
UR	-0,003	-0,005	5,240	6,700	0,900	0,947
U2	-0,836	-44,860	0,891	47,785	0,864	0,605

### 3.3.1. Precipitação

As medidas de precipitação registradas pela EMA foram 0,708% menores que as medidas obtidas pela EMC, com diferença de 0,130 mm. A comparação apresentou 2,602 mm de erro absoluto ( $rRMSE = 14,166\%$ ) entre a EMC e EMA. Ribeiro *et al.* (2017), indicam subestimativa nas medidas registradas pela EMC em relação a EMA em alguns municípios no estado do Piauí. Resultados com maiores valores obtidos pela EMC em comparação a EMA foram encontrados por Sentelhas *et al.* (1997) e Oliveira *et al.* (2010). A precipitação foi o elemento que obteve o melhor desempenho de comparação ( $R^2 = 0,984$ ) e alto índice de concordância (0,995). Resultados semelhantes foram identificados por Sentelhas *et al.* (1997) no município de Piracicaba - SP, apresentando correlação de 0,988 ( $R^2$ ) e concordância de 0,993 ( $d$ ). Oliveira *et al.* (2010) também obtiveram resultados similares ( $R^2 = 0,995$  e  $d = 0,99$ ) no município de Jaboticabal - SP. Carvalho *et al.* (2019) encontraram uma correlação próxima ( $R^2 = 0,972$ ) e índice de concordância superior ( $d = 0,999$ ) na cidade de Teresinha - PI.

### 3.3.2. Temperatura máxima

O termômetro de temperatura máxima do ar, composto por mercúrio, apresentou boa precisão nas medidas diárias comparadas ao sensor eletrônico. As medidas coletadas pela EMA foram 0,091% menores que as medidas registradas pela EMC, com diferença de 0,024 °C. O erro absoluto foi de 0,780 °C ( $rRMSE = 2,909\%$ ). Resultados apresentados por Pereira *et al.* (2008) mostram que a EMC subestimou as medidas da EMA em 0,37%. Ribeiro *et al.* (2017) obtiveram valores maiores de temperatura máxima do ar medidos pela EMA. Sentelhas *et al.* (1997) e Oliveira *et al.* (2010) mostram que existe uma tendência dos dados coletados pela EMC serem maiores que as medidas registradas pela EMA. A Temperatura máxima do ar apresentou bom desempenho de 0,961 ( $R^2$ ) e elevado índice de concordância 0,990 ( $d$ ). Outros estudos apresentam resultados semelhantes realizados por Sentelhas *et al.* (1997), onde obtiveram boa correlação 0,987 ( $R^2$ ) e concordância 0,987 ( $d$ ). Oliveira *et al.* (2010) apresentaram valores próximos também, com correlação de 0,973 ( $R^2$ ) e concordância de 0,97 ( $d$ ). Carvalho *et al.* (2019) encontraram

valores de menor correlação ( $R^2$  0,945) e de concordância maior ( $d$  0,999).

### 3.3.3. Temperatura mínima

As medidas diárias de temperatura mínima do ar obtidas pela EMA foram 6,437% menores que as medidas obtidas pela EMC, com diferença de 1,049 °C. A comparação apresentou 1,650 °C de erro absoluto ( $rRMSE = 10,120\%$ ). Resultados semelhantes foram encontrados por [Sentelhas et al. \(1997\)](#), [Oliveira et al. \(2010\)](#), [Ribeiro et al. \(2017\)](#) e [Pereira et al. \(2008\)](#) com valores maiores de temperatura mínima do ar registrados pela EMC em comparação a EMA. A Temperatura mínima do ar apresentou correlação de 0,844 ( $R^2$ ) e alto índice de concordância ( $d = 0,932$ ), valores superiores foram encontrados por [Sentelhas et al. \(1997\)](#) com alta correlação ( $R^2 = 0,969$ ) e concordância ( $d = 0,991$ ). [Oliveira et al. \(2010\)](#) encontraram valores maiores de correlação ( $R^2 = 0,963$ ) e concordância ( $d = 0,98$ ). [Carvalho et al. \(2019\)](#) chegaram a valores mais próximos de Botucatu - SP, com correlação de 0,869 ( $R^2$ ) e concordância de 0,980 ( $d$ ).

### 3.3.4. Temperatura média

As medidas de temperatura média do ar registradas pela EMA foram 0,080% menores que as medidas obtidas pela EMC, com diferença de 0,017 °C. O erro absoluto foi de 0,971 °C ( $rRMSE = 4,659\%$ ). Resultados semelhantes foram encontrados por [Oliveira et al. \(2010\)](#) e [Sentelhas et al. \(1997\)](#), onde é apresentada uma tendência de valores maiores registrados pela EMC. [Ribeiro et al. \(2017\)](#) atingiram resultados maiores registrados pela EMA no estado do Piauí. A Temperatura média do ar apresentou bom resultado ( $R^2 = 0,917$ ) e alto índice de concordância ( $d = 0,978$ ). Trabalhos realizados por [Sentelhas et al. \(1997\)](#) obtiveram um valor maior de correlação ( $R^2 = 0,953$ ) e menor de concordância ( $d = 0,958$ ). [Oliveira et al. \(2010\)](#) apresentaram valor semelhante de correlação ( $R^2 = 0,968$ ) e maior valor de concordância ( $d = 0,99$ ). [Carvalho et al. \(2019\)](#) chegaram ao resultado mais próximo deste estudo, com correlação ( $R^2 = 0,903$ ) e concordância ( $d = 0,989$ ).

### 3.3.5. Umidade relativa

As medidas de umidade relativa registradas pela EMA foram 0,005% menores que as medidas da EMC, com diferença de 0,003%. A comparação apresentou 5,240% de erro absoluto ( $rRMSE = 6,700\%$ ). Estudos realizados por [Sentelhas et al. \(1997\)](#) e [Pereira et al. \(2008\)](#) encontraram valores elevados registrados na EMC em comparação a EMA. [Oliveira et al. \(2010\)](#) encontraram valores elevados de umidade relativa registrados pela EMA em comparação a EMC. [Ribeiro et al. \(2017\)](#) obteve os dois tipos de resultados em municípios do estado do

Piauí. A umidade relativa foi um modelo que gerou resultados intermediários de correlação ( $R^2 = 0,811$ ) e alta concordância ( $d = 0,947$ ), valores superiores foram encontrados por [Sentelhas et al. \(1997\)](#), onde obtiveram uma correlação ( $R^2 = 0,869$ ) e concordância ( $d = 0,960$ ). [Oliveira et al. \(2010\)](#) encontraram maior correlação ( $R^2 = 0,929$ ) e concordância ( $d = 0,94$ ). [Carvalho et al. \(2019\)](#) chegaram aos valores próximos de correlação ( $R^2 = 0,892$ ) e concordância ( $d = 0,944$ ).

### 3.3.6. Velocidade do vento

As medidas de velocidade do vento registradas na EMA apresentaram menores valores de 44,860% em relação as medidas registradas pela EMC, com diferença de 0,836 m s<sup>-1</sup>. O modelo apresentou 0,891 m s<sup>-1</sup> de erro absoluto ( $rRMSE = 47,785\%$ ). [Pereira et al. \(2008\)](#) obtiveram valores superestimados pela EMA de 5,4% em comparação a EMC. [Ribeiro et al. \(2017\)](#) informam que os maiores valores foram registrados pela EMA em apenas um município e nos outros cinco municípios analisados a EMC apresentou maiores valores. A velocidade média do vento apresentou correlação de 0,746 ( $R^2$ ) com menor desempenho em comparação as demais variáveis analisadas e alto índice de concordância de 0,922 ( $d$ ). Resultados encontrados por [Ribeiro et al. \(2017\)](#) obtiveram correlação de 0,71 ( $R^2$ ). [Carvalho et al. \(2019\)](#) apresentam resultados menores de correlação ( $R^2 = 0,329$ ) e alta concordância ( $d = 0,952$ ).

## 3.4. Estimativa de medidas da EMA com mediadas da EMC

No processo de comparação das medidas entre as estações foram gerados os coeficientes: interceptor ( $a$ ) e inclinação ( $b$ ), que permitiram a criação da equação para estimar as medidas da EMA com medidas da EMC.

Em seguida foi aplicada a equação para estimar os dados e comparado com medidas registradas pela EMA no período de 2015 a 2016, analisando a capacidade de fazer previsão das medidas. A [Tabela 6](#) apresenta os resultados dos indicadores estatísticos após o término das 1000 repetições.

Todas as variáveis analisadas apresentaram ótimo índice de concordância com baixos valores para  $MBE$  e  $rMBE$ . Os valores do desvio padrão foram próximos de zero, apresentando pouca variação do resultado médio obtido, os maiores valores do desvio padrão foram encontrados nos indicadores  $rMBE$  e  $rRMSE$  da precipitação.

As variáveis apresentaram bons resultados de correlação, os menores resultados foram encontrados na umidade relativa e velocidade do vento como já identificado na comparação entre a EMC e EMA. A umidade relativa apresentou maior valor de  $RMSE$  e velocidade do vento maior  $rRMSE$ . A precipitação foi a variável que apresen-

**Tabela 6** - Indicativos estatísticos de medidas da EMA com dados estimados pela equação utilizando medidas da EMC.

Elemento	<i>MBE</i>	<i>rMBE</i>	<i>RMSE</i>	<i>rRMSE</i>	<i>r</i>	<i>d</i>
Prec	-0,012 ± 0,415	-0,020 ± 2,288	2,529 ± 0,518	13,909 ± 2,892	0,991 ± 0,004	0,995 ± 0,002
Tmax	-0,001 ± 0,067	-0,001 ± 0,248	0,778 ± 0,061	2,900 ± 0,003	0,981 ± 0,003	0,990 ± 0,002
Tmin	-0,003 ± 0,115	-0,017 ± 0,704	1,235 ± 0,070	7,580 ± 0,431	0,920 ± 0,012	0,956 ± 0,006
Tmed	0,009 ± 0,085	0,042 ± 0,410	0,965 ± 0,045	4,637 ± 0,215	0,958 ± 0,005	0,978 ± 0,003
UR	-0,030 ± 0,446	-0,039 ± 0,595	5,202 ± 0,250	6,945 ± 0,332	0,900 ± 0,011	0,944 ± 0,006
U <sub>2</sub>	-0,002 ± 0,023	-0,102 ± 1,257	0,299 ± 0,013	16,052 ± 0,693	0,864 ± 0,017	0,922 ± 0,009

tou o segundo maior valor de *rRMSE* e maior desvio padrão dos indicativos.

Os resultados indicam que é seguro utilizar os coeficientes gerados para estimar os dados da EMA a partir de medidas da EMC. Permitindo a criação uma série de dados para a EMA com início em 1971 até os dias atuais.

#### 4. Conclusão

O estudo permitiu identificar as seguintes conclusões:

- Manter uma EMC em funcionamento exige um alto custo operacional, pois necessita de funcionários diariamente monitorando as medidas. Outro problema presente é a falta de instrumentos convencionais de referência para realizar a calibração e substituição em casos de falhas.
- Por causa de falhas apresentadas nos instrumentos da EMC que realizaram a coleta de medidas para este estudo não foi possível realizar uma análise com maior número de medidas.
- A filtragem das medidas coletadas pelo *DRi* permitiu um controle de qualidade sobre as medidas, eliminando dados errôneos relacionados ao mal funcionamento dos instrumentos no processo de aquisição, que também é sujeito a falha humana e eletrônica.
- A precipitação foi o elemento que apresentou a melhor correlação, no entanto, devido a utilização apenas de dias com ocorrência de chuva, foi o elemento com menor número de medidas. A temperatura máxima foi o elemento com maior número de medidas que apresentou melhor resultado. Seguido pela temperatura média que apresentou menor correlação devido a temperatura mínima, causada pelo mal funcionamento do termômetro de mínima da EMC. A umidade relativa apresentou menor valor de correlação devido a diferença de sensibilidade na composição dos instrumentos. A velocidade do vento foi o elemento que obteve a correlação mais baixa. Em vários estudos analisados a velocidade do vento não apresentou bons resultados.
- Todos os elementos apresentaram pouca diferença entre as medidas registradas pela EMC e EMA, apesar da diferença entre as tecnologias presentes em cada instrumento.

- O processo de comparação entre uma EMC e EMA deve ser realizado para cada localidade, verificando se é seguro desativar a EMC.
- Os métodos de reamostragem (Bootstrap) e validação cruzada (K-Fold) foram um diferencial nesta pesquisa, possibilitando a separação dos dados aleatoriamente em 70% para criação do modelo e 30% para validação. A utilização do laço de 1000 repetições apresentou bom resultado de desvio padrão referente a média para todos os elementos analisados.
- A criação de equações permitiu estimar os dados da EMA utilizando medidas da EMC, permitindo a criação de uma série de dados única e homogênea de 1971 até o início de operação da EMA em 2015, se mantendo em operação até os dias atuais.

#### Agradecimento

Os autores agradecem a Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela bolsa de estudos e ao Programa de Pós-Graduação em Agronomia - Engenharia Agrícola da Faculdade de Ciências Agrônomicas (UNESP) de Botucatu, SP, Brasil. A todos os observadores meteorológicos que trabalharam dia a dia durante todos esses anos e possibilitaram a realização desta pesquisa.

#### Referências

- ALMEIDA, H.A.; HERMENEGIDIO, G.M.S. Comparação de dados meteorológicos obtidos por estações meteorológicas convencional e automática. **Revista Brasileira de Climatologia**, v. 12, p. 32-47, 2013.
- ARENALES, S.; DAREZZO, A. **Cálculo Numérico: Aprendizagem com Apoio de Software**. São Paulo: Thomson Learning, 2008.
- CARVALHO, M.W.L.; BASTOS, E.A.; ANDRADE JUNIOR, A.S.; SENTELHAS, P.C. Comparação de dados meteorológicos e estimativa da radiação líquida e evapotranspiração de referência utilizando estações convencional e automática. **Agrometeoros**, v. 27, n. 2, p. 285-292, 2019.
- DAL PAI, A.; ESCOBEDO, J.F.; DAL PAI, E.; OLIVEIRA, A.P.; SOARES, J.; CODATO, G. MEO shadowring method for measuring diffuse solar irradiance: Corrections based on sky cover. **Renewable Energy**, v. 99, p. 754-763, 2016.

- GALINA, M.H.; VERONA, J.A. Fontes de observações meteorológicas no Estado de São Paulo. **Estudos Geográficos**, v. 2, n. 1, p. 107-118, 2004.
- HESTERBERG, T.; MONAGHAN, S.; MOORE, D.S.; CLIPSON, A.; EPSTEIN, R. **Bootstrap Methods and Permutation Tests**. New York: W. H. Freeman and Company, 2003.
- IBGE. Brasil/São Paulo/Botucatu. **Cidades: IBGE, 2019**. Disponível em: <https://cidades.ibge.gov.br/brasil/sp/botucatu/panorama>. Acesso em: 28 fev. 2021.
- MEDEIROS, R.M.; TAVARES, A.; KASSAR, C.B.; SILVA, J.A.S.; SILVA, V.P.R. Metodologia de cálculo da temperatura média do ar: Aplicação para os municípios de Parnaíba, Picos e Gilbuês, PI. **Revista Brasileira de Agricultura Irrigada**, v. 6, n. 4, p. 283-295, 2012.
- OLIVEIRA, A.D.; ALMEIDA, B.M.; CAVALCANTE JUNIOR, E.G.; ESPINOLA SOBRINHO, J.; VIEIRA, R.Y.M. Comparação de dados meteorológicos obtidos por estação Convencional e automática em Jaboticabal-SP. **Revista Caatinga**, v. 23, n. 4, p. 108-114, 2010.
- PEREIRA, L.M.P.; CARAMORI, P.H.; RICCE, W.S.; CAVI-GLIONE, J.H. Análise comparativa de dados meteorológicos obtidos por estação convencional e automática em Londrina - PR. **Ciências Agrárias**, v. 29, n. 2, p. 299-306, 2008.
- R CORE TEAM. **The R Project for Statistical Computing**. 2021. Disponível em <https://www.r-project.org/>. Acesso em 31 out. 2021.
- RANIERO, M.R.; CALCA, M.V. C.; FERNANDO, D.M.Z.; DAL PAI, A. Software for automated reading of sunshine duration by Digital Image Processing. **IEEE Latin America Transactions**, v. 18, n. 9, p. 1599-1605, 2020.
- RIBEIRO, A.A.; ANDRADE JÚNIOR, A.S.; SILVA, E.M.; SIMEÃO, M.; BASTOS E.A. Comparação entre dados meteorológicos obtidos por estações convencionais e automáticas no estado do Piauí, Brasil. **Irriga**, v. 22, n. 2, p. 220-235, 2017.
- RINCÓN, A.; ORIL, J.; FRUTOS, M.; ALVAREZ, L.; BARRIOS, F.P.; GONZÁLEZ, J.A. Bias correction of global irradiance modelled with weather and research forecasting model over Paraguay. **Solar Energy**, v. 170, p. 201-211, 2018.
- ROHANI, A.; TAKI, M.; ABDOLLAHPOUR, M. A novel soft computing model (Gaussian process regression with K-fold cross validation) for daily and monthly solar radiation forecasting (Part: I). **Renewable Energy**, v. 115, p. 411-422, 2018.
- ROSSI, T.J.; ESCOBEDO, J.F.; SANTOS, C.M.; ROSSI, L.R.; SILVA, M.B.P.; DAL PAI, E. Global, diffuse and direct solar radiation of the infrared spectrum in Botucatu/SP/Brazil. **Renewable and Sustainable Energy Reviews**, v. 82, p. 448-459, 2018.
- SENTELHAS, P.C.; MORAES, S.O.; PIEDADE, S.M.S.; PEREIRA, A.R.; ANGELOCCI, L.R.; MARIN, F.R. Análise comparativa de dados meteorológicos obtidos por Estações convencional e automática. **Revista Brasileira de Agrometeorologia**, v. 5, n. 2, p. 215-221, 1997.
- SMAABC. **Sistema de Monitoramento Agrometeorológico: São Paulo**. 2002. Disponível em [https://sma.fundacaoabc.org/climatologia/classificacao\\_climatica/sao\\_paulo](https://sma.fundacaoabc.org/climatologia/classificacao_climatica/sao_paulo). Acesso em 20 ago. 2020.
- STRASSBURGER, A.S.; MENEZES, A.J.E.A.; PERLEBERG, T.D.; EICHOLZ, E.D.; MENEZES, M.E.G.; SCHÖFFEL, E.R. Comparação da temperatura do ar obtida por estação meteorológica convencional e automática. **Revista Brasileira de Meteorologia**, v. 26, n. 2, p. 273-278, 2011.
- WILLMOTT, C.J.; ACKLESON, S.G.; DAVIS, R.E.; FEDDEMA, J.J.; KLINK, K.M.; LEGATES, D.R.; *et al.* Statistics for the evaluation and comparison of models. **Journal of Geophysical Research**, v. 90, n. C5, p. 8995-9005, 1985.