Table of Contents

# **IDENTIFICAÇÃO**

**Título do Plano de trabalho:** Controle de qualidade dos dados da temperatura do ar, para as estações meteorológicas automáticas do INMET na região sul do Brasil

**Nome do Bolsista:** Jonas Barboza Corrêa

**Nome do Orientador:** Jônatan Dupont Tatsch

**Local de execução:** Laboratório de Hidrometeorologia (LHMet - UFSM)

# **1. INTRODUÇÃO**

As observações meteorológicas (e relacionadas ambientais e geofísicas) são feitas por uma variedade de razões. Eles são usados para a preparação em tempo real de análises meteorológicas, previsões e advertências meteorológicas severas, para o estudo do clima, para operações locais dependentes do clima (por exemplo, operações locais de vôo de aeródromo, trabalhos de construção em terra e no mar), para hidrologia e meteorologia agrícola, e para pesquisa em meteorologia e climatologia [(WMO, 2012)](https://library.wmo.int/pmb_ged/wmo_8_en-2012.pdf).

Nos últimos anos, a automação da estação e o aumento das velocidades de transmissão de dados estão em progresso constante. É necessário um controle de qualidade rápido e efetivo para identificação e sinalização de erros ou observações suspeitas para proporcionar acesso rápido à informação e disseminação de observações confiáveis quanto possível aos usuários. Geralmente, os objetivos para o desenvolvimento de sistemas de controle de qualidade são os seguintes: tornar o controle de qualidade mais eficaz e mais próximo do tempo real; identificar erros de calibração, medição e comunicação tão próximos da fonte de observação quanto possível; focar no desenvolvimento automático de algoritmos de controle de qualidade; desenvolver um sistema abrangente de sinalização para indicar o nível de qualidade dos dados; para facilitar aos usuários de dados identificar dados suspeitos e errados e destacar os valores corrigidos [(VEJEN et al., 2002)](https://www.google.com.br/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&cad=rja&uact=8&ved=0ahUKEwjgneCYo5LZAhWNl5AKHWevC-cQFggzMAE&url=https%3A%2F%2Fwww.met.no%2Fpublikasjoner%2Fmet-report%2Fmet-report-2002%2F_%2Fattachment%2Fdownload%2F2fdbdbcf-2ae8-4bb3-86de-f7b5a6f11dae%3Ab322c47c99e9a34086013dfa7e4a915a383d2b42%2FMET-report-08-2002.pdf&usg=AOvVaw2iOII9wUFKYDayOfVnLgwv). Entre os principais tipos de erros, encontram-se os erros aleatórios, os erros sistemáticos, erros grandes e os erros micrometeorológicos.

Os erros aleatórios são distribuídos de forma mais ou menos simétrica em torno de zero e não dependem do valor medido. Erros aleatórios, por vezes, resultam em superestimação e às vezes em subestimação do valor real. Em média, os erros se cancelam mutuamente. Os erros sistemáticos, por outro lado, são distribuídos de forma assimétrica em torno de zero. Em média, esses erros tendem a polarizar o valor medido acima ou abaixo do valor real. Uma razão de erros aleatórios é uma deriva a longo prazo de sensores. Erros grandes (ásperos) são causados por mau funcionamento de dispositivos de medição ou por erros cometidos durante o processamento de dados; Os erros são facilmente detectados por cheques. Os erros micrometeorológicos (representatividade) são o resultado de perturbações em pequena escala ou sistemas meteorológicos que afetam a observação do tempo. Estes sistemas não são completamente observáveis pelo sistema de observação devido à resolução temporal ou espacial do sistema de observação. No entanto, quando tal fenômeno ocorre durante uma observação de rotina, os resultados podem parecer estranhos em comparação com as observações circundantes que ocorrem ao mesmo tempo [(ZAHUMENSKÝ, 2004)](https://www.wmo.int/pages/prog/www/IMOP/meetings/Surface/ET-STMT1_Geneva2004/Doc6.1(2).pdf).

No Brasil, os dados fornecidos, em sua maioria, encontram-se em sua forma bruta, sem que a qualidade dos mesmos seja verificada. Este trabalho consiste em, através da aplicação de certos métodos de controle de qualidade (QC), verificar e rotular tais dados como suspeitos ou não, para que futuramente possam serem melhores aproveitados em seu uso, com um maior grau de confiabilidade.

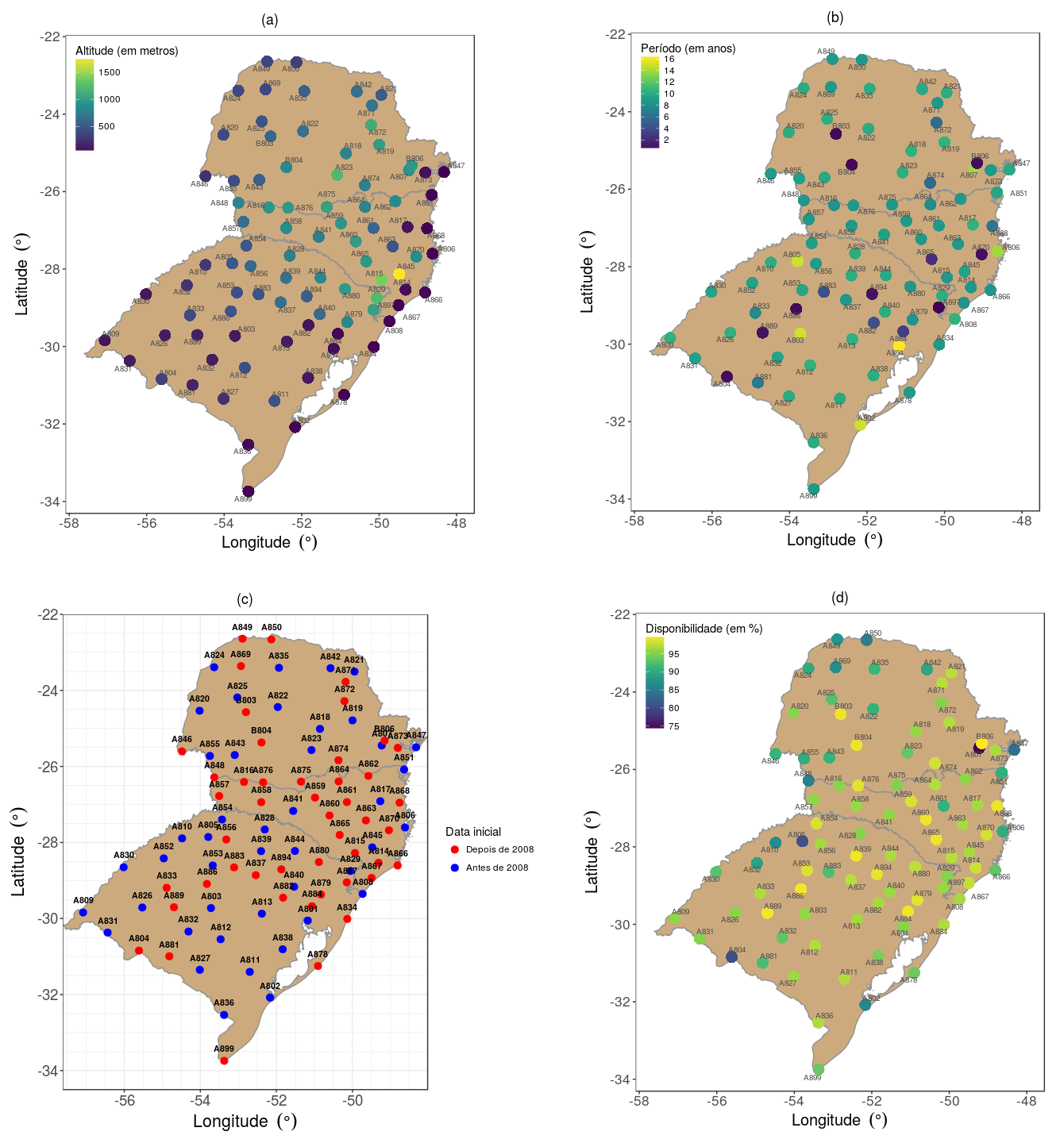
# **2. METODOLOGIA**

A metodologia adotada consiste em duas partes, inicialmente os dados brutos passarão por uma seleção inicial baseada em certos fatores, em seguida, eles serão submetidos à uma série de testes de controle de qualidade, que os qualificarão como suspeitos ou não, além, de obter-se certas informações relevantes a respeito deles.

## **2.1 SELEÇÃO DO PERÍODO DE DADOS**

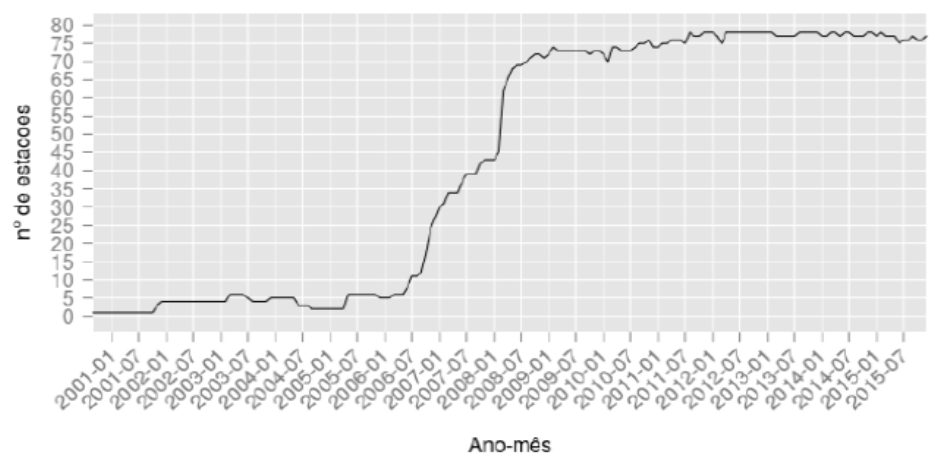
Neste trabalho, serão usados dados horários da Temperatura do Ar (Tar) de 91 Estações Meteorológicas Automáticas (EMAs) do [Instituto Nacional de Meteorologia (INMET)](http://www.inmet.gov.br/portal/), localizadas na região sul do Brasil. Ao todo, os estados do Paraná (PR), Rio Grande do Sul (RS) e Santa Catarina (SC), possuem, respectivamente, 27, 42 e 22 EMAs em funcionamento. Os dados horários de Tar utilizados estão no padrão do [Tempo Universal Coordenado (UTC)](https://pt.wikipedia.org/wiki/Tempo_Universal_Coordenado).

O modelo dos sensores de temperatura usados é o QMH102 da fabricante Vaisala, estes possuem acurácia de ±0,2°C, e são instalados a uma altura de 2 metros acima do solo. As EMAs devem estar em uma área livre de obstruções (como árvores e prédios), em solo gramado, tendo área mínima de 14mx18m, e estando cercada por tela metálica.

Figura 2.1 – (a) Localização das EMAs do INMET no sul do Brasil, (b) Período de dados das EMAs em anos, (c) EMAs com início de funcionamento pré-pós 2008, (d) Disponibilidade dos dados das EMAs em %. 

Dois critérios serão usados para a seleção dos dados que serão utilizados, (i) as estações meteorológicas automáticas devem terem no mínimo quatro anos de dados (podendo serem descontínuos), para haver um tempo mínimo para análise ser consistente, e (ii) o período de análise terá início em 01/01/2008, tendo como data final 31/12/2016. Este ano inicial foi escolhido, pois a partir dele notou-se um bom acrescimo na quantidade de EMAs existente, além de que, a distribuição espacial delas tornou-se mais homogênea. Dois critérios serão usados para a seleção dos dados que serão utilizados, (i) as estações meteorológicas automáticas devem terem no mínimo quatro anos de dados (podendo serem descontínuos), para haver um tempo mínimo para análise ser consistente, e (ii) o período de análise terá início em 01/01/2008, tendo como data final 31/12/2016. Este ano inicial foi escolhido, pois a partir dele notou-se um bom acrescimo na quantidade de EMAs existente, além de que, a distribuição espacial delas tornou-se mais homogênea.

Figura 2.2 – Evolução temporal mensal do número de EMAs.



Fonte: KANNENBERG, 2016

Para que a seleção fosse feita, antes, foi necessário uma regularização das séries temporais de cada EMA, para assim garantir que todas tenham 24 horas em cada dia, e 365 (ou 366 dias, se ano bissexto) em cada ano.

Após a seleção, restaram ao todo 80 EMAs para análise, estando 24 localizadas no estado do Paraná (PR), 36 no estado do Rio Grande do Sul (RS) e 20 no estado de Santa Catarina (SC). No quadro 2.1 é mostrado as EMAs selecionadas que apresentaram o maior período de dados.

Tabela 2.1 – EMAs com maior período de dados em ordem descrescente.

|  |  |  |  |
| --- | --- | --- | --- |
| Código | Estado | Nome | Período (em anos) |
| A801 | RS | Porto Alegre | 16.3 |
| A802 | RS | Rio Grande | 15.1 |
| A803 | RS | Santa Maria | 15.1 |
| A805 | RS | Santo Augusto | 15.1 |
| A806 | SC | Florianópolis-São José | 14.0 |
| A807 | PR | Curitiba | 13.9 |
| A808 | RS | Torres | 10.6 |
| A812 | RS | Caçapava Do Sul | 10.5 |
| A817 | SC | Indaial | 10.5 |
| A818 | PR | Ivaí | 10.5 |
| A819 | PR | Castro | 10.5 |
| A809 | RS | Uruguaiana | 10.3 |
| A813 | RS | Rio Pardo | 10.3 |
| A826 | RS | Alegrete | 10.3 |
| A823 | PR | Inácio Martins | 10.2 |
| A825 | PR | Goioere | 10.2 |
| A829 | RS | São José Dos Ausentes | 10.2 |
| A810 | RS | Santa Rosa | 10.1 |
| A820 | PR | Marechal Cândido Rondon | 10.1 |
| A821 | PR | Joaquim Távora | 10.1 |
| A824 | PR | Icaraíma | 10.1 |
| A828 | RS | Erechim | 10.1 |
| A835 | PR | Maringá | 10.1 |
| A838 | RS | Camaquã | 10.1 |
| A839 | RS | Passo Fundo | 10.1 |
| A840 | RS | Bento Gonçalves | 10.1 |
| A822 | PR | Nova Tebas | 10.0 |
| A827 | RS | Bagé | 10.0 |
| A836 | RS | Jaguarão | 10.0 |
| A811 | RS | Canguçu | 9.9 |
| A842 | PR | Nova Fátima | 9.9 |
| A843 | PR | Dois Vizinhos | 9.8 |
| A844 | RS | Lagoa Vemelha | 9.8 |
| A847 | PR | Ilha Do Mel | 9.8 |
| A851 | SC | Itapoá | 9.6 |
| A853 | RS | Cruz Alta | 9.6 |
| A830 | RS | São Borja | 9.5 |
| A832 | RS | São Gabriel | 9.5 |
| A845 | SC | Morro Da Igreja (Bom Jardim Da Serra) | 9.5 |
| A852 | RS | São Luiz Gonzaga | 9.4 |
| A841 | SC | Joaçaba | 9.3 |
| A831 | RS | Quaraí | 9.2 |
| A855 | PR | Planalto | 9.2 |
| A854 | RS | Frederico Westphalen | 9.1 |
| A846 | PR | Foz Do Iguaçu | 8.9 |
| A856 | RS | Palmeira Das Missões | 8.9 |
| A834 | RS | Tramandaí | 8.8 |
| A837 | RS | Soledade | 8.8 |
| A849 | PR | Diamante Do Norte | 8.8 |
| A850 | PR | Paranapoema | 8.8 |
| A857 | SC | São Miguel Do Oeste | 8.8 |
| A858 | SC | Xanxerê | 8.8 |
| A859 | SC | Caçador | 8.8 |
| A860 | SC | Curitibanos | 8.8 |
| A861 | SC | Rio Do Campo | 8.8 |
| A862 | SC | Rio Negrinho | 8.8 |
| A863 | SC | Ituporanga | 8.8 |
| A869 | PR | Cidade Gaúcha | 8.8 |
| A871 | PR | Japira | 8.8 |
| A873 | PR | Morretes | 8.8 |
| A878 | RS | Mostardas | 8.8 |
| A899 | RS | Santa Vitória Do Palmar (Barra Do Chuí) | 8.8 |
| A815 | SC | São Joaquim | 8.7 |
| A875 | PR | General Carneiro | 8.7 |
| A880 | RS | Vacaria | 8.7 |
| A814 | SC | Urussanga | 8.6 |
| A848 | SC | Dionísio Cerqueira | 8.6 |
| A866 | SC | Laguna (Farol Santa Marta) | 8.6 |
| A876 | PR | Clevelândia | 8.6 |
| A879 | RS | Canela | 8.4 |
| A816 | SC | Novo Horizonte | 8.3 |
| A867 | SC | Araranguá | 8.3 |
| A833 | RS | Santiago | 7.9 |
| A864 | SC | Major Vieira | 7.9 |
| A881 | RS | Dom Pedrito | 6.7 |
| A868 | SC | Itajaí | 6.5 |
| A872 | PR | Ventania | 5.7 |
| A874 | PR | São Mateus Do Sul | 5.7 |
| A882 | RS | Teutônia | 4.2 |
| A883 | RS | Ibirubá | 4.1 |

Na realização deste trabalho, foi utilizado a linguagem de programação [R](https://www.r-project.org/) (Version 1.1.383 – © 2009-2017 RStudio, Inc.) no sistema operacional [Linux Mind](https://linuxmint.com/edition.php?id=204) ("Rosa" Version), com o software livre com ambiente de desenvolvimento integrado, [Rstudio](https://www.rstudio.com).

Os seguintes pacotes foram utilizados: [dplyr](https://cran.r-project.org/web/packages/dplyr/index.html) (WICKHAM, 2017), [ggplot2](https://cran.r-project.org/web/packages/ggplot2/index.html) (WICKHAM, 2016), [kableExtra](https://cran.r-project.org/web/packages/kableExtra/index.html) (ZHU, 2017), [knitr](https://cran.r-project.org/web/packages/knitr/index.html) (XIE, 2017), [lubridate](https://cran.r-project.org/web/packages/lubridate/index.html) (GROLEMUND, 2016), [magrittr](https://cran.r-project.org/web/packages/magrittr/index.html) (BACHE; WICKHAM, 2014), [openair](https://cran.r-project.org/web/packages/openair/index.html) (CARSLAW; ROPKINS, 2017), [padr](https://cran.r-project.org/web/packages/padr/index.html) (THOEN, 2017), [plyr](https://cran.r-project.org/web/packages/plyr/index.html) (WICKHAM, 2016), [raster](https://cran.r-project.org/web/packages/raster/index.html) (HIJMANS, 2016), [scales](https://cran.r-project.org/web/packages/scales/index.html) (WICKHAM, 2017), [stringr](https://cran.r-project.org/web/packages/stringr/index.html) (WICKHAM, 2017), [tidyverse](https://cran.r-project.org/web/packages/tidyverse/index.html) (WICKHAM, 2017).

## **2.2 TESTES DE CONTROLE DE QUALIDADE**

Os testes aplicados aos dados da temperatura do ar, serão dividos em categorias, sendo elas: Limites do Intervalo de Variação, Persistência Temporal, Consistência Interna, Consistência Temporal, Consistência Espacial e Homogeneidade Temporal. Os dados serão considerados suspeitos quando os testes abaixo forem válidos.

### **2.2.1 Limites do Intervalo de Variação (QC1)**

#### 

**(a)** ou [*(ESTÉVEZ el al., 2011)*](https://www.google.com.br/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKEwik6umip5LZAhWGD5AKHYfEAK0QFggrMAA&url=https%3A%2F%2Fwww.researchgate.net%2Fprofile%2FVladimir_Kulchitsky%2Fpost%2FWho_is_assessing_the_quality_of_Net-Atmo_automated_weather_stations%2Fattachment%2F59d63df579197b807799aa57%2FAS%3A422033134755841%401477631948607%2Fdownload%2FEstevez%2Bet%2Bal.%2B2011.pdf)

Onde:

* é a média aritmética simples de (temperatura mínima do ar) e (temperatura máxima do ar);
* e são, respectivamente, os limites mínimo e máximo instrumental;

#### 

**(b)** ou [*(ESTÉVEZ el al., 2011)*](https://www.google.com.br/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKEwik6umip5LZAhWGD5AKHYfEAK0QFggrMAA&url=https%3A%2F%2Fwww.researchgate.net%2Fprofile%2FVladimir_Kulchitsky%2Fpost%2FWho_is_assessing_the_quality_of_Net-Atmo_automated_weather_stations%2Fattachment%2F59d63df579197b807799aa57%2FAS%3A422033134755841%401477631948607%2Fdownload%2FEstevez%2Bet%2Bal.%2B2011.pdf)

Onde:

* e são, respectivamente, os valores mínimo e máximo já registrados pela estação climatológica mais próxima de cada EMA;

### **2.2.2 Persistência Temporal (QC2)**

#### 

**(a)** [*(MEEK; HATFIELD, 1994)*](https://ac.els-cdn.com/0168192394900833/1-s2.0-0168192394900833-main.pdf?_tid=868c718a-0b8d-11e8-8581-00000aab0f26&acdnat=1517956491_41a505ab0f4dff5b801e7e33aab57b66)

Onde:

* é um número natural, correspondente à(s) hora(s) anteriores à ;

### **2.2.3 Consistência Interna (QC3)**

#### 

**(a)** [*(ESTÉVEZ el al., 2011)*](https://www.google.com.br/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKEwik6umip5LZAhWGD5AKHYfEAK0QFggrMAA&url=https%3A%2F%2Fwww.researchgate.net%2Fprofile%2FVladimir_Kulchitsky%2Fpost%2FWho_is_assessing_the_quality_of_Net-Atmo_automated_weather_stations%2Fattachment%2F59d63df579197b807799aa57%2FAS%3A422033134755841%401477631948607%2Fdownload%2FEstevez%2Bet%2Bal.%2B2011.pdf)

Onde:

* e são, respectivamente, os valores mínimo e máximo horário de temperatura;

#### 

**(b)** ou

Onde:

* é a temperatura instantânea do ar, que corresponde à média de um minuto de 12 valores de amostragens medidos a cada cinco segundos;

#### 

**(c)**

Onde:

* é a temperatura média do ponto de orvalho, filtrada apenas com dados que passaram nos testes anteriores de QC3, aplicados para a variável (temperatura média do ponto de orvalho horária);

#### 

**(d)** ou [*(MEEK; HATFIELD, 1994)*](https://ac.els-cdn.com/0168192394900833/1-s2.0-0168192394900833-main.pdf?_tid=868c718a-0b8d-11e8-8581-00000aab0f26&acdnat=1517956491_41a505ab0f4dff5b801e7e33aab57b66)

Onde:

* é média aritmética simples das ;
* é o menor valor de temperatura registrada no dia das ;
* é o maior valor de temperatura registrada no dia das ;

#### 

**(e)** [*(ESTÉVEZ el al., 2011)*](https://www.google.com.br/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKEwik6umip5LZAhWGD5AKHYfEAK0QFggrMAA&url=https%3A%2F%2Fwww.researchgate.net%2Fprofile%2FVladimir_Kulchitsky%2Fpost%2FWho_is_assessing_the_quality_of_Net-Atmo_automated_weather_stations%2Fattachment%2F59d63df579197b807799aa57%2FAS%3A422033134755841%401477631948607%2Fdownload%2FEstevez%2Bet%2Bal.%2B2011.pdf)

Onde:

* é a temperatura mínima do dia anterior;

#### 

**(f)** [*(ESTÉVEZ el al., 2011)*](https://www.google.com.br/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKEwik6umip5LZAhWGD5AKHYfEAK0QFggrMAA&url=https%3A%2F%2Fwww.researchgate.net%2Fprofile%2FVladimir_Kulchitsky%2Fpost%2FWho_is_assessing_the_quality_of_Net-Atmo_automated_weather_stations%2Fattachment%2F59d63df579197b807799aa57%2FAS%3A422033134755841%401477631948607%2Fdownload%2FEstevez%2Bet%2Bal.%2B2011.pdf)

Onde:

* é a temperatura máxima do dia anterior;

### **2.2.4 Consistência Temporal (QC4)**

#### 

**(a)** [*(WMO, 1993)*](https://library.wmo.int/pmb_ged/wmo_305_en.pdf)

Onde:

* são valorores correspondetes a uma determinada hora, sendo eles: 1h, 2h, 3h, 6h e 12h;
* é um valor tolerável de temperatura, que varia com , recebendo os seguintes valores respectivos: 4°C (1h), 7°C (2h), 9°C (3h), 15°C (6h) e 25°C (12h);

#### 

**(b)** Valente & Tatsch

### **2.2.5 Consistência Espacial (QC5)**

### **2.2.6 Homogeneidade Temporal (QC6)**

# **3. RESULTADOS**

## **3.1 Limites do Intervalo de Variação (QC1)**

### 

**(a)** Das 80 EMAs , nenhuma apresentou dados suspeitos.

**(b)** Das 80 EMAs, 66 apresentaram dados suspeitos.

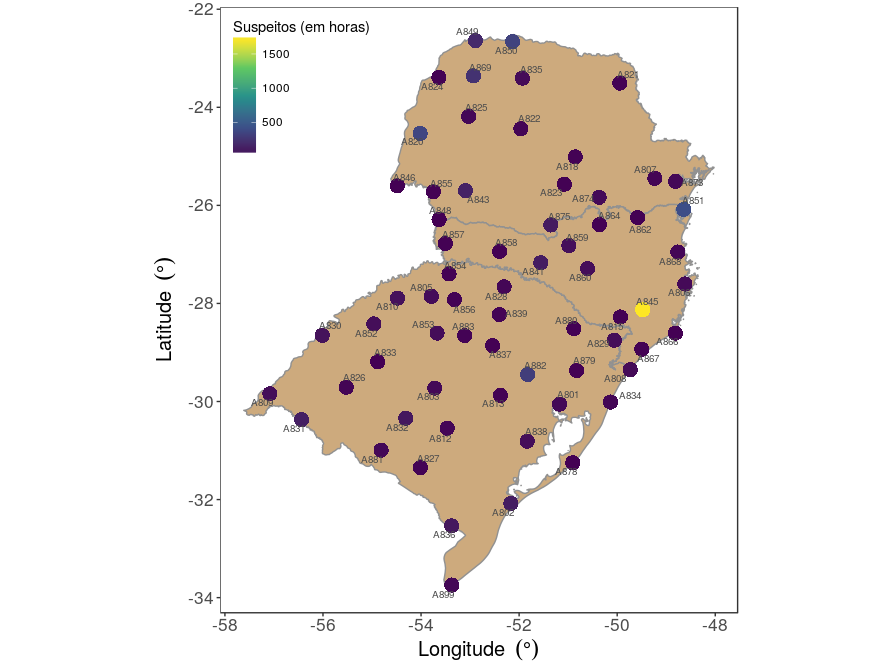
Tabela 3.1 – EMAs com maior número de dados considerados suspeitos no teste QC1b em ordem decrescente.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Código | Estado | Nome | Dados Suspeitos (em horas) | Dados Suspeitos (em %) |
| A845 | SC | Morro Da Igreja (Bom Jardim Da Serra) | 1726 | 2.1872 |
| A851 | SC | Itapoá | 402 | 0.5094 |
| A820 | PR | Marechal Cândido Rondon | 354 | 0.4486 |
| A850 | PR | Paranapoema | 337 | 0.4271 |
| A882 | RS | Teutônia | 306 | 0.3878 |
| A869 | PR | Cidade Gaúcha | 229 | 0.2902 |
| A849 | PR | Diamante Do Norte | 183 | 0.2319 |
| A843 | PR | Dois Vizinhos | 136 | 0.1723 |
| A831 | RS | Quaraí | 122 | 0.1546 |
| A802 | RS | Rio Grande | 109 | 0.1381 |
| A841 | SC | Joaçaba | 103 | 0.1305 |
| A875 | PR | General Carneiro | 97 | 0.1229 |
| A836 | RS | Jaguarão | 86 | 0.1090 |
| A832 | RS | São Gabriel | 84 | 0.1064 |
| A830 | RS | São Borja | 62 | 0.0786 |
| A810 | RS | Santa Rosa | 54 | 0.0684 |
| A859 | SC | Caçador | 51 | 0.0646 |
| A838 | RS | Camaquã | 45 | 0.0570 |
| A852 | RS | São Luiz Gonzaga | 45 | 0.0570 |
| A809 | RS | Uruguaiana | 43 | 0.0545 |
| A829 | RS | São José Dos Ausentes | 43 | 0.0545 |
| A860 | SC | Curitibanos | 43 | 0.0545 |
| A853 | RS | Cruz Alta | 41 | 0.0520 |
| A805 | RS | Santo Augusto | 38 | 0.0482 |
| A835 | PR | Maringá | 34 | 0.0431 |
| A808 | RS | Torres | 32 | 0.0406 |
| A806 | SC | Florianópolis-São José | 31 | 0.0393 |
| A874 | PR | São Mateus Do Sul | 30 | 0.0380 |
| A801 | RS | Porto Alegre | 29 | 0.0367 |
| A803 | RS | Santa Maria | 28 | 0.0355 |
| A825 | PR | Goioere | 25 | 0.0317 |
| A899 | RS | Santa Vitória Do Palmar (Barra Do Chuí) | 25 | 0.0317 |
| A873 | PR | Morretes | 24 | 0.0304 |
| A881 | RS | Dom Pedrito | 23 | 0.0291 |
| A828 | RS | Erechim | 21 | 0.0266 |
| A812 | RS | Caçapava Do Sul | 17 | 0.0215 |
| A834 | RS | Tramandaí | 17 | 0.0215 |
| A868 | SC | Itajaí | 17 | 0.0215 |
| A807 | PR | Curitiba | 13 | 0.0165 |
| A822 | PR | Nova Tebas | 13 | 0.0165 |
| A826 | RS | Alegrete | 13 | 0.0165 |
| A856 | RS | Palmeira Das Missões | 9 | 0.0114 |
| A883 | RS | Ibirubá | 9 | 0.0114 |
| A827 | RS | Bagé | 8 | 0.0101 |
| A818 | PR | Ivaí | 7 | 0.0089 |
| A839 | RS | Passo Fundo | 7 | 0.0089 |
| A821 | PR | Joaquim Távora | 6 | 0.0076 |
| A880 | RS | Vacaria | 6 | 0.0076 |
| A823 | PR | Inácio Martins | 5 | 0.0063 |
| A824 | PR | Icaraíma | 5 | 0.0063 |
| A855 | PR | Planalto | 5 | 0.0063 |
| A813 | RS | Rio Pardo | 4 | 0.0051 |
| A846 | PR | Foz Do Iguaçu | 4 | 0.0051 |
| A854 | RS | Frederico Westphalen | 4 | 0.0051 |
| A857 | SC | São Miguel Do Oeste | 4 | 0.0051 |
| A867 | SC | Araranguá | 4 | 0.0051 |
| A815 | SC | São Joaquim | 3 | 0.0038 |
| A837 | RS | Soledade | 3 | 0.0038 |
| A862 | SC | Rio Negrinho | 3 | 0.0038 |
| A864 | SC | Major Vieira | 3 | 0.0038 |
| A833 | RS | Santiago | 2 | 0.0025 |
| A858 | SC | Xanxerê | 2 | 0.0025 |
| A866 | SC | Laguna (Farol Santa Marta) | 2 | 0.0025 |
| A848 | SC | Dionísio Cerqueira | 1 | 0.0013 |
| A878 | RS | Mostardas | 1 | 0.0013 |
| A879 | RS | Canela | 1 | 0.0013 |

table.intro(  
 directory = '/home/jonas/Documents/jbc-ic/output/qc1b-summary-tavg-data-inmet-2008-2016-4yrs-south.rds',  
 EMAS = emas\_sel  
)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Código | Estado | Nome | Dados Suspeitos (em horas) | Dados Suspeitos (em %) |
| A845 | SC | Morro Da Igreja (Bom Jardim Da Serra) | 1726 | 2.1872 |
| A851 | SC | Itapoá | 402 | 0.5094 |
| A820 | PR | Marechal Cândido Rondon | 354 | 0.4486 |
| A850 | PR | Paranapoema | 337 | 0.4271 |
| A882 | RS | Teutônia | 306 | 0.3878 |
| A869 | PR | Cidade Gaúcha | 229 | 0.2902 |
| A849 | PR | Diamante Do Norte | 183 | 0.2319 |
| A843 | PR | Dois Vizinhos | 136 | 0.1723 |
| A831 | RS | Quaraí | 122 | 0.1546 |
| A802 | RS | Rio Grande | 109 | 0.1381 |
| A841 | SC | Joaçaba | 103 | 0.1305 |
| A875 | PR | General Carneiro | 97 | 0.1229 |
| A836 | RS | Jaguarão | 86 | 0.1090 |
| A832 | RS | São Gabriel | 84 | 0.1064 |
| A830 | RS | São Borja | 62 | 0.0786 |
| A810 | RS | Santa Rosa | 54 | 0.0684 |
| A859 | SC | Caçador | 51 | 0.0646 |
| A838 | RS | Camaquã | 45 | 0.0570 |
| A852 | RS | São Luiz Gonzaga | 45 | 0.0570 |
| A809 | RS | Uruguaiana | 43 | 0.0545 |
| A829 | RS | São José Dos Ausentes | 43 | 0.0545 |
| A860 | SC | Curitibanos | 43 | 0.0545 |
| A853 | RS | Cruz Alta | 41 | 0.0520 |
| A805 | RS | Santo Augusto | 38 | 0.0482 |
| A835 | PR | Maringá | 34 | 0.0431 |
| A808 | RS | Torres | 32 | 0.0406 |
| A806 | SC | Florianópolis-São José | 31 | 0.0393 |
| A874 | PR | São Mateus Do Sul | 30 | 0.0380 |
| A801 | RS | Porto Alegre | 29 | 0.0367 |
| A803 | RS | Santa Maria | 28 | 0.0355 |
| A825 | PR | Goioere | 25 | 0.0317 |
| A899 | RS | Santa Vitória Do Palmar (Barra Do Chuí) | 25 | 0.0317 |
| A873 | PR | Morretes | 24 | 0.0304 |
| A881 | RS | Dom Pedrito | 23 | 0.0291 |
| A828 | RS | Erechim | 21 | 0.0266 |
| A812 | RS | Caçapava Do Sul | 17 | 0.0215 |
| A834 | RS | Tramandaí | 17 | 0.0215 |
| A868 | SC | Itajaí | 17 | 0.0215 |
| A807 | PR | Curitiba | 13 | 0.0165 |
| A822 | PR | Nova Tebas | 13 | 0.0165 |
| A826 | RS | Alegrete | 13 | 0.0165 |
| A856 | RS | Palmeira Das Missões | 9 | 0.0114 |
| A883 | RS | Ibirubá | 9 | 0.0114 |
| A827 | RS | Bagé | 8 | 0.0101 |
| A818 | PR | Ivaí | 7 | 0.0089 |
| A839 | RS | Passo Fundo | 7 | 0.0089 |
| A821 | PR | Joaquim Távora | 6 | 0.0076 |
| A880 | RS | Vacaria | 6 | 0.0076 |
| A823 | PR | Inácio Martins | 5 | 0.0063 |
| A824 | PR | Icaraíma | 5 | 0.0063 |
| A855 | PR | Planalto | 5 | 0.0063 |
| A813 | RS | Rio Pardo | 4 | 0.0051 |
| A846 | PR | Foz Do Iguaçu | 4 | 0.0051 |
| A854 | RS | Frederico Westphalen | 4 | 0.0051 |
| A857 | SC | São Miguel Do Oeste | 4 | 0.0051 |
| A867 | SC | Araranguá | 4 | 0.0051 |
| A815 | SC | São Joaquim | 3 | 0.0038 |
| A837 | RS | Soledade | 3 | 0.0038 |
| A862 | SC | Rio Negrinho | 3 | 0.0038 |
| A864 | SC | Major Vieira | 3 | 0.0038 |
| A833 | RS | Santiago | 2 | 0.0025 |
| A858 | SC | Xanxerê | 2 | 0.0025 |
| A866 | SC | Laguna (Farol Santa Marta) | 2 | 0.0025 |
| A848 | SC | Dionísio Cerqueira | 1 | 0.0013 |
| A878 | RS | Mostardas | 1 | 0.0013 |
| A879 | RS | Canela | 1 | 0.0013 |

Figura 3.1 – Distribuição espacial das EMAs com dados considerados suspeitos no teste QC1b.



## **3.2 Persistência Temporal (QC2)**

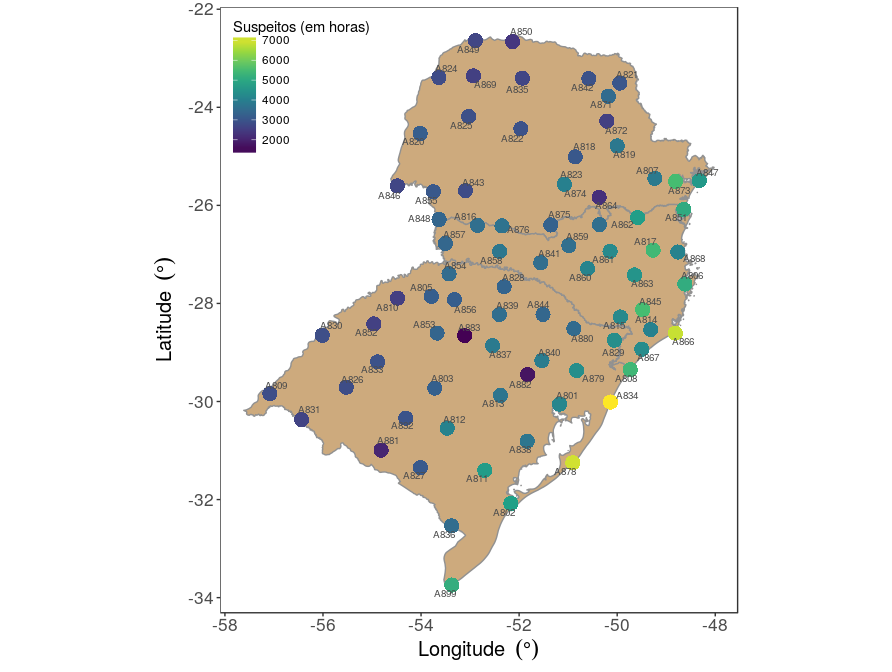
### 

**(a)** Das 80 EMAs , 80 apresentaram dados suspeitos.

Tabela 3.2 – EMAs com maior número de dados considerados suspeitos no teste QC2a em ordem decrescente.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Código | Estado | Nome | Dados Suspeitos (em horas) | Dados Suspeitos (em %) |
| A834 | RS | Tramandaí | 7454 | 9.4460 |
| A878 | RS | Mostardas | 7028 | 8.9061 |
| A866 | SC | Laguna (Farol Santa Marta) | 6933 | 8.7857 |
| A845 | SC | Morro Da Igreja (Bom Jardim Da Serra) | 5604 | 7.1016 |
| A873 | PR | Morretes | 5590 | 7.0838 |
| A817 | SC | Indaial | 5459 | 6.9178 |
| A808 | RS | Torres | 5437 | 6.8900 |
| A899 | RS | Santa Vitória Do Palmar (Barra Do Chuí) | 5180 | 6.5643 |
| A806 | SC | Florianópolis-São José | 5142 | 6.5161 |
| A851 | SC | Itapoá | 5088 | 6.4477 |
| A802 | RS | Rio Grande | 4810 | 6.0954 |
| A862 | SC | Rio Negrinho | 4779 | 6.0561 |
| A811 | RS | Canguçu | 4705 | 5.9623 |
| A847 | PR | Ilha Do Mel | 4646 | 5.8876 |
| A863 | SC | Ituporanga | 4502 | 5.7051 |
| A829 | RS | São José Dos Ausentes | 4363 | 5.5289 |
| A879 | RS | Canela | 4361 | 5.5264 |
| A867 | SC | Araranguá | 4301 | 5.4504 |
| A861 | SC | Rio Do Campo | 4266 | 5.4060 |
| A815 | SC | São Joaquim | 4222 | 5.3503 |
| A801 | RS | Porto Alegre | 4158 | 5.2692 |
| A860 | SC | Curitibanos | 4093 | 5.1868 |
| A812 | RS | Caçapava Do Sul | 4059 | 5.1437 |
| A868 | SC | Itajaí | 4058 | 5.1424 |
| A814 | SC | Urussanga | 4020 | 5.0943 |
| A823 | PR | Inácio Martins | 3999 | 5.0677 |
| A840 | RS | Bento Gonçalves | 3889 | 4.9283 |
| A837 | RS | Soledade | 3849 | 4.8776 |
| A807 | PR | Curitiba | 3841 | 4.8674 |
| A819 | PR | Castro | 3793 | 4.8066 |
| A858 | SC | Xanxerê | 3761 | 4.7661 |
| A838 | RS | Camaquã | 3757 | 4.7610 |
| A813 | RS | Rio Pardo | 3731 | 4.7281 |
| A876 | PR | Clevelândia | 3685 | 4.6698 |
| A859 | SC | Caçador | 3583 | 4.5405 |
| A864 | SC | Major Vieira | 3577 | 4.5329 |
| A816 | SC | Novo Horizonte | 3570 | 4.5240 |
| A839 | RS | Passo Fundo | 3547 | 4.4949 |
| A841 | SC | Joaçaba | 3545 | 4.4923 |
| A836 | RS | Jaguarão | 3521 | 4.4619 |
| A854 | RS | Frederico Westphalen | 3490 | 4.4226 |
| A880 | RS | Vacaria | 3473 | 4.4011 |
| A871 | PR | Japira | 3470 | 4.3973 |
| A844 | RS | Lagoa Vemelha | 3373 | 4.2744 |
| A857 | SC | São Miguel Do Oeste | 3373 | 4.2744 |
| A848 | SC | Dionísio Cerqueira | 3339 | 4.2313 |
| A853 | RS | Cruz Alta | 3333 | 4.2237 |
| A828 | RS | Erechim | 3299 | 4.1806 |
| A875 | PR | General Carneiro | 3250 | 4.1185 |
| A803 | RS | Santa Maria | 3205 | 4.0615 |
| A805 | RS | Santo Augusto | 3158 | 4.0019 |
| A820 | PR | Marechal Cândido Rondon | 3148 | 3.9893 |
| A856 | RS | Palmeira Das Missões | 3139 | 3.9778 |
| A827 | RS | Bagé | 3041 | 3.8537 |
| A855 | PR | Planalto | 3013 | 3.8182 |
| A818 | PR | Ivaí | 3011 | 3.8156 |
| A821 | PR | Joaquim Távora | 2963 | 3.7548 |
| A832 | RS | São Gabriel | 2912 | 3.6902 |
| A842 | PR | Nova Fátima | 2888 | 3.6598 |
| A833 | RS | Santiago | 2878 | 3.6471 |
| A822 | PR | Nova Tebas | 2877 | 3.6458 |
| A825 | PR | Goioere | 2828 | 3.5837 |
| A830 | RS | São Borja | 2805 | 3.5546 |
| A809 | RS | Uruguaiana | 2796 | 3.5432 |
| A826 | RS | Alegrete | 2779 | 3.5216 |
| A843 | PR | Dois Vizinhos | 2758 | 3.4950 |
| A824 | PR | Icaraíma | 2744 | 3.4773 |
| A849 | PR | Diamante Do Norte | 2622 | 3.3227 |
| A835 | PR | Maringá | 2614 | 3.3126 |
| A846 | PR | Foz Do Iguaçu | 2610 | 3.3075 |
| A831 | RS | Quaraí | 2585 | 3.2758 |
| A852 | RS | São Luiz Gonzaga | 2524 | 3.1985 |
| A869 | PR | Cidade Gaúcha | 2489 | 3.1541 |
| A872 | PR | Ventania | 2486 | 3.1503 |
| A810 | RS | Santa Rosa | 2464 | 3.1225 |
| A850 | PR | Paranapoema | 2289 | 2.9007 |
| A874 | PR | São Mateus Do Sul | 2149 | 2.7233 |
| A881 | RS | Dom Pedrito | 1962 | 2.4863 |
| A882 | RS | Teutônia | 1629 | 2.0643 |
| A883 | RS | Ibirubá | 1360 | 1.7234 |

Figura 3.2 – Distribuição espacial das EMAs com dados considerados suspeitos no teste QC2a.



## **3.3 Consistência Interna (QC3)**

### 

**(a)** Das 80 EMAs , 80 apresentaram dados suspeitos.

Tabela 3.3 – EMAs com maior número de dados considerados suspeitos no teste QC3a em ordem decrescente.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Código | Estado | Nome | Dados Suspeitos (em horas) | Dados Suspeitos (em %) |
| A851 | SC | Itapoá | 538 | 0.6818 |
| A862 | SC | Rio Negrinho | 389 | 0.4930 |
| A867 | SC | Araranguá | 328 | 0.4157 |
| A803 | RS | Santa Maria | 272 | 0.3447 |
| A811 | RS | Canguçu | 271 | 0.3434 |
| A834 | RS | Tramandaí | 255 | 0.3231 |
| A817 | SC | Indaial | 240 | 0.3041 |
| A819 | PR | Castro | 240 | 0.3041 |
| A864 | SC | Major Vieira | 234 | 0.2965 |
| A838 | RS | Camaquã | 214 | 0.2712 |
| A861 | SC | Rio Do Campo | 209 | 0.2649 |
| A813 | RS | Rio Pardo | 205 | 0.2598 |
| A868 | SC | Itajaí | 196 | 0.2484 |
| A863 | SC | Ituporanga | 190 | 0.2408 |
| A880 | RS | Vacaria | 183 | 0.2319 |
| A812 | RS | Caçapava Do Sul | 168 | 0.2129 |
| A875 | PR | General Carneiro | 167 | 0.2116 |
| A823 | PR | Inácio Martins | 166 | 0.2104 |
| A857 | SC | São Miguel Do Oeste | 163 | 0.2066 |
| A878 | RS | Mostardas | 156 | 0.1977 |
| A873 | PR | Morretes | 150 | 0.1901 |
| A801 | RS | Porto Alegre | 140 | 0.1774 |
| A805 | RS | Santo Augusto | 140 | 0.1774 |
| A845 | SC | Morro Da Igreja (Bom Jardim Da Serra) | 134 | 0.1698 |
| A855 | PR | Planalto | 130 | 0.1647 |
| A871 | PR | Japira | 128 | 0.1622 |
| A808 | RS | Torres | 126 | 0.1597 |
| A847 | PR | Ilha Do Mel | 124 | 0.1571 |
| A860 | SC | Curitibanos | 124 | 0.1571 |
| A858 | SC | Xanxerê | 122 | 0.1546 |
| A818 | PR | Ivaí | 112 | 0.1419 |
| A836 | RS | Jaguarão | 107 | 0.1356 |
| A876 | PR | Clevelândia | 106 | 0.1343 |
| A833 | RS | Santiago | 104 | 0.1318 |
| A816 | SC | Novo Horizonte | 102 | 0.1293 |
| A828 | RS | Erechim | 101 | 0.1280 |
| A874 | PR | São Mateus Do Sul | 97 | 0.1229 |
| A815 | SC | São Joaquim | 96 | 0.1217 |
| A872 | PR | Ventania | 95 | 0.1204 |
| A830 | RS | São Borja | 94 | 0.1191 |
| A856 | RS | Palmeira Das Missões | 93 | 0.1179 |
| A866 | SC | Laguna (Farol Santa Marta) | 90 | 0.1141 |
| A820 | PR | Marechal Cândido Rondon | 89 | 0.1128 |
| A844 | RS | Lagoa Vemelha | 89 | 0.1128 |
| A829 | RS | São José Dos Ausentes | 88 | 0.1115 |
| A852 | RS | São Luiz Gonzaga | 87 | 0.1102 |
| A882 | RS | Teutônia | 87 | 0.1102 |
| A832 | RS | São Gabriel | 86 | 0.1090 |
| A837 | RS | Soledade | 86 | 0.1090 |
| A899 | RS | Santa Vitória Do Palmar (Barra Do Chuí) | 85 | 0.1077 |
| A826 | RS | Alegrete | 78 | 0.0988 |
| A879 | RS | Canela | 78 | 0.0988 |
| A814 | SC | Urussanga | 75 | 0.0950 |
| A802 | RS | Rio Grande | 70 | 0.0887 |
| A854 | RS | Frederico Westphalen | 67 | 0.0849 |
| A824 | PR | Icaraíma | 66 | 0.0836 |
| A843 | PR | Dois Vizinhos | 63 | 0.0798 |
| A839 | RS | Passo Fundo | 62 | 0.0786 |
| A853 | RS | Cruz Alta | 60 | 0.0760 |
| A848 | SC | Dionísio Cerqueira | 59 | 0.0748 |
| A859 | SC | Caçador | 59 | 0.0748 |
| A807 | PR | Curitiba | 56 | 0.0710 |
| A840 | RS | Bento Gonçalves | 55 | 0.0697 |
| A809 | RS | Uruguaiana | 53 | 0.0672 |
| A849 | PR | Diamante Do Norte | 52 | 0.0659 |
| A846 | PR | Foz Do Iguaçu | 51 | 0.0646 |
| A842 | PR | Nova Fátima | 47 | 0.0596 |
| A810 | RS | Santa Rosa | 45 | 0.0570 |
| A827 | RS | Bagé | 45 | 0.0570 |
| A835 | PR | Maringá | 45 | 0.0570 |
| A825 | PR | Goioere | 44 | 0.0558 |
| A883 | RS | Ibirubá | 44 | 0.0558 |
| A822 | PR | Nova Tebas | 42 | 0.0532 |
| A881 | RS | Dom Pedrito | 39 | 0.0494 |
| A806 | SC | Florianópolis-São José | 37 | 0.0469 |
| A821 | PR | Joaquim Távora | 36 | 0.0456 |
| A831 | RS | Quaraí | 36 | 0.0456 |
| A869 | PR | Cidade Gaúcha | 36 | 0.0456 |
| A841 | SC | Joaçaba | 25 | 0.0317 |
| A850 | PR | Paranapoema | 15 | 0.0190 |

## **3.4 Consistência Temporal (QC4)**

## **3.5 Consistência Espacial (QC5)**

## **3.6 Homogeneidade Temporal (QC6)**

# **4. CONCLUSÃO**

Somando os resultados dos testes, as estações que apresentaram o maior número de dados suspeitos foram A875 - General Carneiro com 10943 horas; A845 – Morro Da Igreja (Bom Jardim Da Serra) com 8465 horas; A834 – Tramandaí com 7825; A819 – Castro com 7778 horas; A862 – Rio Negrinho com 7382 horas; A851 – Itapoá; A878 – Mostardas com 7289; A866 – Laguna (Farol Santa Marta) com 7055; A873 – Morretes com 6595 horas e A831 – Quaraí com 6526 horas.

No geral, embora a maioria da estações meteorológicas automáticas tenham apresentado ao mínimo um teste onde foram detectados dados suspeitos, a porcentagem de dados suspeitos detectada em cada estação, em relação ao número total de horas de dados na mesma, é suficientemente pequeno.

# **5. REFERÊNCIAS BIBLIOGRÁFICAS**

Carslaw, D. C. and K. Ropkins, **(2012) openair --- an R package for air quality data analysis.** Environmental Modelling & Software. Volume 27-28, 52-61.

Edwin Thoen (2017). **padr: Quickly Get Datetime Data Ready for Analysis.** R package version 0.3.0. <https://CRAN.R-project.org/package=padr>

ESTÉVEZ, J. et al. **Guidelines on validation procedures for meteorological data from automatic weather stations.** Journal of Hydrology, p. 147, 2011.

Garrett Grolemund, Hadley Wickham (2011). **Dates and Times Made Easy with lubridate.** Journal of Statistical Software, 40(3), 1-25. URL <http://www.jstatsoft.org/v40/i03/>.

Hadley Wickham, Romain Francois, Lionel Henry and Kirill Müller (2017). **dplyr: A Grammar of Data Manipulation.** R package version 0.7.4. <https://CRAN.R-project.org/package=dplyr>

H. Wickham. **ggplot2: Elegant Graphics for Data Analysis.** Springer-Verlag New York, 2009

Hadley Wickham (2017). **scales: Scale Functions for Visualization.** R package version 0.5.0. <https://CRAN.R-project.org/package=scales>

Hadley Wickham (2017). **stringr: Simple, Consistent Wrappers for Common String Operations.** R package version 1.2.0. <https://CRAN.R-project.org/package=stringr>

Hadley Wickham (2017). **tidyverse: Easily Install and Load 'Tidyverse' Packages.** R package version 1.1.1. <https://CRAN.R-project.org/package=tidyverse>

Hadley Wickham (2011). **The Split-Apply-Combine Strategy for Data Analysis.** Journal of Statistical Software, 40(1), 1-29. URL <http://www.jstatsoft.org/v40/i01/>.

Hao Zhu (2017). **kableExtra: Construct Complex Table with 'kable' and Pipe Syntax.** R package version 0.5.2. <https://CRAN.R-project.org/package=kableExtra>

MEEK, D. W.; HATFIELD, J. L. **Data quality checking for single station meteorological databases.** Agricultural and Forest Meteorology, p. 90-91, 1994.

KANNENBERG, C. **DESAGREGAÇÃO DA TEMPERATURA DO AR DA ESCALA DIÁRIA PARA HORÁRIA PARA O SUL DO BRASIL**, 2016, p. 24. Tese (TCC para obtenção do grau de Bacharel em Meteorologia) – Universidade Federal de Santa Maria, RS, 2016.

Robert J. Hijmans (2016). **raster: Geographic Data Analysis and Modeling.** R package version 2.5-8. <https://CRAN.R-project.org/package=raster>

Stefan Milton Bache and Hadley Wickham (2014). **magrittr: A Forward-Pipe Operator for R.** R package version 1.5. <https://CRAN.R-project.org/package=magrittr>

VEJEN et al. **Quality control of meteorological observations: Automatic Methods Used in the Nordic Countries.** Norwegian Meteorological Institute, p. 11, 2002.

WMO. **GUIDE ON THE GLOBAL DATA-PROCESSING SYSTEM**, p. VI.21, 1993.

WMO. **Guide to Meteorological Instruments and Methods of Observation**, p. 1, 2012.

Yihui Xie (2017). **knitr: A General-Purpose Package for Dynamic Report Generation in R.** R package version 1.17.

ZAHUMENSKÝ, I. **Guidelines on Quality Control Procedures for Data from Automatic Weather Stations.** World Meteorological Organization, p. 3-4, 2004.