



**UNIVERSIDADE DE BRASÍLIA**

**FACULDADE DE AGRONOMIA E MEDICINA VETERINÁRIA**

**AVALIAÇÃO DA QUALIDADE DOS DADOS COLETADOS PELA  
ESTAÇÃO AGROMETEOROLÓGICA AUTOMÁTICA DA  
FAZENDA ÁGUA LIMPA**

**Luana Costa Tavares**

**Brasília, DF**

**Novembro de 2018**



**UNIVERSIDADE DE BRASÍLIA**  
**FACULDADE DE AGRONOMIA E MEDICINA VETERINÁRIA**

**Luana Costa Tavares**

**AVALIAÇÃO DA QUALIDADE DOS DADOS COLETADOS PELA  
ESTAÇÃO AGROMETEOROLÓGICA AUTOMÁTICA DA  
FAZENDA ÁGUA LIMPA**

Monografia apresentada à Faculdade de Agronomia e Medicina Veterinária da Universidade de Brasília – UnB, como parte das exigências do curso de Graduação em Agronomia, para a obtenção do título de Engenheira Agrônoma.

Orientador: Prof. Dr. **SELMA REGINA MAGGIOTTO**

**Brasília, DF**

**Novembro de 2018**



**UNIVERSIDADE DE BRASÍLIA**  
**FACULDADE DE AGRONOMIA E MEDICINA VETERINÁRIA**

**Termo de Aprovação**

Autor: Luana Costa Tavares

Título: Avaliação da qualidade dos dados coletados pela estação Agrometeorológica automática da Fazenda Água Limpa

Monografia de conclusão do Curso de Graduação em Agronomia apresentada à Faculdade de Agronomia e Medicina Veterinária da Universidade de Brasília – UnB.

Aprovado em: Brasília, 29 de novembro de 2018.

Banca examinadora:

---

Profa. Dra. Selma Regina Maggiotto  
Universidade de Brasília – UnB  
Orientadora

---

Prof. Dr. João José da Silva Júnior  
Universidade de Brasília – UnB  
Examinador

---

Profa. Dra. Michelle Souza Vilela  
Universidade de Brasília – UnB  
Examinadora

## AGRADECIMENTO

Primeiramente gostaria de agradecer aos meus pais, Davi e Francisca, que apesar de serem analfabetos e não possuir muitas condições financeiras, sempre me deram a base necessária para os meus estudos, aos meus irmãos e família, ao meu marido Jonathan de Jesus por todo o companheirismo e apoio.

A Deus por me dar saúde e força em todos os momentos da minha vida.

Aos colegas da graduação pelo apoio na vida acadêmica e pelos momentos inesquecíveis vividos dentro e fora da UnB, em especial à Ana Clara Barbosa, Giordana Cruz, Karen Pereira, Jhon Kenedy, Letícia Vieira, Lucas Simioni, Lucas Vitória, Lorrany Silva, Marta Oliveira, Thaís França e Thamires Dutra, que serão os melhores companheiros de profissão.

A todos os professores da FAV-UnB com que tive oportunidade de estudar, aprender e trabalhar. Entre eles, o Prof. Dr. Tairone Leão pela ajuda em linguagem de programação estatísticas no software R.

Também agradeço e dedico esse trabalho à minha orientadora Selma Regina Maggiotto por toda a sua competência, atenção, incentivo, apoio e experiências transferidos na vida acadêmica, pessoal e na elaboração desse trabalho.

TAVARES, L. C. Avaliação da qualidade dos dados coletados pela estação Agrometeorológica automática da Fazenda Água Limpa. 49 p. Monografia (Graduação em Agronomia) – Universidade de Brasília – UnB.

## RESUMO

As estações meteorológicas são uma importante ferramenta de monitoramento climático, mas é necessário estar atento à qualidade dos dados coletados e a possíveis erros e falhas na coleta de dados, que interferem na qualidade final do banco de dados climáticos do local. O objetivo desse trabalho foi verificar a qualidade do banco de dados climáticos da estação Agrometeorológica pertencente à FAV/UnB no período de abril de 2000 a abril de 2018, elaborando a análise no *software* R. A primeira etapa consistiu na análise exploratória dos dados diários, com o auxílio de gráficos, e a segunda etapa foi a análise da qualidade dos dados pela aplicação dos testes de limite (*range*) e de degrau (*step*) em dados horários. Os testes atribuem valores de sinalização aos dados (*flags*), indicando-os como aprovado (*flag* = 0), reprovado (*flag* = 3) e ausente (*flag* = 7). Com os gráficos, foi possível verificar a flutuação diária das variáveis, e observar os períodos com falhas na coleta de dados. Nos testes de qualidade foram observados dados horários reprovados (*flag* = 3), indicando problemas com os equipamentos, e a ausência de dados em alguns períodos (*flag* = 7), explicitando os períodos em que a estação apresentou problemas com o coletor de dados ou com o fornecimento de energia. As análises utilizando programação no *software* R apresentou-se como uma boa alternativa para verificar a qualidade dos dados.

**PALAVRAS-CHAVE:** banco de dados; *software* R; análise exploratória; testes de qualidade de dados.

## LISTA DE FIGURAS

Figura 1. Estação agrometeorológica automática Hobeco, do Laboratório de Agroclimatologia – FAV/UnB, localizada na Fazenda Água Limpa. ....	22
Figura 2. Script 1: Comandos utilizados para executar o teste de range em todos os elementos climáticos do banco de dados da estação meteorológica da FAL/UnB de 2000 a 2018.....	23
Figura 3. Script 2: Comandos utilizados para executar o teste de step para temperatura e umidade relativa do ar do banco de dados da estação meteorológica da FAL/UnB do ano 2000 a 2018.....	24
Figura 4. Flutuação das temperaturas médias, máximas e mínimas diárias durante o período de coleta de dados de abril de 2000 a dezembro de 2010, pela estação automática da FAL (FAV/UnB). ....	28
Figura 5. Flutuação das temperaturas médias, máximas e mínimas diárias durante o período de coleta de dados de janeiro de 2011 a abril de 2018 pela estação automática da FAL (FAV/UnB).....	29
Figura 6. Frequência de ocorrência das temperaturas mínimas, média e máximas diárias nos 18 anos de coleta de dados climáticos realizados na estação agrometeorológica da FAL (FAV/UnB).....	29
Figura 7. Flutuação dos valores diários de umidade relativa médias coletadas no período de abril de 2000 a dezembro de 2005 pela estação automática da FAL (FAV/UnB).31	31
Figura 8. Flutuação dos valores diários de umidade relativa médias, máximas e mínimas diárias durante os anos de janeiro de 2006 a abril de 2018 coletados pela estação automática da FAL (FAV/UnB). ....	32
Figura 9. Precipitação acumulada anual para o período de abril de 2000 a abril de 2018 coletado pela estação meteorológica da FAV/UnB. ....	33
Figura 10. Precipitação pluviométrica acumulada mensal (mm), entre abril de 2000 e abril de 2018, coletado pela estação meteorológica da FAL (FAV/UnB). ....	34
Figura 11. Radiação solar global diária ( $\text{MJ m}^{-2} \text{d}^{-1}$ ) incidente na superfície e radiação solar incidente no topo da atmosfera para período de abril de 2000 a dezembro de 2010, coletado pela estação meteorológica da FAL (FAV/UnB). ....	36
Figura 12. Radiação solar global diária ( $\text{MJ m}^{-2} \text{d}^{-1}$ ) incidente na superfície e radiação solar incidente no topo da atmosfera para período de janeiro de 2011 a abril de 2018, coletado pela estação meteorológica da FAL (FAV/UnB). ....	37

Figura 13. Frequência de ocorrência da velocidade ( $\text{m s}^{-1}$ ) e direção do vento ( $^{\circ}$ ) para o período de abril de 2000 a dezembro de 2010, coletado pela estação meteorológica da FAL (FAV/UnB). .....	38
Figura 14. Frequência de ocorrência da velocidade ( $\text{m s}^{-1}$ ) e direção do vento ( $^{\circ}$ ) para o período de abril de 2000 a abril de 2018, coletado pela estação meteorológica da FAL (FAV/UnB). .....	39

**LISTA DE TABELAS**

Tabela 1. Códigos utilizados na classificação dos dados, conforme apresentado pela Organização Mundial de Meteorologia. ....	16
Tabela 2. Códigos sinalizadores da etapa de validação dos dados no controle de qualidade de um banco de dados meteorológicos. ....	17
Tabela 3. Códigos de sinalização da qualidade dos dados utilizados em Pittigliani (2000). .....	17
Tabela 4. Descrição dos instrumentos utilizados nas estações meteorológicas automáticas da FAL no período de abril de 2000 a abril de 2018. ....	21
Tabela 5. Descrição de alguns comandos de programação utilizados na execução dos testes range e step. ....	25
Tabela 6. Temperatura média ( $T_{med}$ ) anual e valores médios e extremos da temperatura mínima ( $T_{min}$ ) e temperatura máxima ( $T_{max}$ ) anuais coletados pela estação automática da FAL (FAV-UnB) para o período de abril de 2000 a abril de 2018.....	26
Tabela 7. Umidade relativa média ( $UR_{med}$ ) anual e valores médios umidade relativa mínima ( $UR_{min}$ ) e umidade relativa ( $UR_{max}$ ) anuais coletados pela estação automática da FAL (FAV-UnB) para o período de abril de 2000 a abril de 2018.....	30
Tabela 8. Normais Climatológicas da precipitação no período de 1981 a 2010 para a região de Brasília- DF.....	35
Tabela 9. Número de vezes que cada valor sinalizador (flag) foi observado no teste de limites (range), quando aplicado aos dados horários dos diferentes elementos climáticos do banco de dados da FAL, FAV/UnB, compreendendo o período de abril de 2000 a abril de 2018.....	41
Tabela 10. Número de vezes que cada valor sinalizador ( <i>flag</i> ) foi observado no teste de degrau ( <i>step</i> ), quando aplicado aos dados horários de temperatura e umidade relativa do ar do banco de dados da FAL, FAV/UnB, no período de 2000 a 2018. ....	45



## SUMÁRIO

1. INTRODUÇÃO .....	9
2. REVISÃO DE LITERATURA.....	11
2.1- Importância da coleta de dados meteorológicos .....	11
2.2- Estações meteorológicas e agrometeorológicas.....	13
2.3- Controle de qualidade dos dados coletados .....	15
2.4- Objetivo .....	19
2.4-a) Objetivo geral .....	19
2.4-b) Objetivos específicos .....	19
3. MATERIAL E MÉTODOS .....	20
3.1- Caracterização do local e das estações agrometeorológicas .....	20
3.2- Análise exploratória dos dados diários .....	22
3.3- Análise de qualidade dos dados horários .....	23
4. RESULTADOS E DISCUSSÃO .....	26
4.1- Análise exploratória.....	26
4.2- Análise de qualidade .....	39
5. CONCLUSÕES .....	46
CONSIDERAÇÕES FINAIS .....	47
6. REFERÊNCIAS.....	48

## 1. INTRODUÇÃO

O clima é definido por interações atmosféricas, terrestres e oceânicas, que exercem influência sobre a vida na Terra e criam um conjunto de características específicas para cada local e período de tempo, através de combinações dos parâmetros climáticos, como temperatura, precipitação, pressão atmosférica, umidade e vento (WORLD METEOROLOGICAL ORGANIZATION, 2011).

O interesse em entender como se comporta e como se varia o clima ao longo do espaço de tempo e região, foi e é importante para aprender como agir e conviver com as mudanças do ambiente em meio às variações. Referências climáticas datadas do início das civilizações foram encontradas, quando o homem sentiu necessidade de se proteger contra as tempestades, calor, frio e vento, além de necessidade de cultivo quando deixou de ser nômade. De 1800 a.C. foram encontradas escritas hindus; posteriormente, poemas na Grécia antiga e estudos elaborados por volta de 400 a.C. por Hipócrates e em 350 a.C. por Aristóteles, escritas no Antigo Testamento da Bíblia judaico-cristã, placas de barro na antiga Babilônia com conhecimento empírico de previsão do tempo (LIMA; SANTOS, 2009; WORLD METEOROLOGICAL ORGANIZATION, 2011). Esses registros são importantes para apontar como os fenômenos climáticos interferiam na estruturação do local e como se enfrentavam os diversos cenários climáticos, já que esse está sempre passando por alterações ao longo do tempo e espaço.

Além disso, o registro dos dados meteorológicos se faz necessário para entender como ocorre o progresso na obtenção de conhecimento e para a construção um banco de observações que auxilie cada vez no entendimento de como seus parâmetros se comportam. Até o século XX, os fenômenos climáticos, incluindo os mais frequentes, como enchentes decorrentes de chuvas intensas ou degelo, ondas de calor, furacões e etc., eram considerados como manifestações da natureza que não podiam ser previstas (MAURER; SILVA, 2017).

Os elementos mais comumente usados em climatologia são temperatura do ar (incluindo máxima e mínima), precipitação (chuva e todos os tipos de deposição úmida, como granizo, orvalho, geada, neve e neblina de precipitação), umidade do ar, movimento atmosférico (velocidade e direção do vento), pressão atmosférica, evaporação e luz do sol (radiação e número de horas de brilho solar) (WORLD METEOROLOGICAL ORGANIZATION, 2011). Estes são usados como componentes da previsão do tempo, que permitem fazer uma descrição detalhada de ocorrências das situações atmosféricas

atual do local e fazer previsões/ tendências futuras e assim auxiliar nas diversas atividades que precisam dessas informações (PEREIRA; ANGELOCCI; SENTELHAS, 2002).

O arquivamento de informações relacionadas a eventos catastróficos e observações diárias começou a fazer parte de revistas, periódicos e arquivos meteorológicos a partir da metade do século XX. Revistas específicas para a coleta e retenção de informações climatológicas foram criadas, e com a criação da OMM (Organização Meteorológica Mundial) em 1950, houve uma padronização das informações, coleta e arquivamento contínuos de dados a partir da utilização de instrumentos específicos (WORLD METEOROLOGICAL ORGANIZATION, 2011).

As informações climáticas são importantes para todas as áreas de trabalho e em todas as atividades. Engloba desde a saúde, segurança e bem-estar humano até o comportamento e desenvolvimento dos vegetais, animais terrestres e aquáticos. Envolve os ecossistemas, energia, indústria, produção e distribuição de bens, projeto e construção de engenharia, transporte, turismo, recursos hídricos e gestão de desastres, pesca e desenvolvimento costeiro, beneficiando de forma direta e indireta todos esses setores da sociedade (PADILHA; VANZELLA; CASTRO, 2010; WORLD METEOROLOGICAL ORGANIZATION, 2011).

Como as estações meteorológicas coletam todas as informações necessárias para auxiliar no planejamento e execução das atividades, faz-se necessário de haver um banco de dados com os parâmetros medidos e que esses tenham um conjunto de observações confiáveis, necessitando que no processo de coleta de dados, antes da sua divulgação/utilização submetê-los a uma análise de qualidade (FERREIRA et al., 2006).

Desta forma, o presente trabalho começa com uma apresentação sobre a coleta de dados, seguindo por sua análise de qualidade e posteriormente uma investigação sobre os dados que compõem o banco de dados de uma estação agrometeorológica no Distrito Federal, localizada na Fazenda Água Limpa.

## 2. REVISÃO DE LITERATURA

### 2.1- Importância da coleta de dados meteorológicos

As áreas urbanas e rurais possuem interações importantes com a climatologia. Nas áreas urbanas, a grande concentração de pessoas e a alta densidade de infraestrutura, fazem com que seja necessário estar atento ao monitoramento dos elementos climáticos, mas também a outros elementos, como, a poluição do ar, que podem causar um grande impacto no local. Os impactos podem ser provenientes de chuvas fortes, tempestades severas acompanhadas de raios, granizos, ventos fortes, entre outros, que podem causar um impacto significativo na rotina e na infraestrutura (WORLD METEOROLOGICAL ORGANIZATION, 2011). Na área rural, além de sofrer com os mesmos impactos que ocorrem nas cidades, os elementos climáticos podem prejudicar ou beneficiar todas as práticas executadas na agricultura, silvicultura, planejamento, manejo e uso da terra (PEREIRA; ANGELOCCI; SENTELHAS, 2002) por isso se faz necessário um monitoramento contínuo, permitindo uma execução mais segura das atividades.

Na agricultura, a meteorologia agrícola é de fundamental importância para contribuir com o produtor nos momentos de tomada de decisão, já que os elementos climáticos interferem no planejamento e execução de diversas atividades, determinando o nível da produtividade em um certo período. Todas as áreas agrícolas são influenciadas: irrigação, para a definição de quando, quanto e como irrigar; nas áreas de fitopatologia e entomologia, para a compreensão da intensidade de ocorrência e definição de formas de controle de pragas e doenças; na produção e fisiologia vegetal e animal; no preparo e conservação do solo; na colheita, transporte e armazenamentos dos produtos vegetais; no monitoramento e previsão de ocorrências de intempéries, entre outros. As consequências dessas interferências, podem promover diversas situações que desencadeiem grandes impactos sociais e econômicos (PEREIRA; ANGELOCCI; SENTELHAS, 2002).

Assim, os conhecimentos de meteorologia agrícola são necessários não somente para o aumento da produtividade, como também para a preservação dos recursos naturais e melhoria da qualidade dos produtos. Além de serem usados para tomadas de decisões a curto prazo, os planejamentos das atividades de longo prazo devem ser baseados em séries de dados históricos ou tendências climáticas, de forma a minimizar as influências diretas no desenvolvimento das atividades (PEREIRA; ANGELOCCI; SENTELHAS, 2002;

SENTELHAS; MONTEIRO, 2009; WORLD METEOROLOGICAL ORGANIZATION, 2011)

Em um cenário de mudanças constantes, os sistemas agrícolas enfrentam condições adversas de clima desfavorável, oscilações de produção, comercialização, entre outros fatores, para se manter em estado de produtividade com algum resultado. Desta forma, é necessário aumentar a resiliência agrícola para minimizar os efeitos adversos do clima sobre o processo de produção (SENTELHAS; MONTEIRO, 2009). Diversificação de culturas, cultivares, época de plantio, densidade de plantio, utilização de cultivo mínimo e plantio direto, controle integrado de pragas e doenças, uso racional da irrigação, uso adequado do terreno, etc., são exemplos de estratégias que ajudam a manter o ambiente em estado de produção e auxiliam na redução de riscos e perdas, além de permitir que alguns ambientes recuperem o seu estado de produção.

Os principais elementos climáticos que afetam o crescimento, desenvolvimento e produtividade das culturas são: chuva, temperatura do ar, radiação solar, fotoperíodo, umidade do ar e do solo e velocidade do vento (SENTELHAS; MONTEIRO, 2009).

A temperatura afeta vários processos nas plantas, dentre elas a respiração, a transpiração, o repouso vegetativo e duração das fases fenológicas. Já a chuva afeta o crescimento e desenvolvimento da cultura. A radiação solar interfere no processo de fotossíntese e conseqüentemente no processo de acúmulo de carboidratos e biomassa. A umidade influencia no processo de transpiração quando o ambiente está muito quente, e conseqüentemente na quantidade e periodicidade de irrigação, além de uma maior ou menor ocorrência de ataque de microrganismos, enquanto que o vento, dependendo da velocidade, pode prejudicar as culturas quando causa acamamento das plantas, queda de folhas, quebra de galhos, entre outros danos mecânicos, ou beneficiar as culturas, quando renova o CO<sub>2</sub> das plantas em vento mais leves (SENTELHAS; MONTEIRO, 2009). Se todos esses parâmetros estiverem dentro dos limites tolerados pelas culturas, essas poderão se desenvolver, crescer e produzir de forma adequada, utilizando todos os recursos de forma mais eficiente.

## 2.2- Estações meteorológicas e agrometeorológicas

A coleta de dados meteorológicos e agrometeorológicos é feita em estações de monitoramento, que coletam dados destinados a estudos climáticos e de previsão do tempo. Ambos os tipos de estações, podem conter os mesmos equipamentos de mensuração dos elementos climáticos, e o que vai diferenciar uma da outra é a finalidade e/ou propósito da coleta dos dados e o número de elementos climáticos observados (PEREIRA; ANGELOCCI; SENTELHAS, 2002; MAGGIOTTO et al., 2007; WORLD METEOROLOGICAL ORGANIZATION, 2011).

As estações meteorológicas são classificadas em “primeira classe” quando possuem instrumentos para observar todos os elementos meteorológicos, possibilitando uma caracterização do clima local; em “segunda classe” quando se observam os principais elementos de caracterização agrometeorológica, e em “terceira classe” quando se observam somente os elementos de balanço hídrico no solo (PEREIRA et al. 2002).

As estações agrometeorológicas coletam dados de interesse às atividades agrícolas, gerando informações que auxiliam no planejamento de atividades diárias. Podem ser consideradas como de “primeira” ou “segunda” classe, da mesma forma que as estações meteorológicas, dependendo da quantidade de variáveis estudadas (PEREIRA; ANGELOCCI; SENTELHAS, 2002). Essas estações podem ser instaladas em pequenas propriedades rurais e coletar somente dados de temperatura do ar e pluviosidade ou podem fazer uma caracterização do clima do local (PEREIRA; ANGELOCCI; SENTELHAS, 2002; MAGGIOTTO et al., 2007). Quando a estação agrometeorológica coleta somente dados de temperatura e precipitação, pode ser chamada também de Estação Climatológica Ordinária, de acordo com a OMM (WORLD METEOROLOGICAL ORGANIZATION, 2011).

Outra forma de classificação das estações é baseada no sistema de coleta de dados, que pode ser do tipo convencional ou automático (PEREIRA; ANGELOCCI; SENTELHAS, 2002; MAGGIOTTO et al., 2007). As estações meteorológicas convencionais (EMC) necessitam da presença diária de um observador, pois os parâmetros meteorológicos mensurados precisam ser lidos e anotados por ele a cada intervalo de tempo, de forma contínua e uniforme, para que se tenha um banco de dados homogêneo. Os sensores que compõem esse tipo de estação são isolados uns dos outros e necessitam de leitura direta como os termômetros, termohigrógrafo, pluviômetro,

actinógrafo, heliógrafo, anemógrafo e barômetros (MAGGIOTTO et al., 2007). O banco de dados de uma estação convencional é físico, isto é, é formado pelos formulários de leitura dos parâmetros medidos, que devem ser organizados de forma adequada e com cuidado para não haver o extravio de dados que o compõe.

As estações meteorológicas automáticas (EMA), também chamadas de plataforma de coleta de dados (PCD), fazem as medições e o armazenamento dos dados automaticamente. São compostas pelos subsistemas de coleta, controle e armazenamento de dados climáticos, que podem medir os elementos climáticos em alta frequência e armazená-los em intervalos de tempo de acordo com a programação do usuário. As variáveis coletadas nessas estações são armazenadas em um coletor de dados (*datalogger*). Esse coletor de dados é uma unidade de memória que comporta e armazena todas as informações coletadas (PADILHA; VANZELLA; CASTRO, 2010; INMET, 2011). As PCD's são desenvolvidas para medidas mais precisas e para proporcionar armazenamento de dados meteorológicos por um longo período de tempo e a um custo relativamente baixo (PADILHA; VANZELLA; CASTRO, 2010).

O Brasil possui um bom número de estações meteorológicas, públicas e privadas, distribuídas em diversos locais estratégicos representando boa parte do país. O conjunto dessas estações é denominado de rede de estações meteorológicas. As redes de estações possuem um papel relevante na formação de um banco de dados mais coeso e que possam contribuir significativamente em um cenário local, agregando mais informações aos dados já existentes, tornando-os mais confiáveis. As redes de estações que mais se destacam no Brasil são administradas pelo INMET (Instituto Nacional de Meteorologia), que é composta por mais de 500 estações automáticas, INPE (Instituto Nacional de Pesquisa Espaciais), IAC (Instituto Agrônomo de Campinas), IAPAR (Instituto Agrônomo do Paraná), SIMEPAR (Sistema Meteorológico do Paraná) e EMBRAPA (Empresa Brasileira de Pesquisa Agropecuária) (INMET, 2011).

É importante que as estações e suas redes, façam um arquivamento de seus dados, já que essas observações permitem acesso as informações descritivas do clima naquele momento histórico, como ele está se comportando atualmente e como é sua previsão para o cenário futuro, de forma a compreender o cenário de mudanças climáticas globais (WORLD METEOROLOGICAL ORGANIZATION, 2011).

Para a implantação de uma estação agrometeorológica e validação de seus dados é necessário estar dentro de padrões internacionais, definidos pela OMM, para garantir a qualidade e homogeneidade dos dados (SHAFER et al., 2000; WORLD METEOROLOGICAL ORGANIZATION, 2011). Desta forma, deve-se adequar os parâmetros medidos de velocidade e direção do vento à altura de 2 m; e temperatura e umidade relativa a 1,5 m, para equivaler com as observações de outras estações já existentes. A torre com os equipamentos e o abrigo meteorológico devem ficar próximos ao centro de um recinto de 10 m × 10 m, cercada por uma cerca de 1,3 m de altura para proteger a área dos animais e da atividade humana nas proximidades (SHAFER et al., 2000). Além disso, o local de instalação deve ser em terrenos mais planos, a área deve ser bem exposta principalmente no sentido Leste-Oeste e deve-se evitar a proximidade de florestas, árvores, construções e qualquer outro tipo de elemento que possa provocar alguma interferência de sombra ou de condições atmosféricas locais (PEREIRA; ANGELOCCI; SENTELHAS, 2002).

### 2.3- Controle de qualidade dos dados coletados

O objetivo do controle de qualidade de dados meteorológicos é verificar se os valores observados estão condizentes e representativos para o momento das medições. Os registros de dados climáticos podem conter dados não verdadeiros ou mesmo períodos com dados não registrados, devido a falhas humanas, de equipamentos ou interrupção no fornecimento de energia. Os dados devem, portanto, ser submetidos a uma análise para verificar sua qualidade e, quando possível, preencher as falhas (GRIGOLETTI; FLORES; SANTOS, 2016).

Muitas vezes as PCD's são instaladas em locais isolados, e estão sujeitas a sofrer com inúmeros fatores que podem prejudicar a qualidade dos dados coletados, tais como: mau funcionamento dos sensores, problemas de calibração, falhas no fornecimento de energia, erros de transmissão, vandalismo. Estes aspectos negativos podem ser associados à falta da visita diária de um observador, traduzindo a importância da qualificação dos dados. O controle de qualidade irá evitar a difusão de dados incorretos e também auxiliar na tarefa de manutenção de uma rede de estações (FERREIRA et al., 2006; PADILHA; VANZELLA; CASTRO, 2010).



Para o controle de qualidade é necessário o gerenciamento dos dados feito pelo uso de um conjunto de ferramentas e procedimentos, que permitem sua edição, mas mantém sua integridade, para posteriormente serem novamente armazenado de forma a satisfazer seus usos futuros. Além disso, a avaliação dos dados deve incluir alguns testes, criados indicar a qualidade dos mesmos.

A OMM sugere o uso de um código para a sinalização dos dados coletados, composto por dois dígitos, o primeiro indicando o tipo de dado, se original, corrigido, reconstruído ou calculado (Tabela 1), e o segundo dígito indicando o estágio de validação do dado (Tabela 2). Um dos testes de qualidade é a comparação do dado com limites de tolerância superiores e inferiores em todos os elementos climáticos do banco de dados. Esses limites podem ser estabelecidos estatisticamente, com base na inspeção dos dados ou com base nos valores históricos. Outro teste que pode ser feito é o teste de consistência dos dados, que verifica a consistência interna, temporal, espacial e sumarização dos dados. O teste de consistência interna verifica a relação das observações dos elementos coletados com o seu padrão de variação; o teste de consistência temporal verifica a variação do elemento em um intervalo de tempo; o teste de consistência espacial compara as observações climáticas de estação com outras estações da mesma área no mesmo período de tempo, e o teste de consistência de sumarização verifica se há alguma inconsistência numérica de soma e média dos valores calculados presentes no banco de dados (WORLD METEOROLOGICAL ORGANIZATION, 2011).

Tabela 1. Códigos utilizados na classificação dos dados, conforme apresentado pela Organização Mundial de Meteorologia.

<b>Código de tipo de dado</b>	<b>Significado</b>
0	Dado Original
1	Dado Corrigido
2	Dado Reconstruído (interpolação/ estimacão)
3	Dado Calculado

\*Fonte: adaptado WORLD METEOROLOGICAL ORGANIZATION (2011)

Tabela 2. Códigos sinalizadores da etapa de validação dos dados no controle de qualidade de um banco de dados meteorológicos.

<b>Código de estágio de validação do dado</b>	<b>Significado</b>
1	Dado Perdido (não observado/medida não feita)
2	Dado eliminado
3	Dado não controlado
4	Dado possivelmente errado
5	Dado provavelmente errado
6	Dado validado após análise, mas que ainda aceita modificação
7	Dado validado

\*Fonte: adaptado WORLD METEOROLOGICAL ORGANIZATION (2011)

Pittigliani (2000), buscando uma forma de avaliar a qualidade dos dados hidrometeorológicos de uma rede de estações adaptou o sistema proposto (Tabela 3) e concluiu que foi possível fazer a análise de qualidade e identificar dados bons e dados com problemas, através dos testes *range*, *step*, *persistence* e *spatial*. O teste *range* inspecionou os limites climatológicos máximos e mínimos de cada elemento de acordo com a resolução do equipamento de coleta; o teste *step* realizou uma comparação entre dois dados consecutivos, no qual verifica-se para cada observação se a diferença do dado anterior é menor ou maior que um limite pré- estabelecido; o teste *persistence* verificou a constância de um mesmo valor em uma série temporal; e o teste *spatial* realizou uma análise espacial dos dados de um mesmo horário com outras estações da mesma região.

Tabela 3. Códigos de sinalização da qualidade dos dados utilizados em Pittigliani (2000).

<b>Código de sinalização do dado</b>	<b>Significado</b>
0	Dado Aprovado
1	Dado Suspeito
2	Dado com possível erro
3	Dado Reprovado
4	Dado não verificado
6	Não instrumento instalado no local
7	Dado não coletado

Outra forma de avaliação de qualidade de dados meteorológicos é utilizando programas (*softwares*) já disponíveis, como o MADIS (*Meteorological Assimilation Data Ingest System*). Este é um programa que pode ser obtido gratuitamente após o

estabelecimento de uma parceria com o Laboratório de Sistemas de Previsão (FSL – *Forecast Systems Laboratory*), que é vinculado à NOAA (*National Oceanic and Atmospheric Administration*). O MADIS é um sistema de assimilação de dados meteorológicos, que coleta, controla a qualidade e distribui os dados de várias estações meteorológicas automáticas, de forma que é utilizado para controlar mais de 15000 estações nos Estados Unidos (MILLER; BARTH, 2003; FERREIRA et al., 2006; PADILHA; VANZELLA; CASTRO, 2010)

O controle de qualidade realizado pelo MADIS é feito em 3 níveis: o nível 1 consiste nas verificações de validade ou limites absolutos, que estabelece os valores máximos e mínimos para a variável; o nível 2 consiste nas verificações de consistência interna, exigindo leituras de diferentes instrumentos de uma mesma estação; e o nível 3, que consiste na verificação da consistência espacial. Todos os dados são avaliados nos seus devidos níveis, e se as observações não estiverem dentro dos limites estabelecidos, recebem uma sinalização (*flag*), identificando-as como suspeitas (FERREIRA et al., 2006; PADILHA, 2010).

Ferreira et al. (2006) usaram este *software* de qualificação na rede de estações no INPE, vinculando diretamente seus dados à NOAA, devido à concessão do programa fonte ao CPTEC/INPE (Centro de Previsões e Estudos Climáticos). Padilha (2010) avaliou os dados da estação meteorológica automática do Instituto Federal de Santa Catarina (IFSC), usando a licença do LEPTEN/UFSC (Laboratórios de Engenharia de Processos de Conversão e Tecnologia de Energia/ Universidade Federal de Santa Catarina) instalado num *cluster* computacional do IFSC (Instituto Federal de Santa Catarina).

Como forma alternativa a esses programas de análise de qualidade que são restritos, iremos utilizar neste trabalho o ambiente de programação R que utiliza linguagem computacional aberta e uma grande quantidade de procedimentos e inúmeras bibliotecas (*packages*) para executar várias funções estatísticas de manipulação, análise e geração de gráficos e dados, para realizar a análise de qualidade do banco de dados através de diversas funções (*scripts*), combinando e adaptando as formas de análises usadas pela OMM, pelo software MADIS, e por Pittigliani (2000) atribuindo códigos de sinalização.

## 2.4- Objetivo

### 2.4-a) Objetivo geral

Devido a importância que representa a coleta e qualificação dos dados meteorológicos, o objetivo desse trabalho foi avaliar os dados coletados pela estação Agrometeorológica da Fazenda Água Limpa (FAL) pertencente à Faculdade de Agronomia e Medicina Veterinária da Universidade de Brasília (FAV/UnB), para que seus dados publicados na *web* página da faculdade possam ser utilizados de forma confiável para os diversos fins na agricultura e pesquisa.

### 2.4-b) Objetivos específicos

Desenvolver um *script* no Sistema Estatístico R para realizar:

- a) Análise exploratória dos dados diários que foram publicados de 2000 a 2018, utilizando gráficos de dispersão, barras e histogramas;
- b) Testar a qualidade dos dados horários:
  - I. Teste de *range* (limites de leitura) com base na resolução dos equipamentos;
  - II. Teste de *step* (teste de “degrau” ou “variação horária”) com base na variação estabelecida;
- c) Desenvolver um *script* de fácil execução no Sistema Estatístico R, que permitirá a identificação das possíveis falhas do sistema de coleta de dados antes de serem publicados na página da FAV;
- d) Identificar quais são os anos que possuem os registros de dados menos confiáveis;

### 3. MATERIAL E MÉTODOS

#### 3.1- Caracterização do local e das estações agrometeorológicas

A estação agrometeorológica automática pertencente ao Laboratório de Agroclimatologia da Faculdade de Agronomia e Medicina Veterinária (FAV) e está instalada na Fazenda Água Limpa, pertencente à Universidade de Brasília – FAL/UnB, localiza-se na latitude 15°56' S, longitude 47°56' W e altitude 1080 m, na região do Cerrado do Distrito Federal.

A coleta de dados meteorológicos ocorre em frequência horária, desde de abril do ano 2000, e a compilação dos dados diários são divulgados ao público. Nesse período de coleta de dados, o laboratório foi composto por dois modelos de estação automática: a estação da marca Campbell, que coletou dados do ano 2000 a 2010 e a estação da marca Hobeco, que coletou os dados desde 2011 até o presente momento. A diferença de uma estação para outra é mínima, relacionada à resolução de leitura de alguns equipamentos, como se observa na Tabela 4. O equipamento que possui diferença mais significativa de uma estação para outra, é o anemômetro, já que o modelo Young da estação Campbell tem uma velocidade do vento mínima para funcionamento igual a 1 m/s, enquanto que o modelo Vaisala da estação Hobeco precisa de uma velocidade menor do vento, de 0,4 m/s.

No período de funcionamento da estação Campbell, no ano de 2005, de julho a dezembro e no ano de 2006, de 28 de novembro a 31 de dezembro de 2006, há falhas na coleta de dados. O mesmo ocorre na estação Hobeco, no período de 16 de abril a 14 de julho de 2011. Esses períodos sem dados ocorreram devido ao mal funcionamento dos equipamentos, que logo após a sua manutenção voltaram a coletar os dados normalmente.

A estação agrometeorológica é composta pelos seguintes equipamentos: termohigrômetro, anemômetro, piranômetro e pluviômetro (Figura 1). Os dados coletados se destinam a dar suporte a trabalhos de pesquisa que são desenvolvidos no local e fornecer dados a produtores da região. Os dados diários estão disponíveis na página da web: <http://www.fav.unb.br/laboratorios/laboratorio-de-agroclimatologia> .

Tabela 4. Descrição dos instrumentos utilizados nas estações meteorológicas automáticas da FAL no período de abril de 2000 a abril de 2018.

Equipamento	Modelo	Fabricante	Parâmetro Medido	Resolução	Intervalo de coleta	Erro do equipamento	Período de coleta
Estação: Campbell Sci. Inc.							
Termohigrômetro	HMP145	Vaisala	Umidade Relativa Temperatura	0 a 100% -40 a +60°C	1 segundo	± 2% ± 1%	2000-2010
Anemômetro	05103	Young	Velocidade Direção do vento	0 a 100 m/s 0 a 360 °	1 segundo	± 0.3 m/s ± 3°	2000-2010
Piranômetro	LI200X	Licor	Radiação solar global	-	1 segundo	± 2%	2000-2010
Pluviômetro	TB4-L	Campbell Sci. Inc	Precipitação	0 a 700 mm/h	1 segundo	± 2%	2000-2010
Estação: Maws 110 da Hobeco							
Termohigrômetro	HMP155	Vaisala	Umidade Relativa Temperatura	0 a 100% -50 a +60° C	10 segundos	± 1% ± 1%	A partir de 2011
Anemômetro	WMS302	Vaisala	Velocidade Direção do vento	0 a 75 m/s 0 a 360 °	1 segundo	± 0.3 m/s ± 3°	A partir de 2011
Piranômetro	SP Lite	Kipp-Zonen	Radiação solar global	0 a 1500 W/m <sup>2</sup>	10 segundos	± 2%	A partir de 2011
Pluviômetro	QMR102	Hobeco	Precipitação	0 a 254 mm/h	60 segundos	-	A partir de 2011

Fonte: Manuais dos sensores - adaptado pela autora.

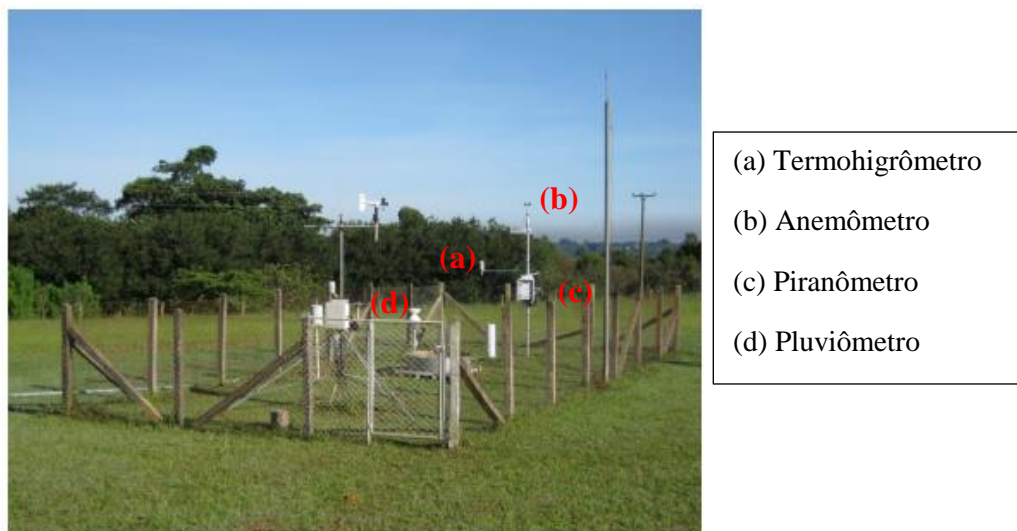


Figura 1. Estação agrometeorológica automática Hobeco, do Laboratório de Agroclimatologia – FAV/UnB, localizada na Fazenda Água Limpa.

O banco de dados usados neste estudo é composto pelos parâmetros climáticos de temperatura média, máxima e mínima ( $T_{med}$ ,  $T_{max}$  e  $T_{min}$ ), umidade relativa do ar média, máxima e mínima ( $UR_{med}$ ,  $UR_{max}$  e  $UR_{min}$ ), velocidade do vento média ( $U_{med}$ ), direção média do vento ( $U_{dir_{med}}$ ), radiação solar global (Rad) e precipitação (Prec), do período de abril de 2000 a abril de 2018, compreendendo uma análise dos dados diários publicados e avaliação de qualidade dos dados horários.

### 3.2- Análise exploratória dos dados diários

Para a análise do banco de dados, utilizou-se a linguagem de programação R, versão 3.4.3 (2017-11-30) do RStudio, e alguns pacotes de análise já disponibilizados. A caracterização inicial dos dados foi realizada por análise exploratória, com a finalidade de identificar a variação dos valores médios, máximos e mínimos dos elementos diários através da elaboração dos gráficos de dispersão, barras e histogramas, e assim identificar as possíveis ocorrências, persistência e sazonalidade dos dados coletados no período de estudo. Para isso fez-se necessário a utilização dos pacotes *gplot2* (WICKHAM, 2016), e os pacotes de formatação de temas, escalas e cores, *ggthemes* (ARNOLD, 2018) e *gridExtra* (AUGUIE, 2017) para a construção dos gráficos de temperatura, umidade, precipitação e radiação. Já para os gráficos de vento, utilizou-se o pacote *openair* (CARSLAW; ROPKINS, 2012), com a função *WindRose* para representar a frequência da velocidade e direção do vento.

### 3.3- Análise de qualidade dos dados horários

A análise de qualidade efetuou-se por meio das combinações dos códigos sugeridos pela OMM, juntamente com o estabelecimento dos limites de tolerâncias realizados no nível 1 software MADIS e a utilização dos códigos sinalizadores de Pittigliani (2000). Nessa etapa, utilizou-se os dados horários dos 18 anos de dados, já que ao trabalhar com dados diários pode levar a interpretações errôneas. Foram realizados os testes de limites (*range*) e o teste de degrau (*step*).

Os valores de tolerância foram baseados na resolução dos equipamentos, apresentados na Tabela 4, realizando assim a análise em todos os elementos climáticos que a estação possui equipamento de leitura. Para a estação Campbell, o sensor de radiação não apresenta a informação de limite da resolução, portanto, foi utilizado como referência 80% da radiação máxima que chega na superfície terrestre ( $1500 \text{ W m}^{-2}$ ), estabelecendo uma radiação máxima de  $1200 \text{ W m}^{-2}$ .

Para realizar o teste de *range* em linguagem R, utilizou-se o código apresentado no *Script 1* (Figura 2), no qual o limite de tolerância varia de acordo com o elemento avaliado e com o período de análise, isto é, do ano 2000-2010, com a Estação Campbell e a partir de 2011 com a estação Hobeco.

```
> X2000_hora$temp_med <- as.numeric(as.character(X2000_hora$temp_med))
> lim_temp <- (-39.6) <= X2000_hora$temp_med & X2000_hora$temp_med <= 60.6
> X2000_hora$nota1_temp_med <- 7L
> X2000_hora$nota1_temp_med [lim_temp] <- 0L
> X2000_hora$nota1_temp_med [!lim_temp] <- 3L
```

Figura 2. *Script 1*: Comandos utilizados para executar o teste de *range* em todos os elementos climáticos do banco de dados da estação meteorológica da FAL/UnB de 2000 a 2018.

O teste de *step* (degrau), que verifica a diferença máxima entre dois valores de dados consecutivos, foi aplicado somente nos elementos umidade e temperatura, já que esses estão menos sujeitos a sofrer com alterações bruscas, enquanto que os elementos precipitação, radiação e vento, são mais variáveis e mais sujeitos a fatores momentâneos,



como chuva passageira, nuvens e rajadas de ventos, respectivamente, não sendo possível assim estabelecer um limite aceitável de variação entre medidas consecutivas.

No estabelecimento dos limites de degrau aceitáveis de temperatura e umidade desse teste, utilizou-se como base o trabalho de Anunciação; Maggioletto; André (2012), que avaliou o comportamento dos dados horários de temperatura e umidade relativa do ar do mês de agosto em Brasília. Neste trabalho verificou-se que a variação horária da temperatura ultrapassou 4,2 °C e de umidade do ar de 15% entre horas consecutivas no início da manhã e vai aumentando de forma suave até atingir o valor máximo do dia e depois volta a cair decrescendo 4,2 °C e 15% no início da noite em horas consecutivas. Tomando como base esses valores e explorando os 18 anos de dados dessa estação, utilizou-se o limite de tolerância para a diferença horária de temperatura de  $\pm 4,5$  °C e  $\pm 15\%$  para a umidade relativa do ar.

Para a execução do teste de *step* em linguagem R, utilizou-se a função *append*, que adiciona valores a um novo vetor, e a função *diff*, que calcula a diferença entre os valores consecutivos. Como a análise é horária, utiliza-se o valor 0 dentro da função *append* para ignorar o primeiro valor de cada arquivo sem perder seu registro. Posteriormente, realiza-se a atribuição dos códigos sinalizadores, de acordo com os critérios estabelecidos. A função *abs*, na atribuição dos *flags* possui a função de exibir o valor absoluto da condição estabelecida (Figura 3).

```
> X2001_hora$temp_med<- as.numeric (as.character(X2001_hora$temp_med))
> dif_temp_med<- diff (X2001_hora$temp_med)
> X2001_hora$dif_temp_med<- append (dif_temp_med, 0)
> X2001_hora$nota_temp_med[abs(X2001_hora$dif_temp_med)<=4.5|abs(X2001_hora$dif
_temp_med)>= (-4.5)] <- 0
> X2001_hora$nota_temp_med[!abs(X2001_hora$dif_temp_med)<=4.5|
!abs(X2001_hora$dif_temp_med)>= (-4.5)] <- 3
> X2001_hora$nota_temp_med [is.na(X2001_hora$dif_temp_med)] <- 7
```

Figura 3. *Script 2*: Comandos utilizados para executar o teste de *step* para temperatura e umidade relativa do ar do banco de dados da estação meteorológica da FAL/UnB do ano 2000 a 2018.

Como a linguagem R é uma linguagem de programação vetorizada, isto é, pode-se processar vetores inteiros de uma só vez, para entender melhor os *scripts* executados nos testes *range* e *step*, é necessário ter atenção há alguns comandos chaves (Tabela 5).

Tabela 5. Descrição de alguns comandos de programação utilizados na execução dos testes *range* e *step*.

Funções/Símbolos	Descrição
>	Linha de comando do R
<-	Armazena um objeto
as.numeric	Interpreta o objeto como numérico
dif_temp_med<-	Objeto que está armazenando a função <i>diff</i>
X2000_hora	Nome e ano do arquivo que está sendo trabalhado
\$temp_med	Coluna do arquivo que está sendo trabalhada
\$nota_temp_med –	Coluna que está sendo criada para armazenar o <i>flag</i>
	Operador lógico OR
!	Operador lógico NOT
<=	Menor ou igual
>=	Maior ou igual
in.na	Indica que há um valor ausente naquela linha

A partir disso, com a execução dos *scripts*, atribuiu-se as sinalizações (*flags*) indicando a qualidade dos dados, no qual

- 0 = Aprovado → Dentro da resolução do equipamento e variação estabelecida
- 3 = Reprovado → Fora da resolução do equipamento e variação estabelecida
- 7 = Dados ausentes → Não há registros de dados

#### 4. RESULTADOS E DISCUSSÃO

##### 4.1- Análise exploratória

A Tabela 6 apresenta os valores de temperaturas médias, mínimas e máximas, do período de abril de 2000 a abril 2018, a partir dos valores diários armazenados. Observa-se a  $T_{med}$  não variou muito ao longo dos anos, ficando entre 19,8 e 20,8°C.

Tabela 6. Temperatura média ( $T_{med}$ ) anual e valores médios e extremos da temperatura mínima ( $T_{min}$ ) e temperatura máxima ( $T_{max}$ ) anuais coletados pela estação automática da FAL (FAV-UnB) para o período de abril de 2000 a abril de 2018.

Ano	$T_{med}$ (°C)	$T_{min}$ (°C)		$T_{max}$ (°C)	
		média	valor extremo (data de ocorrência)	média	valor extremo (data de ocorrência)
2000*	19.8	13.0	1.5 (18/07)	27.9	33.9 (26/08)
2001	20.1	13.8	6.3 (28/06)	27.9	34.2 (12/09)
2002	20.6	14.0	5.7 (28/05)	28.6	34.0 (22/10)
2003	20.3	14.0	4.2 (03/07)	28.0	34.6 (21/10)
2004	20.1	14.1	4.0 (24/07)	27.5	34.0 (30/09)
2005	20.3	15.0	7.8 (05/05)	27.4	30.8 (22/02)
2006	19.8	13.9	4.0 (24/07)	27.2	32.8 (18/09)
2007	20.3	13.4	4.8 (01/08)	28.3	34.9 (24/09)
2008	19.9	13.3	4.1 (10/07)	28.0	36.5 (30/04)
2009	20.2	14.4	5.8 (04/07)	27.7	32.6 (06/03)
2010	20.2	13.4	3.9 (20/08)	28.3	33.2 (06/10)
2011	20.6	14.6	5.5 (05/09)	28.0	34.7 (11/09)
2012	20.3	13.8	4.7 (25/07)	28.0	34.9 (30/10)
2013	20.3	14.1	4.9 (03/08)	28.0	34.1 (22/09)
2014	20.2	13.6	3.5 (09/08)	28.2	36.2 (15/10)
2015	20.8	14.1	5.7 (13/08)	28.9	36.3 (18/10)
2016	20.8	14.1	5.3 (18/06)	29.0	34.3 (16/09)
2017	20.3	13.6	4.9 (03/09)	28.3	35.9 (15/10)
2018**	21.0	15.8	11.9 (23/01)	22.1	22.2 (04/03)

\* abril a dezembro

\*\* janeiro a abril

Os valores extremos anuais de temperatura mínima registrados variaram de 1,5 a 7,8 °C, observados em 2000 e 2015 respectivamente (não foi considerado o valor extremo em 2018). As temperaturas mínimas extremas ocorreram quase sempre entre os meses de junho e agosto. No ano de 2002, a temperatura mínima extrema (5,7 °C) ocorreu em 28 de maio; já no ano de 2011, a menor temperatura registrada foi de 5,5 °C no dia 5 de setembro. Esses valores de temperaturas mais baixas para esses meses não são usualmente observados, mas é importante ressaltar que para o ano de 2011, o valor registrado em setembro pode não ter sido a menor temperatura do ano, já que não houve coleta de dados

nos meses de abril a julho devido a problemas com o coletor de dados. Se não forem considerados os anos com falha na coleta de dados (2005, 2006 e 2011), a média das  $T_{\min}$  diárias variou de 13,0 °C a 14,4 °C; nos anos 2005 e 2018 a médias de  $T_{\min}$  foram de 15,0 °C e 15,8 °C. Mas como não há registro diário completo de temperatura, esses valores são influenciados pela média de dias com registros.

Para os extremos da  $T_{\max}$ , os meses que se concentram as maiores temperaturas foram setembro e outubro. As menores extremas máxima foram em meses atípicos de registros de temperaturas extremas, sendo elas no dia 22 de fevereiro de 2005 (30,8 °C) e 6 de março de 2009 (32,6 °C). Como no ano de 2005 não possui o registro anual completo de dados, esse valor pode ser não representativo para o ano. Já os maiores valores extremos de  $T_{\max}$  foram registrados nos anos de 2008, 2014 e 2015, com valores acima de 36 °C (36,5 °C, 36,3 °C e 36,2 °C, respectivamente), para os meses de abril e outubro. A média da  $T_{\max}$  não variou muito de um ano para o outro, se concentrando entre 27,2 °C e 29,0 °C.

Apesar do ano 2000 apresentar a menor temperatura registrada e o ano de 2008 a maior temperatura, não é possível afirmar se essas foram realmente as menores mínimas e as maiores máximas ocorridas no local nesses 18 anos, já que nos anos de 2005, 2006 e 2011 há falhas de registros de dados. Nesse sentido, de acordo com esse banco de dados, esses são os anos que possuem os menores e maiores extremos de temperatura.

A flutuação diária da temperatura ao longo do ano pode ser observada nas Figura 4 e Figura 5. Verifica-se que no início dos anos, a  $T_{\text{med}}$  do ar não varia muito, ficam na faixa entre 20 e 25 °C. Em geral, a  $T_{\text{med}}$  começa a decair entre os dias 99 e 113, que corresponde a metade do mês de abril. A partir daí ocorre uma maior frequência de temperaturas mínimas abaixo de 15 °C, nos meses de maio a julho, que são os meses mais frios do ano. A partir do dia 211 aproximadamente, que corresponde ao final de julho/início de agosto, até o final de outubro/início de novembro, observa-se registros de temperaturas mais altas, com  $T_{\max}$  acima de 32 °C. Após esse período, as temperaturas médias começam a diminuir e se estabilizarem novamente na faixa de 20 °C a 25 °C.

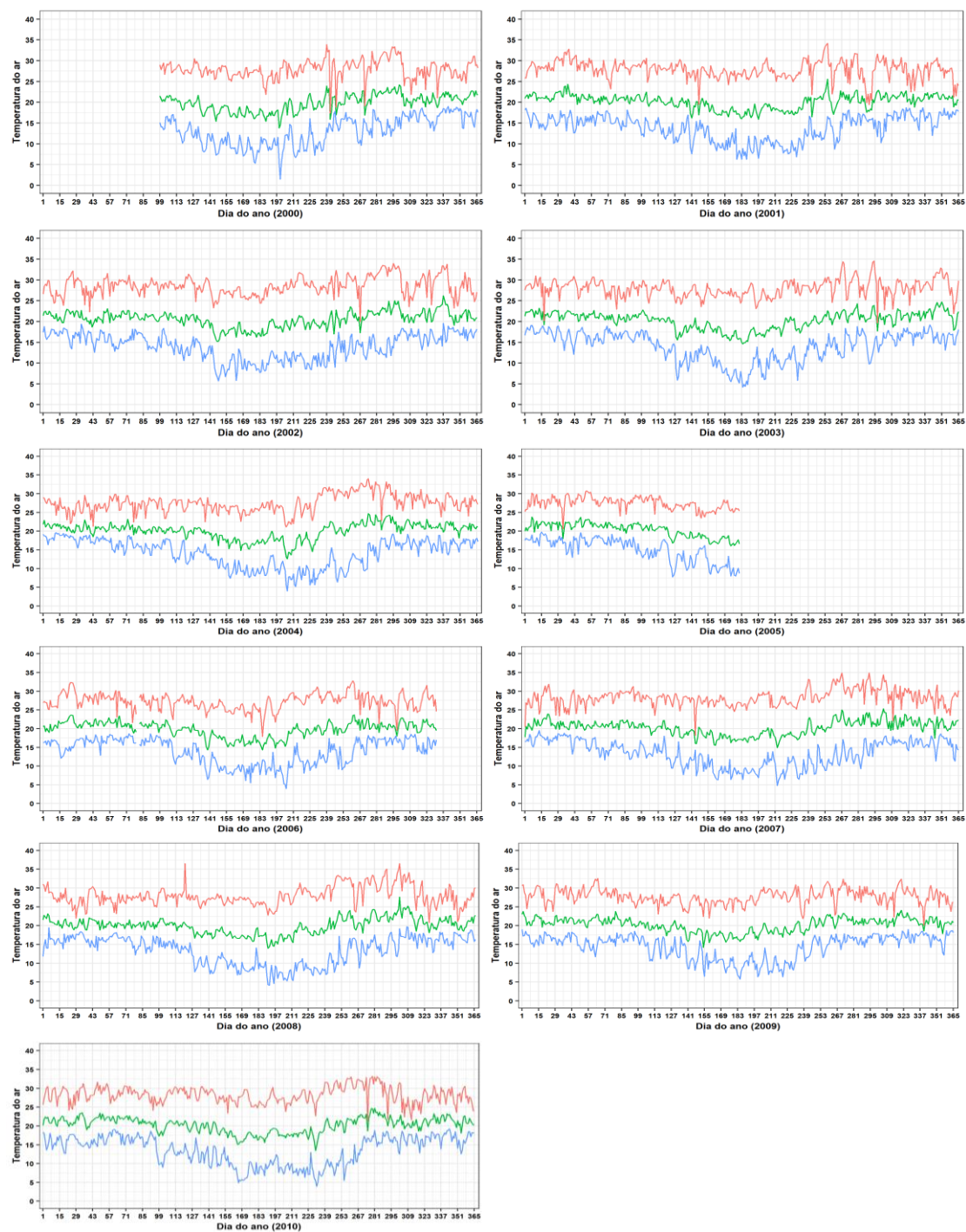


Figura 4. Flutuação das temperaturas médias, máximas e mínimas diárias durante o período de coleta de dados de abril de 2000 a dezembro de 2010, pela estação automática da FAL (FAV/UnB).

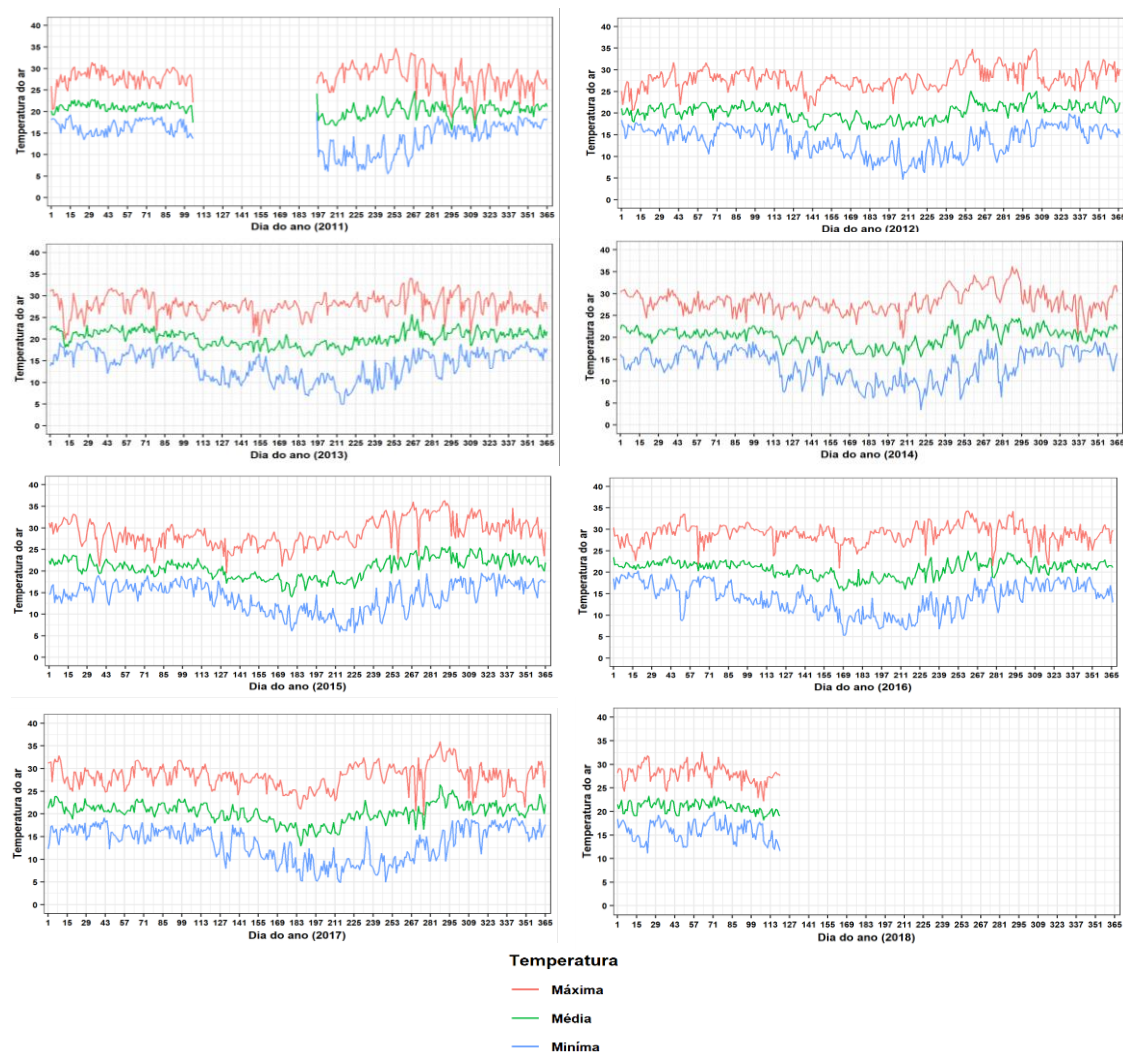


Figura 5. Flutuação das temperaturas médias, máximas e mínimas diárias durante o período de coleta de dados de janeiro de 2011 a abril de 2018 pela estação automática da FAL (FAV/UnB).

A análise da frequência de ocorrência dos registros diários de temperatura ao longo dos 18 anos de coleta de dados na FAL (Figura 6), as maiores frequências de registros  $T_{\min}$  foi entre 17 e 18 °C; entre 21 e 23 °C para  $T_{\text{med}}$ ; e entre 27 e 29 °C para  $T_{\text{max}}$ .

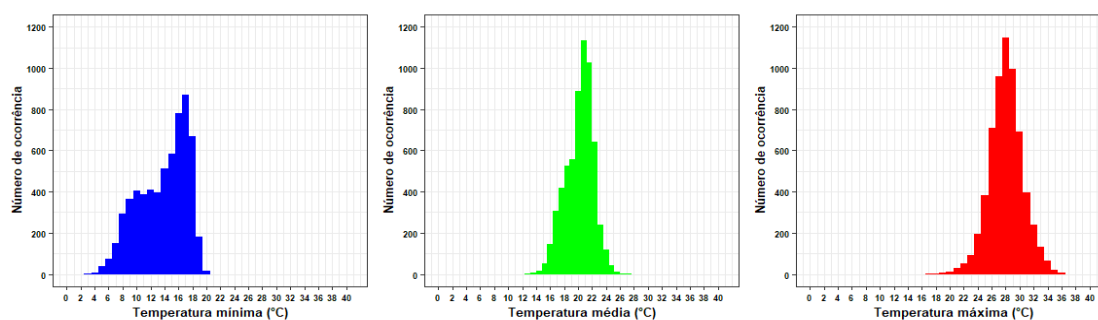


Figura 6. Frequência de ocorrência das temperaturas mínimas, média e máximas diárias nos 18 anos de coleta de dados climáticos realizados na estação agrometeorológica da FAL (FAV/UnB).

As normais climatológicas de 1981-2010 para Brasília apontam uma temperatura média anual de 21,0 °C, média das mínimas de 16,8 °C e média das máximas de 26,6 °C (INMET, 2018). Na estação da FAL, as médias das temperaturas máximas foi 28,1 °C, das temperaturas mínimas foi 13,9 °C, e temperatura média anual foi 20,3 °C, ao longo desses 18 anos de coleta. Essa diferença pode ser oriunda da grande quantidade de períodos que não possuem dados. Ao observar os registros de frequências de ocorrência (Figura 6), a maior quantidade de dados inclui os valores das normais de temperatura (média, mínima e máxima) do INMET para Brasília.

A análise da umidade relativa feita a partir dos valores médios diários de UR indicou que a UR<sub>med</sub> anual variou de 70 a 77 %, quando considerados os anos completos (Tabela 7). Até 2006 somente a informação de UR<sub>med</sub> diária era armazenada, e, portanto, os valores de UR<sub>min</sub> e UR<sub>max</sub> apresentados na Tabela 7 indicam os valores extremos de máxima e mínima da UR<sub>med</sub> diária ao longo dos anos. A menor UR<sub>min</sub> registrada foi no ano de 2004, com 35,8% , enquanto que os anos de 2009 e 2010 registraram dias em que a UR<sub>max</sub> foi igual a 100%.

Tabela 7. Umidade relativa média (UR<sub>med</sub>) anual e valores médios umidade relativa mínima (UR<sub>min</sub>) e umidade relativa (UR<sub>max</sub>) anuais coletados pela estação automática da FAL (FAV-UnB) para o período de abril de 2000 a abril de 2018.

Ano	UR <sub>min</sub>	UR <sub>med</sub>	UR <sub>max</sub>
2000*	44.5	69.7	93.6
2001	43.1	72.3	93.6
2002	38.7	70.5	93.3
2003	40.8	72.9	94.8
2004	35.8	74.1	94.4
2005	69.6	80.2	95.2
2006	46.9	76.9	95.7
2007	38.3	71.1	97.7
2008	38.7	74.3	97.3
2009	51.5	81.6	100.0
2010	42.4	76.8	100.0
2011	37.5	73.8	95.1
2012	42.0	73.6	95.6
2013	46.5	75.6	96.0
2014	39.8	73.7	96.1
2015	38.4	73.6	94.6
2016	41.8	73.0	95.6
2017	36.1	73.0	96.2
2018**	60.6	82.8	96.9

\* abril a dezembro

\*\* janeiro a abril

Nos dados de diários de UR (Figura 7 e Figura 8), observa-se que as  $UR_{\min}$  extremas foram bastante baixos em alguns períodos. Apesar de não possuir um ano completo de registro de dados, o ano de 2011 registrou a menor umidade para esses 18 anos de coleta de dados, atingindo 10% no dia 23 de agosto (dia do ano 234). No ano de 2014, os registros de  $UR_{\min}$  foram acima de 20%, e somente o dia 210, no final de agosto, registrou um valor abaixo, de 18,8%. Esses registros de umidade relativa mais baixa do ar ocorrem em períodos nos quais as temperaturas do ar são mais altas.

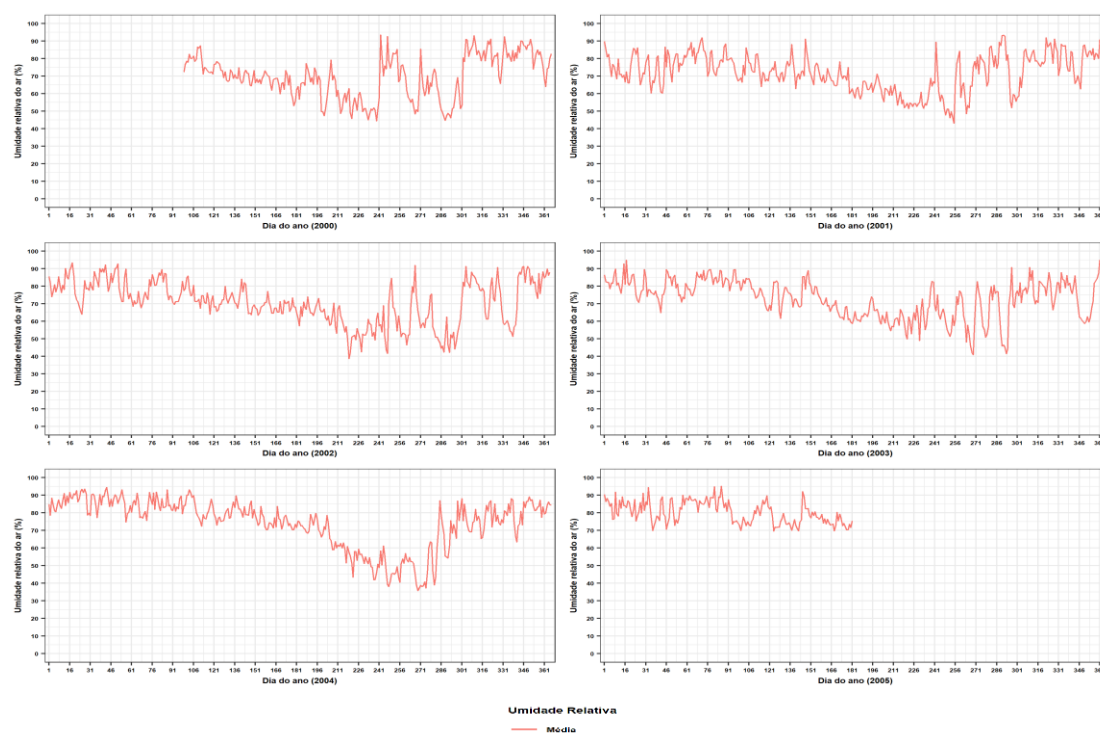


Figura 7. Flutuação dos valores diários de umidade relativa médias coletadas no período de abril de 2000 a dezembro de 2005 pela estação automática da FAL (FAV/UnB).



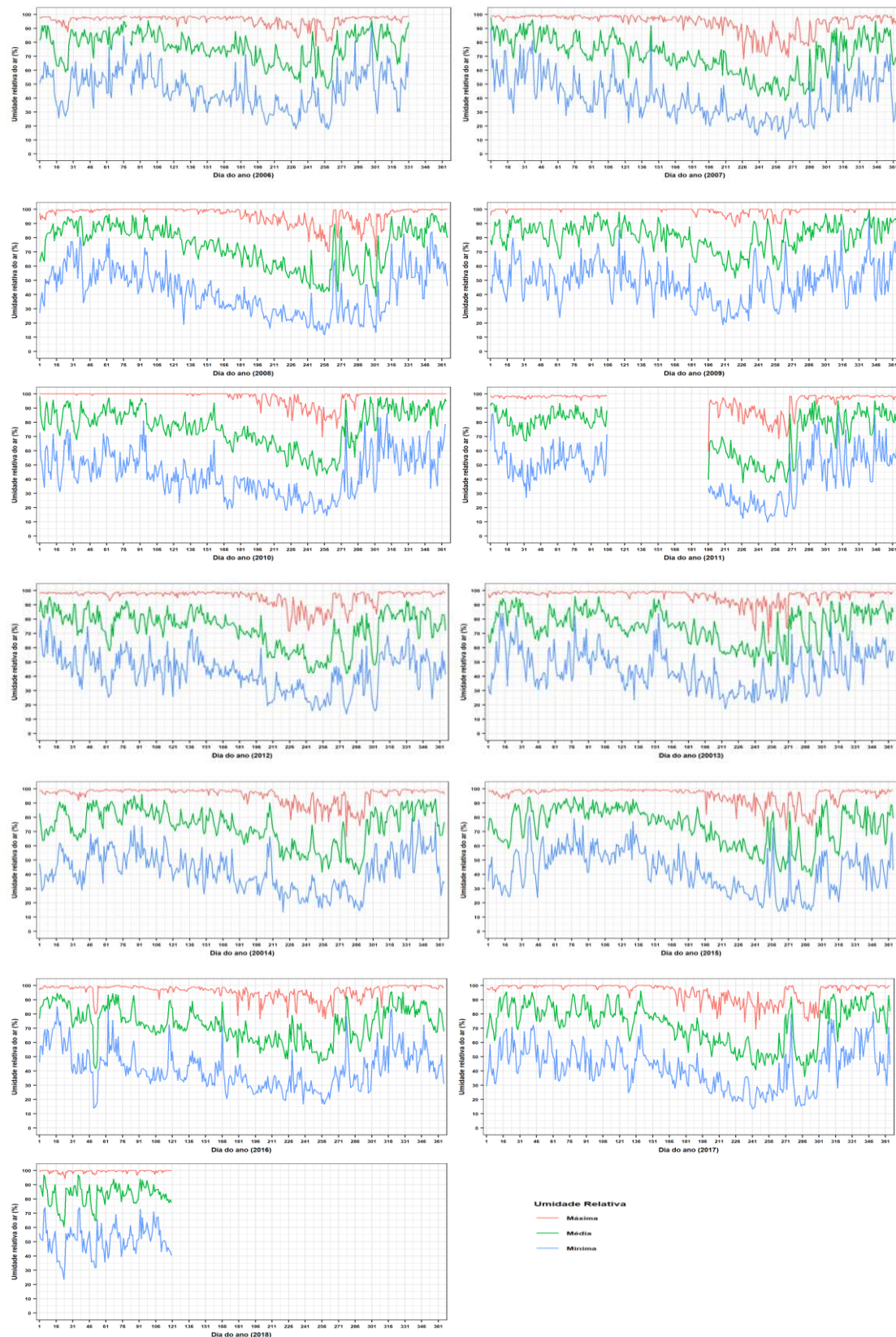


Figura 8. Flutuação dos valores diários de umidade relativa médias, máximas e mínimas diárias durante os anos de janeiro de 2006 a abril de 2018 coletados pela estação automática da FAL (FAV/UnB).

Durante os 18 anos de coleta de dados, observa-se que não houve um padrão de distribuição e frequência de chuva, havendo grande variação da precipitação acumulada de um ano para outro. Observando a **Erro! Fonte de referência não encontrada.**, os anos de 2011, 2013 e 2014 foram os anos que mais se destacaram quanto ao acúmulo de chuva, ultrapassando 1500 mm anuais, enquanto que em 2005, 2007 e 2015 a precipitação acumulada foi de 1090,6 mm, 1043,7 e 1118,6 mm, respectivamente. Apesar de não possuir um ano completo de registro de dados, o ano de 2005 obteve uma maior precipitação acumulada do que todo o ano de 2007. Esses registros estão bem distantes da média acumulada para Brasília, que é de 1477,4 mm, segundo as normais climatológicas de 1981-2010 (INMET, 2018).

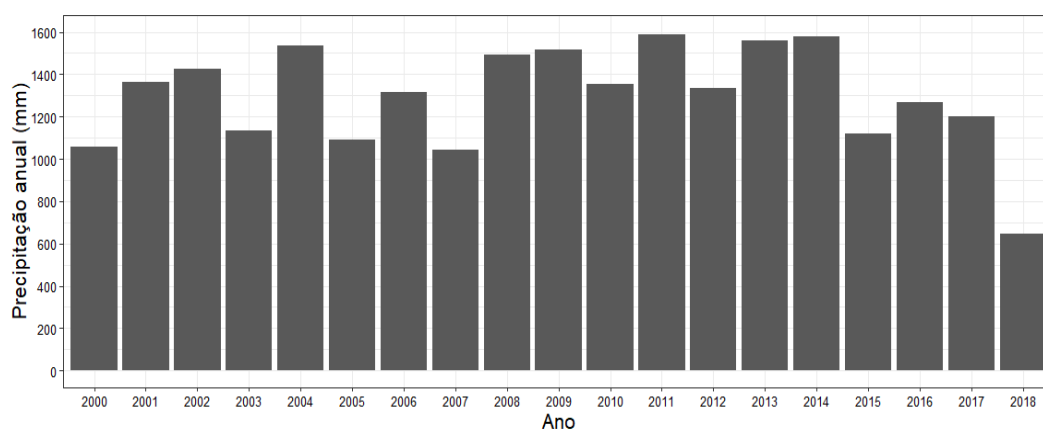


Figura 9. Precipitação acumulada anual para o período de abril de 2000 a abril de 2018 coletado pela estação meteorológica da FAV/UnB.

O baixo acúmulo de precipitação influencia nas atividades rurais e urbanas, de forma que se faz necessária a utilização ainda mais racional dos recursos hídricos, já que essa quantidade de chuva não é suficiente para manter os reservatórios nos níveis de manutenção do consumo.

Coelho et al. (2012), caracterizando a estação chuvosa no Distrito Federal, observaram que a estação chuvosa em Brasília ocorre entre os meses de outubro a abril, e que no período de transição entre as estações chuvosa e seca, ocorrem chuvas irregulares, nos meses de maio a setembro, isto é, esses registros nem sempre ocorrem, variando de ano para ano. Analisando a Figura 10, observa-se os dados coletados estão condizente com a análise de Coelho et al. (2012), já que na estação seca, de maio a setembro, observa-se que a precipitação acumulada mensal não ultrapassou 50 mm, registrando chuvas fora do período somente nos anos 2009 e 2012, quando foram registrados 99 e 59 mm no mês de maio, respectivamente.

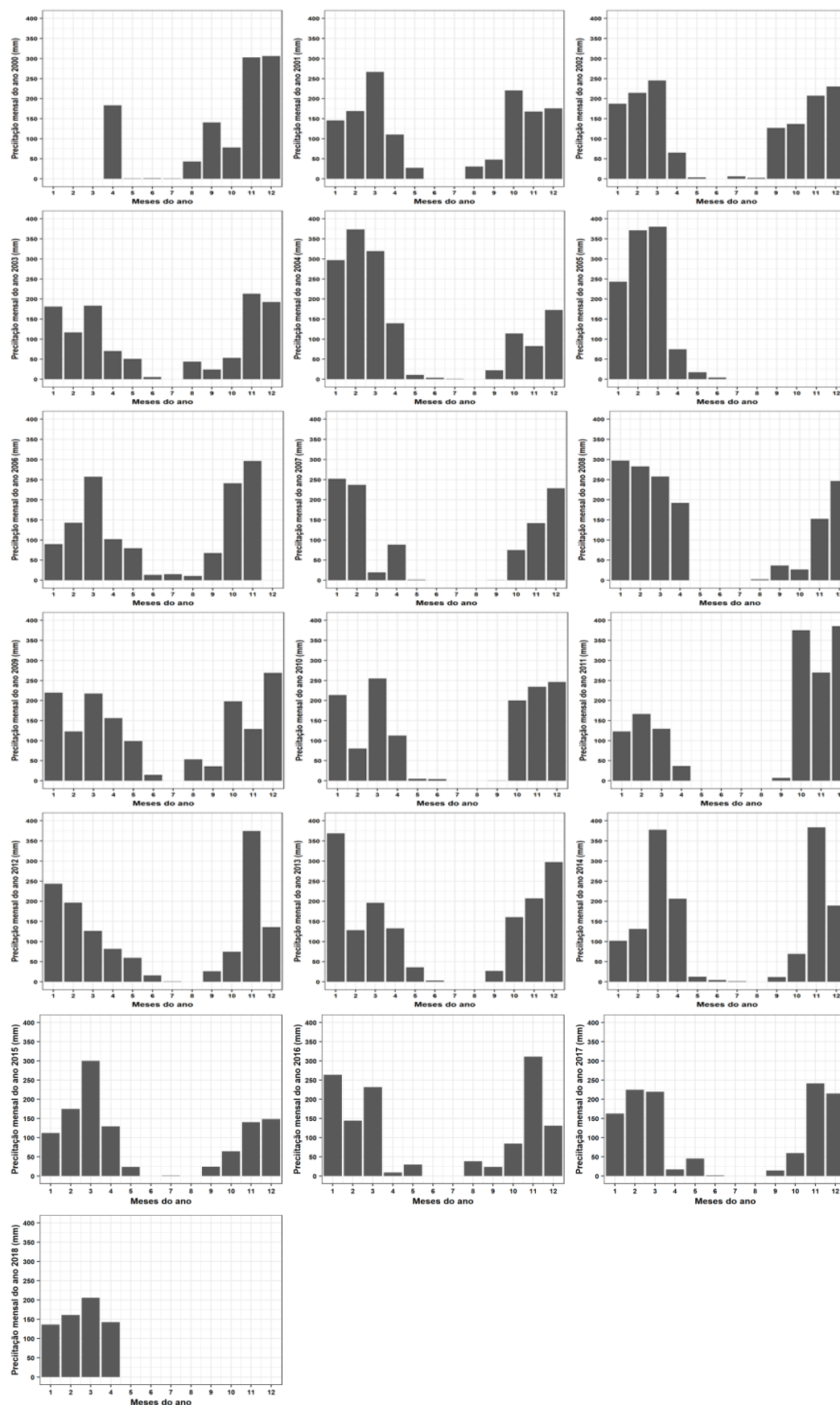


Figura 10. Precipitação pluviométrica acumulada mensal (mm), entre abril de 2000 e abril de 2018, coletado pela estação meteorológica da FAL (FAV/UnB).

Na Figura 10 observa-se também, que o mês de maior precipitação durante a estação chuvosa variou bastante ao longo do período do estudo. O mês de novembro foi o que apresentou maior precipitação nos anos de 2003, 2006, 2012, 2014, 2016 e 2017, variando de 200 a 350 mm mensais. O mês de março foi o mais chuvoso nos anos de 2001, 2002, 2005, 2010, 2015 e 2018, com precipitações entre 200 e 300 mm. O mês de dezembro teve maior precipitação acumulada em 2000, 2009 e 2011, e janeiro nos anos de 2007, 2008 e 2013. Fevereiro registrou maior precipitação mensal somente no ano de 2004. Zoby (1999) comenta que os meses de transição do período seco para o chuvoso e vice-versa são os meses que possuem maiores registros de precipitação, como ocorreu no presente estudo, já que os meses de novembro e março são os que mais se destacaram.

Nos meses de maio de 2009 e 2012, o acúmulo de precipitação apresentou registros bem acima da Normal Climatológica para o mês, que é de 29,7 mm (Tabela 8). Isso também ocorreu para o mês de agosto, nos anos de 2000, 2001, 2003, 2019 e 2016, no qual a precipitação média é 24,1 mm. Como o ano de 2017 apresenta registros mensais de acúmulos pluviométricos abaixo dos registros das normais climatológicas mensais para Brasília, configura-se esse período como ano seco de acordo com Coelho et al. (2012).

Tabela 8. Normais Climatológicas da precipitação no período de 1981 a 2010 para a região de Brasília- DF.

Mês	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
Prec	209.4	183.0	211.8	133.4	29.7	4.9	6.3	24.1	46.6	159.8	226.9	241.5

Fonte: INMET, 2018

Para a análise da variação da radiação solar global diária coletada pela estação da FAL ( $Q_g$ ), comparou-se os valores observados com a irradiância solar extraterrestre ( $Q_o$  – valores tabelados) ao longo do ano. As diferenças observadas ocorrem devido à variação da transmitância atmosférica, dependente principalmente da nebulosidade.

A radiação solar que atinge a superfície no período de maio a setembro (dia do ano 124 a 244) apresentou uma menor oscilação, quando comparada ao restante do ano devido à menor nebulosidade, típica do período (Figura 11 e Figura 12).

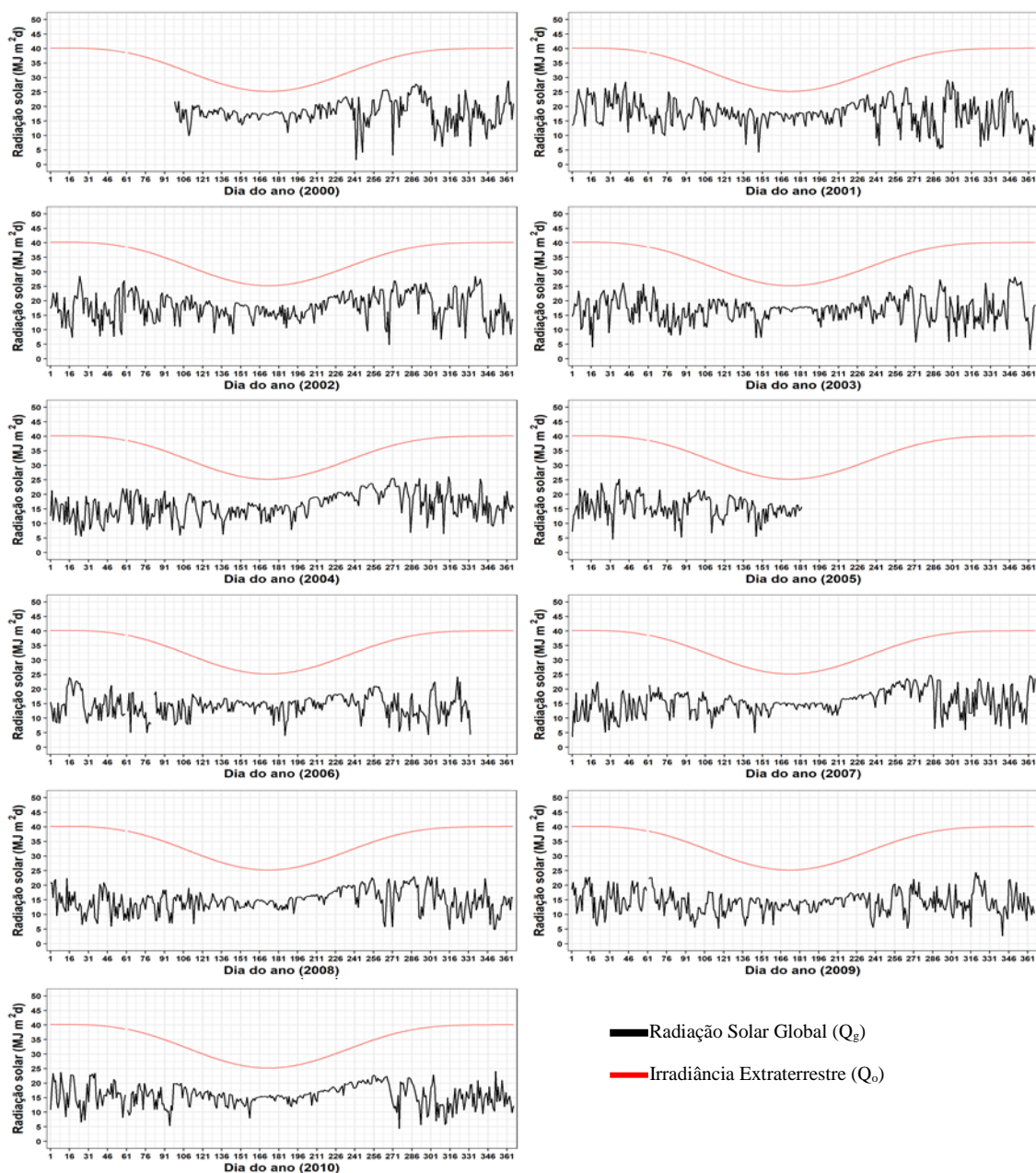


Figura 11. Radiação solar global diária ( $\text{MJ m}^{-2} \text{d}^{-1}$ ) incidente na superfície e radiação solar incidente no topo da atmosfera para período de abril de 2000 a dezembro de 2010, coletado pela estação meteorológica da FAL (FAV/UnB).

No dia 196 do ano 2011, a radiação medida foi de  $30 \text{ MJ m}^{-2} \text{d}^{-1}$ , valor superior a  $Q_0$  para a data (Figura 12). Fisicamente, a medida não pode estar correta, já que na superfície a radiação não pode ser superior à que chega no topo da atmosfera ( $Q_0$ ), o que indica que esse registro é oriundo de um erro de medição do equipamento, já que esse valor foi observado logo após um período no qual houve falhas de coleta de dados.

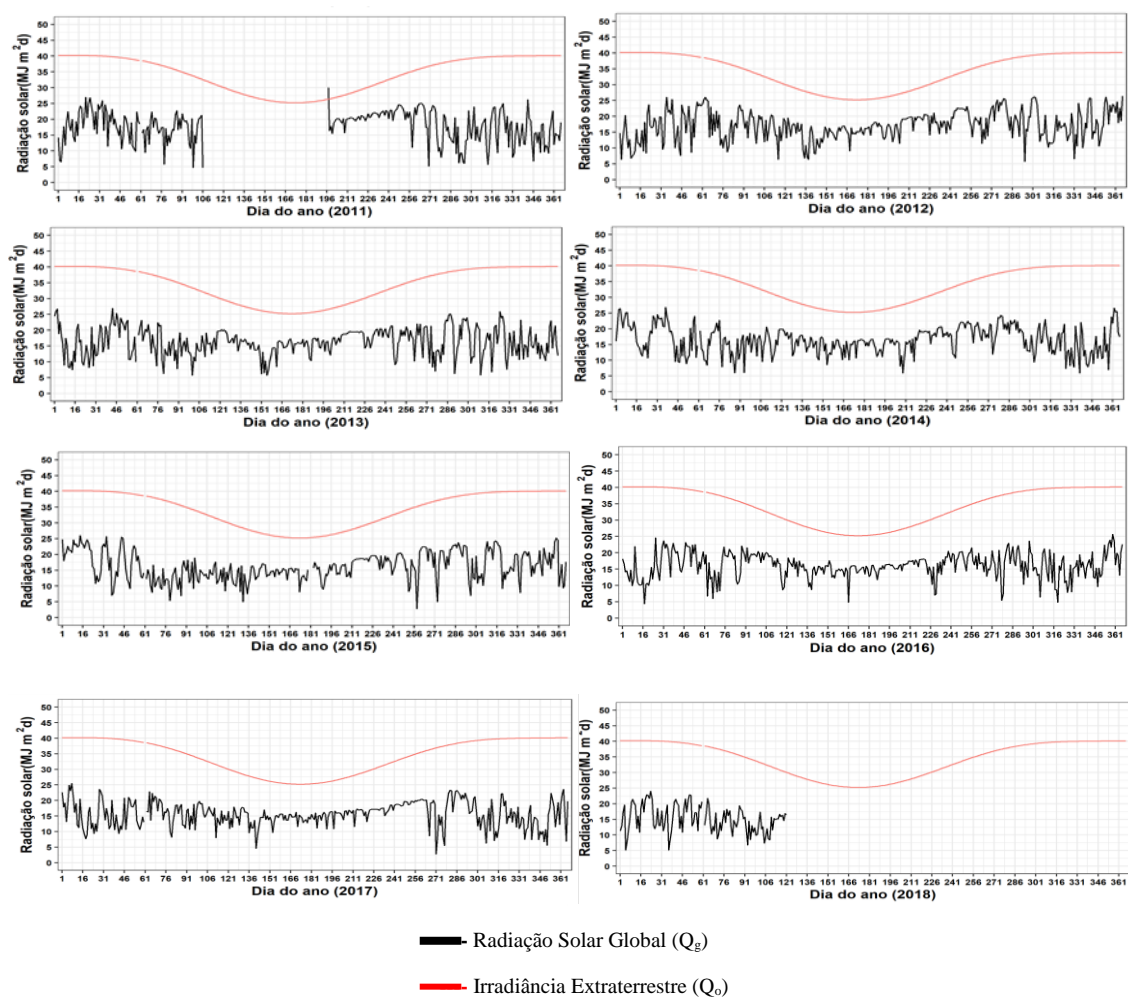


Figura 12. Radiação solar global diária ( $\text{MJ m}^{-2} \text{d}^{-1}$ ) incidente na superfície e radiação solar incidente no topo da atmosfera para período de janeiro de 2011 a abril de 2018, coletado pela estação meteorológica da FAL (FAV/UnB).

A Figura 13 e a Figura 14 apresentam os gráficos de frequência de velocidade e direção do vento, a partir dos valores médios diários. Mesmo os anos que não possuem registro completo das variáveis climáticas (2000, 2005, 2006, 2011 e 2018) pode-se observar uma maior frequência de velocidade do vento entre  $1$  e  $2 \text{ m s}^{-1}$ , com direção predominante SE e S. Observa-se que o ano de 2007 (Figura 13), diferentemente dos outros anos, apresenta direção do vento predominante N e NE. Analisando os outros elementos climáticos (temperatura, umidade, precipitação e radiação), não foram observados registros fora da normalidade dos dados no mesmo período; infere-se, portanto, que para esse período de coleta de dados de direção do vento há algum tipo de erro relacionado ao equipamento.

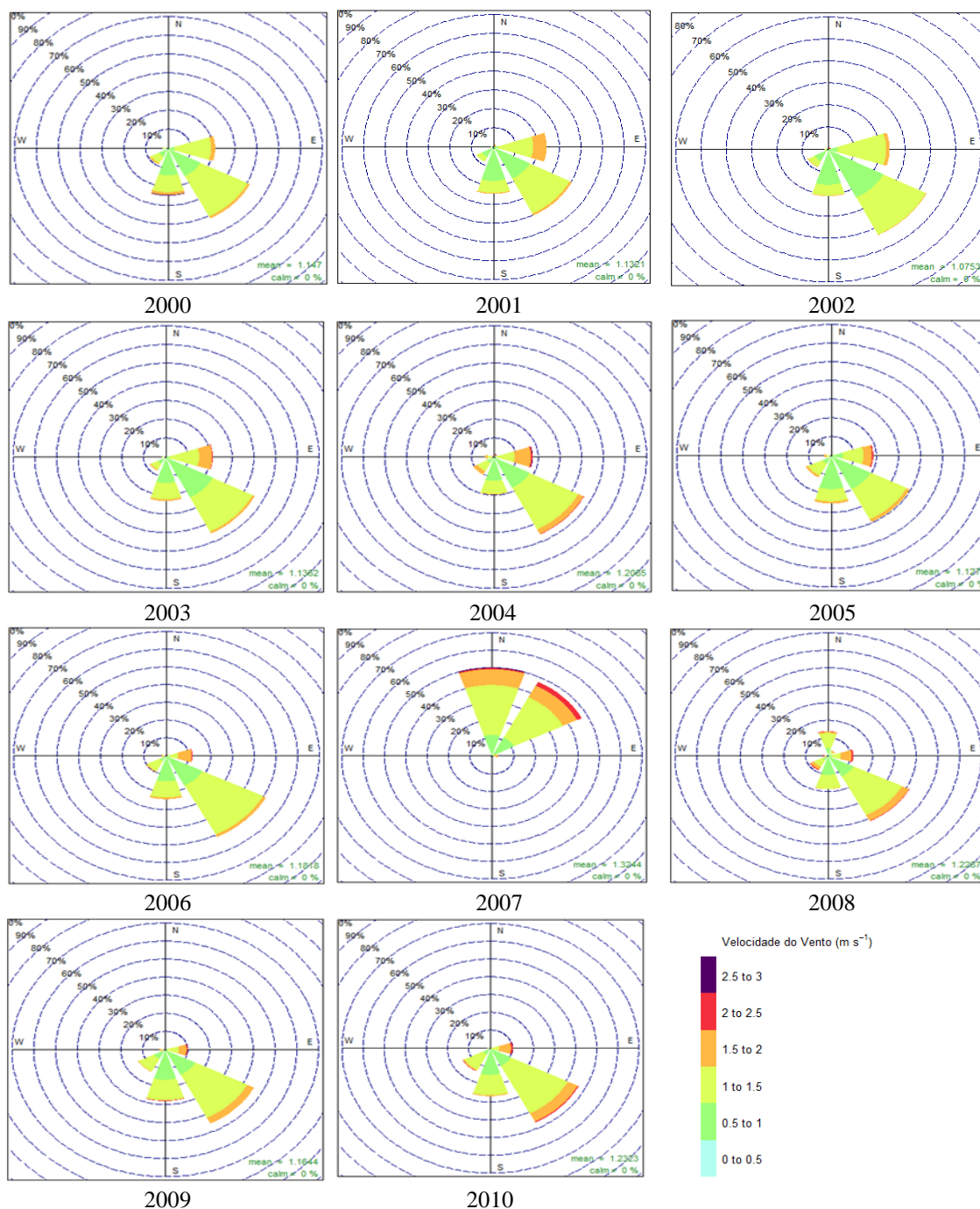


Figura 13. Frequência de ocorrência da velocidade ( $\text{m s}^{-1}$ ) e direção do vento ( $^{\circ}$ ) para o período de abril de 2000 a dezembro de 2010, coletado pela estação meteorológica da FAL (FAV/UnB).

A diferença de equipamentos instalados nas estações nos diferentes períodos de medição, 2000-2010 e 2011-2018, pode ser observada nas Figuras 13 e 14, respectivamente, já que até o ano 2010 o anemômetro utilizado exigia uma velocidade mínima do vento de  $1 \text{ m s}^{-1}$  para começar o seu funcionamento, enquanto que após 2011, o equipamento utilizado permitiu medições, e conseqüentemente médias, de velocidade do vento abaixo de  $0,5 \text{ m s}^{-1}$ .

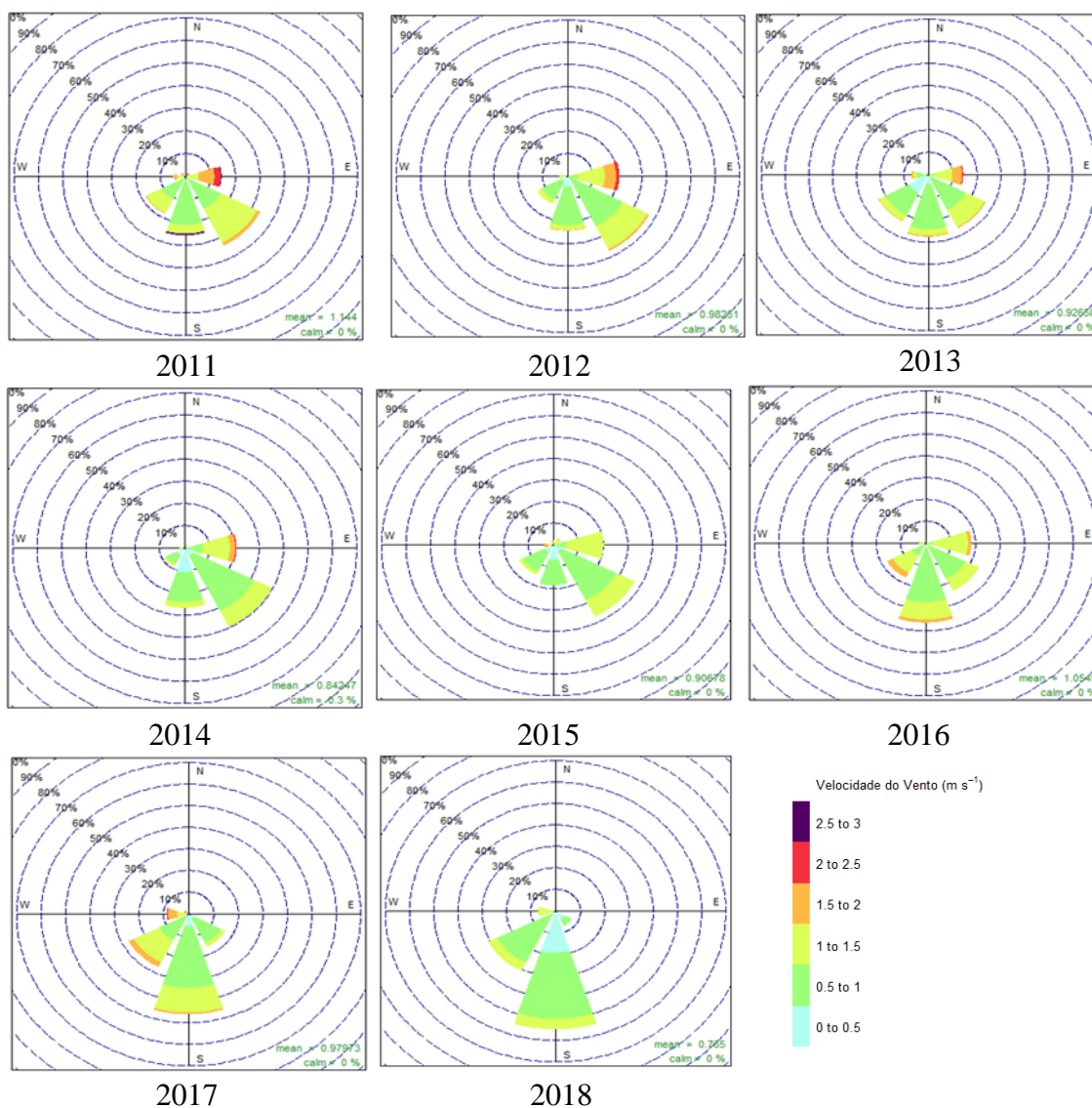


Figura 14. Frequência de ocorrência da velocidade ( $\text{m s}^{-1}$ ) e direção do vento ( $^\circ$ ) para o período de abril de 2000 a abril de 2018, coletado pela estação meteorológica da FAL (FAV/UnB).

Zoby (1999), em estudo da hidrologia da Bacia do Riberão Preto em Sobradinho, registrou uma variação do vento entre  $1,6$  e  $2,0 \text{ m s}^{-1}$ , com uma direção predominante SE/E. Observa-se que os dados da estação agrometeorológica da FAL apresentou menores velocidades do vento do que os resultados obtidos por Zoby (1999), diferença pode ser oriunda da influência da variação altimétrica, da pressão e pelo relevo do terreno.

#### 4.2- Análise de qualidade

As tabelas a seguir apresentam o número de vezes em que os valores sinalizadores dos testes de qualidade foram observados nos dados horários, para os testes realizados, teste de limites (*range*) e o teste de degrau (*step*).



Nos 18 anos de coleta de dados meteorológicos, não houve nenhum registro de temperatura que excedeu a resolução do equipamento (Tabela 9), não registrando portanto nenhum *flag* igual a 3. O que mais se observa é a quantidade de *flags* iguais a 7, principalmente nos anos 2005, 2006 e 2011, indicando a ausência de dados nesses períodos. É importante observar que os valores que obtiveram *flag* igual a 0, ou seja, foram aprovados e foram utilizados nos cálculos de média diária, mesmo sem haver o registro completo das 24 horas de dados.

Observando o resultado do teste de *range* (limites) para a umidade relativa do ar na Tabela 9, os anos 2000 a 2006 há somente dados classificados com *flag* 7 para  $UR_{\min}$  e  $UR_{\max}$ , indicando que nesse período não se realizava o registro desses parâmetros. A partir do ano de 2006 até abril de 2018, quando se tem dados de  $UR_{\text{med, max e min}}$ , observa-se que nesse estudo, o ano de 2008 também possui o *flag* 7, mas somente um registro, indicado que um dia desse ano não possui as 24 horas de registros completos (a falha de coleta de dados ocorreu no dia 26/06/2009, às 9:00 h), ocorrendo também nos outros parâmetros (temperatura, radiação, precipitação e vento). Já os *flags* 7 nos anos 2005, 2006 e 2011 correspondem ao período no qual os equipamentos não estavam funcionando, como já foi citado anteriormente.

Nos anos 2009 e 2010, observa-se registros de *flags* 3 na UR (Tabela 9), indicando que o período possui dados que estão acima do limite de resolução do equipamento, mesmo admitindo a porcentagem de erro (Tabela 4). No ano de 2009, os registros acima do valor aceitável variaram entre 102,1 e 102,8% para  $UR_{\text{med}}$ , 102,1 e 103,1% para  $UR_{\max}$ , e 102,1 e 102,6% para  $UR_{\min}$ . Já no ano de 2010 a variação foi entre 102,1 e 104,1% tanto para  $UR_{\text{med}}$  quanto para  $UR_{\max}$  e  $UR_{\min}$ . Esses valores de UR classificados com *flag* 3 foram registradas nos horários de 00:00 às 7:00 h e também entre as 21:00 às 23:00 h, e com maior frequência nos meses de novembro e dezembro, em ambos anos.

Para as variáveis radiação solar global (Rad), precipitação (Prec), velocidade do vento (Vel Vento) e direção do vento (Dir Vento), os dados registrados estão todos dentro da resolução dos equipamentos, recebendo o *flag* 0, ou *flag* 7 nos períodos já citados acima, quando não houve coleta de dados (Tabela 9).





Observa-se que ocorrem períodos de falhas de coleta de dados de direção do vento, nos anos de 2013, 2014, 2015 e 2016 indicando possíveis falhas no anemômetro (Tabela 9). O maior número de *flags* iguais a 7 nos anos de 2015 e 2016, começando no dia 2 de setembro de 2015 até dia 22 de setembro de 2016, devido ao mau funcionamento do anemômetro da estação Hobeco, que foi substituído e só voltou a coletar dados depois de um ano. Para que não houvesse falha de dados diários por esse longo período, foi necessário reativar o anemômetro da estação Campbell, que forneceu os dados.

Além disso, observa-se que na estação Hobeco apresentou dados classificados com *flag 7* no ano de 2015 também nos elementos temperatura, umidade relativa do ar, precipitação e radiação (Tabela 9). Essa falha de coleta de dados ocorreu devido à falta de bateria do equipamento, que voltou a funcionar depois da substituição desta.

Ao comparar a quantidade de *flags* reprovados das duas estações (*flag 3*), observa-se que a estação Hobeco não possui nenhum registro fora da resolução dos equipamentos, diferente da estação Campbell, que possui *flags 3*, na umidade relativa do ar e direção do vento. A estação Campbell também possui uma maior quantidade de *flags 7*, isto é, períodos horários sem dados. Sem as informações detalhadas sobre os motivos que provocaram os períodos sem coleta de dados, pode-se entender que a estação Hobeco passa uma maior segurança na coleta de dados, porém é importante estar sempre atento a possíveis falhas de funcionamento dos equipamentos, que devem passar por manutenção regular para evitar erros de leitura e períodos sem dados, para não comprometer a qualidade/homogeneidade dos dados que compõem o banco de dados da estação.

Na Figura 12 observa-se que no ano de 2011 há um registro de radiação solar ( $Q_g$ ) acima do limite da radiação extraterrestre ( $Q_o$ ), mas como o teste *range*, leva em conta a radiação solar máxima que chega na Terra e não o período do ano, esse dado não foi eliminado porque está dentro do limite da resolução do equipamento ( $1200 \text{ W m}^{-2}$ ). Desta forma é necessário utilizar outro teste para sinalizar esse dado como “reprovado”.

A Tabela 10 apresenta o resultado da análise de qualidade do teste *step* (degrau) para os elementos temperatura e umidade relativa. Para essas variáveis, observa-se que em todos os anos ocorreu o *flag 3*, indicano a ocorrência de uma variação de temperatura e umidade horária maior que o estabelecido de uma medida horária para a outra.

Observa-se que o aumento súbito da temperatura ocorre em horários específicos, que vão das 8:00 às 10:00 h, e a queda súbita da temperatura também ocorre em horários

específicos, das 17:00 às 19:00 h. E para esses mesmos horários observa-se a maior queda e o maior aumento da umidade relativa do ar, respectivamente. Tais variações de aumento ou queda da temperatura pela manhã e à tarde estão diretamente relacionadas ao aumento e redução da radiação solar.

O ano de 2008 foi o que apresentou o maior número de registros de temperatura e umidade reprovados (com *flag* igual a 3) nesse teste. A maior diferença horária de aumento de temperatura observada foi de 12,1 °C, registrada no dia 9 de agosto de 2014, no horário de 8:00 para as 9:00 h; já o maior decréscimo de temperatura foi de -12,0 °C registrado nos dias 29 de outubro de 2008 e no dia 23 de setembro de 2016, no horário de 16:00 para as 17:00 h, e 15:00 para 16:00 h, respectivamente. Para a umidade, foram observados degraus de variações bem maiores, já que no horário das 14:00 para 15:00 h do dia 11 de novembro de 2016, ocorreu um aumento de 50,5%, e um decréscimo de -51,5 % no dia 11 de maio de 2017, no horário de 10:00 para 11:00 h.

Como cerca de 19,5% dos dados foram reprovados nesse teste, talvez seja interessante fazer um estudo para avaliar o aumento dos limites de tolerância de aumento e decaimento da temperatura e umidade. Cheng et al. (2016), avaliando a qualidade dos dados de temperatura horária de várias estações meteorológicas no Taiwan, utilizaram como referência o valor de degrau igual  $\pm 5$  °C; já Meek et al. (1994), avaliando a qualidade de dados meteorológicos de uma estação nos EUA, estabeleceram o valor do degrau de  $\pm 6$  °C .

A radiação é o principal fator para essa variação rápida de temperatura e umidade relativa que ocorrem tipicamente pela manhã e ao final da tarde. A partir das 8:00 h, há uma grande quantidade de energia chegando à superfície, provocando alterações rápidas nos processos aquecimento e evapotranspiração. Da mesma forma, a queda e interrupção da entrada de energia no sistema provoca quedas súbitas da temperatura e aumento da umidade relativa. Além disso, tais variações rápidas podem ser resultado de ocorrência de chuvas ou da passagem de grandes massas de ar. Esses grandes aumentos ou decréscimos da temperatura e da umidade relativa não são necessariamente resultado de erros de medição, especialmente quando observados nas áreas abertas e rurais.

Tabela 10. Número de vezes que cada valor sinalizador (*flag*) foi observado no teste de degrau (*step*), quando aplicado aos dados horários de temperatura e umidade relativa do ar do banco de dados da FAL, FAV/UnB, no período de 2000 a 2018.

Ano	Flag	T <sub>med</sub>	T <sub>max</sub>	T <sub>min</sub>	UR <sub>med</sub>	UR <sub>max</sub>	UR <sub>min</sub>
2000	0	6268	6182	6181	6181	---	---
	3	173	259	260	260	---	---
	7	159	159	159	159	6600	6600
2001	0	8546	8461	8464	8397	---	---
	3	214	299	296	363	---	---
	7	---	---	---	---	8786	8786
2002	0	8526	8458	8453	8415	---	---
	3	234	302	307	345	---	---
	7	---	---	---	---	8786	8786
2003	0	8549	8483	8474	8424	---	---
	3	211	277	286	336	---	---
	7	---	---	---	---	8786	8786
2004	0	8597	8562	8548	8505	---	---
	3	187	222	236	279	---	---
	7	---	---	---	---	8784	8784
2005	0	4093	4053	4058	4020	---	---
	3	49	89	84	122	---	---
	7	4618	4618	4618	4618	8760	8760
2006	0	7730	7685	7690	7611	7590	7517
	3	177	222	217	296	317	390
	7	853	853	853	853	853	853
2007	0	8512	8455	8450	8391	8342	8319
	3	248	305	310	369	418	441
	7	---	---	---	---	---	---
2008	0	8493	8425	8431	8387	8306	8248
	3	289	357	351	395	476	534
	7	2	2	2	2	2	2
2009	0	8542	8470	8460	8360	8305	8246
	3	175	247	257	357	412	471
	7	43	43	43	43	43	43
2010	0	8507	8447	8455	8362	8296	8246
	3	253	313	305	398	464	514
	7	---	---	---	---	---	---
2011	0	5808	5789	5780	5749	5670	5680
	3	151	170	179	210	289	279
	7	2800	2800	2800	2800	2800	2800
2012	0	8581	8546	8520	8437	8324	8354
	3	203	238	264	347	460	430
	7	---	---	---	---	---	---
2013	0	8543	8526	8506	8394	8323	8313
	3	215	232	252	364	435	445
	7	2	2	2	2	2	2
2014	0	8482	8430	8406	8327	8219	8238
	3	249	301	325	404	512	493
	7	29	29	29	29	29	29
2015	0	8361	8354	8293	8241	8170	8186
	3	292	299	360	412	483	467
	7	107	107	107	107	107	107
2016	0	8397	8345	8323	8216	8130	8148
	3	275	327	349	456	542	524
	7	112	112	112	112	112	112
2017	0	8506	8483	8435	8343	8222	8268
	3	254	277	325	417	538	492
	7	---	---	---	---	---	---
2018	0	2852	2817	2825	2770	2758	2741
	3	29	64	56	111	123	140
	7	---	---	---	---	---	---

## 5. CONCLUSÕES

- O uso do *software* R para realizar análise de qualidade de um banco de dados climáticos é viável, mas para isso é necessário ter conhecimento da linguagem para realizar as análises desejadas;
- A utilização de gráficos permite identificação de falhas no banco de dados e também pode ser utilizada como uma forma de identificar os dados mais extremos e errôneos, que estão fora do limite de tolerância;
- Os anos de 2005 e 2011 são os menos indicados para ser usados em, já que possuem uma grande lacuna de dados. Já no ano de 2007, somente os dados de direção do vento não é indicado para utilização. Em relação aos outros anos, todos os dados são considerados confiáveis para serem utilizados;
- A partir dos resultados obtidos na análise de qualidade de dados meteorológicos do período de abril de 2000 a abril de 2018, pode-se afirmar que os dados divulgados estarão menos sujeitos a estarem errados, já que o método, se adotado, é capaz filtrar a maioria dos dados suspeitos e errados.

## CONSIDERAÇÕES FINAIS

- Apesar da coleta de dados meteorológicos ser automática, é necessário estar sempre atento a possíveis falhas de coletas de dados e também de dados fora do padrão. É importante que ocorram manutenções regulares nos equipamentos para evitar possíveis erros, e assim prejudicar o banco de dados;
- Como a análise de qualidade foi feita somente para uma estação, sem haver a comparação com outras instaladas na região; assim, não é possível afirmar se alguns dados que possuem valores muito extremos e fora do intervalo de frequência de ocorrência estão completamente errados. Nessas situações deve-se suspeitar dos dados, e, antes de usá-los ou reprová-los, comparar com outras estações meteorológicas da região;
- Para melhor análise de qualidade do banco de dados, seria interessante dividir todos os elementos climáticos por períodos, como em estação seca (maio a outubro) e estação chuvosa (novembro a abril), já que para o Cerrado, essas estações são bem definidas e variam bastante nos seus registros. Além disso, é necessário envolver nessa análise de qualidade verificações estatísticas para garantir uma maior confiança e homogeneidade dos dados que compõe o banco de dados da estação meteorológica;



## 6. REFERÊNCIAS

- ANUNCIACÃO, Y. M. T.; MAGGIOTTO, S. R.; ANDRÉ, R. G. B. Análise de dados horários do mês de agosto em Brasília, DF, utilizando curvas empíricas suavizadas. In: Anais do **XVII Congresso Brasileiro de Meteorologia**, Gramado, RS, 2012.
- ARNOLD, J. B. ggthemes: temas extras, escalas e gemas para 'ggplot2'. **Versão do pacote R**, 2018.
- ARSLAW, D. C.; ROPKINS, K. Openair - An R package for air quality data analysis. **Environmental Modelling & Software**, v. 27–28, p. 52–61, 2012.
- CHENG, A. R. et al. Quality control program for real-time hourly temperature observation in Taiwan. **Journal of Atmospheric and Oceanic Technology**, v. 33, n. 5, p. 953–976, 2016.
- COELHO, L. L.; STEINKE, E. T. STEINKE, V. A; Caracterização prévia do início e fim da estação chuvosa no Distrito Federal: estudo de caso da Bacia do Lago Paranoá. **Revista Geonorte**, v. 1, n° 5, Edição Especial 2, p. 441–450, 2012.
- FERREIRA, A. L. T. et al. Implementação de sistema para controle de qualidade dos dados meteorológicos da rede de PCDS do INPE. In: Anais do **XIV Congresso Brasileiro de Meteorologia**, Florianópolis, SC, 2006.
- GRIGOLETTI, G.C.; FLORES, M. G.; SANTOS, J. C. P. Tratamento de dados climáticos de Santa Maria, RS, para análise de desempenho térmico de edificações. **Ambiente Construído**, Porto Alegre, v. 16, n. 1, p. 123–141, 2016.
- INMET. **Rede de Estações Meteorológicas Automáticas do INMET**, 2011. Nota Técnica n°001.
- INMET. **Normais Climatológicas do Brasil (1981-2010)**. <http://www.inmet.gov.br/portal/index.php?r=clima/normaisClimatologicas>. Acesso em: setembro de 2018.
- LIMA, R. A. O.; SANTOS, R. P. **Controle de qualidade de dados meteorológicos**. INPE, São José dos Campos, 2009.
- MAGGIOTTO, S. R. et al. Diretrizes para redes automáticas e telemétricas de superfície. **Revista Brasileira de Recursos Hídricos**, v. 12, n. 4, p. 225–240, 2007.

- MAURER, A.; SILVA, C. O surgimento dos estudos sobre climatologia. *In: VI Fórum de Sustentabilidade*, Corede Alto Jacuí, 2017.
- MEEK, D. W.; HATFIELD, J. L. Data quality checking for single station meteorological databases. **Agricultural and Forest Meteorology**, v. 69, n. 1–2, p. 85–109, 1994.
- MILLER, P. A.; BARTH, M. F. Ingest, Integration, Quality Control, and Distribution of Observations From State Transportation Departments Using Madis. *In: (International Conference) Conference On Interactive Information and Processing Systems for Meteorology, Oceanography and Hydrology*, California, 2003.
- PADILHA, G.; VANZELLA, J.; CASTRO, R. **Avaliação dos dados da Estação Meteorológica Automática do IFSC, a partir de um software de qualificação**, 2010.
- PEREIRA, A. R.; ANGELOCCI, L. R.; SENTELHAS, P. C. **Agrometeorologia: Fundamentos e Aplicações práticas**. Livraria e Editora Agropecuária, Guaira, 2002. 478 p.
- PITTIGLIANI, M. Controle de qualidade de dados hidrometeorológicos do SIMEPAR. *In: Anais do XI Congresso Brasileiro de Meteorologia*, Rio de Janeiro, RJ, 2000.
- SHAFER, M. A. et al. Quality assurance procedures in the Oklahoma Mesonet. **Journal of Atmospheric and Oceanic Technology**, v. 17, n. 4, p. 474–494, 2000.
- SENTELHAS, P. C.; MONTEIRO, J. E. B. A. Informações para uma Agricultura Sustentável. **Agrometeorologia dos Cultivos – O Fator Meteorológico na Produção Agrícola**. INMET (Ed.), cap.1, 10 p.
- WICKHAM, H. **ggplot2: gráficos elegantes para análise de dados**. Springer, 2016.
- WORLD METEOROLOGICAL ORGANIZATION (WMO). **Guide to Climatological Practices WMO-N°100**, 2011.