

Análise estatística da concentração de monóxido de carbono na cidade de Ipatinga, Minas Gerais, Brasil

Statistical analysis of carbon monoxide concentration in the city of Ipatinga, Minas Gerais, Brazil

Análisis estadístico de la concentración de monóxido de carbono en la ciudad de Ipatinga, Minas Gerais, Brasil

Recebido: 17/08/2022 | Revisado: 29/08/2022 | Aceito: 05/09/2022 | Publicado: 12/09/2022

Heitor Cardoso de Brito

ORCID: <https://orcid.org/0000-0002-3028-2112>

Instituto Federal de Educação, Ciência e Tecnologia de Minas Gerais, Brasil

E-mail: heitor.cardoso@ifmg.edu.br

Carlos Alberto Mirez Tarrillo

ORCID: <https://orcid.org/0000-0003-3024-915X>

Universidade Federal dos Vales de Jequitinhonha e Mucuri, Brasil

E-mail: carlos.mirez@ufvjm.edu.br

Carlos Henrique Alexandrino

ORCID: <https://orcid.org/0000-0001-9408-4644>

Universidade Federal dos Vales de Jequitinhonha e Mucuri, Brasil

E-mail: carlos.alexandrino@ufvjm.edu.br

Nathalia Stockler Fialho Soares

ORCID: <https://orcid.org/0000-0002-9667-3060>

Universidade Federal dos Vales de Jequitinhonha e Mucuri, Brasil

E-mail: nathaliastockler17@hotmail.com

Resumo

O presente artigo objetiva encontrar a distribuição estatística que melhor se adere ao conjunto de dados de monóxido de carbono atmosférico (CO_{atm}) registrado em microgramas por metro cúbico ($\mu\text{g}\cdot\text{m}^{-3}$) por quatro estações automáticas de monitoramento da qualidade do ar (medições horárias) instaladas na cidade de Ipatinga, Vale do Aço, Minas Gerais (MG), Brasil. A série histórica de dados compreende os períodos de outubro-2013 a outubro-2017 (4 anos), registrados nos bairros Bom Retiro, Cariru e Cidade Nobre, e de janeiro-2014 a março-2016 (2 anos) medido no bairro Veneza. Para reduzir as enormes discrepâncias entre os dados horários, mudou-se o para a análise de quantias semanais de captação de CO_{atm} . Com a aplicação da análise de variância (ANOVA) após o tratamento estatístico de dados, constatou-se que os dados semanais de CO_{atm} coletados nas três estações juntas no biênio 2015-16 eram estatisticamente iguais, e que a distribuição estatística que melhor se aderiu ao conjunto de dados foi a Beta: $[135.000 + 170.000 * \text{BETA}(1,47, 1,72)] \mu\text{g}\cdot\text{m}^{-3}$. Ao se fazer a aproximação pela distribuição Normal (NORM), obteve-se o seguinte valor: $[\text{NORM}(213.000, 41.100)] \mu\text{g}\cdot\text{m}^{-3}$, onde não se rejeitou a hipótese de que os dados seguiam essa distribuição com as captações de CO_{atm} acumuladas nas três estações analisadas. Dados históricos mostram que a concentração de dióxido de carbono (CO_2) tem permanecido constante entre 172 e 300 partes por milhão (ppm). Mas, esta quantia tem aumentado de forma exponencial nos últimos anos.

Palavras-chave: Monóxido/dióxido de carbono; Distribuição estatística; Monitoramento da qualidade do ar; Meio ambiente.

Abstract

This article aims to find the statistical distribution that best adheres to the set of atmospheric carbon monoxide (CO_{atm}) data recorded in micrograms per cubic meter ($\mu\text{g}\cdot\text{m}^{-3}$) by four automatic air quality monitoring stations (hourly measurements) installed in the city of Ipatinga, Vale do Aço, Minas Gerais (MG), Brazil. The historical data series comprises periods, from October-2013 to October-2017 (4 years), recorded in the neighborhoods Bom Retiro, Cariru and Cidade Nobre, and from January-2014 to March-2016 (2 years) measured in the Venice neighborhood. To reduce the huge discrepancies between the hourly data, the one was moved to the analysis of weekly amounts of CO_{atm} capture. With the application of variance analysis (ANOVA) after statistical data processing, it was found that the weekly data of CO_{atm} collected in the three stations together in the biennium 2015-16 were statistically equal, and that the statistical distribution that best adhered to the data set was beta: $[135,000 + 170,000 * \text{BETA}(1.47, 1.72)] \mu\text{g}\cdot\text{m}^{-3}$. When the approximation was made by normal distribution (NORM), the following value was obtained: $[\text{NORM}(213,000, 41,100)] \mu\text{g}\cdot\text{m}^{-3}$, where the hypothesis that the data followed this distribution with the CO_{atm} accumulated in the three stations analyzed was not

rejected. Historical data show that the concentration of carbon dioxide (CO_2) has remained constant between 172 and 300 parts per million (ppm). But this amount has increased exponentially in recent years.

Keywords: Monoxide/carbon dioxide; Statistical distribution; Air quality monitoring; Environment.

Resumen

Este artículo tiene como objetivo encontrar la distribución estadística que mejor se adapte al conjunto de datos de monóxido de carbono atmosférico (CO_{atm}) registrados en microgramos por metro cúbico ($\mu g \cdot m^{-3}$) por cuatro estaciones automáticas de monitoreo de la calidad del aire (mediciones horarias) instaladas en la ciudad de Ipatinga, Vale do Aço (Valle del Acero), Minas Gerais (MG), Brasil. La serie de datos históricos comprende los períodos de octubre-2013 a octubre-2017 (4 años), registrado en los barrios Bom Retiro, Cariru y Cidade Nobre, y de enero-2014 a marzo-2016 (2 años) medido en el barrio de Venecia. Para reducir las enormes discrepancias entre los datos horarios, se trasladó al análisis de las cantidades semanales de captura de CO_{atm} . Con la aplicación del análisis de varianza (ANOVA) después del procesamiento estadístico de datos, se encontró que los datos semanales de CO_{atm} recopilados en las tres estaciones juntas en el bienio 2015-16 eran estadísticamente iguales, y que la distribución estadística que mejor se adhirió al conjunto de datos fue beta: $[135,000 + 170,000 * BETA(1.47, 1.72)] \mu g \cdot m^{-3}$. Cuando se realizó la aproximación por distribución normal (NORM), se obtuvo el siguiente valor: [NORM (213.000, 41.100)] $\mu g \cdot m^{-3}$, donde no se rechazó la hipótesis de que los datos seguían esta distribución con la captación de CO_{atm} acumulada en las tres estaciones analizadas. Los datos históricos muestran que la concentración de dióxido de carbono (CO_2) se ha mantenido constante entre 172 y 300 partes por millón (ppm). Pero esta cantidad ha aumentado exponencialmente en los últimos años.

Palabras clave: Monóxido/dióxido de carbono; Distribución estadística; Monitoreo de la calidad del aire; Medio ambiente.

1. Introdução

Não é de hoje que o homem busca meios para garantir e melhorar a sua sobrevivência no mundo. Desde a antiguidade o ser humano procura formas mais eficazes para se proteger, vestir-se, locomover-se, comunicar-se, alimentar-se etc. A Primeira Revolução Industrial, iniciada na Inglaterra do século XVIII, alterou drasticamente o meio de produção e sobrevivência humana e trouxe o Capitalismo Industrial, onde a indústria, com o auxílio de máquinas, passou a pagar pela força de trabalho. Nos últimos dois séculos, atividades humanas como o desmatamento, uso de carvão e petróleo para obter energia tem causado, dentre outros, um aumento na concentração atmosférica de gases estufa. O pós-Revolução Industrial ocasionou um aumento exponencial dessa concentração. De acordo com a curva de Keeling, seu valor não tem baixado dos 400 p.p.m desde o ano de 2013 (The Keeling Curve, 2018).

O que o homem não se dava conta (ou não se interessava) era o quanto o planeta viria sofrer com tamanho crescimento industrial e conseqüente geração e lançamento de gases poluentes no ar atmosférico. Para sobreviverem, algas marinhas, algumas bactérias e vegetais capturam o CO_2 para realizar o processo de fotossíntese, eliminando-o da atmosfera, produzindo o seu substrato alimentar e liberando o gás oxigênio como subproduto do processo. Além dos malefícios climáticos causados pelos seus excessos, esses gases também prejudicam a saúde humana.

A Norma Regulamentadora número 15 (NR-15), por exemplo, é a diretriz brasileira que trata de limites de tolerância para esses gases. Nela, para uma jornada de trabalho de até 48 horas semanais, o homem não pode se expor por mais de $43.000 \mu g \cdot m^{-3}$ de monóxido de carbono cujo grau de insalubridade é considerado como sendo de máxima caracterização.

O dióxido de carbono é considerado o gás de efeito estufa que mais contribui para o aquecimento global, principalmente devido a um aumento de 35% na sua emissão após o início da era industrial, sendo a sua emissão no Brasil no ano de 2020 de 2,16 bilhões de toneladas brutas de gás carbônico equivalente (CO_2e), trata-se de 3% do total mundial (Brasil, 2022).

As questões ambientais associadas ao CO_2 ganharam destaque nas últimas décadas, devido à problemática das mudanças climáticas. Assim, algumas estratégias de mitigação das emissões desse gás vêm sendo desenvolvidas e o setor industrial, por sua vez, agregou a necessidade de adotar a sustentabilidade a fim de atender o meio ambiente e se sobressair quanto à vantagem competitiva de mercado (Souza, 2022).

2. Metodologia

Quanto à natureza da pesquisa, trata-se de uma pesquisa aplicada, abordagem quali-quantitativa. Uma pesquisa dessa natureza busca descrever a relação os objetivos e os resultados, e não apenas dados numéricos (Praça, 2015). Quanto aos objetivos, trata-se de uma pesquisa descritiva. A pesquisa descritiva exige do investigador uma série de informações sobre o que deseja pesquisar. Esse tipo de estudo pretende descrever os fatos e fenômenos de determinada realidade (Triviños, 1987).

Os dados de monóxido de carbono atmosférico (CO_{atm}) foram registrados em medições horárias de microgramas por metro cúbico de ar ($\mu\text{g.m}^{-3}$) por quatro estações automáticas de monitoramento da qualidade do ar instaladas nos bairros Bom Retiro, Cariru, Cidade Nobre e Veneza, na cidade de Ipatinga/Mingás Gerais (MG), Brasil, dados em planilha do *software MS Excel*[®]. A responsabilidade pelo monitoramento desses gases e de outros poluentes é da empresa siderúrgica USIMINAS, e as séries históricas analisadas compreendem dois períodos. O primeiro de 2 de outubro de 2013 a 30 de outubro de 2017 (aproximadamente 4 anos) registrados nos bairros Bom Retiro, Cariru e Cidade Nobre, e o segundo de 2014 a 26 de março de 2016 (cerca de 2,0 anos) medido no bairro Veneza. Como os registros de hora em hora causavam grandes distorções/diferenças entre eles, optou-se por realizar o tratamento de dados com as informações originais (ou seja, em horas) e realizar estudo das distribuições estatísticas com dados semanais. A análise estatística inicial de dados foi feita com o *software Microsoft Office Excel*[®] 2013. Já a análise de variância (ANOVA) foi realizada pelo *software* estatístico *IBM SPSS Statistics 22* (ou somente SPSS) e o teste da distribuição que melhor iria se aderir ao conjunto de dados foi realizado por meio da ferramenta computacional *Input Analyzer*, disponível no *software Arena 15.1*[®] (versão estudantil).

3. Resultados e Discussão

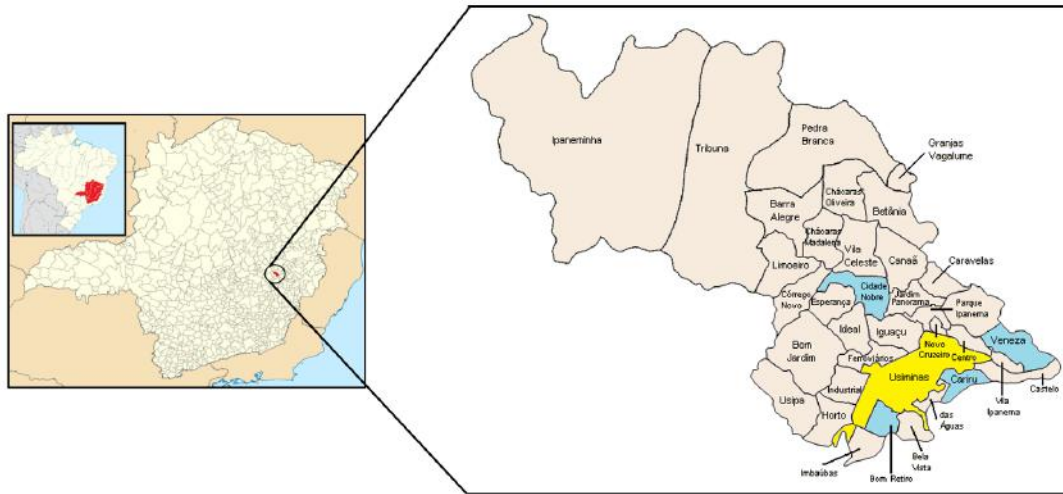
A cidade brasileira de Ipatinga/MG SUL 19°28'6", Oeste 42°32'12" (IBGE, 2018) encontra-se instalada na Região Metropolitana do Vale do Aço. Esta região é um importante polo industrial brasileiro, que abriga duas grandes usinas siderúrgicas e outras grandes empresas de bens de bens de capital (máquinas, metalúrgicas e equipamentos) e de consumo (indústrias alimentícias, papel/celulose etc.), que podem gerar elevada quantidade de poluentes. A Figura 1 mostra a localização do município e dos bairros de Ipatinga em relação ao estado de Minas Gerais, no Brasil.

No intuito de medir a qualidade do ar respirada pelos habitantes da cidade, a empresa siderúrgica USIMINAS instalou ao seu redor quatro estações automáticas de monitoramento da qualidade do ar nos bairros Bom Retiro (EST-1), Cariru (EST-2), Cidade Nobre (EST-3) e Veneza (EST-4). Segundo a Associação Brasileira do Ministério Público do Meio Ambiente – ABRAMPA (2010), essa rede de monitoramento é considerada uma das mais avançadas do país.

Os índices de qualidade do ar são fornecidos em tempo real pela referida empresa e os dados são acompanhados pelo Ministério Público (MP), pela Fundação Estadual do Meio Ambiente (FEAM), pelo Departamento de Meio Ambiente da prefeitura municipal de Ipatinga e pela própria USIMINAS.

De modo geral, as máquinas automáticas fazem leituras de quantidades de ozônio, óxidos de nitrogênio, hidrocarbonetos e de monóxido de carbono. Este último, em especial, é medido a cada 15 minutos pelas estações na unidade de medida em microgramas por metro cúbico de ar ($\mu\text{g.m}^{-3}$), e o sistema central de armazenamento de dados calcula a média horária dessas medições e disponibiliza os dados aos interessados.

Figura 1. Localização do município e dos bairros de Ipatinga/MG, Brasil.



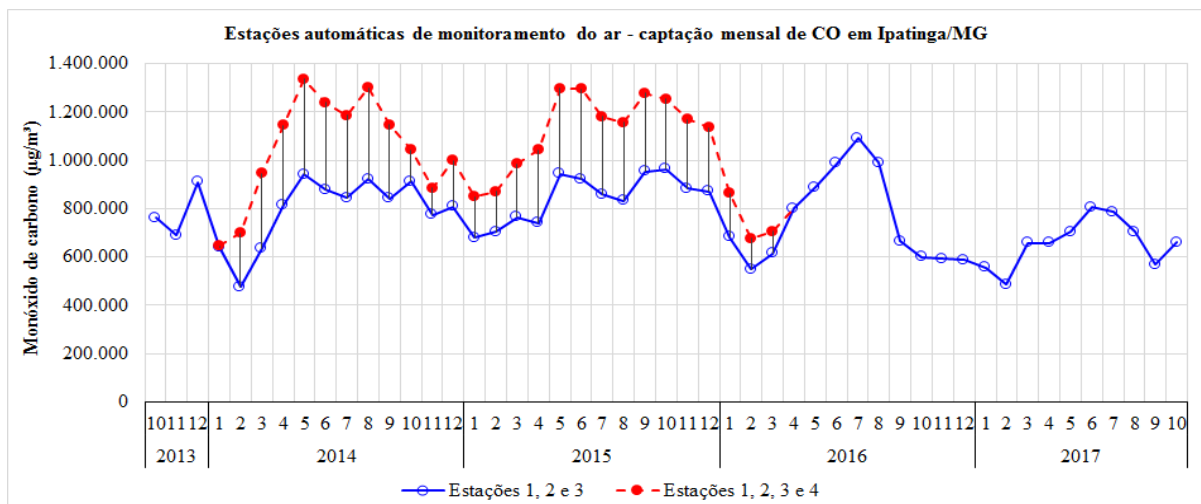
Fonte: Adaptado de Wikipedia (2018).

Ressalta-se, na Figura 1, que os bairros onde se localizam as estações automáticas (marcados em azul) são os que circundam a empresa USIMINAS.

3.1 Análise preliminar de dados das estações automáticas de monitoramento da qualidade do ar

A análise preliminar dos dados de monóxido de carbono atmosférico compreendeu, inicialmente, a visualização e compreensão de seus comportamentos ao longo dos meses para tentar identificar o aparecimento de uma possível tendência ou sazonalidade. A Figura 2 mostra o comportamento inicial dos dados iniciais.

Figura 2. Registros mensais de CO_{am} pelas estações automáticas de monitoramento do ar (comparação dos totais com três e com quatro estações) $\mu g.m^{-3}$



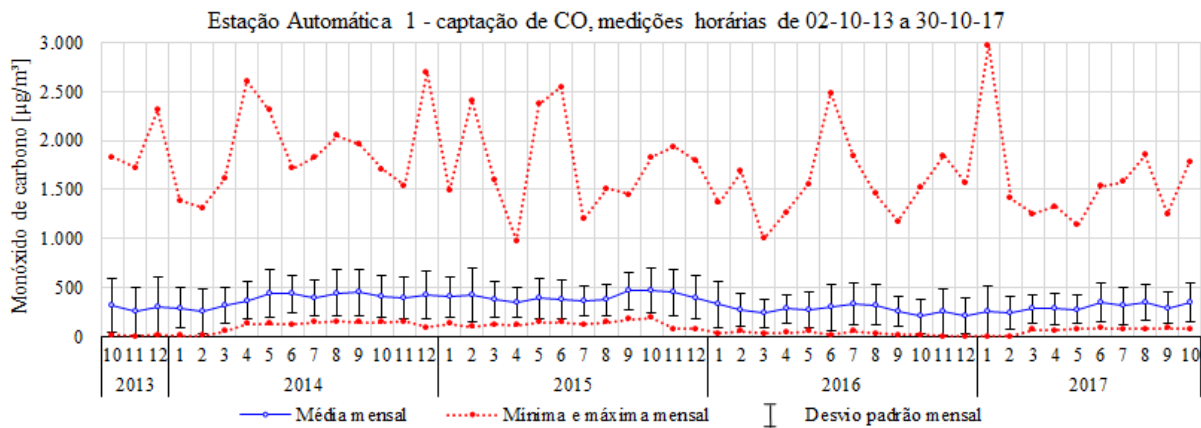
Fonte: Elaboração própria (2018).

A estação de monitoramento nº 4 (bairro Veneza) operou por pouco mais de dois anos (entre os dias de 31 de janeiro de 2014 e 26 de março de 2016), e o sua operação foi descontinuada. Observa-se, na Figura 2, que a inserção de seus valores no conjunto de dados poderia influenciar de forma negativa na elaboração de estudos matemáticos.

Dessa forma, para evitar possíveis discrepâncias nas análises estatísticas, optou-se por descartar o histórico de dados fornecidos pela quarta estação. Ressalta-se que não se sabe os motivos que levaram a descontinuidade deste equipamento no monitoramento de dados, e que a investigação das causas que levaram a isto não faz parte do escopo do estudo.

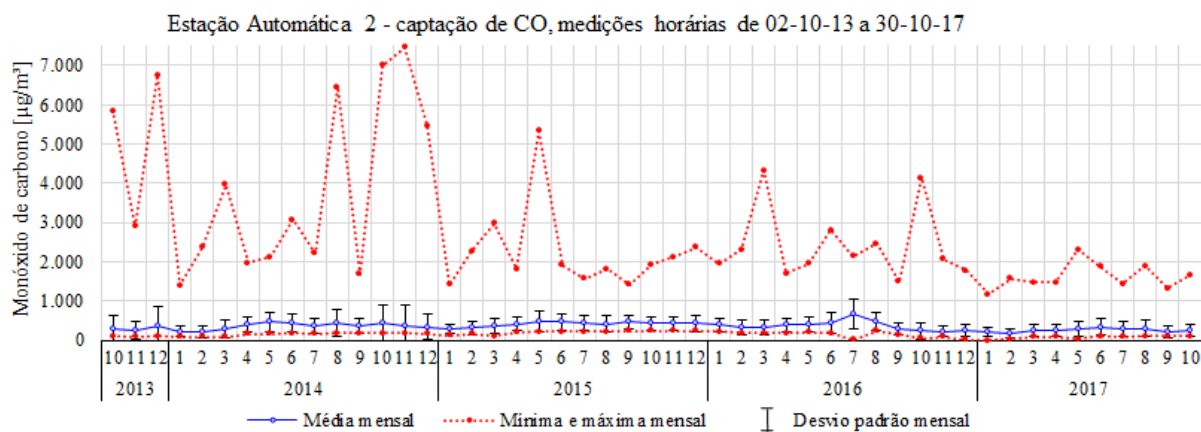
Dando sequência às análises, a etapa seguinte consistiu na visualização e compreensão das captações de CO_{atm} pelas três máquinas restantes. Tal análise compreendeu busca por informações estatísticas como média, desvio padrão, mínima e máxima registradas de hora em hora, e resumido mês a mês. As Figuras 3, 4 e 5 mostram o comportamento dos dados analisados.

Figura 3. Dados estatísticos iniciais das medições horárias de monóxido de carbono registradas no bairro Bom Retiro.



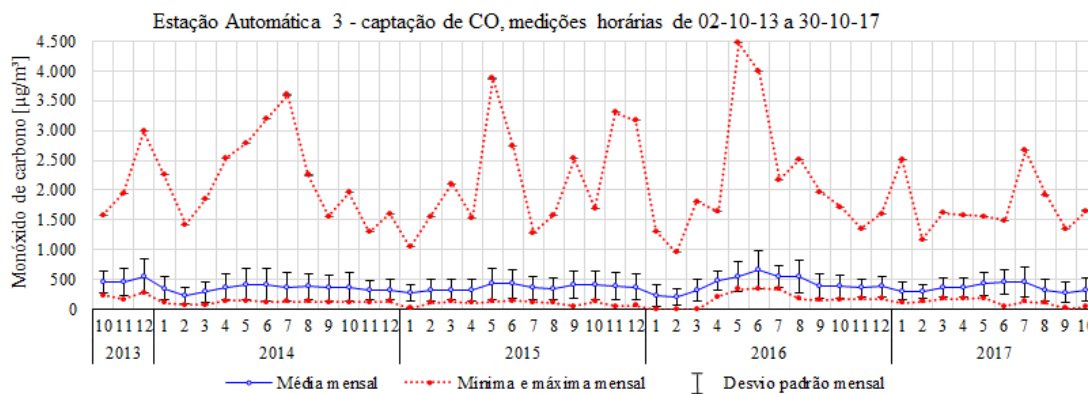
Fonte: Elaboração própria (2018).

Figura 4. Dados estatísticos iniciais das medições horárias de monóxido de carbono registradas no bairro Caruru.



Fonte: Elaboração própria (2018).

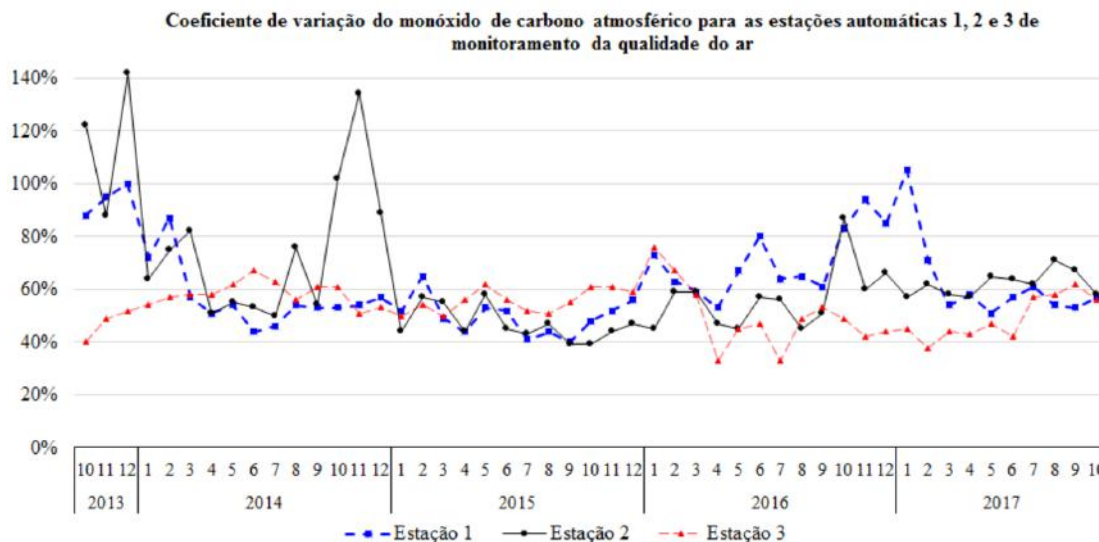
Figura 5. Dados estatísticos iniciais das medições horárias de monóxido de carbono registradas no bairro Cidade Nobre.



Fonte: Elaboração própria (2018).

Pode-se observar, nas Figuras 3, 4 e 5 a ocorrência de grande variabilidade das medições horárias do monóxido de carbono. As máximas mensais diferem (e muito) das médias de cada período e os desvios padrão parece ser bem grandes em relação às médias. Para comprovar esta afirmação, obteve-se o coeficiente de variação (CV) de cada uma das estações. A Figura 6 mostra o comportamento das variações estatísticas sumarizadas do CV horário nas estações automáticas de monitoramento da qualidade do ar selecionadas.

Figura 6. Coeficiente de variação para medições horárias do monóxido de carbono atmosférico nas estações automáticas de monitoramento da qualidade do ar em Ipatinga MG.



Fonte: Elaboração própria (2018).

Verifica-se, na Figura 6, que a estação 03 aparenta apresentar coeficientes de variação mais uniformes ao passo que a estação 01 exibe coeficientes mais altos no final de 2013 e volta a aumentar no final do ano de 2016. Já a estação 02 indica grandes variações no CV nos dois primeiros anos de observação e depois mostra um comportamento mais linear. Maiores detalhes a respeito desses desempenhos horários e sumarizados mensalmente podem ser resumidos na Tabela 1.

Tabela 1. Coeficientes de variação para o monóxido de carbono atmosférico (médias horárias mensais) registrados em quatro estações de monitoramento em Ipatinga/MG.

Ano	Coeficiente de variação (%) para dados médios horários mensais		
	Estação 1 – Bom Retiro	Estação 2 – Cariru	Estação 3 – Cidade Nobre
2013	94	117	47
2014	57	74	58
2015	50	47	56
2016	71	56	50
2017	62	62	49
Média global	62	63	53

Fonte: Elaboração própria (2018).

Em função dessas grandes discrepâncias nos coeficientes de variação calculados para os equipamentos de medição (Tabela 1), tornou-se necessário elaborar um novo critério para mitigar essas amplas dispersões. Uma das metodologias seria a construção de um gráfico *box plot* (ou diagrama em caixa) para verificar a presença de pontos extremos (ou *outliers*); caso surjam, os mesmos devem ser eliminados do modelo.

3.2 Tratamento dos dados de monóxido de carbono coletados nas estações analisadas

No banco de dados, constatou-se a existência de valores não registrados pelos equipamentos. As medições não mensuradas podem se referir a falhas na transmissão dos dados, defeitos diversos, manutenções nos equipamentos dentre outros. Para melhorar a confiabilidade numérica, optou-se por fazer substituições das quantias não computados por médias de acordo com os critérios mostrados na Tabela 2.

Tabela 2. Procedimentos adotados para substituir os dados não coletados pelas estações automáticas de monitoramento da qualidade do ar.

Duração	Procedimento
1 hora	Média de 08 horas anteriores e 08 posteriores ao acontecimento.
2 a 8 horas seguidas	Média móvel das 08 horas anteriores ao acontecimento e 08 horas posteriores, sendo o primeiro número posterior contabilizado a partir da primeira medição validada. A média anterior ao dado da segunda hora não medida em diante não exclui o valor calculado referente ao instante anterior (ou seja, tirou-se a média de oito horas acima dela).
9 a 24 horas	Média móvel das 48 horas posteriores ao acontecimento, sendo o primeiro valor incluído a partir da primeira medição validada.
Acima de 24 horas	Média móvel de 10 dias posteriores (1/3 do mês, ou 240 horas), supondo-se que houve algum grande reparo/manutenção no equipamento.

Fonte: Elaboração própria (2018).

Houve casos em que ocorreram falhas em medições em horários muito próximos uns dos outros. Nesta situação, calculou-se a média anterior ao evento e, na posterior, inseriu-se somente os dados validados (ou seja, excluiu-se a média do dado faltante para não gerar referências circulares no *software MS-Excel*. Exemplo: a estação não realizou uma coleta de dado em um determinado momento e, três horas depois, não a fez novamente; inseriu-se, desta forma, oito pontos de coleta posteriores à primeira medição excluindo-se o segundo ponto de falha. A Tabela 3 mostra a quantidade de horas registradas e não registradas – e seus percentuais equivalentes para cada ano nas três estações automáticas de monitoramento da qualidade do ar.

Tabela 3. Número de horas registradas e não registradas em três estações automáticas de monitoramento da qualidade do ar em Ipatinga/MG.

Descrição	Quantidade e percentual de horas registradas e não registradas – h (%)					
	2013	2014	2015	2016	2017	Total
Estação 1 (Bom Retiro)						
Registradas – qtde	2.174	8.591	8.574	8.543	7.092	34.974
Não registradas – qtde (%)	10 (0,5)	167 (1,9)	185 (2,2)	241 (2,8)	180 (2,5)	783 (2,2)
Estação 2 (Cariru)						
Registradas – qtde	2.176	8.619	8.560	8.267	7.024	34.646
Não registradas – qtde (%)	8 (0,4)	139 (1,6)	199 (2,3)	517 (6,3)	248 (3,5)	1.111 (3,2)
Estação 3 (Cidade Nobre)						
Registradas – qtde	2.181	8.564	8.645	8.606	7.202	35.198
Não registradas – qtde (%)	3 (0,1)	194 (2,3)	114 (1,3)	178 (2,1)	70 (1,0)	559 (1,6)
Estações 1, 2 e 3						
Registradas – qtde	6.531	25.774	25.779	25.416	21.318	104.818
Não registradas – qtde (%)	21 (0,3)	500 (1,9)	498 (1,9)	936 (3,7)	498 (2,3)	2.453 (2,3)

Fonte: Elaboração própria (2018).

Observa-se na Tabela 3, que a configuração inicial dos conjuntos de dados – sem a substituição das medições horárias não registradas – evidenciou a presença de 2.453 horas não registradas e 104.818 horas assinaladas nas três estações escolhidas. Estes valores não catalogados correspondem a cerca de 2,3% do total de dados amostrais.

Ao se observar os percentis totais de horas não registradas nas estações 1, 2 e 3, constata-se que os valores foram, respectivamente, 2,2%, 3,2% e 1,6%. Ou seja, a estação 2, instalada no bairro Cariru, foi a que mais teve problemas com a falta de registros horários. Nota-se também que o maior percentual de dados não colhidos ocorreu no ano de 2016 (6,3%), nesta mesma estação. Pode-se ainda afirmar que o ano de 2013, apesar de ter os registros iniciados no mês de outubro, obteve resultados próximos uns dos outros e abaixo de 0,5%.

Dada a identificação dessas lacunas não preenchidas em função dos dados não coletados pelas máquinas automáticas, a etapa seguinte compreendeu em completar os horários não coletados por valores médios de CO_{atm} de acordo com o procedimento mostrado na Tabela 2. Findada esta etapa, buscou-se eliminar os pontos discrepantes por meio de outros tratamentos estatísticos. Desta vez, a remoção dos *outliers* foi substituída pela média de três dias antes e três dias depois do seu acontecimento. A compilação dos registros antes, durante e depois da realização dos tratamentos é mostrada na Tabela 4.

Tabela 4. Quantidade total de CO_{atm} registrado nas estações automáticas de monitoramento do ar e verificação do número de *outliers* antes, durante e depois dos tratamentos estatísticos

Estação	Total de CO_{atm} emitido ($\mu g.m^{-3}$)			% de aumento $\left(\frac{(3)}{(1)} - 1\right)$	Nº de <i>outliers</i> encontrados		
	Antes (1)	Tratamentos			Antes (A)	Tratamentos	
		Durante (2)	Depois (3)			Durante (B)	Depois (C)
EST-1	11.854.338	12.106.630	Idem	2,13	-	-	-
EST-2	12.215.984	12.603.056	12.539.921	2,65	1	1	-
EST-3	13.522.880	13.746.824	13.662.823	1,03	10	7	-
Total	37.593.202	38.456.510	38.309.374	1,91	11	8	-

Fonte: Elaboração própria (2018).

- (1) Total de CO_{atm} capturado pelas estações automáticas de monitoramento da qualidade do ar.
 - (2) Total de CO_{atm} nas estações após o preenchimento das lacunas não registradas pelos equipamentos.
 - (3) Total de CO_{atm} nas estações após a substituição de dados discrepantes por médias.
- (A) Dados originais, conforme registradas nas estações de monitoramento da qualidade do ar.
 - (B) Substituição de medições não realizadas por médias estipuladas.
 - (C) Remoção dos *outliers* e substituição pela média de três dias antes e três dias depois do seu acontecimento.

Verifica-se, na Tabela 4, que a estação 01 não apresentou a ocorrência de pontos discrepantes nem antes e nem durante a realização dos tratamentos. Após a execução dos procedimentos estatísticos ela teve um aumento de 2,13% na concentração de monóxido de carbono.

Já em relação aos equipamentos 02 e 03, observa-se que o número total de *outliers* reduziu de 11 para 08 após a realização da primeira etapa de tratamento dos dados. Esses oito pontos – um da estação 02 e sete da estação 03 – foram novamente tratados e em decorrência disso, registrou-se aumentos de 2,65% e 1,03% de CO_{atm} . Ressalta-se que o bairro Cariru (EST-2), apesar de ter sido a estação que mais teve problemas com as medições horárias, foi a que apresentou apenas um ponto discrepante antes e após o preenchimento das lacunas por valores médios.

Ainda na Tabela 4, observa-se que a quantidade total de monóxido de carbono atmosférico aumentou em $716.172 \mu\text{g}\cdot\text{m}^{-3}$. Ou seja, o total de CO_{atm} registrado no banco de dados inicial teve um aumento de 1,91% em relação aos $37.593.202 \mu\text{g}\cdot\text{m}^{-3}$ disponibilizados na planilha original.

Passou-se a observar que o registro de informações horárias apresentava grande discrepância entre si. A estação automática número 2, por exemplo, no bairro Cariru, apresentou um pico de mais de $7.000 \mu\text{g}\cdot\text{m}^{-3}$ em novembro-2014 ao passo que a média diária não ultrapassou a casa dos $500 \mu\text{g}\cdot\text{m}^{-3}$. Para contornar esta situação, e dado o grande volume de medições horárias optou-se por reanalisar as medidas, agora transformando/somando as medidas horárias em períodos semanais.

Os motivos da opção de se estudar o comportamento deste período semanal ao invés de somas diárias residem em dois fatos principais:

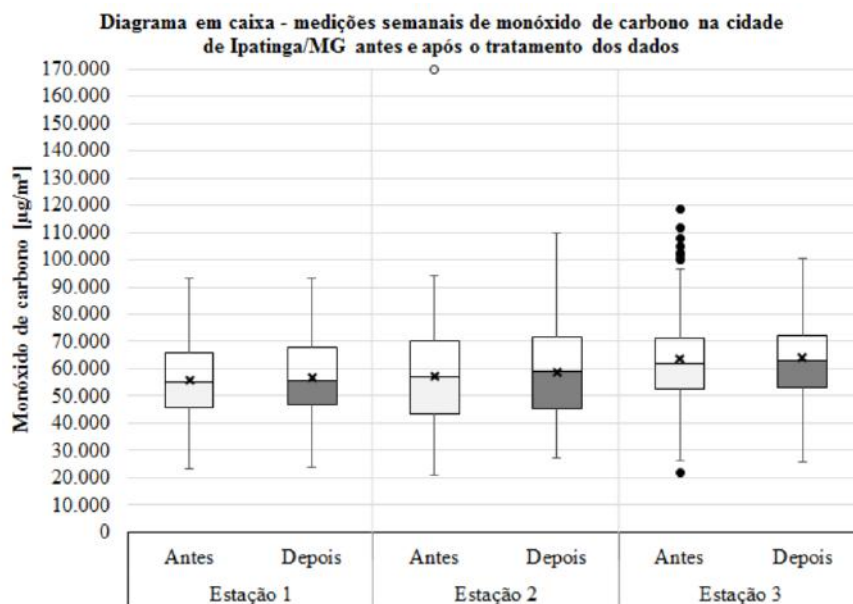
- Possibilitar maior margem de erro em relação à dispersão dos dados e, em consequência, reduzir a quantidade desses valores indesejáveis com o objetivo de não perder a característica original da amostra;
- O monóxido de carbono e outros poluentes podem permanecer na atmosfera por décadas, séculos ou mesmo milhares de anos (Freitas, 2011). Ou seja, o CO_{atm} pode não ser consumido ou transformado em um curtíssimo espaço de tempo (como um mês, por exemplo).

Então, para melhor se adequar à análise de variância – ANOVA, a ser trabalhada posteriormente, realizou-se três procedimentos de tratamentos amostrais, a saber:

1. Análise das quantidades de *outliers* encontrados nos valores originais, antes dos tratamentos estatísticos;
2. Substituição dos dados não coletados por outras medidas estatísticas conforme as regras mostradas na Tabela 2, com o objetivo de reduzir a ocorrência desses pontos. É importante ressaltar que as quantias não colhidas pelas estações automáticas podem ter ocorrido por diversos motivos como, por exemplo, de falhas internas, falhas na transmissão de dados, realização de manutenções nos aparelhos etc.;
3. Análise minuciosa do conjunto de dados através da substituição das quantias discrepantes por outros valores representativos. Como exemplo, pode-se citar a substituição de um *outlier* pela média das últimas 16 horas de medição de modo a eliminar o número indesejável sem prejudicar a amostra.

Compilou-se novamente os dados de CO_{atm} – agora em períodos semanais para verificar o comportamento da dispersão empírica dos dados. A Figura 7 mostra o resultado obtido com o auxílio da elaboração de um diagrama *boxplot*. Ressalta-se que o histórico das medições (agora semanais) sofreram novo tratamento de dados.

Figura 7. *Box plot* para medições semanais de monóxido de carbono atmosférico durante e após o tratamento dos dados.



Fonte: Elaboração própria (2018).

Na Figura 7, observa-se que o impacto do tratamento de dados realizado nas informações coletadas pela estação 01 foi praticamente insignificante. A dispersão total dos dados – ou seja, a diferença entre a máxima e a mínima registradas – reduziu somente 0,9%. Em relação ao comportamento gráfico, nota-se que em ambas as situações o intervalo inter-quartil (IIQ – diferença entre o 1ª e 3º quartil), a média e a mediana permaneceram bem centralizados, mantendo um comportamento bem próximo ao da distribuição Normal (média bem centralizada e valores uniformemente distribuídos). Ao se analisar a parte interna do *boxplot*, constata-se que o tamanho do IIQ aumentou em 5,65% (de 19.956,12 $\mu\text{g}/\text{m}^3$ para 21.083,13 $\mu\text{g}/\text{m}^3$). Ainda em relação à estação 01, verifica-se que o maior valor não *outlier* permaneceu o mesmo, limitando-se a 93.304 $\mu\text{g}/\text{m}^3$, e que o menor valor não *outlier* sofreu um aumento de somente 639,27 $\mu\text{g}/\text{m}^3$ em relação aos 23.167,98 $\mu\text{g}/\text{m}^3$ iniciais.

Já os dados referentes à estação 02 apresentaram mudanças um pouco mais sutis. Nota-se que a sombra inferior reduziu um pouco e que a sombra superior aumentou consideravelmente após o tratamento de dados, fazendo com que a distribuição se tornasse visualmente assimétrica à direita. Esse aumento do tamanho da sombra superior (60,5% ao todo, de 23.836 para 38.260 $\mu\text{g}/\text{m}^3$ e chegando a um limite superior próximo dos 110.000 $\mu\text{g}/\text{m}^3$) se deve aos ajustes realizados para remover um *outlier* extremo e na substituição dos dados não registrados pelo equipamento conforme mencionados anteriormente. Ainda, o intervalo inter-quartil reduziu em 2,74% (de 26.643,24 para 25.912,65 $\mu\text{g}/\text{m}^3$), o valor mínimo encontrado aumentou em 29,05% (de 21.150,31 para 27.295,45 $\mu\text{g}/\text{m}^3$) e a máxima reduziu em 35,42% (de 170.000 para 110.000 $\mu\text{g}/\text{m}^3$ aproximadamente).

Ao se analisar o terceiro aparato de medição, observa-se a ocorrência de aumentos de 4,5% e 10,1% nas sombras inferior e superior, respectivamente. Houve também pequenas oscilações nos valores referentes ao primeiro e terceiro quartis, média e mediana; as diferenças foram, respectivamente, de 788, 1.276, 652 e 934 $\mu\text{g}/\text{m}^3$. Porém, a mudança mais marcante configurada neste diagrama em caixa foi a remoção dos oito pontos discrepantes (um abaixo do limite inferior e sete acima do maior valor não *outlier*), lembrando que a distribuição dos dados em torno da medida central também aparenta ter um comportamento semelhante

ao da distribuição Normal. Ademais, a Tabela 5 traz maiores informações a respeito das observações estatísticas (estatística descritiva) conseguidas com a exploração dos dados referentes às três estações de monitoramento.

Tabela 5. Estatística descritiva a respeito dos dados de CO_{atm} nas estações de monitoramento da qualidade do ar 01, 02 e 03.

Descrição	Estação 1		Estação 2		Estação 3	
	Antes	Depois	Antes	Depois	Antes	Depois
Mínimo (mín)	23.167,98	23.807,25	21.150,31	27.295,45	21.854,61	25.779,87
Média	55.654,17	56.838,64	57.352,04	58.873,06	63.489,59	64.142,04
Mediana	55.295,86	55.790,88	57.249,70	58.812,49	62.040,89	62.975,00
Máximo (máx)	93.304,24	93.304,24	169.895,33	109.726,65	118.369,20	100.615,00
Dispersão total (máx – mín)	70.136,26	69.496,99	148.745,02	82.431,20	96.514,59	74.835,13
% Dispersão (depois ÷ antes)	-	-0,91%	-	-44,58%	-	-22,46%
1º Quartil (Q1)	45.729,98	46.621,76	43.475,02	45.553,58	21.854,61	25.779,87
3º Quartil (Q3)	65.686,10	67.704,89	70.118,26	71.466,23	118.369,20	100.615,00
Intervalo inter quartis (Q3 – Q1)	19.956,12	21.083,13	26.643,24	25.912,65	96.514,59	74.835,13
Erro padrão	991,33	980,62	1.259,03	1.138,15	1.154,36	1.088,01
Modo	#N/D	#N/D	#N/D	#N/D	#N/D	#N/D
Desvio padrão	14.467,94	14.311,64	18.374,99	16.610,70	16.847,39	15.879,03
Curtose	-0,34	-0,41	5,25	-0,48	0,58	-0,13
Assimetria	0,2	0,24	1,05	0,22	0,61	0,39
Intervalo	70.136,26	69.496,99	148.745,02	82.431,20	96.514,59	74.835,13
Contagem	213	213	213	213	213	213
Nível de confiança (95,0%)	1.954,12	1.933,01	2.481,83	2.243,53	2.275,50	2.144,71
Soma	11.854.338	12.106.630	12.215.984	12.539.961	13.523.283	13.662.254
Variância	209.321.324	204.822.903	337.640.308	275.915.335	283.834.681	252.143.507

Fonte: Elaboração própria (2018).

Essa verificação mostra a enorme importância de se realizar uma análise/tratamento inicial dos dados com certo rigor para depois utilizar os dados na obtenção de resultados efetivos. Ressalta-se que, quanto menor a dispersão dos dados, mais significativa é a amostragem. Logo, para o resultado obtido, as conclusões podem apresentar grandes variações nos erros amostrais; para contornar essa situação, deve-se realizar a análise de variância (ANOVA), método este muito mais robusto que a simples obtenção do diagrama em caixa.

3.3 Aplicação da análise de variância nos dados de monóxido de carbono atmosférico

Triola (2013) define a análise de variância (ANOVA) como sendo um método para testar a igualdade de três ou mais médias populacionais por meio da análise das variâncias obtidas nas amostras. Dessa forma, quanto maior for a quantidade de horas não registradas, reduzindo o tamanho amostral, e quanto maior for a frequência de *outliers* registrados nos medidores das estações atmosféricas, maior será a interferência dos dados ao se realizar a ANOVA, daí a importância de se fazer o tratamento dos dados conforme já mostrado na Tabela 2. Poder-se-ia, por exemplo, registrar uma possível rejeição de uma hipótese em função de uma menor quantidade de dados registrados ou de uma maior dispersão dos mesmos em função dos pontos discrepantes.

Para implementar o estudo da ANOVA, teve-se que fazer ajustes nos dados iniciais para análise do modelo em função do número de semanas no ano e de limitações do *software* utilizado. Esses ajustes foram:

1. O calendário Gregoriano leva em consideração o ciclo solar, que é de 365 dias e 08 horas. Adotou-se, neste trabalho um período fixo de 365 dias, ou aproximadamente 52,14 semanas (Galileu, 2016);
2. Cada semana é interpretada no *software* estatístico SPSS como sendo um tratamento. O sistema computacional se limita a calcular até 50 agrupamentos de dados (isto é, 50 semanas). Assim, estipulou-se que:

- A semana tem um período de 08 dias, sendo que o primeiro dia se inicia a partir do primeiro dado amostral (dia 02 de outubro de 2013, correspondente ao 2º dia da 40ª semana, de acordo com o calendário Gregoriano, ou, neste ajuste, à 35ª semana);

- Ao fim e/ou início de cada ano, a semana deve (obrigatoriamente) finalizar com 08 um período de dias. Exemplo 01: a 45ª semana de 2013 compreendeu, nessa nova formulação, os dias 21 a 28 de dezembro daquele ano; para que a 46ª semana não concluísse o ano com somente 03 dias, optou-se por reiniciar a contagem; assim, os dias 29-dezembro-2013 a 05-janeiro-2014 passaram a fazer parte da primeira semana de 2015, evitando-se, assim, que o menor período causasse uma possível interferência na comparação das médias na ANOVA. Exemplo 02: a 46ª semana de 2015 iniciou-se no dia 27 de dezembro daquele ano e finalizou-se no dia 03 de janeiro de 2016;

A necessidade de se adaptar o período semanal como tendo 8 dias em função das limitações do *software* proporcionou encontrar as seguintes quantidades de tratamentos (Tabela 6).

Tabela 6. Número de semanas (tratamentos) encontrados em cada ano.

Estação	Ano					Total
	2013	2014	2015	2016	2017	
1, 2 e 3	11	46	46	45	38	186

Fonte: Elaboração própria (2018).

Na Tabela 6, observa-se que os anos 2013 e 2017 nas estações 01, 02 e 03 apresentaram 11 e 38 tratamentos, respectivamente. Isso se deve ao histórico de dados recebidos uma vez que a coleta de dados realizada compreendeu o período de 02 de outubro de 2013 a 30 de outubro de 2017.

Por se tratar da ocorrência de períodos diferentes, optou-se por realizar diversas análises de variância com o objetivo de encontrar o modelo de dados que obtivesse a maior aderência. Dessa forma, a hipótese inicial H_0 deve ser rejeitada se o *p-value* encontrado for menor que o nível de significância α de 5%. Em complemento, a procura pela formação de subgrupos se deu com a aplicação do teste de Tuckey a também 5% de significância. A formulação do teste de hipóteses para poder interpretar os dados fornecidos pela ANOVA é dado por:

- H_0 : $\mu_1 = \mu_2 = \dots = \mu_j$ – as médias (tratamentos) semanais das estações são estatisticamente iguais.
- H_1 : pelo menos uma das médias difere das demais.

A Tabela 7, mostra os resultados encontrados em função do período observado (tamanho amostral) e da análise do *p-value* juntamente com o teste de Tukey para formação de subgrupos.

Tabela 7. Análise de variância para valores semanais de CO_{atm} .

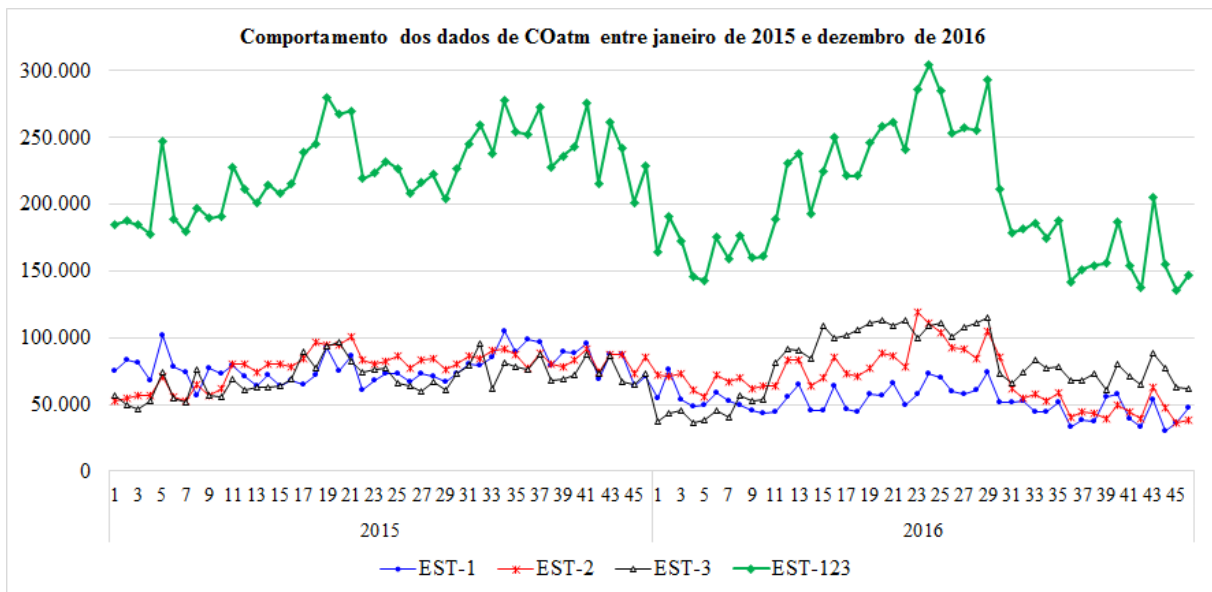
Período	Análise de variância para CO_{atm} <i>p-value</i> encontrado (subgrupos formados)			
	Estação 1	Estação 2	Estação 3	Estações 1, 2 e 3
Out-2013 a out-2017	0,990 (1)	0,028 (3)	0,000 (6)	0,000 (2)
Jan-2014 a out-2017	0,987 (1)	0,041 (3)	0,000 (4)	0,000 (2)
Jan-2014 a dez-2016	1,000 (1)	0,015 (4)	0,000 (4)	0,002 (2)
Jan-2014 a dez-2015	0,287 (1)	0,000 (5)	0,041 (2)	0,001 (3)
Jan-2015 a dez-2016	1,000 (1)	0,606 (1)	0,029 (2)	0,579 (1)

Fonte: Elaboração própria (2018).

Na Tabela 7, o período que obteve maior aderência em relação à análise de variância foi o de janeiro de 2015 a dezembro de 2016. Nesse ensaio, com exceção da estação 03, os demais resultados apresentaram um *p-value* acima de 0,55, inclusive para a somatória dos dados semanais nos três aparatos.

Após esse reajuste, pôde-se dar início a novo análise dos dados ajustados para o período de janeiro de 2015 a dezembro de 2016. A Figura 8 mostra o comportamento geral dos dados ao longo desse período.

Figura 8. Comportamento geral dos dados semanais de CO_{atm} após o ajuste fino dos dados.



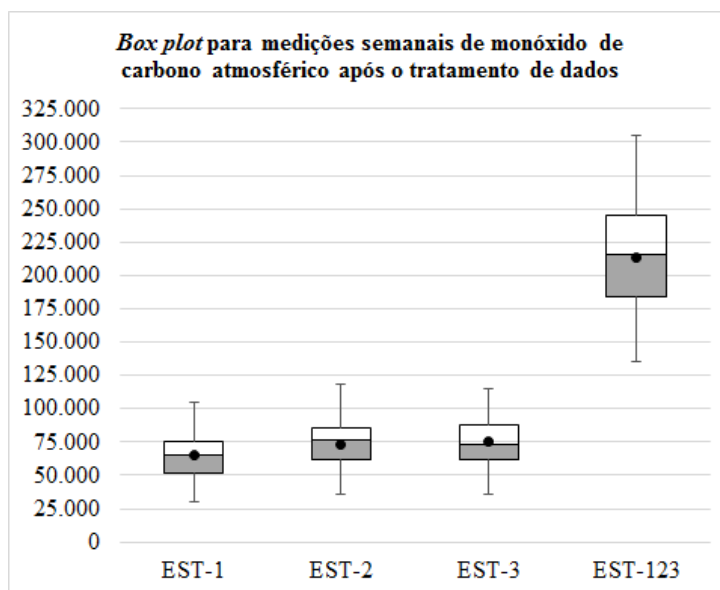
Fonte: Elaboração própria (2018).

Na Figura 8, há maior linearidade dos dados ao longo do ano de 2015 e aumento do volume de CO_{atm} emitido a partir do ano de 2016. Ainda assim, a ANOVA se mostrou favorável à análise dos dados com exceção da estação 03.

3.4 Modelos de previsão de dados para demandas futuras

Realizou-se a análise variância para verificar se havia diferença estatística significativa entre os dados semanais coletados nas três estações de monitoramento da qualidade do ar. Conforme observado, o período que melhor se adaptou ao modelo foi o de dados semanais de janeiro de 2015 a dezembro de 2016. A Figura 9 mostra o *box plot* devidamente atualizado para esse período em vigência.

Figura 9. Gráfico *box plot* para dados semanais de CO_{atm} após o ajuste fino de dados. Período: janeiro-2015 a dezembro-2016.



Fonte: Elaboração própria (2018).

Na Figura 9, os dados aparentam estar igualmente dispersos (média e mediana bem próximas), demonstrando certa proximidade com a distribuição Normal.

Field (2011) propõe a realização do teste de Levene para verificar a homogeneidade das variâncias. O resultado das análises levando-se em consideração o fator ‘Ano’ com um nível de 5% de significância é mostrado na Tabela 8.

Tabela 8. Teste de Levene para igualdade de variâncias dos dados semanais de CO_{atm} .

Estação	Desvio-padrão amostral semanal		Teste de Levene	
	2015	2016	Z	p-value
Estação 1	11.387,70	10.736,06	0,474	0,493
Estação 2	12.449,30	21.402,27	9,573	0,003
Estação 3	12.308,46	24.496,04	24,347	0,000
Estações 1, 2 e 3	28.707,74	47.621,39	17,196	0,000

Fonte: Elaboração própria (2018) com dados do *software* SPSS.

Na Tabela 8, somente a estação 01 não apresentou diferenças significativas nas variâncias de acordo com o teste de Levene. Ou seja, não se rejeitou a hipótese de igualdade de variâncias entre os anos 2015 e 2016 para os dados da estação 01. Já para os demais testes, isso não aconteceu.

O teste não-paramétrico de Kruskal-Wallis para amostras independentes é uma alternativa para a ANOVA de fator único caso as suposições de os dados seguirem a distribuição Normal for violada. Ressalta-se que análise de variância tem como premissas (dentre outras) os fatos de que as populações das quais as amostras são extraídas apresentem distribuições aproximadamente normais e de que as populações apresentem o mesmo desvio-padrão (Mann, 2009).

O teste de Kruskal-Wallis testa as hipóteses de que as populações possuem funções de distribuição iguais (H_0) contra a de que pelo menos duas populações possuem diferentes funções de distribuição (H_1). Dessa forma, a realização desse teste no *software* SPSS evidenciou os seguintes valores (Tabela 9):

Tabela 9. Teste de Kruskal-Wallis para identificação de igualdade de funções de distribuição em dados semanais de CO_{atm} (2015-2016).

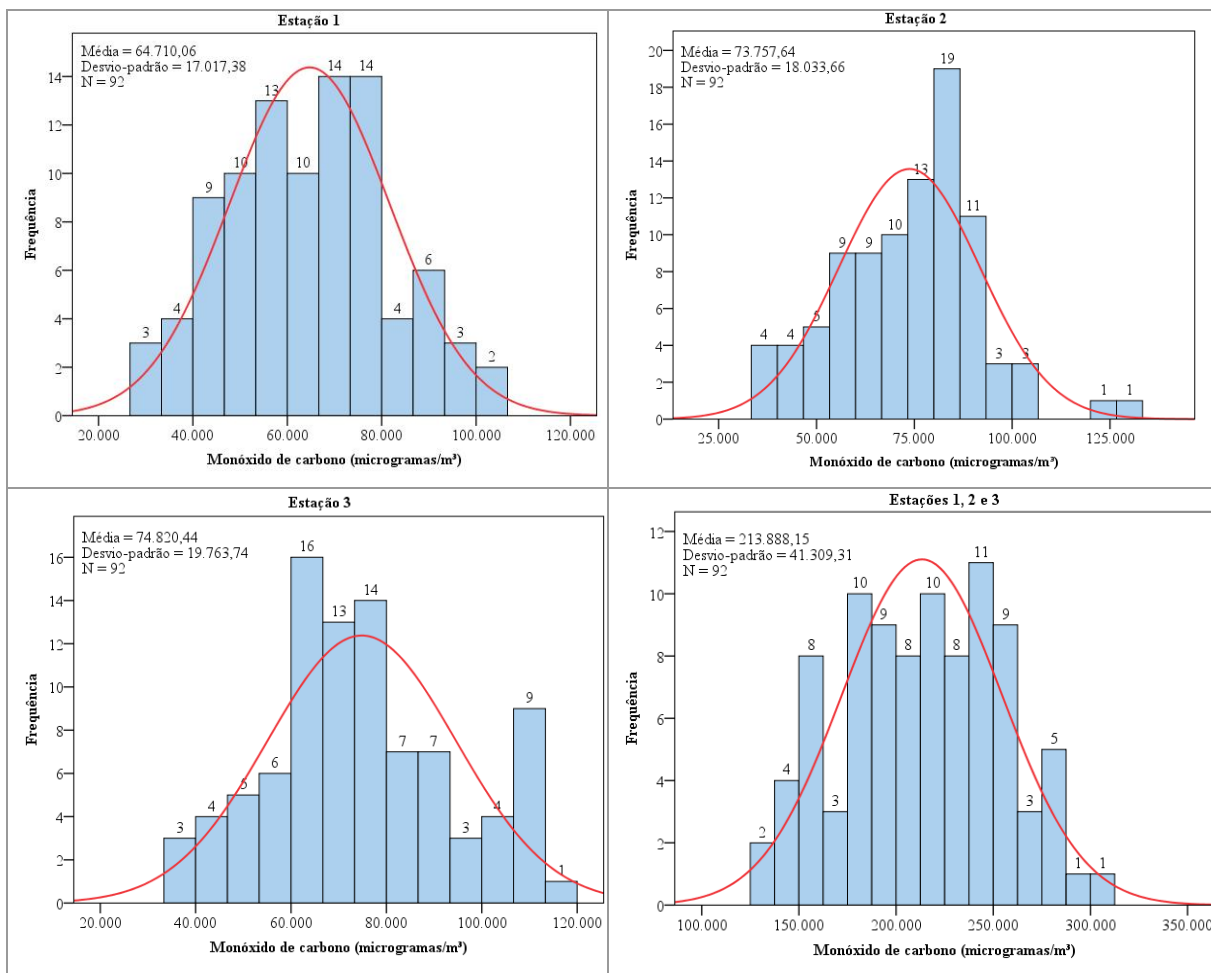
Descrição	Qui-quadrado	p-value
Estação 1	14,99	1,000
Estação 2	43,34	0,542
Estação 3	59,04	0,078
Estações 1, 2 e 3	43,19	0,549

Fonte: Elaboração própria (2018) com dados do *software SPSS*.

De acordo com a Tabela 9, verifica-se, a um nível de 5% de significância, a não rejeição da hipótese nula em todos os quatro casos. Ou seja, constata-se que os dados semanais de CO_{atm} possuem funções de distribuição iguais nas 46 semanas analisadas nos anos de 2015 e 2016.

Ademais, os histogramas referentes aos dados semanais de CO_{atm} para verificar a proximidade com essa distribuição são mostrados na Figura 10. Observa-se, que somente a estação 01 se aproxima do formato da distribuição Normal.

Figura 10. Histograma para os dados semanais de CO_{atm} das estações analisadas (2015-2016).



Fonte: Elaboração própria (2018) com dados do *software SPSS*.

Na Figura 10, mesmo que nenhum dos histogramas ali mostrados não apresenta semelhança com a distribuição Normal, valores como os erros padrão (que estimam a variabilidade das amostras), as máximas, mínimas e desvios-padrão continuam bem

próximos uns dos outros, explicando de certa forma o ajuste do modelo. A análise descritiva referente aos dados fornecidos pelas estações é mostrada na Tabela 10.

Tabela 10. Estatística descritiva para dados semanais de CO_{atm} após ajuste fino de dados.

Estação	N	Mínimo	Máximo	Soma	Média		Desvio padrão	CV
					Estatística	Erro Padrão		
EST_1	92	30.579,63	104.679,85	5.953.325,68	64.710,06	1.774,18	17.017,38	26,3
EST_2	92	36.253,80	118.821,84	6.763.906,01	73.520,72	1.814,88	17.407,68	23,7
EST_3	92	35.942,77	114.974,45	6.883.480,73	74.820,44	2.060,51	19.763,74	26,4
EST_123	92	135.170,14	304.716,46	19.622.509,54	213.288,15	4.306,79	41.309,31	19,4

Fonte: Elaboração própria (2018).

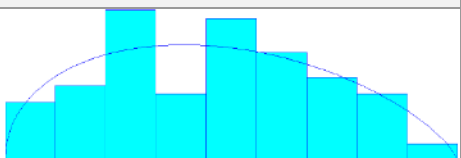
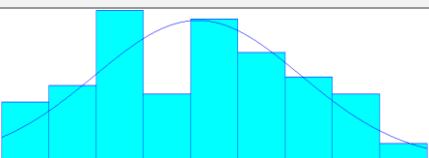
Uma das formas de verificar se os dados estão normalmente distribuídos se dá com o auxílio do teste estatístico *Kolmogorov-Smirnov (K-S)*, disponível em diversos *softwares* estatísticos. O teste *K-S* pode ser aplicado em pequenas amostras, é válido somente para distribuições contínuas e seu teste se baseia nas comparações de probabilidades acumuladas observada e teórica. O teste de hipóteses é elaborado da seguinte forma:

- H_0 : os dados amostrais seguem a distribuição teórica desejada.
- H_1 : os dados amostrais não seguem a distribuição teórica desejada.

A hipótese H_0 é rejeitada ao se comparar os testes calculado e tabelado ($K - S_{calculado} \neq K - S_{tabelado}$) ou por meio da análise do *p-value* encontrado ($p - value < \alpha$), teste este de maior poder estatístico que o mencionado anteriormente.

Assim, para as amostras tratadas, utilizou-se os seguintes parâmetros de entrada: adoção de um nível de significância $\alpha = 5\%$ para $n = 46$ amostras (ou semanas), $K - S_{critico} = 0,1923$ e desejando-se inferir que os dados amostrais coletados semanalmente nas três estações juntas são normalmente distribuídos. Os resultados da análise estatística são mostrados na Tabela 11 (mediante *Input Analyzer*[®]).

Tabela 11. Ajuste de dados à melhor distribuição, aproximação pela Normal e discriminação dos dados (estações 1, 2 e 3).

Descrição	Distribuição Beta	Distribuição Normal
Representação gráfica		
Distribuição	$135.000 + 170.000 * \text{BETA}(1,47, 1,72)$	$\text{NORM}(213.000, 41.100)$
Erro quadrático	0,009245	0,013042
K-S _{calculado}	0,053400	0,104000
<i>p-value</i> correspondente	> 0,15	> 0,15
Valor mínimo		135.170,14
Valor máximo		304.716,46
Coefficiente de variação		19,3%

Fonte: Elaboração própria com dados do *software Arena/Input Analyzer*.

Como resposta ao teste $K-S$ realizado, constata-se, a um nível de 5% de significância, que não há evidências suficientes para provar a rejeição da hipótese H_0 . Ou seja, constata-se que os dados seguem a distribuição $Beta$, e eles podem ser aproximados para a distribuição Normal uma vez que $p\text{-value} < \alpha$ (teste mais robusto) e $K-S_{calculado} < K-S_{tabelado}$ em ambos os casos.

4. Conclusões

O objetivo desse artigo foi o de encontrar a distribuição estatística que melhor se aderisse ao conjunto de dados de monóxido de carbono atmosférico (CO_{atm}) registrado em microgramas por metro cúbico ($\mu g.m^{-3}$) por quatro estações automáticas de monitoramento da qualidade do ar (medições horárias) instaladas na cidade de Ipatinga, Minas Gerais (MG), Brasil. O histórico de dados compreendeu os períodos de 2 de outubro de 2013 a 30 de outubro de 2017 (4,1 anos, aproximadamente) para três das máquinas automáticas e de 31 de janeiro de 2014 a 26 de março de 2016 (cerca de 2,2 anos) para o quarto aparato.

Como a estação de monitoramento da qualidade do ar número 04 (bairro Veneza) teve o seu início após as demais e sua operação paralisada tempos depois, optou-se por removê-la do histórico de dados amostrais. Para as demais estações, iniciou-se a etapa de tratamento dos registros horários de captação de CO_{atm} ; esta fase compreendeu na substituição de dados não registrados pelos equipamentos por valores médios e, em seguida, na remoção de pontos discrepantes (*outliers*).

Essa decisão se mostrou assertiva pois as discrepâncias reduziram bastante com o uso de dados semanais. Elaborou-se um diagrama em caixa para verificar o comportamento dos dados antes e depois do tratamento inicial dos dados nas estações (valores semanais por estação e valores semanais acumulados) e constatou-se que o comportamento gráfico em todos os casos pareceu estar centralizado.

Para o estudo de análise de variância (ANOVA), removeu-se os dados referentes aos anos de 2013 e 2017 por apresentarem, respectivamente, 11 e 38 semanas. Realizou-se alguns cenários comparativos envolvendo as três estações separadamente, em semanas por ano, e depois com os valores totais emitidos pelos três equipamentos, também em semanas por ano. O período que obteve maior aderência em relação à ANOVA foi o de janeiro de 2015 a dezembro de 2016, e com os dados semanais das três estações juntas. Neste cenário verificou-se que os dados de captação CO_{atm} eram estatisticamente iguais ($p\text{-value} = 0,579$ e a formação de um único subgrupo de acordo com o teste de *Tukey*).

Porém, o teste de Levene para analisar a homogeneidade de variâncias apresentou significância estatística ($p\text{-value} = 0,00$). Em outras palavras, o teste de igualdade de variâncias foi rejeitado para o período semanal entre 2015-16 com os dados das três estações automáticas de monitoramento da qualidade do ar.

Em contrapartida, realizou-se o teste de *Kruskal-Wallis* para testar as hipóteses de que as populações possuem funções de distribuição iguais (H_0) contra a de que pelo menos duas populações possuem diferentes funções de distribuição (H_1). Particularmente para as estações 1, 2 e 3 juntas, o $p\text{-value}$ encontrado foi de 0,549. Constatou-se, desta forma, que os dados semanais de CO_{atm} possuem funções de distribuição iguais nas semanas analisadas nos anos de 2015 e 2016.

Ao pesquisar qual distribuição estatística que melhor poderia explicar os dados semanais referentes ao biênio de 2015-16, descobriu-se que a distribuição $Beta$ apresentou o melhor ajuste para o modelo. O teste estatístico de *Kolmogorov-Smirnov* ($K-S$) para verificar se os dados amostrais seguiam a distribuição teórica desejada encontrou um $p\text{-value} > 0,15$ e a distribuição [135.000 + 170.000 * $BETA(1,47, 1,72)$] foi então validada para fornecer valores aleatórios de dados semanais de CO_{atm} envolvendo a captação das três estações automáticas de monitoramento da qualidade do ar.

Ao se fazer a aproximação pela distribuição Normal (NORM), o teste $K-S$ também evidenciou um $p\text{-value} > 0,15$. Tal registro também proporcionou afirmar que os dados seguiam a distribuição $NORM(213.000, 41.100)$ para as três estações. Finalmente, o resultado possibilita o uso de dados aleatórios envolvendo essas distribuições em modelos matemáticos como de Lotka-Volterra.

Para trabalhos futuros, recomenda-se a continuação das análises, pois o monitoramento da qualidade do ar se apresentam como valiosa justificativa para as buscas por tecnologias que amenizem as consequências da poluição atmosférica.

Referências

- Abrampa, Associação Brasileira do Ministério Público do Meio Ambiente. (2010). Cinco meses de monitoramento da qualidade do ar em Ipatinga/MG.–. 28 nov. 2010. <<https://abrampa.jusbrasil.com.br/noticias/2478523/cinco-meses-de-monitoramento-da-qualidade-do-ar-em-ipatinga-mg>>.
- Afonso, C. (2012). *Termodinâmica para Engenharia*. Revisão técnica Felipe Direito. (1ª. ed.) Faculdade de Engenharia da Universidade do Porto. [Porto]: FEUP Edições.
- Alisson, E., Iziq, C., & Lopes, N. (2013). *Mudanças no clima do Brasil até 2100*. Fundação de Amparo à Pesquisa do Estado de São Paulo – FAPESP. 10 set. 2013. <<http://agencia.fapesp.br/mudancas-no-clima-do-brasil-ate-2100/17840/>>.
- Amabis, J. M., & Martho, G. R. (2013). *Biologia em Contexto*. (1ª. ed.): Moderna, 2013.
- Araújo, S. A. do C., & Deminicis, B. B. (2009). Fotoinibição da fotossíntese. *Revista Brasileira de Biociências*. Porto Alegre, 7(4), 463-472, out./dez. 2009. <http://www.ufrgs.br/seerbio/ojs/index.php/rbb/article/view/1009>
- Atkins, P., & Jones, La. (2012). *Princípios de Química: questionando a vida moderna e o meio ambiente*. Tradução técnica Ricardo Bicca Alencastro. (5ª. ed.): Bookman, 2012.
- Barnola, J. M., et al. (1987). Vostok ice core provides 160,000-year record of atmospheric CO₂. *Nature*. v. 329, october 1987, p. 408. <<http://www.nature.com/nature/journal/v329/n6138/pdf/329408a0.pdf>>.
- Barrère, M. (1992). Terra, Patrimônio Comum: a ciência a serviço do meio ambiente e do desenvolvimento. Tradução Estela dos Santos Abreu. São Paulo: Nobel.
- Bassanezi, R. C. (2016). *Equações Diferenciais Ordinárias: um curso introdutório*. Centro de Matemática, Computação e Cognição, Universidade Federal do ABC – UFABC, Santo André, 19 fev. 2016. Apostila. <<http://gradmat.ufabc.edu.br/disciplinas/listas/iedo/notasdeaulas/equacoes-diferenciais-ordinarias-rodney.pdf>>.
- Björkman, O. (1981). *Responses to Different Quantum Flux Densities*. In: Lange O. L., Nobel, P. S., Osmond, C. B., & Ziegler H. (eds) *Physiological Plant Ecology I*, v. 12. A. Springer, Berlin, Heidelberg, 1981. p. 57-107. <https://link.springer.com/chapter/10.1007%2F978-3-642-68090-8_4>.
- Boyce, W. E., & Diprima, R. C. (2015). *Equações Diferenciais Elementares e Problemas de Valores de Contorno*. Tradução e revisão técnica Valéria de Magalhães Iorio. (10ª.ed.): LTC.
- Bumbeers, F. (2022). Como será o clima da Terra em 2100? *Revista Galileu*. Disponível em: <<https://revistagalileu.globo.com/Ciencia/noticia/2015/06/como-sera-o-clima-da-terra-em-2100.html>>.
- BRASIL, Ministério do Meio Ambiente (2022). Efeito Estufa e Aquecimento Global. <http://www.mma.gov.br/informma/item/195-efeito-estufa-e-aquecimento-global>.
- Collares, C. F. (2011). *Testes post-hoc para análise de variância (ANOVA): um tutorial básico*. Blog do C.F.C. 03 fev. <<http://carloscollares.blogspot.com/2011/02/testes-post-hoc-para-analise-de.html>>.
- Conti, F. (2009). *Teste de hipóteses e significância*. Instituto de Ciências Biológicas, Universidade Federal do Pará – UFPa, Belém, 3 out. 2009. <<http://www.cultura.ufpa.br/dicas/biome/biopdf/biotestes.pdf>>.
- Faria, B. (2017). *Comparações múltiplas – teste de Tukey*. ABG Consultoria. 25 set. <http://www.abgconsultoria.com.br/blog/como-interpretar-uma-analise-de-variancia-anova/>.
- Faria, B. (2017). Como interpretar uma análise de variância (ANOVA)? ABG Consultoria. 10 jul. 2017. <<http://www.abgconsultoria.com.br/blog/como-interpretar-uma-analise-de-variancia-anova/>>
- Field, A. (2009). *Descobindo a Estatística com o SPSS*. Tradução Lorí Viali. (2ª. ed.): Artmed,
- Fonseca, J. J. S. (2002). *Metodologia da Pesquisa Científica*. Fortaleza: Universidade Estadual do Ceará, 2002. Apostila.
- Freund, J. E. (2006). *Estatística Aplicada: economia, administração e contabilidade*. Tradução Claus Ivo Doering. (11ª. ed.): Bookman.
- Freitas, C. M. *Um Equilíbrio Delicado: crise ambiental e a saúde no planeta*: Garamond, 2011.
- Gil, A. C. (2010). *Como Elaborar Projetos de Pesquisa*. (5ª. ed.): Atlas.
- Goklany, I. M. (2000). Potential consequences of increasing atmospheric CO₂ concentration compared to other environmental problems. *Technology*, v. 75, p. 189-213, abr. 2000. <<https://pdfs.semanticscholar.org/f5bfigurab/eb844e5eae378047aa825298fc98a4ad4d5a.pdf>>.
- Jean-Claude, K., & Pierre-André, H. *The Consequences of an Increase of the Atmospheric CO₂ Concentration*, a global survey of our current state of knowledge. In: WORLD ENGINEER'S CONVENTION. Geneva, 2011.
- Joos, F. (2009). *Climate and carbon cycle representation in Integrated Assessment Models*. 23 dez. <<http://hassler-j.ies.su.se/courses/climate/joos09.pdf>>.

- Kirschbaum, M. U. F. (1994). *The Sensitivity of C3 photosynthesis to increasing CO₂ concentration: a theoretical analysis of its dependence on temperature and background CO₂ concentration*. In: PLANT, CELL AND ENVIRONMENT, v. 17, 747-754.
- Kozakevich, D. N. (2003). Seção 8: equações diferenciais não lineares. Departamento de Matemática, Centro de Física-Matemática, Universidade Federal de Santa Catarina - UFSC, Florianópolis, 10 jun. Apostila. <http://mtm.ufsc.br/~daniel/sem1_05/edo/farlow/sec8.pdf>.
- Mann, P. S. (2012). *Introdução à Estatística*. Tradução Eduardo Benedito Curtolo; Teresa Cristina Padilha de Souza. [Reimpr.]. Rio de Janeiro: LTC.
- Marcel, G. (2014). Plantas C3, C4 e CAM: fixação de carbono. 15 fev. <<https://www.euquerobiologia.com.br/2014/02/plantas-c3-c4-e-cam-fixacao-de-carbono.html>>.
- Martinez, C. A., et al. (2015). Respostas das plantas ao incremento atmosférico de dióxido de carbono e da temperatura. *Revista de Geografia Física*. [Recife], v. 8, p. 635-650, 10 nov. <<https://www.researchgate.net/publication/287804710>>.
- Montgomery, D. C., & Runger, G. C. (2009). *Estatística Aplicada e Probabilidade para Engenheiros*. Tradução e revisão técnica Verônica Calado. (4ª. ed.): LTC.
- Morettin, P. A., & Bussab, W. O. (2013). *Estatística Básica*. (8ª. ed.): Saraiva.
- Morrison, J. I. L., & Lawlor, D. W. (2002). Interactions between increasing CO₂ concentration and temperature on plant growth. *Plant, Cell & Environment*, v. 22, p. 659-682.
- Neftel, A., et al. (1985). Evidence from polar ice cores for the increase in atmospheric CO₂ in the past two centuries. *Nature International Journal of Science*. v. 315, 45-47, 02 maio.
- Praça, F. S. G. (2015). Metodologia da pesquisa científica: organização estrutural e os desafios para redigir o trabalho de conclusão. *Revista Eletrônica Diálogos Acadêmicos*, 8(1), 72-87.
- Souza, A. P. R. (2022). Avaliação de impacto ambiental do processo de obtenção de compósito do tipo nanocarbonato de cálcio/lignina kraft a partir da captura de CO₂. Dissertação de Mestrado. <https://www.alice.cnptia.embrapa.br/handle/doc/1145264>.
- Triviños, A. N. S. (1987). *Introdução à pesquisa em ciências sociais: a pesquisa qualitativa em educação*: Atlas.