



MINISTÉRIO DA CIÊNCIA, TECNOLOGIA, INOVAÇÕES E COMUNICAÇÕES
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

**AVALIAÇÃO DOS DADOS ESPACIAIS DE MP₁₀ REGIONAIS E SUA
COMPARAÇÃO COM DADOS OBSERVADOS PARA A REGIÃO DO
GRANDE PORTO – PORTUGAL**

Bárbara Pavani Biju

Trabalho final apresentado a disciplina de
Introdução ao Geoprocessamento em
formato de artigo.

Professor: Dr. Antônio Miguel Vieira
Monteiro

AValiação DOS DADOS ESPACIAIS DE MP₁₀ REGIONAIS E SUA COMPARAÇÃO COM DADOS OBSERVADOS PARA A REGIÃO DO GRANDE PORTO - PORTUGAL

Barbara Pavani Biju¹

¹ Faculdade de Ciências da Universidade do Porto

Rua do Campo Alegre s/n, 4169-007 Porto

RESUMO

A poluição atmosférica é a causa da morte precoce de milhares de pessoas na Europa e manutenção da qualidade do ar e o monitoramento da concentração dos poluentes se faz necessária. Principalmente os materiais particulados (MP), que tem efeito deletério na saúde humana. Assim, diferentes técnicas e metodologias são empregadas para avaliar a concentração dos poluentes atmosféricos principalmente em áreas urbanas, desde estações de monitoramento, sensoriamento e técnicas de geoprocessamento. As redes de medição são amplamente empregadas, mas devido as suas limitações é necessário inferir os valores medidos através de modelos e métodos geoestatísticos, como a krigagem. Contudo, o sensoriamento remoto tem demonstrado um papel importante no monitoramento do MP e ajudado no desenvolvimento de produtos sobre análise de qualidade do ar para a Europa. Por fim, este trabalho tem o objetivo de verificar a qualidade dos dados de análise de qualidade do ar, correlacionando com os valores medidos in-situ e interpolados a partir da kringagem. Apesar das limitações, comparação dos valores estimados da kringagem com os de análise de qualidade do ar permitiu constatar que os produtos gerados pelo serviço de monitoramento atmosférico Copérnico (CAMs), possam ser de grande valia no controle da qualidade do ar em locais onde as estações de monitoramento de qualidade do ar não estão disponíveis.

Palavras-chave: material particulado; poluição do ar, kringagem, qualidade do ar.

Abstract

Air pollution is the cause of early death for thousands of people in Europe, so the maintenance of air quality and monitoring pollutants concentration is necessary. Primarily Particulate Matter (PM), which has a deleterious effect on human health. Thus, different techniques and methodologies are used to assess the concentration of atmospheric pollutants, mainly in urban areas, such as monitoring stations, remote sensing, and geoprocessing techniques. air quality monitoring networks are widely used, but due to their limitations it is necessary to infer measured values through geostatistical models and methods, such as kriging. However, remote sensing has shown an important role in PM monitoring and helped the development of air quality analysis products for Europe. Finally, this work aims to verify the quality of air quality analysis data, correlating with the values measured in-situ and interpolated from the kriging. Despite the limitations, comparing the estimated values of the kriging with those of the air quality analysis allowed us to verify that the products generated by the Copernicus atmospheric monitoring service (CAMs) can be of great value in the control of air quality in places where the stations air quality monitoring devices are not available.

Keywords: particulate matter; air pollution, kriging, air quality, geoprocessing

1. INTRODUÇÃO

A poluição atmosférica afeta a saúde humana e o meio ambiente, principalmente nas áreas urbanas. Em 2018, aproximadamente 0.4 milhões de pessoas morreram prematuramente por toda Europa, devida a longa exposição aos poluentes atmosféricos, como ozônio (O₃), dióxido de nitrogênio (NO₂), dióxido de enxofre (SO₂), e material particulado (MP) (; WORLD HEALTH ORGANIZATION, 2005; GABRIEL et al., 2020), o último de grande preocupação devido aos efeitos deletérios à saúde humana. Todos os poluentes mencionados impactam diretamente na qualidade do ar, e é de costume avaliar sua concentração, principalmente das partículas com diâmetro menor que 1 µm (PM₁), 2.5 µm (PM_{2.5}) ou 10 µm (PM₁₀), por causa do impacto na saúde (WHO REGIONAL OFFICE FOR EUROPE, 2016)

Hoje em dia, diferentes técnicas são empregadas para avaliar a concentração dos poluentes em áreas urbanas, incluindo redes de medição de monitoramento, técnicas de geoprocessamento, sensoriamento remoto. Todas estas ferramentas podem ser associadas umas com as outras e com modelos matemáticos. Imagens de diferentes satélites tem sido empregue ao redor do mundo para monitorar a qualidade do ar, especialmente quando estações de monitoramento não estão disponíveis.(ALVAREZ-MENDOZA; TEODORO; RAMIREZ-CANDO, 2019; OLIVEIRA et al., 2020)

Tradicionalmente, o estudo dos MPs acontece através da rede de estações, mas os custos associados com a sua instalação e manutenção, é alto. Apesar das estações de monitoramento serem amplamente utilizadas para avaliar a qualidade do ar de uma região pois fornecem as informações de concentração de forma mais confiável e há certos problemas relacionados com a cobertura de dados (YANG et al., 2017), por exemplo, a distância da area estudada e a distância das estações, fazendo-se necessária a extrapolação.

A extrapolação pode ser feita a partir de interpoladores, que são divididos em duas categorias, determinísticos e estocásticos (MILITINO; UGARTE; PÉREZ-GOYA, 2018). O primeiro gera superfícies de acordo com a semelhança espaciais da amostra, não levando em consideração a variabilidade espacial. Já os estocásticos utilizam a teoria da probabilidade, incorporando critérios estatísticos na determinação do peso atribuído aos

pontos amostrais no cálculo das interpolações (BURROUGH E MCDONEEL, 1998; MAZZINI E SCHETTINI, 2009)

O sensoriamento remoto tem desempenhado um papel importante no monitoramento do MP, demonstrando a capacidade de ultrapassar os limites devido as informações descontínuas providas pelas estações de monitoramento, tanto no espaço quanto no tempo (ROUX et al., 2020). Contudo, as estações disponíveis ainda são utilizadas como referência para validar os dados obtidos através do sensoriamento remoto, como é o caso dos produtos do serviço de monitoramento atmosférico Copernicus, (CAMS). Os produtos CAMS provém análises diárias e previsão da qualidade do ar global, como também para a qualidade do ar para a Europa.

Os dados atmosféricos fornecidos pela CAMs são baseados no conjunto *ensemble*, elaborado a partir de nove modelos matemáticos de previsão e qualidade do ar. Um conjunto de mediana é calculada a partir dos resultados individuais destes nove modelos, pois os produtos *ensemble*, tem um melhor desempenho que os produtos dos modelos individuais. Além do uso dos modelos, a observação *in-situ* é usada para avaliar e garantir a qualidade dos dados para MP₁₀ (COPERNICUS ATMOSPHERE MONITORING SERVICES, 2021).

Estes dados provêm das redes de monitoramento de qualidade do ar dos países pertencentes à União Europeia, como é o caso de Portugal. Devido as diretivas europeias sobre a qualidade do ar, os países têm que emitir anualmente o Inventário Anual de Emissões Atmosféricas, como também disponibilizar os dados de concentração dos principais poluentes atmosféricos para o monitoramento da poluição do ar (EUROPEAN ENVIRONMENTAL AGENCY, 2020)

Todavia, nem todos os dados coletados pelas estações de monitoramento e disponibilizados em Portugal, são utilizados para avaliar a qualidade dos produtos CAMs para PM₁₀. Somente dados das estações situadas nas grandes cidades são utilizados na comparação, como no caso da cidade do Porto. Assim, este trabalho pretende analisar os dados de análise de qualidade do ar do CAMs e comparar com os dados observados adquiridos da rede de monitoramento de qualidade do ar local para o grande Porto.

2. Materiais e métodos

2.1. Área de estudo

A região selecionada para este trabalho, foi a região metropolitana da cidade do Porto, situada no norte de Portugal. O Porto é uma cidade costeira, de grande importância para a economia de Portugal e é considerada património cultural da humanidade. Além da importância apresentada, a região foi escolhida devido ao número de estações da rede de medição contidas na região e que monitoram a qualidade do ar, densidade urbana, em especial o MP₁₀ (Figura 01). Os dados espaciais, como por exemplo, limite administrativo, para a região de estudo foram obtidos no site Registro Nacional de Dados Geográficos (SNIG) de Portugal.

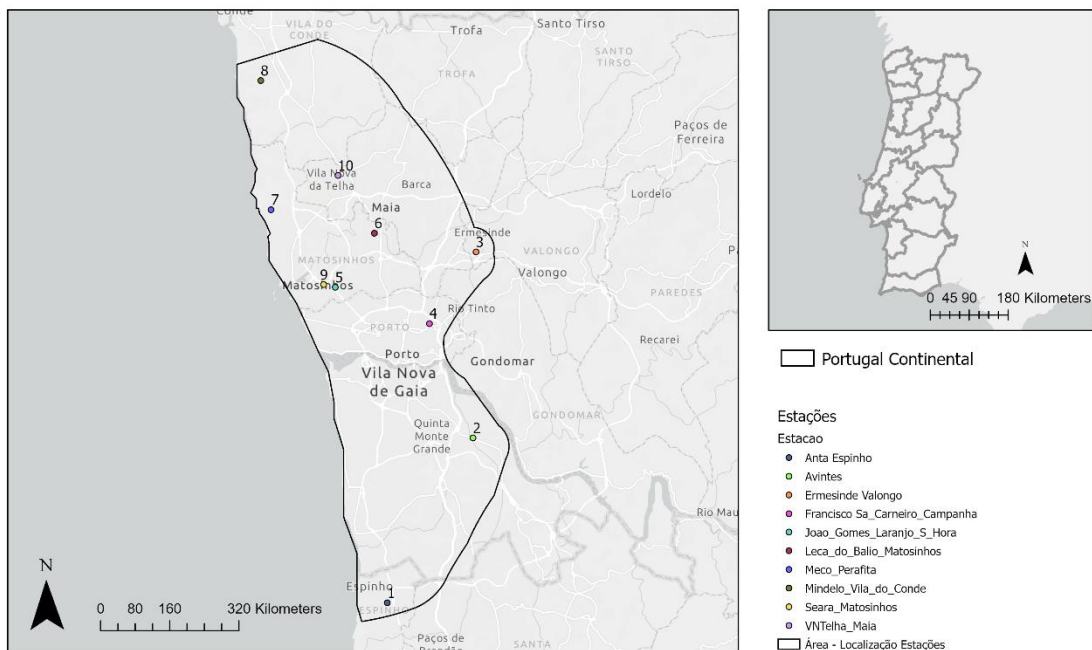


Figura 01 – Localização Area de Estudo e Estações

Elaboração: Autora, 2021

Em Portugal, o órgão responsável pela gestão e avaliação da qualidade do ar e a rede de medição é a Agência Ambiental Portuguesa (APA). Através do Decreto-lei nº 102/2010 que transpõe a diretiva Diretiva 2008/50/CE do parlamento europeu - relativa aos objetivos fixados para qualidade do ar ambiente, levando em consideração as normas e orientações

da Organização Mundial da Saúde. Fixa os valores limite (VL) para monitorização dos poluentes, isto é, um nível é fixado com o intuito de reduzir os efeitos nocivos na saúde e no ambiente. Há dois VL nacionais, a média anual e a média horária. Para PM_{10} , o VL anual é de $40\mu g/m^3$ e a média diária $50\mu g/m^3$, não podendo exceder mais de 35 vezes por ano.

A rede de medição do Norte é constituída por 22 estações. Dentre elas, 10 estações que constam na figura acima, estão situadas na zona denominada Porto Litoral. A medição da concentração do PM_{10} nestas estações é feita de hora em hora, posteriormente os dados são tratados estatisticamente e disponibilizados na plataforma de informações de qualidade do ar – QualAr, onde pode-se obter a concentração média horária e a concentração média diária anual. Os dados estão disponíveis desde o ano de 2001 até 2019. Para 2020, ainda não há dados das médias disponíveis na plataforma, portanto os anos selecionados para este trabalho foram os anos de 2018 e 2019, pois a qualidade dos dados fora verificada pela equipe responsável da agência ambiental.

2.2. Dados espaciais – CAMS

O CAMS produz análise da qualidade do ar e previsão diária para a União Europeia em uma resolução espacial significativamente maior ($0.1^\circ \times 0.1^\circ - 10\text{ km} \times 10\text{ km}$). Os dados para MP_{10} foram obtidos através da plataforma “*Atmosphere Data Store*”, para os anos de 2018 e 2019, utilizando o modelo “*essemble*”, para superfície da area de estudo.

2.3. Análise Descritiva

Os dados obtidos do PM_{10} para 10 estações na plataforma QualAr, são disponibilizados em ficheiros no formato *comma-separated values* (csv) e estão desconsolidados. Portanto, é necessário o seu tratamento prévio para que se possa descrever os aspectos e características importantes deste conjunto de dados. Todo o tratamento estatístico foi feito no software R, devido ao grande volume de dados diários. Para cada estação obteve-se um arquivo com 8760 observações. Para cada uma das 10 estações foram calculadas a média, mediana, variância, desvio padrão, coeficiente de variação, coeficiente de assimetria, coeficiente de curtose, quartil inferior e superior para os anos de 2018 e 2019. Também foi verificada a normalidade dos dados através de histogramas e do teste de Shapiro-Wilk. Esta etapa também se faz necessária para o uso posterior dos dados para interpolação

como também no software de Sistema de Informação Geográfica escolhido para este trabalho.

2.4. Krigagem

Com os resultados obtidos da análise estatística e exploratória dos dados, foram elaboradas duas tabelas com as coordenadas geográficas e os atributos para cada uma das 10 estações deste trabalho, para os anos de 2018 e 2019. Com estes dados foi possível efetuar a técnica de interpolação mais adequada, a krigagem. A técnica é constituída por várias etapas, sendo elas: estatística descritiva (feita previamente), análise do semivariograma, validação e por fim, a krigagem.

Primeiramente foram feitos os cálculos, ajustes e seleção do semivariograma, para obter-se a variabilidade espacial da concentração de PM_{10} para a área de estudo, onde foi definido o modelo do semivariograma. Esta etapa, como as posteriores foram efetuadas no software de SIG, utilizando as coordenadas inseridas tabela mencionada acima. Foram testados diferentes modelos para verificar o que melhor se adequava a distribuição amostral para a área de estudo.

Em seguida foi efetuada a validação cruzada para verificar qual modelo provém a melhor predição, também indica se o modelo e se os parâmetros associados a ele são razoáveis, esta validação foi feita para todos os modelos do semivariogramas testados. Primeiramente remove-se um ou mais dados locais e prevê-se o seu valor, utilizando os restantes dos dados dos outros locais. Assim, pode-se olhar os valores e comparar os valores estimados com o observado para aquele local, obtendo informações uteis sobre a qualidade do modelo de krigagem (ESRI, 2015)

Com o modelo definido, com a validação cruzada efetuada em sequência foi feita a estimativa da krigagem. Foram testados diferentes métodos de krigagem, obtendo-se um mapa de valores estimados e outro de variância, no caso o mapa de incerteza. Esta variância (incerteza) da interpolação, considera além da localização dos pontos de dados, mas também sua dispersão local (YAMAMOTO, 2000), ou seja, identifica locais ou regiões onde a amostragem pode ser melhorada. Tendo como dado de saída um mapa de estimativa e variância tanto para o ano de 2018, quanto para 2019.

2.5. Análise comparativa

A análise comparativa entre os produtos do CAMS e a interpolação para ambos os anos, foram efetuadas no software de SIG. Pode-se efetuar a estatística básica para os raster e foram gerados os dados relativos concentração de PM₁₀ para a área de estudo. A partir destes valores foram feitas as análises estatísticas para verificar a correlação entre os valores obtidos da interpolação, da variância com os produtos CAMs. Para isso, foi feita a correlação através da estatística de coleta de banda. Para efetuar este teste é necessário atender certos pressupostos como, normalidade dos dados, relação linear e ausência de outliers. A normalidade foi verificada através do histograma gerado no próprio software de SIG, foi verificada a presença de outliers e a relação linear entre as variáveis.

3. RESULTADOS E DISCUSSÕES

Durante a fase de análise exploratória dos dados, fora verificado que há uma descontinuidade da medição em certas estações, como constatado em (ROUX et al., 2020). Isto pode ter ocorrido por diferentes fatores externos, como por exemplo, manutenção da estação. Mas na maioria dos casos, a medição foi feita de forma contínua, permitindo que as análises necessárias fossem efetuadas. Na tabela (tabela 01) abaixo são apresentados os resultados da estatística básica para os dados das 10 estações tanto para 2018, quanto para 2019.

Parâmetros	2018	2019
Quantidade Estações	10	10
Média	18.05	18.77
Variância	11	3.81
Desvio Padrão	3.33	1.95
Coefficiente de Variação (%)	10	18
Coefficiente de assimetria	-0.56	0.2
Coefficiente de Curtose	-0.807	- 1.63
Mediana	18.59	18.66
Mínimo	11.84	16.11
Máximo	22.81	21.67
Quartil inferior	17.68	17.20
Quartil superior	19.22	20.41
Valor p*	0.9816	0.6742

Tabela 01 – Estatística descritiva
Elaboração: Autora, 2021

Para o ano de 2018, a distribuição da concentração de PM_{10} é levemente negativa e assimétrica, o grau de curtose indica uma distribuição com cauda mais leve do que a normal. A média e a mediana possuem valores próximos, o que indica que a distribuição da variável é aproximadamente assimétrica (figura 03). Já para 2019, a distribuição é positiva e assimétrica, o grau de curtose indica que a distribuição também com cauda mais leve que a normal, a media e mediana possuem valores muito próximos, indicando também que a distribuição da variável é aproximadamente assimétrica (figura 04). Estas análises são de extrema importância, pois pode haver erros que devem ser corrigidos ou eliminados. Também foi feito o teste Shapiro-Wilk de normalidade para ambos os anos, onde constatou-se que a distribuição dos dados tem distribuição normal ($p > 0.05$).

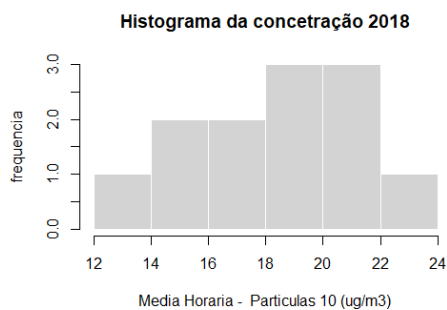


Figura 03 – Histograma dados 2018
Elaboração: Autora, 2021

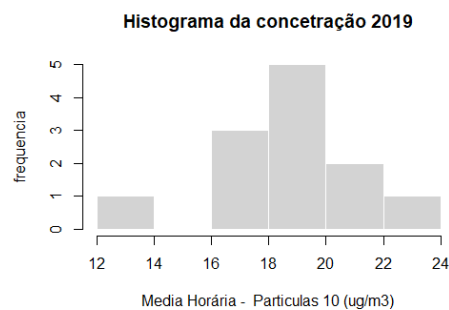


Figura 04 – Histograma dados 2019
Elaboração: Autora, 2021

Foram testados diferentes modelos de semivariogramas para as variáveis, a análise estrutural dos dados é necessária para verificar a autocorrelação e da dissimilaridade entre os pares observados através do semivariograma (CAMARGO; DRUCK; CÂMARA, 2004), para os anos de 2018 e 2019. Foram feitos os ajustes necessários para modelagem do semivariograma a fim de entender o comportamento estrutural das variáveis que são objetos de estudo deste trabalho. Para os dados de 2018 sobre o MP_{10} obteve-se o seguinte semivariograma omnidirecional e seu modelo de ajuste. O modelo mostrado na figura 05 tem os seguintes parâmetros: Estrutura tipo Circular; Efeito Pepita (C_0) = 1,2 e Alcance = 15125,51.

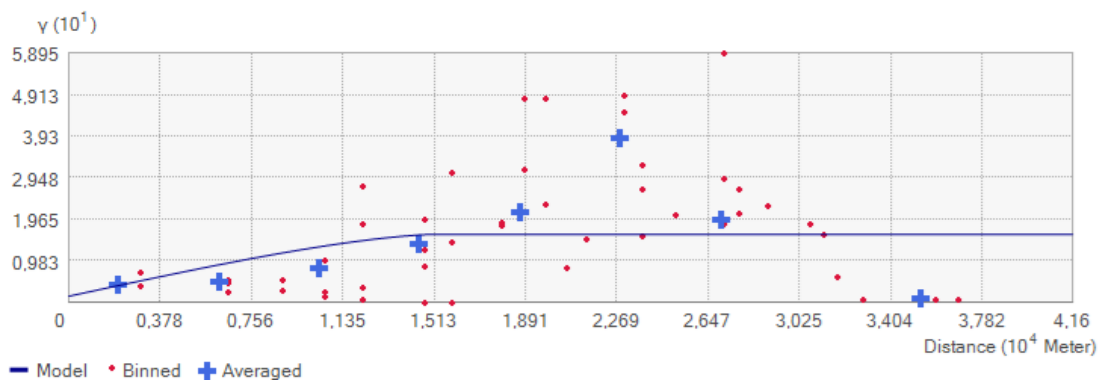


Figura 05: Semivariograma – MP10 ano 2018
 Elaboração: Autora, 2021

Para o ano de 2019, o semivariograma omnidirecional mais adequado e seu modelo de ajuste exibidos na figura 06 possui os seguintes parâmetros: Estrutura tipo Gaussiana; Efeito Pepita (C_0) = 0,252; e Alcance = 11500,59.

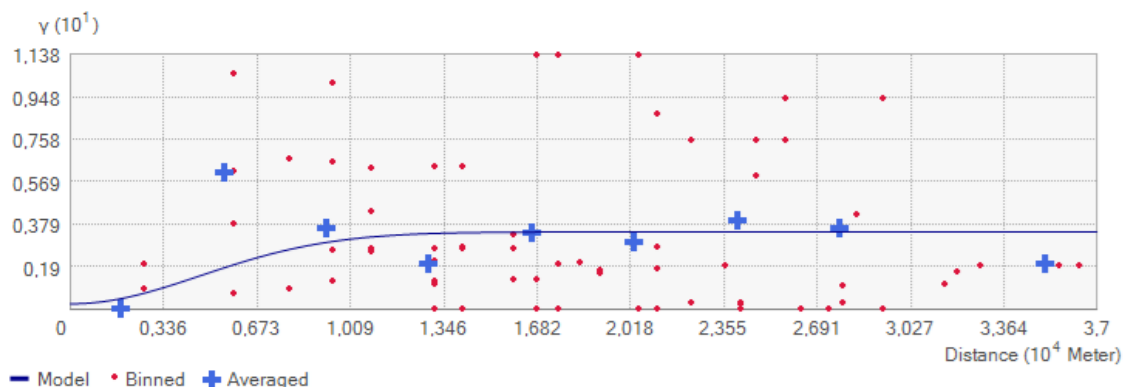


Figura 06: Semivariograma – MP10 ano de 2019
 Elaboração: Autora, 2021

O modelo escolhido para representar a estrutura de correlação espacial para ser utilizada na inferência de krigagem, foi o modelo com o menor erro padrão da estimativa, no caso o método geoestatístico definido foi Krigagem ordinária. O uso deste método se deve a sua simplicidade, pois ele utiliza o modelo fornecido pelo variograma e por fornecer a incerteza associada a estimativa, por meio da variância (YAMAMOTO, 2010).

Ao mesmo tempo, foi efetuada a validação cruzada a partir do método “*leave one out*”, ou verificação cruzada, para verificar a adaptação do modelo aos dados utilizados na interpolação. O método consiste em remover um único ponto do conjunto de dados e utiliza os pontos restante para prever para o valor para o local onde o ponto foi removido, em seguida este valor depois é comparado com valor medido e as análises estatísticas necessárias são geradas para determinar a acurácia da predição (ESRI, 2015). Os resultados obtidos estão descritos abaixo (tabela 02).

Parâmetros	2018	2019
Número de amostras utilizadas	10	10
Média	0,98	0,0876
Raiz Quadrada da Média	3,49	1,71
Média Padronizada (erro)	0,16	0,007
Raiz Quadrada da Média Padronizada	1,03	1,07
Erro Padrão Médio	3,39	1,66

Tabela 02 – Dados obtidos da interpolação
Elaboração: Autora, 2021

Esta escolha se deu, pois, o método foi que melhor se ajustou e que resultou em um menor erro de predição. No caso, a raiz quadrada da média padronizada deve estar próxima a zero, assim a predição dos erros padronizada é válida. Se for maior que um, a variabilidade das predições está subestimada, menor que um, a variabilidade das predições está superestimada. Também deve levar em consideração outros dados, como por exemplo, o valor do erro padrão médio. Se este valor está próximo ao valor da raiz quadrada média, a avaliação da variabilidade da predição está feita de forma correta, se o erro padrão médio é maior, há uma superestimação dos dados na predição. Mas se for menor que a raiz quadrada média, há uma subestimação da estimativa da variabilidade (ESRI, 2015).

Neste caso, a predição dos erros padronizados para ambos os anos está subestimada - acima de 1, mas os valores são muito próximos a zero. Também foi constatado para ambos os anos uma subestimação da variabilidade da predição, pois está abaixo do valor da raiz quadrada média. Outras medidas que indicam a escolha deste método é a média padronizada, qual deve estar próxima a zero e foi constatada para ambos os anos. Nos outros métodos, estes valores ficaram distantes de zero, demonstrando que a predição possivelmente estava enviesada.

A partir destes dados, foram gerados os produtos da estimativa e da variância da krigagem (incerteza) para os respectivos anos de 2018 e 2019 e comparados com os produtos CAMS para os respectivos anos. Primeiramente, foi feita a reamostragem e alteração do tamanho das células (50mx 50m, valor correspondente ao do dado interpolado) dos dados matriciais CAMS para posterior comparação com os dados matriciais provenientes da interpolação e validação cruzada. Na imagem abaixo (figura 07), podemos ver os produtos matriciais gerados para o ano de 2018. A primeira imagem corresponde a estimativa dos valores de PM₁₀, a segunda corresponde ao produto CAMS e a terceira é variância da estimativa para o ano de 2018.

Material Particulado MP10 - 2018

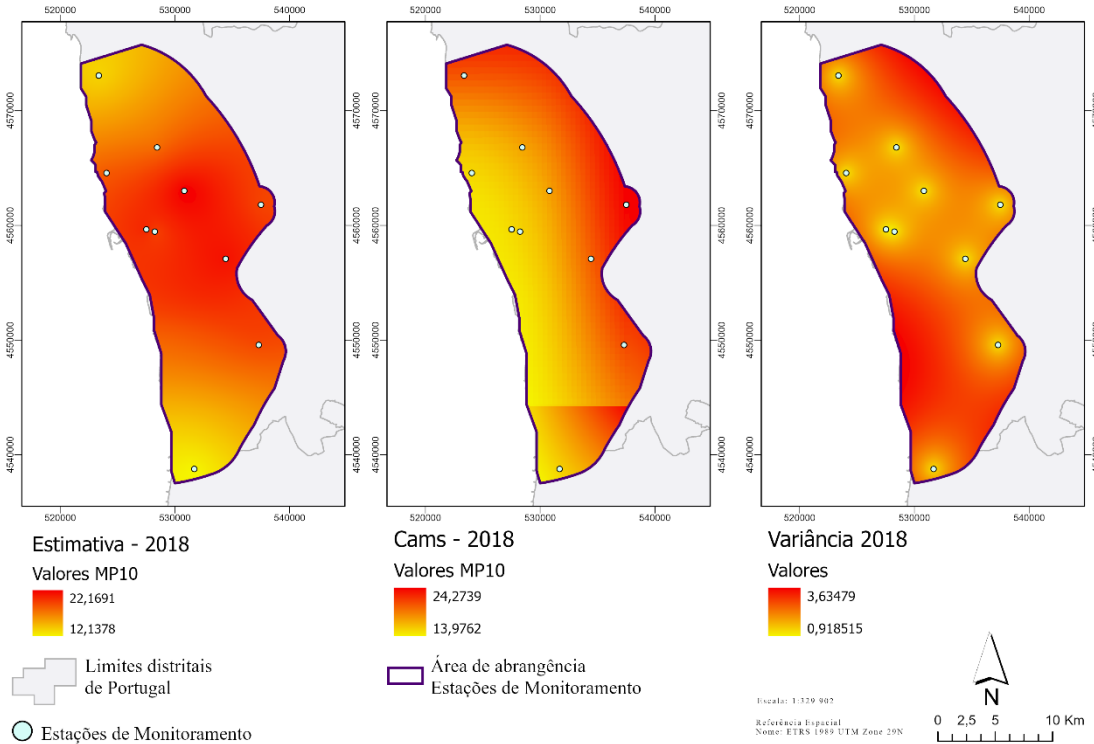


Figura 07: Mapa MP10 – 2018
Elaboração: Autora, 2021

Podemos perceber que o maior grau de variância, a incerteza da estimativa ocorre onde não há estações próximas. Este é um comportamento esperado, já que reflete a baixa variância sobre os pontos e alta entre os pontos de dados estimados. Mas ainda sim, o valor desta incerteza é baixo, comparado aos valores estimados. A estimativa para 2018 apresentou valores mínimos maiores que os do produto CAMs e o valor máximo, menor. Estes valores encontram nos valores máximos e mínimos medidos diretamente através das estações da região de estudo. Estes dados foram validados pela APA, portanto podemos considerar que são fiáveis, lembrando que há incertezas para certas regiões da área de estudo, há uma variância entre os valores estimados e medidos.

Mas os resultados do produto CAMS a partir das análises efetuadas através dos modelos que o constituem, apresentaram diferenças ao longo da área de estudo quando comparado com a estimativa. Mais ao centro da área de estudo, está localizado o centro da cidade do Porto. Onde há uma maior densidade urbana, comércio e consequentemente um maior tráfego. Portanto, é esperado uma maior concentração de MP₁₀ do que em outras áreas, como pode ser visto na estimativa feita para 2018. Já o produto CAMS nos traz dados diferentes, mostra para esta região valores abaixo dos medidos nas estações de

monitoramento de qualidade do ar, principalmente onde encontram-se as duas estações mais próximas dentro da área de estudo. Ademais, comparando os dados CAMS com a variância da estimativa, vemos que em certos locais onde há uma menor variância, é onde os dados da análise de qualidade do ar do CAMS diferem da estimativa obtida através da interpolação.

Para o ano de 2019, o comportamento foi similar, como pode ser visto na imagem abaixo (figura 08). Neste caso, a variância sobre os pontos onde encontram-se as estações de monitoramento é baixa, e a variância entre os pontos de dados estimados, é mais alta. Mas comparado com o ano de 2018, a variância foi menor. Mas ainda sim, o valor da incerteza não é discrepante. A partir destes dados, podemos comparar com os produtos CAMS também apresentado na figura 08. A estimativa para o ano de 2019, apresenta um valor mínimo menor que o do produto CAMS e o valor máximo, maior. Novamente, isto ocorre onde estão localizadas as estações de monitoramento.

Material Particulado MP10 - 2019

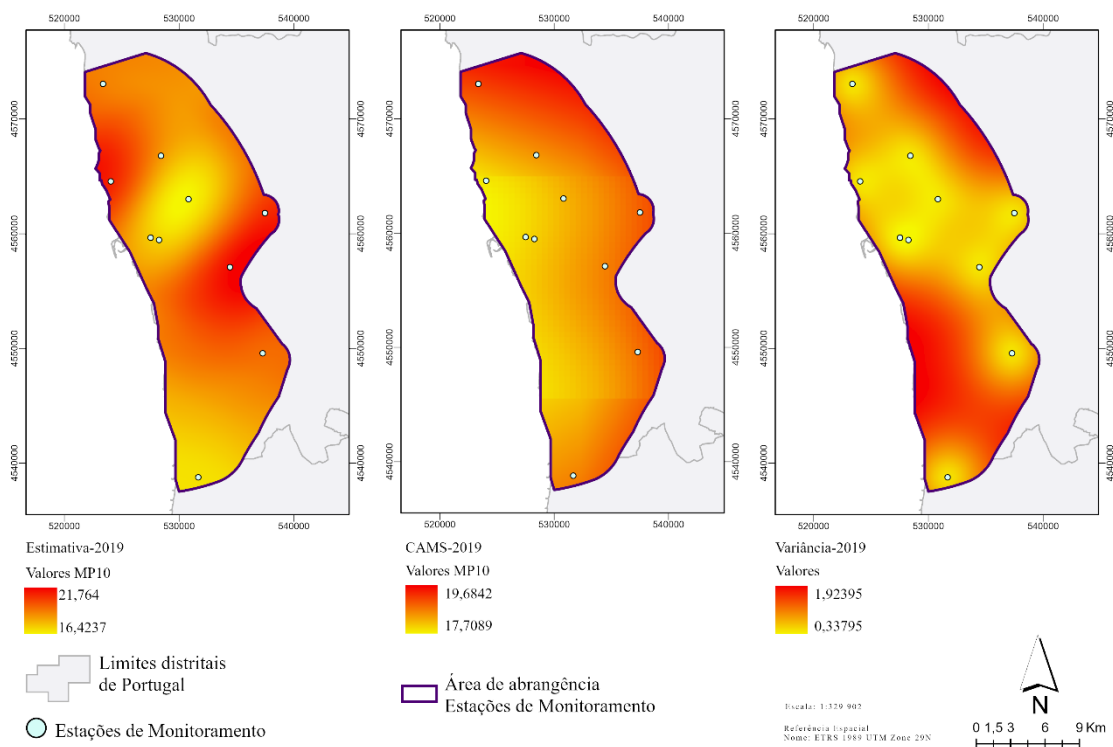


Figura 08: Mapa MP10 – 2019
Elaboração: Autora, 2021

Os valores estimados a partir da interpolação para o ano de 2019, mostram uma maior concentração MP₁₀ - nas regiões em vermelho mais intenso, principalmente na faixa litorânea mais ao norte. Neste local encontra-se o Porto de Leça da Palmeira, qual

pertence a cidade de Matosinhos, mas abastece a toda a região. Possivelmente esta maior concentração ocorreu por causa de um tráfego marítimo mais intenso durante o ano de 2019. Em contrapartida, os dados da análise do CAMS estimou para a mesma área valores mais baixos quando comparados com os valores obtidos através das estações de monitoramento de qualidade do ar. Pode ser verificado mais ao sul da área de estudo, valores similares aos encontrados pelos dados da CAMS. Para este local foram estimados valores de concentração entre o mínimo (valor no local da estação) e $18 \mu\text{m}^3$, através da interpolação. Mas deve-se levar em consideração a variação dos valores estimados para toda a área. Outro detalhe, os valores de concentração obtidos pelos produtos CAMS são feitos a partir de diferentes modelos, que também tem uma leve variância seus resultados.

Como podemos ver, há uma diferença entre os valores resultantes da interpolação e dos dados CAMS. Para verificar se há correlação entre eles, foi calculada a estatística para os dados matriciais do ano de 2018 e 2019, utilizando a ferramenta “*band statistic collection*”. A partir deste método pode ser computada a matriz de correlação, além dos parâmetros básicos de estatística para cada camada (tabela 03).

Parâmetros Estatísticos	2018			2019		
	Estimativa	CAMS	Variância	Estimativa	CAMS	Variância
Mínimo	12,1378	13,9762	0,9185	16,4237	17,7089	0,3379
Máximo	22,1691	24,2739	3,6348	21,7640	19,6842	1,9240
Média	18,0992	19,7791	2,4058	18,8229	18,8628	0,9945
Desvio Padrão	2,4573	2,4878	0,5529	1,1844	0,4217	0,4211

Tabela 03: parâmetros estatísticos básicos
Elaboração: Autora, 2021

A partir destes parâmetros estatísticos, podemos ver que o desvio padrão para o ano de 2018, tanto para os dados interpolados quanto para os dados CAMS, tem um maior desvio padrão (valores próximos) quando comparado com o ano de 2019. Os valores da média tanto para o a Estimativa quanto para o dado CAMS, são similares. Mas o desvio padrão do CAMS é menor que da estimativa, estando este mais próximo da média. Mas, os valores máximos do CAMS não condizem com os valores máximos coletados nas estações e nem os valores mínimos.

Em seguida, foi obtida a matriz de correlação. Esta matriz mostra a correlação entre duas camadas, ou seja, é uma medida de dependência entre as camadas. Uma correlação positiva indica uma relação direta entre duas camadas, como quando os valores das células de uma camada aumentam, os valores das células de outra camada também tendem a aumentar. Uma correlação negativa significa que uma variável muda

inversamente à outra. Uma correlação de zero significa que duas camadas são independentes uma da outra (ESRI, 2021). A matriz de correlação obtida para 2018, pode ser vista abaixo (tabela 04).

Camadas	Estimativa	CAMS	Variância
Estimativa	1,00000	-0,14321	-0,35028
CAMS	-0,14321	1,00000	0,07819
Variância	-0,35028	0,07819	1,00000

Tabela 04: Matriz de Correlação
Elaboração: Autora, 2021

A correlação entre a camada “Estimativa” e “CAMS” é negativa, mas fraca, pois está próximo de zero. Já a correlação entre a “Estimativa” e “Variância” é esperado, pois como visto acima na figura 07 as regiões onde há uma menor variância, são as mais próximas das estações, então espera-se que o valor seja menor. Mas também podemos constar uma correlação positiva, mas fraca - muito próxima a zero, entre as camadas “Variância” e “CAMS”, isso pode ter ocorrido por causa das áreas onde o CAMS tem valores diferentes da interpolação, por isso se correlaciona positivamente com a camada “Variância” e não com a “Estimativa”.

No ano de 2019, as camadas apresentaram correlação negativa entre elas (tabela 05), mas muito fraca e muito próxima a zero. Principalmente entre a camada “CAMS” e “Estimativa”, ou seja, pode-se dizer que não há dependência entre elas. A correlação negativa, mas muito fraca, pode ser explicada pelos diferentes valores obtidos nas regiões onde há estações, se olharmos para a figura 08 novamente, podemos ver que há locais onde os valores são maiores na camada “Estimativa” e menores na camada “CAMS” e vice-versa, como também pode ser influência da reamostragem dos dados “CAMS”.

Camadas	Estimativa	CAMS	Variância
Estimativa	1,00000	-0,03941	-0,08754
CAMS	-0,03941	1,00000	-0,04605
Variância	1,00000	-0,03941	-0,08754

Tabela 05: Matriz de Correlação – 2019
Elaboração: Autora, 2021

As análises estatísticas dos dados matriciais só foram possíveis pois os dados atendem os requisitos para que se possa efetuar o teste adequado. A amostra tem distribuição normal, homogênea, independentes, estes parâmetros foram verificados antes de efetuar a análise espacial.

4. CONSIDERAÇÕES FINAIS

Com os resultados obtidos através das diferentes etapas para estimar a concentração para a área de estudo, o método de interpolação utilizado demonstrou que os valores não foram tão diferentes, quando comparado ao CAMs. Os dados obtidos através do CAMs para a região de estudo tem valores mínimos e máximos que diferem da concentração encontrada para o MP_{10} através da krigagem. Mas isso ocorre pois são utilizadas apenas 3 estações na região do Porto, para validação destes produtos, e estas encontram-se próximas ao litoral. Outro detalhe importante é que uma das estações está situada próxima a uma refinaria e ao aeroporto do Porto, e uma outra próxima do Porto de Leça da Palmeira, o que contribuiu para valores mais altos nas estimativas dos anos 2018 e 2019. O que não foi identificado nas análises de qualidade do ar do CAMS.

Portanto, o valor obtido através da interpolação para ambos os anos, possui diferentes valores máximos e mínimos do produto CAMS, principalmente onde estão localizadas as estações de monitoramento da rede QualAr. A análise da qualidade do ar do produto CAMS demonstra ter uma distribuição homogênea da concentração para a área de estudo, mas os valores onde a mancha urbana é mais densa estão abaixo do medido nas estações e isto é visto mesmo com os valores médios das estações, sem a krigagem. Mas também é necessário verificar a descontinuidade dos dados obtidos das estações, quais podem contribuir para valores menores ou maiores de concentração para um certo período de tempo, influenciando no valor da concentração média anual de PM_{10} na estação.

Uma análise mais detalhada dos dados da rede de medição de qualidade do ar deve ser efetuada para verificar se há esta influência ou não na média horária anual. Talvez, um período menor de medição deva ser considerado em futuras análises, a fim de verificar se os valores mínimos e máximos diferem de um dado para o outro. Ademais, os resultados da krigagem deste estudo, são de um modelo isotrópico. A escolha pelo modelo isotrópico foi feita pois simplifica a modelagem, quando em eventos naturais esta isotropia é rara, o que pode ser considerada mais uma limitação para este estudo. Apesar dos poucos dados disponíveis para estimação da krigagem, temos que levar em consideração que a variância da krigagem se torna mais fraca perto dos locais onde encontram-se as estações, e mais forte onde os dados foram estimados, consistente com os valores obtidos do QualAr para os respectivos anos, demonstrando a validade do método aplicado.

Os resultados parecem indicar que os produtos CAMs têm valores aproximados dos dados medidos *in situ*, principalmente quando consideramos os valores obtidos na análise estatísticas como a média e o desvio padrão. Sabemos que as estações de monitoramento, são mais precisas e de extrema importância para monitorar a concentração dos diferentes poluentes atmosféricos visando a proteção da saúde humana e do meio ambiente, principalmente em áreas urbanas. Mas com a melhora e o avanço dos modelos de análise de qualidade do ar associados ao sensoriamento remoto, produtos como o CAMS são de grande valia para implantação de políticas públicas e monitoramento principalmente onde não há estações disponíveis ou há falta de dados. Assim, podemos verificar que a medição *in situ* continua a ser importante tanto para melhorar e validar dados oriundos do sensoriamento remoto, como também para controle da qualidade do ar.

REFERÊNCIAS

ALVAREZ-MENDOZA, C. I.; TEODORO, A.; RAMIREZ-CANDO, L. Spatial estimation of surface ozone concentrations in Quito Ecuador with remote sensing data, air pollution measurements and meteorological variables. **Environmental Monitoring and Assessment**, v. 191, n. 3, 2019.

Burrough, P.A. Mcdoneel, R. Principles of Geographical Information Systems. **Spatial information systems and geostatistics**. New York: Oxford University Press, 1998

CAMARGO, E.; DRUCK, S.; CÂMARA, G. Análise de Superfícies por Geoestatística Linear. In: DRUCK, S. et al. (Eds.). . **Análise Espacial de Dados Geográficos**. Brasília: EMBRAPA, 2004.

COPERNICUS ATMOSPHERE MONITORING SERVICES. **Data | Copernicus**. Disponível em: <<https://atmosphere.copernicus.eu/data>>. Acesso em: 3 jun. 2021.

ESRI. **ArcGIS Pro Resources | Tutorials, Documentation, Videos & More**. Disponível em: <<https://www.esri.com/en-us/arcgis/products/arcgis-pro/resources>>. Acesso em: 11 maio. 2021.

EUROPEAN ENVIRONMENTAL AGENCY. **Air quality in Europe - 2020 report**. Copenhagen: [s.n.]. Disponível em: <<http://europa.eu>>.

GABRIEL, M. F. et al. Assessment of indoor air conditions in households of Portuguese families with newborn children. Implementation of the HEALS IAQ checklist. **Environmental Research**, v. 182, p. 108966, 2020.

Mazzini, P. L. F.; Schettini, C. A. F. Avaliação de metodologias de interpolação espacial aplicadas a dados hidrográficos costeiros quase sinóticos. **Brazilian Journal of Aquatic Science and Technology**, v. 13, n. 1, p. 53-64, 2009.

MILITINO, A. F.; UGARTE, M. D.; PÉREZ-GOYA, U. An introduction to the spatio-temporal analysis of satellite remote sensing data for geostatisticians. In: **Handbook of Mathematical Geosciences: Fifty Years of IAMG**. [s.l.] Springer International Publishing, 2018. p. 239–253.

OLIVEIRA, M. et al. Geospatial Analysis of Environmental Atmospheric Risk Factors in Neurodegenerative Diseases: A Systematic Review. **International Journal of Environmental Research and Public Health**, v. 17, n. 22, 13 nov. 2020.

ROUX, E. et al. Toward an early warning system for health issues related to particulate matter exposure in brazil: The feasibility of using global pm2.5 concentration forecast products. **Remote Sensing**, v. 12, n. 24, p. 1–45, 2 dez. 2020.

WHO REGIONAL OFFICE FOR EUROPE. Health Risk Assessment of air pollution. p. 1–40, 2016.

WORLD HEALTH ORGANIZATION, W. WHO Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide: Global update 2005. p. 1–21, 2005.

YAMAMOTO, Jorge Kazuo. An alternative measure of the reliability of ordinary kriging estimates. **Mathematical Geology**, 2000, 32.4: 489-509.

YANG, X. et al. Development of PM 2.5 and NO 2 models in a LUR framework incorporating satellite remote sensing and air quality model data in Pearl River Delta region, China. **Environmental Pollution**, v. 226, jul. 2017.