Interpolação espacial de variáveis ambientais e aerossóis na região da Bacia Amazônica próxima a Manaus-AM

Renato O. Miyaji[1], Lucas O. Bauer[2], Victor M. Ferrari[1], Felipe V. de Almeida[1], Pedro L. P. Corrêa[1], Luciana V. Rizzo[2]

[1]Escola Politécnica – Universidade de São Paulo (USP)

[2]Universidade Federal de São Paulo (UNIFESP)

{re.miyaji,ferrari.victor99,felipe.valencia.almeida,pedro.correa}@usp.br

{bauer,lrizzo}@unifesp.br

Abstract. *In the Amazon Forest, Manaus city is considered an ideal laboratory for studies on the effects of human activies on terrestrial ecosystems and climate in a tropical forest. Through the GOAmazon 2014/15 project, researchers were able to collect atmospherical data, regarding polutants and meteorological variables. In this context, this project aimed to collect and treat air pollutant and meteorological data, comparing spatial interpolation methods, such as linear, through splines and nearest neighbor, and choosing the best one to map the variables in the region downwind of Manaus city and create new knowledge about Manaus plume and its impact to the environment. These will enable the development of new experiments.*

Resumo. *Localizada na floresta amazônica, a cidade de Manaus é considerada um laboratório ideal para o estudo dos efeitos da ação antrópica no clima e nos ecossistemas terrestres em uma floresta tropical. O projeto GOAmazon 2014/15 coletou dados atmosféricos de poluentes e variáveis meteorológicas na região. Neste projeto, buscou-se realizar a coleta e o tratamento dos dados, comparando métodos de interpolação espacial, como linear, por splines e nearest neighbor, e selecionando o com o melhor desempenho para estabelecer um mapeamento das variáveis na região vento abaixo da cidade de Manaus e agregar conhecimentos sobre pluma de Manaus e seus impactos no ambiente, que possibilitarão o desenvolvimento de novos experimentos.*

1. Introdução

Os ciclos hidrológicos e energéticos que se desenvolvem em florestas tropicais, como a amazônica, são fundamentais para a manutenção do clima e da circulação atmosférica da Terra. Porém, a influência das ações antrópicas nessas regiões ainda é pouco estudada [Martin et al. 2017]. O projeto *GOAmazon* 2014/15 foi desenvolvido na região entre Manaus e Manacapuru (AM) pelo ARM (*Atmospheric Radiation Measurement*), programa de pesquisa vinculado ao governo dos Estados Unidos da América, em parceria com instituições brasileiras, como a Universidade de São Paulo (USP), a Universidade do Estado do Amazonas (UEA) e o Instituto Nacional de Pesquisas Espaciais (INPE). Esse possuía como objetivo principal entender e quantificar os impactos da pluma de Manaus (AM) sobre as concentrações de poluentes e sobre as interações entre aerossóis, nuvens e precipitação em uma região de floresta tropical [Martin et al. 2016].

No centro da floresta amazônica, a cidade de Manaus (AM) apresentou-se como um laboratório único para o estudo do efeito da influência do ser humano no clima, nos ecossistemas terrestres e na qualidade do ar em um contexto de floresta tropical [Martin et al. 2017]. Para isso, coletou-se dados de aerossóis, de nuvens, além de meteorológicos. A coleta foi realizada por meio de nove estações de pesquisa terrestres situadas nas proximidades de Manaus (AM). Além dessas, os dados também foram coletados através de voos realizados por aeronaves pertencentes ao ARM. Uma delas, um G-159 *Gulfstream* que realizou voos em baixa altitude, apresentada na Figura 1. Foram realizados dois períodos de operação intensiva durante a estação seca e a chuvosa.

Figura 1. Aeronave utilizada para coleta de dados (esquerda). Trajetórias percorridas pela aeronave (linhas vermelhas) e localização das estações terrestres (pinos amarelos) (direita) [Martin et al. 2017]

Para que análises mais aprofundadas possam ser realizadas com os dados obtidos, mostra-se necessária a aplicação de técnicas que permitam a extração de novas informações a partir dos dados, gerando conhecimento. Apesar dos dados coletados pelas estações terrestres apresentarem uma alta resolução temporal, essa é baixa em termos espaciais. Assim, o objetivo deste trabalho é a aplicação de métodos de interpolação espacial utilizando os dados do repositório do projeto *GOAmazon* 2014/15 para a geração de novas informações, através do mapeamento de variáveis ambientais e poluentes atmosféricos em uma grande área. A partir desse, será possível desenvolver experimentos, como o da modelagem de distribuição de espécies, visando determinar a influência das variáveis ambientais e de poluentes atmosféricos nos ecossistemas terrestres da região de Manaus e Manacapuru (AM).

2. Trabalhos Relacionados

Análises, como as de [Martin et al. 2017], foram realizadas a partir dos dados coletados pelas estações terrestres e aeronaves do projeto *GOAmazon* 2014/15. Nela, houve um enfoque na visualização e interpretação de variáveis, para se observar a influência da pluma de Manaus (AM) na região. Já no trabalho de [Cirino et al. 2018], utilizou-se os dados de apenas duas estações terrestres para estudar os efeitos da interação das emissões biogênicas e antrópicas. Assim, foi possível identificar e caracterizar o transporte da pluma de Manaus (AM).

Em relação à interpolação espacial, [Wong et al. 2004] realizou uma comparação entre métodos para mapear variáveis atmosféricas, como ozônio (O_3) e o material particulado (PM_{10}), a partir de dados de sensores localizados ao longo do território norte-americano. Foram testadas quatro diferentes técnicas: média espacial, o algoritmo de

nearest neighbor, o inverso da potência das distâncias (*IDW*) e a krigagem. Segundo os autores, não houve uma grande diferença nos resultados obtidos a partir das técnicas testadas. Porém, mostrou-se uma preferência pela krigagem, por conta de sua capacidade de manter a fidelidade aos dados originais. Além dessas técnicas, interpolações lineares, por *splines* e funções de base radial foram aplicadas com sucesso [Myers 1994].

Tanto no trabalho de [Martin et al. 2017], quanto no de [Cirino et al. 2018], nota-se uma limitação referente à localização dos dados disponíveis, que se restringem às estações terrestres e às trajetórias percorridas pelas aeronaves. Também não foi possível localizar na literatura trabalhos que aplicassem técnicas de interpolação espacial nos dados do projeto *GOAmazon* 2014/15. Caso existissem dados que abrangessem uma área maior, análises mais detalhadas a respeito da concentração de aerossóis poderiam ser feitas. Nesse contexto, a aplicação de métodos, como o de interpolação espacial, mostram-se de grande utilidade.

3. Metodologia

Para a realização do projeto, inicialmente, foi necessário obter os dados coletados pelas aeronaves do ARM durante o projeto *GOAmazon* 2014/15. Em seguida, realizou-se o tratamento e a filtragem dos dados, de modo a permitir a manipulação desses e a posterior aplicação do algoritmo de interpolação espacial. Os dados coletados a partir dos repositórios do ARM foram: a temperatura, as concentrações de monóxido de carbono (CO), ozônio (O_3), óxidos de nitrogênio (NO_x), metano (CH_4), compostos orgânicos voláteis, como o isopreno e a acetonitrila e a concentração numérica de partículas (CPC). Para este projeto, o escopo foi restrito aos dados coletados durante a estação seca, durante a qual foram realizados 19 voos entre setembro e outubro de 2014.

3.1. Estratégias para Tratamento e Filtragem dos Dados e Desenvolvimento dos Códigos

O projeto foi desenvolvido utilizando a linguagem *Python*, por conta da grande variedade de bibliotecas existentes, através da aplicação *web open source Jupyter Notebook* [JupyterTeam 2015]. As bibliotecas utilizadas para o desenvolvimento deste projeto foram: *Pandas* [PandasDevelopmentTeam 2021], *Matplotlib* [Hunter et al. 2012], *NumPy* [ScipyCommunity 2021] e *SciPy* [SciPyDevelopers 2021]. Para o projeto, utilizou-se os dados coletados pela aeronave G-1 durante seu segundo período de operação intensiva: IOP 2, na estação seca. O fluxograma utilizado para manipulação dos dados pode ser visto na Figura 2.

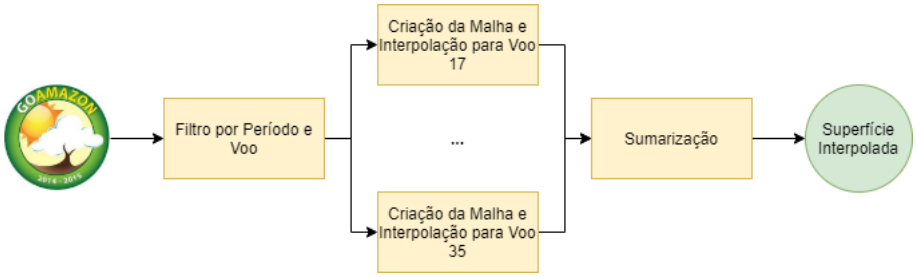

Figura 2. Etapas da manipulação de dados utilizada

Inicialmente, separou-se os dados de interesse, realizando uma separação por voos. Para cada voo, utilizou-se apenas os dados referentes à variável de interesse a ser interpolada, além das coordenadas geográficas (latitude e longitude) e da altitude. Então, aplicou-se o primeiro filtro: foram consideradas apenas as medições realizadas abaixo da altitude de 1500 m, em razão da espessura média da Camada Limite Planetária (CLP) para evitar o uso de dados representativos da troposfera livre. A CLP é a camada da atmosfera mais próxima da superfície, influenciada pelo atrito, pela turbulência e por emissões de poluentes provenientes da superfície terrestre [Ahrens 2012]. Sua espessura é bastante variável ao longo do espaço e do tempo, mas apresenta valores médios de 1100 m de altitude em áreas florestais e de 1650 m em áreas de pastagem na Amazônia [Fisch et al. 2004]. A maior parte dos dados foram coletados em altitudes abaixo de 1500 m (64 % do total). Esse filtro foi aplicado de modo a possibilitar a caracterização da distribuição espacial dos poluentes próxima à superfície.

Posteriormente, criou-se a malha de interpolação, definida a partir dos valores máximos e mínimos de latitude e longitude entre todos os voos presentes para o mesmo período de operação. Adotou-se valores de discretização da malha suficientemente pequenos para a precisão desejada, dada pelo maior intervalo entre coordenadas geográficas de dois pontos coletados de modo subsequente, na ordem de 10^{-3} °. Com a malha pronta, realizou-se a interpolação para cada um dos voos.

A partir dos valores interpolados, utilizou-se o método mais adequado para sumarizar cada variável durante o período de operação intensiva. Utilizou-se o seguinte filtro de frequência: cada simplexo na malha deveria possuir valor diferente de "NaN" (*Not a Number*) em pelo menos metade dos voos para poder ser considerado. Ainda retirou-se os valores extremos, por serem considerados *outliers* através da regra do intervalo interquartil [Upton and Cook 1996]. Aplicando o procedimento para cada um dos simplexos da malha, foi possível sumarizar cada variável. Adotou-se a média como método de sumarização para todas as variáveis. Por fim, utilizou-se a biblioteca *Matplotlib* para gerar a visualização dos dados.

3.2. Manipulação de Dados

A primeira fase de manipulação de dados foi feita através da separação dos dados do projeto *GOAmazon* 2014/15 entre os períodos de operação e, posteriormente, entre os diversos voos realizados pela aeronave G-1. Para isso, utilizou-se as funções de filtragem da biblioteca *Pandas*. Para cada voo, aplicou-se o filtro de altitude citado anteriormente, também utilizando funções de filtragem. Então, determinou-se os valores máximos e mínimos tanto de latitude, quanto de longitude, com o uso das funções implementadas em *Python*. Esses valores foram utilizados posteriormente, junto aos valores de discretização das coordenadas definidos, para a criação da malha para interpolação. Isso foi feito utilizando a função *meshgrid()* da biblioteca *NumPy*.

Na sequência, utilizou-se o comando *interpolate.griddata()* da biblioteca *SciPy*. A partir da definição do hiperparâmetro de método desse, pode-se variar a técnica de interpolação espacial a ser aplicada. Utilizou-se três diferentes métodos, o primeiro deles faz o uso da função *interpolate.LinearNDInterpolator()*, que aplica uma interpolação linear baricêntrica segmentada em cada simplexo da malha, obtido a partir de pontos conhecidos pela triangulação de Delaunay [SciPyDevelopers 2021]. Além disso, também

aplicou-se uma interpolação através de *splines*, ou seja, uma interpolação segmentada por meio de funções polinomiais de terceiro grau. Por fim, utilizou-se também a função *interpolate.NearestNDInterpolator()*, que aplica o algoritmo de *nearest neighbor* para realizar a interpolação espacial [SciPyDevelopers 2021].

Com as superfícies de interpolação obtidas para cada voo, desenvolveu-se uma função para o cálculo da sumarização da variável, sujeita aos filtros de frequência e dispersão citados anteriormente.

3.3. Estratégia de Comparação de Métodos de Interpolação

Com o objetivo de se validar os resultados obtidos através da interpolação espacial, aplicou-se uma validação cruzada nos *datasets*. Através do método *k-fold* com $k = 3$, o *dataset* de cada voo utilizado para a interpolação é dividido de forma aleatória em três partes distintas. Então, utiliza-se duas dessas para realizar a interpolação. A partir dos dados da parcela restante, é possível validar os resultados, calculando os erros entre a interpolação obtida para aqueles pontos e os valores observados. O procedimento é repetido sucessivamente até que cada parte seja utilizada para validação. Ao final, calcula-se a média dos erros obtidos para cada iteração [Anguita et al. 2012].

Por meio do método de *k-fold*, foi possível calcular diferentes funções de custo, como o erro médio absoluto (MAE - *Mean Absolute Error*), o erro médio absoluto percentual (MAPE - *Mean Absolute Percentage Error*) e a raiz quadrada do erro quadrático médio (RMSE - *Root-Mean-Square Error*). Através dessas foi possível avaliar a acurácia da interpolação, comparando os valores interpolados com os valores observados no *dataset* de validação. Quanto menores os valores das funções, o método utilizado apresenta maior acurácia. A principal função a ser minimizada foi o RMSE, por apresentar uma sensibilidade maior a erros mais significativos [Mancuso 2013].

Para se comparar os diferentes métodos de interpolação aplicados e determinar o ideal, avaliou-se as funções de custo citadas acima. Além disso, também foram considerados critérios, como a suavidade das curvas e a fidelidade aos dados originais. Adicionalmente, utilizou-se uma outra forma de validação dos métodos de interpolação espacial utilizados, comparando os resultados obtidos com a média dos valores medidos pelas estações de pesquisa terrestres, como a T2 e a T3.

4. Resultados

Para cada uma das variáveis obtidas, aplicou-se a metodologia descrita anteriormente. Utilizou-se três diferentes métodos de interpolação espacial: linear, por *splines* e pelo algoritmo de *nearest neighbor*. Para avaliar cada um desses, utilizou-se uma validação cruzada, através do método de *3-folds*, calculando métricas de erros.

Para a concentração de monóxido de carbono (CO) em partes por milhão, tanto em relação ao MAE, quanto ao RMSE, o método de interpolação por *splines* apresentou os maiores erros médios: 0,0209 e 0,0478 de MAE e RMSE, respectivamente. Já entre os demais métodos, o linear apresentou consistentemente erros menores em comparação com o *nearest neighbor*: 0,0152 de MAE e 0,0300 de RMSE para o linear, enquanto o *nearest neighbor* obteve 0,0165 e 0,0332, respectivamente. O que representa uma redução de 7,8% em MAE e de 9,6% em RMSE. Também foi calculado o MAPE médio para todas as variáveis, notando o mesmo comportamento: o método de interpolação linear

apresentou um MAPE médio menor, de 9,29%, em comparação com os demais: 10,09% para o método por *splines* e de 12,75% para o método *nearest neighbor*.

Na Figura 3, pode-se observar a interpolação obtida para a concentração de CO em partes por milhão durante o voo 18 pelos três diferentes métodos aplicados. Nela, nota-se que a interpolação por *splines* apresenta a maior suavidade das curvas. Em seguida, os resultados por interpolação linear possuem a segunda maior suavidade. Já a interpolação por *nearest neighbor* apresenta pouca suavidade e limiares bruscos entre as regiões. No entanto, para as três técnicas, é perceptível que os valores próximos à trajetória real da aeronave são semelhantes.

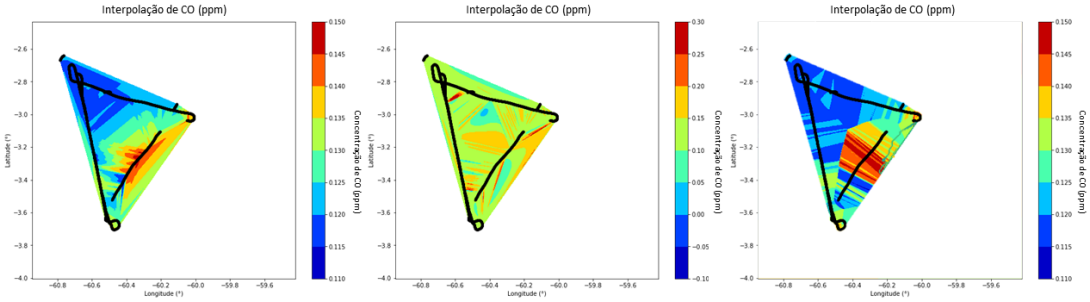

Figura 3. Comparação da suavidade das curvas para cada método de interpolação: linear (esquerda), por *splines* (centro) e por *nearest neighbor* (direita). Linhas em preto correspondem à trajetória percorrida pela aeronave durante o voo

Como outra forma de validação, também comparou-se os resultados médios para todos os voos obtidos por meio da interpolação na região acima das estações de pesquisa terrestres com os valores médios medidos por elas durante o mesmo período e horário. Foi possível obter os dados da concentração de CO e de O_3 para duas estações: T2 e T3. Para elas, foi calculado o MAPE para os diferentes métodos de interpolação. Notou-se que a interpolação por *spline* possuí o maior MAPE de 41,69 % , assim como possuía os maiores erros na validação cruzada. Porém, neste caso, o erro da interpolação pelo algoritmo de *nearest neighbor* é menor: 12,07 %, enquanto o linear é de 14,08 %.

Vale ressaltar que em todos os casos, os valores obtidos por meio da interpolação subestimaram as medições obtidas pelas estações terrestres. Além disso, erros um pouco mais elevados eram esperados devido à diferença da altitude da coleta dos dados de estações terrestres e de aeronaves. No entanto, a validação não é comprometida, pois a turbulência na CLP irá promover uma mistura contínua vertical de calor, umidade e momento de maneira uniforme [Stull 1988]. Portanto, os valores das variáveis ambientais tendem a ser uniformes em função da altitude na camada de mistura da CLP, porém podem haver discrepâncias devido a fenômenos meteorológicos como inversões térmicas.

Considerando os resultados das comparações realizadas, optou-se por utilizar o método de interpolação espacial linear, por conta dos menores erros (MAE, RMSE e MAPE) obtidos através da validação cruzada. Apesar de apresentar erros levemente superiores em relação à interpolação por *nearest neighbor* para a validação pelos dados das estações terrestres, nota-se que seus resultados apresentam uma suavidade maior nas curvas, tornando-os mais semelhantes com o esperado na realidade. Assim, aplicou-se a interpolação espacial linear. Para cada uma das variáveis, obteve-se a superfície de

interpolação para cada voo. Então, sumarizou-se para cada variável durante a estação seca. Na Figura 4, pode-se observar os resultados da sumarização.

Analisou-se os resultados do método de interpolação linear, apresentados na Figura 4. A interpolação de NO_X apresenta concentrações superiores aos resultados da simulação com o modelo de transporte químico de alta resolução, realizada por [Shrivastava et al. 2019]. No entanto, os dados utilizados pelos autores são da estação chuvosa, ao passo que os dados deste trabalho são da estação seca. Os valores obtidos na interpolação são da ordem de magnitude de 5 a 20 ppb, com valores máximos observados restritos as latitudes e longitudes mais próximas de 60º e 3º, respectivamente, enquanto que o trabalho supracitado tem como valor máximo simulado de 5 ppb.

Em relação à interpolação de O_3 deste trabalho, os valores de concentração são mais próximos em magnitude e as distribuições espaciais se assemelham em comparação com [Shrivastava et al. 2019]. Observa-se o aumento dos valores de concentração de O_3 em áreas mais distantes de Manaus (AM), isso pode ser explicado devido a um regime de produção de O_3 limitado por compostos orgânicos voláteis (COV's) na pluma de poluição de Manaus (AM), que ao se aproximar das áreas florestais, favorece o consumo de NO_X e o aumento da produção de O_3. Assim como discutido por [Orlando et al. 2010], a redução de NO_X em um regime de produção onde a razão COV's/NO_X é limitada por COV's, a redução de NO_X leva ao aumento da produção de O_3.

A interpolação de isopreno apresenta valores máximos de concentração em regiões mais afastadas de Manaus (AM), resultado esperado em função das emissões de isopreno terem origem majoritária biogênica e de seu consumo na pluma de poluição para a produção de poluentes secundários como o O_3 e aerossóis orgânicos secundários. As interpolações de temperatura, CO e CPC apresentam valores máximos em Manaus (AM) e ocorre uma redução desses valores com o aumento da distância da área urbana, assim como a pluma de poluição descrita por [Martin et al. 2017]. O CO ainda apresenta valores máximos próximos às áreas de desmatamento, porém apresenta valores mínimos onde se é observado valores mais elevados de acetonitrila, traçador de queimadas, cuja interpolação tem como resultado os maiores valores de concentração em regiões próximas às áreas de queimadas na estação seca de 2014 e 2015, de acordo com os mapas de áreas queimadas do Programa Queimadas do INPE [INPE 2021].

A interpolação de CH_4 apresentou valores próximos aos valores obtidos no trabalho de [Beck 2012]. Os valores máximos próximos a área urbana de Manaus (AM) podem estar relacionados a emissões biogênicas oriundas da decomposição de resíduos sólidos despejados em corpos d'água e/ou gases emitidos de esgotos a céu aberto [Beck 2012]. Porém, não é possível observar todos os valores máximos de CH_4 em regiões florestais relacionados com os valores máximos de acetonitrila e de CO, pois observou-se no trabalho de [Beck 2012] a associação da ocorrência de queimadas ao aumento de CH_4, com o uso de CO como traçador de queima de biomassa. As principais fontes de metano são naturais, destaca-se áreas inundadas [Basso 2016], mas os seus valores máximos podem estar relacionados também com os corpos d'água que ocorrem na área, como o Rio Negro.

5. Conclusão e Trabalhos Futuros

Este trabalho utilizou os dados referentes a variáveis ambientais e de aerossóis do projeto *GOAmazon* 2014/15 para os quais desenvolveu-se, na linguagem *Python*, uma metodolo-

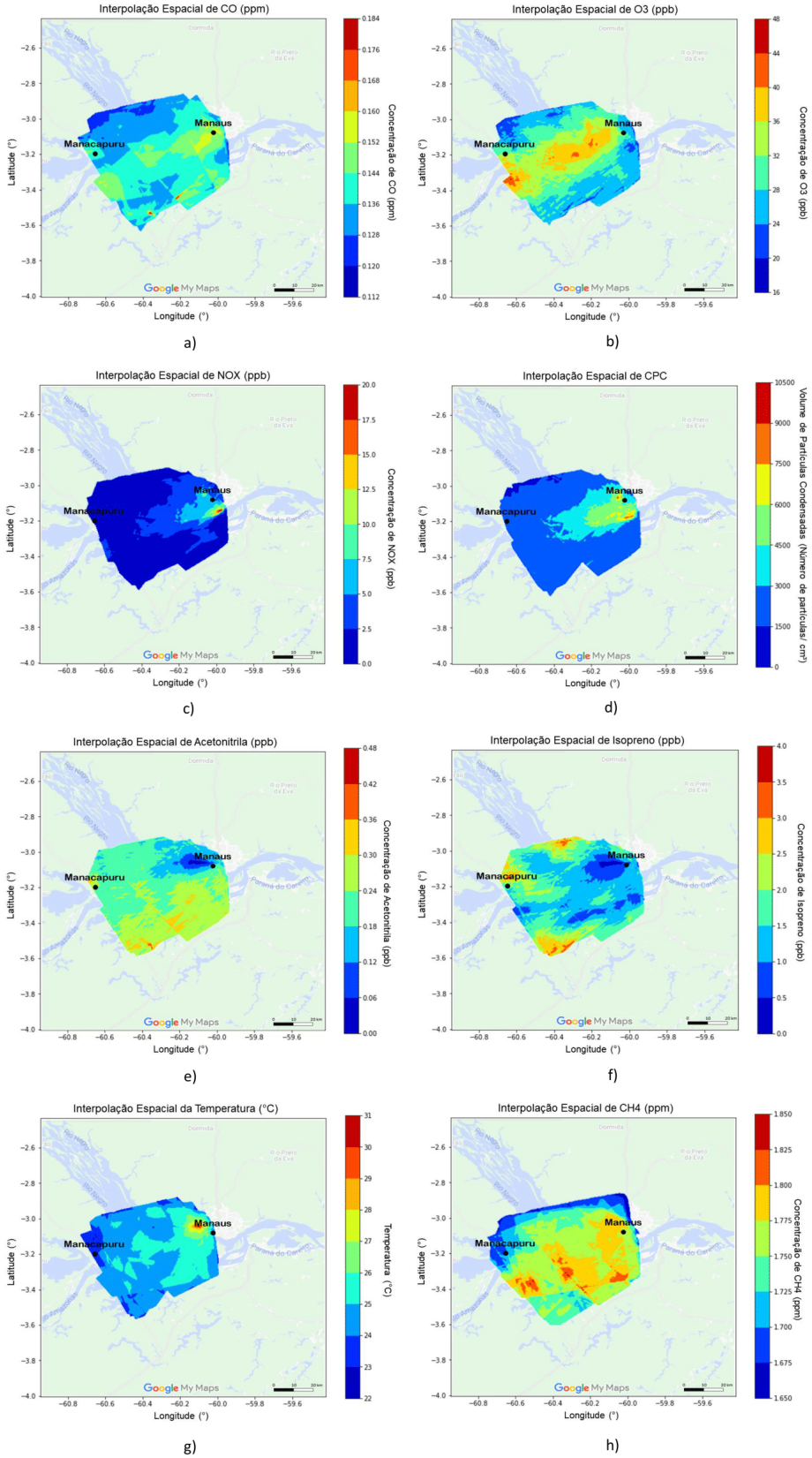

Figura 4. Superfícies de interpolação média para a) CO (ppm) b) O_3 (ppb) c) NO_X (ppb) d) CPC (Partículas/cm^3) e) Acetonitrila (ppb) f) Isopreno (ppb) g) Temperatura (°C) h) CH_4 (ppm)

gia para a aplicação da interpolação espacial. Foram avaliadas três diferentes técnicas de interpolação: linear, por *splines* e pelo algoritmo de *nearest neighbor*, utilizando métricas, como o MAPE. Com o uso de uma interpolação linear baricêntrica, foi possível obter um MAPE médio de 9,29 %, calculado a partir da validação cruzada pelo método de *3-folds*. Além disso, os resultados apresentaram elevada suavidade das curvas, representando bem a realidade.

As superfícies interpoladas obtidas nesse trabalho serão fundamentais para que novos experimentos a respeito da região da Bacia Amazônica sejam realizados, como da modelagem de distribuição de espécies. Os *datasets* gerados para este trabalho estão disponíveis para acesso pela comunidade científica em [Miyaji et al. 2021]. Para o desenvolvimento de trabalhos futuros, sugere-se utilizar os dados provenientes das estações de pesquisa terrestres do projeto *GOAmazon 2014/15*, além dos dados coletados pelas aeronaves, para a aplicação da interpolação espacial. Assim, será possível obter uma superfície de interpolação que compreenda uma área maior. Ademais, sugere-se a avaliação de outras técnicas de interpolação espacial que não foram aplicadas neste trabalho, como a krigagem e o inverso da potência das distâncias (*IDW*).

Agradecimentos

Este trabalho foi possível devido ao apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), através de uma bolsa do programa PIBIC (2020/21 - 1745) e ao Projeto Temático da FAPESP ”Ciclos de vida e nuvens de aerossóis na Amazônia”(2017/ 17047-0).

Referências

Ahrens, C. D. (2012). Meteorology today: an introduction to weather, climate, and the environment. *Cengage Learning*, pages 225–227.

Anguita, A., Ghelardoni, L., Ghio, A., Oneto, L., and Ridella, S. (2012). The ‘k’ in k-fold cross validation. *European Symposium on Artificial Networks, Computational Intelligence and Machine Learning 2012 proceedings*.

Basso, L. S. (2016). Seasonality and internal variability of ch4 fluxes from the eastern amazon basin inferred from atmospheric mole fraction profiles. *Journal of Geophysical Research: Atmospheres*, 121(1):168–184.

Beck, V. (2012). Methane airborne measurements and comparison to global models during barca. *Journal of Geophysical Research: Atmospheres*, 117(15).

Cirino, G., Brito, J., Barbosa, H. M., Rizzo, L., Tundev, P., de Sa, S. S., Jimenez, J. L., Palm, B., Carbone, S., Lavric, J. V., Souza, R. A., Wolff, S., Walter, D., Tota, J., Oliveira, M. B., Martin, S. T., and Artaxo, P. (2018). Observations of manaus urban plume evolution and interaction with biogenic emissions in goamazon 2014/5. *Atmospheric Environment*, 191:513–524.

Fisch, G., Tota, J., and Machado, L. (2004). The convective boundary layer over pasture and forest in amazonia. *Theor Appl Climatol*, 78:47–59.

Hunter, J., Dale, D., Firing, E., Droettboom, M., and the Matplotlib development team (2012). Matplotlib documentation. https://matplotlib.org/stable/users/index.html. Acesso em: 2021-03-01.

INPE (2021). Programa queimadas. http://queimadas.dgi.inpe.br/queimadas/portal. Acesso em: 2021-03-01.

JupyterTeam (2015). The jupyter notebook. https://jupyter-notebook.readthedocs.io/en/stable/. Acesso em: 2021-03-01.

Mancuso, A. C. (2013). Uma investigação do desempenho de métodos de combinação de previsões: Simulada e aplicada. *Universidade Federal do Rio Grande do Sul*.

Martin, S. T., Artaxo, P., Machado, L., Manzi, A. O., Souza, R. A. F. d., Schumacher, C., Wang, J., Biscaro, T., Brito, J., Calheiros, A., et al. (2017). The green ocean amazon experiment (goamazon2014/5) observes pollution affecting gases, aerosols, clouds, and rainfall over the rain forest. *Bulletin of the American Meteorological Society*, 98(5):981–997.

Martin, S. T., Artaxo, P., Machado, L. A. T., Manzi, A. O., Souza, R. A. F. d., Schumacher, C., Wang, J., Andreae, M. O., Barbosa, H., Fan, J., et al. (2016). Introduction: observations and modeling of the green ocean amazon (goamazon2014/5). *Atmospheric Chemistry and Physics*, 16(8):4785–4797.

Miyaji, R. O., Bauer, L. O., Corrêa, P. L. P., Almeida, F. V., Rizzo, L., and Ferrari, V. M. (2021). Interpolação espacial de variáveis ambientais e aerossóis na bacia amazônica. Zenodo. DOI: 10.5281/zenodo.4820954.

Myers, D. E. (1994). Spatial interpolation: an overview. *Geoderma*, 62:17–28.

Orlando, J. P., Alvim, D. S., Yamazaki, A., Corrêa, S. M., and Gatti, L. V. (2010). Ozone precursors for the são paulo metropolitan area. *The Science of the Total Environment*, 408(7):1612–1620.

PandasDevelopmentTeam (2021). Pandas documentation. https://pandas.pydata.org/docs/. Acesso em: 2021-03-01.

ScipyCommunity (2021). Numpy v1.20 manual. https://numpy.org/doc/stable/. Acesso em: 2021-03-01.

SciPyDevelopers (2021). Scipy documentation. https://www.scipy.org/docs.html. Acesso em: 2021-03-01.

Shrivastava, M., Andreae, M. O., Artaxo, P., Barborsa, H. M., Berg, L. K., Brito, J., Ching, J., Isaacman-VanWertz, G., Yee, L. D., Ynoue, R., Zaveri, R. A., Zelenyuk, A., and Zhao, C. (2019). Urban pollution greatly enhances formation of natural aerosols over the amazon rainforest. *Nature Communications*, 10:1–12.

Stull, R. B. (1988). An introduction to boundary layer meteorology. *Springer Science and Business Media*, pages 1–14.

Upton, G. and Cook, I. (1996). *Understanding Statistics*. Oxford University Press, Oxford.

Wong, D. W., Yuang, L., and Perlin, S. A. (2004). Comparison of spatial interpolation methods for the estimation of air quality data. *Journal of Exposure Analysis and Environmental Epidemiology*, 14(5):404–415.