



UNIVERSIDADE FEDERAL DE SANTA CATARINA  
CAMPUS REITOR JOÃO DAVID FERREIRA LIMA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA AMBIENTAL  
NÍVEL MESTRADO

Otávio Nunes dos Santos

Adaptação do método *Potential Source Contribution Function*: Incorporação de campo de  
concentração de poluentes

Florianópolis  
2021

Otávio Nunes dos Santos

Adaptação do método *Potential Source Contribution Function*: Incorporação de campo de concentração de poluentes.

Dissertação submetida ao Programa de Pós-Graduação em Engenharia Ambiental da Universidade Federal de Santa Catarina para a obtenção do título de Mestre em Engenharia Ambiental.

Orientador: Prof. Leonardo Hoinaski, Dr.

Florianópolis  
2021

Ficha de identificação da obra elaborada pelo autor,  
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Santos, Otávio Nunes dos Santos

Atualização do método Potential Source Contribution  
Function: Incorporação de campo bidimensional de  
concentração de poluentes. / Otávio Nunes dos Santos Santos  
; orientador, Leonardo Hoinaski Hoinaski, 2021.  
84 p.

Dissertação (mestrado) - Universidade Federal de Santa  
Catarina, Centro Tecnológico, Programa de Pós-Graduação em  
Engenharia Ambiental, Florianópolis, 2021.

Inclui referências.

1. Engenharia Ambiental. 2. Poluição do ar . 3.  
Identificação de fontes de emissão. 4. Modelos receptores.  
5. Campos de concentração de poluentes. I. Hoinaski,  
Leonardo Hoinaski. II. Universidade Federal de Santa  
Catarina. Programa de Pós-Graduação em Engenharia Ambiental.  
III. Título.

Otávio Nunes dos Santos

Adaptação do método *Potential Source Contribution Function*: Incorporação de campo de concentração de poluentes.

O presente trabalho em nível de mestrado foi avaliado e aprovado por banca examinadora composta pelos seguintes membros:

---

Prof. Pedro Luiz Borges Chaffe, Dr.  
Membro da Banca  
Universidade Federal de Santa Catarina

---

Prof (a). Simone Andréa Pozza, Dr.(a)  
Membro da Banca  
Universidade Estadual de Campinas

---

Prof (a). Taciana Toledo de Almeida Albuquerque, Dr.(a)  
Membro da Banca  
Universidade Federal de Minas Gerais

Certificamos que esta é a versão original e final do trabalho de conclusão que foi julgado adequado para obtenção do título de mestre em Engenharia Ambiental.

---

Prof. Leonardo Hoinaski, Dr.  
Coordenação do Programa de Pós-Graduação em Engenharia Ambiental

---

Prof. Leonardo Hoinaski, Dr.  
Orientador

Florianópolis  
2021



## RESUMO

A identificação de fontes de poluição do ar é uma tarefa desafiadora devido aos efeitos de emissores locais e transfronteiriços. Embora a *potential source contribution function* (PSCF) seja uma ferramenta útil para análise de trajetória e identificação de fontes de emissão, ela pode superestimar o impacto na poluição do ar de áreas menos poluídas pelas quais um evento de trajetória passou. Este trabalho propõe a incorporação de campos de concentração de poluentes atmosféricos a partir de dados de reanálise para melhorar o desempenho do método PSCF. O 3D-PSCF-CONC é baseado em uma versão tridimensional do método PSCF (3D-PSCF), e foi elaborado para evitar a superestimação de áreas menos poluídas. Foi utilizado como caso de teste a Região Metropolitana de São Paulo (RMSP) no Brasil, que é reconhecida pelos problemas de qualidade do ar, para comparar os resultados dos métodos PSCF, 3D-PSCF e 3D-PSCF-CONC. Concentrações horárias de monóxido de carbono (CO) e dióxido de enxofre (SO<sub>2</sub>) de 2015 a 2019 da estação de monitoramento Marginal Tietê – Ponte Remédio, da Companhia Ambiental do Estado de São Paulo (CETESB), foram utilizadas como dados de entrada para aplicação dos métodos. As concentrações médias diárias de CO e SO<sub>2</sub> foram utilizadas nos cálculos de probabilidade. Um total de 1826 trajetórias reversas de massa de ar do HYSPLIT versão 4 foram processadas usando os três métodos. As concentrações superficiais de CO e SO<sub>2</sub> do MERRA-2, com resolução de 0,5 ° x 0,625 °, foram incorporadas ao 3D-PSCF-CONC para corrigir o cálculo da contribuição potencial com base no campo de concentração do poluente atmosférico. Os resultados revelaram que o 3D-PSCF-CONC alcançou uma identificação mais consistente das fontes locais e regionais, ao invés de enfatizar principalmente as fontes regionais como os outros métodos PSCF. O 3D-PSCF-CONC reduziu a contribuição potencial de áreas menos poluídas e com alta precipitação. A análise sugere uma contribuição de fontes veiculares próximas à área receptora na RMSP e o transporte de longo alcance de emissões industriais e queima de biomassa. Este trabalho apresenta uma ferramenta para a compreensão do processo de poluição do ar e identificação de fontes.

**Palavras-chave:** Poluição do ar, identificação de fontes, trajetórias reversas, HYSPLIT, PSCF, MERRA-2.

## ABSTRACT

The identification of air pollution sources is a challenging task due to the confounding effects of local and transboundary emitters. While the Potential Source Contribution Function (PSCF) is a useful tool for trajectory analysis and source identification, it can overestimate the impact of less polluted areas in which a trajectory event passed through. In this work, we propose incorporating air pollutant concentration fields from reanalysis data to improve the PSCF method's performance that avoids the overcontribution of less polluted areas. We develop the 3D-PSCF-CONC based on an updated three-dimensional version of the PSCF method (3D-PSCF). We used as a test case the Metropolitan Area of São Paulo (MASP) in Brazil, one of the most polluted regions in the world, for comparing the results of the conventional and the new PSCF method. Hourly carbon monoxide (CO) and sulfur dioxide (SO<sub>2</sub>) concentrations from 2015 to 2019 of monitoring stations were analyzed. The daily average concentrations of CO and SO<sub>2</sub> were used as a threshold on the probability calculations. An overall of 1826 backward trajectories from HYSPLIT version 4 was processed using the three models. CO and SO<sub>2</sub> surface concentrations from MERRA-2, with a resolution of 0.5° x 0.625° was incorporated in the 3D-PSCF-CONC to correct the potential contribution calculation based on the air pollutant concentration field. Our results revealed that the 3D-PSCF-CONC reached a more consistent identification of local and regional sources, rather than emphasizing mostly the regional sources as the other PSCF methods. The 3D-PSCF-CONC reduced the potential contribution of less polluted and high precipitation areas. The analysis suggests a contribution of vehicular sources close to the receptor site in MASP and long-range transport of industrial emissions and biomass burning from the land clearing of sugarcane production. This work presents an important tool for understanding the air pollution process and source identification.

**Keywords:** Air pollution, source identification, backward trajectories, HYSPLIT, PSCF, MERRA-2.

## LISTA DE FIGURAS

<b>Figura 1.</b> Ilustração esquemática dos conceitos da 2D-PSCF (a) e 3D-PSCF (b, c) em função do deslocamento vertical das massas de ar (Kim et al., 2016). .....	11
<b>Figura 2.</b> Localização do local de estudo na RMSP (Marginal Tietê-Pte. Remédios). Concentrações de CO (a) e SO <sub>2</sub> (b) na estação de monitoramento da RMSP. Uso do solo proveniente do projeto MAPBIOMAS (2019). .....	16
<b>Figura 3.</b> Ilustração esquemática do conceito do método 3D-PSCF-CONC. a) Representa as trajetórias reversas que convergem para RMSP utilizadas no estudo, uma trajetória reversa está destacada em branco para exemplificar as etapas do método. b) cada ponto de trajetória está associado, no tempo, a uma grade de concentração de reanálise atmosférica do MERRA-2. Em c) é apresentada a célula da grade em que está localizado o ponto de trajetória, ao qual será atribuído a concentração. e) Os valores da 3D-PSCF-CONC.....	19
<b>Figura 4.</b> Classes de uso e cobertura do solo do projeto MAPBIOMAS (esquerda). Grade com resolução de 0.5° x 0.625°, que coincide com a resolução das grades dos métodos PSCF empregados neste estudo (direita). .....	23
<b>Figura 5.</b> Mapa da emissão de CO pelo setor de transporte no ano de 2015 (EDGAR, 2017). .....	24
<b>Figura 6.</b> Comparação dos métodos PSCF: a) 2D-PSCF, b) 3D-PSCF, c) 3D-PSCF-SAT, e variabilidade dos valores de PSCF em cada uso do solo do domínio por onde as trajetórias deslocaram. ....	28
<b>Figura 7.</b> Variabilidade dos valores de PSCF em cada uso do solo do domínio por onde as trajetórias se deslocaram.....	28
<b>Figura 8.</b> Distribuição espacial do fluxo médio de CO e TQL média (no topo). Os valores de PSCF x fluxo de concentração de CO são categorizados pela concentração de CO e TQL média do MERRA-2.....	30
<b>Figura 9.</b> Correlação de Spearman entre a distribuição espacial dos valores 3D-PSCF-CONC e as emissões de CO das fontes de emissão do inventário EDGAR.....	32
<b>Figura 10.</b> Correlação entre CO x frota de veículos leves (a) e CO x consumo de gasolina em São Paulo (b). A concentração de CO - MERRA-2 (M2T1NXCHM V5.12.4), frota de veículos leves (CETESB, 2019d), e consumo de gasolina (ANP, 2020). .....	33
<b>Figura 11.</b> Análise do método 3D-PSCF-CONC na RMSP usando dois períodos (a) 2015 a 2017 e (b) 2018 a 2019 para verificar o efeito da redução das emissões da queima de biomassa sobre as contribuições de Barretos. Correlação entre as concentrações de CO e FRP.....	34
<b>Figura A.</b> Resumo esquemático dos módulos utilizados para implementação do método 3D-PSCF-CONC. ....	48



**Figura B.** Comparação entre os métodos PSCF para fontes de SO<sub>2</sub> na RMSP: a) 2D-PSCF, b) 3D-PSCF, c) 3D-PSCF-CONC. Correlação entre os valores de PSCF e concentração de SO<sub>2</sub> (µg.m<sup>-3</sup>) do MERRA-2. Correlação entre os valores de PSCF e precipitação média (mm.h<sup>-1</sup>).  
..... **Erro! Indicador não definido.**62

**Figura C1.** Correlação entre o inventário de emissões EDGAR e o métodos 2D-PSCF. Correlação de Spearman entre a distribuição espacial dos valores 2D-PSCF e as emissões de CO do transporte rodoviário sem ressuspensão do solo, queima de resíduos agrícolas, produção de ferro e aço, combustão para manufatura e refinarias de petróleo e indústria de transformação  
..... 71

**Figura C2.** Correlação entre o inventário de emissões EDGAR e o métodos 3D-PSCF. Correlação de Spearman entre a distribuição espacial dos valores 3D-PSCF e as emissões de CO do transporte rodoviário sem ressuspensão do solo, queima de resíduos agrícolas, produção de ferro e aço, combustão para manufatura e refinarias de petróleo e indústria de transformação  
..... 73

**Figura D.** Análise temporal (a) e correlação de Spearman (b) das concentrações de CO e FRP na região de Barretos/SP..... 75

## LISTA DE ABREVIATURAS E SIGLAS

ARL	Air Resources Laboratory
BC	Black Carbon
CETESB	Companhia Ambiental do Estado de São Paulo
CH <sub>4</sub>	Metano
CO	Monóxido de Carbono
CONAMA	Conselho Nacional do Meio Ambiente
CWT	Concentration-Weighted Trajectory
EDGAR	Emissions Database for Global Atmospheric Research
FIRMS	Fire Information for Resource Management System
FRP	Fire Radioactive Power
GDAS	Global Data Assimilation System
GEOS-5	Goddard Earth Observing System Model
GMAO	Global Modeling and Assimilation Office
HYSPLIT	Hybrid Single-Particle Lagrangian Integrated Trajectory Model
IBGE	Instituto Brasileiro de Geografia e Estatística
IPCC	Intergovernmental Panel on Climate Change
MERRA-2	Modern-Era Retrospective Analysis for Research and Applications – Version 2
MODIS	Moderate-Resolution Imaging Spectroradiometer
MP <sub>10</sub>	Material Particulado Inalável
MP <sub>2,5</sub>	Material Particulado Fino
MVP	Médio Vale do Paraíba
NASA	National Aeronautics and Space Administration
netCDF	Network Common Data Form
NH <sub>3</sub>	Amônia
NMVOCs	Compostos orgânicos voláteis não-metano
NOAA	National Oceanic and Atmospheric Administration
NO <sub>x</sub>	Óxidos de nitrogênio
ONU	Organização das Nações Unidas
PMF	Positive Matrix Factorization
PSCF	Potential Source Contribution Function
RMSP	Região Metropolitana de São Paulo
RTWC	Residence Time Weighted Concentration
SO <sub>2</sub>	Dióxido de enxofre
SSW	Simple Subset Wizard
TRA	Trajectory Regression Analysis



## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>5</b>
1.1	OBJETIVO GERAL	7
1.2	OBJETIVOS ESPECÍFICOS	7
<b>2</b>	<b>REFERENCIAL TEÓRICO</b>	<b>8</b>
2.1	POLUIÇÃO ATMOSFÉRICA TRANSFRONTEIRIÇA	8
2.2	<i>POTENTIAL SOURCE CONTRIBUTION FUNCTION</i> (PSCF)	9
2.3	ASSIMILAÇÃO DE REANÁLISE ATMOSFÉRICA APLICADA A POLUIÇÃO DO AR	12
2.4	INVENTÁRIO DE FONTES DE EMISSÃO	13
<b>3</b>	<b>MATERIAL E MÉTODOS</b>	<b>15</b>
3.1	ÁREAS DE ESTUDO	15
3.2	MÉTODOS 2D-PSCF E 3D-PSCF	16
3.3	DESENVOLVIMENTO DO MÉTODO 3D-PSCF-CONC	18
3.4	DADOS DE ENTRADA DOS MÉTODOS	20
3.4.1	Monitoramento da qualidade do ar	20
3.4.2	Dados de trajetórias reversas de massa de ar	21
3.4.3	Campos de concentração de poluentes atmosféricos	21
3.5	ANÁLISES DE CONSISTÊNCIA DOS MÉTODOS	22
3.5.1	Dados de uso e cobertura do solo	22
3.5.2	Dados de chuva	23
3.5.3	Dados de fontes de emissão de poluentes atmosféricos	24
3.5.4	Dados de frota veicular e consumo de combustível	25
3.5.5	Dados de potência radiativa do fogo	25
<b>4</b>	<b>RESULTADOS</b>	<b>27</b>
4.1	COMPARAÇÃO ENTRE OS MÉTODOS PSCF	27
4.2	ANÁLISE DE CONSISTÊNCIA DOS MÉTODOS PSCF	28
4.2.1	Uso e ocupação do solo	28
4.2.2	Fluxo de poluentes e deposição úmida	29
4.2.3	Identificação das fontes de emissões atmosféricas	30
<b>5</b>	<b>CONCLUSÃO</b>	<b>35</b>
<b>6</b>	<b>REFERÊNCIAS BIBLIOGRÁFICAS</b>	<b>37</b>
	APÊNDICE A: IMPLEMENTAÇÃO DO MÉTODO 3D-PSCF-CONC	49
	APÊNDICE B: COMPARAÇÃO DOS MÉTODOS PSCF NA RMSP - SO <sub>2</sub>	69
	APÊNDICE C: ANÁLISE ESPACIAL ENTRE O INVENTÁRIO DE EMISSÕES EDGAR E OS MÉTODOS PSCF	71
	APÊNDICE D: AVALIAÇÃO DA INFLUÊNCIA DA QUEIMA DE BIOMASSA SOBRE	

AS CONCENTRAÇÕES DE CO EM BARRETOS/SP.....75



## 1 INTRODUÇÃO

A poluição do ar representa um sério risco ambiental para a saúde humana e para os ecossistemas em geral. A maior parte da população mundial está exposta a níveis de qualidade do ar que excedem as recomendações da Organização Mundial da Saúde (OMS) (WHO, 2016). Para reduzir as concentrações de poluentes na atmosfera, as fontes de poluição do ar devem ser identificadas. Além das fontes locais e regionais, o transporte de poluentes de longo alcance pode contribuir significativamente para a deterioração da qualidade do ar de uma determinada região (Vet et al., 2008; Jeong et al., 2011; Zhang et al., 2012; Stojić et al., 2017; Yim et al., 2019; Dimitriou et al., 2021).

Para auxiliar na identificação das fontes de poluição do ar, algumas técnicas de análise da trajetória das massas de ar estão sendo aprimoradas ao longo do tempo (Hsu et al., 2003; Xu et al., 2006; Chalbot et al., 2013; Zong et al., 2018; Shanavas et al., 2020). Dentre elas, a função de contribuição potencial da fonte (*Potential Source Contribution Function* - PSCF) é um dos métodos mais utilizados para indicar regiões onde estão localizadas potenciais fontes de emissão de poluentes atmosféricos, devido à natureza condicional da função de probabilidade (Punsompong; Chantara, 2018; Potier et al., 2019; Shen; Cao, 2020; Wei et al., 2020). O método PSCF utiliza trajetórias reversas das massas de ar para calcular a probabilidade de uma área geográfica específica contribuir para a poluição em um local receptor com base no tempo de residência da massa de ar sobre aquela região (Ashbaugh et al., 1985).

A PSCF convencional, ou bidimensional, (2D-PSCF) não considera o deslocamento vertical da massa de ar e a variação na altura de injeção de poluentes na atmosfera de acordo com o tipo de fonte de emissão. Nesse caso, o método pode superestimar a contribuição de regiões por onde a trajetória da massa de ar se desloca acima da injeção de poluentes. Estas limitações foram corrigidas por uma versão tridimensional do método denominada 3D-PSCF (Kim et al., 2016). No caso da 3D-PSCF, como o cálculo da contribuição considera a altura de injeção dos poluentes, é possível detectar se uma fonte de emissão tem o potencial de contaminar ou não a trajetória. Os estudos de Kim et al., (2016), Kim et al., (2019) e Dimitriou et al., (2021) aplicaram a versão 3D-PSCF e verificaram a influência da variação na altura de injeção de poluentes sobre a contribuição de determinada fonte.

Entretanto, a 3D-PSCF é insensível ao campo de concentração de poluentes atmosféricos. Enquanto se desloca em direção ao local receptor, a trajetória de massa de ar percorre regiões com diferentes níveis de atividades poluidoras, desde áreas com alta densidade de florestas até metrópoles urbanas. Nesse caso, a massa de ar é exposta a diferentes campos de concentração

de poluentes durante seu deslocamento. Ignorar os campos de concentração de poluentes em diferentes regiões pode ocasionar a superestimativa da contribuição de áreas com baixa emissão de poluentes ou aquelas onde os mecanismos de remoção de poluentes ocorrem com mais frequência. Após a emissão, um dos principais mecanismos que controlam as concentrações de poluentes na atmosfera é a deposição úmida (Garrett et al., 2006).

Esse trabalho propõe uma atualização do método 3D-PSCF, através da incorporação dos campos de concentração de poluentes atmosféricos do conjunto de dados MERRA-2. A nova versão denominada 3D-PSCF-CONC possibilita reduzir as superestimativas de regiões pouco poluídas, devido a inclusão de coeficientes de ponderação da concentração no cálculo de probabilidades. Para a aplicação dos métodos PSCF foi selecionada a Região Metropolitana de São Paulo (RMSP), localizada na região sudeste do Brasil. A RMSP é uma das metrópoles com maior potencial de poluição do ar do mundo, principalmente devido às emissões veiculares locais (Andrade et al., 2017).



## 1.1 OBJETIVO GERAL

Tornar o método 3D-PSCF sensível ao campo de concentração de poluentes atmosféricos.

## 1.2 OBJETIVOS ESPECÍFICOS

- Objetivo específico 1: Comparar o desempenho dos métodos PSCF na identificação das regiões que contribuem para as concentrações de poluentes atmosféricos do local receptor.
  - Hipótese: Os métodos 2D-PSCF e 3D-PSCF tendem a superestimar as contribuições das regiões por serem insensíveis aos campos de concentração de poluentes.
- Objetivo específico 2: Avaliar a consistência dos métodos PSCF utilizando dados de uso do solo, fluxo de poluentes, deposição úmida e inventário de emissões atmosféricas.
  - Hipótese: Regiões urbanizadas, onde há maior fluxo de poluentes tem maior probabilidade de contribuir para a poluição do ar.

Objetivo específico 3: Identificação de fontes de emissão de poluentes atmosféricos através de inventário de emissões atmosféricas.

- Hipótese: A incorporação dos campos de concentração de poluentes permite ao método 3D-PSCF-CONC identificar com maior eficiência as fontes de emissão de poluentes atmosféricos.

## 2 REFERENCIAL TEÓRICO

### 2.1 POLUIÇÃO ATMOSFÉRICA TRANSFRONTEIRIÇA

A migração da poluição do ar através de fronteiras estaduais, federais e/ou continentais pode representar uma ameaça à qualidade de vida e causar danos ambientais irreparáveis (Merrifield, 1990). Além das fontes locais, a qualidade do ar ambiente de cada cidade pode ser afetada por poluentes originados fora da cidade e transportados até o local receptor. Utilizando uma combinação de modelos globais, Zhang et al. (2017) descobriram que o Material Particulado fino (MP<sub>2,5</sub>) gerado a partir de atividades industriais na China está associado a mais de 3.100 mortes prematuras na Europa Ocidental e nos Estados Unidos da América (EUA). Jaffe et al. (2003) e Timonen et al. (2013) mostram que episódios de poluição atmosférica oriundas da Ásia influenciam a qualidade do ar nos EUA. Yamashita et al. (2016) apresentaram que há um efeito significativo da poluição do ar transfronteiriça sobre a deposição ácida na Costa do Mar do Japão, na região central do Japão.

Identificar as principais fontes de poluentes atmosféricos é essencial para gestão da qualidade do ar. A partir da caracterização da origem das emissões, é possível estimar o impacto ambiental dos principais poluidores e suas graves consequências (Longhurst et al., 2009; Isley e Taylor, 2018). A Região Metropolitana de São Paulo (RMSP) é uma megacidade com população de 21 milhões de habitantes (ONU, 2018; IBGE, 2017), e um exemplo clássico de degradação da qualidade do ar. A principal fonte de poluentes da RMSP é oriunda da atividade veicular, uma vez que possui frota de aproximadamente 7 milhões de veículos. De acordo com a Companhia Ambiental do Estado de São Paulo (CETESB, 2017), a frota veicular é responsável por aproximadamente 97% das emissões de CO. Os processos industriais e a queima de biomassa são, também, fontes responsáveis pelo aumento significativo de poluentes na atmosfera (CETESB, 2017).

Além das fontes locais, a qualidade do ar da RMSP é afetada pelo transporte de longo alcance de poluentes atmosféricos. Fontes externas como queima de cana-de-açúcar e incêndios florestais contribuem para episódios agudos de poluição na RMSP (França et al., 2014). O Brasil é o maior produtor mundial de cana-de-açúcar, com uma área de cultivo de 7,8 milhões de hectares, correspondendo a 2,3% das terras aráveis do país (<http://www.agricultura.gov.br/servicos-e-sistemas/sistemas/Sapcana>). A queima de resíduo de cana-de-açúcar durante o período de maio a outubro tem efeitos claros na qualidade do ar local da RMSP, o cultivo da cultura é a principal atividade agropecuária do estado de São Paulo. Souza et al., (2014)

identificaram que o transporte de longo alcance do material particulado durante a queima da cana-de-açúcar contribuiu para aumento nas concentrações de  $MP_{2,5}$  e de Material Particulado inalável ( $MP_{10}$ ) na cidade de São Paulo. Souto-Oliveira et al., (2016) identificaram a região sudeste (costeira) e noroeste (agrícola) do estado de São Paulo, como origens do transporte de carbono negro até a RMSP.

## 2.2 *POTENTIAL SOURCE CONTRIBUTION FUNCTION (PSCF)*

Para reduzir as concentrações de poluentes na atmosfera, as fontes de poluição devem ser identificadas, as estimativas de emissão realizadas e estratégias eficazes de gestão desenvolvidas (Seinfeld e Pandis, 1998). Para ajudar a localizar fontes de poluentes, modelos receptores que combinam a meteorologia com as concentrações medidas em um local de amostragem foram desenvolvidos. Estimativas do deslocamento da massa de ar até a chegada ao ponto receptor são utilizadas para criar um cenário dos locais onde os poluentes são emitidos (Hopke et al., 1994).

A modelagem de dispersão atmosférica descreve o transporte dos poluentes de um ponto emissor até os locais receptores (*Forward*), bem como o caminho inverso, de um local receptor até a sua possível origem (*Backward*). Estas trajetórias são calculadas por diversos modelos matemáticos, sendo o *Hybrid Single-Particle Lagrangian Integrated Trajectory Model* (HYSPLIT) o mais amplamente difundido (Poirot et al, 2001; Ara Begum et al, 2005; Draxler et al, 2005; Crawford et al, 2007; Gildemeister et al, 2007). Desenvolvido em conjunto pelo *Air Resources Laboratory* (ARL), pertencente ao *National Oceanic and Atmospheric Administration* (NOAA) e pelo Departamento Australiano de Meteorologia, o HYSPLIT tem como uma das aplicações mais comuns a análise de trajetória reversa para determinação da origem das massas de ar e estabelecimento das relações entre fonte de emissão e receptor (Stein et al., 2015).

Exemplos de modelos receptores híbridos que utilizam o HYSPLIT incluem *Trajectory Regression Analysis* (TRA) (Ashbaugh et al., 1985; Poirot e Wishinski, 1986), *Positive Matrix Factorization* (PMF), *Concentration-Weighted Trajectory* (CWT) (Seibert et al., 1994), *Residence Time Weighted Concentration* (RTWC) (Stohl e Andreas 1996) e *Potential Source Contribution Function* (PSCF) (Ashbaugh et al., 1985; Zeng e Hopke, 1989; Hopke et al., 1993).

Dentre os modelos citados, a PSCF tem sido a principal ferramenta utilizada para identificar os locais de origem e as vias de transporte de poluentes atmosféricos (Zeng e Hopke,

1989; Gao et al., 1994; Hsu et al., 2003; Hwang e Hopke, 2007). Gao et al. (1994) definem a PSCF como a probabilidade condicional de que uma parcela de ar com um nível de concentração de poluente acima de um valor de critério chegue a um local receptor após ter passado por uma área geográfica específica. O estudo de Cheng e Lin (2001) avaliaram o desempenho do método PSCF para eventos de fumaça na América Central usando dados observados em 1998 e descobriram que o método pode determinar com precisão os locais das fontes de fumaça e as vias de transporte dos poluentes. Dimitriou e Kassomenos (2016) identificaram, através da PSCF, que episódios agudos de  $MP_{10}$  em Avignon, sudeste da França, estavam fortemente associados à influência das massas de ar oriundas do norte da Itália e o sul da Alemanha. Já na região do Delta do Yangtze, na China, os resultados da PSCF indicaram que os aerossóis que impactam a qualidade do ar são oriundos principalmente de fontes locais (Fan et al., 2019). Nguyen et al. (2015) avaliaram as influências de plumas de fumaça transportadas por longo alcance nas concentrações de carbono orgânico medidas em Gosan, na Coreia do Sul, usando PSCF.

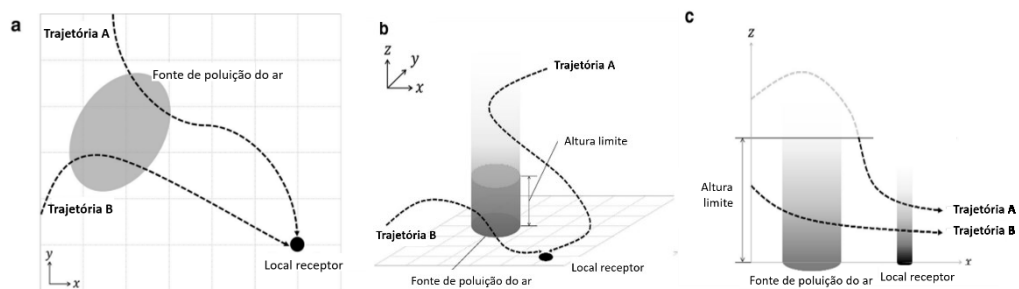
Algumas técnicas foram desenvolvidas e aplicadas para reduzir a incerteza na estimativa da PSCF, como por exemplo a função de ponderação (*Weighted Function*) (Hopke et al. 1995; Cheng et al. 1993b). No cálculo da PSCF, algumas células do domínio podem conter poucos pontos de trajetórias, ou seja, pequeno tempo de residência da massa de ar sobre a região. Caso estes pontos correspondam a trajetórias de eventos poluídos, a PSCF atribuirá alta influência deste local sobre a contribuição no receptor, porém este resultado é inconsistente e está vinculado a uma limitação do método (Hopke et al. 1995; Cheng et al. 1993b). Por exemplo, supondo que existem três células A, B e C, com o mesmo valor de PSCF igual 0,4, mas com diferentes números de pontos de trajetórias passando em cada uma 5, 30 e 900, respectivamente. Os valores de 0,4 nas células A, B e C terão diferentes desvios padrão de 0.548, 0.498 e 0.490, respectivamente. Para reduzir este tipo de incerteza, a *Weighted Function* é utilizada (Polissar et al., 2001; Heo et al., 2009).

Além da ponderação pelo tempo de residência, a PSCF foi atualizada também quanto ao deslocamento vertical da massa de ar, para reduzir as incertezas do método. A altura de injeção de poluentes na atmosfera pode variar de acordo com o tipo de fonte de emissão. Para incêndios florestais, por exemplo, In e Kim (2010) identificaram que a altura de injeção de poluentes pode alcançar um nível vertical de 5 km. Já as alturas de injeção em pequenas queimadas podem ser menores do que em incêndios florestais, devido ao menor fluxo de calor liberado. Para diferenciar os efeitos da altura de injeção de poluentes na atmosfera, Kim et al., (2016)

desenvolveram um novo método, a função tridimensional de contribuição potencial da fonte (3D-PSCF).

O 3D-PSCF adota uma altitude limite nos cálculos do PSCF. A análise convencional assume que as parcelas de ar que passam por uma célula contendo uma fonte de emissão são indiscriminadamente afetadas pela poluição do ar naquela célula. Por outro lado, a análise 3D-PSCF considera que a parcela de ar não é afetada pela poluição, caso passe pela célula a uma altitude superior à altura limite. A Figura 1 ilustra os conceitos da 2D-PSCF (a) e 3D-PSCF (b, c). De acordo com a análise 2D-PSCF, a trajetória **A** e a trajetória **B** têm a mesma possibilidade de serem contaminadas por uma fonte de poluentes atmosféricos. A análise 3D-PSCF distingue a trajetória **A** da trajetória **B**. A trajetória **A** pode não entrar em contato com os poluentes da região de origem por passar acima da altura limite. A parte da trajetória **A** que passa acima da altura limite (cinza) não é contada no cálculo 3D-PSCF. A Trajetória **B** e a parte preta da trajetória **A** são contadas como válidas porque passam abaixo da altura limite, e possuem maior probabilidade de contaminação.

**Figura 1.** Ilustração esquemática dos conceitos da 2D-PSCF (a) e 3D-PSCF (b, c) em função do deslocamento vertical das massas de ar (adaptada de Kim et al., 2016).



O método 3D-PSCF foi aplicado em Seul para avaliar a contribuição das regiões vizinhas sobre a qualidade do ar da capital sul-coreana. Kim et al., (2019) identificaram forte relação entre a queima de combustível de biomassa consumida na Coreia do Norte, Nordeste da China, Norte da China e Leste da China., e episódios agudos de poluição do ar em Seul. Dimitriou et al., (2021) identificaram que eventos de  $MP_{10}$  acima do limite diário, estabelecido pela União Europeia, em Atenas, na Grécia, estavam associados as intrusões diretas de poeira do Saara do Norte da África e especialmente da Líbia e do Egito. O transporte de  $MP_{10}$  ocorreu principalmente através de massas ar deslocando-se a altura inferior a 1 km de altitude.

Apesar da redução nas incertezas das estimativas geradas a partir da atualização da 3D-PSCF, os métodos seguem insensíveis a processos relevantes relacionados a dinâmica de poluentes na atmosfera. Uma das limitações dos métodos é que células do domínio onde há

intensa emissão de poluentes atmosféricos podem apresentar baixos valores de PSCF, isto porque os métodos baseiam suas estimativas apenas no tempo de residência da massa de ar (Cheng et al., 2015; Zachary et al., 2018). Os métodos ignoram a possibilidade de incremento e/ou diluição das concentrações da massa de ar ao se deslocar por diferentes regiões em direção ao local receptor. Como resultado, pode ser difícil distinguir fontes moderadas de fontes intensas, por isso é necessário introduzir à análise a assimilação dos campos de concentração de poluentes atmosféricos.

### 2.3 ASSIMILAÇÃO DE REANÁLISE ATMOSFÉRICA APLICADA A POLUIÇÃO DO AR

Redes de monitoramento da qualidade do ar são peças essenciais para gestão ambiental, pois cumprem o papel preponderante de mensurar a concentração de poluentes atmosféricos (WHO, 2016). Atualmente, existem redes de monitoramento da qualidade do ar em todo o mundo, mas geralmente oferecem cobertura geográfica esparsa. Por este motivo, muitos estudos surgiram com o objetivo de estimar as concentrações de poluentes a partir de medições de satélite (Molod et al., 2015; Reichle et al., 2017). O sensoriamento remoto tem a vantagem de fornecer cobertura global e observações de longo prazo em regiões de difícil acesso e implementação de redes de monitoramento.

Dentre os satélites utilizados para análises atmosféricas, a Análise Retrospectiva da Era Moderna para Pesquisa e Aplicações, versão 2 (*The Modern-Era Retrospective Analysis for Research and Applications - MERRA-2*) vem sendo amplamente empregado para auxiliar na compreensão das dinâmicas e impactos dos poluentes atmosféricos (Bali et al., 2017; Gelaro et al., 2017; Qin et al., 2019; Sun et al., 2019). Coordenado pelo Escritório de Modelagem e Assimilação Global (GMAO) da Administração Nacional de Aeronáutica e Espaço dos Estados Unidos (NASA), o MERRA-2 fornece produtos de reanálise, ou seja, após estimativa prévia do satélite os dados são submetidos a um reprocessamento através de observações meteorológicas (vento, temperatura, umidade, etc.), normalmente abrangendo um segmento estendido do registro de dados históricos. Além disso o sistema de assimilação utiliza uma grade regular com resolução de  $0,625^\circ$  longitude  $\times$   $0,5^\circ$  latitude para interpolação dos dados de saída.

O MERRA-2 utiliza o modelo do sistema de assimilação de dados do sistema de observação da Terra Goddard (GEOS-5) versão 5.12.4, que é capaz de usar tipos de dados de instrumentos de radiação infravermelho hiperespectral de sondas de micro-ondas (Bosilovich et al. 2016) e fornecer os melhores dados em tempo real. O núcleo principal do sistema de

reanálise MERRA-2 é o modelo atmosférico GEOS-5 (Rienecker et al., 2008; Buchard et al., 2015) e o esquema de análise de interpolação estatística de ponto de grade (Kleist et al., 2009). GEOS-5 usou as observações de aerossol com correção de polarização de sensores de satélite, incluindo MODIS, Multi-angle Imaging SpectroRadiometer e Advance Very High Resolution Radiometer (Bosilovich et al., 2015). Os diferentes tipos de aerossóis (BC, poeira, sulfato e sal marinho) são simulados usando o módulo Goddard Chemistry Aerosol Radiation and Transport (GOCART) no GEOS-5 (Buchard et al., 2016). O modelo GEOS-5 fornece os dados de 1980 até a apresentação em grade horária a distribuição de 72 níveis verticais com a camada superior em 0,01 hPa.

Com base em medições ao nível do solo e nos dados de reanálise do MERRA-2, Qin et al., (2019) investigaram a variação espaço-temporal do carbono negro (BC) em Pequim, na China. A partir da distribuição espacial da concentração do carbono negro estimado pelo MERRA-2, das trajetórias reversas de massa de ar e dos inventários de emissões regionais foi possível identificar as emissões de transporte em Pequim e emissões residenciais em regiões vizinhas como fontes importantes de BC em Pequim.

## 2.4 INVENTÁRIO DE FONTES DE EMISSÃO

Atualmente, as ações de melhoria da qualidade do ar estão crescendo rapidamente e são quase obrigatórias, especialmente em áreas urbanas e industrializadas onde a poluição do ar se tornou uma consequência indesejável das atividades humanas (Gurjar et al. 2018; Andrade et al. 2017; Gallagher et al. 2015, Landrigan et al. 2018). No que diz respeito ao uso de abordagens numéricas para análise do clima e da qualidade do ar, o uso de inventários de emissões é obrigatório. Eles são dados fundamentais para que os modelos possam simular o presente e prever os cenários de emissão futuros, além de auxiliar na identificação de importantes fontes emissoras. A disponibilidade de dados de alta qualidade é um passo importante para o sucesso das ações de controle da poluição (Ohara et al. 2007; D'Avignon et al. 2010).

A quantificação das emissões atmosféricas é uma etapa fundamental para observar a variabilidade e tendências da composição química da atmosfera (GEIA, <http://www.geiacenter.org/>). Para lidar com aspectos regionais e globais, os inventários de emissões acoplados a modelos de trajetórias e medições de poluição do ar podem ser ferramentas úteis no fornecimento de informações sobre impactos locais e regionais na qualidade do ar. O Banco de Dados de Emissões para Pesquisa Atmosférica Global (*Emissions Database for Global Atmospheric Research* - EDGAR) é um inventário de emissões globais

que fornece emissões de gases de efeito estufa e poluentes atmosféricos por país e em uma grade espacial de  $0,1^\circ \times 0,1^\circ$  (Janssens-Maenhout et al. 2019; Crippa et al. 2018). A versão mais recente do EDGAR, 5.0, cobre uma grande quantidade de informações, possuindo 45 anos de dados (1970 - 2015), 26 setores de emissão, 226 países e 9 substâncias. Além disso, as emissões são representadas como séries temporais e mapas de grade de emissão.

A poluição por  $MP_{10}$  na região de Hauts-de-France, no norte da França, foi investigada utilizando a análise de trajetórias combinada aos dados de emissão europeus do EDGAR (Potier et al., 2019). Os resultados mostraram que alguns países afastados ao leste francês (Polônia, Eslováquia, República Tcheca) poderiam ter um impacto forte, mas raro, nos níveis de  $MP_{10}$  na região de estudo, enquanto alguns países europeus mais próximos (Bélgica, Holanda, Alemanha) apresentaram impacto frequente e predominante. O Sudeste Asiático é responsável por aproximadamente 15% das florestas tropicais do mundo e também é suscetível à queima de biomassa (Stibig et al., 2014). Yin et al., (2019) investigaram a distribuição espaço-temporal da queima de biomassa no sudeste da Ásia continental e seus efeitos na qualidade do ar entre 2001 e 2016. Utilizando a emissão mensal de  $MP_{2,5}$  estimado pelo EDGAR como suporte para a análise PMF, os autores comprovam que a queima de biomassa é a causa dominante do comprometimento da qualidade do ar no sudeste da Ásia continental.



### 3 MATERIAL E MÉTODOS

#### 3.1 ÁREAS DE ESTUDO

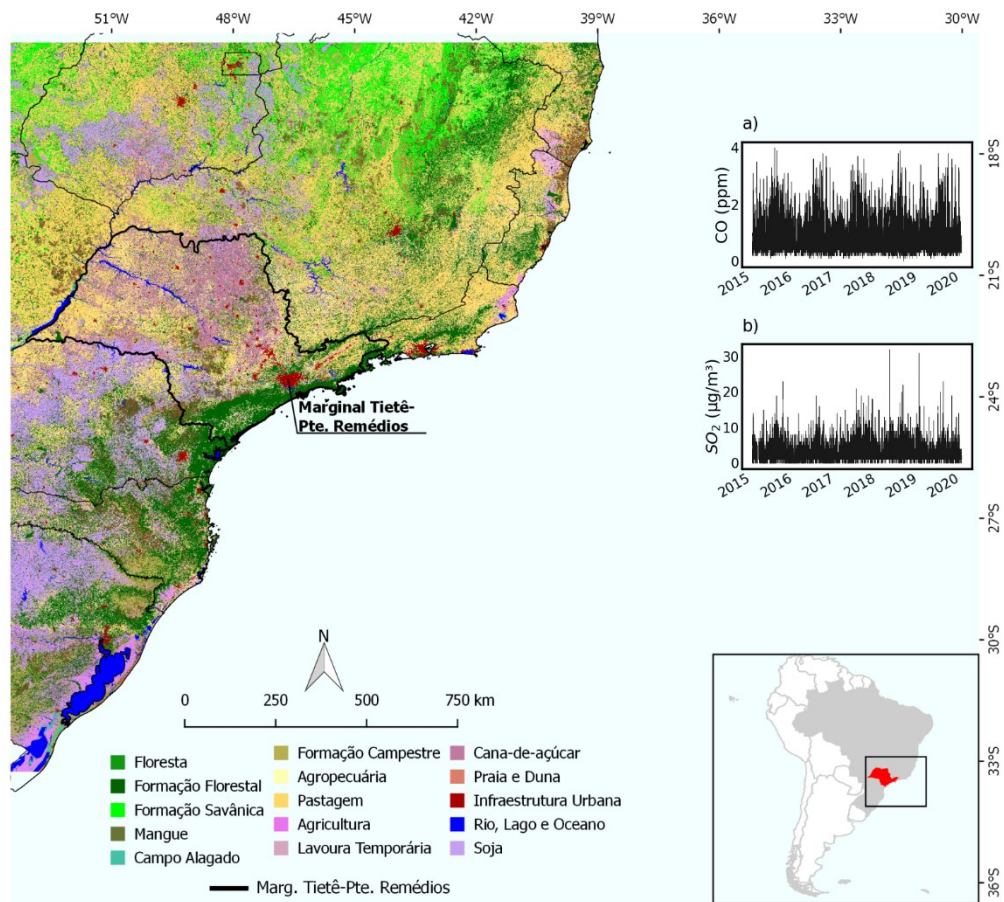
Neste trabalho, foi utilizada a Região Metropolitana de São Paulo (RMSP) como área principal para comparação da consistência dos métodos PSCF. Localizada na região sudeste do Brasil, a RMSP é considerada a região econômica mais importante da América do Sul. Com aproximadamente 21 milhões de habitantes, a metrópole brasileira é a quinta mais populosa do mundo, e a primeira da América Latina (UN, 2016). Segundo o Instituto Brasileiro de Geografia e Estatística (IBGE), a RMSP representou 20% do Produto Interno Bruto (PIB) brasileiro em 2017 (IBGE, 2017). A organização espacial da RMSP, seja com relação à circulação das fontes móveis ou a disposição das fontes estacionárias, juntamente com a topografia e as condições climáticas da região são os fatores determinantes da qualidade do ar (CETESB, 2004). A área urbana da região está localizada 700–900 metros acima do nível do mar, com cumes que se elevam até 1200 m. Sua topografia geral é bastante complexa e o fluxo de ar é fortemente influenciado pelas circulações locais. Aliado a isso, está também o próprio processo de desenvolvimento urbano desordenado ocorrido na região a partir dos anos 50, processo esse que aumentou o número de fontes de emissão, levando a uma situação crítica em relação a qualidade do ar (Carmo, 1995; CETESB, 2005, Andrade et al., 2017).

Com relação as fontes de emissão atmosférica, a RMSP é caracterizada pelas altas emissões locais de fontes móveis. Com frota de cerca de 7 milhões de veículos movidos a gasolina, etanol hidratado e/ou diesel, o tráfego rodoviário é responsável por 97,5% das emissões de CO, 79% das emissões de hidrocarbonetos (HC), 67,5% das emissões de óxido de nitrogênio (NO<sub>x</sub>), 22% das emissões de SO<sub>2</sub> e 40% das emissões de MP<sub>10</sub> na RMSP (CETESB, 2015). Ainda, Andrade et al., (2017) reportaram sobre as influências externas nas concentrações locais do RMSP, principalmente associadas a queima de biomassa no período de maio a agosto na região norte e nordeste do estado de São Paulo; e as emissões industriais originadas na região de Cubatão/SP.

Os gases CO e SO<sub>2</sub> foram utilizados na análise dos métodos PSCF por serem poluentes primários, ou seja, são emitidos diretamente pelas fontes de emissão. Suas concentrações no ambiente não dependem de reações químicas, o que possibilita identificar com maior precisão as fontes de emissão. O CO é resultante da queima incompleta de combustíveis de origem orgânica (combustíveis fósseis, biomassa, etc.), e em geral é encontrado em maiores concentrações nas cidades, emitido principalmente por veículos automotores. Já as principais fontes de emissão de SO<sub>2</sub> são a queima de combustíveis fósseis e atividades industriais (refino

de petróleo, metalurgia, etc.) (Seinfeld e Pandis, 1998). A Figura 2 apresenta a localização das áreas de estudo, e as concentrações de CO e SO<sub>2</sub> durante período de 2015 a 2019 na estação de monitoramento Marginal Tietê-Ponte Remédios (RMSP) (CETESB, 2019a).

**Figura 2.** Localização da estação de monitoramento Marginal Tietê-Pte. Remédios. Concentrações de CO (a) e SO<sub>2</sub> (b) (CETESB, 2019a) nas estações de monitoramento. Uso do solo proveniente do projeto MAPBIOMAS (2019).



O efeito do transporte transfronteiriço de poluentes na qualidade do ar de regiões receptoras pode ser reduzido caso as emissões locais apresentem elevadas magnitudes, como na RMSP. Ainda assim, identificar o impacto de uma massa de ar externa em um evento de poluição pode ser mais uma ferramenta para gestão da qualidade do ar.

### 3.2 MÉTODOS 2D-PSCF E 3D-PSCF

O método 2D-PSCF calcula um campo de probabilidade que pode ser associado às potenciais fontes de poluição do ar que contribuem para um local receptor específico. O método é

geralmente usado para investigar os eventos de qualidade do ar acima de um determinado limite no local do receptor (Lucey et al., 2001; Hsu; Holsen; Hopke, 2003; Zhao et al., 2019). Para minimizar as incertezas e permitir uma comparação adequada entre os métodos é aplicado uma função de ponderação  $W_{i,j}$  multiplicada aos valores PSCF (Equação 2).

$$W_{i,j} = \begin{cases} 1.0 & (3n_{ave} < n_{ij}) \\ 0.7 & (1.5n_{ave} < n_{ij} < 3n_{ave}) \\ 0.4 & (n_{ave} < n_{ij} < 1.5n_{ave}) \\ 0.2 & (n_{ij} < n_{ave}) \end{cases} \quad (2)$$

onde  $n_{ave}$  é o número médio de pontos de trajetória entre as células  $i,j$  não vazias;  $i$  e  $j$  correspondem a latitude e longitude, respectivamente, do ponto.

A 2D-PSCF é uma probabilidade condicional de que uma parcela de ar que passou pela  $i, j^a$  célula teve uma alta concentração ao chegar ao ponto final da trajetória (Equação 3).

$$P_{i,j}^{2D} = \frac{m_{i,j}}{n_{i,j}} W_{i,j} \quad (3)$$

onde  $n_{i,j}$  é o número total de pontos de trajetórias de massa de ar que passaram pela  $i, j^a$  célula, e  $m_{i,j}$  é o número de pontos de trajetória na  $i, j^a$  célula associados aos eventos em que a concentração no local do receptor excedeu o valor de critério.

A 2D-PSCF possui algumas lacunas devido à incerteza associada ao pequeno número de pontos de trajetória. Além disso, o método não consegue distinguir trajetórias que passam em altitudes diferentes. Por este motivo, Kim et al. (2016) desenvolveram o 3D-PSCF para distinguir os tipos de fontes de acordo com a altura de injeção de poluentes na atmosfera, enquanto Zeng e Hopke (1989) propõem uma função de ponderação para reduzir a incerteza associada ao pequeno número de trajetórias. A Equação 4 apresenta o método 3D-PSCF.

$$P_{i,j}^{3D} = \frac{m'_{i,j}}{n'_{i,j}} W_{i,j} \quad (4)$$

onde  $n_{i,j}'$  é o número total de pontos trajetórias que caem na  $i, j^a$  célula abaixo da altura limite, e  $n_{i,j}'$  é o número de pontos de trajetórias que passam pela  $i, j^a$  célula abaixo da altura limiar

para a qual as trajetórias correspondentes chegam ao local do receptor com concentração superior ao valor crítico pré-especificado.

### 3.3 DESENVOLVIMENTO DO MÉTODO 3D-PSCF-CONC

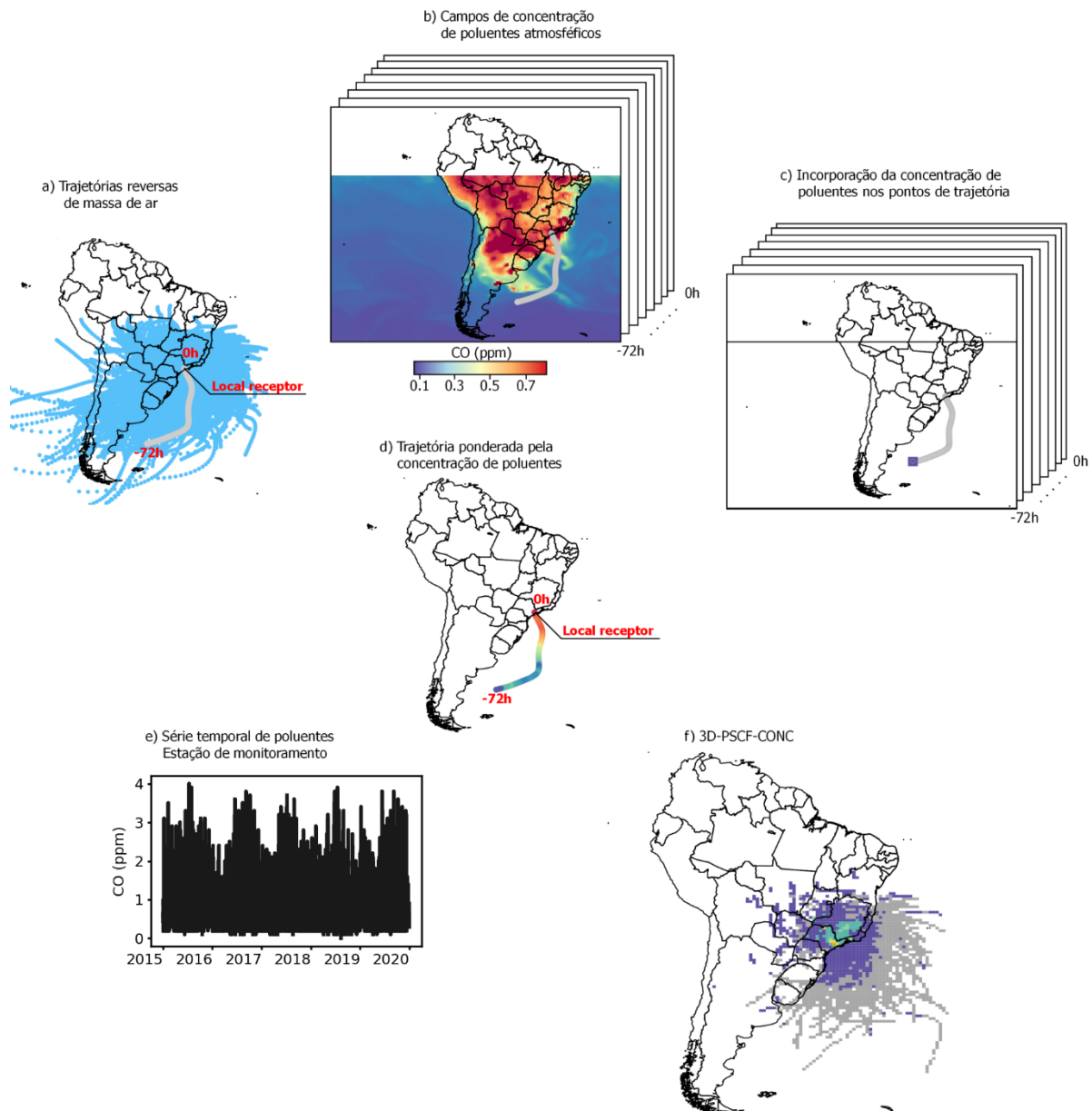
O 3D-PSCF-CONC incorpora uma função de ponderação no método 3D-PSCF com base no campo de concentração de poluentes de reanálises atmosféricas. O 3D-PSCF-CONC é calculado por meio da Equação 5. O Apêndice A apresenta a descrição do código implementado em Python.

$$P_{i,j}^{3D-CONC} = \frac{m'_{i,j}}{n'_{i,j}} W_{i,j} \frac{\bar{C}_{i,j}}{\bar{C}_{rs}} \quad (5)$$

onde,  $\bar{C}_{ij}$  e  $m'_{ij}$  são a concentração média e o tempo de residência total na  $i, j^a$  célula, quando a concentração no local receptor excedeu o valor do critério.  $\bar{C}_{rs}$  e  $n'_{ij}$  são, respectivamente, a concentração média na célula localizada no local receptor medida pela reanálise atmosférica e o tempo total de residência na célula  $i, j^a$  para o conjunto total de trajetórias. O valor máximo de  $\frac{\bar{C}_{ij}}{\bar{C}_{rs}}$  é limitado por 1, uma vez que a componente faz parte de uma função de probabilidade condicional.

A Figura 3 ilustra o conceito do 3D-PSCF-CONC, onde a) representa as trajetórias reversas que convergem para RMSF utilizadas no estudo, uma trajetória reversa está destacada em branco para exemplificar as etapas do método. b) cada ponto de trajetória está associado, no tempo, a uma grade de concentração de reanálise atmosférica do MERRA-2. Em c) é apresentada a célula da grade em que está localizado o ponto de trajetória, ao qual será atribuído a concentração. Após todos os 73 pontos terem suas concentrações definidas a trajetória estará apta a ser utilizada no cálculo da 3D-PSCF-CONC. O valor limite utilizado para definição dos eventos de poluição é obtido a partir da série temporal das concentrações do poluente na estação de monitoramento e), e então a 3D-PSCF-CONC é calculada.

**Figura 3.** Ilustração esquemática do conceito do método 3D-PSCF-CONC. a) Representa as trajetórias reversas que convergem para RMSP utilizadas no estudo, uma trajetória reversa está destacada em branco para exemplificar as etapas do método. b) cada ponto de trajetória está associado, no tempo, a uma grade de concentração de reanálise atmosférica do MERRA-2. Em c) é apresentada a célula da grade em que está localizado o ponto de trajetória, ao qual será atribuído a concentração. e) Os valores da 3D-PSCF-CONC.



O conceito de 3D-PSCF-CONC é fisicamente explicado pelos três elementos da função de probabilidade condicional na Equação 5. O primeiro elemento de PSCF calcula a probabilidade de que uma massa de ar chegue ao receptor após residir em uma célula geográfica específica. O segundo elemento corrige a incerteza associada ao pequeno número de pontos de trajetória, que uma célula específica teria. O terceiro elemento representa a probabilidade de que uma

célula aumente a concentração do local receptor. Se a concentração de uma célula pela qual a parcela de ar passou for menor do que o receptor, essa área provavelmente contribui para diluir a poluição ao invés de aumentar a concentração no local do receptor. Neste estudo, não definimos uma altura limite fixa nos métodos 3D-PSCF e 3D-PSCF-CONC, a altura limite para aplicação dos métodos foi definida a partir da camada limite, que é a região da atmosfera onde os poluentes são emitidos e os processos de transporte e difusão ocorrem (Fan et al., 2019).

Para avaliar a contribuição potencial de áreas menos poluídas, uma análise de fluxo de concentração foi utilizada. Neste trabalho, o fluxo de concentração ( $\text{Fluxo}_{i,j}$  (%)) foi calculado nas células para cada trajetória por onde a massa de ar se deslocou, do ponto original ao receptor, conforme a Equação 6. Para avaliar a variabilidade espacial, uma grade de fluxo foi calculada através da média dos fluxos em cada célula. Este procedimento é utilizado para avaliar a proporção de aumento/diminuição da concentração do poluente atmosférico ao longo da trajetória da origem ao receptor, em relação à concentração do receptor.

Também foi analisado a associação entre o PSCF, o fluxo de CO e a precipitação média oriundo do conjunto de dados de reanálise do MERRA-2.

$$\text{Fluxo}_{i,j} (\%) = \frac{\bar{C}_{i,j} - \bar{C}_{i,j-1}}{\bar{C}_{rs}} \quad (6)$$

onde  $\bar{C}_{i,j}$  é a concentração média na  $i, j^a$  célula,  $\bar{C}_{i,j-1}$  é a concentração média anterior a  $i, j^a$  célula e  $\bar{C}_{rs}$  é a concentração da célula do local receptor.

### 3.4 DADOS DE ENTRADA DOS MÉTODOS

#### 3.4.1 Monitoramento da qualidade do ar

A Agência Ambiental do Estado de São Paulo - CETESB, possui 62 estações de monitoramento da qualidade do ar no Estado de São Paulo (CETESB, 2019b). A estação de monitoramento da Marginal Tietê-Ponte Remédios (-23.518578 S, -46.743148 W) foi selecionada na RMSP devido à disponibilidade e qualidade dos dados (completude dos dados superior a 95%). As concentrações horárias de CO e SO<sub>2</sub> de 2015 a 2019 foram obtidas a partir da plataforma QUALAR. O QUALAR é uma plataforma on-line da CETESB, que possibilita a consulta dos dados de poluentes coletados por suas estações, tanto da rede automática como da rede manual, através do endereço eletrônico (<http://qualar.cetesb.sp.gov.br/qualar/>). A partir da série temporal da concentração de CO (Figura 2.a) e SO<sub>2</sub> (Figura 2.b) no local do receptor,

foram determinados os valores limites de 0,7 ppm e 2,49  $\mu\text{g}\cdot\text{m}^{-3}$  para o cálculo da PSCF. Estes valores representam a concentração média diária de CO e SO<sub>2</sub> no período de 5 anos. Logo, episódios críticos de poluição do ar foram definidos como os dias em que a concentração excede a concentração média diária do poluente.

#### 3.4.2 Dados de trajetórias reversas de massa de ar

As trajetórias reversas de massa de ar (backward) foram geradas utilizando o Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT) versão 4, desenvolvido pela US National Oceanic and Air Administration (NOAA) (Draxler; Hess, 1998; Stein et al., 2015). A partir da plataforma online de simulações do NOAA foram geradas trajetórias para cada dia de 2015 a 2019. Com duração de 72 horas e resolução temporal de 1 hora, foram geradas, no total, 133.298 pontos de trajetória.

Dados meteorológicos do Sistema Global de Assimilação de Dados (GDAS) com resolução espacial de  $1,0^\circ \times 1,0^\circ$  do site da NOAA (<http://ready.arl.noaa.gov/gdas1.php>) são empregados ao HYSPLIT. Definimos a altitude mínima final como 500 m. Conforme sugerido por Hafner e Hites (2003) e Begum et al. (2005), acima de 500 m o vento está bem representado na camada limite. Para 3D-PSCF e 3D-PSCF-CONC, a altura limite para aplicação dos métodos foi definida a partir da camada limite, que é a região da atmosfera onde os poluentes são emitidos e os processos de transporte e difusão ocorrem (Fan et al., 2019). Portanto, foi considerado apenas os pontos de trajetórias situados abaixo da camada limite. Foi possível atualizar o limite ao longo do deslocamento da trajetória, pois o HYSPLIT calcula a altura da camada limite em cada intervalo de tempo.

#### 3.4.3 Campos de concentração de poluentes atmosféricos

Campos de concentração de superfície horária de CO (ppb) (M2T1NXCHM V5.12.4) e SO<sub>2</sub> ( $\mu\text{g}\cdot\text{m}^{-3}$ ) (M2T1NXAER V5.12.4) estimados pela Modern-Era Retrospective Analysis for Research and Applications - MERRA-2 foram utilizados neste estudo para assimilação da concentração apenas no método 3D-PSCF-CONC (Gelaro et al., 2017; Randles et al., 2017). Os dados foram obtidos por meio da plataforma Simple Subset Wizard (SSW), desenvolvida e mantida pela Administração Nacional de Aeronáutica e Espaço (*National Aeronautics and Space Administration* - NASA) (Acker et al., 2014).

O domínio do estudo foi selecionado de acordo com o alcance das trajetórias, sendo estas as coordenadas (Figura 2.b): -65.8, -120.7, -6.7, -18.8. Devido a capacidade de armazenar dados multidimensionais (multivariáveis), foi selecionado o formato Network Common Data Form (netCDF) para a execução deste trabalho.

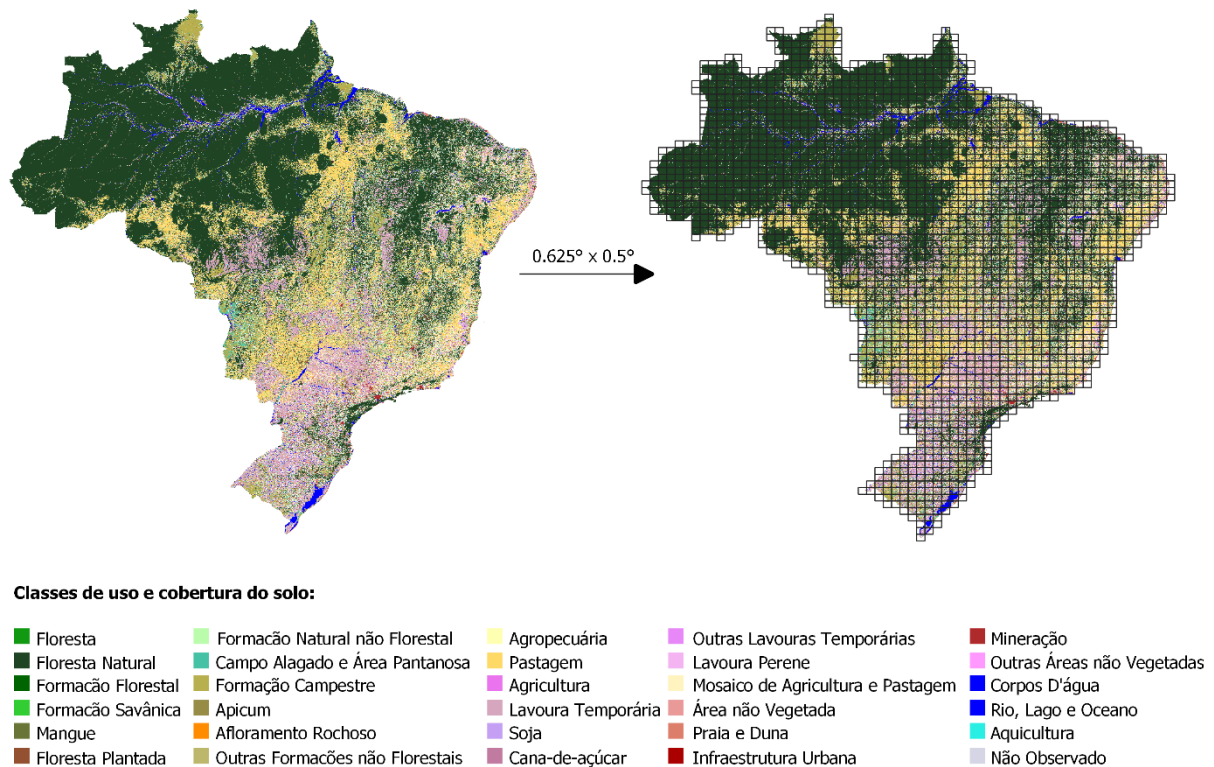
### 3.5 ANÁLISES DE CONSISTÊNCIA DOS MÉTODOS

#### 3.5.1 Dados de uso e cobertura do solo

Uma análise da distribuição espacial entre os valores da PSCF e seus respectivos usos do solo foi realizada para verificar a consistência dos métodos. Espera-se que as áreas urbanas e outras regiões emissoras em potencial forneçam os maiores valores de PSCF. Foram utilizados dados do Projeto Brasileiro de Mapeamento Anual de Uso e Cobertura do Solo (MapBiomias) para identificar a composição de pastagens, áreas urbanas, corpos d'água e outras feições superficiais no domínio PSCF. O MapBiomias usa imagens de satélite LandSat séries 5, 7 e 8, com resolução espacial de 30 metros para o Sistema de Estimativa de Emissões de Gases de Efeito Estufa do Observatório do Clima. Os dados foram obtidos por meio da ferramenta MapBiomias (versão 4.0), que opera a partir da plataforma Google Earth Engine e está disponível online, gratuitamente, em <http://mapbiomas.org> (MAPBIOMAS, 2019). Os mosaicos das classes presentes no MapBiomias estão em formato Geographical Tagged Image Format File (GeoTIFF). Com auxílio da biblioteca rasterio do software Python, foi possível extrair informações de cada pixel das imagens do produto para locais especificados, neste caso, a grade utilizada para PSCF (Figura 4).



**Figura 4.** Classes de uso e cobertura do solo do projeto MAPBIOMAS (esquerda). Grade com resolução de  $0.5^\circ$  x  $0.625^\circ$ , que coincide com a resolução das grades dos métodos PSCF empregados neste estudo (direita).



Após o recorte da área de interesse, o mesmo domínio utilizado no MERRA-2, foram avaliadas as classes com registros de pontos de trajetórias entre os anos de 2015 a 2019. Para esta análise foi utilizada em Python função *sv.contains*, da biblioteca *shapely.vectorized*, com a qual foi possível identificar o número de pontos de trajetórias sobre cada classe. Foram constatadas treze classes de uso da terra definidas pelo MapBiomias para o domínio de estudo, correspondente as áreas de formação florestal, formação savanica, mangue, campos, pastagem, agricultura, praias e dunas, infraestrutura urbana e corpos d'água.

### 3.5.2 Dados de chuva

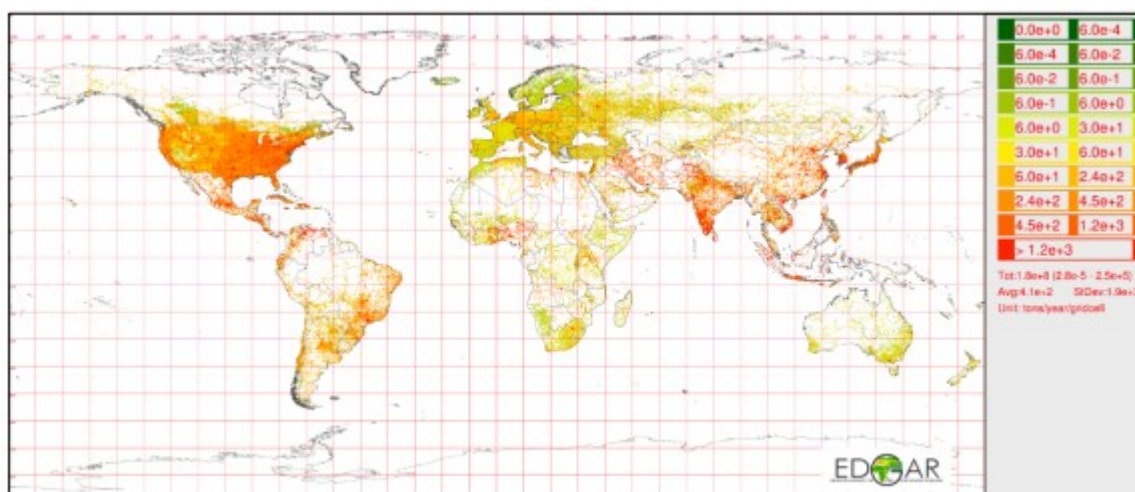
Após a emissão, um dos principais mecanismos que controlam as concentrações de poluentes na atmosfera é a deposição úmida (Garrett et al., 2006). Dados horários de precipitação total ( $\text{mm.h}^{-1}$ ) (*Total Precipitable Liquid Water - TQL*) (M2T1NXSLV V5.12.4) do MERRA-2 foram utilizados para dar suporte à validação dos métodos. Foi realizada a correlação de Spearman (Wilks, 2019). Entre a concentração média de TQL, no período de 2015 a 2019, e as contribuições estimadas pelos métodos PSCF para cada célula das grades.

Além disso, os dados foram utilizados como ponderação na análise de fluxo de poluentes e PSCF.

### 3.5.3 Dados de fontes de emissão de poluentes atmosféricos

Para avaliar a consistência dos resultados obtidos pelos métodos PSCF, a variabilidade espacial das contribuições foi comparada aos dados de fontes de emissão de poluentes do EDGAR. O desenvolvimento atual do EDGAR é um projeto conjunto do Centro de Pesquisa Conjunta (JRC) da Comissão Europeia e da Agência de Avaliação Ambiental da Holanda (PBL) (CRIPPA et al., 2020). O inventário compila emissões antropogênicas e tendências baseadas em estatísticas internacionais e fatores de emissão, mapeando dados históricos sobre as emissões, para os seguintes poluentes: Dióxido de enxofre (SO<sub>2</sub>), Óxidos de nitrogênio (NO<sub>x</sub>), Metano (CH<sub>4</sub>), CO, compostos orgânicos voláteis não-metano (NMVOCs), Amônia (NH<sub>3</sub>), MP<sub>10</sub>, MP<sub>2.5</sub>, BC. A versão v5.0 do EDGAR do inventário possui metodologia consistente para o cálculo de séries temporais de emissão, contendo dados de emissão de poluentes atmosféricos de 1970 até 2015, com resolução espacial de 0.1°×0.1° (THUNIS et al., 2021). A Figura 5 corresponde a um mapa anual de emissões de CO no setor de transporte rodoviário para o ano de 2015, representando um exemplo de *gridmap* do EDGAR.

**Figura 5.** Mapa da emissão de CO pelo setor de transporte no ano de 2015 (EDGAR, 2017).



A emissão anual de superfície do transporte rodoviário sem ressuspensão de solo, queima de resíduos agrícolas, produção de ferro e aço, combustão para manufatura, refinarias de petróleo e indústria de transformação para 2015 foi usada para subsidiar a identificação das fontes de emissão pelos métodos PSCF e avaliar a consistência dos métodos.

#### 3.5.4 Dados de frota veicular e consumo de combustível

Dados da frota veicular circulante no estado de São Paulo, estimados pela CETESB (<https://cetesb.sp.gov.br/veicular/relatorios-e-publicacoes/>, 2019 - Frota Circulante), foram utilizados para dar suporte à identificação das fontes de emissão localizadas nas regiões indicadas pelos métodos PSCF. O levantamento contabiliza o número de veículos circulantes em cada município do estado São Paulo de 1979 a 2019, de acordo com a categoria e tipo de combustível utilizados. Os dados de frota veicular foram correlacionados as concentrações de poluentes do MERRA-2 nas células das grades para avaliar a influência dos veículos na poluição do ar. O período utilizado nesta análise foi de 1980, ano que inicia a série de dados do MERRA-2, a 2019. Este intervalo foi selecionado para aumentar a consistência da análise, uma vez que os dados de frota veicular são anuais e aplicá-los ao período de 2015 a 2019 reduziria a confiabilidade nos resultados devido ao pequeno tamanho amostral.

Para avaliar a influência da queima de diferentes tipos de combustíveis fósseis sobre a qualidade do ar, dados mensais da venda de combustíveis na cidade de São Paulo de janeiro de 2015 a dezembro de 2019, fornecidas pela Agência Nacional de Petróleo (ANP, <https://www.gov.br/anp/pt-br>), foram integrados a análise. Assim como para frota veicular, os dados de volume de combustível comercializados pelas distribuidoras foram correlacionados com a concentração de poluentes do MERRA-2.

#### 3.5.5 Dados de potência radiativa do fogo

Para identificação da queima de biomassa como uma possível fonte de contaminação do ar, foram utilizados dados de potência radiativa do fogo (*fire radiative power* - FRP) do Sistema de Informação de Incêndio para Gerenciamento de Recursos (*Fire Information for Resource Management System* - FIRMS) (<https://firms.modaps.eosdis.nasa.gov/>), da NASA. Através do instrumento de espectrorradiômetro de imagem de resolução moderada (*Moderate Resolution Imaging Spectroradiometer* - MODIS), o FIRMS fornece os dados de caracterização de incêndios ativos em tempo real, disponíveis em resolução espacial de 1km dentro de 3 horas de observação. A bordo dos satélites Aqua e Terra, o MODIS é um instrumento de coleta de dados que explora a atmosfera, o oceano e a terra (Shen; Cao, 2020).

Neste estudo, os hotspots do MODIS de 01 de janeiro de 2015 a 31 de dezembro de 2019 foram agregados em médias mensais de FRP nas regiões de análise. Isto para caracterizar o uso sazonal do fogo em culturas durante a etapa de colheita (IEA, 2020).

## 4 RESULTADOS

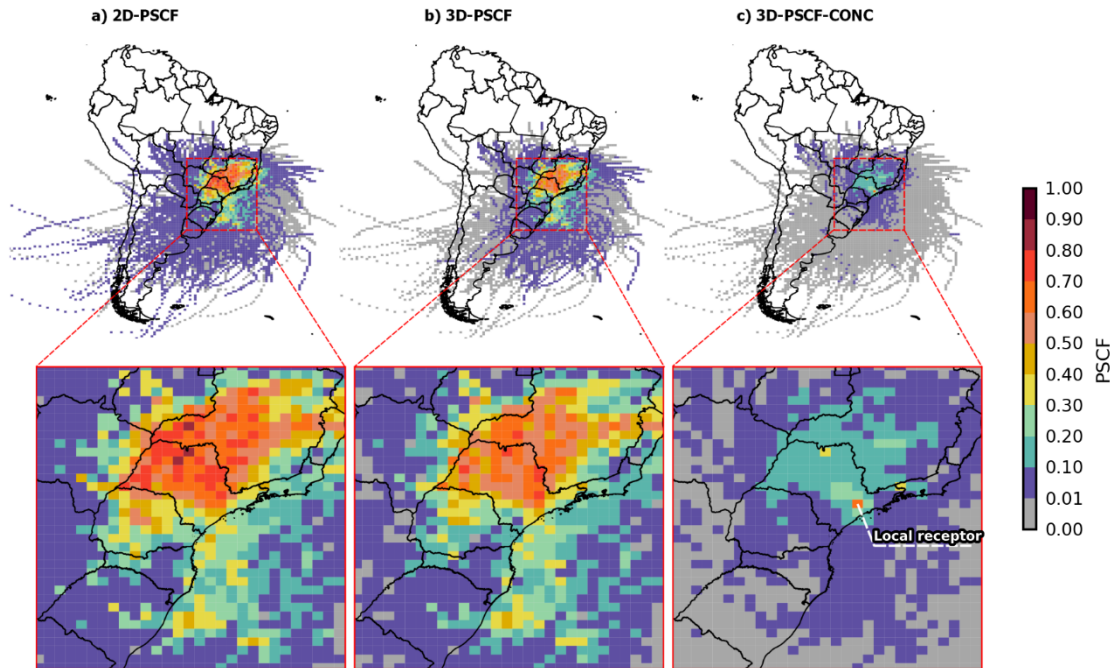
### 4.1 COMPARAÇÃO ENTRE OS MÉTODOS PSCF

Dentre os métodos avaliados, o 2D-PSCF apresentou os maiores valores de probabilidade de contribuição (Figura 6). Os resultados do 2D-PSCF indicaram que diversas regiões dentro do domínio podem contribuir significativamente para a poluição do ar no local receptor (Figura 6.a). A superestimativa gerada pelo 2D-PSCF é esperada, uma vez que não leva em conta a altura de injeção e a concentração de poluentes na atmosfera enquanto a trajetória se move em direção ao receptor. Além disso, o 2D-PSCF não destacou a contribuição local da megacidade de São Paulo, que fica na mesma área do local receptor e é conhecida pelos problemas de qualidade do ar (Andrade et al., 2017).

A distribuição espacial das estimativas 3D-PSCF apresentou semelhanças com a 2D-PSCF. No entanto, a inclusão da análise tridimensional do método 3D-PSCF reduziu a contribuição do transporte de longo alcance, especialmente a noroeste do domínio (Figura 6.b). A 3D-PSCF filtrou as trajetórias abaixo da altura de injeção de poluentes na atmosfera como potencialmente poluídas. Embora o 3D-PSCF tenha ajustado o cálculo de probabilidade, o método ainda identificou regiões distantes do local do receptor como os principais contribuintes.

A 3D-PSCF-CONC alcançou resultados diferentes em comparação as versões anteriores do PSCF (Figura 6.c). O método destacou a RMSP com o principal potencial de contribuição, mostrando consistência com o cenário real. Além disso, O 3D-PSCF-CONC também identificou as fontes regionais a noroeste e nordeste do domínio.

**Figura 6.** Comparação dos métodos PSCF: a) 2D-PSCF, b) 3D-PSCF, c) 3D-PSCF-CONC, e variabilidade dos valores de PSCF em cada uso do solo do domínio por onde as trajetórias deslocaram.

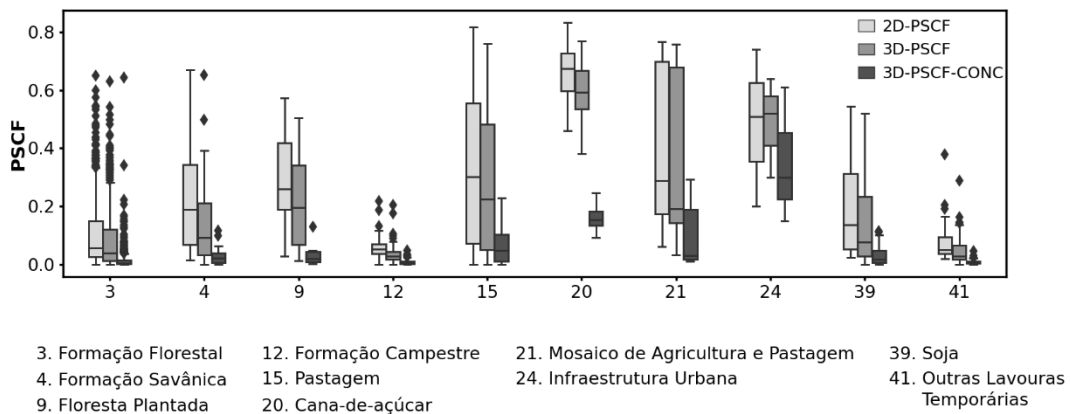


## 4.2 ANÁLISE DE CONSISTÊNCIA DOS MÉTODOS PSCF

### 4.2.1 Uso e ocupação do solo

Para identificar as classes de uso e ocupação do solo com maior probabilidade de contribuição na poluição do ar na RMSR, atribui-se as células do domínio a classe predominante (maior área em km<sup>2</sup>) em cada uma delas (Figura 7).

**Figura 7.** Variabilidade dos valores de PSCF em cada uso do solo do domínio por onde as trajetórias se deslocaram.

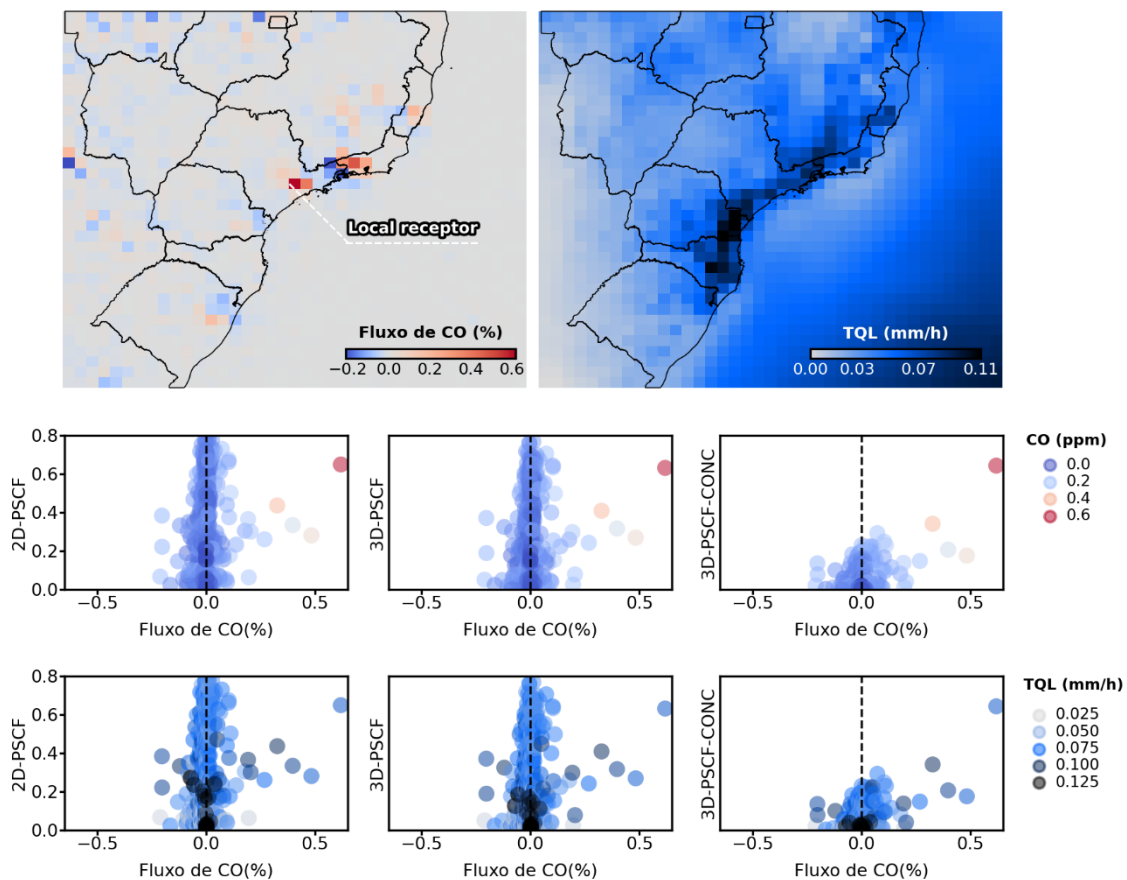


Os três métodos apresentaram padrões consistentes, identificando pastagens, produção de cana-de-açúcar, mosaico de agricultura e pastagem; e infraestrutura urbana como os principais tipos de uso do solo que contribuem associados a potencial contribuição na RMSP. Entre as classes de uso do solo, a cana-de-açúcar esteve associada aos maiores valores de contribuição nos métodos 2D-PSCF e 3D-PSCF. Apesar de ser conhecida como uma importante fonte de poluição do ar no domínio de modelagem, a queima de biomassa associada à produção de cana-de-açúcar ocorre principalmente no inverno (de maio a outubro) (CONAB, 2021). Como a emissão é pequena nas outras estações do ano, conseqüentemente a contribuição das células onde ocorre a produção da cana-de-açúcar é reduzida analisando o ano todo. Ao ponderar a probabilidade de contribuição através da concentração de poluentes, a 3D-PSCF-CONC reduz a influência das regiões onde há a o cultivo da cana-de-açúcar fora da temporada de colheita. Além disso, a 3D-PSCF-CONC reduziu significativamente a contribuição potencial da maioria dos usos do solo, exceto para infraestrutura urbana o que, mais uma vez, apresenta consistência com o cenário real, pois é classe de uso onde estão localizadas as principais fontes de emissão de poluentes atmosféricos (Kumar et al., 2016).

#### 4.2.2 Fluxo de poluentes e deposição úmida

A consistência dos métodos também foi avaliada comparando a variabilidade espacial dos valores de PSCF, fluxo de CO e TQL. A Figura 8 apresenta a distribuição espacial do fluxo médio de CO e sua relação com os resultados da PSCF. O fluxo negativo indica regiões onde as concentrações de CO diminuíram enquanto a massa de ar se deslocou em direção ao receptor. Áreas de alto fluxo, alta concentração e baixa precipitação têm maior probabilidade de contribuir para a poluição do ar. Portanto, espera-se encontrar altos valores de PSCF no lado direito dos gráficos de dispersão e marcadores circulares categorizados com altos valores de concentração de CO e baixos valores de TQL.

**Figura 8.** Distribuição espacial do fluxo médio de CO e TQL média. Os valores são categorizados pela concentração de CO e TQL média do MERRA-2.



Como pode ser observado, a 3D-PSCF-CONC alcançou os resultados mais consistentes, onde maiores valores de PSCF correspondem a áreas com maior fluxo de CO, maior concentração de CO e baixa TQL (Figura 8). Embora o CO tenha baixa solubilidade em água (Garrett et al., 2006), as células com menores valores de contribuição da 3D-PSCF-CONC coincidiram com áreas com maiores valores de TQL. Já os métodos 2D-PSCF e 3D-PSCF atribuíram contribuições de alto potencial para áreas com baixa concentração de CO, baixo fluxo de médio CO (próximo de zero) e altos valores de TQL. Quando testado usando a concentração de  $\text{SO}_2$ , que tem alta solubilidade em água (Zhang et al., 2018), a 3D-PSCF-CONC confirmou sua consistência (Apêndice B). Nossos resultados sugerem que 3D-PSCF-CONC pode representar melhor a contribuição de áreas menos poluídas, bem como a influência da deposição úmida.

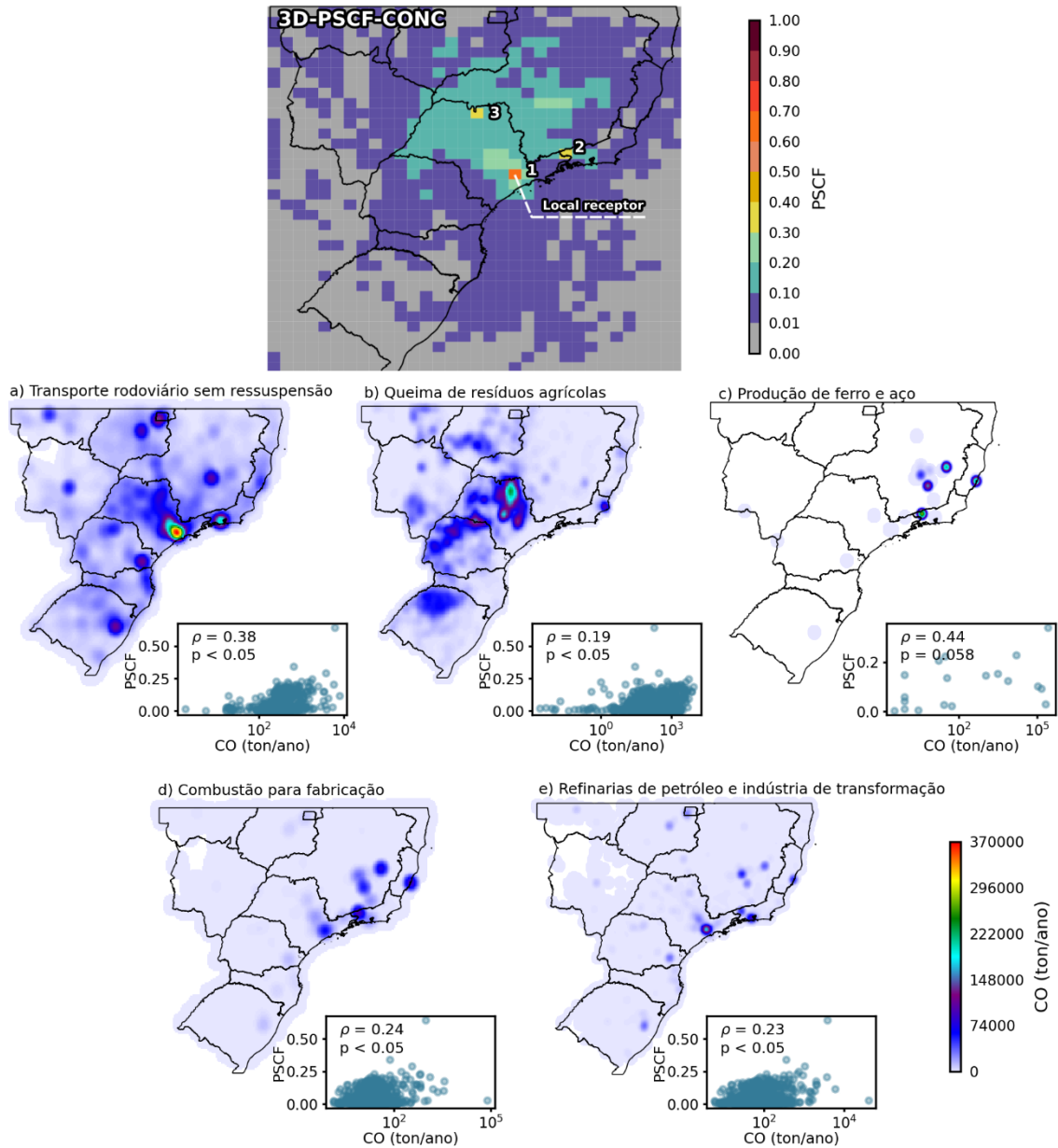
#### 4.2.3 Identificação das fontes de emissões atmosféricas

A 3D-PSCF-CONC identificou três áreas principais de contribuição potencial: (1) a RMSP; (2) o Médio Vale do Paraíba (MVP), no Estado do Rio de Janeiro; (3) e a Região de



Barretos no Norte do Estado de São Paulo. A RMSP é mundialmente conhecida pelos problemas de qualidade do ar, atribuídos principalmente ao transporte rodoviário e as emissões industriais (Andrade et al., 2017). A MVP é uma área intensamente industrializada, e onde está instalada uma importante indústria siderúrgica, a Companhia Siderúrgica Nacional (CSN) (INEA, 2018). As regiões norte e nordeste do Estado de São Paulo são caracterizadas pela alta produção agropecuária, sendo Barretos uma das regiões agrícolas com a maior produção de cana-de-açúcar do Brasil. Durante muitos anos, a queima do bagaço da cana-de-açúcar foi utilizada como uma prática comum para aumentar a produtividade da cana-de-açúcar em Barretos, prática esta que vem sendo substituída pela mecanização durante a colheita nos últimos anos (CONAB, 2021; Da Silva et al., 2021). Todas as fontes mencionadas foram caracterizadas pelo inventário global EDGAR (Figura 9).

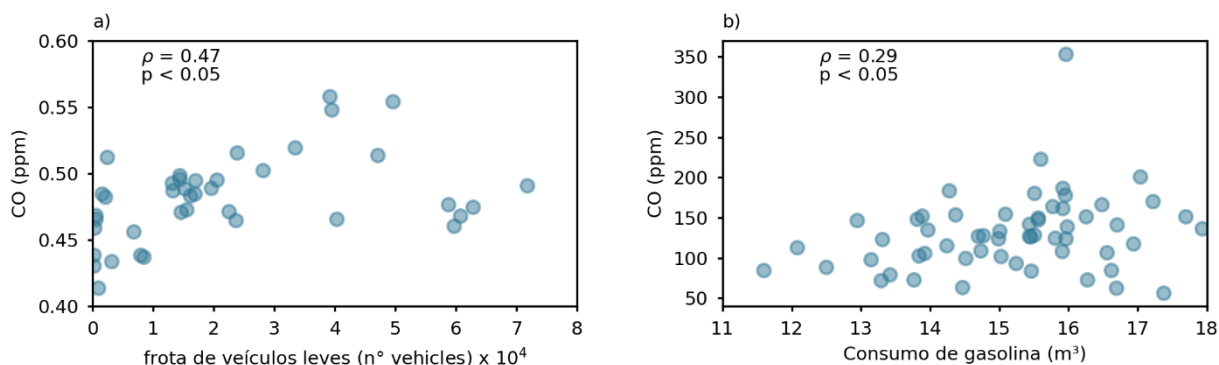
**Figura 9.** Correlação de Spearman entre a distribuição espacial dos valores 3D-PSCF-CONC e as emissões de CO das fontes de emissão do inventário EDGAR.



As estimativas 3D-PSCF-CONC estão espacialmente correlacionadas às emissões do transporte rodoviário sem ressuspensão do solo ( $\rho = 0,38$   $p < 0,05$ ), queima de resíduos agrícolas ( $\rho = 0,19$   $p < 0,05$ ), produção de ferro e aço ( $\rho = 0,44$   $p = 0,058$ ), combustão para a indústria transformadora ( $\rho = 0,24$   $p < 0,05$ ) e refinarias de petróleo e indústria de transformação ( $\rho = 0,23$   $p < 0,05$ ). Tanto o 2D-PSCF quanto o 3D-PSCF não realizaram boas correlações como o método proposto (Apêndice C). A análise realizada através do inventário de emissões EDGAR confirma que o 3D-PSCF-CONC foi capaz de fornecer a melhor identificação das fontes de emissão.

Pode-se observar que o transporte rodoviário do inventário do EDGAR coincide no espaço com a contribuição mais importante calculada pelo 3D-PSCF-CONC. Uma análise da concentração de CO do MERRA-2 na RMSP, frota de veículos leves e consumo de combustível na mesma região apresentou uma correlação positiva entre as variáveis (Figura 10). Dados da frota circulante anual da CETESB de 1980 a 2019, consumo mensal de gasolina de 2015 a 2019 (ANP, 2020) e dados de concentração de CO do MERRA-2 (M2T1NXCHM V5.12.4) na cidade de São Paulo, foram utilizados nesta análise. Os veículos leves são as principais fontes de emissão de monóxido de carbono e hidrocarbonetos, sendo os automóveis a gasolina os maiores emissores de CO (34%). Apesar de o tamanho da frota de veículos a gasolina ser menor do que o da frota de veículos flex, por exemplo, as emissões desse primeiro segmento são maiores em razão da maior idade média dos veículos a gasolina (CETESB, 2020a).

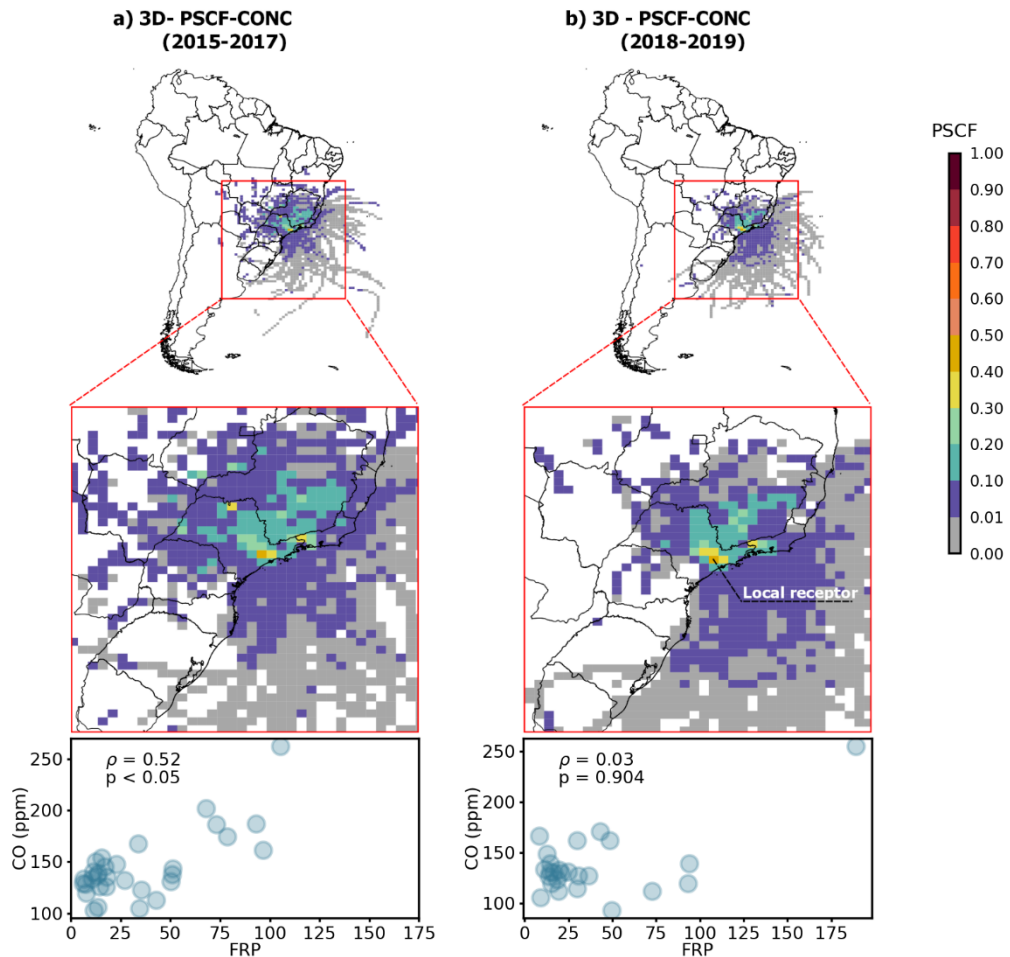
**Figura 10.** Correlação entre CO x frota de veículos leves (a) e CO x consumo de gasolina em São Paulo (b). A concentração de CO - MERRA-2 (M2T1NXCHM V5.12.4), frota de veículos leves (CETESB, 2019d), e consumo de gasolina (ANP, 2020).



O segundo maior valor de contribuição identificada pela 3D-PSCF-CONC ocorreu na mesma área onde está instalada a CSN, enquanto no terceiro maior valor de contribuição estão localizadas as emissões de CO identificadas pelo EDGAR como provenientes da queima de resíduos agrícolas. Conforme demonstrado por estudos anteriores (Vasconcellos et al., 2010; Souza et al., 2014), a queima do resíduo da cana-de-açúcar ocorre no período de maio a outubro. No mesmo período, o poder radiativo do fogo (FRP) atinge os níveis mais elevados na região de Barretos (Apêndice D). Foi verificada uma correlação positiva entre CO e FRP ( $\rho = 0,47$   $p < 0,05$ ), reforçando a influência desta fonte no local receptor. A queima da cana-de-açúcar vem diminuindo no Estado de São Paulo, e em 2018 houve uma redução drástica dessa emissão (IEA, 2020). Para verificar o efeito desta redução das emissões de CO oriundas da queima de biomassa, o método 3D-PSCF-CONC foi aplicado utilizando dois períodos distintos, de 2015

a 2017 e de 2018 a 2019 (Figura 11). Além disso, foi apresentado a correlação entre a concentração de CO e o FRP para ambos os períodos.

**Figura 11.** Análise do método 3D-PSCF-CONC na RMSP usando dois períodos (a) 2015 a 2017 e (b) 2018 a 2019 para verificar o efeito da redução das emissões da queima de biomassa sobre as contribuições de Barretos. Correlação entre as concentrações de CO e FRP.



Pode-se observar que o coeficiente de correlação entre FRP e concentração de CO em Barretos apresenta significância estatística apenas no período de 2015-2017 ( $\rho = 0,52$   $p < 0,05$ ). Os resultados confirmam que a influência da queima da cana-de-açúcar na concentração de CO em Barretos foi reduzida no período mais recente. Além disso, os resultados do 3D-PSCF-CONC no período 2015-2017 são maiores em Barretos do que em 2018-2019. Isso demonstra que a influência da queima do resíduo da cana-de-açúcar no local receptor era maior antes da redução desta emissão.

## 5 CONCLUSÃO

Este trabalho apresenta um novo método que incorpora dados de concentração de reanálises atmosféricas ao cálculo da PSCF. O 3D-PSCF-CONC inclui um fator de ponderação para corrigir a contribuição potencial de áreas menos poluídas. Na análise realizada na RMSP, foi observada uma superestimativa da contribuição de áreas distantes do local receptor pelos métodos 2D-PSCF e 3D-PSCF. Além disso, os métodos não destacaram a RMSP como uma potencial área de contribuição. Os resultados revelaram que o 3D-PSCF-CONC identifica de maneira mais consistente as fontes locais e regionais, ao invés de enfatizar principalmente as fontes regionais como os outros métodos PSCF. O 3D-PSCF-CONC apontou a influência do RMSP, Médio Vale do Paraíba e da região de Barretos como as principais áreas contribuintes para a poluição do ar em São Paulo.

Para avaliar a consistência dos métodos PSCF, foram utilizados dados do tipo de uso do solo, fluxo de poluentes e precipitação. Os três métodos identificaram as pastagens, produção de cana-de-açúcar, mosaico de agricultura e pastagem e infraestrutura urbana como os principais tipos de uso do solo que contribuem para poluição na RMSP. No entanto, os métodos 2D-PSCF e 3D-PSCF atribuíram as maiores contribuições à produção de cana-de-açúcar, enquanto o 3D-PSCF-CONC destacou as emissões oriundas da infraestrutura urbana como sendo a que possui maior potencial de contribuição, apresentando consistência com o cenário real. Além disso, o 3D-PSCF-CONC apresentou melhor adequabilidade ao analisar a influência dos fluxos de poluentes e do processo de deposição úmida. O método atribuiu maior potencial de contribuição a regiões onde há maior fluxo de poluentes e menor ocorrência de chuva.

O 3D-PSCF-CONC demonstrou bom desempenho também na identificação das fontes de emissão de poluentes atmosféricos. Os resultados dos 3 métodos PSCF foram avaliados através da análise espacial entre os valores de contribuição e as taxas de fontes de emissão do EDGAR. A 3D-PSCF-CONC apresentou os maiores valores de correlação de Spearman entre a PSCF e as taxas de emissão, apontando o transporte rodoviário na RMSP e a queima de biomassa na região de Barretos como as principais fontes de contaminação da qualidade do ar. Os veículos leves movidos a gasolina foram identificados como a principal fonte de CO na RMSP, enquanto a queima da palha da cana-de-açúcar, durante o período de colheita, foi identificada como a principal responsável pela emissão de CO em Barretos, sobre tudo nos anos de 2015 a 2017.

O novo método apresentado neste trabalho pode ser usado como uma ferramenta importante para a identificação de fontes e para o entendimento do processo de poluição do ar. Recomenda-se que para pesquisas futuras seja avaliada as incertezas associadas ao método.

Análise de sensibilidade do conjunto de dados do MERRA-2 devem ser conduzidos na região de estudo, assim como a utilização de trajetórias de massa de ar com diferentes tempos de duração para avaliar o efeito da distância entre a fonte potencial e o local receptor. O método 3D-PSCF-CONC pode ser utilizado também para investigar fontes de emissões e processos atmosféricos através de diferentes conjuntos de dados de inventário de emissões, bem como o uso de dados de alta resolução como o do Sentinel-5p.

## 6 REFERÊNCIAS BIBLIOGRÁFICAS

Acker, J., Soebiyanto, R., Kiang, R., Kempler, S., 2014. Geo-Information Use of the NASA Giovanni Data System for Geospatial Public Health Research: Example of Weather-Influenza Connection. *ISPRS Int. J. Geo-Inf* 3, 1372–1386. <https://doi.org/10.3390/ijgi3041372>

Andrade, M. de F., Kumar, P., de Freitas, E.D., Ynoue, R.Y., Martins, J., Martins, L.D., Nogueira, T., Perez-Martinez, P., de Miranda, R.M., Albuquerque, T., Gonçalves, F.L.T., Oyama, B., Zhang, Y., 2017. Air quality in the megacity of São Paulo: Evolution over the last 30 years and future perspectives. *Atmos. Environ.* 159, 66–82. <https://doi.org/10.1016/j.atmosenv.2017.03.051>

ANP, 2020. Petroleum, Natural Gas, and Biofuels National Agency, Brazil. <http://www.anp.gov.br> (accessed 20.07.2020).

Ara Begum, B., Kim, E., Jeong, C.-H., Lee, D.-W., Hopke, P.K., 2005. Evaluation of the potential source contribution function using the 2002 Quebec forest fire episode. *Atmos. Environ.* 39, 3719–3724. <https://doi.org/10.1016/j.atmosenv.2005.03.008>

Ashbaugh, L.L., Malm, W.C., Sadeh, W.Z., 1985. A residence time probability analysis of sulfur concentrations at grand Canyon National Park. *Atmos. Environ.* 19, 1263–1270. [https://doi.org/10.1016/0004-6981\(85\)90256-2](https://doi.org/10.1016/0004-6981(85)90256-2)

Bali, K., Mishra, A.K., Singh, S., 2017. Impact of anomalous forest fire on aerosol radiative forcing and snow cover over Himalayan region. *Atmos. Environ.* 150, 264–275. <https://doi.org/10.1016/j.atmosenv.2016.11.061>

Begum, B.A., Biswas, S.K., Kim, E., Hopke, P.K., Khaliqzaman, M., 2005. Investigation of sources of atmospheric aerosol at a hot spot area in Dhaka, Bangladesh. *J. Air Waste Manage. Assoc.* 55, 227–240. <https://doi.org/10.1080/10473289.2005.10464606>

Bosilovich, M. G., and Coauthors, 2016: MERRA-2: Initial evaluation of the climate. NASA/TM-2015-104606, Vol. 43, NASA Global Modeling and Assimilation Office, 139 pp. [Available online at <https://gmao.gsfc.nasa.gov/pubs/tm/docs/Bosilovich803.pdf>.]

Buchard, V., and Coauthors, 2015: Using the OMI aerosol index and absorption aerosol optical depth to evaluate the NASA MERRA Aerosol Reanalysis. *Atmos. Chem. Phys.*, 15, 5743–5760, doi:10.5194/acp-15-5743-2015.

CETESB, 2012. Companhia de Tecnologia de Saneamento Ambiental - Qualidade do ar no Estado de São Paulo 2011. Disponível em <https://cetesb.sp.gov.br/ar/wp-content/uploads/sites/28/2013/12/Relatorio-Ar-2011-1.zip>

CETESB, 2015. Relatório da Qualidade do ar, 2015. Disponível em: <<https://cetesb.sp.gov.br/ar/wp-content/uploads/sites/28/2013/12/RQAR-2015.pdf>>. (accessed em: 15 fev. 2020).

CETESB, 2016. Relatório da Qualidade do ar, 2016. Disponível em: <<https://cetesb.sp.gov.br/ar/wp-content/uploads/sites/28/2013/12/relatorio-ar-2016.pdf>>. (accessed em: 15 fev. 2020).

CETESB, 2017. Relatório da Qualidade do ar, 2017. Disponível em: <<https://cetesb.sp.gov.br/ar/wp-content/uploads/sites/28/2019/05/Relat%C3%B3rio-de-Qualidade-do-Ar-2017.pdf>>. (accessed em: 15 fev. 2020).

CETESB, 2018. Relatório da Qualidade do ar, 2018. Disponível em: <<https://cetesb.sp.gov.br/ar/wp-content/uploads/sites/28/2019/07/Relat%C3%B3rio-de-Qualidade-do-Ar-2018.pdf>>. (accessed em: 15 fev. 2020).

CETESB, 2019a. QUALAR Acesso ao Sistema, 2019. Disponível em: <https://cetesb.sp.gov.br/ar/qualar/>>. (accessed em: 15 fev. 2020).

CETESB, 2019b. Saiba mais sobre a qualidade do ar. Disponível em: <<https://cetesb.sp.gov.br/blog/2019/06/18/saiba-mais-sobre-a-qualidade-do-ar/>>. (accessed 07.02.2021).

CETESB, 2020a. Relatório da Qualidade do ar, 2019. Disponível em: <<https://cetesb.sp.gov.br/ar/wp-content/uploads/sites/28/2020/07/Relatório-de-Qualidade-do-Ar-2019.pdf>>. (accessed em: 3 mar. 2020).

CETESB, 2020b. Relatórios de Emissões Veiculares no Estado São Paulo: Relatório 2019. Disponível em: <<https://cetesb.sp.gov.br/ar/publicacoes-relatorios/>>. (accessed 02.11.2020).

CETESB, 2020c. Relatórios de Emissões Veiculares no Estado São Paulo: Frota Circulante no Estado, 2019. Disponível em: <<https://cetesb.sp.gov.br/veicular/relatorios-e-publicacoes/>>. (accessed 02.10.2020).

Chalbot, M.C., Lianou, M., Vei, I.C., Kotronarou, A., Kavouras, I.G., 2013. Spatial attribution of sulfate and dust aerosol sources in an urban area using receptor modeling coupled with Lagrangian trajectories. *Atmos. Pollut. Res.* 4, 346–353. <https://doi.org/10.5094/APR.2013.039>

Cheng, I., Xu, X., Zhang, L., 2015. Overview of receptor-based source apportionment studies for speciated atmospheric mercury. *Atmos. Chem. Phys.* 15, 7877–7895. <https://doi.org/10.5194/ACP-15-7877-2015>

Cheng, M.-D., Lin, C.-J., 2001. Receptor modeling for smoke of 1998 biomass burning in Central America. *J. Geophys. Res. Atmos.* 106, 22871–22886. <https://doi.org/10.1029/2001JD900024>

CONAB, 2021. Acompanhamento de safra brasileira cana-de-açúcar - safra 2019/2020. Companhia Nacional de Abastecimento, v. 6, p. 58, 2019. <https://www.conab.gov.br/info-agro/safra/cana/boletim-da-safra-de-cana-de-acucar>. (accessed 21.09.2020)

Crawford, J., Chambers, S., Cohen, D.D., Dyer, L., Wang, T., Zahorowski, W., 2007. Receptor modelling using Positive Matrix Factorisation, back trajectories and Radon-222. *Atmos. Environ.* 41, 6823–6837. <https://doi.org/10.1016/j.atmosenv.2007.04.048>

Crippa, M., Guizzardi, D., Muntean, M., Schaaf, E., Dentener, F., Van Aardenne, J.A., Monni, S., Doering, U., Olivier, J.G.J., Pagliari, V., Janssens-Maenhout, G., 2018. Gridded emissions



of air pollutants for the period 1970-2012 within EDGAR v4.3.2. *Earth Syst. Sci. Data*. <https://doi.org/10.5194/essd-10-1987-2018>

D'Avignon A, Carloni FA, La Rovere EL, Dubeux CBS, 2010. Emission inventory: an urban public policy instrument and benchmark. *Energy Policy* 38(9):4838–4847. <https://doi.org/10.1016/j.enpol.2009.10.002>

da Silva, G.J., Berg, E.C., Calijuri, M.L., dos Santos, V.J., Lorentz, J.F., Carmo Alves, S. do, 2021. Aptitude of areas planned for sugarcane cultivation expansion in the state of São Paulo, Brazil: a study based on climate change effects. *Agric. Ecosyst. Environ.* 305, 107164. <https://doi.org/10.1016/j.agee.2020.107164>

Dimitriou, K., Grivas, G., Liakakou, E., Gerasopoulos, E., Mihalopoulos, N., 2021. Assessing the contribution of regional sources to urban air pollution by applying 3D-PSCF modeling. *Atmos. Res.* 248, 105187. <https://doi.org/10.1016/j.atmosres.2020.105187>

Dimitriou, K., Kassomenos, P., 2016. Combining AOT, Angstrom Exponent and PM concentration data, with PSCF model, to distinguish fine and coarse aerosol intrusions in Southern France. *Atmos. Res.* 172–173, 74–82. <https://doi.org/10.1016/j.atmosres.2016.01.002>

Draxler, R. R.; Hess, G. D., 1998. An overview of the HYSPLIT\_4 modelling system for trajectories, dispersion and deposition. *Australian Meteorological Magazine*, v. 47, n. 4, p. 295–308. <https://www.arl.noaa.gov/documents/reports/MetMag.pdf>

Eduardo Souto-Oliveira, C., De Fátima Andrade, M., Kumar, P., Juliano Da Silva Lopes, F., Babinski, M., Landulfo, E., 2016. Effect of vehicular traffic, remote sources and new particle formation on the activation properties of cloud condensation nuclei in the megacity of São Paulo, Brazil. *Atmos. Chem. Phys.* 16, 14635–14656. <https://doi.org/10.5194/acp-16-14635-2016>

Fan, W., Qin, K., Xu, J., Yuan, L., Li, D., Jin, Z., Zhang, K., 2019. Aerosol vertical distribution and sources estimation at a site of the Yangtze River Delta region of China. *Atmos. Res.* 217, 128–136. <https://doi.org/10.1016/j.atmosres.2018.11.002>

FEEMA, 2004. Fundação Estadual de Engenharia do Meio Ambiente - Inventário de fontes emissoras de poluentes atmosféricos da Região Metropolitana do Rio de Janeiro. Rio de Janeiro, 2004.

França, D., Longo, K., Rudorff, B., Aguiar, D., Freitas, S., Stockler, R., Pereira, G., 2014. Pre-harvest sugarcane burning emission inventories based on remote sensing data in the state of São paulo, Brazil. *Atmos. Environ.* 99, 446–456. <https://doi.org/10.1016/j.atmosenv.2014.10.010>

Gallagher J, Baldauf R, Fuller CH, Kumar P, Gill LW, McNabola A, 2015. Passive methods for improving air quality in the built environment: a review of porous and solid barriers. *Atmos Environ* 120: 61–70. <https://doi.org/10.1016/j.atmosenv.2015.08.075>

Gao, N., Cheng, M.D., Hopke, P.K., 1994. Receptor modeling of airborne ionic species collected in SCAQS. *Atmos. Environ.* 28, 1447–1470. [https://doi.org/10.1016/1352-2310\(94\)90207-0](https://doi.org/10.1016/1352-2310(94)90207-0)

Garrett, T.J., Avey, L., Palmer, P.I., Stohl, A., Neuman, J.A., Brock, C.A., Ryerson, T.B., Holloway, J.S., 2006. Quantifying wet scavenging processes in aircraft observations of nitric acid and cloud condensation nuclei. *J. Geophys. Res. Atmos.* 111, 23–51. <https://doi.org/10.1029/2006JD007416>

Gelaro, R., McCarty, W., Suárez, M.J., Todling, R., Molod, A., Takacs, L., Randles, C.A., Darmenov, A., Bosilovich, M.G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A.M., Gu, W., Kim, G.K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J.E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S.D., Sienkiewicz, M., Zhao, B., 2017. The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). *J. Clim.* 30, 5419–5454. <https://doi.org/10.1175/JCLI-D-16-0758.1>

Gildemeister, A.E., Hopke, P.K., Kim, E., 2007. Sources of fine urban particulate matter in Detroit, MI. *Chemosphere* 69, 1064–1074. <https://doi.org/10.1016/J.CHEMOSPHERE.2007.04.027>

Gurjar BR, Ohara T, Khare M, Kulshrestha P, Tyagi V, Nagpure AS, 2018. South Asian perspective: a case of urban air pollution and potential for climate co-benefits in India, In *Mainstreaming Climate Co-Benefits in Indian Cities*(pp. 77–98). Springer, Singapore. [https://doi.org/10.1007/978-981-10-5816-5\\_3](https://doi.org/10.1007/978-981-10-5816-5_3)

Hafner, W. D.; Hites, R. A., 2003. Potential sources of pesticides, PCBs, and PAHs to the atmosphere of the Great Lakes. *Environmental Science and Technology*, v. 37, n. 17, p. 3764–3773. <https://pubmed.ncbi.nlm.nih.gov/12967094>

Heo, J.B., Hopke, P.K., Yi, S.M., 2009. Source apportionment of PM<sub>2.5</sub> in seoul, korea. *Atmos. Chem. Phys.* 9, 4957–4971. <https://doi.org/10.5194/ACP-9-4957-2009>

Hopke, P. K.; Gao, N.; Cheng, M. D., 1993. Combining chemical and meteorological data to infer source areas of airborne pollutants. *Chemometrics and Intelligent Laboratory Systems*, v. 19, n. 2, p. 187–199. [https://doi.org/10.1016/0169-7439\(93\)80103-O](https://doi.org/10.1016/0169-7439(93)80103-O)

Hopke, P.K., Gao, N., Cheng, M.D., 1993. Combining chemical and meteorological data to infer source areas of airborne pollutants. *Chemom. Intell. Lab. Syst.* 19, 187–199. [https://doi.org/10.1016/0169-7439\(93\)80103-O](https://doi.org/10.1016/0169-7439(93)80103-O)

Hsu, Y. K.; Holsen, T. M.; Hopke, P. K., 2003. Comparison of hybrid receptor models to locate PCB sources in Chicago. *Atmospheric Environment*, v. 37, n. 4, p. 545–562. [https://doi.org/10.1016/S1352-2310\(02\)00886-5](https://doi.org/10.1016/S1352-2310(02)00886-5)

Hwang, I., Hopke, P.K., 2007. Estimation of source apportionment and potential source locations of PM<sub>2.5</sub> at a west coastal IMPROVE site. *Atmos. Environ.* 3, 506–518. <https://doi.org/10.1016/J.ATMOSENV.2006.08.043>

IBGE, 2017. Instituto Brasileiro de Geografia e Estatística (Brazilian Institute of Geography and Statistics). <http://cod.ibge.gov.br/QHF>. (accessed 15.12.2020)

IEA, 2020. Instituto de Economia Agrícola (IEA), Colheita Mecanizada da Cana-de-açúcar Atinge 95,3% das Áreas Produtivas do Estado de São Paulo na Safra Agrícola 2018/19. <http://www.iea.sp.gov.br/out/TerTexto.php?codTexto=14825> (accessed 18.06.2021)

In, H.J., Kim, Y.P., 2010. Estimation of the aerosol optical thickness distribution in the Northeast Asian forest fire episode in May 2003: Possible missing emissions. *Atmos. Res.* 98, 261–273. <https://doi.org/10.1016/j.atmosres.2010.09.009>

INEA, 2018. Relatório da qualidade do ar do Estado do Rio de Janeiro – ano base 2018, p. 140, 2020. Rio de Janeiro. Disponível em: < <http://www.inea.rj.gov.br/wp-content/uploads/2020/11/relatorio-qualidade-ar-2018.pdf>>

Isley, C.F., Taylor, M.P., 2018. Air quality management in the Pacific Islands: A review of past performance and implications for future directions. *Environ. Sci. Policy* 84, 26–33. <https://doi.org/10.1016/j.envsci.2018.02.013>

Jaffe, D., McKendry, I., Anderson, T., Price, H., 2003. Six ‘new’ episodes of trans-Pacific transport of air pollutants. *Atmos. Environ.* 37, 391–404. [https://doi.org/10.1016/S1352-2310\(02\)00862-2](https://doi.org/10.1016/S1352-2310(02)00862-2)

Janssens-Maenhout, G., Crippa, M., Guizzardi, D., Muntean, M., Schaaf, E., Dentener, F., Bergamaschi, P., Pagliari, V., Olivier, J.G.J., Peters, J.A.H.W., Van Aardenne, J.A., Monni, S., Doering, U., Roxana Petrescu, A.M., Solazzo, E., Oreggioni, G.D., 2019. EDGAR v4.3.2 Global Atlas of the three major greenhouse gas emissions for the period 1970-2012. *Earth Syst. Sci. Data* 11, 959–1002. <https://doi.org/10.5194/essd-11-959-2019>

Jeong, U., Kim, J., Lee, H., Jung, J., Kim, Y.J., Song, C.H., Koo, J.-H., 2011. Estimation of the contributions of long range transported aerosol in East Asia to carbonaceous aerosol and PM concentrations in Seoul, Korea using highly time resolved measurements: a PSCF model approach. *J. Environ. Monit.* 13, 1905–1918. <https://doi.org/10.1039/C0EM00659A>

Kim, I.S., Lee, J.Y., Wee, D., Kim, Y.P., 2019. Estimation of the contribution of biomass fuel burning activities in North Korea to the air quality in Seoul, South Korea: Application of the 3D-PSCF method. *Atmos. Res.* 230, 104628. <https://doi.org/10.1016/j.atmosres.2019.104628>

Kim, I.S., Wee, D., Kim, Y.P., Lee, J.Y., 2016. Development and application of three-dimensional potential source contribution function (3D-PSCF). *Environ. Sci. Pollut. Res.* 23, 16946–16954. <https://doi.org/10.1007/s11356-016-6787-x>

Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, R. M. Errico, and R. Yang, 2009. Improving incremental balance in the GSI 3DVAR analysis system. *Mon. Wea. Rev.*, 137, 1046– 1060, doi:10.1175/2008MWR2623.1.

Kumar, P., de Fatima Andrade, M., Ynoue, R.Y., Fornaro, A., de Freitas, E.D., Martins, J., Martins, L.D., Albuquerque, T., Zhang, Y., Morawska, L., 2016. New directions: From biofuels to wood stoves: The modern and ancient air quality challenges in the megacity of São Paulo. *Atmos. Environ.* <https://doi.org/10.1016/j.atmosenv.2016.05.059>

Landrigan PJ, Fuller R, Acosta NJ, Adeyi O, Arnold R, Baldé AB et al, 2018. The Lancet Commission on pollution and health. *Lancet* 391(10119):462–512. [https://doi.org/10.1016/S0140-6736\(17\)32345-0](https://doi.org/10.1016/S0140-6736(17)32345-0)

Longhurst, J.W.S., Irwin, J.G., Chatterton, T.J., Hayes, E.T., Leksmono, N.S., Symons, J.K., 2009. The development of effects-based air quality management regimes. *Atmos. Environ.* 43, 64–78. <https://doi.org/10.1016/j.atmosenv.2008.09.050>

Lucey, D., Hadjiiski, L., Hopke, P.K., Scudlark, J.R., Church, T., 2001. Identification of sources of pollutants in precipitation measured at the mid-Atlantic US coast using potential source contribution function (PSCF). *Atmos. Environ.* 35, 3979–3986. [https://doi.org/10.1016/S1352-2310\(01\)00185-6](https://doi.org/10.1016/S1352-2310(01)00185-6)

MAPBIOMAS, 2020. Projeto MapBiomass - Coleção [versão 4.0] da Série Anual de Mapas de Cobertura e Uso de Solo do Brasil. <http://plataforma.mapbiomas.org/map#coverage>. (accessed 15.12. 2020)

Martins, L.D., Wikuats, C.F.H., Capucim, M.N., de Almeida, D.S., da Costa, S.C., Albuquerque, T., Barreto Carvalho, V.S., de Freitas, E.D., de Fátima Andrade, M., Martins, J.A., 2017. Extreme value analysis of air pollution data and their comparison between two large urban regions of South America. *Weather Clim. Extrem.* 18, 44–54. <https://doi.org/10.1016/j.wace.2017.10.004>

Merrifield, J., 1990. A Critical Overview of the Evolutionary Approach to Air Pollution Abatement Policy. *J. Policy Anal. Manag.* 9, 367. <https://doi.org/10.2307/3325281>

Molod, A., Takacs, L., Suarez, M., Bacmeister, J., 2015. Development of the GEOS-5 atmospheric general circulation model: Evolution from MERRA to MERRA2. *Geosci. Model Dev.* 8, 1339–1356. <https://doi.org/10.5194/gmd-8-1339-2015>

Nakazato, R.K., Rinaldi, M.C.S., Domingos, M., 2015. Will technological modernization for power generation at an oil refinery diminish the risks from air pollution to the Atlantic Rainforest in Cubatão, SE Brazil? *Environ. Pollut.* 196, 489–496. <https://doi.org/10.1016/j.envpol.2014.05.011>

Nguyen, T.T.N., Bui, H.Q., Pham, H. V., Luu, H. V., Man, C.D., Pham, H.N., Le, H.T., Nguyen, T.T., 2015. Particulate matter concentration mapping from MODIS satellite data: A Vietnamese case study. *Environ. Res. Lett.* 10. <https://doi.org/10.1088/1748-9326/10/9/095016>

Ohara T, Akimoto H, Kurokawa J, Horii N, Yamaji K, Yan X, Hayasaka T, 2007. An Asian emission inventory of anthropogenic emission sources for the period 1980-2020. *Atmos Chem Phys* 7:4419–4444.

ONU, 2018. UN World Urbanization Prospects. (The Revision 2018) <<https://population.un.org/wup/Publications/Files/WUP2018-Report.pdf>>

Poirot, R.L., Wishinski, P.R., 1986. Visibility, sulfate and air mass history associated with the summertime aerosol in northern Vermont. *Atmos. Environ.* 20, 1457–1469. [https://doi.org/10.1016/0004-6981\(86\)90018-1](https://doi.org/10.1016/0004-6981(86)90018-1)

Polissar, A. V., Hopke, P.K., Harris, J.M., 2001. Source regions for atmospheric aerosol measured at Barrow, Alaska. *Environ. Sci. Technol.* 35, 4214–4226. <https://doi.org/10.1021/ES0107529>

Potier, E., Waked, A., Bourin, A., Minvielle, F., Péré, J.C., Perdrix, E., Michoud, V., Riffault, V., Alleman, L.Y., Sauvage, S., 2019. Characterizing the regional contribution to PM 10 pollution over northern France using two complementary approaches: Chemistry transport and trajectory-based receptor models. *Atmos. Res.* 223, 1–14. <https://doi.org/10.1016/j.atmosres.2019.03.002>

Potier, E., Waked, A., Bourin, A., Minvielle, F., Péré, J.C., Perdrix, E., Michoud, V., Riffault, V., Alleman, L.Y., Sauvage, S., 2019. Characterizing the regional contribution to PM 10 pollution over northern France using two complementary approaches: Chemistry transport and trajectory-based receptor models. *Atmos. Res.* 223, 1–14. <https://doi.org/10.1016/j.atmosres.2019.03.002>

Punsompong, P.; Chantara, S., 2018. Identification of potential sources of PM10 pollution from biomass burning in northern Thailand using statistical analysis of trajectories. *Atmospheric Pollution Research*, v. 9, n. 6, p. 1038–1051. <http://dx.doi.org/10.1016/j.apr.2018.04.003>

Qin, W., Zhang, Y., Chen, J., Yu, Q., Cheng, S., Li, W., Liu, X., Tian, H., 2019. Variation, sources and historical trend of black carbon in Beijing, China based on ground observation and MERRA-2 reanalysis data. *Environ. Pollut.* 245, 853–863. <https://doi.org/10.1016/j.envpol.2018.11.063>

Randles, C.A., da Silva, A.M., Buchard, V., Colarco, P.R., Darmenov, A., Govindaraju, R., Smirnov, A., Holben, B., Ferrare, R., Hair, J., Shinozuka, Y., Flynn, C.J., 2017. The MERRA-2 aerosol reanalysis, 1980 onward. Part I: System description and data assimilation evaluation. *J. Clim.* 30, 6823–6850. <https://doi.org/10.1175/JCLI-D-16-0609.1>

Reichle, R.H., Liu, Q., Koster, R.D., Draper, C.S., Mahanama, S.P.P., Partyka, G.S., 2017. Land Surface Precipitation in MERRA-2. *J. Clim.* 30, 1643–1664. <https://doi.org/10.1175/JCLI-D-16-0570.1>

Richard L. Poirot, Paul R. Wishinski, Philip K. Hopke\*, and Polissar, A. V., 2001. Comparative Application of Multiple Receptor Methods To Identify Aerosol Sources in Northern Vermont. *Environ. Sci. Technol.* 36, 820. <https://doi.org/10.1021/ES011442T>

Rienecker, M. M., and Coauthors, 2008: The GEOS-5 Data Assimilation System—Documentation of versions 5.0.1, 5.1.0, and 5.2.0. Technical Report Series on Global Modeling and Data Assimilation, Vol. 27, NASA Tech. Rep. NASA/TM-2008-104606, 118 pp. [Available online at <https://gmao.gsfc.nasa.gov/pubs/docs/Rienecker369.pdf>.]

Seinfeld, J.H., Pandis, S.N., 1998. *Atmospheric Chemistry and Physics: from Air Pollution to Climate Change*. John Wiley and Sons, New York, NY, pp. 74–5, 1056–7

Shanavas, A.K., Zhou, C., Menon, R., Hopke, P.K., 2020. PM10 source identification using the trajectory based potential source apportionment (TraPSA) toolkit at Kochi, India. *Atmos. Pollut. Res.* 11, 1535–1542. <https://doi.org/10.1016/j.apr.2020.06.019>

Shen, J.; Cao, N., 2020. Comprehensive observation and analysis of aerosol optical properties and vertical distribution in Nanjing, China. *Atmospheric Environment*, v. 239, p. 117767. <https://doi.org/10.1016/j.atmosenv.2020.117767>

Souza, D.Z., Vasconcellos, P.C., Lee, H., Aurela, M., Saarnio, K., Teinilä, K., Hillamo, R., 2014. Composition of PM<sub>2.5</sub> and PM<sub>10</sub> Collected at Urban Sites in Brazil. *Aerosol Air Qual. Res.* 14, 168–176. <https://doi.org/10.4209/aaqr.2013.03.0071>

Stein, A.F., Draxler, R.R., Rolph, G.D., Stunder, B.J.B., Cohen, M.D., Ngan, F., 2015. Noaa's hysplit atmospheric transport and dispersion modeling system. *Bull. Am. Meteorol. Soc.* 96, 2059–2077. <https://doi.org/10.1175/BAMS-D-14-00110.1>

Stibig, H.J., Achard, F., Carboni, S., Ra?si, R., Miettinen, J., 2014. Change in tropical forest cover of Southeast Asia from 1990 to 2010. *Biogeosciences* 11 (2), 247e258.

Stohl, A., Andreas, 1996. Trajectory statistics-A new method to establish source-receptor relationships of air pollutants and its application to the transport of particulate sulfate in Europe. *AtmEn* 30, 579–587. [https://doi.org/10.1016/1352-2310\(95\)00314-2](https://doi.org/10.1016/1352-2310(95)00314-2)

Stojić, A., Stanišić Stojić, S., 2017. The innovative concept of three-dimensional hybrid receptor modeling. *Atmos. Environ.* 164, 216–223. <https://doi.org/10.1016/j.atmosenv.2017.06.009>

Sun, E., Xu, X., Che, H., Tang, Z., Gui, K., An, L., Lu, C., Shi, G., 2019. Variation in MERRA-2 aerosol optical depth and absorption aerosol optical depth over China from 1980 to 2017. *J. Atmos. Solar-Terrestrial Phys.* 186, 8–19. <https://doi.org/10.1016/j.jastp.2019.01.019>

Timonen, H., Wigder, N., Jaffe, D., 2013. Influence of background particulate matter (PM) on Urban air quality in the Pacific Northwest. *J. Environ. Manage.* 129, 333–340. <https://doi.org/10.1016/j.jenvman.2013.07.023>

UN, 2016. United Nations, Department of Economic and Social Affairs, Population Division. <http://esa.un.org/unpd/wup/highlights/wup2016-highlights.pdf>. (accessed 04.02.2020).

Vasconcellos, P.C., Souza, D.Z., Sanchez-Ccoyllo, O., Bustillos, J.O. V., Lee, H., Santos, F.C., Nascimento, K.H., Araújo, M.P., Saarnio, K., Teinilä, K., Hillamo, R., 2010. Determination of anthropogenic and biogenic compounds on atmospheric aerosol collected in urban, biomass burning and forest areas in São Paulo, Brazil. *Sci. Total Environ.* 408, 5836–5844. <https://doi.org/10.1016/j.scitotenv.2010.08.012>

Vet, R., Ro, C.-U., 2008. Contribution of Canada–United States transboundary transport to wet deposition of sulphur and nitrogen oxides—A mass balance approach. *Atmos. Environ.* 42, 2518–2529. <https://doi.org/10.1016/J.ATMOSENV.2007.12.034>

Wei, W., Zang, J., Wang, X., Cheng, S., 2020. Peroxyacetyl nitrate (PAN) in the border of Beijing, Tianjin and Hebei of China: Concentration, source apportionment and photochemical pollution assessment. *Atmos. Res.* 246, 105106. <https://doi.org/10.1016/j.atmosres.2020.105106>

WHO, 2016. World Health Organization, Air pollution levels rising in many of the world's poorest cities. p. 10–13, 2016. <https://www.who.int/news/item/12-05-2016-air-pollution-levels-rising-in-many-of-the-world-s-poorest-cities>. (accessed 07.08.2020)

WILKS, D. S., 2013. Statistical methods in the atmospheric sciences. [s.l: s.n.]. v. 53  
Xu, J., DuBois, D., Pitchford, M., Green, M., Etyemezian, V., 2006. Attribution of sulfate aerosols in Federal Class I areas of the western United States based on trajectory regression analysis. *Atmos. Environ.* 40, 3433–3447. <https://doi.org/10.1016/j.atmosenv.2006.02.009>

Yamashita, N., Sase, H., Ohizumi, T., Kurokawa, J., Ohara, T., Morino, Y., Kuribayashi, M., Ohta, S., Kaneko, S., Hayashi, K., Fukuhara, H., Hakamata, T., 2016. Mapping the relative risk of surface water acidification based on cumulative acid deposition over the past 25 years in Japan. *J. For. Res.* 21, 115–124. <https://doi.org/10.1007/s10310-016-0523-8>

Yim, S.H.L., Hou, X., Guo, J., Yang, Y., 2019. Contribution of local emissions and transboundary air pollution to air quality in Hong Kong during El Niño-Southern Oscillation and heatwaves. *Atmos. Res.* 218, 50–58. <https://doi.org/10.1016/j.atmosres.2018.10.021>

Yin, S., Wang, X., Zhang, X., Guo, M., Miura, M., Xiao, Y., 2019. Influence of biomass burning on local air pollution in mainland Southeast Asia from 2001 to 2016. *Environ. Pollut.* 254, 112949. <https://doi.org/10.1016/j.envpol.2019.07.117>

Zachary, M., Yin, L., Zacharia, M., Zachary, M., Yin, L., Zacharia, M., 2018. Application of PSCF and CWT to Identify Potential Sources of Aerosol Optical Depth in ICIPE Mbita. *Open Access Libr. J.* 5, 1–12. <https://doi.org/10.4236/OALIB.1104487>

Zeng, Y.; Hopke, P. K., 1989. A study of the sources of acid precipitation in Ontario, Canada. *Atmospheric Environment (1967)*, v. 23, n. 7, p. 1499–1509. [https://doi.org/10.1016/0004-6981\(89\)90409-5](https://doi.org/10.1016/0004-6981(89)90409-5)

Zhang, H., Li, J., Ying, Q., Yu, J.Z., Wu, D., Cheng, Y., He, K., Jiang, J., 2012. Source apportionment of PM 2.5 nitrate and sulfate in China using a source-oriented chemical transport model. *Atmos. Environ.* 62, 228–242. <https://doi.org/10.1016/j.atmosenv.2012.08.014>

Zhang, Q., Jiang, X., Tong, D., Davis, S.J., Zhao, H., Geng, G., Feng, T., Zheng, B., Lu, Z., Streets, D.G., Ni, R., Brauer, M., Van Donkelaar, A., Martin, R. V., Huo, H., Liu, Z., Pan, D., Kan, H., Yan, Y., Lin, J., He, K., Guan, D., 2017. Transboundary health impacts of transported global air pollution and international trade. *Nature* 543, 705–709. <https://doi.org/10.1038/nature21712>

Zhang, X., Zhao, L., Cheng, M., Liu, H., Wang, Z., Wu, X., Yu, H., 2018. Long-term changes in wet nitrogen and sulfur deposition in Nanjing. *Atmos. Environ.* 195, 104–111. <https://doi.org/10.1016/j.atmosenv.2018.09.048>

Zhao, S., Yu, Y., Qin, D., Yin, D., Dong, L., He, J., 2019. Analyses of regional pollution and transportation of PM 2.5 and ozone in the city clusters of Sichuan Basin, China. *Atmos. Pollut. Res.* 10, 374–385. <https://doi.org/10.1016/j.apr.2018.08.014>

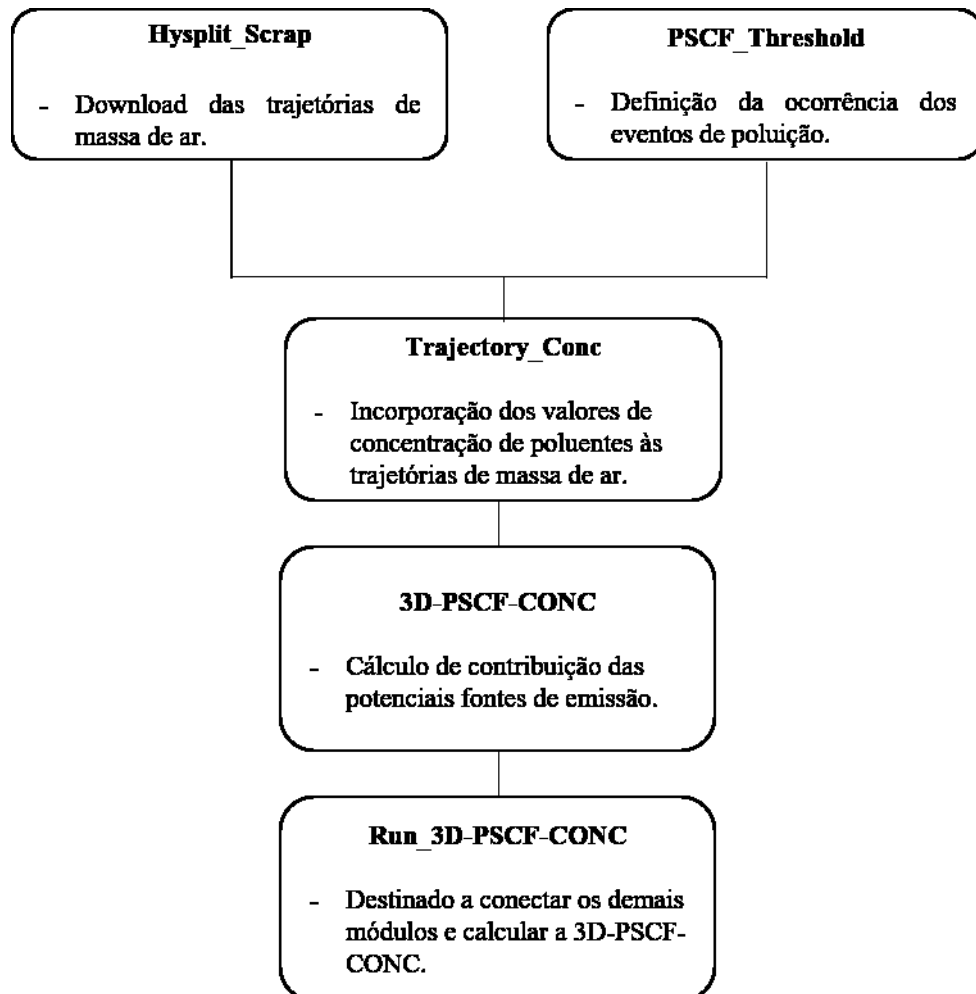
Zong, Z., Wang, X., Tian, C., Chen, Y., Fu, S., Qu, L., Ji, L., Li, J., Zhang, G., 2018. PMF and PSCF based source apportionment of PM<sub>2.5</sub> at a regional background site in North China. *Atmos. Res.* 203, 207–215. <https://doi.org/10.1016/j.atmosres.2017.12.013>





## APÊNDICE A: IMPLEMENTAÇÃO DO MÉTODO 3D-PSCF-CONC.

Figura A1. Resumo esquemático dos módulos utilizados para implementação do método 3D-PSCF-CONC.



## A1. Módulo Hysplit\_Scrap para automatização das requisições de trajetórias no site NOAA.

```

%% ----- Importação de bibliotecas
from selenium import webdriver
from selenium.webdriver.support.ui import WebDriverWait
from selenium.webdriver.support import expected_conditions as EC
from selenium.webdriver.common.by import By
from webdriver_manager.chrome import ChromeDriverManager
from bs4 import BeautifulSoup

%% ----- Download dos dados de trajetórias do site NOAA

class Hysplit_Scrap:

    def __init__(self, latitude, longitude, downloadFolder):
        self.latitude = latitude
        self.longitude = longitude
        self.downloadFolder = downloadFolder

    def runHysplit(self):
        webdriver = self.ConfiguracaoWebDriver()
        driverURL = self.Conexao_NOAA(webdriver)
        gdas = self.Obter_gdasFile(driverURL)
        self.Download_dados(gdas, driverURL)

    def ConfiguracaoWebDriver(self):
        chromeOptions = webdriver.ChromeOptions()
        chromeOptions.add_argument("--incognito")
        prefs = {"download.default_directory" : self.downloadFolder}
        chromeOptions.add_experimental_option("prefs",prefs)
        driver = webdriver.Chrome(ChromeDriverManager().install(),
        chrome_options=chromeOptions)
        return driver

    def Conexao_NOAA(self, driver):
        driver.get('https://www.ready.noaa.gov/hypub-bin/trajasrc.pl')
        lat = driver.find_element_by_id("LatId")
        lat.send_keys(self.latitude)
        lon = driver.find_element_by_id("LonId")
        lon.send_keys(self.longitude)

        driver.find_element_by_xpath("/html/body/div/table/tbody/tr/td/div[
4]/div[2]/form/div/div/table[2]/tbody/tr[1]/td[2]/select/option[2]"
).click()

        driver.find_element_by_xpath("/html/body/div/table/tbody/tr/td/div[
4]/div[2]/form/div/center/input[2]").click()
        return driver

    def Obter_gdasFile(self, driver):
        gdas = driver.find_element_by_name("mfile")
        mask = [word.strip().split(' ')[0] for word in
        list(gdas.text.split('\n'))]
        mask = mask[mask.index('gdas1.dec19.w5'):mask.index(
        'gdas1.feb19.w3')+1]
        mask.reverse()
        return mask

```

```

def Inputs_Hysplit(self, driver, ano, mes ,dia, meteoCheck):
    year = driver.find_element_by_name("Start year")
    year.send_keys(ano)
    month = driver.find_element_by_name("Start month")
    month.send_keys(mes)
    day = driver.find_element_by_name("Start day")
    day.send_keys(dia)
    hora = driver.find_element_by_name("Start hour")
    hora.send_keys("00") # trajetórias chegando ao local receptor
                        meia noite (00:00)
    duracao = driver.find_element_by_name("duration")
    duracao.clear()
    duracao.send_keys("72") # duracao das trajetórias de 72horas

    driver.find_element_by_xpath("/html/body/div/table/tbody/tr/td/div[
4]/form/div[3]/table[2]/tbody/tr[10]/td[3]/input").click()

    # --- Selecionar variáveis meteorológicas:
    if meteoCheck == True:

        driver.find_element_by_xpath("/html/body/div/table/tbody/tr/td/
div[4]/form/div[3]/table[2]/tbody/tr[25]/td[2]/input[1]").click
        ()
        driver.find_element_by_xpath("/html/body/div/table/tbody/tr/td/
div[4]/form/div[3]/table[2]/tbody/tr[25]/td[2]/input[2]").click
        ()
        driver.find_element_by_xpath("/html/body/div/table/tbody/tr/td/
div[4]/form/div[3]/table[2]/tbody/tr[25]/td[2]/input[3]").click
        ()
        driver.find_element_by_xpath("/html/body/div/table/tbody/tr/td/
div[4]/form/div[3]/table[2]/tbody/tr[25]/td[2]/input[4]").click
        ()
        driver.find_element_by_xpath("/html/body/div/table/tbody/tr/td/
div[4]/form/div[3]/table[2]/tbody/tr[25]/td[2]/input[5]").click
        ()
        driver.find_element_by_xpath("/html/body/div/table/tbody/tr/td/
div[4]/form/div[3]/table[2]/tbody/tr[25]/td[2]/input[6]").click
        ()
        driver.find_element_by_xpath("/html/body/div/table/tbody/tr/td/
div[4]/form/div[3]/table[2]/tbody/tr[25]/td[2]/input[7]").click
        ()

    driver.find_element_by_xpath("/html/body/div/table/tbody/tr/td/div[
4]/form/table/tbody/tr/td/input").click()

def Outputs_Hysplit(self, driver, ano, mes ,dias, meteoCheck):
    self.Inputs_Hysplit(driver, ano, mes ,dias, meteoCheck)
    objetoXpath = "/html/body/div/table/tbody/tr/td/div[4]/table[1]
/tbody/tr/td/font/ul/li[2]/b/a"
    outputFile = EC.presence_of_element_located((By.XPATH,
objetoXpath))
    WebDriverWait(driver, 60).until(outputFile)
    outputFile = driver.find_element_by_xpath(objetoXpath).
get_attribute("href")

    file = "http://www.ready.noaa.gov"+outputFile.split('"')[1]

```

```

driver.execute_script("window.open('');")
driver.switch_to.window(driver.window_handles[1])
driver.get(file)
content = driver.page_source
soup = BeautifulSoup(content, "html.parser")

f = open(self.downloadFolder+'/'+ano+mes+dias+'.txt', 'a')
f.write(soup.text)
f.close()

driver.close()
driver.switch_to.window(driver.window_handles[0])
driver.back()

def Download_dados(self, datesWeek, driver):
    for files in datesWeek:
        meteoCheck = True
        print('\n'+files)
        gdas = driver.find_element_by_name("mfile")
        gdas.send_keys(files)

        driver.find_element_by_xpath("/html/body/div/table/tbody/tr/td/
div[4]/form/center/input").click()

        driver.find_element_by_xpath("/html/body/div/table/tbody/tr/td/
div[4]/form/div[2]/table[1]/tbody/tr[2]/
td[2]/input").click()

        dias = driver.find_element_by_name("Start day")
        dias = [word.strip().split(' ')[0] for word in
list(dias.text.split('\n'))]

        ano = ['20'+driver.find_element_by_name("Start
year").get_attribute("value")]*len(dias)

        mes = [driver.find_element_by_name("Start month").
get_attribute("value")]*len(dias)

        for idx in range(0, len(dias)):
            self.Outputs_Hysplit(driver, ano[idx], mes[idx]
, dias[idx], meteoCheck)

            print(dias[idx]+'/'+mes[idx]+'/'+ano[idx])
            meteoCheck = False
        driver.back()

```



## A2. Módulo PSCF\_Threshold para definição da ocorrência de eventos poluídos.

```

#%% ----- Importação de bibliotecas
import glob
import pandas as pd
import numpy as np

#%% ----- Definição da ocorrência de eventos poluídos

class PSCF_Threshold:

    def __init__(self, file, pol):
        self.file = file
        self.pol = pol

    def Threshold_date(self):

        # ---- Upload dos dados
        dados = pd.concat([pd.read_excel(file) for file in
                           glob.glob(self.file)], ignore_index = True)
        dados = dados.sort_values(by = 'Data')

        dados = dados[['Data', self.pol]]
        dados = dados.set_index('Data')
        dados_daily = dados.resample('D').mean() # Threshold definido como
                                                a média diária do período

        # ---- Definição do threshold para eventos poluídos
        eventosPoluidos = dados_daily[dados_daily[self.pol] >
                                       dados_daily[self.pol].mean()]

        eventosPoluidos = dados_daily.loc[eventosPoluidos.index]
        eventosPoluidos['Date'] = [str(data.year)+str(data.month).
                                   rjust(2, '0')+str(data.day).rjust(2, '0')
                                   for data in eventosPoluidos.index]

        datas = eventosPoluidos.groupby('Date').agg({'Date': 'first'})
        return datas

```

### A3. Módulo Trajectory\_Conc para cálculo da concentração de poluentes nas trajetórias de massa de ar.

```

%% ----- Importação de bibliotecas
import os
import glob
import pandas as pd
import numpy as np
import xarray as xr
import geopandas as gpd
from termcolor import colored
from shapely.geometry import Polygon
import shapely.vectorized as sv
import rioarray
from shapely.geometry import mapping
from datetime import datetime, timedelta

%% ----- Cálculo de concentração nas trajetórias de massa
de ar

class Trajectory_Conc:

    def __init__(self, pathMERRA, rs, pathTrajs, pol, typeCol, dateIn,
                 dateFin):

        self.pathMERRA = pathMERRA
        self.pathTrajs = pathTrajs
        self.dateIn = dateIn
        self.dateFin = dateFin
        self.rs = rs
        self.pol = pol
        self.typeCol = typeCol

# --- Checagem dos dados de trajetória

    def ChecagemArquivos(self):

        def Check1():

            idx = pd.DataFrame(index = pd.date_range(self.dateIn+'-01-01
                                                    00:00:00',self.dateFin+'-12-31 00:00:00'))

            trajDays = [traj.split('.')[0] for traj in
                        os.listdir(self.pathTrajs)]

            trajDay = pd.DataFrame(index = pd.to_datetime(trajDays))
            trajDay['mask'] = 'datas ok!!!'
            validacao = idx.join(trajDay)
            validacao = validacao.reset_index()

            if len(np.where(validacao.isna())[0]) > 0:
                missVal = validacao.loc[np.where(validacao.isna())[0].item()]['index']

            if missVal: print(colored('\nDados faltantes: %s'%missVal),
                               'red')

```



```

        value = input("\nOs arquivos HYSPLIT foram corretamente
                      baixados? (s, n): ")

        if value == 'não': print('\nBaixar as trajetórias dos dias
                                indicados acima.')
    else: print(colored('\nTodas as trajetórias foram baixadas
                        corretamente!!!', 'green'))
    return trajDays

def Check2():
    cols = {2:'ano', 3:'mes', 4:'dia', 5:'hora', 8:'horaRev',
           9:'latitude', 10:'longitude', 11: 'Starting leve (m)',
           12: 'Pressure (kPa)', 13: 'THETA', 14: 'AIR_TEMP (K)',
           15: 'RAINFALL (mm/h)', 16:'MIXDEPTH (m)', 17:'RELHUMID
           (%)', 18:'TERR_HEIGHT (m)', 19:'SUN_FLUX (W/m²)'}

    files = []
    for file in os.listdir(self.pathTrajs):
        trajFile = pd.read_fwf(self.pathTrajs+'/'+file,
                               header = None, skiprows=
                               len(pd.read_fwf(self.pathTrajs+'/'+file))-72,
                               usecols = cols.keys(), names = cols.values())

        dateFormat = ['20'+str(ano)+str(mes).
                      rjust(2, '0')+str(dia).
                      rjust(2, '0')+str(hora).rjust(2, '0')
                      for ano, mes, dia, hora in zip(trajFile.ano,
                                                      trajFile.mes, trajFile.dia, trajFile.hora)]

        trajFile['dateFormat'] = dateFormat

        files.append(trajFile)

    datas = ['20'+str(date.ano[0])+str(date.mes[0]).
             rjust(2, '0')+str(date.dia[0]).rjust(2, '0')
             for date in files]

    datasRep = list(set([int(x) for x in datas if
                        datas.count(x) > 1]))

    if len(datasRep) > 0:
        print('\nVerificar arquivos, datas estão duplicadas em: ')
        datasRep.sort()
        for p in datasRep: print(colored(p, 'red'))
    return datas, files

def Check3(trajDays, datas):
    if trajDays==datas: print(colored('\nValidação concluída!!! ',
                                       'green'))

trajDays = Check1()
datas, files = Check2()
Check3(trajDays, datas)
return files

```

```

# --- Definição do domínio de modelagem com base no alcance
#       do conjunto de trajetórias

def Coords_MERRA2(self, traajs_list):
    traajs = pd.concat(traajs_list)
    xmin = min(traajs.longitude); xmax = max(traajs.longitude)
    ymin = min(traajs.latitude); ymax = max(traajs.latitude)
    return xmin, xmax, ymin, ymax

# --- Definição dos 4 arquivos .nc do MERRA-2
#       para casa pto de trajetória (73h)

def Extract_MERRA2(self, datas, band):

    file_0H = datas
    file_0H='MERRA2/'+self.rs+'/'+self.pol+'/MERRA2_400.tavg1_2d_'+
            self.typeCol+'_Nx.'+file_0H+'.nc4.nc4'
    day_0h = xr.open_dataset(file_0H)[band][:]
    print(file_0H)

    file_24H = (pd.to_datetime(datas) - timedelta(hours=24)).
                strftime('%Y%m%d')
    file_24H = 'MERRA2/'+self.rs+'/'+self.pol+'/MERRA2_400.tavg1_2d_'+
            self.typeCol+'_Nx.'+file_24H+'.nc4.nc4'
    day_24h = xr.open_dataset(file_24H)[band][:]
    print(file_24H)

    file_48H = (pd.to_datetime(datas) - timedelta(hours=48)).
                strftime('%Y%m%d')
    file_48H = 'MERRA-2/'+self.rs+'/'+self.pol+'/MERRA2_400.tavg1_2d_'+
            self.typeCol+'_Nx.'+file_48H+'.nc4.nc4'
    day_48h = xr.open_dataset(file_48H)[band][:]
    print(file_48H)

    file_72H = (pd.to_datetime(datas) - timedelta(hours=72)).
                strftime('%Y%m%d')
    file_72H = 'MERRA-2/'+self.rs+'/'+self.pol+'/MERRA2_400.tavg1_2d_'+
            self.typeCol+'_Nx.'+file_72H+'.nc4.nc4'
    day_72h = xr.open_dataset(file_72H)[band][:]
    print(file_72H)

    mergedConc = xr.merge([day_0h, day_24h, day_48h, day_72h])
    mergedConc = mergedConc[band][:73]
    return mergedConc

# --- Utilização dos campos de concentração de poluentes MERRA-2

def Grid_MERRA2(self, band):

    file = glob.glob(self.pathMERRA+self.rs+'/'+self.pol+'/*.nc4')[0]
    nc4_files = xr.open_dataset(file)
    nc4_files.rio.set_spatial_dims(x_dim="lon", y_dim="lat",
                                  inplace=True)
    nc4_files.rio.write_crs("epsg:4326", inplace=True)

    trajet = self.ChecagemArquivos().copy()

```

```

trajet = pd.concat(trajet, ignore_index = True)
trajet = gpd.GeoDataFrame(trajet, crs = 'EPSG:4326',
                        geometry=gpd.points_from_xy(trajet.
                        longitude, trajet.latitude))

clipped = nc4_files.rio.clip(trajet.geometry.apply(mapping),
                        trajet.crs, drop=True)

co = np.array(clipped[band][0]).flatten()
lon, lat = np.meshgrid(clipped['lon'], clipped['lat'])
lon = lon.flatten(); lat = lat.flatten()
coords = pd.DataFrame({self.pol:co, 'lon':lon, 'lat':lat}).dropna()

spatialX = 0.625/2; spatialY = 0.5/2
maskGeoMerra = gpd.GeoDataFrame(geometry = [Polygon([(x - spatialX,
                                                    y - spatialY),
                                                    (x - spatialX, y + spatialY),
                                                    (x + spatialX, y + spatialY),
                                                    (x + spatialX, y - spatialY)])
                                            for x, y in zip (coords.lon,
                                                            coords.lat)]), crs = 'EPSG:4326')

return maskGeoMerra

# --- Separando trajetórias totais das associadas a eventos de poluição

def Trajetorias_Threshold(self, traj_n, eventosPoluidos):
    dates = pd.DataFrame(index = [pd.to_datetime(traj.split('.')[0])
                                for traj in os.listdir(self.pathTrajs)])
    mask = dates.join(eventosPoluidos).reset_index().dropna()
    trajs_Thres = [traj_n[idx] for idx in list(mask.index)]
    return trajs_Thres

# --- Cálculo da concentração de poluentes nas trajetórias

def fromMERRA_toTrajs(self, trajs_list, band):
    dados = []

    for idx, traj in enumerate(trajs_list):
        datas = '20'+str(traj.ano[0])+str(traj.mes[0]).
                rjust(2, '0')+str(traj.dia[0]).rjust(2, '0')
        print(colored('\nCarregando Trajetória %s de %s: %s' %(idx,
                                                                len(trajs_list),
                                                                datas), 'red'))

        merra = self.Extract_MERRA2(datas, band)
        traj = traj.reset_index(drop = True)
        mask = [np.array(merra[i].sel(lat = traj.latitude[i],
                                    lon=traj.longitude[i], method="nearest")) for
                i in range(0, len(traj))]

        traj['Concentration'] = mask
        traj['Concentration'] = traj['Concentration'].astype(float)
        dados.append(traj)

    return dados

```



```

grid_n = gpd.GeoDataFrame(grid_n).
    rename(columns = {'index':'Residence_n',
                      'Concentration':'Concentration_n'}).
    reset_index(drop = True)

# ----- Cálculo do tempo de residência do conjunto de trajetórias
#          associado a eventos de poluição sobre as células da grade

grid_m = gpd.sjoin(gridMERRA, traj_mTotal)

grid_m['hRevs'] = [(datetime.strptime(dates, '%Y%m%d%H') -
                    timedelta(hours = (2*int(hour) - int(hours
Rev))))).strftime('%Y%m%d%H') for dates, hoursRev,
hour in zip(grid_m.dateFormat, grid_m.horaRev,
grid_m.hora)]

traj_hRev = pd.concat(traj_n)
gridRev = traj_hRev[traj_hRev['dateFormat'].
                    isin(grid_m['hRevs'])].loc[0]

gridRev= gridRev.rename(columns = {'dateFormat':'hRevs'})

grid_m = pd.merge(grid_m, gridRev, how = 'left',
                  on = 'hRevs', indicator=True).
                  set_index(grid_m.index)

grid_m.reset_index(inplace = True)
grid_m = grid_m.groupby('index').agg({'RAINFALL (mm/h)_x': 'mean',
                                     'index': 'count',
                                     'Concentration_x': 'mean',
                                     'Concentration_y': 'mean',
                                     'geometry': 'first'})

grid_m = gpd.GeoDataFrame(grid_m).
    rename(columns = {'index':'Residence_m',
                      'Concentration_x':'Concentration_m',
                      'Concentration_y':'Concentration_rs',
                      'RAINFALL (mm/h)_x':'RAINFALL (mm/h)_m'}).
    reset_index(drop = True)

# ----- PSCF-3D-CONC
pscf3DCONC = pd.merge(grid_n, grid_m, on = 'geometry')

pscf3DCONC['W(i,j)'] = 0.0

# ----- Weighted_func

n_ave = np.mean(grid_n['Residence_n'])
for idx in range(0, len(pscf3DCONC)):

    if pscf3DCONC['Residence_n'][idx] > 3*n_ave:
        pscf3DCONC['W(i,j)'][idx] = 1
    if 3*n_ave > pscf3DCONC['Residence_n'][idx] > 1.5*n_ave:
        pscf3DCONC['W(i,j)'][idx] = 0.7
    if 1.5*n_ave > pscf3DCONC['Residence_n'][idx] > n_ave:
        pscf3DCONC['W(i,j)'][idx] = 0.4

```

```
if n_ave > pscf3DCONC['Residence_n'][idx]:
    pscf3DCONC['W(i,j)'][idx] = 0.1

pscf3DCONC['Concentration_m'] = np.where(pscf3DCONC['Concentration_m'] > pscf3DCONC['Concentration_rs'],
    pscf3DCONC['Concentration_rs'],
    pscf3DCONC['Concentration_m'])

pscf3DCONC['3D-PSCF-CONC'] = ((pscf3DCONC['Residence_m']*pscf3DCONC['Concentration_m'])/
    (pscf3DCONC['Residence_n']*pscf3DCONC['Concentration_rs']))*pscf3DCONC['W(i,j)']

print(colored('*** PROCESSO FINALIZADO: '+id_pscf+ ' ***', 'red'))

return pscf3DCONC
```

#### A4. Cálculo de contribuição pelo método 3D-PSCF-CONC para RMSP.

```

### ----- Importação dos módulos
from webScraping import Hysplit_Scrap
from trajConc import Trajectory_Conc
from PSCF_METHOD import Method_3DPSCF_CONC
from Threshold_analysis import PSCF_Threshold

### ----- Importação da Classe PSCF_Threshold

# ---- Definição dos eventos de poluição

threshold_date = PSCF_Threshold('CETESB/co.csv', 'CO').Threshold_date()

### ----- Importação da Classe Hysplit_Scrap

# ---- Dados de trajetória convergindo para Marginal Tietê-Ponte Remédios
webScrap = Hysplit_Scrap("-23.518", "-46.743", 'trajetoriasSP')
webScrap.runHysplit()

### ----- Importação da Classe Trajectory_Conc

dadosTraj = Trajectory_Conc('MERRA-2/', 'SP', 'trajetoriasSP', 'CO', 'chm',
'2015', '2019')

# --- Checagem do download das trajetórias
traj_n = dadosTraj.ChecagemArquivos()

# --- Definição das dimensões do domínio do campo de concentração (MERRA-2)
coordsMERRA = dadosTraj.Coords_MERRA2(traj_n) # xmin, xmax, ymin, ymax
gridMERRA = dadosTraj.Grid_MERRA2('COSC')

# --- Cálculo de concentração de poluentes nas trajetórias de massa de ar
traj_n = dadosTraj.fromMERRA_toTrajs(traj_n, 'COSC')
traj_m = dadosTraj.Trajetorias_Threshold(traj_n, threshold_date)

### ----- Importação da Classe Method_3DPSCF_CONC

contribuicoes = Method_3DPSCF_CONC(traj_n, traj_m, gridMERRA)

# --- Selecionando as trajetórias a partir da altura de injeção de poluen-
tes
traj_n3D, traj_m3D = contribuicoes.Analisy_3D()

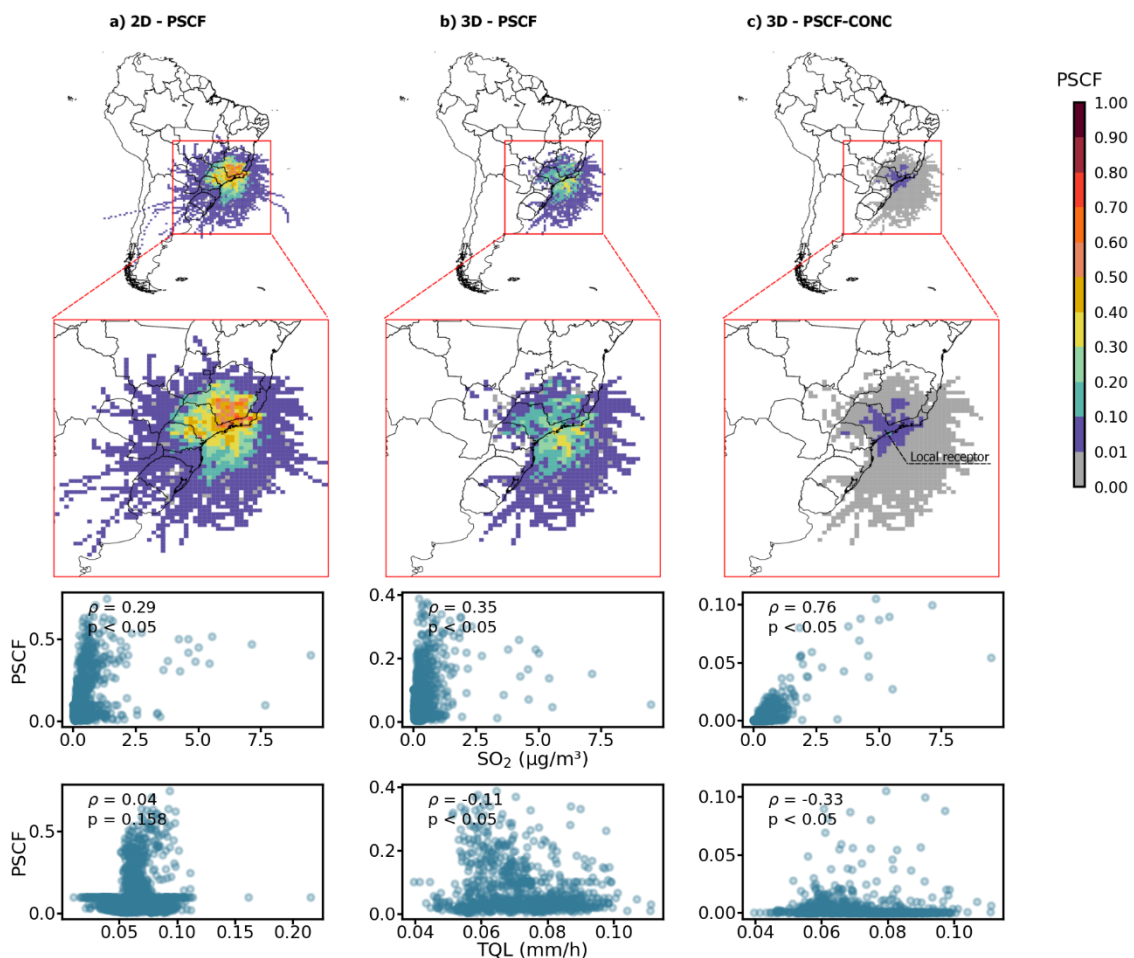
# --- Calculando as contribuições pelo método 3D-PSCF-CONC
pscf3DCONC = contribuicoes.PSCF_3DCONC(traj_n3D, traj_m3D)

```

## APÊNDICE B: COMPARAÇÃO DOS MÉTODOS PSCF NA RMSP - SO<sub>2</sub>

Os métodos 2D-PSCF e 3D-PSCF tendem a superestimar a contribuição de regiões distantes da RMSP quando avaliadas as concentrações de SO<sub>2</sub> (Figura B). O primeiro método identificou regiões em Minas Gerais e Rio de Janeiro contendo os maiores valores de PSCF (Figura B.a). O método 3D-PSCF reduziu a influência destas regiões, porém atribuiu às emissões oceânicas a principal contribuição de SO<sub>2</sub> (Figura B.b). Através do ajuste no cálculo de probabilidade, a 3D-PSCF-CONC identificou na Baixada Santista, localizada no litoral sul de São Paulo, a principal contribuição regional (Figura B.c). No Estado de São Paulo, destacam-se algumas áreas críticas em termos de poluição do ar, especialmente os polos industriais, que possuem relevante emissão de SO<sub>2</sub> (Nakazato et al., 2015).

**Figura B.** Comparação entre os métodos PSCF para fontes de SO<sub>2</sub> na RMSP: a) 2D-PSCF, b) 3D-PSCF, c) 3D-PSCF-CONC. Correlação entre os valores de PSCF e concentração de SO<sub>2</sub> ( $\mu\text{g}\cdot\text{m}^{-3}$ ) do MERRA-2. Correlação entre os valores de PSCF e precipitação média ( $\text{mm}\cdot\text{h}^{-1}$ ).



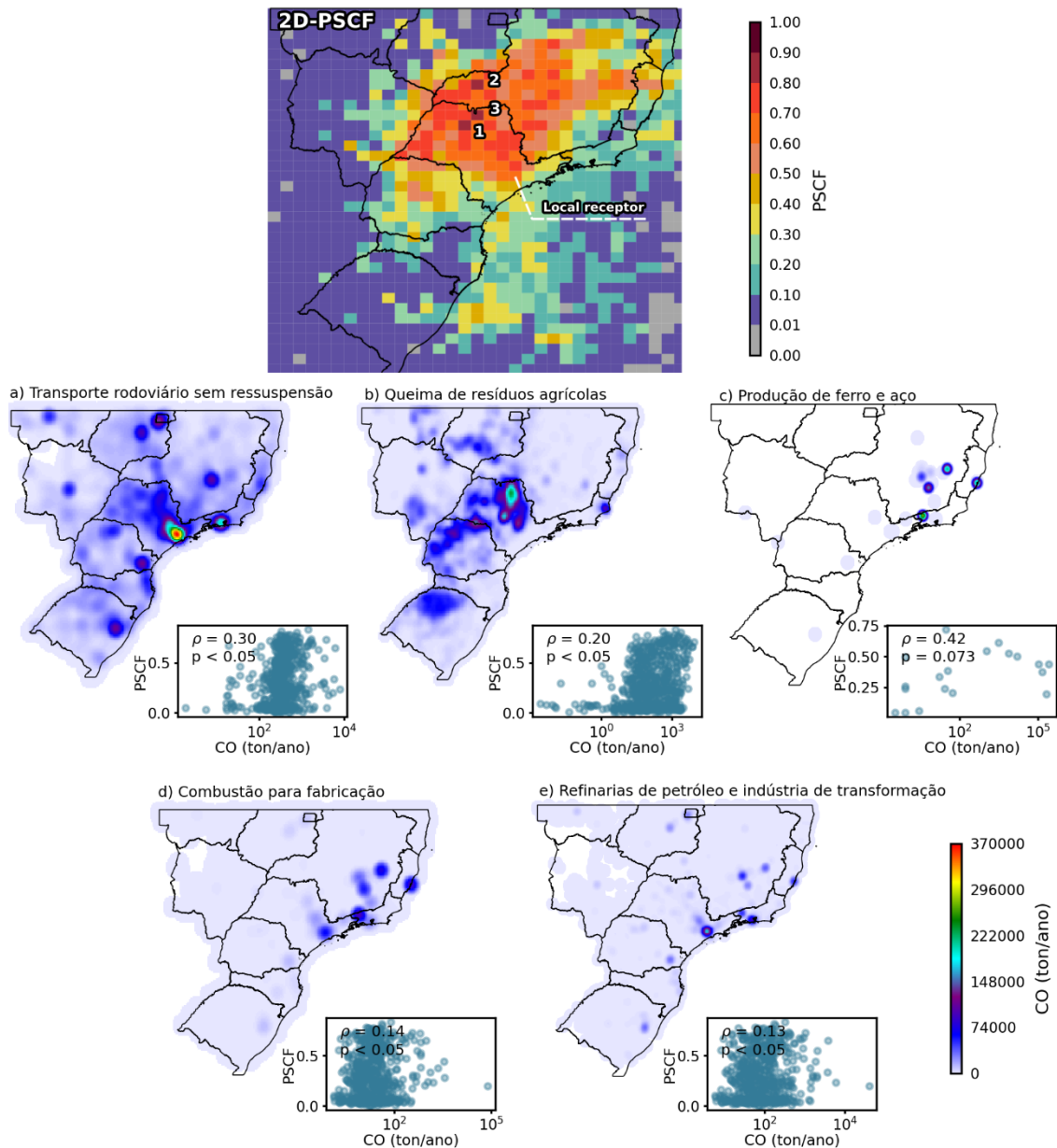


A CETESB possui seis estações de monitoramento da qualidade do ar instaladas na Baixada Santista, sendo elas Cubatão - Centro, Cubatão – Vale do Mogi, Cubatão - Vila Paris, Santos, Santos - Ponta da Praia e Guarujá - Vicente de Carvalho. Para o período de modelagem do estudo, 2015 a 2019, as estações localizadas em Cubatão apresentaram as maiores concentrações média diárias e médias anuais de SO<sub>2</sub>, dentre todas as estações monitoradas pela CETESB no estado de São Paulo (CETESB, 2015; CETESB 2016; CETESB 2017; CETESB 2018; CETESB 2019). A refinaria de petróleo está entre as indústrias mais poluentes da região de Cubatão (CETESB, 2012), emitindo teores de SO<sub>2</sub> superiores às demais fontes industriais, devido aos seus processos de refino de petróleo e também ao seu sistema de geração de energia e vapor.

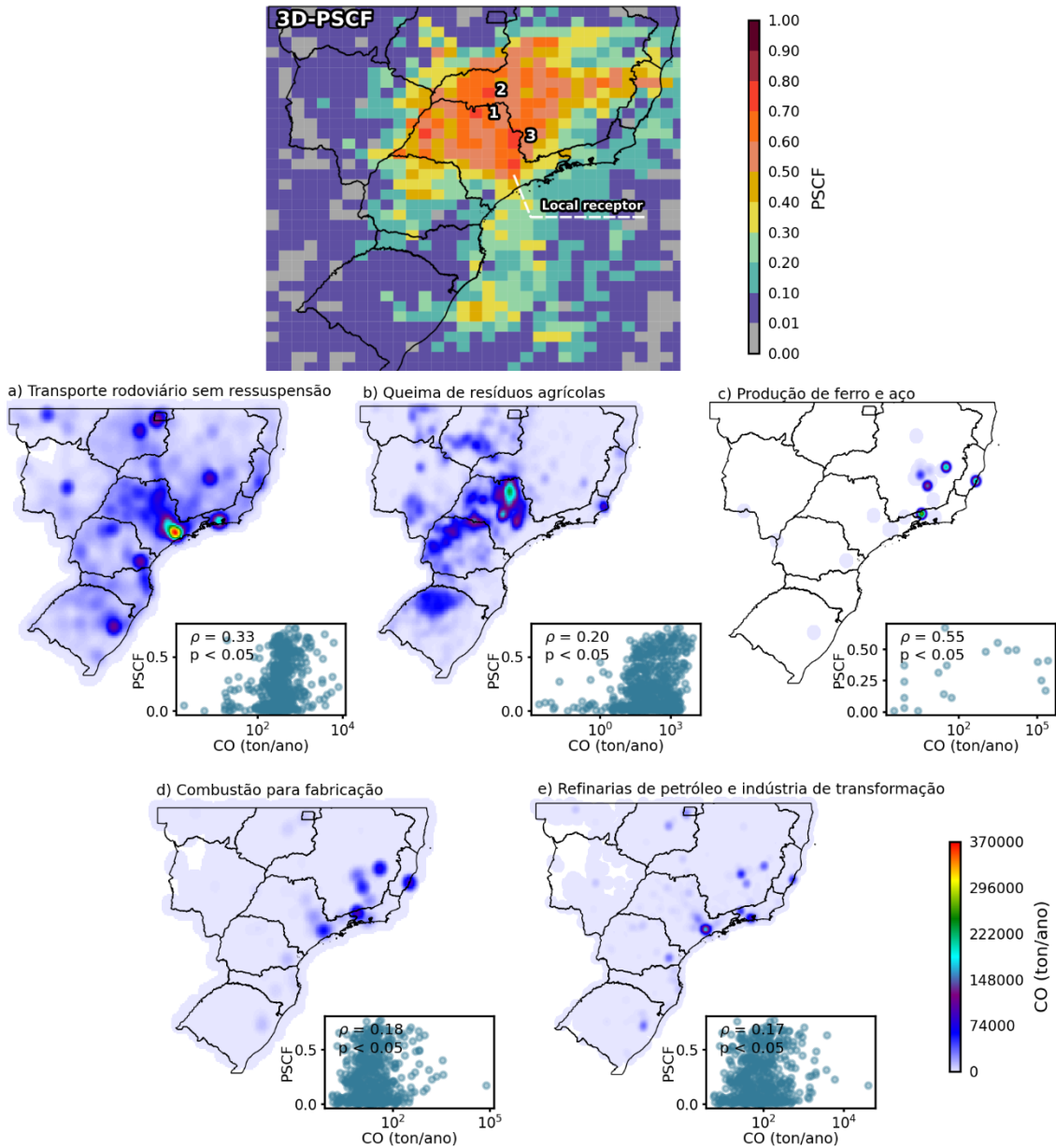
Além de refinar a identificação da localização das fontes de emissão, os resultados sugerem que a 3D-PSCF-CONC pode representar melhor a contribuição de áreas menos poluídas, bem como a influência da deposição úmida. O método apresentou as correlações mais significativas tanto entre a concentração de SO<sub>2</sub> e valores de PSCF ( $\rho = 0,74$  p <0,05), quanto entre a precipitação e valores de PSCF ( $\rho = -0,53$  p <0,05).

## APÊNDICE C: ANÁLISE ESPACIAL ENTRE O INVENTÁRIO DE EMISSÕES EDGAR E OS MÉTODOS PSCF CONVENCIONAS.

**Figura C1.** Correlação entre o inventário de emissões EDGAR e o método 2D-PSCF. Correlação de Spearman entre a distribuição espacial dos valores 2D-PSCF e as emissões de CO do transporte rodoviário sem ressuspensão do solo, queima de resíduos agrícolas, produção de ferro e aço, combustão para manufatura e refinarias de petróleo e indústria de transformação.



**Figura C2.** Correlação entre o inventário de emissões EDGAR e o método 3D-PSCF. Correlação de Spearman entre a distribuição espacial dos valores 3D-PSCF e as emissões de CO do transporte rodoviário sem ressuspensão do solo, queima de resíduos agrícolas, produção de ferro e aço, combustão para manufatura e refinarias de petróleo e indústria de transformação.



## APÊNDICE D: AVALIAÇÃO DA INFLUÊNCIA DA QUEIMA DE BIOMASSA SOBRE AS CONCENTRAÇÕES DE CO.

**Figura D1.** Análise temporal (a) e correlação de Spearman (b) das concentrações de CO e FRP na região de Barretos/SP.

