Improving automated emission quantification of TROPOMI methane plumes using Eulerian and Lagrangian models

SRON Netherlands Institute for Space Research

Pratik Rindhe

Delft University of Technology

SRON

TUDelft

Improving automated emission quantification of TROPOMI methane plumes using Eulerian and Lagrangian models

by

Pratik Rindhe

to obtain the degree of Master of Science in Aerospace Engineering, Aerodynamics and wind energy
at the Delft University of Technology,
to be defended publicly on Thursday July 14, 2022 at 13:30.

Student number:	5253268	
Project duration:	October 11, 2021 – July 14, 2022	
Thesis committee:	Dr. I. C. Dedoussi,	TU Delft, Thesis supervisor
	Dr. S. Naus,	SRON, Main supervisor
	Dr. J. D. Maasakkers,	SRON, Supervisor
	Prof. Dr. I. Aben,	SRON, Supervisor
	Prof. Dr. D. Simons,	TU Delft, Committee chair
	Prof. Dr. S.J. Hulshoff,	TU Delft, Committee member

An electronic version of this thesis is available at `http://repository.tudelft.nl/`.

Cover: Sentinel-5 Precursor satellite [1]

TU Delft

Preface

This report presents the research work done in the context of my Masters Thesis in Aerospace Engineering at the Delft University of Technology. I was a part of the TROPOMI CH_4 level 4 team as a thesis student at SRON Netherlands Institute for Space Research from 11th October 2021 to 14th July 2022. My thesis work started at SRON at a time when only hybrid work was allowed during the pandemic. At that time, I used to go to the SRON office in Leiden once or twice a week. I am grateful to the SRON management for excellently organising the hybrid mode of working. The biweekly online coffee meetings on Mondays were beneficial to get an idea about the things happening at SRON. I am also grateful to SRON for organising the Science Days conference in June 2022, where I got an extensive overview of the research happening at SRON from the scientists. I also want to thank the TU Delft administration and staff for providing academic and non-academic support during this project.

I want to thank my supervisors Stijn, Bram, and Ilse from SRON and Irene from TU Delft for providing me with this opportunity. I am also grateful to them for the detailed feedback they gave on my literature study report and the thesis report; without this feedback, it would not have been possible to achieve the level of the report I could reach today. I am thankful to Stijn for outlining this research's very clear and detailed structure before the start of this thesis. I am grateful to Stijn, Bram and Ilse for their excellent guidance and for providing their valuable feedback in the meetings, which gave me some positive constraints and helped me keep my research on track. I am also thankful to Irene for handling supervision from the TU Delft side and for giving valuable feedback after the monthly meetings, which helped in the improvement of this project. I am grateful to the chair of my thesis committee, Dr Dick Simons, for providing his suggestions during the kick-off, mid-term and green light meetings. I want to thank my thesis committee in advance for taking the time to read my thesis report.

Mainly, I want to thank my main supervisor Stijn for the close supervision of this project and excellent feedback after the meetings, which helped develop my communication abilities and research skills. Whenever I was stuck with a question, he never hesitated to answer it, which I am grateful for. I am also thankful to Stijn for spending his time and taking an extra step to provide me with the data and material required for this project.

I want to thank other researchers and colleagues I met at SRON during my internship as well as my thesis for creating a positive and healthy work environment that kept me motivated throughout this project, and a special mention to Pieter for initiating the coffee breaks: Berend, Elena, Fabrizio, Filippo, Gijs, Gourav, Ishita, Ivar, Lodewijck, Mitchell, Pankaj, Pieter, Sahana, and Sudhanshu. I am also thankful to Manu, who provided some ideas for the outline of this project before it started.

How can I not mention my roommate/chef Siddharth for tolerating my constant talks about what happened today at SRON while cooking in the kitchen together and especially for taking the initiative to cook tasty meals when I was busy with work, for which I am very grateful. I would like to thank Aravind for being a part of our after-dinner puzzle-solving sessions and giving me some tips about after-graduation life, which I will always keep in mind. I want to thank Srijit, without whom our weekends and cricket matches would have been boring. I would also like to thank my other six roommates, Julio, and Revanth, for tolerating my technical discussions and keeping a genuine interest in my project.

Finally, I would like to thank my parents, Anil and Kirti, and my brother Prasanna for giving me the emotional support I required and patiently waiting to see me in person for two long years.

Pratik Rindhe
Delft, July 2022

Abstract

Since the pre-industrial era, methane has contributed as much as 0.5°C to global warming. With a global warming potential of 28-34 times that of CO_2 over 100 years, and with 80 times the warming power of CO_2 in the first 20 years, methane has a large impact on climate change. Measurements from the satellite instrument TROPOMI help inform on the global atmospheric content of methane. We frequently observe methane plumes in TROPOMI data. These plumes come from super-emitters, and those provide a significant but uncertain contribution to global methane emissions, and they are relatively easy to mitigate. The TROPOMI automated plume detection algorithm makes plume detection and emission quantification possible around the world, which is infeasible without automation. However, the current TROPOMI plume emission quantification algorithm makes use of mass balance methods to estimate the methane emissions of point sources. These approaches provide reasonable but highly uncertain estimates. Hence, in this thesis, we aim to improve the accuracy of the automated emission quantification algorithm with the help of more sophisticated techniques based on atmospheric transport models.

First, this study analysed mass balance methods by quantifying emissions from synthetic plumes having known emission rates. Synthetic plumes were generated using the WRF and FLEXPART atmospheric transport models. A classification algorithm was developed to segregate synthetic plumes into different categories based on their geometries to determine challenging plume emission quantification scenarios for the mass balance methods. We found that mass balance methods produce uncertain emission estimates due to several inherent limitations like missing plume pixels, performance under low wind speed conditions, and no utilisation of three-dimensional wind speeds. Next, the atmospheric transport model based plume scaling approach was analysed by quantifying emissions from synthetic plumes. This analysis revealed that the plume scaling approach could overcome several inherent limitations of the mass balance methods and reduce the uncertainty of plume emission quantification by nearly 10%. Finally, the plume scaling approach was applied to TROPOMI plumes. The FLEXPART model was observed to be most suitable for replicating TROPOMI plumes. With the help of results obtained from this research, a decision tree algorithm was developed. This decision tree can choose the most suitable plume emission quantification method between the mass balance methods and the plume scaling approach for a given TROPOMI plume, maintaining a balance between ease of use of the mass balance methods and accuracy of the plume scaling approach.

Key Points

- Mass balance methods are uncertain due to several inherent limitations that include missing plume pixels, performance under low wind speed conditions, and a too simplified way of accounting for the transport of particles by using wind at one location at one moment in time.
- Atmospheric transport model based plume scaling approach can overcome several limitations of mass balance methods in order to improve emission quantification of TROPOMI plumes.
- Mass balance methods and atmospheric transport model based plume emission quantification approach can be used in a balanced way to improve emission quantification of TROPOMI plumes.

Keywords

Methane, CH4, TROPOMI, Plume emission quantification, Mass balance methods, Atmospheric transport models, FLEXPART, WRF-CHEM, Plume scaling approach

Nomenclature

Physics Constants

M_{CH_4}	molecular weight of methane	$16.043 \times 10^{-3}\,\mathrm{kg\,mol^{-1}}$
M_{air}	mean molecular weight of air	$28.96 \times 10^{-3}\,\mathrm{kg\,mol^{-1}}$
p_0	pressure at sea level	$1013 \times 10^{2}\,\mathrm{Pa}$

Other Symbols

α_1, α_2	coefficients of parametric wind speed equation for IME from the LES model
α_3, α_4	coefficients of parametric wind speed equation for IME from the WRF model
β	coefficient of parametric wind speed equation for CSF
$\Delta\Omega$	total vertical column mass
$\Delta\Omega_{air}$	weight of column of dry air
ΔX	enhancement in dry-air mole fraction of methane
η	vertical coordinate
$\hat{x}$	optimized emission values
$\mathbf{v}$	velocity vector
ρ_i	concentration in terms of mass density
σ	pressure coordinate
τ	residence time of methane
CO_2	carbon dioxide
T_1, T_2, T_3, T_4	threshold values
$XCH_{4model,i}$	model outputted methane dry air mole fraction
$XCH_{4prior,i}$	prior profile of methane vertical column
$XCH_{4total\ column}$	column-averaged methane dry air mole fraction
θ	solar zenith angle
θ_r	satellite viewing angle
A_j	area of pixel j
$A_{col,i}$	averaging kernel
F_i	mass flux
IME	integrated mass enhancement
K	Jacobian matrix
k	vertical level
L	length or extent of plume
p_k	pressure at vertical level k
p_s	pressure at the surface

p_{top}	pressure at topmost vertical level
Q	emission rate
Q_S	known emission rate of synthetic plume
Q_T	unknown emission rate of TROPOMI plume
S_A	prior covariance matrix
S_o	error covariance matrix
t	time
$U(x, y)$	wind speed at location (x, y)
U_{eff}	effective wind speed
U_{10}	10 [m] wind speed
$V_{\text{air,dry}}$	dry air column
X	measured dry-air mole fraction of methane
x_a	prior emission rate
X_b	background dry-air mole fraction of methane
X_S	mass enhancements in synthetic plume
X_T	mass enhancements in TROPOMI
z_s	local terrain elevation
u	velocity along x-direction
v	velocity along y-direction
w	velocity along z-direction

Abbreviations

CAMS Copernicus Atmosphere Monitoring Service

CFL Courant-Friedrichs-Lewy

CSF Cross Sectional Flux

ECMWF European Centre for Medium-Range Weather Forecasts

EPA United States Environmental Protection Agency

ESA European Space Agency

FLEXPART FLEXible PARTicle dispersion model

GFS Global Forecast System

GOSAT Greenhouse Gases Observing Satellite

GWP Global Warming Potential

IME Integrated Mass Enhancement

IPCC Intergovernmental Panel on Climate Change

JAXA Japanese Space Agency

KNMI Royal Netherlands Meteorological Institute

LES Large Eddy Simulation

LPDM Lagrangian Particle Dispersion Model

NCEP National Centers for Environmental Prediction

NIR Near Infrared

NOAA National Oceanic and Atmospheric Administration

NSO Netherlands Space Office

OMI Ozone Monitoring Instrument

PBL Planetary Boundary Layer

RMSE Root Mean Square Error

S5P Sentinel-5 Precursor

SCIAMACHY SCanning Imaging Absorption spectroMeter for Atmospheric CartograpHY

SRON Netherlands Institute for Space Research

SWIR Shortwave Infrared

TIR Thermal Infrared

TROPOMI TROPOspheric Monitoring Instrument

UNFCCC United Nations Framework Convention on Climate Change

UVNS Ultraviolet Near-infrared Shortwave-infrared

VIIRS Visible Infrared Imaging Radiometer Suite

WRF Weather Research and Forecasting Model

Contents

List of Figures

List of Tables

1

Introduction

According to the IPCC report, methane levels in the atmosphere today are higher than at any time in the past 800,000 years [2]. Since the pre-industrial era, methane has contributed to global warming as much as 0.5°C and is the second-largest contributor after CO_2. Methane also has a shorter atmospheric lifetime (12 years [3]) than CO_2 (∼100 years [3]); hence, reducing the methane concentrations in the atmosphere can provide short-term benefits in the global climate situation and it can be done at low cost [4].

Anthropogenic methane emissions contribute almost 60% to the global methane emissions [5]. Anthropogenic emissions occur from a variety of sources like oil and gas infrastructures, coal mines, landfills, agricultural activities, wastewater treatment, and some industrial processes [6]. Anthropogenic methane emissions can be regulated, and this gives an opportunity to mitigate a large part of the methane emissions. Mitigation of methane emissions requires accurate identification of emission locations and quantification of emission rates.

Satellites provide a way to collect high-level details over remote areas, which is difficult to achieve using any other method. Spectral data collected by satellite instruments like SCIAMACHY, OMI, TROPOMI, etc., have been used to detect atmospheric methane. In this study, we use the spectral data collected by the TROPOMI satellite instrument to trace atmospheric methane. TROPOMI has a daily global coverage and a spatial resolution of roughly 7 [km] x 5.5 [km]. Frequently, methane plumes can be seen in the TROPOMI images. Their first quantification estimates are usually obtained using simple mass balance approaches. The mass balance methods tend to be uncertain since they do not account for the three-dimensional transport of gas particles. Hence, there is a need for a more sophisticated quantification approach that could account for particle transport. To fill this research gap and improve the TROPOMI plume emission quantification algorithm, we plan to use the atmospheric transport models.

Atmospheric transport models account for the three-dimensional transport of the particles (or air parcels). There are two distinct categories of atmospheric transport models: Eulerian and Lagrangian. These are based on the frame of reference used for solving the transport equations. An Eulerian transport model uses a frame of reference fixed to the computational grid, whereas the Lagrangian transport model uses a frame of reference fixed to the transported particles. In this research, we use the WRF-CHEM model based on the Eulerian frame of reference and the FLEXPART model based on the Lagrangian frame of reference.

The entire research timeline is divided into four phases. The first phase focuses on analysing mass balance methods and identifying their limitations. To achieve this, we apply mass balance methods to the synthetic plumes with known emission rates generated by atmospheric transport models. This provides some benchmark results for later comparisons. The second and third phases involve quantifying synthetic plumes with atmospheric transport models and evaluating the quantification uncertainties. For quantifying plumes using atmospheric transport models, the plume scaling approach is considered in this study. Effects of using different meteorological data, different physics settings (planetary bound-

ary layer schemes), and different atmospheric transport models on the results of the plume scaling approach are investigated. Quantification results from the plume scaling approach are compared to mass balance methods to see whether plume emission quantification is improved or not. The final phase involves testing the plume scaling approach on real TROPOMI plumes. Based on the results from all four phases, a decision tree algorithm is proposed that selects the most suitable plume emission quantification method for a given TROPOMI plume. Finally, this study discusses the possibility of integrating an atmospheric transport model based plume emission quantification approach with the TROPOMI automated plume detection algorithm.

The structure of the thesis report is presented in Figure 1.1. This study starts with gathering the theoretical background required for this research which is presented in Chapter 2. A detailed research plan describing four phases of this study is presented in Chapter 3. In the first three phases of this study, we analyse mass balance methods and the plume scaling approach by quantifying synthetic plumes. The setup used for generating synthetic plumes is described in Chapter 4. Setups of plume emission quantification methods considered in this study are explained in Chapter 5. Some pre-analysis steps are required before assessing the plume emission quantification methods. These pre-analysis steps are described in Chapter 6. In the final phase, we test the plume emission quantification approaches on real TROPOMI plumes. Results obtained from all four phases of this study are provided in Chapter 7. Finally, conclusions and future recommendations from this study are discussed in chapters 8 and 9, respectively.

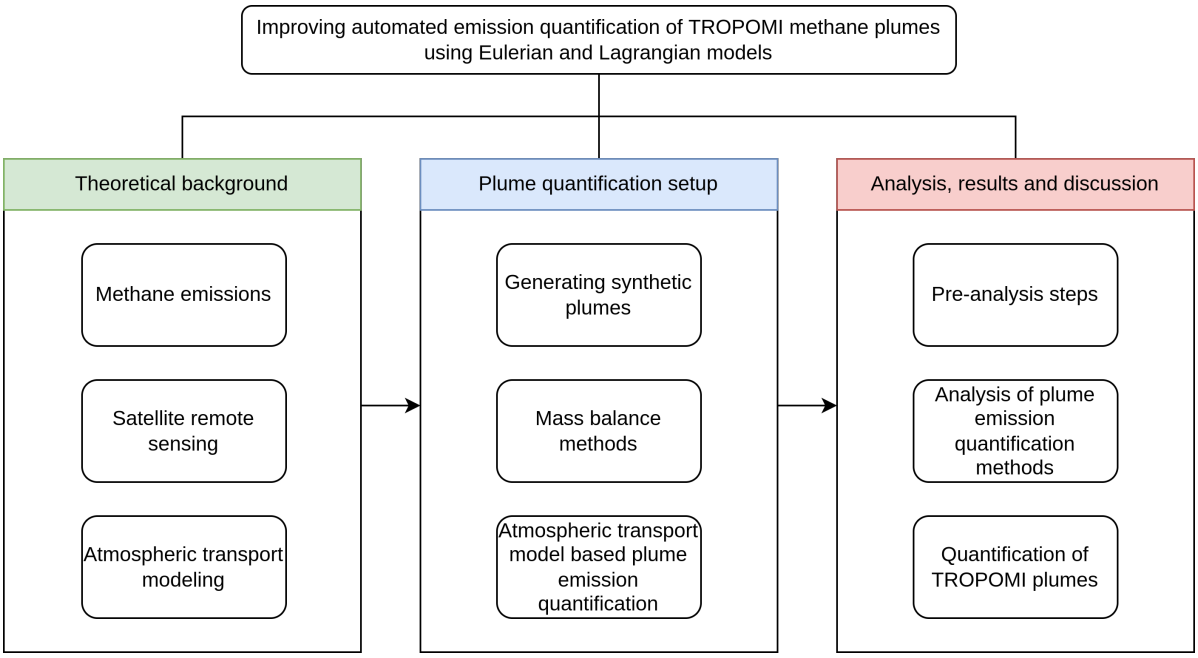


Figure 1.1: Structure of the thesis report. This report is divided into three major parts: theoretical background, plume emission quantification setup, and results and discussion. Theoretical background is provided in Chapter 2. Plume emission quantification setup is explained in Chapter 4 and Chapter 5. Analysis, and results and discussion are presented in Chapter 6 and Chapter 7, respectively.

2

Theoretical background

2.1. Global atmospheric methane concentrations

Figure 2.1 shows global monthly averaged atmospheric methane concentrations from the year 1983 to the year 2022. It can be seen that the methane concentrations in the atmosphere have risen over the past 30 years. The rise in the methane concentrations slowed down from the period 1999 to 2006, but thereafter, the concentrations rapidly increased.

Compared to carbon dioxide, methane can trap more amount of energy, and it has a shorter lifetime [2]. Therefore, methane has a global warming potential (GWP) of 28-36 over 100 years [2], which means methane can absorb 28-36 times more energy than carbon dioxide over a scale of nearly 100 years. Please note that methane's global warming potential accounts for indirect effects, such as methane's reaction with chemicals in the atmosphere to form ozone, an important greenhouse gas.

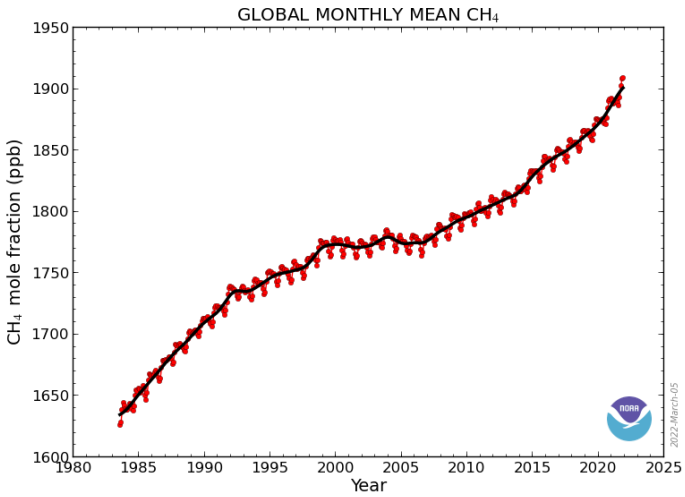


Figure 2.1: The graph represents the NOAA time-series of global monthly averaged atmospheric methane dry air mole fractions (in bold black line) from the year 1983 to the year 2022 [7]. NOAA has measured methane concentrations since 1983 at multiple air sampling sites distributed globally [8]. Methane concentrations are reported in terms of dry air mole fraction. The dry air mole fraction is the ratio of the number of methane molecules to the total number of molecules in the sample. The dry air mole fraction is given in units nmol/mol or particles per billion (ppb). Red dots represent globally and monthly averaged methane concentration values in the middle of each month. The bold black line represents the twelve-month mean without the effects of the average seasonal cycle. (Accessed: 19-03-2022)

The Paris agreement aims to limit the global temperature rise to well below 2°C compared to pre-industrial levels by 2050. According to an IPCC special report on the impacts of global warming published in 2018, the global temperature rise touched the 1°C mark in 2017. If this trend continues, the

1.5°C mark will be reached by 2040 [9]. Following the discussions presented in the previous and this paragraph, it is crucial to mitigate the atmospheric methane concentrations.

2.2. Sources of methane

Methane is emitted by various sources and broken down by various sinks. The most prominent sources of methane include agriculture, wetlands, fossil fuel production and use, biomass and biofuel burning and waste management, and other natural sources [10]. Some prominent methane sinks are atmospheric chemical reactions, and soil uptake [10]. The breakdown of the methane sources and sinks can be seen in Figure 2.2.

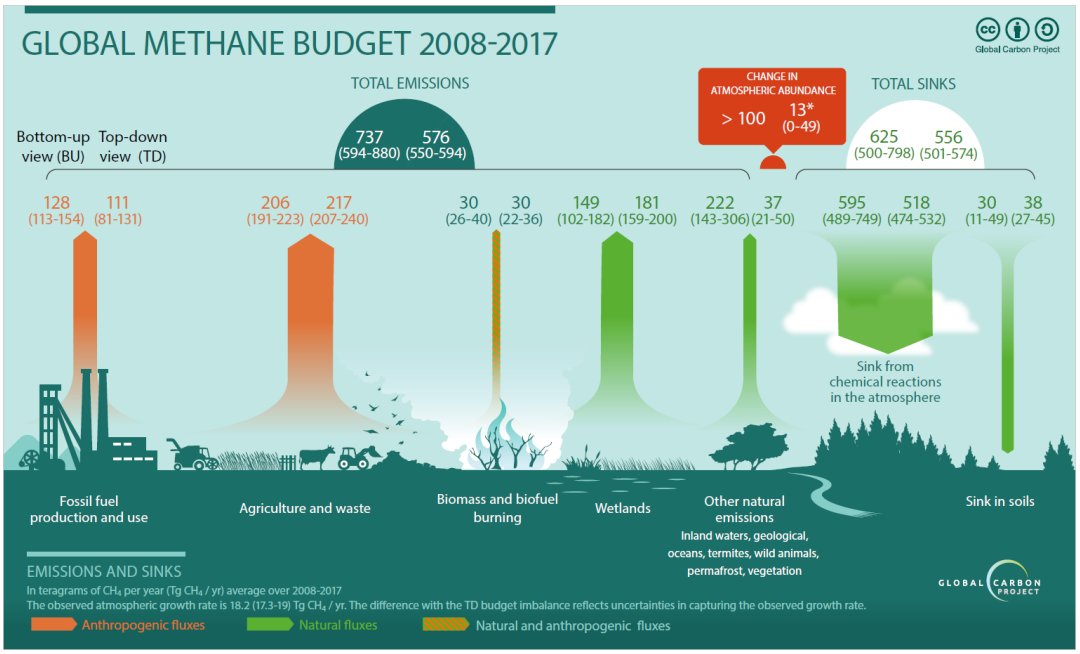


Figure 2.2: Global methane budget 2008-2017 depicting prominent methane sources and sinks [11]. Methane sources include fossil fuel production and use, agriculture and waste, biomass and biofuel burning, and wetlands. Methane sinks include chemical reactions in the atmosphere and uptake by soil. Methane sources are categorised into anthropogenic and natural sources. In the figure, anthropogenic methane sources are represented in red colour, and natural methane sources are represented in green colour. (Accessed: 19-03-2022)

The following subsections focus on emissions from the most prominent sources of methane. Sinks of atmospheric methane remain out of scope for this study and are not included in the following subsections.

2.2.1. Agriculture

The agricultural sector releases a large amount of methane into the atmosphere [12, 13, 14]. Methane is produced in agriculture by rice cultivation, ruminant livestock, manure management, and residue burning [15].

Methane is produced in rice cultivations due to the flooded soil, which creates a favourable environment for the anaerobic (oxygen-deprived) decomposition of organic matter. In ruminant livestock, methane is mostly produced by the micro-bacterial fermentation of food inside their stomach. Methane production from animal manures is also an enteric fermentation process, which happens mostly when the manures are stored. Incomplete combustion of residual biomass also produces methane, but compared to other agriculture sources, it is small in quantity [16].

Agriculture is the largest anthropogenic methane emission source. Only the ruminant livestock emissions account for approximately one-third of anthropogenic methane emissions. Due to the population growth, economic developments, change in eating habits, and rise in food demands, methane emissions from the agricultural sector have seen an unprecedented rise [2]. With the global population approaching 10 billion by 2050, it is expected that food demands will increase by 70 per cent [17].

There are ample opportunities available for mitigating methane emissions from the agricultural sector. Methane emissions from rice production can be reduced by changing fertilizer practices, residue management, keeping the soil dry in the off-rice season, and avoiding waterlogging [16]. Methane emissions from Ruminant livestock can be reduced by improving feeding practices, using dietary additives, and longer-term management changes [16]. Methane emissions from manure can be reduced by cooling and closing the storage tanks and by capturing the produced methane [16].

2.2.2. Waste

Methane emissions from waste largely occur from the oxygen-deprived decomposition of organic material at the solid waste disposal sites and wastewater handling plants [18]. The time needed for the decomposition of organic material depends on the conditions at the waste site. It can take less than a year or more than 100 years.

Methane emissions from waste are classified under anthropogenic emissions. Methane emissions from solid waste disposal in industrialized countries are showing a decreasing trend. On the other hand, methane emissions in developing countries are expected to rise due to the increase in landfill management practices and improved economy [19]. Good landfill management (often lacking in developing countries) can reduce methane emissions by a lot.

Municipal and industrial waste statistics have improved over the past decade in many countries. However, precisely monitoring and quantifying the methane emissions is challenging due to the dynamic nature of the waste disposal sites and changing waste management practices [20]. Scientists are working to fill the missing piece of knowledge about the dynamics of waste disposal sites.

2.2.3. Fossil fuel production and use

The fossil fuel sector is one of the major sources of methane emissions, accounting for nearly 40% of anthropogenic methane emissions worldwide. Methane emissions from fossil fuels occur during extraction of oil, natural gas and coal, but, depending on the fuel type, also during transport and usage [21].

Natural gas is mainly methane. During the extraction of oil, natural gas, and coal, methane is released due to fugitive emissions and leakages in the equipment [21]. Methane is released in the transportation of natural gas due to leakages in the compressor, pipelines, and storage sites [21]. During the use of natural gas, methane is emitted due to incomplete combustion [21]. Several abandoned wells of oil, natural gas, and coal emit methane because of improper sealing [21].

Strong economic growth in developing countries like India and China and an increase in the population are causing a substantial increase in energy demands [22]. At least till 2030, fossil fuels will remain the most important energy source, and their use is expected to grow [23].

It is estimated that nearly 70 per cent of the current emissions from oil and gas facilities are technically feasible to prevent. Around 45 per cent of the emissions could be mitigated at no extra cost by capturing the leaking gas because the cost of captured gas is higher than the cost of the measures taken. The percentage of captured gas is expected to rise in the future due to the increasing prices of fossil fuels. Therefore, limiting methane emissions from the oil and gas sectors can be cost-effective [4].

2.2.4. Biomass and biofuel burning

Biomass and biofuel burning is a small but significant source of atmospheric methane. Methane can be emitted due to incomplete combustion of woodlands, agricultural waste, and savanna [24]. Biomass and biofuel burning involves natural (e.g. lightning-induced) burning and anthropogenic periodic burning of woodlands and grassland to create land for agriculture and other activities.

Under ideal circumstances of complete combustion, biomass and biofuel burning produces CO_2 and water, but, in reality, biomass and biofuel burning happens with incomplete combustion [24]. Incomplete combustion of biomass and biofuel produces other carbon species, and one of them is methane [24].

Natural fires cause large biomass burning events, but anthropogenic biomass burning accounts for nearly 90% of the vegetation burned [25]. In the last 100 years, biomass and biofuel burning events

have increased on a global scale. It is predicted that rising atmospheric temperatures due to global warming will lead to an increase in biomass burning events [25].

Biomass burning happens in an uncontrolled way. To reduce the methane emissions from this source, the burning needs to be reduced. Some of the biomass burnings are required to retain the savanna and create agricultural lands. Currently, research work is going on to predict and detect biomass burning events.

2.2.5. Wetlands

Wetlands contribute to the Earth's natural water management processes. Wetlands contain a complex system of aquatic life, plants and soil that help in water cleansing. Wetlands are the largest natural and biogenic source of methane emissions contributing up to one-third of the total natural and anthropogenic methane emissions [26].

Methane is produced due to the anaerobic decomposition of plants and other organic matter by the soil microbes. In wetlands, the release of methane takes place through three known mechanisms: diffusion [27], aerenchyma [27], and ebullition [28]. Through diffusion, methane is transported through water bodies and soil to the atmosphere. Aerenchyma refers to the process of transport of gases through tubes in the tissues of a particular kind of plant. Ebullition is the release of large entrapped bubbles of methane into the atmosphere.

Wetlands are ecosystems that are very sensitive to climate change [29]. Wetlands can be affected by increased temperatures, storms, floods, precipitation, and droughts [29]. Moreover, due to the population growth and increase in economic developments, wetlands are converted into lands for infrastructure developments [30].

Understanding the circumstances under which methane is released from wetlands could give solutions to mitigate methane emissions from wetlands [30]. Due to the large area covered by wetlands, they are challenging to study. The biological distribution of methane-generating microbe communities in the wetlands and the metabolic processes that drive the generation of methane at the scale of the ecosystem are poorly known due to the lack of observational information [31]. Climate change also affects wetland methane emissions, but due to a lack of information and uncertainties, it is difficult to determine how the wetlands respond to the changing climate [29].

2.3. Detection and control of atmospheric methane

Anthropogenic methane sources contribute to almost 60% of the global methane emissions [5], which makes them a good target for mitigation. Anthropogenic methane sources like fossil fuel infrastructure and waste management facilities show room for cost-effective mitigation [4], as described in the Subsections 2.2.2 and 2.2.3. Therefore, this study will mainly refer to methane emissions from fossil fuel facilities and waste management.

To detect atmospheric concentrations of methane, there are many different approaches. These approaches are mainly distinguished into two categories: bottom-up approaches and top-down approaches. In general, bottom-up approaches involve the use of emission factors and production information from the methane emission sites. Top-down approaches rely heavily on the atmospheric information of methane concentrations, and they contain atmospheric transport.

To mitigate methane emissions from anthropogenic sources, two key parameters are required. These parameters are the location of emission and emission rate. In general, emission inventories could give information about the locations and under what circumstances most of the methane is emitted from the fossil fuel facilities and waste management. Using the bottom-up emission information, sometimes maps are created that provide a spatial representation of methane emissions [32]. These maps help monitor greenhouse gas emissions, and their data could inform on the expected locations of emissions.

There are several gaps in the bottom-up emission estimates. Under the United Nations Framework Convention on Climate Change (UNFCCC), some nations report sector-wise methane emissions. Most nations only provide an overall number for methane emissions for the entire nation, and sector-wise disaggregation of this data induces uncertainties [32]. In some cases, the bottom-up emission infor-

mation might not be recorded [32]. For example, in Algeria, not a lot of bottom-up methane emission measurements are available. There are some cases where the information is recorded but might be restricted to use. In addition, direct measurements are difficult at the locations where some oil and gas facilities are. With the limited and uncertain bottom-up information, it is difficult to get an integrated picture of methane emissions. Therefore, there is a need for independent constraints from a different approach. Top-down measurements using satellite instruments can be used to place some independent constraints and overcome many limitations of the challenging in-situ measurements.

Large anthropogenic methane emissions are frequently spotted in the measurements of satellite instruments. These emissions often originate from a large number of point sources (Box 2.1) spread around the surface of Earth [33].

Box 2.1: Point source

A methane point source is a single identifiable source of methane emission with a very less or negligible area on the Earth's surface. Large methane emissions from point sources are often observed in satellite measurements. Methane emissions from fossil fuel facilities and waste management are often considered point source emissions.

Quantification of methane emissions from point sources can help in identifying and improving the understanding of methane sources. Currently, there are some challenges in accurately quantifying methane emissions seen in satellite instrument measurements. Specifications of these challenges are discussed in the following sections.

2.4. Measuring atmospheric methane concentrations from space

Some part of the solar radiation hitting the Earth's surface is reflected in the atmosphere. This reflected radiation is called solar backscatter. Earth's surface also emits radiation. Certain gases absorb infrared radiation from the solar spectrum (incoming and backscattered radiation) and the Earth emission spectrum. These gases are known as greenhouse gases. Methane is a greenhouse gas, and it absorbs some wavelengths of infrared radiation.

Figure 2.3 represents optical depth of methane (red line) against several wavelengths. Optical depth describes the amount of light absorbed when it travels through an absorbing medium. Atmospheric methane absorbs radiation belonging to shortwave infrared (SWIR) wavelengths 1.65 [μm] and 2.3 [μm], and thermal infrared (TIR, not represented in the Figure 2.3) wavelength 8 [μm] [34]. This absorption feature of methane is used in the measurements done by satellite instruments [35].

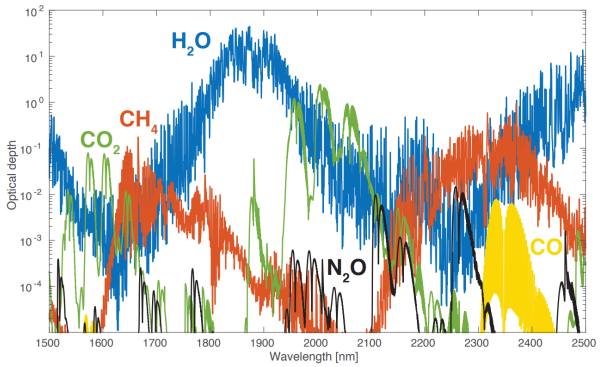


Figure 2.3: This figure represents optical depth of various gases for different wavelengths. The spectral resolution of this data is 0.1 nm. Optical depth for methane is marked by red line. Optical depth is unit less, wavelength is given in nm. This experiment was performed when the surface concentration of methane was 1.9 ppm. [34]

Figure 2.4 shows shortwave infrared and thermal infrared radiations' sensitivities to methane as a function of atmospheric pressure for a typical satellite instrument. Methane emissions from sources like fossil fuel facilities and waste management reside mostly in the troposphere; hence uniform satellite instrument sensitivity in the troposphere is desirable. From Figure 2.4, shortwave infrared measure-

ments are uniformly sensitive to methane in the troposphere; hence, shortwave infrared radiations are more suitable for measuring methane in the troposphere.

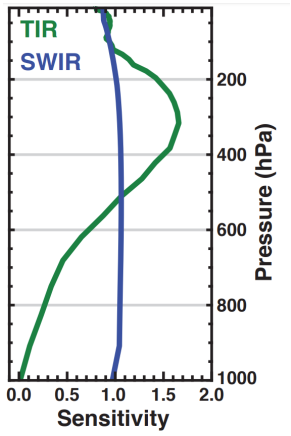


Figure 2.4: Shortwave infrared and thermal infrared sensitivities to methane for a typical satellite instrument as a function of atmospheric pressure [34]. The blue line shows sensitivity with respect to shortwave infrared radiation. The green line shows sensitivity with respect to thermal infrared radiation. The sensitivities are adapted from [36].

Figure 2.5 shows the solar radiation and Earth radiation spectrum. It can be noticed that the shortwave infrared wavelengths absorbed by methane (1.65 [µm] and 2.3 [µm]) are predominantly present in the solar spectrum. Therefore, for further discussions, we will focus on solar radiation.

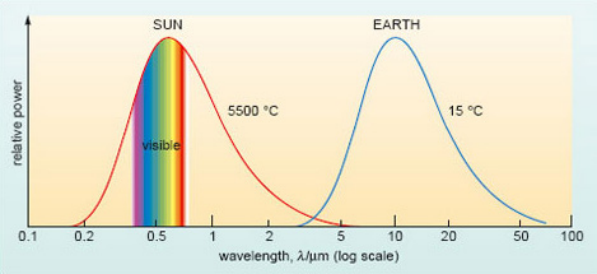


Figure 2.5: This figure shows spectrum of solar radiation intercepted by Earth (in red curve), and Earth's emission spectrum (in blue curve). Both the spectra are plotted against wavelength in a log scale. Please note that solar spectrum is relatively more powerful than Earth's emission spectrum, but for representation purpose, solar spectrum has been simplified.

Figure 2.6 shows a typical configuration of a satellite measuring shortwave infrared radiation in solar backscatter. Measured spectral data contain information about how much radiation in SWIR is absorbed in the atmosphere. The measured spectral data can change with the viewing geometry. Viewing geometry is defined by the solar zenith angle θ and satellite viewing angle θ_r. (Figure 2.6). The measured spectral data is fitted to a model to derive the total vertical column density of methane [37]. This model takes care of the satellite viewing geometry. The unit of retrieved vertical column density is [molecules/cm^2].

Satellite instruments that measure solar backscatter require the surface to be reflective. Therefore, the measurements are largely limited to land. Clouds reflect the solar radiation back to space. This prevents the methane below the cloud from getting detected. Therefore, cloudless scenes are preferable for measuring atmospheric methane concentrations. There are various satellite instruments that are used for measuring atmospheric methane concentrations using SWIR information. Some of the most prominent satellite instruments used for measuring atmospheric methane are SCIAMACHY, GOSAT, and TROPOMI. In this study, we are using the TROPOMI instrument for measuring atmospheric methane. The following text gives some details about SCIAMACHY, GOSAT, and TROPOMI.

SCIAMACHY

SCIAMACHY (SCanning Imaging Absorption SpectroMeter for Atmospheric CHartographY) aboard

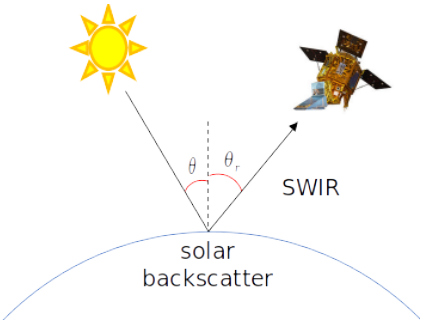


Figure 2.6: Satellite configuration for methane observations in the SWIR band. θ is the incidence angle of solar radiation at the Earth's surface. θ_r is the satellite observation angle. Angles θ and θ_r describe satellite viewing geometry. Solar backscatter is the radiation reflected and radiated by the Earth's surface which is later detected by the satellite instrument.

the ENVISAT satellite was launched in March 2002. It had a spatial resolution of 30 [km] × 60 [km], and required 6 days time for global coverage [37]. The fraction of successful methane concentration retrievals from SCIAMACHY was said to be around 9%, largely due to cloud cover limitations [38]. SCIAMACHY was originally designed to operate at the 2300 [nm] SWIR band, but due to the formation of an ice layer on the detector, the 1650 [nm] SWIR band was used. SCIAMACHY detected radiation in the spectral band range 1630 [nm] - 1670 [nm] with spectral resolution of 1.4 [nm] [37].

GOSAT

GOSAT (Greenhouse Gases Observing Satellite) was developed by JAXA (Japanese space agency) and was launched in January 2009. It has a spatial resolution of 10 [km] × 10 [km], and requires 3 days time for global coverage [39]. Measurements collected by GOSAT are sparse. GOSAT's fraction of successful methane concentration retrieval is nearly 17%, largely due to cloud cover limitations [40]. GOSAT uses a 1650 [nm] SWIR band for methane concentration retrieval. Range of this spectral band is 1630 [nm] - 1700 [nm] with spectral resolution of 0.06 [nm] [39].

TROPOMI

TROPOMI (TROPOspheric Monitoring Instrument) was developed by European Space Agency (ESA) in collaboration with Netherlands Institute for Space Research (SRON), Royal Netherlands Meteorological Institute (KNMI), Netherlands Space Office (NSO), and other institutes. It was launched in October 2017. TROPOMI has a spatial resolution of approximately 7 [km] × 7 [km] [41], and has a daily global coverage. TROPOMI can measure SWIR wavelengths from 2305 [nm] to 2385 [nm] with a spectral resolution of 0.25 [nm] [41].

SCIAMACHY and GOSAT were launched almost a decade earlier and have coarser spatial resolution and lower temporal coverage compared to TROPOMI. TROPOMI's higher spatial resolution can give better information about the spatial distribution of methane. Cloudy scenes are a big challenge for the SCIAMACHY, GOSAT and TROPOMI instruments. TROPOMI is somewhat better able to identify cloudy scenes using measurements from the VIIRS instrument [42].

SCIAMACHY and GOSAT operate at 1650 [nm] SWIR band, and TROPOMI operates at 2300 [nm] SWIR band. From Figure 2.3, it can be noticed that the 2300 [nm] absorption band is more powerful than the 1650 [nm] absorption band. Hence, the 2300 [nm] absorption band can give us better quality measurements.

Due to daily global coverage, exclusion of cloudy scenes using VIIRS, and use of a stronger absorption band in SWIR, TROPOMI is suitable for measuring global methane concentrations. There are other satellite instruments with a better spatial resolution than TROPOMI. Most of them are used in a targeted mode, which means they are pointed to observe some specific targets. Hence, most of the satellite instruments with better spatial resolution than TROPOMI lack daily global coverage and can also have issues with the signal-to-noise ratio [34].

2.5. Atmospheric methane concentration retrieval using TROPOMI

In the previous section, three satellite instruments were discussed: SCIAMACHY, GOSAT, and TROPOMI. TROPOMI instrument has been developed more recently, and it is suitable to monitor global methane emissions due to its daily global coverage and exclusion of cloudy scenes. This section discusses how TROPOMI collects information in orbit, what spectral information is captured by TROPOMI, and how that spectral information is converted to atmospheric methane concentrations.

TROPOMI instrument studies several chemical species, including methane, in the troposphere. The orbital path of TROPOMI is represented in Figure 2.7. TROPOMI is carried by Sentinel-5 Precursor (S-5 P) satellite. S-5 P follows a sun-synchronous orbit, with a mean overpass local time at 13:30.

TROPOMI operates in a push-broom configuration. In this configuration, a line of sensors simultaneously captures radiation over a strip of land. The land covered by this line of sensors is approximately 2600 [km] wide. Each sensor in the line of sensors captures a subdivision of a 2600 [km] wide strip. TROPOMI covers approximately 7 [km] along the flight direction in a 1-second flight. So, the dimensions of the land strip covered by a 1-second flight are approximately 2600 [km] $\times$ 7 [km]. After this 1 second flight, TROPOMI moves ahead, and measurements over a new strip of land are taken.

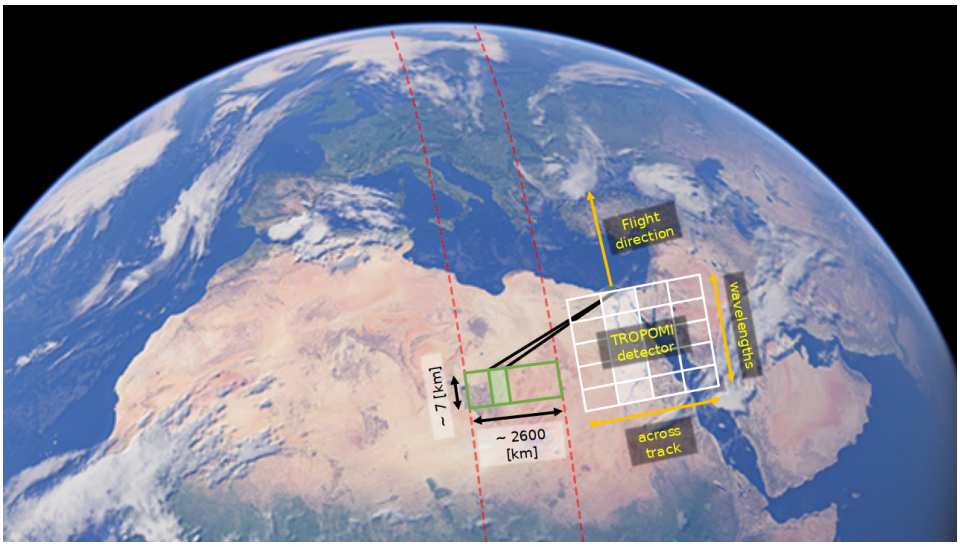


Figure 2.7: TROPOMI measurement push-broom configuration. Red curves indicate the swath covered by TROPOMI. In push-broom configuration, a line of sensors arranged perpendicular to the direction of the flight path is used. This line of sensors captures the information in the strip marked by a green rectangle on the surface of the earth. Each sensor along this line records a subdivision of this green strip. The spatial dimension of this green strip for TROPOMI is approximately 2600 [km] $\times$ 7 [km].

TROPOMI contains a nadir viewing grating UVNS spectrometer [42]. Nadir is the point/area on the Earth's surface that is directly below the satellite instrument. UVNS is the abbreviation for ultraviolet (UV), near-infrared (N), and short wave infrared (S) bands. Figure 2.8 represents the TROPOMI detector bandwidths (in white regions). TROPOMI's ultraviolet and visible region includes wavelengths from 270 [nm] to 500 [nm]. TROPOMI's near-infrared region covers the wavelengths between 675 [nm] to 775 [nm], and TROPOMI's short wave infrared region includes the wavelengths between 2305 [nm] to 2385 [nm].

Measured spectral data is fitted to a model to derive the total vertical column density of methane [37]. This model takes care of the satellite viewing geometry because the measured spectral data is sensitive to the viewing geometry. This model also takes care of variations in surface albedo and aerosol optical depth. The unit of retrieved total vertical column density is [molecules/cm^2]. The vertical column density of methane is sensitive to changes in topography, weather, and surface pressure. To remove this dependence, vertical column density is converted to a dry air column average mole fraction. Dry air column average mole fraction is not dependent on changes in pressure, and humidity [34]. The conversion is done by dividing the vertical column density of methane by the vertical column density of dry air.

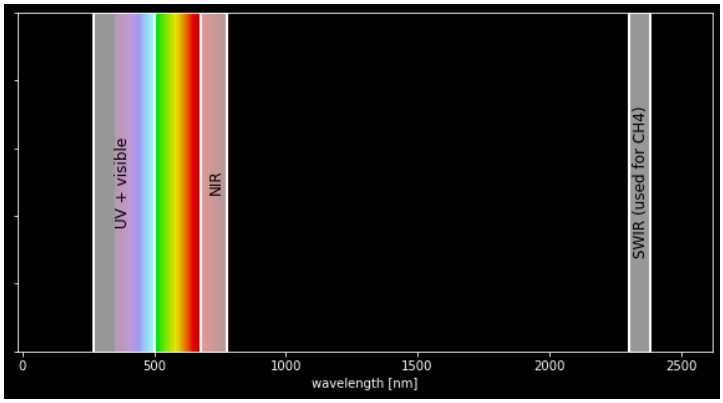


Figure 2.8: Regions in the light spectrum detected by the TROPOMI instrument (marked by transparent white vertical strips). TROPOMI can detect three regions in this spectrum: UV and visible, near-infrared (NIR), and shortwave infrared (SWIR). SWIR band is used for atmospheric methane concentration measurements.

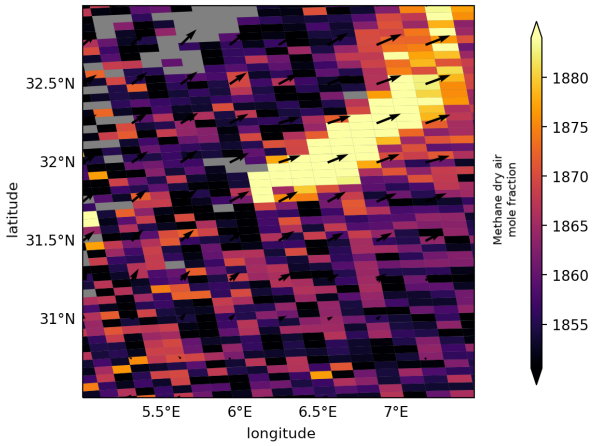


Figure 2.9: This figure shows an example of methane concentrations retrieved by TROPOMI. Colorbar shows dry air mole fraction of methane in [ppb]. the x-axis represents longitude, and the y-axis represents latitude. Longitude and latitude are given in degrees. Black arrows indicate wind direction. The lengths of black arrows indicate wind speed. The wind data was retrieved from GEOS-FP 10 [m] wind fields.

Figure 2.9 shows an example of dry air mole fraction of methane retrieved by TROPOMI at a particular location. Pixels in Figure 2.9 contain information about mean vertical methane concentrations. In this figure, pixels with high concentrations align nicely with the wind direction. In this case, the highly enhanced pixels indicate a methane plume (Box 2.2).

Box 2.2: Methane plume

A methane plume is a cloud of methane often originating from a point source. In most cases, methane plumes are distinctly identifiable in TROPOMI images due to their features. These features include clusters of pixels having large concentrations compared to surrounding/background pixels and orientation along the wind direction.

Accidental leakages in the fossil fuel infrastructure from even a single-point failure can release an extremely large quantity of methane into the atmosphere [43]. Large methane emissions from point sources like fossil fuel infrastructure and waste (Chapter 2.1) frequently form plumes. Accidental methane emission events occurring from fossil fuel infrastructure are often not quantified and remain unreported to bottom-up inventories. The company involved in the accidental emission of methane is not held accountable, and this is especially the case when the emission locations happen to be in remote areas [43]. Satellite instruments with daily global coverage like TROPOMI can be used to measure atmospheric methane in remote areas and have a higher chance of detecting accidental

methane emissions from fossil fuel infrastructure as well as methane emissions from waste management. Plumes observed in TROPOMI measurements can be used to estimate methane emission rate from point sources. Hence, the TROPOMI instrument can be used to put independent constraints on emission measurements that are difficult to do using bottom-up methods.

Methane plumes are often sighted in the TROPOMI images. TROPOMI has a daily global coverage, and manually identifying methane plumes in TROPOMI images can be cumbersome work. To automate the process of plume detection in the TROPOMI images, an automated algorithm was developed by Schuit [44]. The following subsection discusses the TROPOMI automated plume detection algorithm.

2.5.1. TROPOMI automated plume detection algorithm

According to a study performed by Schuit [44], a machine learning model was deemed to be suitable for detecting methane plumes in the TROPOMI data. This algorithm provided a way to make the operation of methane plume detection autonomous with acceptable accuracy. This study looked at the consequences of using a binary classification approach. This approach showed that the plumes with low source emission rates are harder to distinguish from the noise. The binary classification approach was termed to be the most suitable approach, while it may lead to erroneous predictions for the difficult classification scenarios.

TROPOMI measurements may contain artefacts (Box 2.3) due to ground features, surface albedo, aerosols, etc. This problem was also encountered in the study of the autonomous plume detection algorithm, which decreased the performance of the algorithm.

Box 2.3: Artefact

Artefacts can come from surface albedo variations or from aerosols. They can show up as methane enhancements while actually there is no methane there. The TROPOMI retrieval algorithm partly addresses this but is not able to fully resolve the problem. The following text explains artefacts from surface albedo variations and aerosols.

Surface albedo is the fraction of sunlight reflected back to space by the surface of Earth. Variations on the surface of Earth-like changes in terrain (from desert to rocky) and vegetation can affect the surface albedo. Hence, surface albedo is not uniform over the surface of Earth. Variations in surface albedo can distort the methane concentration measurements taken by the TROPOMI instrument, which collects solar backscatter.

Aerosols are the suspended fine solid particles or droplets of liquids in the atmosphere. Aerosols introduce uncertainties in the light path due to the scattering of light. Also, large aerosols like dust particles can block the solar backscattered radiations. Hence the presence of aerosols affects the satellite measurements of methane.

To get rid of artefacts, an automated feature engineering algorithm was developed by Schuit [44]. The TROPOMI automated plume detection algorithm is now being used for global monitoring of methane plumes. Several additions were made to this algorithm since its development with the use of a larger training dataset which improves the detection accuracy of the algorithm in challenging scenarios.

Quantification of plumes identified in TROPOMI images is essential to get an estimate of the emission rate of the source. Quantification methods based on simple mass balance are used in the automated algorithm to quantify emissions corresponding to each plume. The following section discusses TROPOMI plume emission quantification based on simple mass-balance methods.

2.6. Plume emission quantification methods used with the TROPOMI automated plume detection algorithm

As explained in section 2.4, satellite instruments measure the vertical column density of atmospheric methane using solar backscattered radiation in the shortwave infrared (SWIR) region. These measurements are converted to a dry air mole fraction [ppb] to make them independent of pressure and

humidity. The dry air mole fraction of methane [ppb] can be converted to the total vertical column mass of methane using the following expression.

$$\Delta\Omega = \frac{M_{\mathrm{CH_4}}}{M_{\mathrm{air}}}\Omega_{\mathrm{air}}\Delta X \tag{2.1}$$

In the eq. (2.1), $\Delta\Omega$ is the total vertical column mass or vertical column mass enhancement in [kg/m^2], $M_{\mathrm{CH_4}}$ is the molar mass of methane in [kg/mol], M_{air} is the molar mass of dry air in [kg/mol], Ω_{air} is the weight of column of dry air in [kg/m^2], and ΔX is the measured dry air mole fraction of methane in [ppb] minus the background dry air mole fraction of methane (X-X$_b$).

The emission rate of methane can be obtained using mass conservation or mass balance principle. Currently, TROPOMI automated plume detection algorithm uses the following mass balance approaches: Integrated Mass Enhancement (IME) method and Cross-Sectional Flux (CSF) method [33]. The following sections explain IME and CSF TROPOMI plume emission quantification methods.

2.6.1. Integrated Mass Enhancement (IME) method

IME was given by Frankenberg et al. [45] and is used in the TROPOMI automated plume detection algorithm to quantify plumes. First, the total mass is calculated by integrating the mass of all the source pixels (Box 2.4). The obtained total mass is also called Integrated Mass Enhancement, hence the name of this method.

$$IME = \sum_{j=1}^{N} \Delta\Omega_j A_j \tag{2.2}$$

In the above equation, $\Delta\Omega_j$ is the total vertical column mass of pixel number j in [kg/m^2], and A_j is the area of pixel number j in [m^2].

Box 2.4: Source pixels

Source pixels are the pixels which are part of the plume. These pixels are distinguishable from other pixels due to their high concentration values. In ideal situations, source pixels can be separated from the rest of the pixels. The remaining pixels are categorised as background pixels. In real TROPOMI images, separating source pixels is challenging due to missing pixels or noise in the image. Different methods are available to separate the source pixels from the background pixels.

Later, the total mass is divided by the residence time of a plume to obtain the source emission rate. The empirical relation between the Integrated Mass Enhancement (IME) and the source emission rate is shown below.

$$Q = \frac{IME}{\tau} \tag{2.3}$$

In the eq. (2.3), τ is the residence time of methane plume. The residence time of methane is the time for which methane is travelled in the atmosphere after it is released from the source. The residence time can be approximated by the distance travelled by the plume (characterised by the plume length) divided by the effective wind speed.

$$\tau = \frac{L}{U_{\mathrm{eff}}} \tag{2.4}$$

In the above equation, L is the length or extent of plume in [m] and U_{eff} is the effective wind speed in [m/s]. Varon et al. [33] calculated the plume extent L by masking the plume and taking the square root of the plume mask ($\sqrt{A_m}$). A different approach can also be applied to calculate plume extent.

The effective wind speed can be calculated with various empirical formulations. Varon et al. [33] use empirical relation based on Large Eddy Simulation (LES) results. This empirical expression is given in the following equation.

$$U_{\text{eff}} = \alpha_1 \log U_{10} + \alpha_2 \tag{2.5}$$

In eq. (2.5), α_1 and α_2 are the regression coefficients given by the range 0.9-1.1 and a value of 0.6 [m/s], respectively, and U_{10} is the wind speed measured 10 [m] above the surface [33]. This setup is valid for images with pixel resolution of 50×50 [m^2].

For TROPOMI images, a study done at SRON found a linear empirical relation between U_{eff} and U_{10}. This empirical relation is shown in the following equation.

$$U_{\text{eff}} = \alpha_3 U_{10} + \alpha_4 \tag{2.6}$$

In eq. (2.6), α_3 and α_4 are 0.444 and 0.278, respectively. The empirical relations described in this section are case-specific and might not hold true for a different setup or different satellite instrument. Empirical relations presented in this section are simple, which means there is an option to derive empirical relations for a different setup.

2.6.2. Cross-Sectional Flux (CSF) method

In the CSF method, plume emission quantification is done by calculating flux through the cross-section(s) of the plume (orthogonal to the downwind direction). A flux, in this case, is the flow of mass through a reference geometry. The source rate is equal to the integration of the product of wind speed U and the column enhancements over a cross-section perpendicular to the downwind direction of the plume. Here downwind direction of the plume refers to the plume direction or plume orientation and not the wind direction, as the wind direction might not always be the same as the plume direction due to reasons mentioned later in this chapter. The following equation shows this relationship.

$$Q = \int_{-\infty}^{+\infty} U(x,y)\Delta\Omega(x,y)\,\mathrm{d}y \tag{2.7}$$

In the above equation, $\Delta\Omega(x,y)$ is the total vertical column mass of pixel located downwind at a distance x from the source and y along the cross-section line. $U(x,y)$ is the wind speed at the location x downwind and y along the cross-section line. Pixels along cross-section lines are selected based on threshold criteria. Usually, this criterion is based on a multiple of standard deviation values (for example, 1.8σ, σ is the standard deviation).

The integration is approximate by summation of the product of effective velocity (U_{eff}) and column enhancements over a cross-section, as shown in the following equation. In practice, the source rate is computed for multiple cross-sections, and later an average is taken.

$$Q = \sum \Delta\Omega(x,y) U_{\text{eff}}\,\mathrm{d}y \tag{2.8}$$

The effective wind speed can be calculated by various empirical relations. Varon et al. [33] found the regression coefficient β used in equation 2.9 to be in the range 1.3-1.5. Similar to IME, the empirical relations described in this section might vary for a different satellite instrument.

$$U_{eff} = \beta U_{10} \tag{2.9}$$

The IME and CSF source rate quantification methods produced promising results for the GHGsat measurements [33], but there are several limitations to using IME and CSF methods. These limitations can be divided into solvable limitations and non-solvable limitations. The following paragraphs discuss the solvable limitations of the mass balance methods.

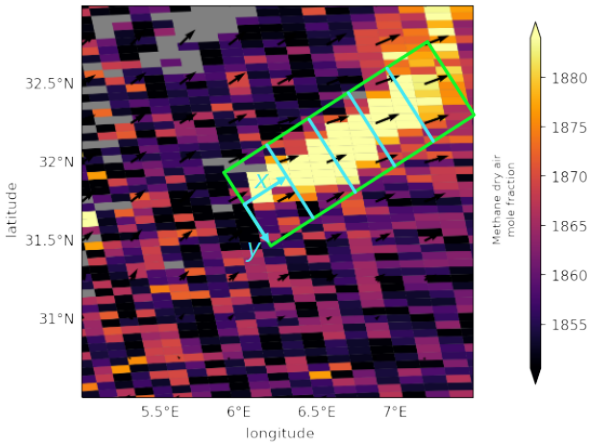


Figure 2.10: This figure illustrates the CSF method. The green box starts from the source location and is oriented in the downwind direction. Cross-section lines are marked in cyan colour. x axis is oriented along the downwind direction, and y axis is oriented along the cross-section lines. Description for the TROPOMI image in the background can be seen in Figure 2.9.

Parametric coefficients in the effective wind speed equation are not universal and dependent on a particular setup. This means parametric coefficients need to be re-calibrated for a new setup. Similarly, limitations linked to the use of uniform effective wind speed over the plume, plume mask, calculation of plume extent and more can be resolved by developing some advanced IME and CSF methods. On the other hand, some limitations of the mass balance methods are inherent and cannot be fixed. These non-fixable limitations are discussed in the following paragraph.

The inherent limitations of mass balance methods include missing plume pixels, non-characterisation of wind speed in the vertical direction, low wind speed conditions and more. This paragraph discusses some of the inherent limitations of the mass balance methods. TROPOMI images might have missing pixels due to cloud coverage or noise due to various factors. Missing or noisy pixel information causes an underestimation or overestimation of plume mass that affects the mass balance quantification and makes the estimates uncertain. The wind velocity $U(x, y)$ is not characterised due to the non-inclusion of variation of velocity in the vertical direction. The mass balance methods tend to perform poorly in low wind speed conditions [33].

Due to the fact that the inherent limitations of simple mass balance methods cannot be fixed, there will always be some uncertainty associated with the mass balance methods. Therefore, there is a need for some physics-informed approach that could overcome most of the inherent limitations of simple mass balance methods. Atmospheric transport models can account for three-dimensional particle transport in the atmosphere and can be used to reconstruct methane plumes. Therefore, in this study, we are trying to improve TROPOMI plume emission quantification with the help of atmospheric transport models. The following sections discuss atmospheric transport models used in this study and the generation and quantification of the TROPOMI plumes using atmospheric transport models.

2.7. Atmospheric transport models

The TROPOMI automated plume detection algorithm often captures plumes originating from point sources such as fossil fuel infrastructure and waste management. The automated detection algorithm currently uses the IME and CSF mass balance methods to quantify the detected plumes. The mass balance methods tend to produce uncertain plume emission quantifications due to several limitations. Hence, there is a need for a robust method that overcomes most of the inherent limitations of the mass balance methods. In this study, we exploit atmospheric transport models to improve the TROPOMI plume emission quantification in the automated detection algorithm. This section gives a brief overview of atmospheric transport models before discussing how the TROPOMI plumes can be quantified with the atmospheric transport models.

The origin of atmospheric transport models dates back to the early twentieth century. The earliest mod-

els were based on simplified chemistry and transport. One of the main objectives of developing these models was to predict climate change. Compared to the models present today, models in the early-mid twentieth century could not sample all the possible intricacies of a complex atmospheric system. Therefore, they were very uncertain in their predictions. This gave rise to the need for the development of more reliable and robust models. In 1914, Bjerknes proposed to use physics-based deterministic models for weather forecasting [46]. The physics-based models involve complex equations that cannot be solved analytically; hence models based on numerical methods were developed. This laid the foundation for the development of atmospheric transport models.

The atmospheric transport models are categorised into the Eulerian models and the Lagrangian models based on the approach used for particle transport. The Eulerian transport model uses a frame of reference fixed to the computational grid, whereas the Lagrangian transport model uses a frame of reference fixed to the transported particles. The following sections provide some ideas about how particle transport takes place in the Eulerian and Lagrangian approaches.

2.7.1. The Eulerian approach

In the Eulerian approach, a fixed computational grid is used for performing numerical calculations. Figure 2.11 shows flow passing through a fixed computational grid cell centred at (x,y,z). In Figure 2.11, F_i [kg m^{-2} s^{-1}] is the mass flux of the species passing through walls of i^{th} cell. Here, flux is defined as the rate of mass passing across a given area.

In the Eulerian model environment, several different processes are involved. These processes are taken care of by the transport and local terms. Equation 2.10 represents the sum of transport and local terms used for describing different processes in the Eulerian framework.

$$\begin{aligned}\frac{\partial \rho_i}{\partial t} = &\left[\frac{\partial \rho_i}{\partial t}\right]_{\text{advection}} + \left[\frac{\partial \rho_i}{\partial t}\right]_{\text{turbulent mixing}} + \left[\frac{\partial \rho_i}{\partial t}\right]_{\text{convection}} + \left[\frac{\partial \rho_i}{\partial t}\right]_{\text{scavenging}} \\ &+ \left[\frac{\partial \rho_i}{\partial t}\right]_{\text{chemistry}} + \left[\frac{\partial \rho_i}{\partial t}\right]_{\text{emissions}} + \left[\frac{\partial \rho_i}{\partial t}\right]_{\text{dry deposition}}\end{aligned} \tag{2.10}$$

In the above equation, ρ_i is the concentration measured in terms of mass density [kg m^{-3}], and t is time [s]. The mass flux F_i can be written as the product of mass density ρ_i and averaged velocity vector **v**.

$$F_i = \rho_i \mathbf{v} \tag{2.11}$$

In equation 2.10, the advection term, the turbulent mixing term, and the convection term are mainly responsible for particle transport and hence called transport terms. We are using the atmospheric transport model for particle transport; hence we are considering the transport terms. Apart from the transport terms, the chemistry term is considered in this study to account for several atmospheric reactions. The rest of the terms are local terms and do not contribute to particle transport. The following text discusses the transport terms used in this study.

Advection is the transport of species by the wind on a model scale. A model scale depends on the size of the computational grid used in the Eulerian model. The change of mass concentration in a grid cell per unit time is given by the difference of flow rate into the grid cell and flow rate out of the grid cell per unit volume of the grid cell. The flow rate is obtained by multiplying flux with the area through which flux is passing. The following equation describes the temporal change in mass concentration due to advection in the x direction.

$$\left[\frac{\partial \rho_i}{\partial t}\right]_{\text{advection},x} = \frac{[F_i^x(x - dx/2) - F_i^x(x + dx/2)]\, dy\, dz}{dx\, dy\, dz} = -\frac{\partial F_i^x}{\partial x} = -\frac{\partial(\rho_i u)}{\partial x} \tag{2.12}$$

Similarly, temporal changes in mass concentrations in the y and z directions can be obtained using the above equation.

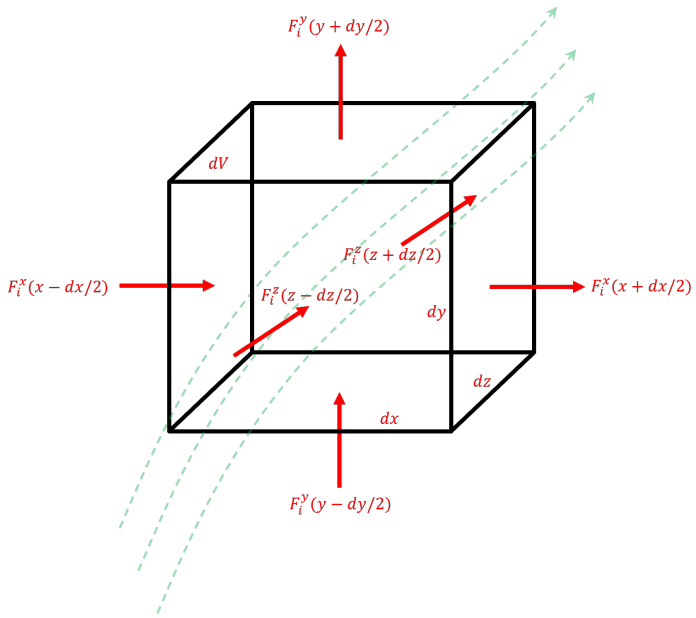


Figure 2.11: Eulerian framework depicting flow (green dotted lines) going through a grid cell with volume dV and the split of the flux F_i in 3 dimensions (red arrows). The dimensions of this grid cell are *dx*, *dy*, and *dz*. The orientation of the x axis is from left to right, the orientation of the y axis is from bottom to top, and the orientation of the z axis is from out of the page to into the page.

Adding the components along x, y, and z directions, we get the final form of the advection term.

$$\left[\frac{\partial \rho_i}{\partial t}\right]\text{advection} = -\frac{\partial F_i^x}{\partial x} - \frac{\partial F_i^y}{\partial y} - \frac{\partial F_i^z}{\partial z} = -\frac{\partial(\rho_i u)}{\partial x} - \frac{\partial(\rho_i v)}{\partial y} - \frac{\partial(\rho_i w)}{\partial z} \tag{2.13}$$

It can be noticed that equation 2.13 makes use of wind velocity components in three dimensions (u,v,w). Hence, the Eulerian model uses three-dimensional wind velocity information, unlike the simple mass balance approaches seen in the previous chapter.

The advection term takes care of the model level scales that can be resolved by the size of the computational grid. A typical Eulerian model computational grid has a resolution of 10^0 - 10^1 [km]. Scales in the atmospheric boundary layer are considered to be small and can often not be resolved by the computational grid [47]. Small scales which can not be resolved by the computational grid are often termed turbulent scales (Box 2.5). The atmospheric boundary layer is typically 1-2 [km] in height (Figure 2.12), and the size of the turbulent scales inside the atmospheric boundary layer ranges from 10^{-2} to 10^0 [km] [48]. These fluctuating scales have zero value for the time ensemble mean. However, turbulent scales' contribution to the transport of species is often high [47]. Generally, transport models apply parametric turbulent mixing in the boundary layer of the planet to take care of the small scales.

Box 2.5: Turbulence

Turbulence is characterised by chaotic changes in the motion of the fluid. Some complex terms present in physics-based transport equations are related to turbulence. These complex terms cannot be solved analytically and are linked to small scales. Therefore, numerical models often parameterise complex turbulence terms.

Turbulent mixing is mainly responsible for the vertical transport of the species in the planetary boundary layer [47]. Moreover, the turbulent mixing term stabilises the numerical simulation [50]. On the other hand, the convection term takes care of the vertical transport outside the planetary boundary layer. For small-scale models, the convection term is based on the advection principle. For large-scale models, the convective term is parameterised.

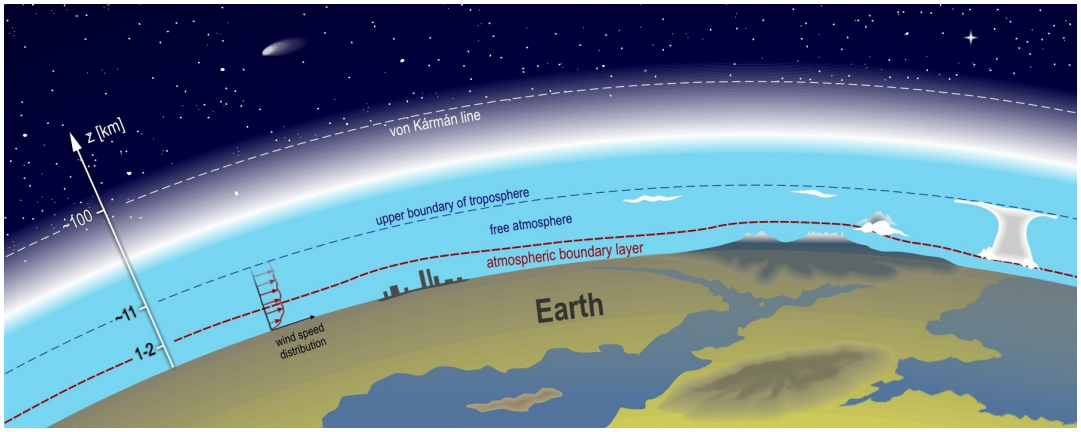


Figure 2.12: This figure shows the atmospheric boundary layer (region below the red dotted line). The upper boundary of the troposphere is marked by a blue dotted line. The white vertical arrow indicates height from the surface of the earth. Please note that the figure is not drawn to scale. [49]

This section provided an overview of the principle behind the Eulerian models. The Eulerian models use a fixed computational grid for performing numerical calculations. There is another category of models that traces the particles to compute their atmospheric concentrations. These models are termed the Lagrangian models. The following section provides an overview of the principle behind the Lagrangian models.

2.7.2. The Lagrangian approach

In the Lagrangian model, chemical concentrations are obtained by following the trajectory of particles displaced by air motions [47]. There are different types of Lagrangian models: a model that tracks a single particle displaced by the mean velocity with neglected turbulence, a model that tracks multiple particles confined in a box, a model with Gaussian puff describing turbulent dispersion, and a particle dispersion model (LPDM) with turbulence generated by a stochastic process.

The LPDMs simulate particle transport with advection and turbulent diffusion. The advection and turbulent diffusion terms in LPDM are based on sophisticated parametric stochastic equations [51]. LPDMs have the ability to naturally deal with turbulent diffusion, unlike the other Lagrangian models [51]. Due to their several benefits, most atmospheric models use LPDM to simulate the Lagrangian particle transport.

Most of the Lagrangian models, including LPDM, can simulate particle transport in forward or backward mode. In the forward mode, particles are released from one or more source locations, and their transport takes place using fluid motions in the model domain (refer to Figure 2.13). In the backward mode, particles released at a receptor are used to find influences of upwind locations at the receptor point (refer to Figure 2.13). The backward simulation mode can potentially help in locating upwind source regions from atmospheric measurements, in contrast to the forward mode, which might require many simulations to locate the source.

2.7.3. Advantages and disadvantages of the Lagrangian approach over the Eulerian approach

Advantages and disadvantages of using the Lagrangian framework over the Eulerian framework can be divided into two categories: advantages and disadvantages associated with the models' setup and advantages and disadvantages associated with working of the models.

Some advantages of using the Lagrangian model over the Eulerian model related to the setup are discussed in this paragraph. In the Lagrangian model, since there is no computational grid, the time step is not limited by the Courant–Friedrichs–Lewy (CFL) condition. On the other hand, the time step in the Eulerian model is restricted due to the CFL condition. The CFL condition is the relation between the time step, grid cell size, and velocity, which is often used to check the stability of a numerical system. Due to the reasons mentioned above, large time steps can be used in the Lagrangian model, which can

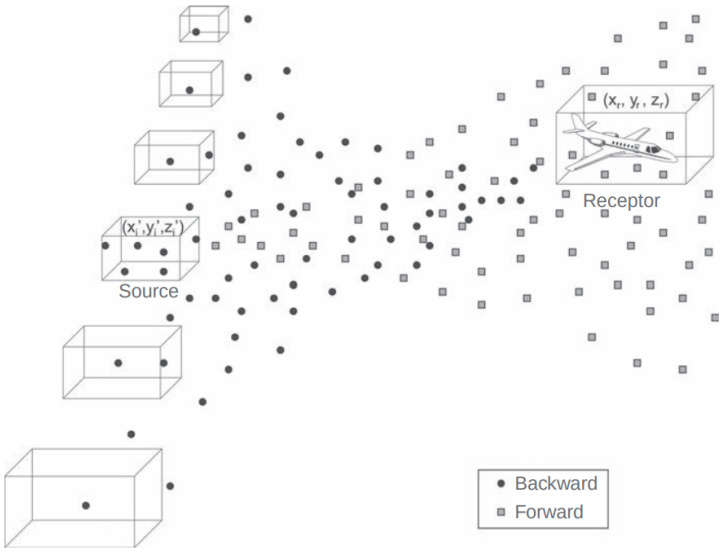


Figure 2.13: This figure illustrates the forward and backward Lagrangian models. Black circles represent particles used for the backward Lagrangian model run, and grey boxes represent particles used in forward Lagrangian model run. In the forward Lagrangian model run, particles move from the source to the receptor, and in the backward Lagrangian model run particles move from the receptor to the source. [47]

reduce the computational time. In the Lagrangian model, resolution in the area of interest can easily be increased by increasing the number of particles. On the other hand, to increase the resolution in the Eulerian model, the size of the grid cell needs to be reduced, which increases computational costs for the Eulerian model. In the Lagrangian model, particle trajectories are independent; hence performing parallel computation is easier on the Lagrangian model. On the other hand, in the Eulerian model, grid cells are not independent, and it makes parallel computation challenging.

Some disadvantages of using the Lagrangian model over the Eulerian model related to the setup are discussed in this paragraph. The Lagrangian model offers non-uniform coverage of the 3-dimensional domain; hence there might be some unsampled regions in the domain. On the other hand, the Eulerian model performs calculations over a uniformly distributed 3-dimensional grid over the computational domain; hence there are no unsampled regions in the computational domain.

Some advantages of using the Lagrangian model over the Eulerian model related to the working of the model are discussed in this paragraph. In the Lagrangian framework, there is no numerical diffusion; hence the sharp gradients in tracer concentrations are preserved. Numerical diffusion is the numerical error that originates due to the first-order approximation of a spatial partial derivative. Since no spatial computational grid is used in the Lagrangian model, there is no numerical diffusion. Numerical diffusion causes the reduction of sharp peaks or edges. Hence, a sharp drop in concentration values between grid cells might not be preserved in the Eulerian model.

Some disadvantages of using the Lagrangian model over the Eulerian model related to the working of the model are discussed in this paragraph. The Lagrangian model generates small-scale features due to insufficient particle mixing. These small-scale features are usually dissipated in the atmosphere due to turbulent mixing but are present in the Lagrangian model's results, whereas these unwanted small-scale features are not present in the Eulerian modelś results. The Lagrangian model tends to neglect the non-linearities present in the transport process, which are to some extent preserved by the Eulerian model [47].

This study exploits both the Eulerian and Lagrangian frameworks for TROPOMI plume emission quantification. Both the frameworks have some positive points and negative points, and this study will check and compare the ability of the Eulerian and Lagrangian frameworks to quantify the TROPOMI plumes.

2.8. WRF-CHEM and FLEXPART

In the previous section, two distinct frameworks of atmospheric transport models were discussed. There are various computational tools or software available that can simulate particle transport using the Eulerian as well as the Lagrangian model. At SRON, the Weather Research Forecast model coupled with chemistry (WRF-CHEM) is used for simulating the Eulerian particle transport. For simulating the Lagrangian particle transport, the FLEXible PARTicle (FLEXPART) model is chosen in this study. The FLEXPART model was selected due to the availability of literature and its ability to simulate using the LPDM. The following subsections provide a short overview of the WRF-CHEM and FLEXPART models.

2.8.1. Short overview of the WRF-CHEM model

The Weather Research and Forecasting model (WRF) [52] is an open-source Eulerian atmospheric modelling system used for numerical weather prediction and research related to it. The WRF model has been employed for research at universities, and government laboratories and commercial applications by industries [53]. The WRF model is written in Fortran language and can be built in any UNIX-based operating system using any Fortran compiler, and this model can also be operated in a parallel computing framework [53].

In this study, the atmospheric transport simulations are performed on a system having a WRF model integrated with a chemistry package (WRF-CHEM). WRF-CHEM is an inline chemistry solver coupled with the dynamical solver of WRF. Henceforth, the WRF-CHEM model will be referred to as WRF in this report.

Running atmospheric simulations with WRF consists of three phases: preprocessing, configuration, and model run. A WRF preprocessing system was designed for preprocessing the meteorological data (ECMWF [54], and NCEP [55]) (Box 2.6), which is supplied to the dynamical solver (used in the third phase) of WRF. The second phase deals with the configurations of the model domain, supplying the input data, and providing the initial conditions. The third phase deals with running the forecast model using a dynamical solver.

Box 2.6: Meteorological data

WRF-CHEM and FLEXPART use meteorological wind data as input since they are offline models. Offline models use the archived output from a separate meteorological model. The wind field can be either in a GRiB or GRiB2 format. We are acquiring the wind fields from the ECMWF [54] and GFS (NCEP) [55] numerical weather prediction database because of their compatibility with both the WRF and FLEXPART models. The meteorological grid is distributed over latitude, longitude and the model levels (heights). The meteorological data often needs to be preprocessed before using it with the atmospheric transport models.

The WRF solver uses various formulations for turbulent mixing. As seen previously, turbulent mixing is related to vertical mixing in the planetary boundary layer. In WRF, vertical mixing is considered dominant in the boundary layer and is parametrised with the help of planetary boundary layer (PBL) physics. Various PBL schemes are listed in the WRF user manual [56]. The WRF model supports multiprocessing for parallel simulations. The WRF model can output the grid concentrations and some meteorological variables in the NetCDF format. This NetCDF file is used later in the postprocessing routine.

2.8.2. Short overview of the FLEXPART model

The FLEXible PARTicle model, abbreviated as FLEXPART, is a tool used for atmospheric transport modelling and analysis using the stochastic version of the Lagrangian approach or LPDM. It originated almost ten years ago to simulate the long-range dispersion of pollution particles from a point source. This model calculates the trajectories of the particle parcels to analyse atmospheric transport and diffusion of tracer gases like methane.

The first version of the FLEXPART was developed by Eckhardt and Stohl in the 2000s. The second ver-

sion of FLEXPART was validated with the data from three tracer experiments conducted by Stohl [57]. The next versions of FLEXPART saw the addition of convective schemes and boundary layer parametric schemes. The capabilities of FLEXPART have been validated with various air pollution transport studies [58], [59], and [60]. FLEXPART is based on Fortran 95 language and can work on the compilers like gnu, Absoft and Portland Group. FLEXPART can model the particle trajectories in a forward and backward manner. FLEXPART also supports multiprocessing for parallel simulations.

Similar to the WRF model, a FLEXPART model run consists of three phases. The FLEXPART model uses a preprocessing routine to check the meteorological data, and this preprocessing routine computes vertical velocities if they are not present in the meteorological data. Then, the FLEXPART model reads the input settings from a directory, which contains several command files. These command files have settings that define the FLEXPART solver run. A more detailed description of these command files can be seen in the FLEXPART used guide document [61].

FLEXPART uses a parametric model for simulating the atmospheric boundary layer. This parametric model calculates the velocities in the unit [m/s] and not in the η coordinate (Box 2.7). Therefore, to avoid the complexity of transforming the coordinates every time step, the coordinate data is linearly interpolated from the ECMWF model levels to the Cartesian terrain-following system [61]. Hence, the vertical coordinate system present in the FLEXPART output is a Cartesian terrain-following system.

Box 2.7: Eta coordinate

Messinger, in 1984 introduced a vertical coordinate for taking care of the mountains and altitude variations on the land [62]. Meteorological data required for the WRF and FLEXPART models are usually distributed in the η vertical coordinate. This coordinate is called the η coordinate and is represented in the following equation.

$$\eta = \sigma \frac{p_{ref}(z_s) - p_{top}}{p_{ref}(0) - p_{top}} \tag{2.14}$$

In eq. (2.14), the term σ is calculated with the help of the following equation.

$$\sigma = \frac{p - p_{top}}{p_s - p_{top}} \tag{2.15}$$

In eq. (2.14) and eq. (2.15), z_s is the local terrain elevation, z is the geometrical height, p is the pressure at a particular vertical level, p_s is the pressure at the surface, p_{top} is the pressure at the topmost vertical level, $p_{ref}(z_s)$ is the reference pressure, and $p_{ref}(0) = 0$.

Meteorological data fields are often available in the eta coordinate system.

The FLEXPART model outputs the grid concentrations of the particles over a uniform grid defined in the input settings. This file is by default in binary format. Users can define the units of the output (mass [kg] or mixing ratio [pptv]) by adjusting the input settings. The FLEXPART model can also output in NetCDF format, which is more desirable for this study.

2.9. Generating plumes using atmospheric transport models

Atmospheric transport models use physics-informed equations and account for the three-dimensional transport of chemical species, which was lacking in the simple mass balance methods. We use this ability of atmospheric transport models to improve the TROPOMI plume emission quantification algorithm. We can generate synthetic plumes using atmospheric transport models. These synthetic plumes can be matched with the TROPOMI plumes. If a synthetic plume agrees well with the TROPOMI plume, we can estimate the emission rate of the TROPOMI plume based on that synthetic plume. The previous sentences briefly explained the process of quantifying emissions from TROPOMI plumes with atmospheric transport models, but before going into details of how this quantification process works, a methodology to generate synthetic plumes using atmospheric transport models is discussed in this

section.

We are focusing on methane emissions from point sources like fossil fuel infrastructure and waste management. Both the WRF and FLEXPART models can simulate methane emissions from a point source. For generating plumes in the WRF and FLEXPART models, the location and emission rate of a persistent point source of methane can be provided. After the simulation is complete, the dry air mole fraction of methane and some other meteorological parameters are stored in an output file. Please note that the FLEXPART model does not use the computational grid for simulation, but the outputs are stored in a three-dimensional grid. Methane dry air mole fraction outputs from the WRF and FLEXPART models span over the longitude, latitude, and vertical coordinates. The FLEXPART model internally converts the hybrid sigma pressure vertical coordinates (Box 2.8) to the height over sea level in metres and uses this coordinate system for boundary-layer parametrisation. Hence, the outputs obtained from the FLEXPART model follow an altitude [m] based vertical coordinate system. On the other hand, the WRF model gives the output in the hybrid sigma pressure coordinates.

Box 2.8: Hybrid sigma-pressure vertical coordinate system

Atmospheric transport models often use a hybrid sigma-pressure coordinate system. The pressure at vertical level k is called p_k and is calculated using the following equation.

$$p_k = A_k p_0 + B_k p_s \tag{2.16}$$

In eq. (2.16), coefficients A_k and B_k are dependent on the vertical level k, and the term p_0 is the pressure at the sea level (1013 hPa).

Since the hybrid sigma pressure coordinate system is more suitable for postprocessing, the FLEXPART output is mapped to a hybrid sigma pressure coordinate system using a separate routine.

The postprocessing involves the conversion of three-dimensional methane dry air mole fraction output data to a TROPOMI-like image. This can be done with the help of an algorithm known as the resampling algorithm. First, latitude-longitude distributed data in each vertical level is interpolated to match the TROPOMI pixel resolution. For this interpolation, the resampling algorithm uses an actual TROPOMI image to get the location, size and orientation of the pixels. Next, the model output vertical levels are interpolated to the TROPOMI vertical levels. There are, in total, 12 vertical levels retrieved in the TROPOMI measurements. Please note that despite retrieving 12 vertical levels in the measurement, TROPOMI virtually has no methane vertical profile information in the measurement [63]. Therefore, the TROPOMI methane data is given as a column-averaged dry-air mole fraction of methane or a 2-dimensional image (refer to Section 2.5).

In the next step, information from the vertical levels is converted to column-averaged dry-air mole fraction of methane. The relation between vertical distribution of methane ($\mathrm{XCH_4}_{\mathrm{model,i}}$) and column-averaged dry-air mole fraction of methane ($\mathrm{XCH_4}_{\mathrm{total\ column}}$) can be given by the following equation.

$$\mathrm{XCH_4}_{\mathrm{total\ column}} = \sum_{i=1}^{n} \left(A_{col,i} \mathrm{XCH_4}_{\mathrm{model,i}} + (1 - A_{col,i}) \mathrm{XCH_4}_{\mathrm{prior,i}} \right) / V_{\mathrm{air,dry}} \tag{2.17}$$

In the above equation, i represents a number of a vertical level, and $A_{col,i}$ is an averaging kernel. The averaging kernel describes the sensitivity of retrieved methane information from a column to changes to the methane profile in the column (Figure 2.4). The term ($\mathrm{XCH_4}_{\mathrm{prior,i}}$) describes the prior profile of methane in the vertical column. The term ($V_{\mathrm{air,dry}}$) is the dry air column. The TROPOMI measurement files usually contain information about the prior methane profile in columns, the averaging kernel for the columns, and the amount of dry air in columns. The value of ($\mathrm{XCH_4}_{\mathrm{total\ column}}$) is obtained for all the TROPOMI pixels distributed in latitude-longitude. Therefore, the quality of the results after this step depends on the coverage of pixels in the TROPOMI image. Finally, a TROPOMI-like image with a synthetic plume is constructed.

Synthetic plumes generated by the atmospheric transport models can be used to reconstruct and quantify the TROPOMI plumes. A detailed description of TROPOMI plume emission quantification using atmospheric transport models is provided in the next section.

2.10. Quantifying emissions from plumes using atmospheric transport models

In the situations where the TROPOMI coverage is bad, mass balance methods perform poorly, but synthetic plumes generated by atmospheric transport models can help quantify the TROPOMI plumes. In 2018, a published study revealed extreme methane leakage from a natural gas well blowout (uncontrolled release of gas) event in Ohio, USA [43]. This study used the WRF model to perform atmospheric transport of methane released from the blowout site. The WRF simulations were performed closest to the TROPOMI overpass time, sampled at TROPOMI pixels, and later scaled to match the TROPOMI measured methane enhancements. The WRF results showed similar plume enhancements as seen in the TROPOMI image.

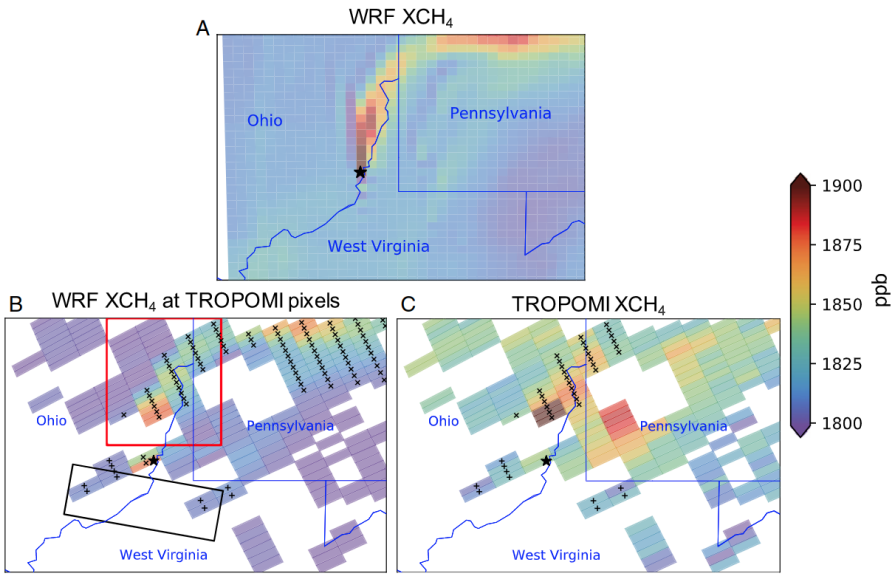


Figure 2.14: This figure represents a comparison of the WRF synthetic plume with the TROPOMI observed methane emissions from the Ohio natural gas well blowout site [43]. Image (A) shows the column-averaged mixing ratios of methane near the TROPOMI overpass time simulated with the WRF model. In the image (A), a synthetic plume can be seen. Image (B) shows the WRF results sampled at TROPOMI pixels. In the image (B), the rectangle in the south region (upwind direction from blowout site) of the blowout site contains the background pixels. Background pixels are used to determine excess enhancements in the plume pixels compared to the surroundings. Pixels used from the background region are marked with plus symbols. The red box in the downwind direction from the blowout site marks the pixels closest to the blowout site. The pixels used for the emission quantification are marked with cross symbols. Image (C) represents the measurements taken by the TROPOMI instrument over the blowout site. In the image (C), the pixels used for the emission quantification are marked with crosses.

The emission rate of the TROPOMI plume is quantified by matching the synthetic plume to the TROPOMI observations. The scaling used for this purpose is given in the following equation.

$$Q_T = Q_S \times \frac{X_T}{X_S} \tag{2.18}$$

In the above equation, X_T represents the methane enhancements seen in the TROPOMI image, and X_S represents the methane enhancements seen in the synthetic plume image. X_S was calculated by adding methane enhancements due to blowout, EPA emissions, and boundary conditions. EPA is the national anthropogenic emissions gridded inventory of the United States. The EPA emissions were added to account for the emissions not related to the blowout. The boundary conditions took care of the methane transported from outside the computational domain. Only the pixels close to the blowout

site were considered for quantification to minimise the uncertainty present in the WRF simulated results. The terms X_S and X_T were obtained after subtracting background pixel enhancements from the blowout pixel enhancements.

In another study published in 2021, methane emissions from a natural gas blowout that happened on 30 August 2019 near Louisiana, USA, were quantified with the help of the WRF simulations [64]. In this study, emissions were estimated by the analytical solution of the Bayesian inverse problem. The following equation represents an analytical solution to the Bayesian inverse problem.

$$\hat{x} = \mathbf{x_A} + \mathbf{S_A}\mathbf{K}^T(\mathbf{K}\mathbf{S_A}\mathbf{K}^T + \mathbf{S_o})^{-1}(\mathbf{y} - \mathbf{K}\mathbf{x_A}) \tag{2.19}$$

In the above equation, $\hat{x}$ gives us independent contributions of optimised emission values from the blowout, EPA inventories, and boundary and initial conditions from CAMS (Copernicus Atmosphere Monitoring Service). $\mathbf{x_A}$ is the prior, and it is taken as the mean of reported emission rates from the blowout. $\mathbf{K}$ is the Jacobin matrix constructed by running perturbation simulations for each contribution in $\hat{x}$. $\mathbf{S_A}$ is the prior error covariance matrix, and it is obtained by using error contributions from blowout emissions, emission inventories, and boundary conditions. $\mathbf{y}$ contains TROPOMI measurements. $\mathbf{S_o}$ is the error covariance matrix for observations, and it computes the error by calculating the standard deviation of the difference between the prior model and the observations.

In the previous paragraphs, two approaches used for quantifying the TROPOMI plumes using the WRF synthetic plumes were discussed. The first approach uses linear scaling to scale the synthetic plumes to match them to the TROPOMI plumes. The scaling parameter is based on the enhancements in the TROPOMI image and the enhancements in the synthetic plume image. The enhancements are calculated by taking an average dry air mole fraction of methane over the plume, which is equivalent to calculating integrated mass enhancement (IME) over the plume. The second approach uses the analytical solution to the Bayesian inverse problem. The analytical solution used in the second approach minimises the error between the TROPOMI enhancements and the synthetic plume enhancements, similar to minimising the root mean square error to reach an optimal solution.

The previous paragraphs described TROPOMI plume emission quantification using synthetic plumes generated by the WRF model. The FLEXPART model can generate synthetic plumes in forward run mode, which can be utilised for plume emission quantification. The FLEXPART model can also run in the backward mode, which gives us another option for plume emission quantification. The following paragraph describes how the backward run of the FLEXPART model can be used to quantify the plumes.

In backward mode, the FLEXPART model simulates particle transport from the receptor to the source (refer to Figure 2.13). In the FLEXPART backward run, a unit dry air mole fraction of species can be provided at the receptor as an input. After the backward run is complete, we get a relationship between sources and receptors in the form of a matrix [65]. The obtained relationship tells us emission contribution from each source to achieve a unit dry air mole fraction of methane at the receptor. The sources can be distributed over a three-dimensional grid or a surface grid. The surface grid is more suitable for this study since methane emissions from fossil fuel infrastructure and waste management occur near the surface. Since the relationship is linear, we can scale the emission rate at the sources to match the methane dry air mole fraction measured at the receptor [66]. In this study, we can try to match the methane dry air mole fractions observed in the TROPOMI image. We get a 2-dimensional image from TROPOMI, but we need coordinates of the receptor in terms of latitude, longitude and vertical levels. TROPOMI does not reveal the methane distribution in the vertical columns. Hence, we need to derive the vertical distribution of methane from the TROPOMI image.

The plume scaling method described in the Ohio blowout study is less mathematically intensive and much easier to set up than the Bayesian inverse solution method used in the Louisiana blowout study and the backward Lagrangian model method described in the previous paragraph. We aim to compare the plume emission quantification results of the atmospheric transport models to the mass balance methods and see whether the atmospheric transport model-based plume emission quantification approach overcomes some of the mass balance methods' inherent limitations.

3

Research Plan

The following research questions and objectives are formulated for this Master thesis project:

3.1. Research Question(s)

The most important research question for this thesis is:

> How can Eulerian and Lagrangian atmospheric transport models be used to improve the emission quantification of methane plumes detected by the TROPOMI automated plume detection algorithm?

The following sub-questions will help in answering the main research question:

3.1.1. Analysis of mass balance methods (Phase I):

The leading research question for this part is: How well can emissions be quantified through mass-balance methods?

Sub-questions:

In the mass balance approach, how is the emission estimate affected by the input parameters (plume mask filtering criteria, effective wind speed)?

Under which (meteorological) conditions do a mass balance approach performs best, or worst?

What are the limitations of using the mass balance approach for plume emission quantification?

3.1.2. Internal analysis of the WRF and FLEXPART models (Phase II):

The leading research question for this part is: How well can the (known) emissions of a simulated plume be quantified based on an independent simulation with the same model but with different settings (e.g. transport parametrisations and source location)?

Sub-questions:

How sensitive are the plumes produced by atmospheric transport models to the input parameters?

What are the computational costs of using atmospheric transport models for simulations?

What (meteorological) parameters are predictors for the uncertainty of each method?

3.1.3. Cross-validation of the WRF and FLEXPART models (Phase III):

Phase III is similar to phase II and has the same research questions, except that the emissions of a simulated plume are quantified with a different transport model.

3.1.4. Implementation on the TROPOMI cases (Phase IV):

The leading research question for this part is: Which emission quantification algorithm performs better when implemented on actual TROPOMI data?

Sub-questions:

Which model is best able to reproduce the plumes seen in TROPOMI data?

Based on our findings from all four phases, what does it mean for the potential implementation to the TROPOMI automated plume detection algorithm?

3.2. Approach

This section describes an approach to answer the research questions given in Section 3.1. A flowchart illustrating the proposed research approach is given in Figure 3.1. This figure shows three blocks titled Phase I, Phase II & III and Phase IV. Phase I is dedicated to testing mass balance approaches on the simulated synthetic plumes. Phase II & III are dedicated to testing and (cross) validating atmospheric transport models, and Phase IV will be used for applying WRF and FLEXPART on TROPOMI plumes. Based on the transport models based approach, we will provide an outlook on the implementation on the operational algorithms. The following subsections will discuss plans for the four phases in detail.

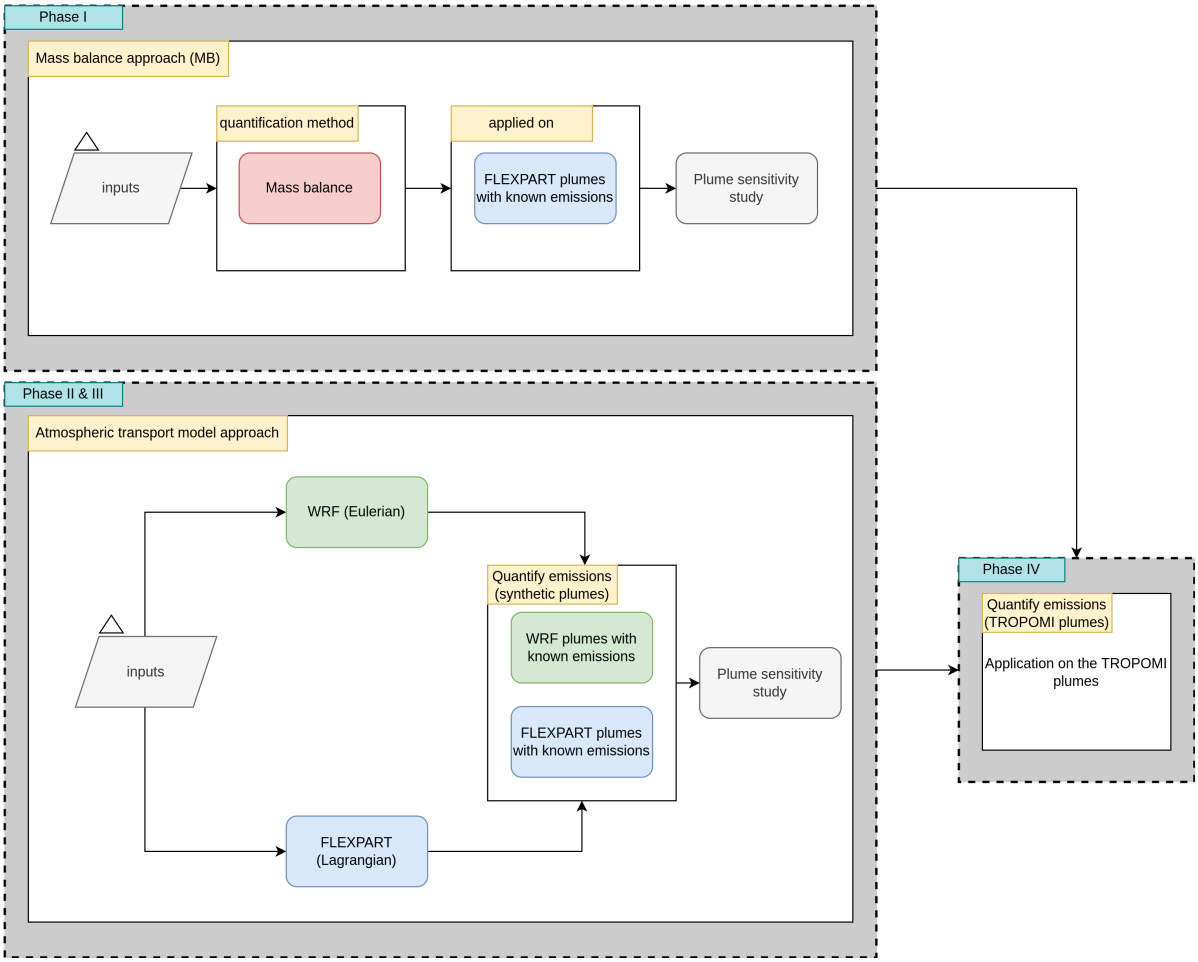


Figure 3.1: Flow diagram depicting the research approach

3.2.1. Phase I: Analysis of mass balance methods

The first phase focuses on setting up the WRF and FLEXPART models for generating synthetic plumes and using the mass balance approach for their quantification. The WRF model was chosen for this research following the suggestions from the researchers at SRON who have experience with it in the past.

By providing relevant input data to the researcher operating WRF, synthetic plumes will be obtained with time-forward simulations. The FLEXPART model was selected for this research due to its simplicity and availability of instruction material. Particle dispersion simulations in a time-forward manner will be performed in the FLEXPART model to obtain synthetic plumes.

Later, mass balance approaches will be applied to the WRF and FLEXPART-generated synthetic plumes for quantifying emission rates. The outcome from Integrated Mass Enhancement (IME) method (a simple mass balance approach) is dependent on filtering criteria (plume mask). Moreover, outcomes from the IME and Cross-sectional Flux (CSF) are dependent on an empirical relation used for computing effective wind speed. Currently, the empirical relation for effective wind speed is based on a high-resolution LES model, which does not hold well for low-resolution TROPOMI data. Hence, by varying the filtering criteria and the empirical relation, we will obtain uncertainties related to the mass balance approaches. The purpose of this exercise is to 1) validate the implemented mass balance approaches and 2) observe the limitations of using mass balance approaches.

3.2.2. Phase II: Internal analysis of the WRF and FLEXPART models

The second phase will deal with analysing the WRF model and FLEXPART model internally. To be more specific, the WRF model will be used to generate a target (synthetic) plume for an independent WRF model simulation (with a different set of inputs). Based on the best match between the target plume and plume obtained from an independent WRF run, the emission rate of the source will be determined. Since the emission rate of the synthetic plume is known, the emission rate we determine from the simulation can be compared to the true emission rate for obtaining the first error estimate. Similar operations will be performed for the FLEXPART model.

Currently, we are considering only the parameterised boundary layer and wind field data as important input parameters that can be varied. In later stages, more input parameters might be considered. A sensitivity study based on variations in the input parameters will be performed on the WRF and FLEXPART models. Sensitivity analysis will be one of the ways to quantify uncertainties in WRF and FLEXPART models. For comparing the results obtained from WRF and Lagrangian models, the following metrics will be used:

Comparative study based on plume shape

The simulated plume and target plume will be compared based on visual cues like alignment, size and enhancements in the pixels. A more sophisticated method involving cross-correlation can also be applied to find the best match.

Comparative study based on uncertainties in the outcomes

As mentioned previously, predictions from the WRF-CHEM model depend on parametrised boundary layer (empirical model) and meteorological data (4-dimensional wind field). Boundary layer parameters affect the calculation of vertical wind velocity, which, in turn, has a larger influence on the outcome of the WRF-CHEM model. Also, using different meteorological data (ERA5 instead of GEOSFP) or sampling the data by considering a different time affects the results. These factors will be used while quantifying uncertainties in the WRF-CHEM model. Similar to WRF-CHEM, FLEXPART also requires parametrised boundary layer equations and meteorological data. Hence, the uncertainty quantification methods used for the WRF-CHEM will be implemented for FLEXPART too. Additionally, FLEXPART can perform forward as well as backward particle trajectory simulations. The effects of these approaches on the predictions will also be studied.

Comparative study based on the computational efficiency of models

In this study, the WRF and FLEXPART methods will be compared based on the computation requirements. Large computational requirements are undesirable, and if needed, suitable approximations will be made to reduce the computational cost. Plans for the third and the fourth phases will depend on the results from the first and second phases (what uncertainty quantification metrics could be relevant in this study, what input parameters play a larger role in the sensitivity of the results, etc.) hence, they are made flexible for now.

3.2.3. Phase III: Cross-validation of the WRF and FLEXPART models

This phase is dedicated to performing cross-validations of the WRF and FLEXPART models. Here, the synthetic plume from the FLEXPART model will be used as a target plume for the WRF simulation and vice versa. Comparison metrics mentioned in the previous phase will be used for comparing the performance of the models (if the internal analysis phase gives us new error metrics for predicting uncertainties, we will also include them in this phase).

3.2.4. Phase IV: Applying the atmospheric transport model-based quantification approach to the TROPOMI plumes

In the final phase of the thesis, the WRF and FLEXPART models will be applied to real TROPOMI plumes. The performance of these models will be compared based on how well these models will be able to reproduce the TROPOMI plumes and other error metrics discussed in the previous parts.

4

Synthetic plumes

In Section 2.6, we mentioned that mass balance methods are currently used for emission quantification of plumes detected by the TROPOMI automated plume detection algorithm. Mass balance methods provide emission quantifications with relatively large uncertainties. Our goal is to improve plume emission quantification using atmospheric transport models. For this purpose, we first analyse mass balance methods to determine their limitations. Later, we analyse plume emission quantification using atmospheric transport models and see whether they can overcome some of the limitations of mass balance methods to improve plume emission quantification.

To analyse a particular plume emission quantification approach, we need multiple plumes with known emission locations and emission rates. The TROPOMI automated plume detection algorithm frequently captures methane plumes over the globe, but we often do not know those plumes' actual emission locations and emission rates. Atmospheric transport models can simulate atmospheric particle transport, and hence they can be used to generate plumes. We can transform atmospheric transport model outputs into TROPOMI-like plumes, and they can be used to analyse plume emission quantification approaches. Furthermore, atmospheric transport models give us flexibility in selecting emission locations and emission rates of plumes. It allows us to study how different regions or plumes with different emission rates affect plume emission quantification. Therefore, we use atmospheric transport models to generate multiple plumes required to analyse plume emission quantification methods.

We use the WRF and FLEXPART atmospheric transport models to generate plumes. The WRF model uses an Eulerian frame of reference to simulate particle transport, and the FLEXPART model uses a Lagrangian frame of reference to simulate particle transport (Section 2.7). We call the plumes generated by atmospheric transport models synthetic plumes in this study. This chapter discusses the setup used to generate synthetic plumes using atmospheric transport models, which will be used in Chapter 5 to assess how well the different emission quantification methods are performing.

4.1. General settings of atmospheric transport models

One component of this research is to analyse how different meteorological data fields, different physics settings (planetary boundary layer schemes) in atmospheric transport models, and using different atmospheric transport models affect the generated synthetic plumes (Chapter 3). This section mainly focuses on these settings of atmospheric transport models. In Figure 4.1, we present a general flowchart that illustrates the generation of synthetic plumes using the WRF and FLEXPART atmospheric transport models.

FLEXPART is an offline atmospheric transport model, and it requires pre-processed meteorological data. In contrast, WRF is an online model that uses pre-processed meteorological data as boundary conditions, but the meteorological model takes care of the transport within the domain. Both the WRF and FLEXPART models require pre-processed meteorological data to carry out particle transport simulations. Raw meteorological data might lack vertical wind velocities. Hence, the raw meteorological data often needs to be prepossessed to compute the vertical velocities before it is supplied to a model.

For the WRF model, we use NCEP meteorological fields [55] with 6-hour temporal resolution, 21 vertical levels, and $1^{\circ} \times 1^{\circ}$ horizontal spatial resolution. For the FLEXPART model, we use the previously mentioned NCEP meteorological fields as well as ECMWF meteorological fields [54] with 1-hour temporal resolution, 138 vertical levels, and $1^{\circ} \times 1^{\circ}$ horizontal spatial resolution. The selection of NCEP and/or ECMWF meteorological fields were largely based on availability and compatibility with the atmospheric transport models.

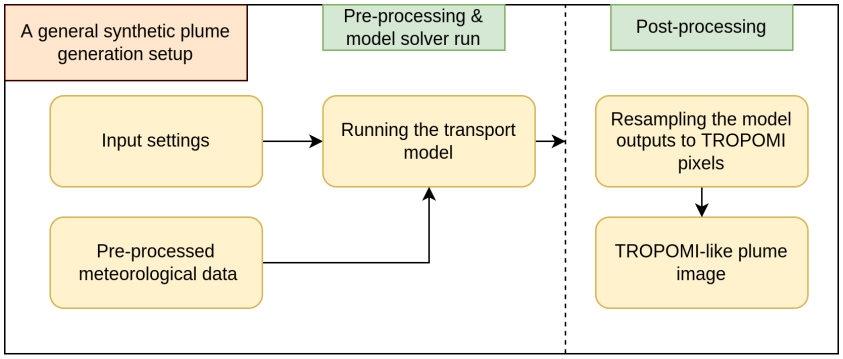


Figure 4.1: General setup of an atmospheric transport model to generate synthetic plumes. This figure depicts three stages involved in a typical atmospheric transport model simulation (on the left-hand side of the dotted line) which are a) pre-processing of meteorological data b) setting input parameters, and c) running the transport model. The right-hand side of this figure shows a post-processing step after the model solver run is complete.

The way meteorological data is used in the WRF and FLEXPART models is different. The WRF model computes its own wind fields, which are nudged to the pre-processed meteorological data at the outer domain boundaries, whereas the FLEXPART model makes use of meteorological wind speeds directly to carry out particle transport. Hence, we expect that using different meteorological data; we get synthetic plumes that are different. Apart from pre-processed meteorological data, both the atmospheric transport models require several input settings. Some input settings are very crucial for this study and are common for both the WRF and FLEXPART models. These input settings are described in the following paragraphs.

One common setting for both models is the definition of the computational domain. The WRF model transports particles in an Eulerian frame of reference, and it requires a computational grid to perform calculations at every time step. For WRF, we define an inner domain with a grid having finer resolution (0.03°) than a typical TROPOMI pixel resolution (0.05°) and surround it with two outer grids with coarser resolutions. We kept the resolution of the (inner) grids finer than a typical TROPOMI pixel resolution since we later transform the atmospheric transport model outputs to create a TROPOMI-like image, and this involves interpolations of the outputs. On the other hand, for the FLEXPART model, we require a computational grid only at the last time step to distribute output methane concentrations, and the resolution of the computational grid does not affect particle transport simulations. We keep the horizontal resolution of the FLEXPART grid finer (0.02°) than a typical TROPOMI pixel resolution (0.05°). Examples of the WRF inner domain and the FLEXPART domain are provided in Figure 4.3.

Next, we define emissions in the WRF and FLEXPART models to generate synthetic plumes. In the WRF model, we define persistently emitting point sources with emission rates of 7.23 [t/hr], and in the FLEXPART model, we define persistently emitting point sources with emission rates of 40 [t/hr]. We assume emissions from point sources in atmospheric transport models are scalable, and this is generally true if the problem does not include chemical reactions in the atmosphere; hence output concentrations of methane can be multiplied with some constant value to scale the input emission rate. Scaling of emissions from point sources is necessary to obtain synthetic plumes with varying emission rates, which will be required for the analysis of the mass balance methods.

A common physics setting required for both the transport models is the planetary boundary layer parameterisation scheme. We consider parametric boundary layer setting important for the atmospheric transport models since it has a strong influence on the vertical transport of air parcels, which can affect the plume geometries. We select two separate planetary boundary layer (PBL) schemes in

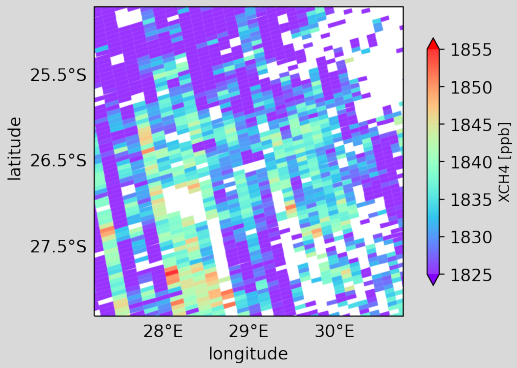
the WRF model that are most easily available: the Yonsei University Scheme (YSU) and the Mellor-Yamada-Janjic Scheme (MYJ). Similarly, we select two separate planetary boundary layer schemes in the FLEXPART model that are available: the Gaussian turbulence parametrisation and the Skewed turbulence parametrisation. The PBL schemes in the WRF model are different from the PBL schemes in the FLEXPART model since the same schemes are not available in both models.

After the input settings are defined, atmospheric transport model simulations can be performed to generate synthetic plumes. The outputs from atmospheric transport models contain methane concentrations distributed over a three-dimensional grid. We post-process the three-dimensional methane concentration outputs using an algorithm to obtain two-dimensional TROPOMI-like images. The algorithm used for this purpose is called the resampling algorithm. The resampling algorithm was discussed in Section 2.9, and it is used to convert the methane concentrations distributed over multiple vertical atmospheric levels to an image having TROPOMI-like pixels. In the next section, we compare TROPOMI-like images containing synthetic plumes generated with the WRF and FLEXPART atmospheric transport models to an actual TROPOMI image with a plume.

4.2. Comparison between TROPOMI plume and synthetic plume

Figure 4.2 shows an actual TROPOMI image on the left side, a resampled synthetic plume image generated by the WRF-CHEM model in the middle, and a resampled synthetic plume image generated by the FLEXPART model on the right side. When we compare a TROPOMI plume to synthetic plumes generated with atmospheric transport models, several differences can be noticed. A major difference is the presence of noise. TROPOMI measurements are affected by instrumental noise, artefacts, correlated noise, presence of stripes in the data, and proximity to bad coverage region (costs and cloud edges) (Box 4.1). The presence of noise in TROPOMI images affects the selection of plume pixels, and plumes having low emission rates might be challenging to identify within the noise. We want to replicate details seen in the TROPOMI images in synthetic plume images. However, it is challenging to replicate the noise present in TROPOMI images to synthetic plume images because there is not any definite pattern of that noise. We can partially account for TROPOMI noise by adding Gaussian noise to synthetic plume images. We do this in the analysis of mass balance methods.

Box 4.1: TROPOMI noise



This figure shows a typical TROPOMI image containing measurements taken on 2020-09-26. Some noise can be seen in the figure. The presence of noise affects the measurements in the TROPOMI pixels, which makes plume masking challenging and adds uncertainty to the emission estimates. Noise in the TROPOMI images can be found using visual information like random ups and downs in the methane concentration values around the mean concentration value, similarity to ground artefacts, and correlation with aerosols or albedo etc.

Fluctuations in the temperature of detectors used on the TROPOMI instrument can also introduce noise in the measurements, which is a part of the instrumental noise. It was also seen that a detector could give faulty readings, which introduces stripes in the data, which can also be seen in the image given in this box. A de-striping algorithm was developed at SRON to take care of stripes present in the data. The presence of clouds affects the readings, and pixels that fall over clouds are often missing from the data. Similarly, retrieving data over water bodies is difficult because there is very less solar backscatter, and those pixels are also missing from the data. Pixels that fall at the edge of clouds or water bodies might not be removed, and those pixels will have some influence due to clouds or water bodies. Usually, we see that pixels in the proximity of bad coverage regions (coasts and cloud edges) have

very low methane concentration values.

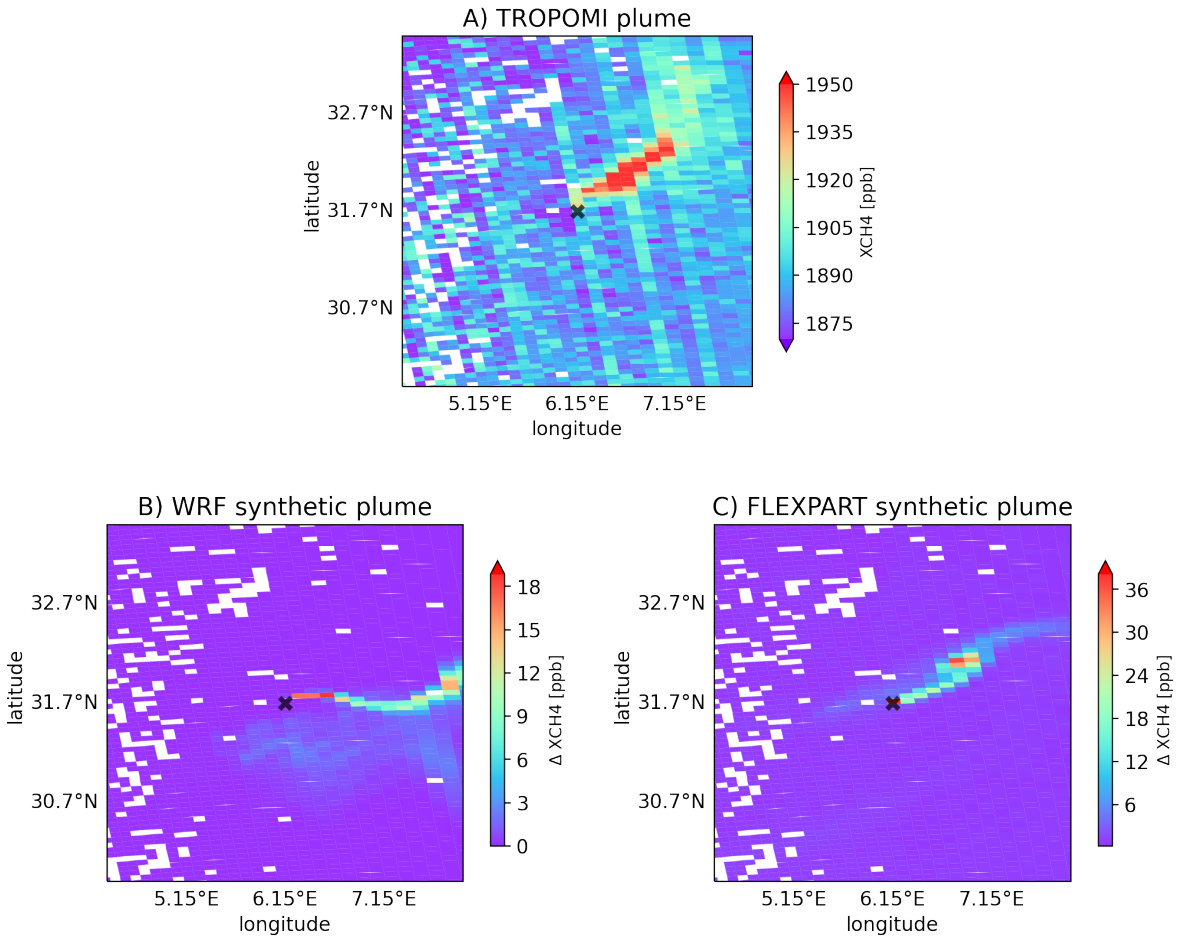


Figure 4.2: This figure shows A) an example of a TROPOMI plume image for measurements taken on 03-01-2020. Pixels of this image contain dry air mole fraction of methane. There are some white/transparent pixels in this image which indicate the non-availability of TROPOMI measurements due to cloud coverage. B) An example of a TROPOMI sampled synthetic plume generated with the WRF atmospheric transport model with an emission rate of 7.23 [t/hr]. Pixels in this image represent enhancement in the dry-air mole fraction of methane with respect to the background dry-air mole fraction of methane. C) An example of TROPOMI sampled synthetic plume generated with the FLEXPART atmospheric transport model with emission rate 7.23 [t/hr]. Similar to WRF synthetic plume image, pixels represent enhancement in dry-air mole fraction of methane. The black cross markers indicate the emission locations. Missing pixels in all the images are the same, this is due to the resampling algorithm.

The previous sections discussed how we obtain TROPOMI-like (synthetic) plumes using the WRF-CHEM and FLEXPART models. The resampling algorithm projects the outputs obtained from atmospheric transport models to TROPOMI pixels using the sizes, locations and orientations of those TROPOMI pixels, and this involves interpolation. Some TROPOMI images might have missing pixels due to cloud coverage, and the TROPOMI-like images produced by the resampling algorithm also have the same missing pixels. Hence, the locations and days on which synthetic plumes are generated are constrained by the coverage (available pixels) in TROPOMI measurements. If TROPOMI coverage is very bad (a large number of missing pixels in TROPOMI measurements), TROPOMI-like images generated by the resampling algorithm will also show bad coverage. Keeping this in mind, we select several locations and days having good TROPOMI coverage for simulating synthetic plumes.

4.3. Selection of locations and days for synthetic plumes

We notice frequent methane plumes in TROPOMI measurements over Algeria and over a region near Johannesburg in South Africa which contains a SASOL plant. The latter will be referred to as the Sasol region in this report. From manual investigation, we found that coverage of TROPOMI over the aforementioned regions is good. Hence, in this study, we pick a region in Algeria and a region near Sasol to generate synthetic plumes. The average elevation of land in the Algeria region is 1200 [m], whereas the average elevation of land in the Sasol region is 1700 [m]. Furthermore, the Sasol region sees more rain than the Algeria region, and the weather patterns of these regions are different. Generating synthetic plumes over these two regions may help to understand the effects of different surface elevations and weather synoptic on synthetic plumes and on their emission quantification.

Figure 4.3 shows the WRF and FLEXPART domains over Algeria and Sasol. We pick several emission sites inside these domains and perform separate methane release simulations from each site to generate separate synthetic plumes. Days for simulations are selected based on days with good TROPOMI coverage (> 75%) over the domains. From the atmospheric transport model simulations, we obtain an inventory of synthetic plumes for multiple days in 2020. This inventory is later used for analysing plume emission quantification methods.

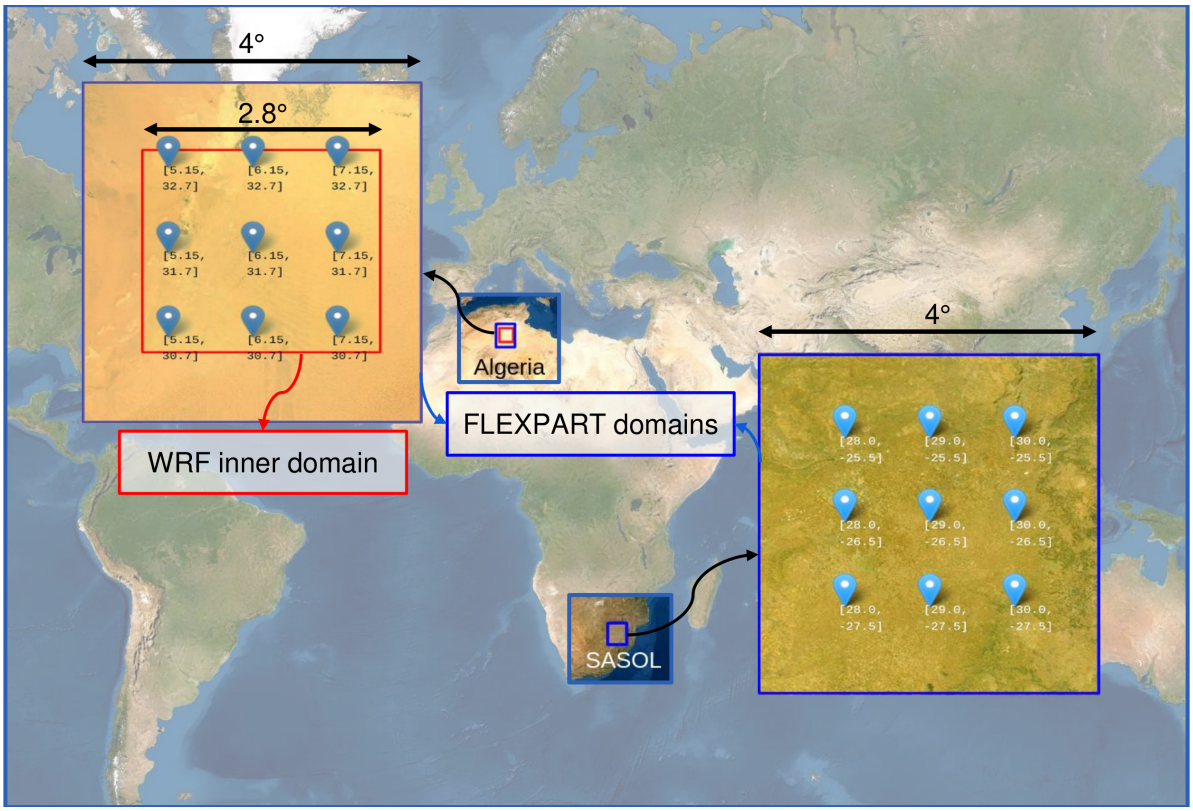


Figure 4.3: Square computational domains defined for the WRF and FLEXPART models over Algeria and around SASOL represented over a satellite image (folium map). The red box over Algeria shows the inner domain defined for the WRF model. The top left coordinates (longitude, latitude) of the red box are (4.86°, 32.92°), and the bottom right coordinates of the red box are (7.67°, 30.53°). Please note that we use the WRF model to simulate synthetic plumes over Algeria only. The blue boxes over Algeria and around SASOL show FLEXPART domains. For Algeria, the top-left coordinates of the blue box are (4.15°, 33.7°), and the bottom right coordinates of the blue box are (8.15°, 29.7°). For Sasol, the top-left coordinates of the blue box are (27°, -24.5°), and the bottom right coordinates of the blue box are (31°, -28.5°). The Blue location markers in zoomed-in pictures indicate several plume emission sites used in the models. Please note that these emission sites are just for representation purposes and do not include all the emission sites considered for this study. A detailed list of emission sites used for generating synthetic plumes is provided in appendix A.2. Coordinates mentioned with the markers represent the longitude and latitude of a particular location in degrees.

In phase I of this study (phases of this study are described in Section 3), we aim to investigate the limitations of mass balance methods by quantifying synthetic plumes generated by the setup described

in this section. For phase I, we consider eight days in 2020 and nine emission sites over the Algeria region, eight days (different than what we considered for Algeria) and nine emission sites over the Sasol region for generating synthetic plumes. Later, in phases II and III, we analyse how atmospheric transport models can overcome the limitations of mass balance methods in plume emission quantification. For phases II and III, we extend the synthetic plume data by considering nearly 200 days (or TROPOMI orbits) over 2020. Please note that for phases II and III, we relaxed the TROPOMI coverage threshold to 30%, which was 75% in phase I. We present the setup of the mass balance methods and atmospheric transport model based plume emission quantification approach in the following chapter.

5

Plume emission quantification methods

In the previous chapter, we discussed the setup used to generate synthetic plumes using atmospheric transport models. An assessment of the quality of mass balance methods and atmospheric transport model based plume emission quantification approaches will now be done by quantifying the emission for the various synthetic plumes. This chapter discusses how we apply the different emission quantification methods to synthetic plumes.

5.1. Plume emission quantification using mass balance methods

The TROPOMI automated plume detection algorithm uses the Integrated Mass Enhancement (IME) and the Cross-Sectional Flux (CSF) method to quantify emissions of methane plumes (see Section 2.6). The IME and CSF methods use the mass conservation principle in 2-dimensions to quantify emissions of plumes. This can be done in different ways, which results in different implementations of the IME and CSF methods. In this study, we consider four implementations of the IME method and two implementations of the CSF method. These implementations are described in the following subsection.

5.1.1. Various implementations of IME and CSF methods

One way to isolate plume pixels from the background pixels is to determine the standard deviation in the methane values across the studied domain. All data below the mean plus some multiple of standard deviation are considered part of the background. A study by Schuit [44] found that filtering values below mean plus 1.8 times standard deviation worked well for TROPOMI methane plumes. TROPOMI images contain noise (Box 4.1), and if we use only a standard deviation filter, there is a chance we might consider the out-lying high-value pixels (that are not part of the plume) as part of the plume. To overcome this problem and put more constraints on the selection of plume pixels, several implementations of mass balance methods based on the selection of plume pixels are developed at SRON. In this analysis, we consider four implementations of the IME method and two implementations of the CSF method. The four implementations of the IME method are: IME plume dilation method, IME plume box method, IME rotated plume box method and IME plume circle method. Two implementations of the CSF method are the CSF method without rotation and the CSF method with rotation. With different implementations of the IME and CSF methods, we can study the effects of applying different plume masks (selection of plume pixels) on mass balance method estimates. The following paragraphs provide brief details about these implementations.

IME plume dilation method

To consider only the pixels connected to the plume and not the out-lying high-value pixels, the plume dilation method was developed. This method uses an iterative method to mask the pixels. Initially, all the pixels are left out of the mask. In the first iteration, a mask is dilated in the outward direction by one pixel from the highest methane column value. In the newly dilated part of the mask, if the value of a pixel is below a certain threshold standard deviation, then that pixel is kept out of the mask. Only the

pixels in the dilated part of the mask having values that are above the threshold are considered part of the plume. This process is repeated till convergence is reached, or a maximum number of iterations is met. Iterations of IME plume dilation method are illustrated in Figure 5.1.

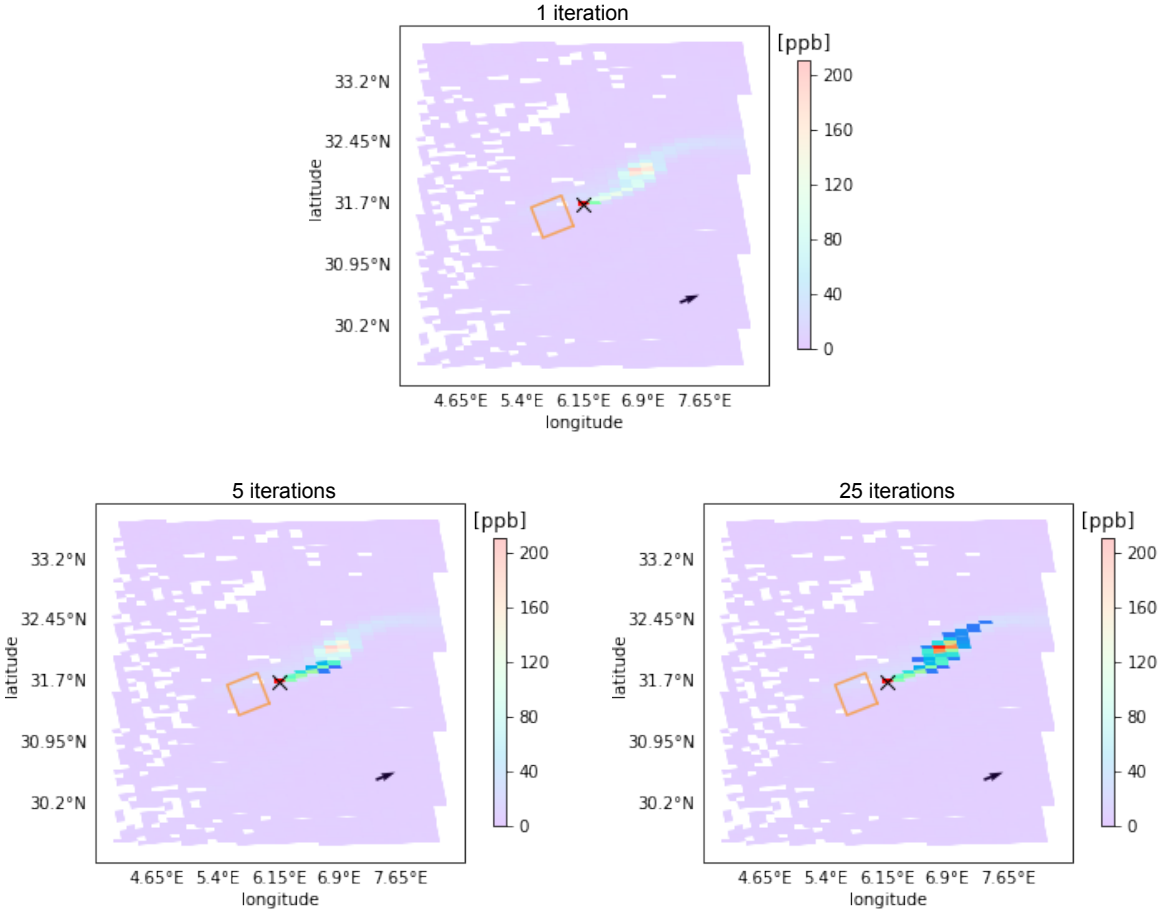


Figure 5.1: Illustrations of IME plume dilation method on a FLEXPART generated and TROPOMI-resampled plume image. The top image shows plume mask after 1 iteration of plume dilation algorithm, the bottom left image shows plume mask after 5 iterations of plume dilation algorithm, and the bottom right image shows plume mask after 25 iterations of plume dilation algorithm. Colorbar shows excess dry-air mole fraction of methane in [ppb]. The black cross is the plume source location. The black arrow indicates wind direction at plume source location, and length of the arrow indicates 10 [m] wind speed at the plume source location obtained from the NCEP meteorological data. Background pixels are made semi-transparent to highlight plume pixels. An orange box in the upwind direction of the source location is used for calculating background dry-air mole fraction of methane, which is required in plume emission quantification process.

IME plume box method

IME plume box method is illustrated in Figure 5.2. In this method, an area is defined with the help of a box that starts at the plume source location and is oriented in a downwind direction from the source location. Pixels that fall inside the box and have values higher than a certain threshold standard deviation value are considered part of the plume.

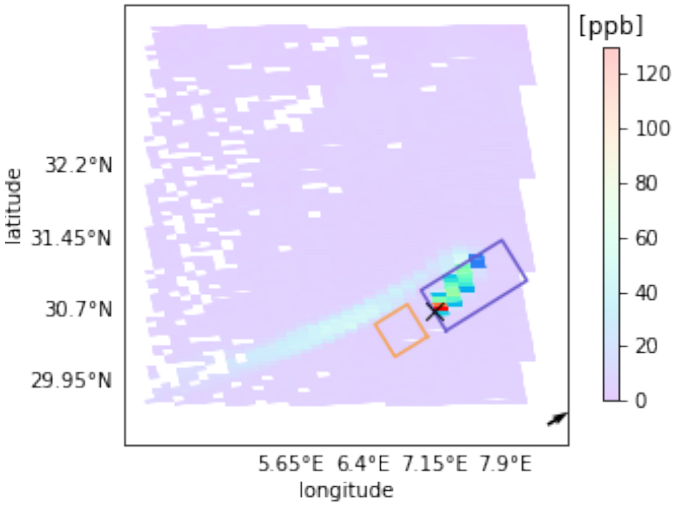


Figure 5.2: This figure illustrates IME plume box method. The blue box in this figure indicates a region from which plume pixels are picked. The blue box is aligned with the wind direction. All pixels outside the blue box are considered non-plume/background pixels. Similar to plume dilation method, orange box indicates region from which background pixels are picked, black arrow indicates wind speed and direction, and black cross indicates plume source location.

IME rotated plume box method

The IME plume box method is limited by the applicability of the wind-direction at the source location across the full plume. Therefore the IME rotated plume box method aligns the box towards highly enhanced plume pixels. IME rotated plume box method is illustrated in Figure 5.3.

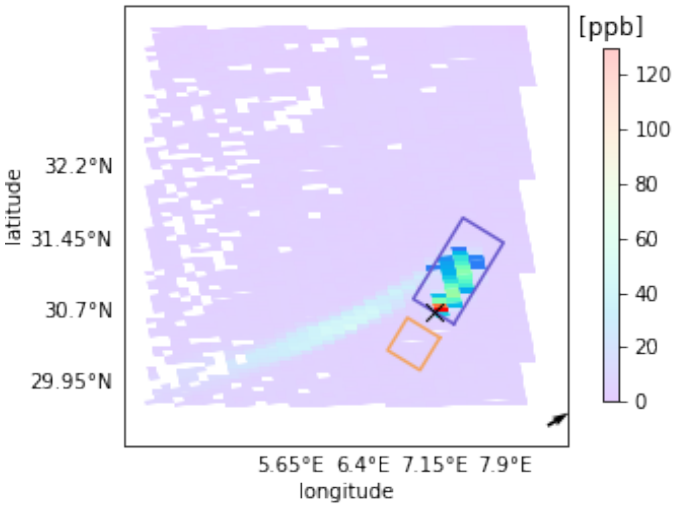


Figure 5.3: This figure illustrates IME rotated plume box method. The description of this figure is the same as what we described for IME plume box method in Figure 5.2.

IME plume circle method

Plumes might have curved geometries, and using a plume box might be limiting for those cases. Hence, the IME plume circle method was developed. IME plume circle method is illustrated in Figure 5.4. In this method, instead of a box, a circle is drawn around the source location of a plume. All pixels that have values higher than a standard deviation threshold inside the circle are considered part of the plume. The rest of the pixels are masked and not considered part of the plume.

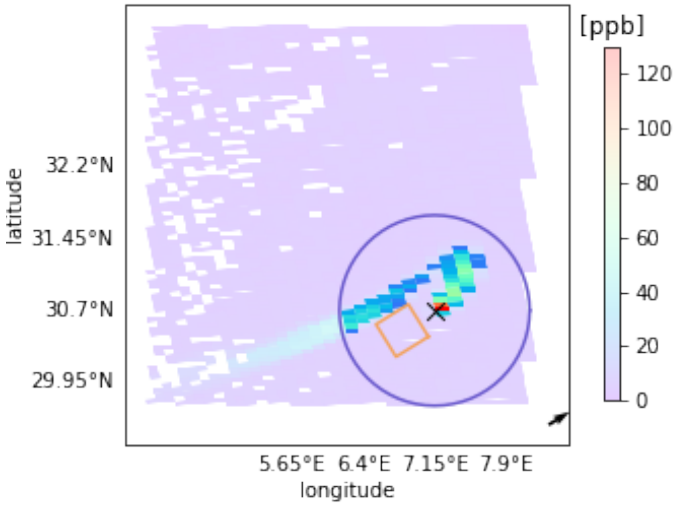


Figure 5.4: This figure illustrates IME plume circle method. Blue circle in this figure indicates a region that contains plume pixels. All pixels outside the blue circle are considered non-plume/background pixels. Similar to plume dilation method, orange box indicates region from which background pixels are picked, black arrow indicates wind speed and direction, and black cross indicates plume source location.

CSF method without rotation

CSF method without rotation is illustrated in Figure 5.5. In this method, transect lines are oriented perpendicular to the source location wind direction. The first transect line near the source/emission location is not considered in the quantification process. Only the transect lines that fall over pixels with values higher than a certain standard deviation threshold and having more than 70% of their lengths falling over pixels are considered for plume emission quantification.

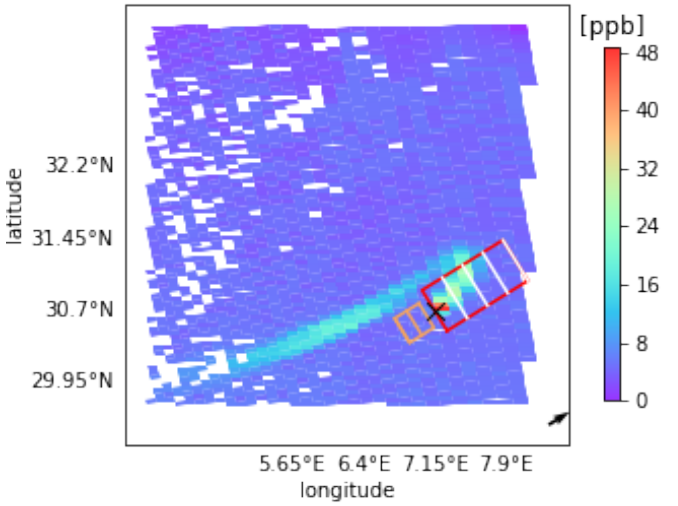


Figure 5.5: This figure illustrates CSF method without rotation. The red box indicates region from which plume pixels are selected. White transect lines are oriented perpendicular to the wind direction at the plume source. The orange box in the upwind direction of plume source indicates background box, which also has transect lines. Similar to the plume box method, the black cross indicates plume source location, and the black arrow indicates wind direction and speed at the plume source location. For illustration purposes, only 5 transect lines are shown, in practice, we are using 20-25 transect lines since we noticed convergence in quantification results with those values.

CSF method with rotation

CSF method with rotation is illustrated in Figure 5.6. The principle behind this method is similar to the principle of the CSF method without rotation, with the only difference being that the transect line box is now rotated such that it aligns with the plume instead of the source location wind direction.

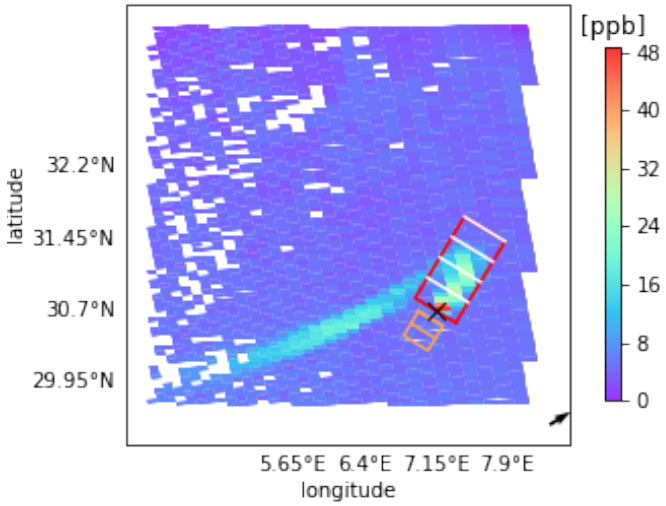


Figure 5.6: This figure illustrates CSF method with rotation. Rest of the description of this figure is similar to Figure 5.5.

We use the previously mentioned implementations of mass balance methods to quantify synthetic plumes. Mass balance methods require some inputs before performing plume emission quantification. These inputs and a general setup of mass balance methods are described in the following subsection.

5.1.2. Setup of mass balance methods

We aim to determine several limitations of mass balance methods by quantifying multiple synthetic plumes. Figure 5.7 shows a flowchart for synthetic plume emission quantification using mass balance methods. The flowchart is divided into three parts: inputs, application of mass balance methods, and outputs. The obtained outputs are later analysed to determine the limitations of mass balance methods. Main inputs required for mass balance methods include TROPOMI sampled synthetic plume, plume source location, satellite-specific parameters or values used in the settings of mass balance methods, and 10 [m] wind field data. These inputs are explained in the following paragraphs.

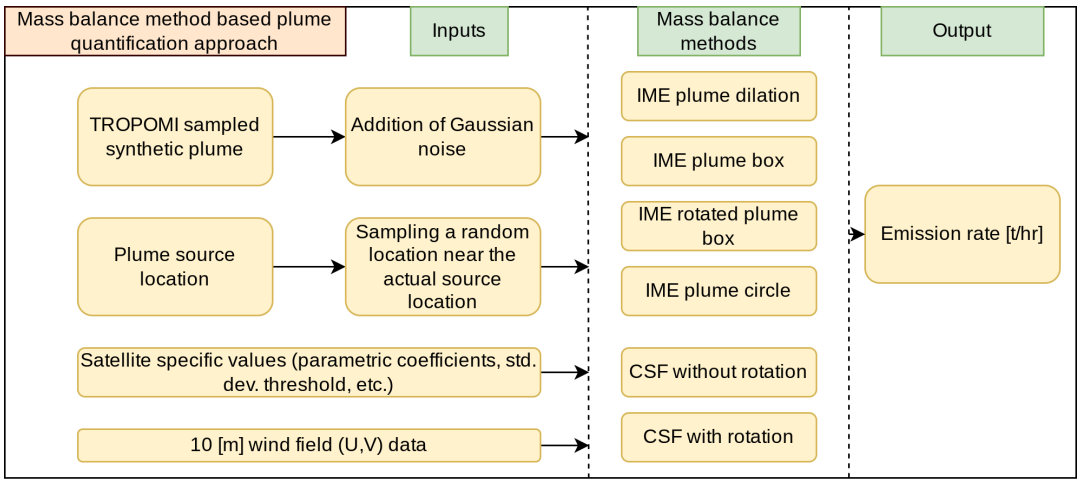


Figure 5.7: This figure shows a flowchart for quantifying synthetic plumes' emissions using implementations of mass balance methods. This flowchart is divided into three parts separated by dotted lines, and those three parts correspond to: inputs, application of mass balance methods, and outputs. This setup outputs the emission rate estimate of synthetic plumes obtained from the different implementations of the mass balance methods.

TROPOMI resampled synthetic plume

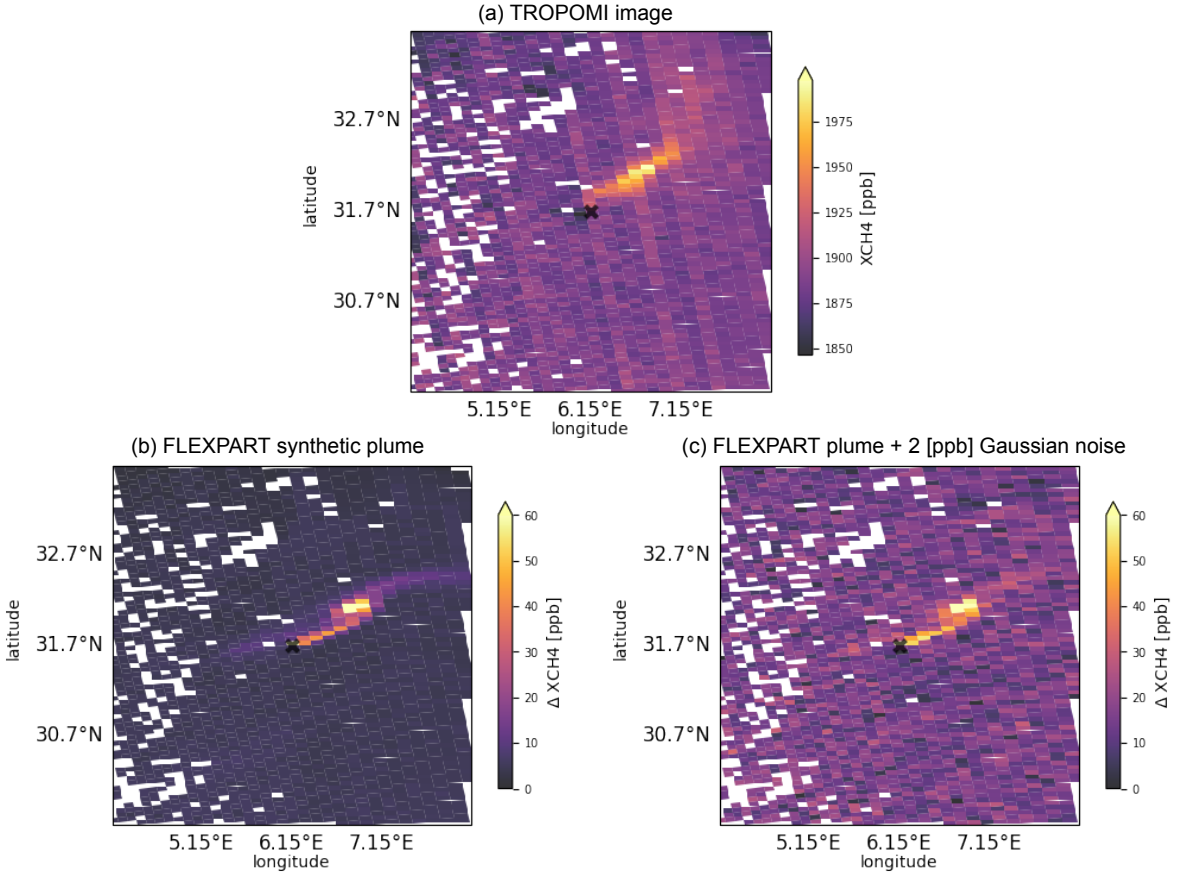


Figure 5.8: This figure shows (a) Methane total column concentrations as observed by TROPOMI over an area in Algeria on 2020-01-03, (b) Corresponding methane total column enhancements as simulated by the FLEXPART model, and (c) as in (b) but including 2 [ppb] Gaussian noise.

We obtain a three-dimensional distribution of methane from atmospheric transport model simulations. This output is resampled to match TROPOMI observations. In Figure 5.8, we can see an actual TROPOMI image at the top and a TROPOMI-resampled synthetic plume image on the bottom left side. It can be noticed that the TROPOMI observations contain noise. Some factors that can potentially affect measurements in TROPOMI pixels are artefacts, correlated noise, random noise, stripes in data, and proximity to bad coverage regions (Box 4.1). A typical resampled FLEXPART output (Figure 5.8) does not include TROPOMI noise effects. We partially account for uncertainty due to TROPOMI noise by adding Gaussian noise in resampled synthetic plume images. Although adding Gaussian noise can not replicate all the effects of the TROPOMI noise, it is the best we can do at this point. A resampled synthetic plume image with Gaussian noise can be seen on the bottom right side in Figure 5.8.

To know how much Gaussian noise should be added to resampled synthetic plume images, we tried to estimate the amount of noise present in TROPOMI images. We chose several TROPOMI plume cases and filtered out plume pixels using a standard deviation filter. Then, we computed the standard deviation of values in the remaining pixels to estimate the noise. From several TROPOMI scenes, we could obtain a noise estimate of 6.2 ± 2.4 [ppb]. Therefore, we decide to add Gaussian noise of 5 [ppb] to synthetic plumes.

Plume source location

Most of the implementations of IME and CSF mass balance methods require source locations of the plumes. We often do not know the accurate source location of TROPOMI plumes, and we have to rely on estimated source locations in the plume emission quantification process. Estimated source locations

add uncertainty to the quantification results. We try to account for this uncertainty by randomly sampling multiple locations around the known source location of a synthetic plume. We sample source locations inside a region defined by a circle of radius 0.1° (the approximate size of a TROPOMI pixel) centred at the actual source location of the synthetic plume (Figure 5.9). Choosing the radius of the circle equal to the approximate size of a TROPOMI pixel might not completely account for the uncertainty due to the selection of source location, but it can give us a first-order estimate of the uncertainty. We found that around 50 random samples are enough to cover the circular region uniformly and to reach convergence in mass balance quantification results.

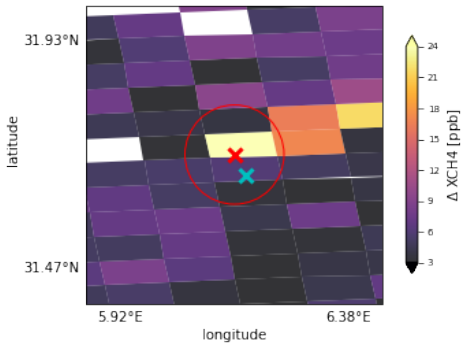


Figure 5.9: This figure shows a zoomed-in part of a FLEXPART generated synthetic plume with an actual plume source location in red marker, and randomly picked plume source location in the vicinity of actual plume source location in cyan marker. The vicinity is defined by a red circle of radius equal to 0.1 degrees (approximate TROPOMI pixel size).

Satellite specific values

Measurements from different satellite instruments might have different resolutions. IME and CSF mass balance methods use several parameters that are dependent on the resolution of pixels and hence are specific to a particular satellite instrument. In this study, we use measurements taken by the TROPOMI satellite instrument, and hence we need to set some parameters that are linked to the resolution of the TROPOMI instrument. The satellite specific parameters include parametric coefficients to calculate effective wind velocity (Section 2.6), threshold for plume mask, and dimensions of plume box/ dimensions of plume circle/ number of transect lines/ number of iterations for dilation.

Effective wind velocity is used to describe wind speed over plume (see Section 2.6), and it was mentioned that coefficients for effective wind speed equation are different for IME and CSF methods. A past study done at SRON found a linear relationship (equation 5.1) between 10 [m] wind data U_{10} and effective wind velocity U_{eff}. In the past study, a calibration process was performed to obtain coefficients of the linear relation, which involved the use of model-simulated plumes and meteorological wind speeds. For the IME method, α_1 was found to be 0.444, and α_2 was found to be 0.278. For the CSF method, α_1 was found to be 1, and α_2 was found to be 0. We check whether these coefficients are suitable for this study by performing recalibration using the model-simulated plume data we generated for this study.

$$U_{eff} = \alpha_1 U_{10} + \alpha_2 \tag{5.1}$$

We set some threshold to isolate plume pixels from non-plume or background pixels. A study done by Schuit [44] found that a threshold of mean plus 1.8 times the standard deviation to separate plume pixels with the background pixels works well for the TROPOMI images. We are using the same threshold in our study.

We set some parameters that define the dimensions of the plume box/ plume circle, number of transect lines for the CSF method, maximum number of iterations for the IME plume dilation method, etc. All these parameters and their values that were found to be suitable for this study are summarised in Table 5.1.

Table 5.1: TROPOMI instrument specific parameters used in the mass balance methods for quantifying synthetic plumes. Please note that these settings were determined from manual observations.

Parameter name	Value	Unit
Maximum number of iterations for dilation[1]	25	-
Plume box length[2,3,5,6]	1	degree
Plume box width[2,3,5,6]	0.5	degree
Background box length[2,3,5,6]	0.4	degree
Background box width[2,3,5,6]	0.4	degree
Plume circle radius[4]	1	degree
Number of transect lines for CSF[5,6]	20	-
Minimum coverage fraction for transect line[5,6]	0.7	-
Methods corresponding to superscripts:		
[1]: IME plume dilation method	[4]: IME plume circle method	
[2]: IME plume box method	[5]: CSF without rotation method	
[3]: IME rotated plume box method	[6]: CSF with rotation method	

wind field data

10 [m] wind field data contain horizontal velocity components (U, V) at 10 [m] height from the surface. We use ECMWF and NCEP meteorological datasets to generate synthetic plumes. 10 [m] wind speeds from the same two datasets are used to obtain effective wind speeds.

5.2. Plume emission quantification using atmospheric transport models

Section 2.10 discussed several atmospheric transport model based plume emission quantification methods, which include: the plume scaling approach used in a study performed to quantify emissions from a gas well blowout event that happened in Ohio [43], the Bayesian inverse solution approach used in a study performed to quantify emissions from a natural gas well blowout event that happened in Louisiana [64] and backward Lagrangian model simulation approach. The plume scaling approach was found to be less complex, easier to set up, and closer to the previously used mass balance approaches. Hence, we consider the plume scaling approach in this study. The following subsection discusses the setup used for quantifying emissions from plumes using the plume scaling approach.

5.2.1. Setup of plume emission quantification with atmospheric transport models

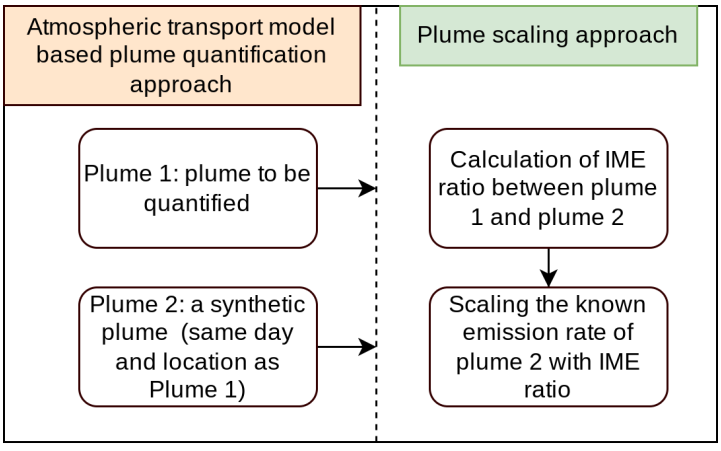


Figure 5.10: This figure shows a flowchart of plume emission quantification using atmospheric transport model based plume scaling approach. On the left side of the dotted line, inputs are mentioned. The plume scaling approach requires a plume to be quantified and a synthetic plume generated for the same day and location. IMEs of both plumes are calculated and are used to scale the known emission rate of the synthetic plume to estimate the emission rate of the other plume.

Figure 5.10 shows a simple flowchart of how the plume scaling approach works. In the plume scaling approach, we use a plume generated with an atmospheric transport model having a known emission rate to quantify another plume. This other plume could be a TROPOMI plume or a synthetic plume generated with different settings of an atmospheric transport model. In Section 2.10, it was discussed that the plume scaling approach involves a calculation of integrated mass enhancement (IME) for the plume that is being quantified as well as for the plume that is used to quantify another plume. The ratio of integrated mass enhancements of both plumes is later used to scale the known emission rate of a plume to estimate the emission rate of another plume.

There are different ways of applying the plume scaling approach based on how we obtain IMEs of plumes. We can obtain IME by using all pixels inside the plume mask, or we can obtain IME by considering only the common pixels between both plumes. The approach where a common part between plumes is used to obtain IME is quite similar to the one used in a study performed to quantify gas well blowout emissions in Ohio [43]. The Ohio study considered all the TROPOMI plume pixels that were part of the synthetic plumes mask, so only the synthetic plume mask was needed. In our common pixel approach, we consider overlapping pixels between two plume masks, and we discard the rest of the pixels; hence our common pixel approach varies slightly from the Ohio study approach. In Figure 5.11, the top two images show a case where all pixels inside plume masks are considered for calculating IME, whereas the bottom two images show a case where only the common part between two plume masks is considered for calculating IME. In our study, we compare both variations of the plume scaling approach.

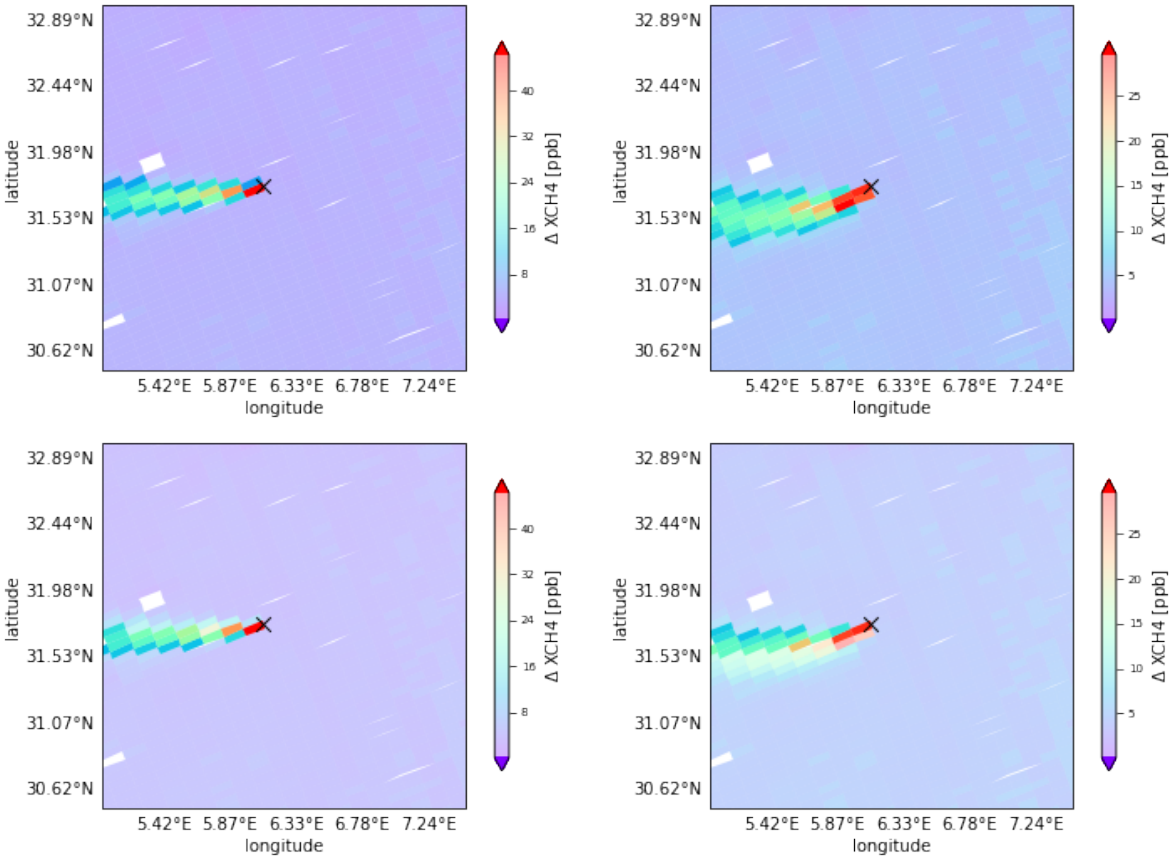


Figure 5.11: This figure illustrates two variations of the plume scaling approach. The two plumes at the top show a variation of the plume scaling approach where all plume pixels are considered for IME calculation, and the two plumes at the bottom show a variation of the plume scaling approach where only the common pixels between two plumes are considered for IME calculation. All the plumes shown here are generated with the FLEXPART model over the Algeria region. Please note that plumes used for illustrating variations of the plume scaling approach are generated with different meteorological data. The black cross indicates the source location of synthetic plumes. The colorbar shows the dry-air mole fraction of methane in [ppb].

The emission rates of our synthetic plumes are known. If we try to quantify a synthetic plume with

another synthetic plume having the same emission rate, under ideal circumstances, ratio of their IMEs should be equal to 1. The use of different settings of atmospheric transport models can influence the shape, size, and orientation of plumes. Plume images have missing pixels, and plumes with different geometry might not have the same missing pixels, and hence they might have different IME values. Due to these reasons and more, the IME ratio deviates from unity, and by studying this deviation, we investigate uncertainties in the plume scaling approach due to different settings of atmospheric transport models.

6

Pre-analysis steps

Several methods used for emission quantification of TROPOMI plumes were discussed in Chapter 5. Before analysing these quantification methods, we had to perform several pre-analysis steps. These pre-analysis steps are discussed in this chapter.

6.1. Pre-analysis steps for mass balance methods

Different TROPOMI plumes have different shapes, sizes, and orientations depending on meteorological conditions, or they might lack coverage due to the presence of clouds or aerosol effects etc. Mass balance methods are constrained by geometries (due to plume masking) and coverage of pixels over plumes, and hence mass balance methods might not perform uniformly well for different types of plumes. Classifying plumes according to their geometrical features or coverage might help us in identifying challenging cases for the mass balance methods. Therefore, before we quantify synthetic plumes using mass balance methods, we segregate plumes into different categories.

We segregate synthetic plumes into four categories. These include three categories based on plume geometries and coverage where according to our past observations, mass balance methods might not perform well and a category that contains standard or normal plume cases. The first category is blob-like plumes. Blob-like plumes do not really show a clear plume-like structure but are relatively localised enhancements. They often occur due to low wind speed conditions where pollution accumulates. In the past, we have observed several cases of mass balance methods not performing well in low wind speed conditions. The next category is curved plumes. Curved plumes are a result of large variations in the wind speed directions with respect to time or place. Some implementations of mass balance methods considered in this study use a plume box to identify plume pixels. Plumes with curved geometries might be challenging to quantify for several implementations of mass balance methods. Furthermore, overly curved plumes might have overlapping regions, and mass balance methods are not suitable to quantify plumes having overlapping regions. This is because mass balance methods assume 2-dimensional mass conservation, and the overlapping regions disturb the mass conservation. The next category is related to plumes with bad coverage. Pixels in the TROPOMI images are sometimes missing due to the presence of clouds or other factors, and since we are resampling synthetic plume pixels to TROPOMI pixels, we are missing the same pixels in synthetic plume images. We tried to choose days for generating synthetic plumes where we had good TROPOMI coverage, but a plume might coincidentally lie in regions where a lot of pixels are missing. A large number of missing plume pixels might lead to uncertainties in the mass balance quantification of plumes since these methods are based on adding up the total plume mass, which does not work if parts are missing. Plumes which do not belong to the blob-like, curved or bad coverage plume categories are classified as normal plumes in this study. Figure 6.1 shows examples of plumes belonging to each of the four categories.

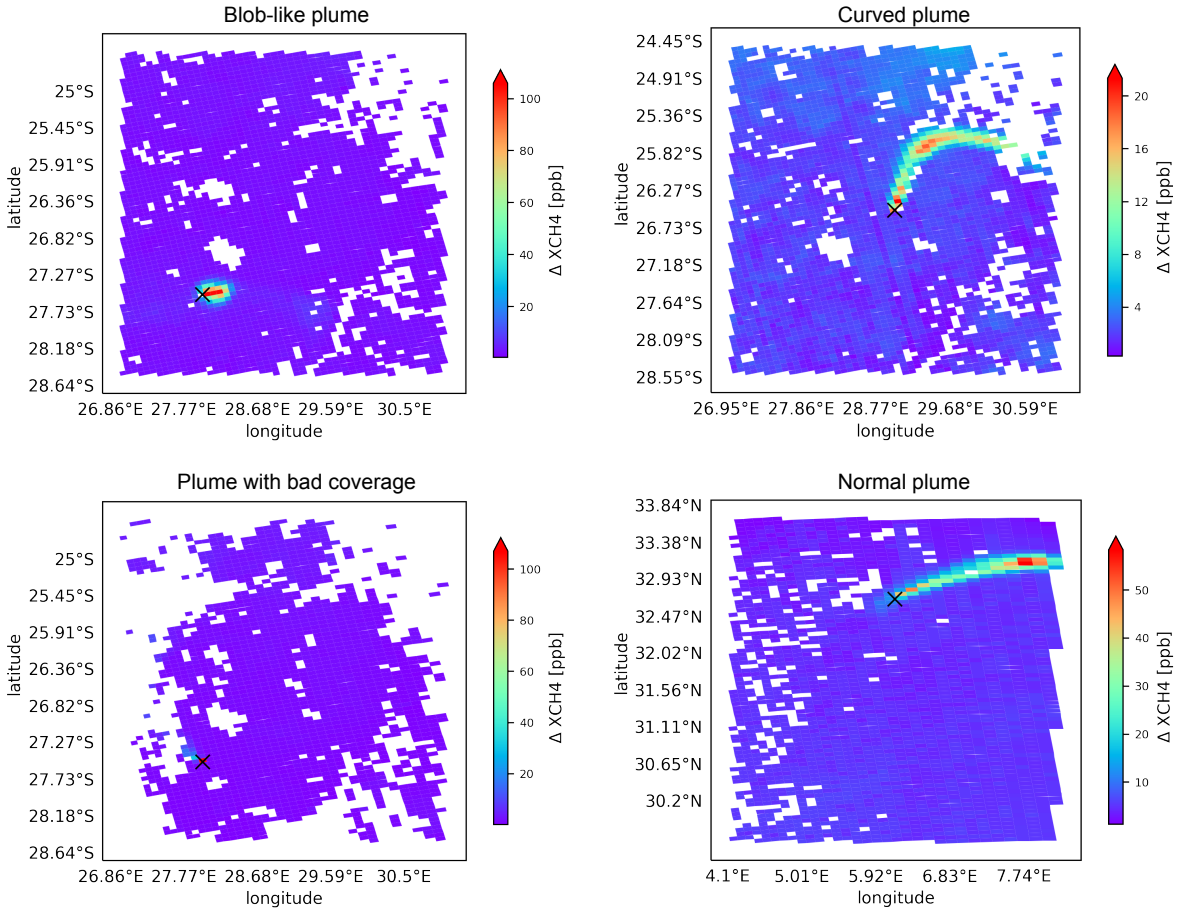


Figure 6.1: Examples of plumes belonging to different categories (from top left to bottom right): 1) blob-like plume (generated over the SASOL region on 18-05-2020), 2) curved plume (generated over the SASOL region on 18-05-2020), 3) bad coverage plume (generated over the SASOL region on 21-04-2020), 4) normal plume (generated over the Algeria region on 03-01-2020). All plumes are generated with the FLEXPART model and resampled to match TROPOMI pixel resolution. Please note that these images contain the same missing pixels as the TROPOMI measurements. The black cross marker in the images indicates plume source location.

Manually segregating a large number of synthetic plumes into different categories is a time-consuming process, and to solve this problem, we developed a plume classification algorithm. This algorithm will also be useful later for the analysis of atmospheric transport model based plume emission quantification approaches. This classification algorithm is based on a decision tree classifier that uses several metrics based on geometrical features and coverage information of plumes to identify which category a particular plume belongs to. To extract geometrical features from plumes, we use two metrics: distance between extreme ends of a plume and overall length of a plume. A schematic depicting the geometrical features of a plume is shown in Figure 6.2.

To extract the geometrical features from plumes, we developed an algorithm called the plume tracer algorithm. For a given image with a plume, the plume tracer algorithm highlights plume pixels using a threshold of mean plus 1.8 times the standard deviation of all pixels. The plume tracer algorithm then detects extreme points of the identified plume and calculates the distance between those points in degrees. Next, the plume tracer algorithm fits a spline through highly enhanced pixels between the extreme points of a plume. The total length of this spline gives the total length of a plume in degrees. An example of a plume traced by the plume tracer algorithm is given in Figure 6.2. Extracted geometrical features by the plume tracer algorithm are used in the plume classification algorithm. We noticed that the plume tracer algorithm performs well for complex plume geometries and even if some plume pixels are missing. Plumes with overly curved geometries might be challenging. Some parameters for spline fitting in the plume tracer algorithm are decided based on a manual investigation; fine-tuning these

parameters might improve the plume tracer algorithm.

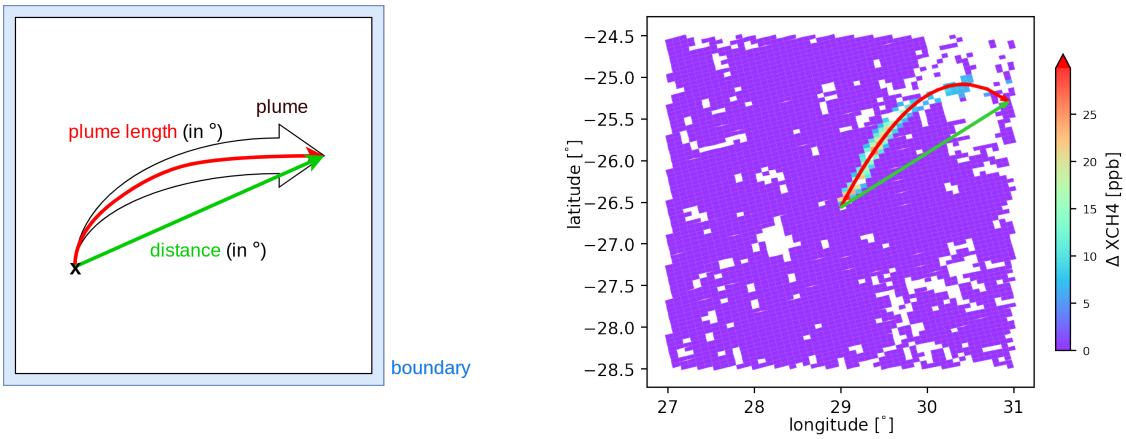


Figure 6.2: On the left side of this figure we show a schematic image of a plume in form of an arrow inside a region defined by blue boundary. In this image, green line depicts distance between extreme ends of plume in degrees. Red arrow depicts overall length of plume in degrees. Distance between extreme ends of plume, and overall length of plume are some of the metrics used in the classification algorithm. On the right side of this figure we show an example of plume traced by the plume tracer algorithm. This algorithm gives us distance between extreme ends of a plume (marked by green arrow by the algorithm) and total length of a plume (marked by red arrow by the algorithm) in degrees.

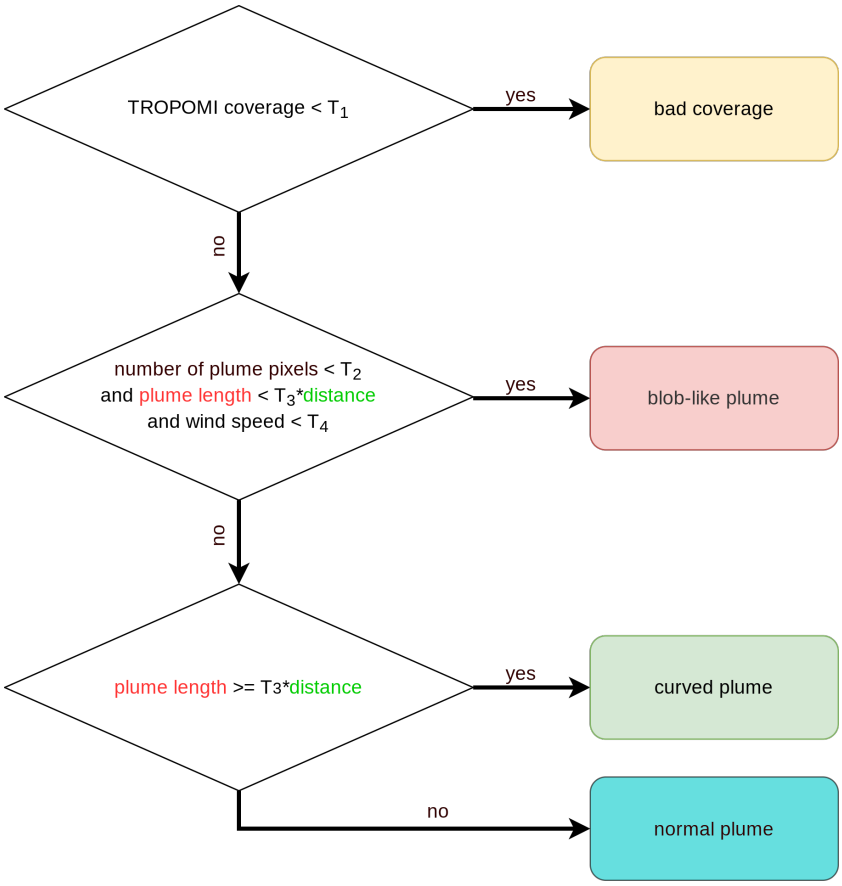


Figure 6.3: This figure shows a decision tree classifier used in this study to classify synthetic plumes. Plumes are classified into four categories: plumes with bad coverage, blob-like plumes, curved plumes, and normal plumes. This decision tree uses several threshold parameters T_1, T_2, T_3, and T_4 in the process of segregating the plumes. The parameter T_1 is the threshold percentage of coverage below which a plume is classified as a bad coverage plume. The parameters T_2, T_3, and T_4 are the threshold number of plume pixels, threshold ratio between total length and distance, and threshold wind speed for identifying blob-like plumes. Parameter T_3 is again used for identifying curved plumes.

The flowchart of the plume classification algorithm is shown in Figure 6.3. The plume classification algorithm makes use of several metrics to segregate plumes. These metrics include coverage or availability of pixels near plume region, number of pixels that belong to a plume, a ratio of the distance between extreme ends of plume and overall length of a plume, and wind speed at plume source location. The following paragraphs discuss how each metric is used in identifying the class of plumes.

Identification of plumes with bad coverage is made by calculating the percentage of available pixels in the quadrant of an image where the maximum portion of the plume lies. If the percentage of available pixels is lower than a certain threshold (T_1), then that plume image is classified under the bad coverage category.

Blob-like plumes are identified by checking three conditions. The first condition checks if the number of plume pixels is smaller than a certain threshold number (T_2) since blob-like plumes are short and usually have less number of pixels. The second condition checks if the ratio between the overall length of the plume (red line in Figure 6.2) and the distance between extreme points of the plume (green line in Figure 6.2) is smaller than a certain threshold (T_3) since, for very short plumes, length of plume and distance between extreme points of plume is nearly equal. This condition was added to avoid classifying curved plumes as blob-like plumes. The third condition checks if the wind speed is below a certain value (T_4) since blob-like plumes are often formed due to low wind speed conditions.

Curved plumes are separated from the remaining plumes by checking if the ratio between length of plume and the distance between extreme points of plume is bigger than a certain threshold. Finally, plumes that do not satisfy any of the previously mentioned conditions are classified as normal plumes.

The previous paragraphs mentioned that we use some threshold parameters in the plume classification algorithm. We obtain values of those threshold parameters using an iterative optimisation process. This iterative process uses an optimisation algorithm that tries to obtain the best set of threshold parameters in order to maximise classification accuracy. We use 54 synthetic plumes generated over the Algeria region using ECMWF meteorological data for optimising the threshold parameters. We manually label each plume in the training data and use those labels to compute the classification accuracy of the algorithm. We calculate classification accuracy using three metrics: true positives (TP), false positives (FP), and false negatives (FN) (Box 6.1). For maximum classification accuracy, we try to maximise the mean of true positives for all categories and minimise the mean of false positives and false negatives for all categories. The optimisation process stops once a maximum classification accuracy is reached. The optimisation algorithm could achieve 85.33% classification accuracy.

Box 6.1: Metrics used for calculating classification accuracy

In this study, we use true positives (TP), false positives (FP), and false negatives (FN) to calculate the classification accuracy of the plume classification algorithm. True positive is the case when the classification algorithm correctly predicts positive class of plume. An example of a true positive result is: algorithm predicts a manually labelled normal plume as a normal plume. False positive is the case when the classification algorithm incorrectly predicts positive class of plume. An example of a false positive result is: algorithm predicts a blob-like plume as a normal plume. False negative is the case when the classification algorithm incorrectly predicts negative class of plume. An example of a false negative is: algorithm predicts a normal plume as a blob-like plume. We also use true positive rate to analyse the classification results. The true positive rate is the ratio of true positive and the addition of true positive and false negative. Having a good true positive rate indicates the accuracy of the classification algorithm in correctly identifying the class or category of plumes.

The threshold parameters obtained after completing the optimisation process are listed in Table 6.1.

Table 6.1: This table presents the threshold parameters obtained after optimization process is complete. Threshold parameter T_1 is used for identifying bad coverage plumes. Threshold parameters T_2, T_3, and T_4 are used for identifying the blob-like plumes. Threshold parameter T_3 is again used for identifying curved plumes.

Parameter	Value	Description
T_1	< 70%	Threshold percentage coverage
T_2	< 8	Threshold number of pixels
T_3	> 1.12	Threshold ratio between total length and distance
T_4	< 1.95 [m/s]	Threshold wind speed

We test the threshold parameters mentioned in Table 6.1 on several synthetic plumes. Some representative cases of manually classified plumes were presented in Figure 6.1. Results of classification algorithm for the cases mentioned in Figure 6.1 are presented in Figure 6.4. For these cases, the classification algorithm works well, but there are also some cases where this algorithm does not perform well. We summarise the results of the classification algorithm applied on plumes generated over a different region or with different meteorological data in Table 6.2.

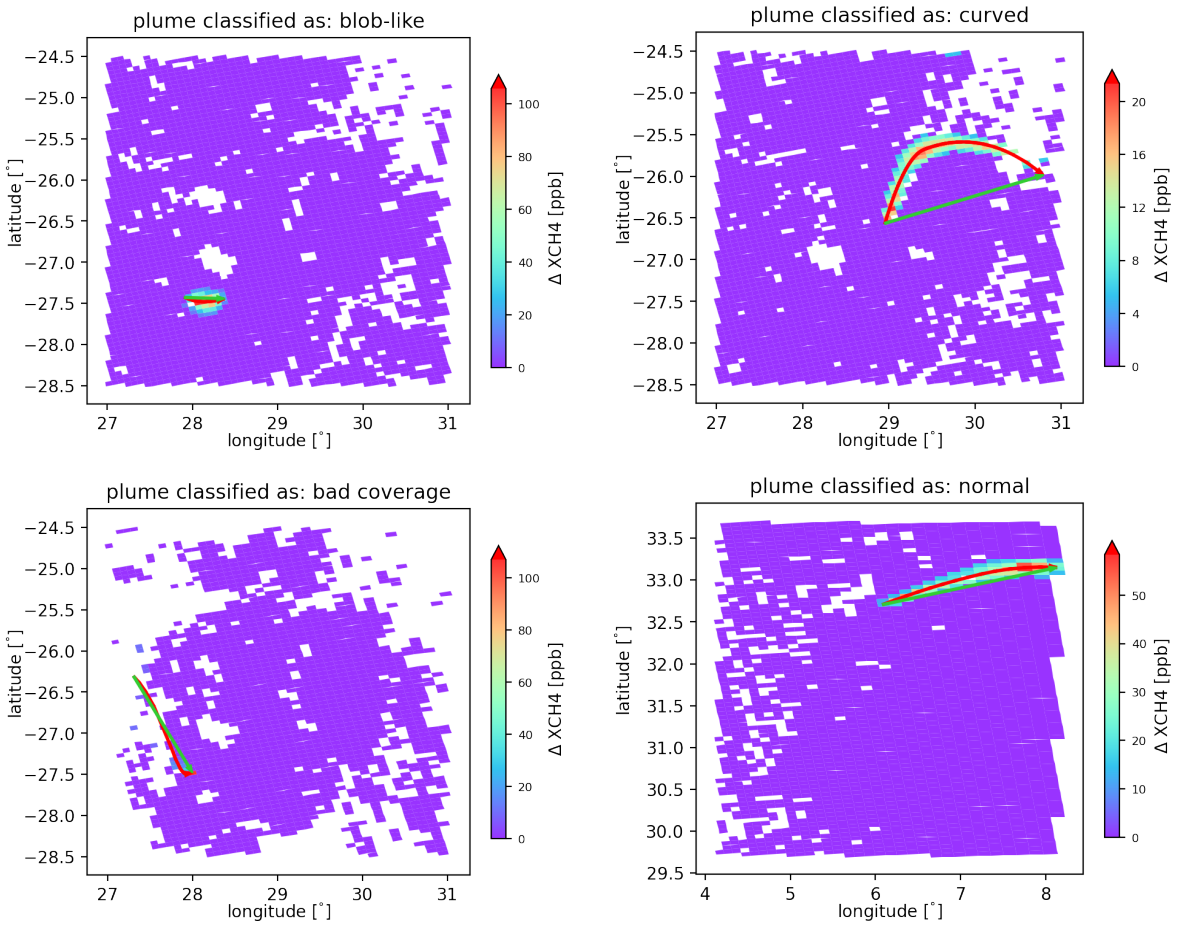


Figure 6.4: Results from plume classification algorithm for examples mentioned in Figure 6.1 (from top left to bottom right): 1) blob-like plumes, 2) curved plumes, 3) bad coverage plumes, 4) normal plumes. All the plumes are generated with the FLEXPART model and resampled to match the TROPOMI pixel resolution. The red and green arrows in the images are used for extracting plume geometrical features.

Table 6.2: Classification results for various sets of plumes based on the threshold parameters obtained using 75% Algeria-ECMWF synthetic plumes. In this table, results are shown for the following categories of plumes: blob-like plumes, curved plumes, plumes with bad coverage, and normal plumes. Three metrics are presented for each category of plume: true positives (TP), false positives (FP), and false negatives (FN).

Plume specifications	Blob-like	Curved	Bad coverage	Normal
Location - meteo data	TP | FP | FN	TP | FP | FN	TP | FP | FN	TP | FP | FN
Algeria - ECMWF (remaining 25%)	1 | 0 | 0	2 | 1 | 0	0 | 0 | 0	14 | 0 | 1
Algeria - NCEP	1 | 0 | 2	11 | 1 | 5	9 | 0 | 0	43 | 7 | 1
Sasol - ECMWF	3 | 1 | 1	16 | 7 | 4	9 | 0 | 6	32 | 5 | 1
Sasol - NCEP	0 | 0 | 1	10 | 10 | 6	9 | 0 | 4	35 | 7 | 7
Total	5 | 1 | 4	39 | 19 | 15	27 | 0 | 10	124 | 19 | 10

Normal plumes account for almost 57% of the tested synthetic plumes, and the classification algorithm could identify them with a 93% true positive rate. The number of false positive cases is also crucial to assess the performance of the classification algorithm since we do not want to include falsely labelled plumes in any category. For normal plumes, the percentage of false positive cases was nearly 15%. The classification algorithm might have some challenges in identifying blob-like plumes, and we get a low true positive rate, but the results for blob-like plumes had very few false positives. Curved plumes and plumes with bad coverage were classified with nearly 75% true positive rates. For the curved plumes category, we see a high number of false positives. For plumes with a bad coverage category, the classification algorithm identified no false positives.

The classification algorithm can be improved in the future with the help of more advanced metrics, especially to increase the true positive rate for classifying blob-like plumes and to reduce false positives in the curved plumes category. After the classification of synthetic plumes is done, we quantify emissions of plumes in each category using various implementations of mass balance methods described in Section 5.1. The emission quantification results are analysed in Chapter 7.

Some pre-analysis steps are required before assessing the atmospheric transport model based plume scaling approach. These pre-analysis steps are discussed in the following section.

6.2. Pre-analysis steps for atmospheric transport model based plume emission quantification

As mentioned in Section 5.2.1, the plume scaling approach uses the ratio of integrated mass enhancements (IMEs) to scale the known emission rate of a synthetic plume to estimate the emission rate of another plume. Using only the IME ratio as a plume scaling factor is limiting since it does not account for the length or diffusion of plumes. Two plumes might not have the same atmospheric transport times. Plumes that are emitted and transported for longer times have higher IME values, lengths, and diffusion than the plumes that are emitted and transported for shorter times. In this study, we account for different atmospheric transport times of plumes by incorporating parametric plume length (L) in the calculation of the scaling factor. The new scaling factor is $(IME_1/L_1)/(IME_2/L_2)$, where 1 and 2 refer to the plume to be quantified and the plume used for quantifying another plume, respectively. The parametric plume length can be calculated by taking the square root of the total area occupied by the plume pixels. This definition of the parametric plume length was originally used with the IME mass balance method to account for the extent of the plume.

In this analysis, we generate synthetic plumes with the same emission rates. As mentioned previously, under ideal circumstances, the IME/L ratio for these two synthetic plumes is unity. In practice, missing pixels and different geometries of plumes affect the IME/L values, and hence the IME/L ratio deviates from unity. Investigation of this deviation in the IME/L ratio can tell us about uncertainties related to the plume scaling approach. In Figure 6.5, several cases with IME/L ratios close to unity are highlighted. In the first two cases, both the plume have different geometries but have very few missing pixels, and for these cases, the plume scaling approach works well.

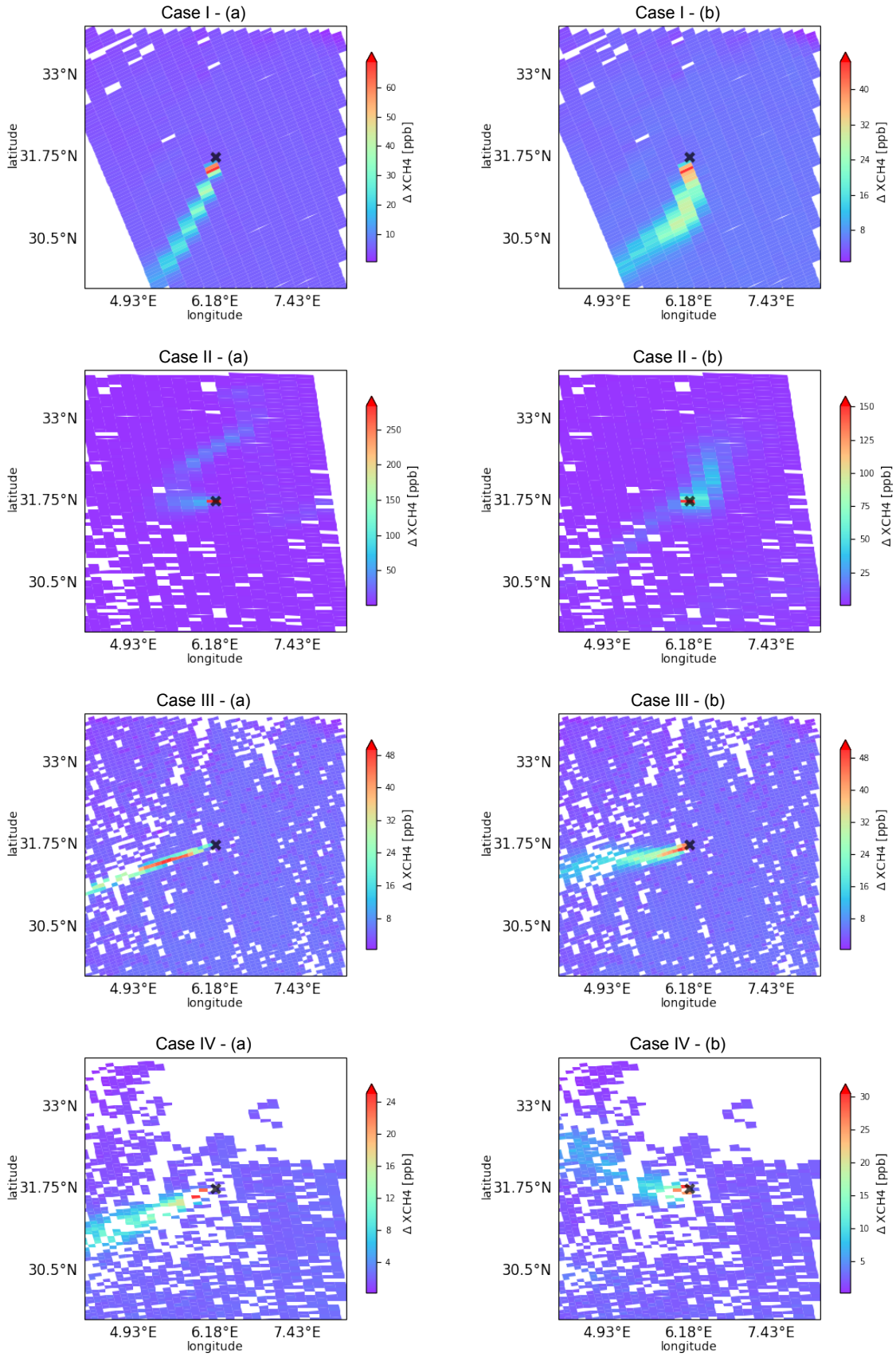


Figure 6.5: This figure shows several representative cases where the plume scaling approach works well. The plume scaling approach is said to be working well if the ratio of IME/L of two plumes is close to 1. The IME/L ratio for cases I, II, III, and IV, are 0.97, 1.06, 1.08, and 1.10, respectively. For a particular case, the left side image (having (a) in the title) contains TROPOMI resampled plume generated with FLEXPART using ECMWF meteorological data, and the right side image ((having (b) in the title)) contains TROPOMI resampled plumes generated with FLEXPART using NCEP meteorological data. Black cross markers indicate plume source locations.

In the third case of Figure 6.5, both the plumes have similar geometry, similar orientations and similar missing pixels, and here the plume scaling approach works well. Furthermore, there can be some cases where the plumes might have different orientations and different missing pixels, but for those cases, the plume scaling approach might coincidentally work well, as shown in the fourth case.

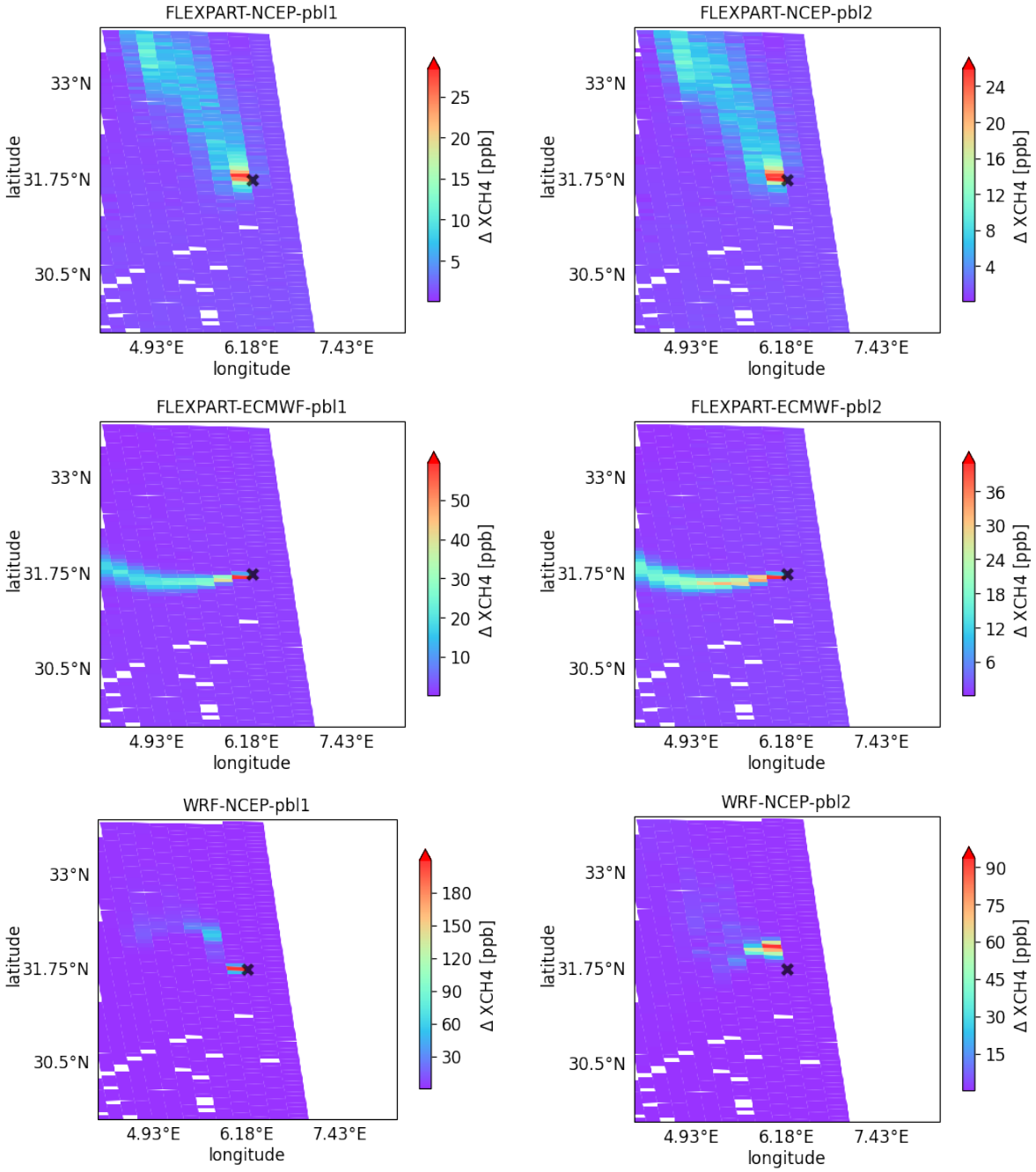


Figure 6.6: This figure shows resampled synthetic plumes generated using FLEXPART and WRF models for 29-08-2020 using different meteorological datasets (NCEP/ECMWF) with different planetary boundary layer schemes (pbl1/pbl2). The planetary boundary layer schemes used with the FLEXPART model are different than the planetary boundary layer schemes used with the WRF model.

The accuracy of the plume scaling approach relies on the similarity between two plumes, especially if there are several missing plume pixels. If both plumes show similar diffusion, orientations, and shapes such that both the plumes are missing almost the same pixels, we can still estimate the emission rate of a plume with good accuracy using the scaling approach. This is an advantage we get by using the plume scaling approach over mass balance methods.

In practice, if we want to quantify an actual TROPOMI plume using a synthetic plume with the plume scaling approach with reasonable accuracy, we need both the plumes to be similar to each other. Getting a synthetic plume that exactly matches the TROPOMI plume is challenging due to several uncertainties in meteorological conditions or atmospheric transport model physics. This can also be observed when we simulate two synthetic plumes using different meteorological data, different atmospheric models, or with a different physics setting (planetary boundary layer scheme) (Figure 6.6 and Figure 6.7). Sometimes using more options of planetary boundary layer schemes helps in getting a plume that matches well to the plume to be quantified.

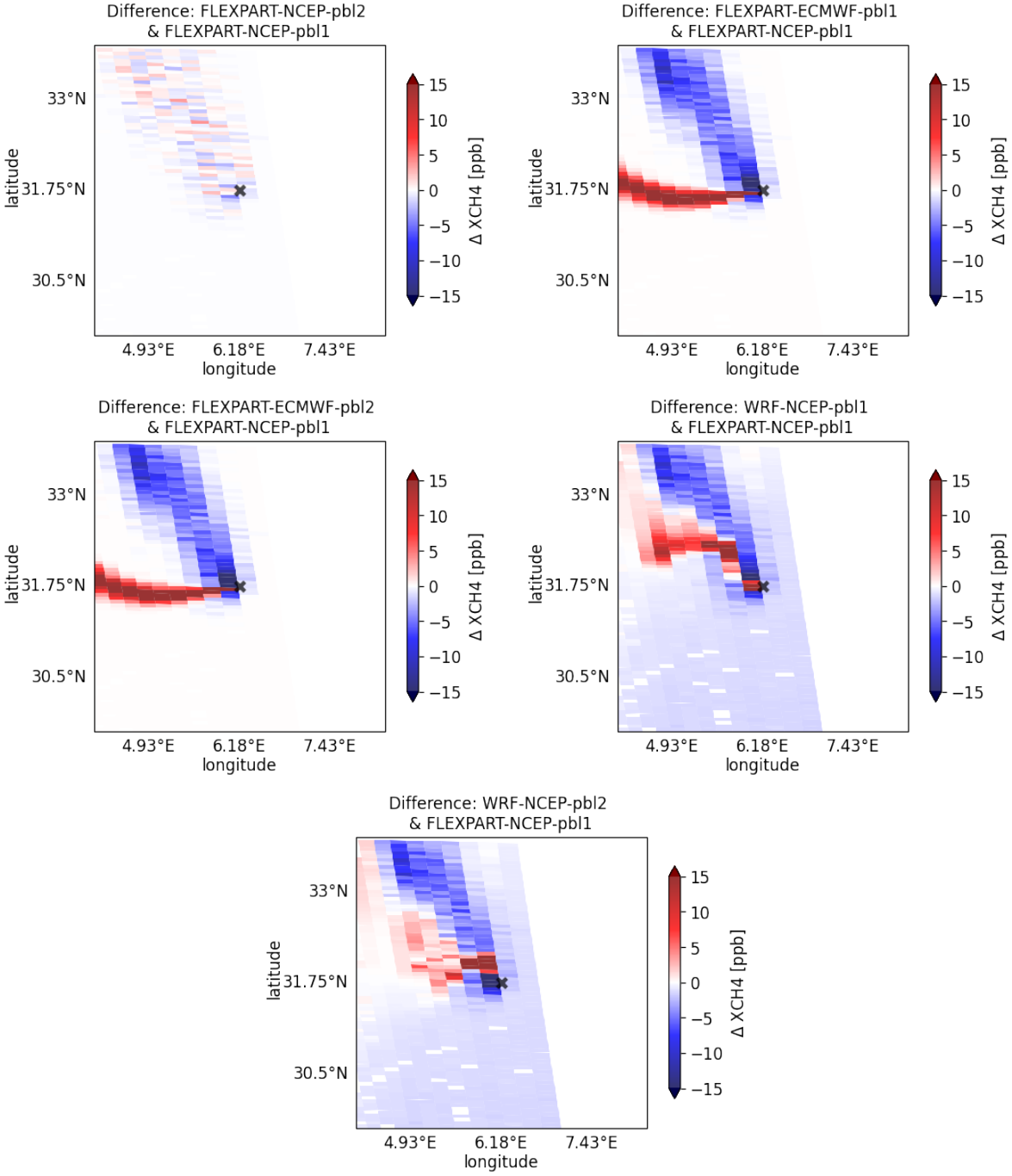


Figure 6.7: This figure shows the difference between two synthetic plumes having same emission rates generated over the same location and on the same day. One of the two synthetic plumes is generated using the FLEXPART-NCEP-pbl1 setup and the other synthetic plume belongs to one of the other settings mentioned in Figure 6.6.

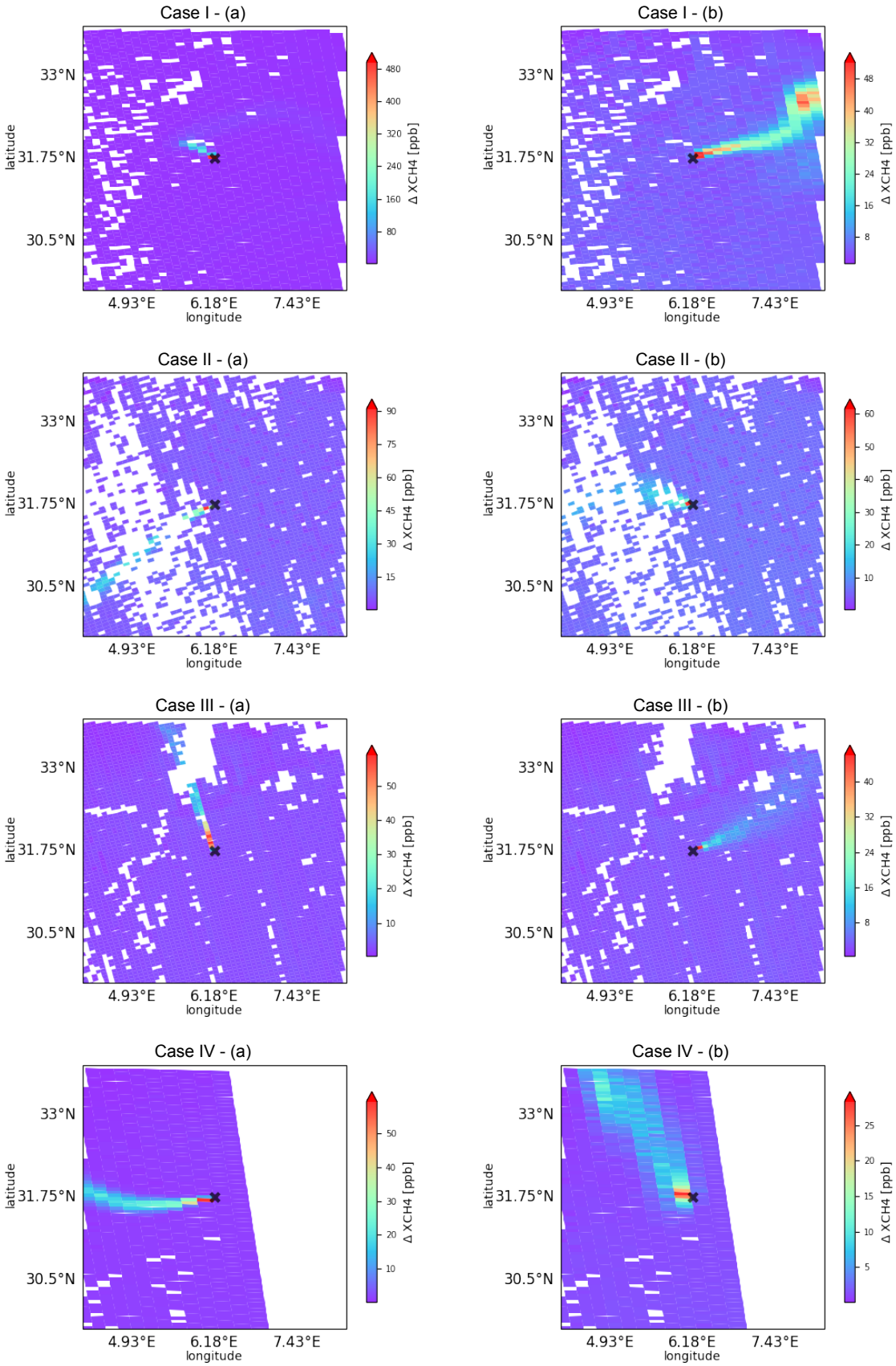


Figure 6.8: This figure shows four representative challenging cases for the plume scaling approach. For a particular case, the left side image (having (a) in the title) contains TROPOMI resampled plume generated with FLEXPART using ECMWF meteorological data, and the right side image ((having (b) in the title)) contains TROPOMI resampled plumes generated with FLEXPART using NCEP meteorological data. Case I contains plumes with very different lengths, and the IME/L ratio is 2.09. Case II contains plumes with a lot of missing pixels, and the IME/L ratio for this case is 1.73. Case III contains plumes having very different orientations, and the left plume falls in the bad coverage region. IME/L ratio for this case is 1.97. Case IV contains plumes with very different diffusions, which makes it difficult to find an optimum plume mask. IME/L ratio for this case is 1.57.

Since the application of the plume scaling approach to cases where plumes are very different from each other has a high tendency to produce poor results, we need to set some constraints to allow only the cases where plumes are not very different from each other. Hence, we perform some pre-analysis steps to identify and filter challenging cases for the plume scaling approach. For this purpose, we first manually investigate several cases where the ratio of IME/L largely deviates from unity. Some manually identified challenging scenarios are shown in Figure 6.8.

The first challenging scenario is when the length or diffusion/ width of plumes are very different. Missing pixels over short plumes have a larger impact on the IME/L of a plume than missing pixels over a large plume. Obtaining an optimum plume mask for plumes with large diffusion is challenging. The second challenging scenario is when the orientations of plumes are very different. Because different orientations can cause plumes to have different missing pixels, which changes the IME/L of plumes, and hence the ratio deviates from unity. The third challenging scenario is when a lot of plume pixels are missing.

To identify and filter the previously mentioned challenging scenarios from a large number of plumes in an automated way, we developed a filtering algorithm. This algorithm uses several metrics to identify plumes with a significant difference in sizes and orientations as well as a large number of missing pixels over plumes. To identify cases where plume length and diffusion are very different, we take a ratio of parametric lengths of plumes. Parametric plume length can be calculated by taking the square root of the area occupied by all pixels inside the plume mask. Parametric plume lengths can partially account for the length of plume and diffusion of plume. If the ratio is above or below a particular threshold, we label that case as plumes with a large difference in sizes. To determine cases where the orientations of plumes are very different, we calculate the distance between plume centres. The plume centre is a middle point of a plume tracer passing through a plume (plume tracer algorithm is described in Section 6.1). If the distance between two plume centres is greater than a particular threshold, we label that case as plumes with a large difference in orientation. Finally, we label plumes with a certain percentage of missing pixels as 'plumes with bad coverage'. Please note that here we calculate the percentage of missing pixels over a plume, whereas, in the pre-analysis of mass balance methods, we calculated the percentage of missing pixels in a quadrant of an image where the plume lies. Calculating missing pixels over a plume is more accurate than calculating missing pixels in a region of an image, but we could not use this new way of calculating missing pixels in the pre-analysis steps of mass balance methods due to time constraints. For further analysis, we calculate missing pixels over plume using the newly developed method. While calculating the percentage of missing pixels, we give more weight to the pixels near the plume source since synthetic plumes have pixels with high concentrations near the plume source and missing those pixels can largely impact the calculated integrated mass. With TROPOMI plume images, the pixels with the highest enhancements might not lie near the plume source; in those cases, the plume filtering algorithm can be modified accordingly. The threshold parameters used for identifying challenging cases are mentioned in Table 6.3.

The plume filtering algorithm is not very stringent, and it allows for variation in the geometry of plumes up to a certain extent. There is a trade-off between the accuracy of the plume scaling approach and the number of plumes that can be considered to quantify with the plume scaling approach.

Table 6.3: Threshold parameters used for identifying challenging cases for the atmospheric transport model based plume scaling approach. These parameters were fixed from manual investigation of several challenging cases for plume scaling approach.

Parameter description	Value
Threshold for size ratio between two plumes	$> 0.7, < 1.5$
Threshold for distance between plume centres	$< 0.9^{\circ}$
Threshold for coverage over plume	$> 30\%$

We filter out the challenging cases using an algorithm that uses threshold parameters mentioned in Table 6.3. The filtering removes almost 50% of the synthetic plumes. The remaining plumes are quantified using the plume scaling approach, and the results are discussed in the next chapter.

7

Results and discussion

We used atmospheric transport models to generate a large number of synthetic plumes over different regions and different dates. In this chapter, we will estimate the emissions of these synthetic plumes A) using mass balance methods and identify their limitations, and B) with atmospheric transport model based plume scaling approach and see if it improves the plume emission quantification. We also investigate how different settings of atmospheric transport models affect synthetic plumes and later find which of these settings are most suitable to replicate TROPOMI plumes to get emissions. Finally, we discuss how atmospheric transport models can be used to improve TROPOMI plume emission quantification, and we give an idea about the integration of atmospheric transport based plume emission quantification method with the TROPOMI automated plume detection algorithm.

7.1. Assessment of mass balance methods

Mass balance methods need several input settings to quantify plumes (Section 5.7). These input settings include several parametric coefficients that are specific to a particular satellite instrument. The settings we use in this study belong to the TROPOMI satellite instrument. Some fine-tuning of these input settings can help in achieving better plume emission quantification results. Hence, we first try to identify if there is any scope for improving the settings of mass balance methods. For this purpose, we examine bias in the results. The presence of highly positive or negative bias can give an indication that the settings of mass balance methods can be changed to improve the plume emission quantification results. Bias is obtained by calculating the difference between mass balance calculated emission rates and actual emission rates. We plot the bias for six implementations of mass balance methods (4 for IME and 2 for CSF), each applied to the four categories of plumes (blob-like, curved, bad coverage, and normal plumes). For all the plumes in a particular category quantified with a particular variation of the mass balance method, bias is presented in the form of a boxplot in Figure 7.1. Please note that we use pre-calibrated settings of mass balance methods in Figure 7.1.

Figure 7.1 shows the biases for all plumes generated with ECMWF meteorological data over the Sasol region. This result is representative for the biases we find for all generated plumes (Appendix A.4). Normal plumes are the best-case scenario to quantify for mass balance methods, and they comprise almost 57% of the total plumes. Hence, we initially look at the normal plumes category in order to investigate the bias due to U_{eff} calibration.

To correct for any small amount of bias and to double-check if the settings of mass balance methods are appropriate, we perform recalibration of parametric coefficients used in the IME and CSF effective wind speed equations (Section 2.6). The recalibration process involves plotting a regression curve between 10 [m] wind speeds and the effective wind speed needed for achieving actual emission rates of plumes. To calculate the effective wind speed needed to achieve the actual emission rate, we perform mass balance plume emission quantification of synthetic plumes with a unit effective wind speed to obtain the emission rate per unit effective wind speed. Next, we take the ratio between the actual emission rate of the synthetic plume and the calculated emission rate per unit effective wind speed. This ratio

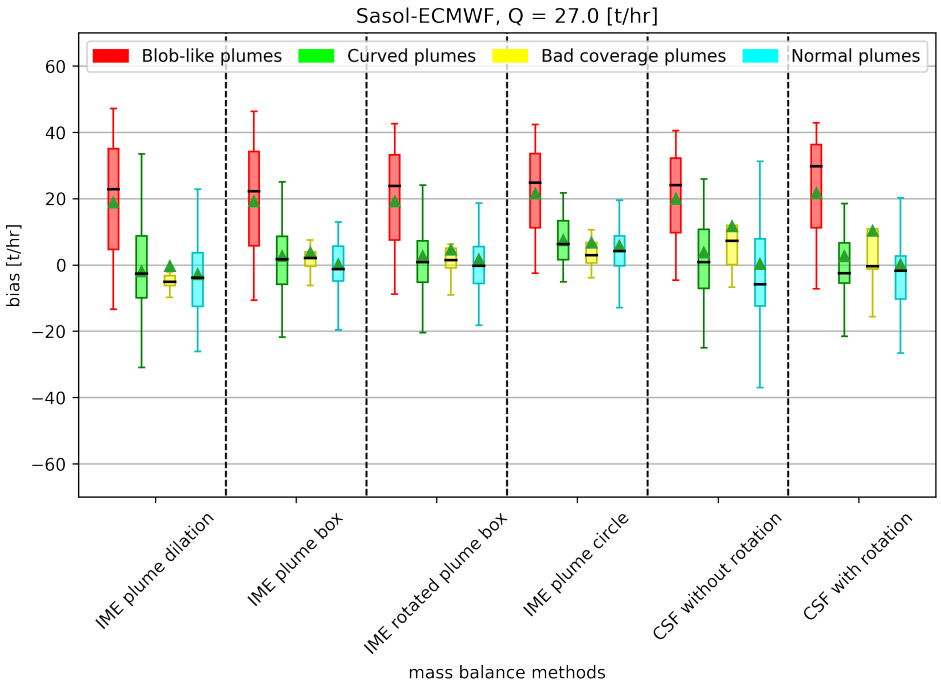


Figure 7.1: Box plots containing the bias corresponding to mass balance quantification of plumes with known emission rate of 27 [t/hr] generated over Sasol region using the ECMWF meteorological data. Each box plot represents biases calculated for plumes belonging to a particular class of plumes. We are quantifying plumes using 4 implementations of IME method and 2 implementations of CSF method. For each box plot, median of data is represented by a black bold line, mean of data is represented by a green triangle, box contains data between 25^{th} percentile and 75^{th} percentile which is also called interquartile range, and whiskers are placed at 1.5 times interquartile range from the ends of box.

gives us the effective wind speed required to achieve the actual emission rates of plumes.

For the recalibration process, we use a linear relationship between effective wind speed and 10 [m] wind velocity at the source location of plumes. Coefficients of this linear relation are obtained by fitting a line through several data points corresponding to synthetic plumes belonging to the normal plume category. We noticed several outliers while performing recalibration. To get rid of these outliers, we use a median filter. This median filter uses median difference, which is the median of the difference between data value and the median of data. We remove data points that are above or below 2.2 times the median difference from the median of data. The linear regression fits for four implementations of the IME mass balance method, and two implementations of the CSF mass balance method are provided in Figure 7.2.

The previously used parametric coefficients for IME parametric wind speed equation were [0.444,0.278], and for CSF parametric wind speed equation were [1,0]. After the recalibration process, new parametric coefficients for IME and CSF methods are [0.457, 0.467] and [1.144, 0.146], respectively. These coefficients are taken from the plots that show the least RMSE in the regression fit. The newly obtained parametric coefficients seem close but higher than the ones used in the model before the recalibration process. We suspect these higher values might be due to the difference in the diffusion of synthetic plumes used in the previous and this calibration processes or results are biased due to median filtering. Since the newly obtained coefficients are close to the previously used coefficients, and our synthetic plume dataset to obtain new parametric coefficients was limited to make any conclusive remarks on parametric coefficients, we chose to use the previously used parametric coefficients for the further analysis of mass balance methods.

In the next step, we investigate uncertainties in mass balance method based plume emission quantification. For this purpose, we plot root mean square errors (RMSE) of plume emission quantification results of mass balance methods. Similar to the bias plots, we present RMSE results as box plots belonging to a particular variation of the mass balance method and to a particular category of plume for

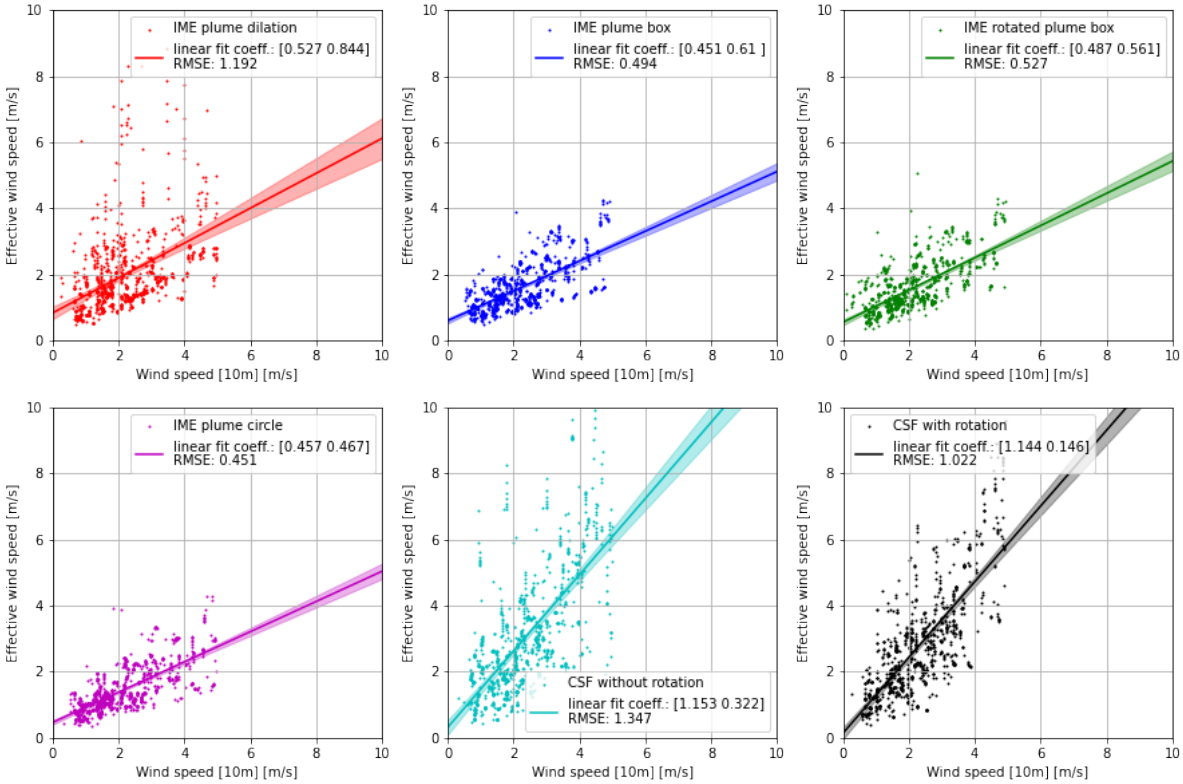


Figure 7.2: Recalibration results for four implementations of the IME mass balance method and two implementations of the CSF mass balance method. Recalibration is performed by plotting a linear regression curve between the effective wind speed needed to achieve the actual emission rate and wind speeds at [10m]. The recalibration plots consider synthetic plumes generated over Algeria and Sasol using ECMWF as well as NCEP meteorological data. The transparent colour regions in the plots represent a 99% confidence interval. Each dot in a plot represents a synthetic plume belonging to the normal plume category. Linear fit coefficients and root mean square error values are provided inside the legend boxes for each plot.

all the plumes generated over the Sasol region using ECMWF meteorological data as a representative case (Figure 7.3). A similar trend can be seen in the RMSE plots corresponding to plumes generated over different regions or with different meteorological data (appendix A.4.2).

From the RMSE plot, all implementations of mass balance methods show systematically high uncertainty in quantifying blob-like plumes. We initially expected mass balance methods to show high uncertainty towards curved plumes, but the performance for curved plumes looks similar to normal plumes. For bad coverage plumes, there is a consistently large gap between the median line (black line in the box plots) and mean triangle (green triangle in the box plots). This indicates that quantification of the plume with bad coverage creates more outliers than plumes from any other category.

Next, we examine how uncertainties in the plume emission quantification estimates of mass balance methods vary with emission rates of plumes. It was mentioned that we scale emission rates of synthetic plumes by multiplying atmospheric transport model output concentrations with some scalar value. Using this process, we obtain synthetic plumes with emission rates varying from 1 [t/hr] to 40 [t/hr]. To give more insights into these emission rates, methane released by all the oil and gas facilities in the UK is around 23 [t/hr] [67]. So, a 40 [t/hr] emission rate plume is very large. We add 5 [ppb] Gaussian noise to these synthetic plumes posterior to scaling their emission rates. Figure 7.4 shows relative uncertainty versus actual emission rates of synthetic plumes. Relative uncertainty is obtained by dividing the median value of RMSE by the actual emission rate and converting it to a percentage. Figure 7.4 shows that CSF method has a systematically higher uncertainty compared to the IME method. Furthermore, with decreasing emission rates of the plumes, the relative uncertainty increases, since noise present in pixels starts to dominate. Among all implementations of the IME method, the IME plume dilation method produced the most uncertain plume emission quantification results, and uncertainty re-

sults from other IME implementations were close to each other except at low emission rates, where the IME plume circle method performed poorly. There is not any noticeable difference between different implementations of the CSF method when we compare the uncertainties.

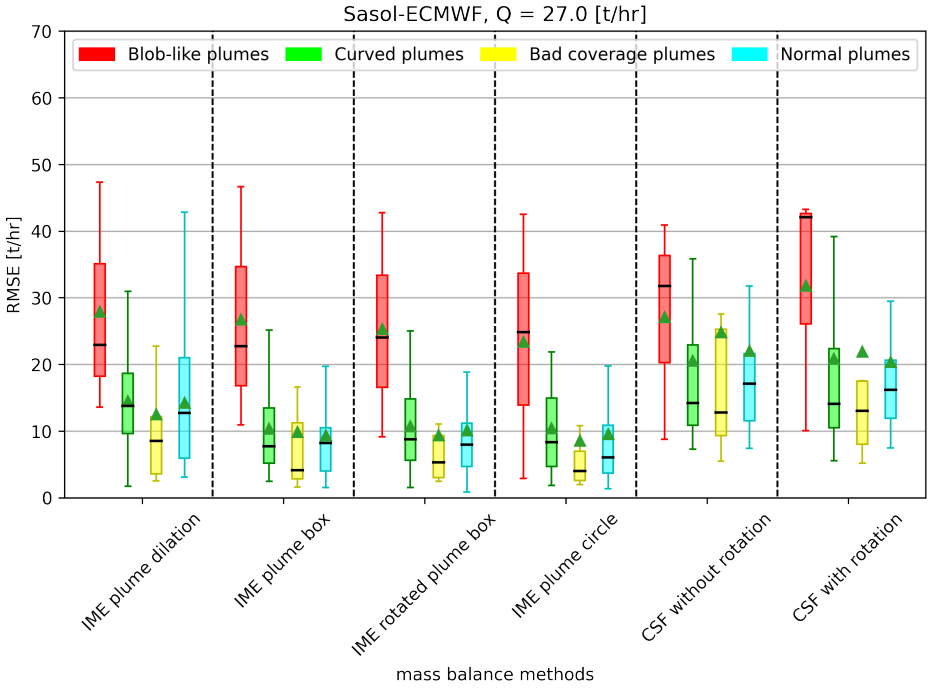


Figure 7.3: This figure shows several box plots for RMSE corresponding to mass balance quantification of plumes with known emission rates of 27 [t/hr] generated over Sasol region using the ECMWF meteorological data. Rest of the description of this plot is similar to the bias plot shown in Figure 7.1.

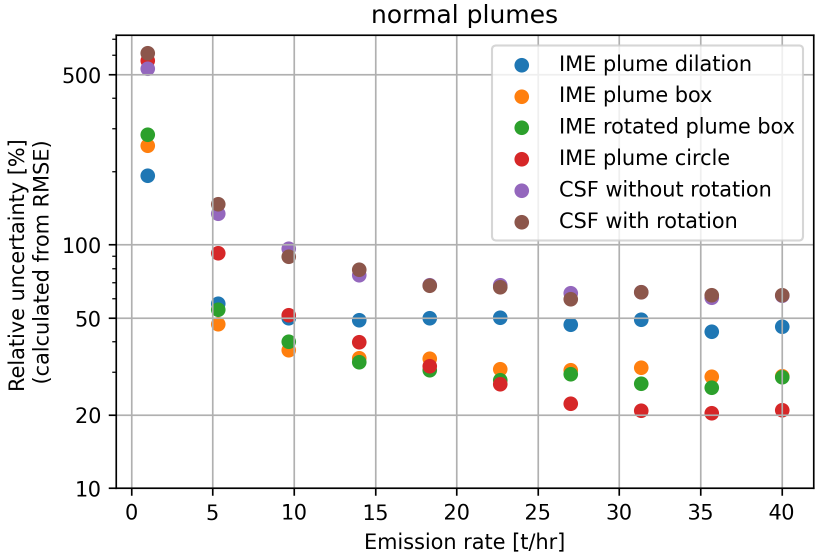


Figure 7.4: This figure shows variation in the percentage of RMSE (RMSE divided by actual emission rate times 100) with respect to the actual emission rate for different classes of plumes. This result is obtained by considering plumes belonging to the normal plume category generated over the Sasol region using the ECMWF meteorological data. Please note that for better representation of results, the y-axis is made logarithmic.

We try to make synthetic plumes as close as possible to TROPOMI plumes, but there are still some

gaps like noise in TROPOMI images. Noise in TROPOMI image affects the masking of plume pixels which affects plume emission quantification. The results we present in this section do not account for these gaps and hence are not sufficient to give concrete remarks on TROPOMI plume emission quantification using mass balance methods. These results are obtained with the main focus of creating benchmarks for the later part of this study. From these benchmarks, we identified several limitations of mass balance methods. These limitations are divided into limitations that can be solved and limitations that cannot be solved. The following paragraph discusses limitations of mass balance methods that can be solved.

Mass balance methods mentioned in this study use several satellite-specific parameters. These satellite-specific parameters include parametric coefficients used in the effective wind speed equation and plume mask filtering criteria. There might be some systematic bias in mass balance quantification due to the use of parameters that were tuned for using a different setup for a different study. This bias can partially be fixed with the recalibration of parametric coefficients in effective wind speed equations. Furthermore, mass balance methods considered in this study use uniform effective wind speed over the plume. Some advanced methods can be developed in the future that accounts for time history or vertical variation of wind velocities, but the development of these advanced methods remains out of scope for this study. Another limiting factor for the mass balance method is the plume mask setting for the IME method. Applying plume masks to synthetic plumes having a low emission rate with 5 ppb Gaussian noise is challenging, and for TROPOMI plumes, it is even more difficult.

The previous paragraph discussed some limitations of mass balance methods that can be solved by recalibration process or by developing some advanced IME and CSF methods. Mass balance methods have several inherent limitations which can not be solved. Inherent limitations of mass balance methods include missing plume pixels, performance under low wind speed conditions, and a too simplified way of accounting for the transport of particles by using wind at one location at one moment in time. TROPOMI images have missing pixels due to cloud coverage, aerosol effects, etc. Mass balance methods calculate the total mass over a plume, and if some part of the plume is missing, the emission estimates will not be accurate. From the results obtained in this section, plumes having bad coverage produced higher outliers than other categories of plumes. Furthermore, blob-like plumes, which often occur due to low wind speed conditions, are challenging to quantify for mass balance methods. Mass balance methods cannot utilise three-dimensional wind speeds in the plume emission quantification process. We aim to overcome some of the previously mentioned inherent limitations of mass balance methods using atmospheric transport models. In the next section, we present our analysis of the quantification of synthetic plumes with atmospheric transport model based plume emission quantification approaches.

7.2. Assessment of plume emission quantification with atmospheric transport models

The automated plume detection algorithm frequently identifies methane plumes in TROPOMI measurements. To improve the emission quantification of these plumes, we aim to use atmospheric transport models. The atmospheric transport model based plume scaling approach is less complex to implement for multiple plumes compared to other methods discussed in Section 5.2.1. In the plume scaling approach, we quantify a plume by scaling the known emission rate of a separate synthetic plume generated at the same location and time as the plume to be quantified. In the previous section, we quantified synthetic plumes using mass balance methods and created some benchmark results. In order to improve plume emission quantification, the plume scaling approach needs to perform better than the benchmark results of mass balance methods. Since synthetic plumes are used in the plume scaling approach, uncertainties of the plume scaling approach are linked to the settings of atmospheric transport models. To investigate these uncertainties, we investigate the effects of different meteorological datasets, different model physics, and different transport models on synthetic plumes.

It was observed that different settings of atmospheric transport models could produce plumes with different geometrical features (Figure 6.6). Since there are missing pixels, the difference in shape/-size/orientation of plumes can lead to uncertainties in emission estimates given by the plume scaling approach. We analyse uncertainties related to the plume scaling approach by quantifying synthetic

plumes using other synthetic plumes generated with a different meteorological dataset or a different planetary boundary layer scheme, or a different atmospheric transport model. Figure 7.5 shows uncertainties in the plume scaling approach. Uncertainty for plumes generated with the different meteorological datasets (27.7%) is comparable to the uncertainty due to different atmospheric transport models (27.9%). Using different planetary boundary layer schemes in the WRF model has a higher impact on the IME of plumes (21.6%) than using different planetary boundary layer schemes in the FLEXPART model (4.5%). Please note that the planetary boundary layer schemes we considered for the FLEXPART model are different from the planetary boundary layer schemes considered for the WRF model.

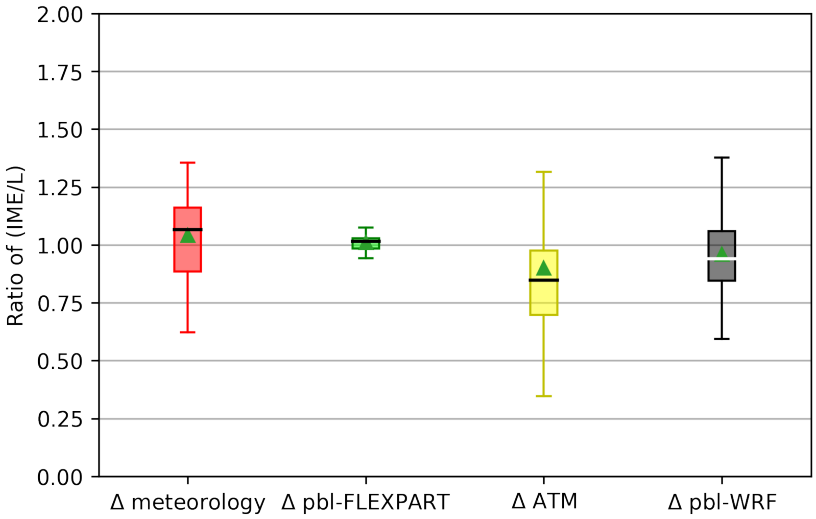


Figure 7.5: This figure shows box plots for IME/L ratios for two plumes generated with: 1) different meteorological datasets in FLEXPART, 2) different pbl schemes in FLEXPART, 3) different atmospheric transport models, and 4) different pbl schemes in WRF. Boxplot belonging to plumes generated with different atmospheric transport models show some negative bias. This bias might be due to non-optimal plume masking of FLEXPART and WRF plumes having different diffusion.

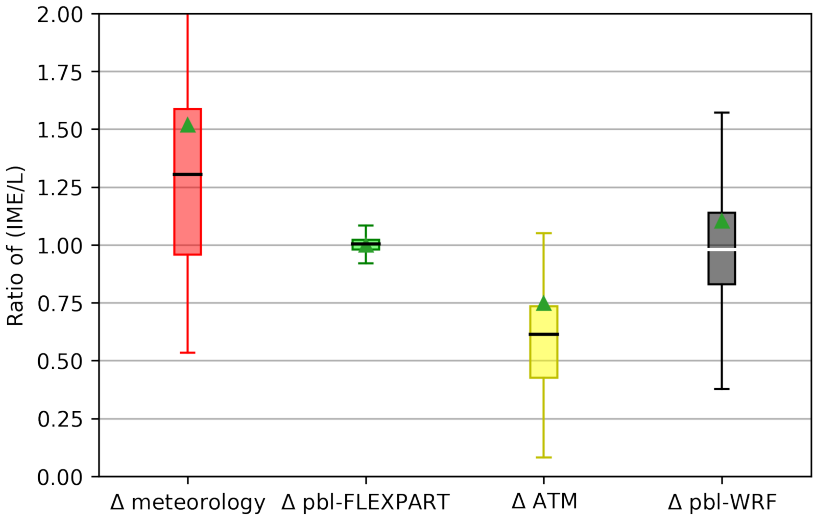


Figure 7.6: This figure shows box plots of IME/L ratios computed by considering only the common pixels between two plumes instead all the pixels over plume. Rest of the description of this plot is the same as the plot shown in Figure 7.5

We described a variation of plume scaling approach in Section 5.2. This variation belongs to a study performed to estimate emissions from a well blowout event that happened in Ohio [43] with the help of TROPOMI measurements. In this variation, we consider common pixels between two plumes (Figure

5.11) instead of all plume pixels. Figure 7.6 shows results from plume scaling approach that considers common plume pixels. Boxplots in this figure show higher uncertainty (62.7%, 30.8%, 30.8%) than boxplots in Figure 7.5 (27.7%, 27.9%, 21.6%). Using only the common part between two plumes to calculate IME/L is more limiting, and it puts a stricter requirement on the similarity between two plumes compared to the variation where all plume pixels are considered in plume emission quantification. If the diffusions of plumes are different or if plumes are slightly dissimilar, the plume scaling approach with common plume pixels performs poorly. Hence, for further analysis, we use the plume scaling approach that considers all plume pixels inside the plume mask.

7.2.1. Comparison between mass balance method and atmospheric transport model based plume emission quantification approach

Figure 7.7 shows the IME method estimates for synthetic plumes which were considered in the previous subsection (Figure 7.5) as a representative result from the mass balance methods. To compare the uncertainties of mass balance methods and atmospheric transport models, we compare interquartile ranges. Interquartile ranges can give an idea about the spread of plume emission quantification estimates, so a smaller interquartile range is better. Values of interquartile ranges for various plume emission quantification methods are listed in Table 7.1 and visualised in Figure 7.8. From these interquartile ranges, the plume scaling approach with all plume pixels can be seen producing consistently less uncertain results than the mass balance method. The idea behind using atmospheric transport models to improve plume emission quantification is that if IME quantifications of the two plumes are wrong in a similar way, we can still estimate the emission rate of a plume using the known emission rate of another plume with reasonable accuracy. From this and the previous subsection, the atmospheric transport model based plume scaling approach with all plume pixels can be seen producing results with less uncertainty (approximately 10% less) and less bias than mass balance methods.

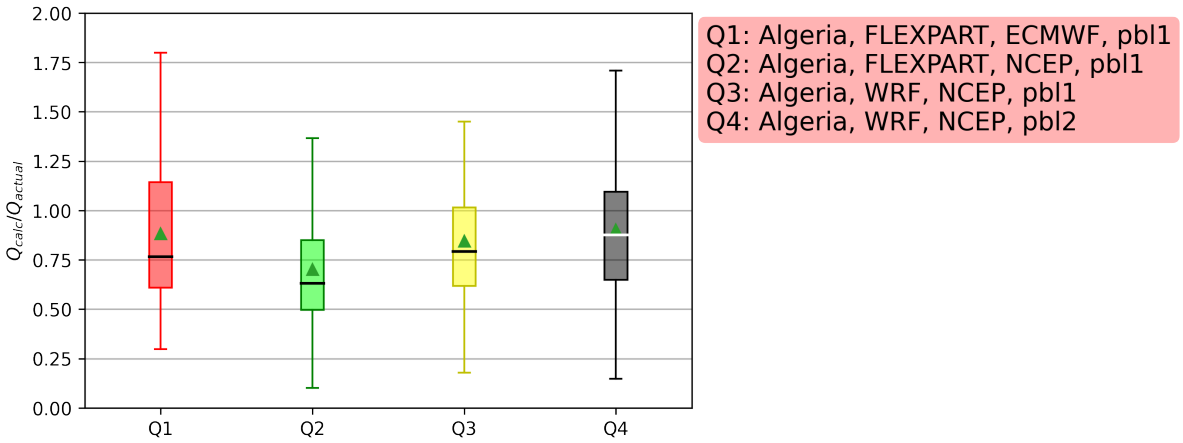


Figure 7.7: Ratio of mass balance calculated emission rates (Q_{calc}) with true emission rates (Q_{actual}) for four types of synthetic plumes: FLEXPART synthetic plumes generated using ECMWF meteo data and pbl scheme 1, FLEXPART synthetic plumes generated using NCEP meteo data and pbl scheme 1, WRF synthetic plumes generated with NCEP meteo data and pbl scheme 1, and WRF synthetic plume generated with NCEP meteo data and pbl scheme 2. Please note that there is a negative bias in the results since we consider synthetic plume images having missing pixels between 0%-70%.

Table 7.1: This table presents interquartile ranges (IQR) of box plots presented in the analysis of atmospheric transport models (Figure 7.5 and 7.6), and mass balance methods (Figure 7.7) for same set of plumes. Please note that pbl1 and pbl2 schemes mentioned with the FLEXPART model are different than the one mentioned with the WRF model. The results given in this table are visualised in Figure 7.1.

Case	Plume	Quantified with	IQR
A)	FLEXPART-ECMWF-pbl1	FLEXPART-ECMWF-pbl2 (Δpbl-FLEX, all pixels)	0.043
		FLEXPART-NCEP-pbl1 (Δmeteo, all pixels)	0.277
		FLEXPART-ECMWF-pbl2 (Δpbl-FLEX, common pixels)	0.042
		FLEXPART-NCEP-pbl1 (Δmeteo, common pixels)	0.627
		Mass balance	0.533
B)	FLEXPART-NCEP-pbl1	FLEXPART-NCEP-pbl2 (Δpbl-FLEX, all pixels)	0.045
		WRF-NCEP-pbl1 (ΔATM, all pixels)	0.279
		FLEXPART-NCEP-pbl2 (Δpbl-FLEX, common pixels)	0.041
		WRF-NCEP-pbl1 (ΔATM, common pixels)	0.308
		Mass balance	0.353
C)	WRF-NCEP-pbl1	WRF-NCEP-pbl2 (Δpbl-WRF, all pixels)	0.216
		WRF-NCEP-pbl2 (Δpbl-WRF, common pixels)	0.308
		Mass balance	0.396

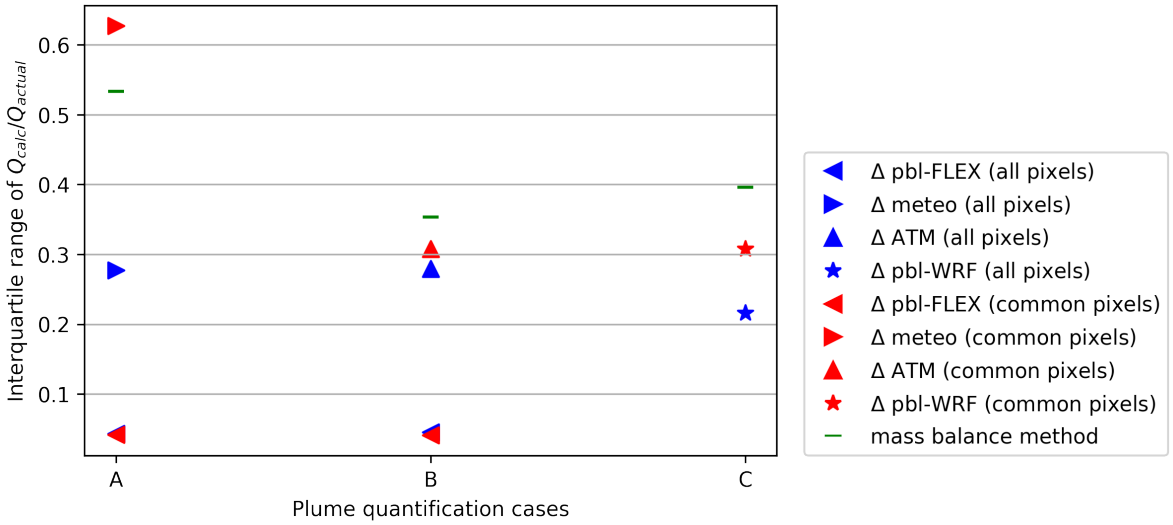


Figure 7.8: This figure shows interquartile ranges of boxplots for different cases listed in Table 7.1. Interquartile range for a particular plume scaling approach is the gap between 25th percentile and 75th percentile of boxplot containing Q_{calc}/Q_{actual} values.

In the analysis of mass balance methods (Section 7.1), we segregated plumes into different categories based on their geometries and checked how mass balance methods perform for each category. From that analysis, we found that quantifying blob-like plumes and plumes with missing pixels/ bad coverage using mass balance methods was challenging. We now apply the plume scaling approach to synthetic plumes segregated into different categories. Figure 7.9 shows a side-to-side comparison between performance of plume scaling approach and mass balance methods. Performance is calculated by measuring deviation of (Q_{calc}/Q_{actual}) from unity (equation 7.1), where Q_{calc} is the calculated emission rate and Q_{actual} is the known emission rate. The smaller the deviation from unity, the better the performance. The plume scaling approach can be seen performing better for blob-like plumes, plumes with bad coverage, and normal plumes compared to the mass balance method.

$$\text{performance} = 1 - \text{abs}\left(1 - \frac{Q_{calc}}{Q_{actual}}\right) \tag{7.1}$$

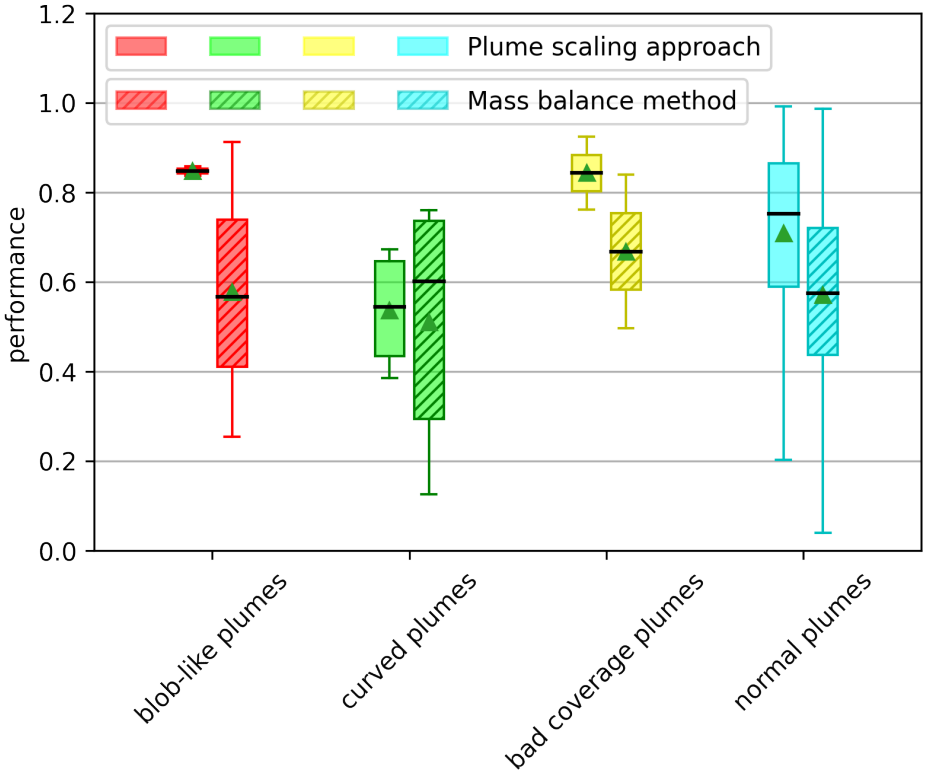


Figure 7.9: Comparison between the performance of plume scaling approach and mass balance method. Performance is calculated by measuring the deviation of Q_{calc}/Q_{actual} from unity. Q_{calc} is the calculated emission rate, and Q_{actual} is the actual emission rate of a synthetic plume. The threshold for bad coverage plumes is set as 40% coverage (missing pixels in the quadrant of a scene where plume lies, not missing pixels over plume), which is different than the threshold mentioned in Table 6.3. The rest of the thresholds are kept the same.

Missing pixels over plume is a limiting factor for the mass balance method. We compare the performance of the plume scaling approach and mass balance method for synthetic plumes having a different number of missing pixels (Figure 7.10).

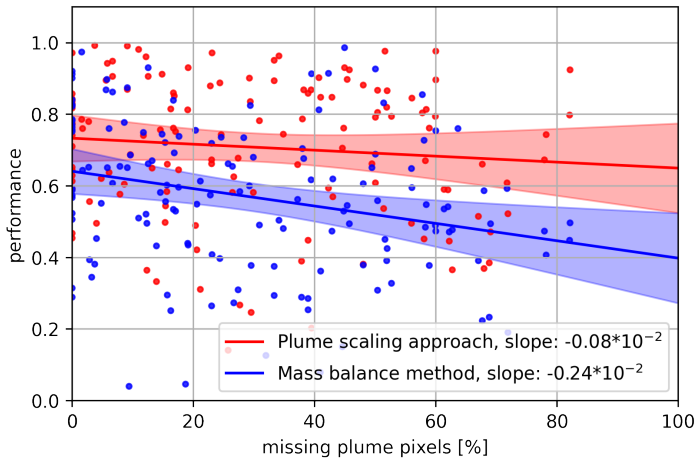


Figure 7.10: Variation in performance of plume scaling approach and mass balance method versus percentage of missing pixels over the plume. Each point represents a synthetic plume. Performance of the mass balance method and the plume scaling approach is calculated by measuring the deviation of Q_{calc}/Q_{actual} from unity. Q_{calc}/Q_{actual} ratio of 1 is the best performing case, and its deviation from 1 means degradation in performance. The Blue and red lines shown in this plot are linear curve fit through all the data points corresponding to that colour. The transparent colour regions in the plots represent a 99% confidence interval.

The performance of mass balance methods drops faster than the plume scaling approach for increasing percentage of missing plume pixels; for, e.g., at 0% missing pixels, the performance of the plume scaling approach is approximately 10% better than the mass balance method, and the difference increases to approximately 25% at 80% missing pixels. As expected, this result implies that the added value of the use of a transport model increases for plumes which are missing more pixels.

In this section, we compared the performance of the plume scaling approach to the performance of the mass balance method by quantifying synthetic plumes (Figure 7.9). It was found that the plume scaling approach can overcome several inherent limitations of mass balance methods like poor performance for blob-like plumes due to low wind speeds and missing plume pixels.

7.2.2. Performance indicators for plume scaling approach

Our aim is to know how well atmospheric transport models can represent TROPOMI plumes. We do not know the actual emission rates of TROPOMI plumes, and hence we cannot use the previously used method to calculate performance. We need some metrics that can be used as performance indicators for the plume scaling approach to be used in cases where the actual emission rates of plumes are unknown. In the previous sections, we discussed the plume scaling approach performing better for cases where plumes had similar geometries and orientations (Figure 6.5). We can use synthetic plume data and measure the similarity between two plumes and see how it correlates with the performance of the plume scaling approach. To quantify the similarity between a synthetic plume and a TROPOMI plume, we use two metrics. The first metric is the overlap between plume masks of TROPOMI plumes and synthetic plumes. The second metric is the correlation between two plumes. To obtain similarity, we calculate an average of overlap and correlation ((overlap + correlation)/2). Please note that correlation can range between -1 and 1, and in this analysis, we rescale the correlation values to fit between 0 and 1 using the equation (1+correlation)/2. Next, we plot similarity between two plumes versus performance (equation 7.1) of plume scaling approach (Figure 7.11).

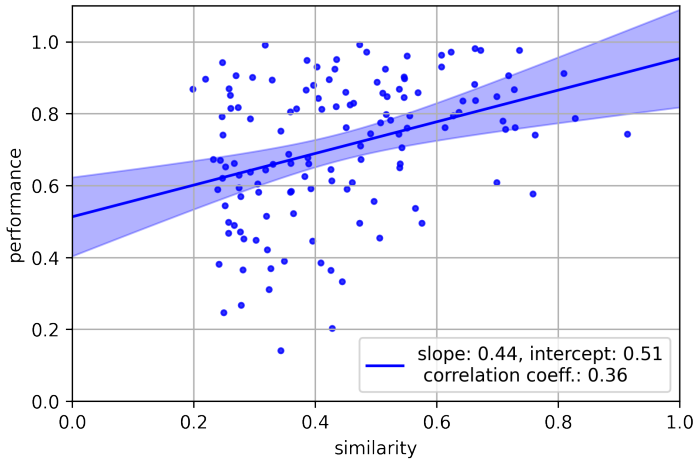


Figure 7.11: Variation in performance of plume scaling approach versus similarity between plumes. Each point represents a synthetic plume quantified using another synthetic plume. The similarity is calculated by averaging overlap and correlation between two synthetic plumes. Performance of the plume scaling approach is calculated by measuring the deviation of Q_{calc}/Q_{actual} from unity. The blue line is a linear curve fit through all the data points. This fit indicates that the performance of the plume scaling approach increases with an increasing similarity between two synthetic plumes. The transparent colour region in the plot represents a 99% confidence interval.

From Figure 7.11, we can see that performance for simulations with similar plumes is always good, while performance for simulations with dissimilar plumes is sometimes good, sometimes bad. Hence, the performance of the plume scaling approach on average increases with the similarity between two synthetic plumes. Hence, we use similarity as a proxy from performance for the plume scaling approach where actual emission rates of plumes are unknown. In the following section, we test the plume scaling approach on TROPOMI plumes, and we use similarity as a performance indicator to gauge the performance of the plume scaling approach.

7.3. Quantification of TROPOMI plumes

Our objective is to investigate how well TROPOMI plumes can be quantified using atmospheric transport models using the plume scaling approach. The plume scaling approach requires synthetic plumes generated by atmospheric transport models. We consider synthetic plumes generated with the following settings of atmospheric transport models to quantify TROPOMI plumes: FLEXPART-ECMWF-pbl1, FLEXPART-NCEP-pbl1, WRF-NCEP-pbl1, and WRF-NCEP-pbl2. We excluded FLEXPART-ECMWF/NCEP-pbl2 settings since it has been found that these plumes are almost identical to FLEXPART-ECMWF/NCEP-pbl1 generated plumes (Figure 6.6).

For this analysis, we considered 50 TROPOMI plumes observed over the Algeria region in 2020. Emission locations of these TROPOMI plumes were roughly known from earlier studies at SRON, but in practice, when the TROPOMI automated plume detection algorithm detects a plume, we often do not know their actual emission locations. Hence, a source localisation method might be required for TROPOMI plumes. For TROPOMI plumes, due to the presence of noise, getting good similarity with a synthetic plume is challenging. However, we expect the trend of similarity and performance (Figure 7.11) to be similar for TROPOMI plume cases. To compare different settings of atmospheric transport models for their ability to replicate TROPOMI plumes, we compare the similarity values. Figure 7.12 presents this comparison.

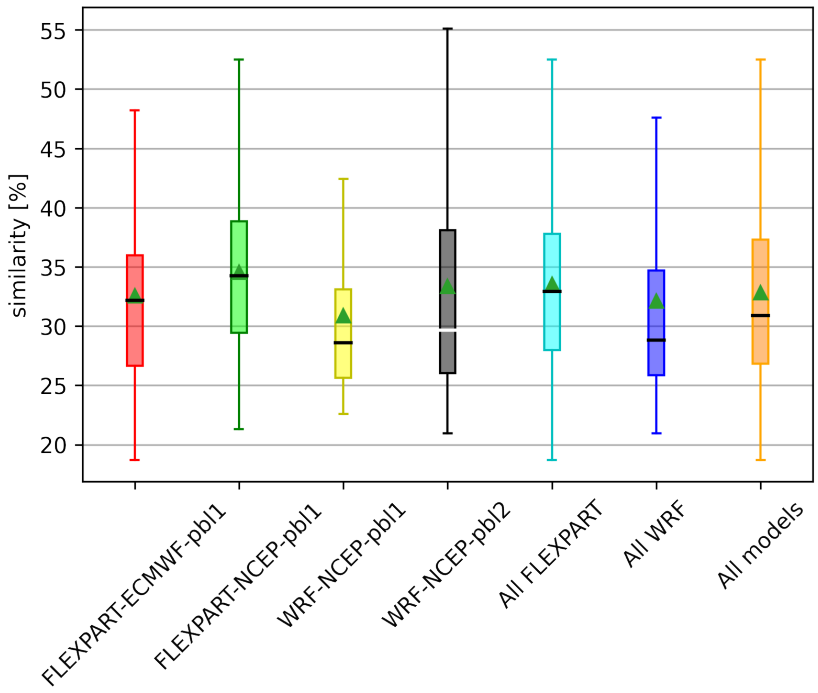


Figure 7.12: Similarity values between synthetic plumes and TROPOMI plumes. The first four boxplots correspond to four different settings of atmospheric transport models. The next two boxplots correspond to FLEXPART and WRF atmospheric transport models and are supersets of first-second and third-fourth boxplots, respectively. The final boxplot is the superset of the first four boxplots.

From Figure 7.12, the FLEXPART generated synthetic plumes show slightly better similarity to TROPOMI plumes compared to the WRF generated synthetic plumes, but it is not very significant. Hence, we also compare the computational costs of generating synthetic plumes with different atmospheric transport models. The computational costs required for the WRF model are 1-2 orders of magnitude higher than the FLEXPART model. Based on the similarity metric and the computational costs, we consider the FLEXPART model more suitable to replicate the TROPOMI plumes.

TROPOMI plume emission quantification using the plume scaling approach requires synthetic plumes, and generating synthetic plumes increases the complexity of the setup, whereas implementing a mass balance method for plume emission quantification is much easier, but it comes at the cost of accuracy.

We could make use of the mass balance method to quantify TROPOMI plume cases where they perform similarly to the plume scaling approach. We found that blob-like plumes (plumes with short length) and plume images with a large number of missing pixels are challenging to quantify for the mass balance method (Figure 7.3). We see if plumes having large lengths and plumes with fewer missing pixels can be quantified with reasonable accuracy with the mass balance method. The performance of the mass balance method for plumes with few missing pixels is close to the atmospheric transport model approach (Figure 7.10). Plumes with larger lengths are often classified as normal plumes, and these cases can be quantified with the mass balance method with reasonable accuracy (Figure 7.3). Hence, we can quantify TROPOMI plumes that have few missing pixels or that are lengthier than a certain threshold using the mass balance method, and in doing so, we will not lose a lot of accuracy and save some computational complexities. On the other hand, TROPOMI plumes that have large missing pixels or shorter lengths can be quantified well with the plume scaling approach. To automate the selection between the mass balance method and the plume scaling approach, we developed a decision tree algorithm (Figure 7.13).

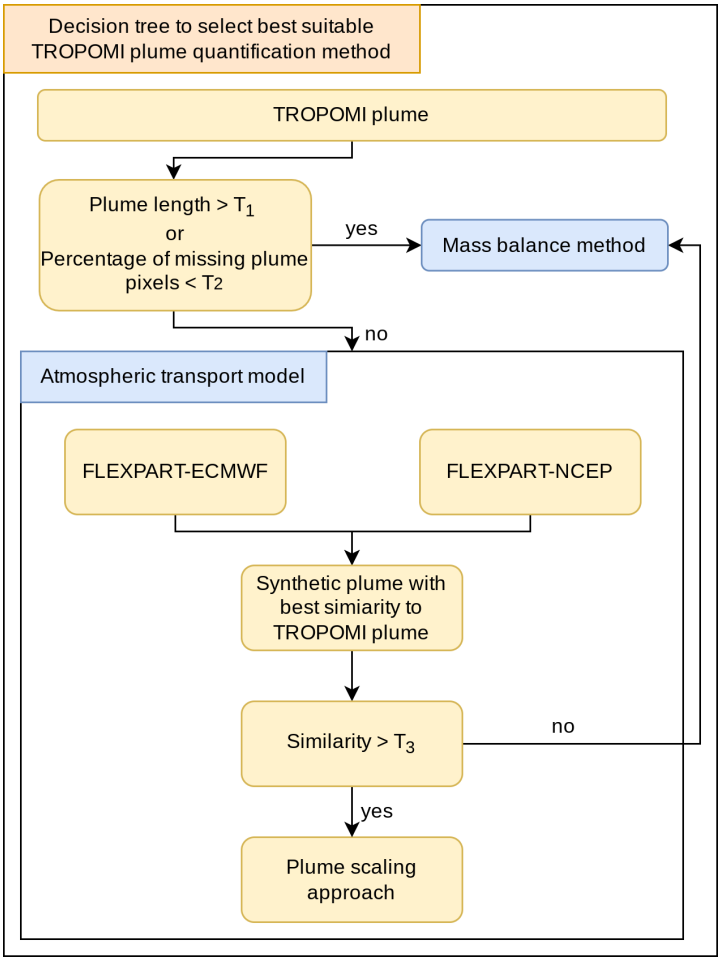


Figure 7.13: Flowchart for decision tree which decides best suitable plume emission quantification model for a given TROPOMI plume. This decision tree algorithm uses three threshold parameters: T_1, T_2 and T_3. T_1 is the threshold for the length of a plume, T_2 is the threshold for the percentage of missing plume pixels, and T_3 is the threshold for the similarity between two plumes.

TROPOMI plumes that are longer than a certain threshold (T_1) or have missing pixels below a certain threshold (T_3) are quantified with the mass balance method. TROPOMI plumes that do not satisfy these conditions are quantified using the plume scaling approach. In the decision tree algorithm, for the plume scaling approach, we had to find the most suitable atmospheric transport model setup for generating synthetic plumes to quantify TROPOMI plumes. We investigated four different settings of atmospheric transport models in this section. It was found that with the FLEXPART model, we can get

a slightly better similarity to the TROPOMI plumes. Furthermore, the WRF model takes around 3200 hours to simulate plumes for an entire year, and the FLEXPART model takes close to 100 hours to simulate plumes over an entire year. Due to slightly better similarity with TROPOMI plumes and less computational expense, the FLEXPART model seems more suitable to quantify TROPOMI plumes. With the FLEXPART model, we have a choice of using either the ECMWF meteorological data or the NCEP meteorological data. From Figure 7.12, performance of FLEXPART-ECMWF-pbl1 setup was close to the performance of FLEXPART-NCEP-pbl1 setup. Instead of choosing any one setup prior to generating synthetic plumes, we decided to take a posterior choice. With this choice, we select either FLEXPART-ECMWF or FLEXPART-NCEP synthetic plume based on the highest similarity to the TROPOMI plume. If the similarity value is below a certain threshold, we go back to the mass balance method for emission quantification.

We illustrate how the decision tree algorithm works and what are the impacts of decisions made by the decision tree on TROPOMI plume emission quantification. For this purpose, we quantify emissions from the TROPOMI plumes using the mass balance method and the plume scaling approach that uses the FLEXPART model and two different meteorological datasets. The plume emission quantification results are given in Figure 7.14.

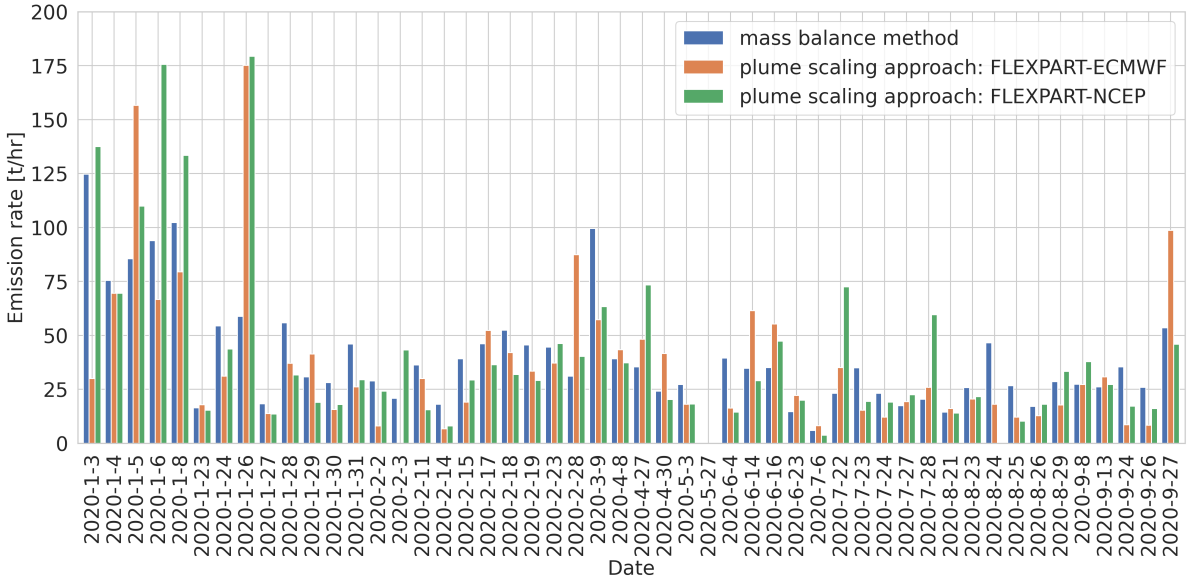


Figure 7.14: Emissions of 50 TROPOMI plumes observed over Algeria in 2020 quantified using 1) the IME mass balance method, 2) the plume scaling approach that uses FLEXPART-ECMWF synthetic plumes, and 3) the plume scaling approach that uses FLEXPART-NCEP synthetic plumes. Please note that the x-axis is not linear. Some quantifications are missing from the plot due to very bad TROPOMI coverage.

We use Figure 7.14 to determine how sensitive the TROPOMI plume emission quantifications are to the threshold parameters T_1 and T_2 used in the decision tree algorithm. The results from this sensitivity analysis are presented in Figure 7.15 and Figure 7.16. The difference between the emission rates obtained by the plume scaling approach and the mass balance method increases with increasing missing pixels and decreasing plume length. We do not know the actual emission rates of TROPOMI plumes. Still, based on our synthetic plume study in phases II and III, the plume scaling approach should give emission estimates closer to the truth than the mass balance method. If we keep the threshold of missing plume pixels too low or the threshold for plume length too high, it will increase the computational cost since we will use the plume scaling approach on many TROPOMI plumes. Hence, we need to select suitable threshold parameters in the decision tree algorithm to improve the emission estimates and, at the same time, keep the computational expenses low.

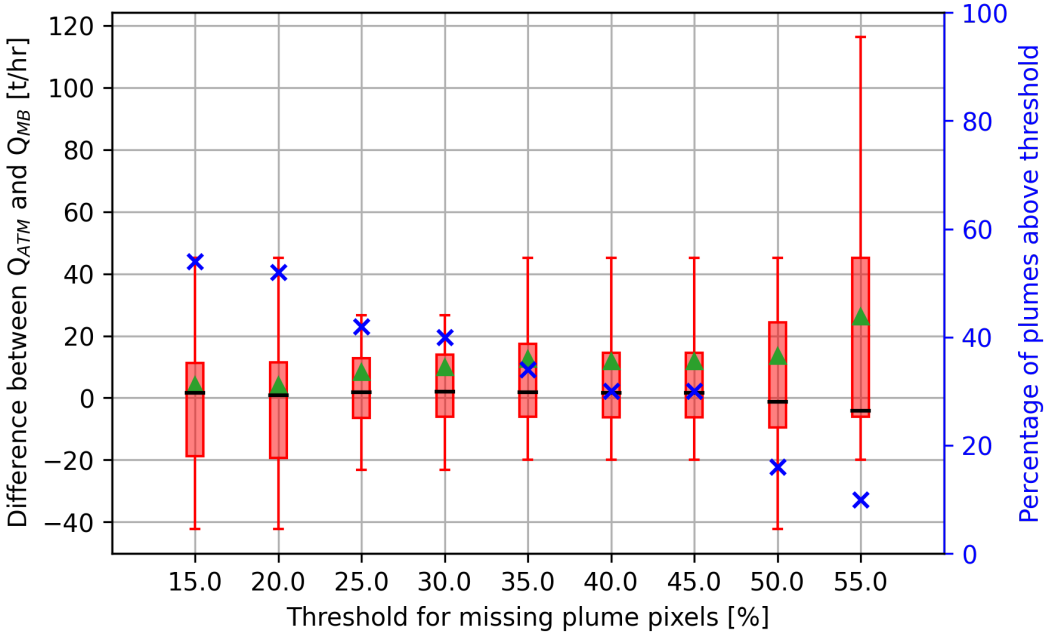


Figure 7.15: Difference between emission rates calculated by the atmospheric transport model based plume scaling approach and the mass balance method versus the threshold for missing plume pixels (sensitivity study). The blue cross markers indicate the percentage of plumes above the threshold for missing plume pixels. Emissions of all the TROPOMI plumes above the threshold are quantified with the plume scaling approach. Each boxplot represents the spread of the difference between Q_{ATM} and Q_{MB}. The black horizontal lines in the boxplots represent median values, and the green triangles represent mean values.

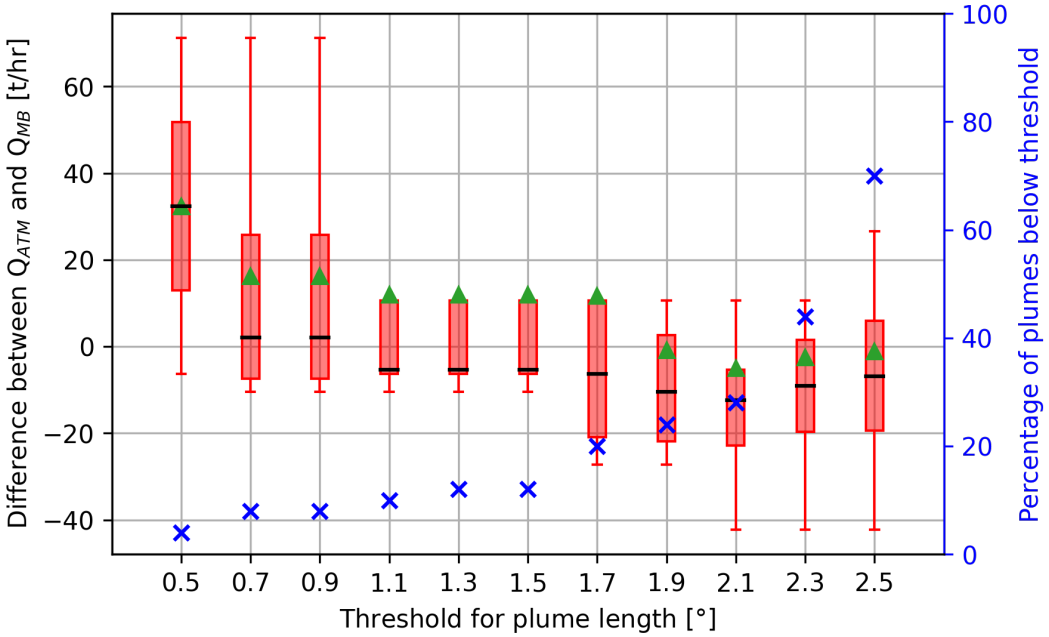


Figure 7.16: Difference between emission rates calculated by the atmospheric transport model based plume scaling approach and the mass balance method versus the threshold for plume length (sensitivity study). The blue cross markers indicate the percentage of plumes below the threshold for plume length. Emissions of all the TROPOMI plumes below the threshold are quantified with the plume scaling approach. Each boxplot represents the spread of the difference between Q_{ATM} and Q_{MB}. The black horizontal lines in the boxplots represent median values, and the green triangles represent mean values.

We now give an example of how the TROPOMI plume emission quantification is affected by making a certain choice of threshold parameters used in the decision tree algorithm. The threshold parameters used for this example are mentioned in Table 7.2. The decision tree algorithm selects the mass balance method for quantifying emissions from TROPOMI plumes that are longer than 1.5° or have less than 30% missing pixels (around 60% TROPOMI plumes). The mass balance method performed similarly to the plume scaling method for those plumes. Also, the mass balance methods are easier to implement and computationally faster (approximately one order of magnitude) than the plume scaling approach.

On the other hand, if the TROPOMI plumes are shorter than 1.5° and have more than 30% missing pixels, the performance of the mass balance method can drop by 10-20%, but the performance of the plume scaling approach does not drop by a lot. Hence, with the decision tree algorithm, we can improve the emission quantification performance by 10-20% for the cases that are challenging for the mass balance methods. For the rest of the cases, we do not use the plume scaling approach, which saves computational time by one order of magnitude.

Table 7.2: This table presents the threshold parameters used for an example showcasing working of the decision tree algorithm.

Parameter	Description	Value
T_1	Threshold for plume length	1.5°
T_2	Threshold for percentage of missing plume pixels	30%
T_3	Threshold for similarity between plumes	30%

From the previous example, we showcased the trade-off between ease of use and accuracy while choosing between the mass balance method and plume scaling approach in the decision tree algorithm. This project is targeted to explore atmospheric transport model based plume emission quantification in order to improve the plume emission quantification routine of the TROPOMI automated plume detection algorithm. We have developed a decision tree which selects the most suitable plume emission quantification approach between the mass balance method and plume scaling method for a given TROPOMI plume. Integration of the decision tree in the TROPOMI automated plume detection algorithm pipeline might give us some advantages, but it also comes with several challenges. With the plume scaling approach, we can reduce uncertainty and bias in plume emission quantification estimates for the cases that are challenging for the mass balance method. We established that plumes with short lengths (blob-like plumes) and plumes having missing pixels are challenging to quantify for the mass balance method. Though blob-like plumes are not very common (Table 6.2), plumes with missing pixels are frequently observed (around 40% of TROPOMI plumes considered in this analysis had missing pixels above 30%) and are still a major limiting factor for the mass balance method. These plumes can be quantified well with the plume scaling approach. On the other hand, atmospheric transport models used for generating synthetic plumes require meteorological data that often needs to be obtained from external sources, which is time-consuming. There is a possibility that plumes generated for the plume scaling approach might not match well with the TROPOMI plumes, and in that case, we need to go back to the mass balance method for plume emission quantification. It is still challenging to make any prior decision on which settings of the atmospheric transport model are most suitable to reproduce a TROPOMI plume. There are also some common challenges for the mass balance method and plume scaling approach. Both plume emission quantification methods require the source location of plumes. TROPOMI automated plume detection algorithm can give an estimated source location, but it is often not precise, which makes plume emission quantifications uncertain. Furthermore, both plume emission quantification methods rely on plume mask settings. Masking TROPOMI plume pixels is challenging due to noise, and this affects the emission quantification of plumes.

If the decision tree algorithm is successfully integrated with the TROPOMI automated plume detection algorithm, we can get a better idea about the impact of methane sources around the globe due to improved emission estimates. Only improving the quantification of methane emissions from super-emitters around the globe is not sufficient to reduce anthropogenic methane emissions, but it can positively benefit the process of methane emission regulation by putting better independent constraints on emissions.

8

Conclusion

The main research question of this thesis, as outlined in Chapter 3 is:

> **How can Eulerian and Lagrangian atmospheric transport models be used to improve the emission quantification of methane plumes detected by the TROPOMI automated plume detection algorithm?**

From this research work, it can be concluded that atmospheric transport models can be used to overcome several inherent limitations of the mass balance methods and reduce the uncertainty of plume emission quantification by nearly 10%. There is a trade-off between ease of use and accuracy when using mass balance methods and atmospheric transport models to quantify emissions from TROPOMI plumes. An optimum way for quantifying emissions from TROPOMI plumes while maintaining a balance between ease of use and accuracy was described by this study in the form of a decision tree algorithm. The main research objective of this study has been met. In the following paragraphs, we answer the leading research questions and sub-questions associated with all four phases of this study.

The literature study performed for this research highlighted that anthropogenic methane emissions from fossil fuel facilities and waste management could be reduced in a cost-effective way. Global bottom-up emission inventories provide uncertain information about methane emissions from fossil fuel facilities and waste management, and top-down measurements using satellite remote sensing can be used to place independent constraints and derive an integrated picture of methane emissions. Methane plumes are frequently observed in the TROPOMI data, and to feasibly identify plumes in that daily global data, an automated plume detection algorithm was developed. The TROPOMI automated plume detection algorithm uses mass balance methods. However, mass balance methods provide emissions with relatively large uncertainties. Atmospheric transport models can be used to overcome some limitations of mass balance methods and improve the TROPOMI plume emission quantification. Atmospheric transport models are categorized into the Eulerian and Lagrangian models. We consider the WRF and FLEXPART models to simulate particle transport in the Eulerian and Lagrangian frames of reference, respectively. We can generate synthetic plumes using the WRF-CHEM and FLEXPART models, which allows us to simulate TROPOMI plumes. It was found that the atmospheric transport models based plume scaling approach is relatively less complex and easier to set up compared to other atmospheric transport models based approaches. The plume scaling approach is also closer to the mass balance approaches, and hence plume scaling approach is considered in this study.

To assess the plume emission quantification approaches, an inventory of TROPOMI-like synthetic plumes with known emission rates and emission locations was generated using the atmospheric transport models. However, replicating details like noise seen in TROPOMI images to synthetic plume images was challenging. TROPOMI noise was partially taken care of by introducing Gaussian noise in synthetic plume images.

We classified synthetic plumes according to their geometrical features and coverage to identify challenging cases for mass balance methods. Four categories were found suitable for the classification of

synthetic plumes: normal plumes, curved plumes, plumes with bad coverage, and blob-like plumes. A decision tree classifier was developed that used geometrical features of plumes and regional missing pixel information to classify plumes. The plume tracer algorithm was used to extract geometrical features from plumes, which worked for complex plume cases and even for plumes having several missing pixels. Most of the classified plumes belonged to normal plumes, while blob-like plumes occurred the least frequently. The classification algorithm classified normal plumes, curved plumes, plumes with bad coverage, and blob-like plumes with true positive rates of 93%, 75%, 75%, and 56%, respectively.

Mass balance methods might require tuning of settings before quantifying plumes to reduce biases present in the results. To fix this bias, recalibration of parametric coefficients used in effective wind speed can be performed. Plume mask settings also affect mass balance estimates. Four implementations of the IME method and two implementations of the CSF method were investigated in this research. Each variation had a different way of selecting/masking plume pixels. IME plume dilation method produced the most uncertain results among different implementations of the IME method since missing pixels in plumes posed a challenge for the IME plume dilation method. Plumes having higher emission rates (> 15 [t/hr]) were easier to mask than plumes with low emission rates (< 15 [t/hr]) for background noise of 5 [ppb]. Blob-like plumes and plumes with missing pixels are challenging to quantify with mass balance methods. Limitations of mass balance methods can be divided into limitations that can be addressed and limitations that cannot be solved. Limitations due to parametric coefficients in effective wind speed equation can be solved using recalibration, and limitations due to plume mask and uniform effective wind speed can be solved by developing advanced IME and CSF methods. Limitations of mass balance methods that cannot be solved include inherent limitations due to low wind speed conditions, missing plume pixels, and no inclusion of three-dimensional particle transport.

The plume scaling approach was exploited to overcome some inherent limitations of mass balance methods. The plume scaling approach uses synthetic plumes to quantify other plumes. However, two plumes generated with different settings of atmospheric transport models are often not the same. This can be observed when we simulate plumes using different meteorological datasets, planetary boundary layer schemes, or different atmospheric transport models. The shapes and orientation of plumes are highly sensitive to meteorological datasets used for generating them. Plumes are less sensitive to changes in planetary boundary layer schemes in the FLEXPART model than planetary boundary layer schemes in the WRF model due to the way meteorological data is used with these models. Using different atmospheric transport models affects the diffusion of plumes, but it often does not affect their orientation.

Synthetic plumes generated by the WRF (Eulerian) model for an entire year took close to 3200 hours. In contrast, the FLEXPART (Lagrangian) model took close to 100 hours, which is one magnitude less than the WRF model.

A manual investigation was performed to identify several challenging cases for the plume scaling approach. The plume scaling approach can work well if two plumes have similar geometry (size, shape, orientation, diffusion) and missing pixels. Filtering the challenging cases using a filtering algorithm removed almost 50% of plumes, but the filtering criteria can be relaxed in exchange for accuracy of plume emission quantification. Quantification of remaining plumes using the plume scaling approach produced results that were nearly 10% less uncertain than mass balance estimates. The plume scaling approach performed 10% better than mass balance methods in quantifying blob-like plumes, plumes with bad coverage and normal plumes. We investigated another variation of the plume scaling approach where only common pixels between two plumes were considered in quantification instead of all plume pixels; this variation puts stringent constraints on the similarity between two plumes, which produced more uncertain results than the variation that considered all plume pixels.

Actual emission rates of TROPOMI plumes are often unknown, and some metrics to gauge the performance of the plume scaling approach had to be developed. It was found that similarity (obtained by averaging the percentage of overlap and correlation between plumes) can be used as a proxy for the performance of the plume scaling approach. Plumes generated with the FLEXPART model had slightly better similarity (approximately 5% better) to TROPOMI plumes than those generated with the WRF model, and the FLEXPART model is approximately one order of magnitude faster than the WRF model; hence the FLEXPART model is more suitable for replicating TROPOMI plumes.

We developed a decision tree algorithm that selects the most suitable method to quantify a given TROPOMI plume. This decision tree chooses between the mass balance method and the plume scaling approach based on the length of the TROPOMI plume or missing pixels in the TROPOMI plume. With the decision tree algorithm, we can target the plume scaling approach to quantify emissions from the plumes that are challenging for the mass balance methods and improve the estimates by nearly 10%. Integration of the decision tree algorithm to the automated plume detection algorithm can be challenging due to the complexities of atmospheric transport models.

9

Recommendations for future research

Accounting for noise seen in TROPOMI images is challenging, and we only partially accounted for randomly distributed noise in TROPOMI images by artificially adding Gaussian distributed noise in synthetic plume images. In the future, sampled noise from the parts of real TROPOMI images that do not belong to the TROPOMI as well as synthetic plume masks can be superimposed on synthetic plume images, and by doing this, we might better account for real TROPOMI noise.

We used synthetic plumes generated over Algeria and Sasol regions in 2020 to investigate the effects of different regions on mass balance plume emission quantification. This analysis is limited due to the use of only two regions. In the future, more regions can be included in this analysis to investigate any effects of meteorological conditions, weather synoptic, etc., on mass balance plume emission quantification.

We considered several variations of IME and CSF in the analysis of mass balance methods. There are some more sophisticated variations of IME and CSF mass balance methods, like the one which uses the plume tracer algorithm to trace plumes and draws a plume box around them. These variations can be analysed in the future to extend the analysis of mass balance methods.

To identify plume pixels, we used plume mask filtering criteria of mean plus 1.8 times the standard deviation of all pixels. This criterion is not sufficient to find an optimum mask for plumes that have different diffusions, which adds uncertainty to plume emission quantification results. Hence, some advanced methods can be developed in the future to find optimum plume masks for differently diffused plumes. One way to mask the plumes would be to use the plume tracer algorithm to track the plume and then dilate the plume mask from the plume tracer line outwards to capture plume pixels.

Although mass balance methods have several limitations, we still include them in the decision tree algorithm due to their feasibility, speed, and reasonable accuracy in quantifying lengthy plumes with few missing pixels. In the settings of mass balance methods, we used an effective wind speed equation that considered uniform wind speed for plume. Uniform effective wind speed does not account for variations in wind speed direction or magnitude over time or space and can produce uncertain mass balance plume emission quantification results. In the future, some advanced methods can be developed that can account for variations in wind speed by considering time history and spatial distribution of wind speeds. Moreover, developing this method takes mass balance methods closer to the atmospheric transport models.

This study analysed the atmospheric transport model based plume scaling approach. We chose the plume scaling approach due to its feasibility of implementation on a large number of plumes. Another plume emission quantification method described by a study performed to investigate methane emissions from well blowout in Louisiana [64] is more sophisticated than the plume scaling approach. The Louisiana well blowout study method optimises concentrations in synthetic plumes to account for the emission rate of TROPOMI plume, and it is most likely to produce better results than the plume scaling approach. In the future, there is an opportunity to analyse the plume emission quantification method

described in the Louisianan well blowout study and compare its performance to the plume scaling approach.

Backward trajectory simulations using the Lagrangian transport model can also be used to quantify TROPOMI plumes. In backward trajectory simulations, TROPOMI plume pixels become the starting point of the simulation. Backward trajectory simulations might perform better than forward trajectory simulations since we can better account for concentrations seen in the TROPOMI pixels, and we do not rely on the similarity between two plumes. Due to the time limitations of this project, this direction was not explored. In the future, there is an opportunity to analyse plume emission quantification using backward trajectory simulations.

We considered 50 TROPOMI plumes observed over the Algeria region in 2020 to test the performance of the plume scaling approach and generate a decision tree algorithm which selects the best setup to quantify TROPOMI plumes. Our TROPOMI plume dataset was statistically limiting to derive any general conclusions on the efficacy of the plume scaling approach. In the future, the TROPOMI plume dataset might be expanded to obtain more statistically robust results, which might help in expanding and improving the decision tree algorithm. Furthermore, we considered TROPOMI plumes that had known approximate emission locations. When TROPOMI automated plume detection algorithm detects a plume, its emission location is often not exactly known. A source localisation method based on Lagrangian backward trajectory simulations can be developed in the future to help with estimating the source locations of TROPOMI plumes.

Bibliography

[1] *S5P*. 2022. URL: https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-5p (visited on 04/24/2022).

[2] Lenny Bernstein et al. *IPCC, 2007: climate change 2007: synthesis report*. 2008.

[3] Intergovernmental Panel On Climate Change. "Climate change 2007: the physical science basis". In: *Agenda* 6.07 (2007), p. 333.

[4] Thomas Lauvaux et al. "Global assessment of oil and gas methane ultra-emitters". In: *Science* 375.6580 (2022), pp. 557–561.

[5] Stefanie Kirschke et al. "Three decades of global methane sources and sinks". In: *Nature geoscience* 6.10 (2013), pp. 813–823.

[6] *Anthropogenic methane sources*. https://www.epa.gov/gmi/importance-methane. Accessed: 2021-10-06.

[7] Ed Dlugokencky. *NOAA/GML*. 2022. URL: gml.noaa.gov/ccgg/trends_ch4/ (visited on 03/19/2022).

[8] Edward J Dlugokencky et al. "The growth rate and distribution of atmospheric methane". In: *Journal of Geophysical Research: Atmospheres* 99.D8 (1994), pp. 17021–17043.

[9] Intergovernmental Panel on Climate Change. *Global warming of 1.5° C: an IPCC special report on the impacts of global warming of 1.5° C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty*. Intergovernmental Panel on Climate Change, 2018.

[10] Hinrich Schaefer. "On the causes and consequences of recent trends in atmospheric methane". In: *Current Climate Change Reports* 5.4 (2019), pp. 259–274.

[11] Marielle Saunois et al. "The global methane budget 2000–2017". In: *Earth system science data* 12.3 (2020), pp. 1561–1623.

[12] CV Cole et al. "Global estimates of potential mitigation of greenhouse gas emissions by agriculture". In: *Nutrient cycling in Agroecosystems* 49.1 (1997), pp. 221–228.

[13] David J Griggs and Maria Noguer. "Climate change 2001: the scientific basis. Contribution of working group I to the third assessment report of the intergovernmental panel on climate change". In: *Weather* 57.8 (2002), pp. 267–269.

[14] Keith Paustian et al. "Agricultural mitigation of greenhouse gases: science and policy options". In: *2001 Conference Proceedings, First National Conference on Carbon Sequestration*. Washington, DC: Conference on Carbon Sequestration. 2001.

[15] AR Mosier et al. "Mitigating agricultural emissions of methane". In: *Climatic change* 40.1 (1998), pp. 39–80.

[16] Pete Smith, Dave Reay, and Jo Smith. "Agricultural methane emissions and the potential formitigation". In: *Philosophical Transactions of the Royal Society A* 379.2210 (2021), p. 20200451.

[17] Janet Ranganathan et al. "How to sustainably feed 10 billion people by 2050, in 21 charts". In: (2018).

[18] Liste Des Participants. "Good practice guidance and uncertainty management in national greenhouse gas inventories". In: *Order* (2001).

[19] Jens E Frøiland Jensen and R Pipatti. "CH4 emissions from solid waste disposal". In: *Good practice guidance and uncertainty management in national greenhouse gas inventories* (2000), pp. 419–439.

[20] Daniel H Cusworth et al. “Using remote sensing to detect, validate, and quantify methane emissions from California solid waste operations”. In: *Environmental Research Letters* 15.5 (2020), p. 054012.

[21] Thomas Bruckner et al. “Energy systems”. In: (2014).

[22] Bharat Raj Singh and Onkar Singh. *Global trends of fossil fuel reserves and climate change in the 21st century*. Vol. 8. chapter, 2012.

[23] Iea. *Oil market report - 2009 archives – analysis*. URL: https://www.iea.org/reports/oil-market-report-2009-archives.

[24] J. Levine et al. “Biomass Burning”. In: (Feb. 1993). DOI: 10.1007/978-3-642-84605-2_14.

[25] Emilio Chuvieco Salinero. *Wildland fire danger estimation and mapping: the role of remote sensing data*. Vol. 4. World Scientific, 2003.

[26] Hanqin Tian et al. “The terrestrial biosphere as a net source of greenhouse gases to the atmosphere”. In: *Nature* 531.7593 (2016), pp. 225–228.

[27] J Couwenberg et al. “Methane emissions from peat soils (organic soils, histosols): facts, MRV-ability, emission factors.” In: *Methane emissions from peat soils (organic soils, histosols): facts, MRV-ability, emission factors.* (2009).

[28] Paul H Glaser et al. “Surface deformations as indicators of deep ebullition fluxes in a large northern peatland”. In: *Global Biogeochemical Cycles* 18.1 (2004).

[29] Shokoufeh Salimi, Suhad AAAN Almuktar, and Miklas Scholz. “Impact of climate change on wetland ecosystems: A critical review of experimental wetlands”. In: *Journal of Environmental Management* 286 (2021), p. 112160.

[30] Cain Silvey et al. “Plant Species and Hydrology as Controls on Constructed Wetland Methane Fluxes”. In: *Soil Science Society of America Journal* 83.3 (2019), pp. 848–855.

[31] Zhen Zhang et al. “Emerging role of wetland methane emissions in driving 21st century climate change”. In: *Proceedings of the National Academy of Sciences* 114.36 (2017), pp. 9647–9652.

[32] Tia R Scarpelli et al. “A global gridded (0.1× 0.1) inventory of methane emissions from oil, gas, and coal exploitation based on national reports to the United Nations Framework Convention on Climate Change”. In: *Earth System Science Data* 12.1 (2020), pp. 563–575.

[33] Daniel J Varon et al. “Quantifying methane point sources from fine-scale satellite observations of atmospheric methane plumes”. In: *Atmospheric Measurement Techniques* 11.10 (2018), pp. 5673–5686.

[34] Daniel J Jacob et al. “Satellite observations of atmospheric methane and their value for quantifying methane emissions”. In: *Atmospheric Chemistry and Physics* 16.22 (2016), pp. 14371–14396.

[35] O Hasekamp et al. “Algorithm theoretical baseline document for Sentinel-5 Precursor methane retrieval”. In: *Netherlands Institute for Space Research* (2019), p. 67.

[36] JR Worden et al. “Quantifying lower tropospheric methane concentrations using GOSAT near-IR and TES thermal IR measurements”. In: *Atmospheric Measurement Techniques* 8.8 (2015), pp. 3433–3445.

[37] C Frankenberg et al. “Satellite chartography of atmospheric methane from SCIAMACHY on board ENVISAT: Analysis of the years 2003 and 2004”. In: *Journal of Geophysical Research: Atmospheres* 111.D7 (2006).

[38] C Frankenberg et al. “Global column-averaged methane mixing ratios from 2003 to 2009 as derived from SCIAMACHY: Trends and variability”. In: *Journal of Geophysical Research: Atmospheres* 116.D4 (2011).

[39] Akihiko Kuze et al. “Update on GOSAT TANSO-FTS performance, operations, and data products after more than 6 years in space”. In: *Atmospheric Measurement Techniques* 9.6 (2016), pp. 2445–2461.

[40] Robert Parker et al. “Methane observations from the Greenhouse Gases Observing SATellite: Comparison to ground-based TCCON data and model calculations”. In: *Geophysical Research Letters* 38.15 (2011).

[41] A Butz et al. “TROPOMI aboard Sentinel-5 Precursor: Prospective performance of CH4 retrievals for aerosol and cirrus loaded atmospheres”. In: *Remote sensing of environment* 120 (2012), pp. 267–276.

[42] JP Veefkind et al. “TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications”. In: *Remote sensing of environment* 120 (2012), pp. 70–83.

[43] Sudhanshu Pandey et al. “Satellite observations reveal extreme methane leakage from a natural gas well blowout”. In: *Proceedings of the National Academy of Sciences* 116.52 (2019), pp. 26376–26381.

[44] Berend Schuit. “Automated detection of methane emissions in TROPOMI data using Neural Networks”. In: (2021).

[45] Christian Frankenberg et al. “Airborne methane remote measurements reveal heavy-tail flux distribution in Four Corners region”. In: *Proceedings of the national academy of sciences* 113.35 (2016), pp. 9734–9739.

[46] V Bjerknes. “Das problem der wettervorhersage, betrachtet vom standpunkte der mechanik und der physik, translation by y. mintz: The problem of weather forecasting as a problem in mechanics and physics. los angeles, 1954”. In: *Meteor. Zeit* 21 (1904), pp. 1–7.

[47] Guy P Brasseur and Daniel J Jacob. *Modeling of atmospheric chemistry*. Cambridge University Press, 2017.

[48] Barry H Lynn et al. “An examination of the impact of grid spacing on WRF simulations of wintertime precipitation in the mid-Atlantic United States”. In: *Weather and Forecasting* 35.6 (2020), pp. 2317–2343.

[49] Pedro López Gallego. “CFD study of the influence of mountain and valley shapes on wind propagation”. B.S. thesis. Universitat Politècnica de Catalunya, 2019.

[50] Huw C Davies. “Limitations of some common lateral boundary schemes used in regional NWP models”. In: *Monthly Weather Review* 111.5 (1983), pp. 1002–1012.

[51] John Lin et al. *Lagrangian modeling of the atmosphere*. John Wiley & Sons, 2013.

[52] William C Skamarock and Joseph B Klemp. “A time-split nonhydrostatic atmospheric model for weather research and forecasting applications”. In: *Journal of computational physics* 227.7 (2008), pp. 3465–3485.

[53] Jordan G Powers et al. “The weather research and forecasting model: Overview, system efforts, and future directions”. In: *Bulletin of the American Meteorological Society* 98.8 (2017), pp. 1717–1737.

[54] Hans Hersbach et al. “The ERA5 global reanalysis”. In: *Quarterly Journal of the Royal Meteorological Society* 146.730 (2020), pp. 1999–2049.

[55] National Centers for Environmental Prediction, National Weather Service, NOAA, U.S. Department of Commerce. *NCEP FNL Operational Model Global Tropospheric Analyses, continuing from July 1999*. Boulder CO, 2000. URL: `https://doi.org/10.5065/D6M043C6`.

[56] William C Skamarock et al. “A description of the advanced research WRF model version 4”. In: *National Center for Atmospheric Research: Boulder, CO, USA* (2019), p. 145.

[57] Andreas Stohl, Markus Hittenberger, and Gerhard Wotawa. “Validation of the Lagrangian particle dispersion model FLEXPART against large-scale tracer experiment data”. In: *Atmospheric Environment* 32.24 (1998), pp. 4245–4264.

[58] Andreas Stohl and Thomas Trickl. “A textbook example of long-range transport: Simultaneous observation of ozone maxima of stratospheric and North American origin in the free troposphere over Europe”. In: *Journal of Geophysical Research: Atmospheres* 104.D23 (1999), pp. 30445–30462.

[59] Caroline Forster et al. “Transport of boreal forest fire emissions from Canada to Europe”. In: *Journal of Geophysical Research: Atmospheres* 106.D19 (2001), pp. 22887–22906.

[60] N Spichtinger et al. “Satellite detection of a continental-scale plume of nitrogen oxides from boreal forest fires”. In: *Geophysical Research Letters* 28.24 (2001), pp. 4579–4582.

[61] Ignacio Pisso et al. “The Lagrangian particle dispersion model FLEXPART version 10.4”. In: *Geoscientific Model Development* 12.12 (2019), pp. 4955–4997.

[62] F Mesinger. “A blocking technique for representation of mountains in atmospheric models”. In: *Rivista di Meteorologia Aeronautica* 44.1-4 (1984), pp. 195–202.

[63] Haili Hu et al. “The operational methane retrieval algorithm for TROPOMI”. In: *Atmospheric Measurement Techniques* 9.11 (2016), pp. 5423–5440.

[64] Joannes D Maasakkers et al. “Reconstructing and quantifying methane emissions from the full duration of a 38-day natural gas well blowout using space-based observations”. In: *Remote Sensing of Environment* 270 (2022), p. 112755.

[65] P Seibert and A Frank. “Source-receptor matrix calculation with a Lagrangian particle dispersion model in backward mode”. In: *Atmospheric Chemistry and Physics* 4.1 (2004), pp. 51–63.

[66] Andreas Stohl et al. “A backward modeling study of intercontinental pollution transport using aircraft measurements”. In: *Journal of Geophysical Research: Atmospheres* 108.D12 (2003).

[67] URL: `https://unfccc.int/process-and-meetings/transparency-and-reporting/reporting-and-review-under-the-convention/greenhouse-gas-inventories-annex-i-parties/submissions/national-inventory-submissions-2017`.

A

Appendix

A.1. *flex extract* settings to retrieve meteorological data

Table A.1: Typical flex extract file settings used in this study to retrieve the ECMWF meteorological data.

Option	Value	Description
DAY1	20200101	Start date to fetch
DAY2	20200102	End date to fetch
DATE_CHUNK	1	Chunk of days to be fetched together
JOB_CHUNK	1	Number of jobs to be done together
DTIME	1	Time resolution
TYPE	AN AN ..	Analysis or forecast
TIME	00 01 ..	Time declaration
STEP	00 00 ..	Intermediate time steps
ACCTYPE	FC	Field type
ACCTIME	06/18	Forecast starting time
ACCMAXSTEP	12	Maximum forecast step
M_GRID	1000	Grid resolution*1000
M_LEFT	-170000	Left grid boundary
M_LOWER	-80000	Lower grid boundary
M_UPPER	80000	Upper grid boundary
M_RIGHT	170000	Right grid boundary
M_LEVELIST	1/to/137	Vertical hybrid sigma pressure levels
M_GAUSS	0	To retrieve divergence of wind field
M_ETA	1	To retrieve horizontal wind fields etadot
M_FORMAT	GRIB2	Format of the meteo data
M_ADDPAR	/186/187/188/235/139/39	Additional parameters to be retrieved
INPUTDIR	wind_data/2020_01_01_ECMWF_highres	Input directory
OUTPUTDIR	wind_data/2020_01_01_ECMWF_highres	Output directory
PREFIX	EA	Prefix for the output meteo files

A.2. Locations and TROPOMI orbits used for generating synthetic plume

A.2.1. Analysis of mass balance methods

Table A.2: Emission locations of synthetic plumes over Algeria that are used for the analysis of mass balance methods. Please note that multiple plumes are generated over these regions for the year 2020.

Region	Lat	Lon
Algeria	5.15	30.7
Algeria	6.15	30.7
Algeria	7.15	30.7
Algeria	5.15	31.7
Algeria	6.15	31.7
Algeria	7.15	31.7
Algeria	5.15	32.7
Algeria	6.15	32.7
Algeria	7.15	32.7

Table A.3: Emission locations of synthetic plumes over Sasol that are used for the analysis of mass balance methods. Please note that multiple plumes are generated over these regions for the year 2020.

Region	Lat	Lon
Sasol	28	-27.5
Sasol	29	-27.5
Sasol	30	-27.5
Sasol	28	-26.5
Sasol	29	-26.5
Sasol	30	-26.5
Sasol	28	-25.5
Sasol	29	-25.5
Sasol	30	-25.5

Table A.4: TROPOMI orbits used for generating synthetic plumes over Algeria for the analysis of mass balance methods.

Year	Month	Date	Orbit
2020	1	3	11522
2020	1	4	11536
2020	1	5	11550
2020	1	28	11876
2020	1	29	11891
2020	1	30	11905
2020	1	31	11919
2020	2	1	11933
2020	2	2	11947
2020	1	8	11593
2020	12	23	16558
2020	12	26	16601

Table A.5: TROPOMI orbits used for generating synthetic plumes over Sasol for the analysis of mass balance methods.

Year	Month	Date	Orbit
2020	4	21	13068
2020	5	3	13238
2020	5	12	13366
2020	5	18	13451
2020	5	23	13522
2020	5	28	13593
2020	6	2	13664
2020	6	3	13678
2020	6	4	13692

A.2.2. Analysis of atmospheric transport models

Table A.6: Emission locations of synthetic plumes over Sasol that are used for the analysis of atmospheric transport model based plume emission quantification approach. Please note that multiple plumes are generated over these regions for the year 2020.

Region	Lat	Lon
Algeria	6.1810	31.7541
Algeria	5.9651	31.6163
Algeria	5.9921	31.7771
Algeria	5.9651	31.7312
Algeria	5.9381	31.7541
Algeria	6.1810	31.8688
Algeria	5.9651	31.1556
Algeria	5.9111	31.6623
Algeria	5.9111	31.7541

TROPOMI orbits: 11493, 11508, 11522, 11536, 11550, 11564, 11578, 11593, 11749, 11763, 11777, 11791, 11805, 11820, 11848, 11862, 11876, 11891, 11905, 11919, 11933, 11947, 11962, 11976, 11990, 12004, 12018, 12047, 12075, 12118, 12132, 12146, 12160, 12174, 12189, 12203, 12217, 12231, 12245, 12259, 12288, 12302, 12316, 12330, 12345, 12359, 12387, 12416, 12430, 12444, 12458, 12472, 12486, 12501, 12515, 12529, 12643, 12671, 12699, 12728, 12756, 12770, 12784, 12799, 12813, 12827, 12841, 12855, 12884, 12997, 13011, 13068, 13111, 13139, 13153, 13182, 13196, 13224, 13238, 13253, 13267, 13338, 13352, 13451, 13465, 13480, 13494, 13508, 13579, 13593, 13636, 13664, 13692, 13707, 13735, 13820, 13834, 13848, 13863, 13877, 13919, 13934, 13948, 13962, 13976, 13990, 14019, 14033, 14075, 14090, 14104, 14132, 14146, 14175, 14203, 14217, 14232, 14246, 14260, 14288, 14345, 14359, 14373, 14388, 14402, 14444, 14459, 14473, 14487, 14501, 14529, 14544, 14558, 14572, 14586, 14600, 14615, 14629, 14643, 14657, 14671, 14686, 14700, 14714, 14728, 14742, 14756, 14771, 14785, 14799, 14813, 14827, 14842, 14856, 14870, 14884, 14898, 14913, 14927, 14984, 14998, 15012, 15026, 15040, 15054, 15097, 15125, 15154, 15182, 15196, 15210, 15225, 15239, 15281, 15296, 15310, 15324, 15338, 15352, 15452, 15466, 15480, 15494, 15508, 15523, 15537, 15551, 15579, 15594, 15608, 15622, 15636, 15650, 15664, 15679, 15693, 15721, 15735

A.3. Sanity check on FLEXPART outputs

We performed a sanity check on the output of FLEXPART. In this check, we calculated the total amount of methane mass present in the output grid after each time step of FLEXPART model run and compared it to the given emission rate. In order to calculate the total amount of methane present in the output grid, we used the following equations.

If the output is in [pptv]

$$\Omega_{CH_4} = \sum_{j=1}^{m}\sum_{i=1}^{n}\left(XCH4[pptv]_i * dp_i * A_j * \frac{M_{CH_4}}{g * M_{air}}\right) \tag{A.1}$$

If the output is in [ng/m^3]

$$\Omega_{CH_4} = \sum_{j=1}^{m}\sum_{i=1}^{n}\left(XCH4[ng/m^3]_i * dh_i * A_j\right) \tag{A.2}$$

where:
- i = certain vertical level
- n = highest vertical level
- j = certain horizontal grid point
- m = last horizontal grid point
- M_{CH_4} = Molar mass of methane
- M_{air} = Molar mass of air
- g = gravitational acceleration constant
- $XCH4_i$ = methane concentration or dry-air mole fraction
- dp_i = pressure difference between vertical levels i and $i-1$
- dh_i = height difference between vertical levels i and $i-1$
- Ω_{CH_4} = total mass of methane

Figure A.1 shows total amount of methane present in the output grid for plumes generated using ECMWF and NCEP meteorological data at several time steps. It can be seen that till 20 hours ECMWF and NCEP results were in agreement with the emission rate line. After that time step, ECMWF methane plume starts to go outside the computational domain, and hence there is a dip in the total amount of methane. After this dip, the ECMWF plume attains equilibrium where the amount of methane entering the domain is equal to the amount of methane removed from the domain. For the NCEP methane plumes, a dip in total mass can be observed around 50 hours, where it seems like a large chunk of NCEP methane plume starts going outside the computational domain. The NCEP plume does not attain equilibrium till 60 hours.

The analysis presented in the previous paragraph indicates that the output grid concentrations of FLEXPART agrees well with the input emission rate.

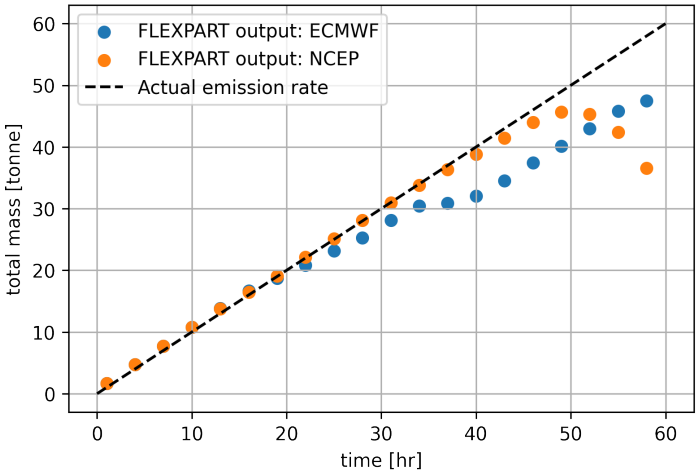


Figure A.1: This figure shows the calculated total mass of methane inside the three dimensional output grid obtained from the FLEXPART model at several time steps. Curves corresponding to plumes generated with two different meteorological datasets are presented in this figure. The blue curve corresponds to the ECMWF plume, and the orange curve corresponds to the NCEP plume. For comparison with the emission rate provided as input to the model, a black dotted line is given in the figure.

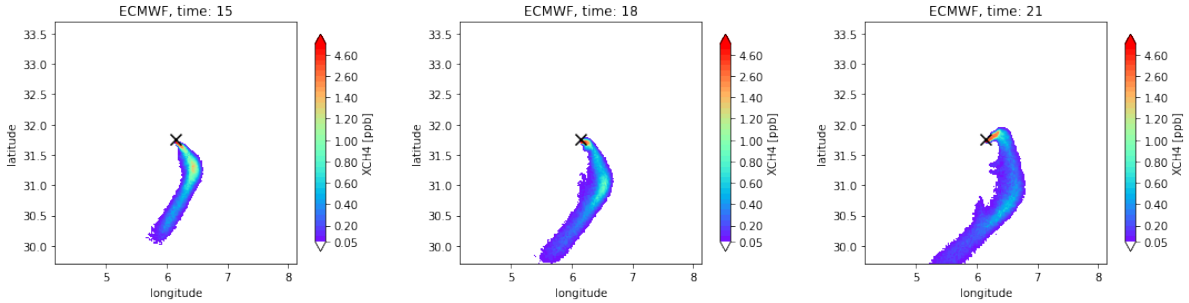


Figure A.2: This figure shows pressure weighted vertical average of three dimensional methane concentrations obtained from the FLEXPART model run that used ECMWF meteorological data. Results after 15, 18, and 21 hours are shown in this image. Around 18 hours, the plume starts going outside the computational domain, which can be seen in form of a dip in ECMWF curve shown in Figure A.1. Please note that this is not a TROPOMI resampled plume and these results are just used for verification of the FLEXPART model outputs.

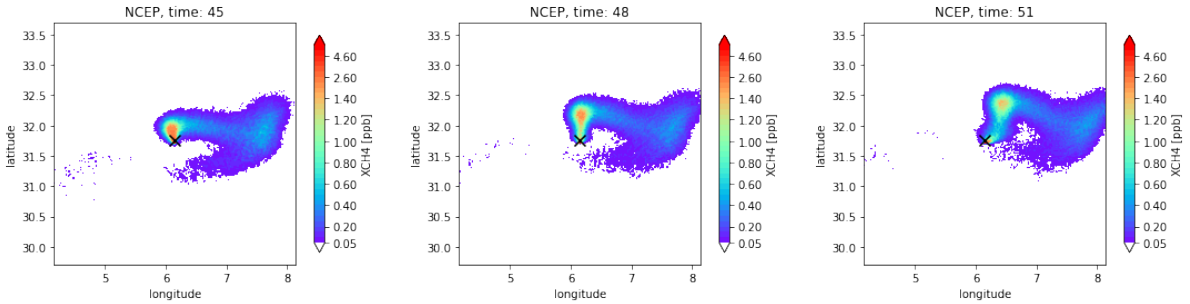


Figure A.3: This figure shows pressure weighted vertical average of three dimensional methane concentrations obtained from the FLEXPART model run that used NCEP meteorological data. Results after 45, 48, and 51 hours are shown in this image. Around 48 hours, the plume starts going outside the computational domain, which can be seen in form of a dip in NCEP curve shown in Figure A.1. Please note that this is not a TROPOMI resampled plume and these results are just used for verification of the FLEXPART model outputs.

A.4. Phase I: analysis of mass balance methods

A.4.1. Bias plots

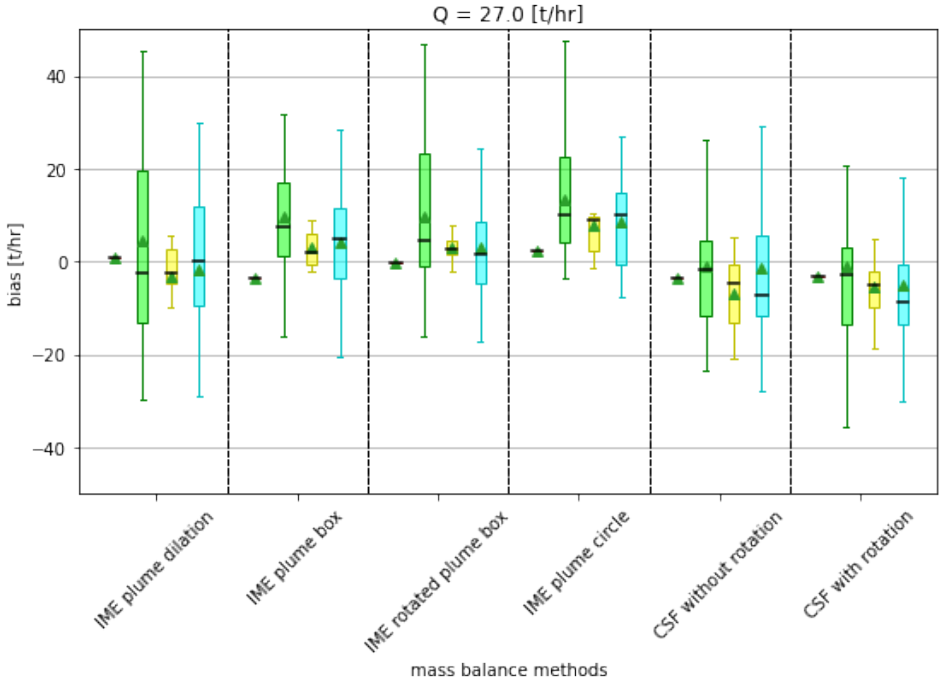


Figure A.4: This figure shows several box plots containing bias corresponding to mass balance quantification of plumes with a known emission rate of 27 [t/hr] generated over Algeria region using the ECMWF meteorological data. Each box plot represents biases calculated for plumes belonging to a particular class of plumes. Each class of plume is assigned a colour. The colours corresponding to classes of plumes are mentioned in the box at the top of the plot. We are quantifying plumes using four variations of the IME method and two variations of the CSF method, which are separated by vertical dotted lines in this plot (for variations of IME and CSF methods, refer to Section 5.1). For each box plot, a median of data is represented by a bold black line, the mean of data is represented by a green triangle, the box contains data between 25th percentile and 75th percentile, which is also called interquartile range, and whiskers are placed at 1.5 times interquartile range from the ends of the box.

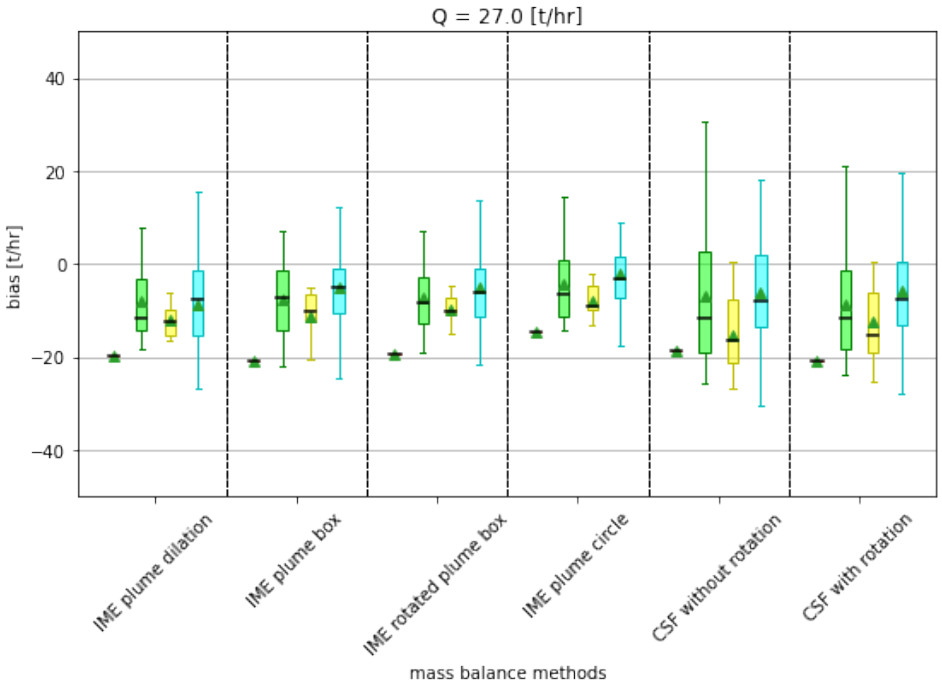


Figure A.5: This figure shows several box plots containing bias corresponding to mass balance quantification of plumes with known emission rate of 27 [t/hr] generated over Algeria region using the NCEP meteorological data. Rest of the description of this figure is same as Figure A.4.

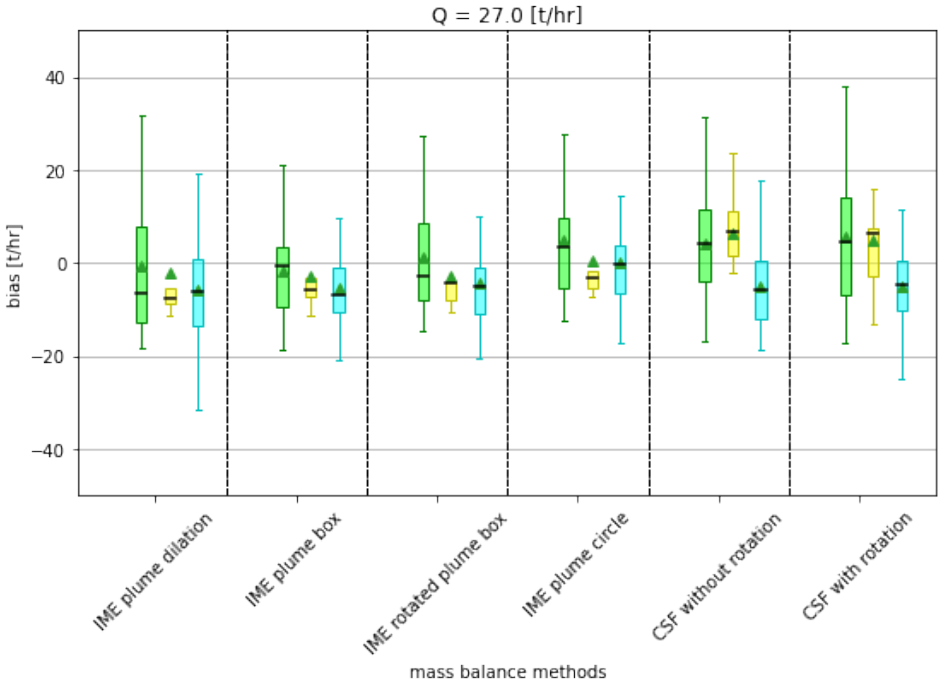


Figure A.6: This figure shows several box plots containing bias corresponding to mass balance quantification of plumes with known emission rate of 27 [t/hr] generated over Sasol region using the NCEP meteorological data. Rest of the description of this figure is same as Figure A.4.

A.4.2. RMSE plots

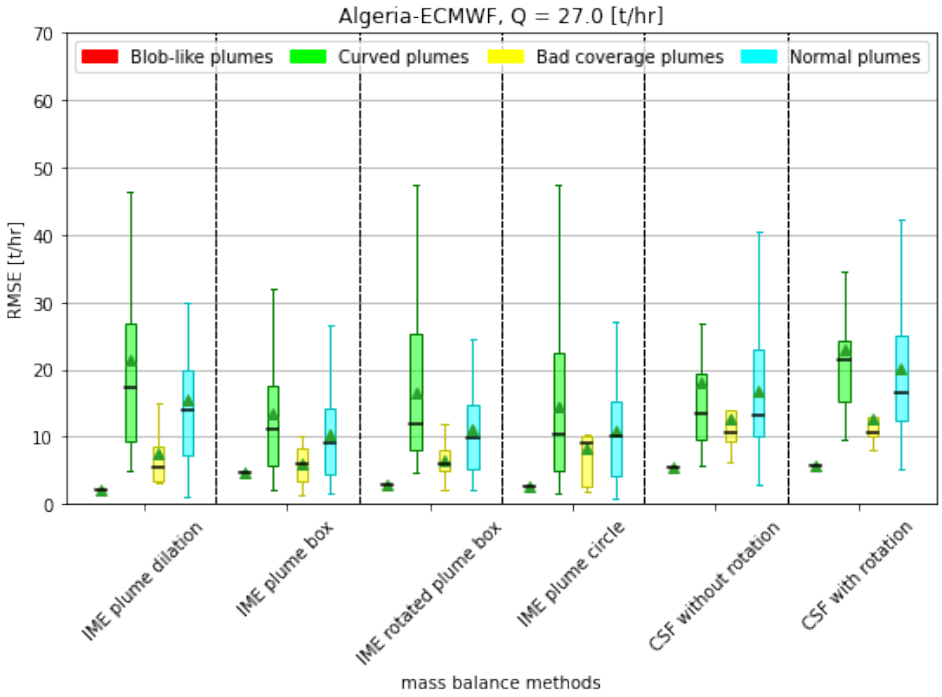


Figure A.7: This figure shows several box plots containing RMSE corresponding to mass balance quantification of plumes with known emission rate of 27 [t/hr] generated over Algeria region using the ECMWF meteorological data. Rest of the description of this figure is same as Figure A.4.

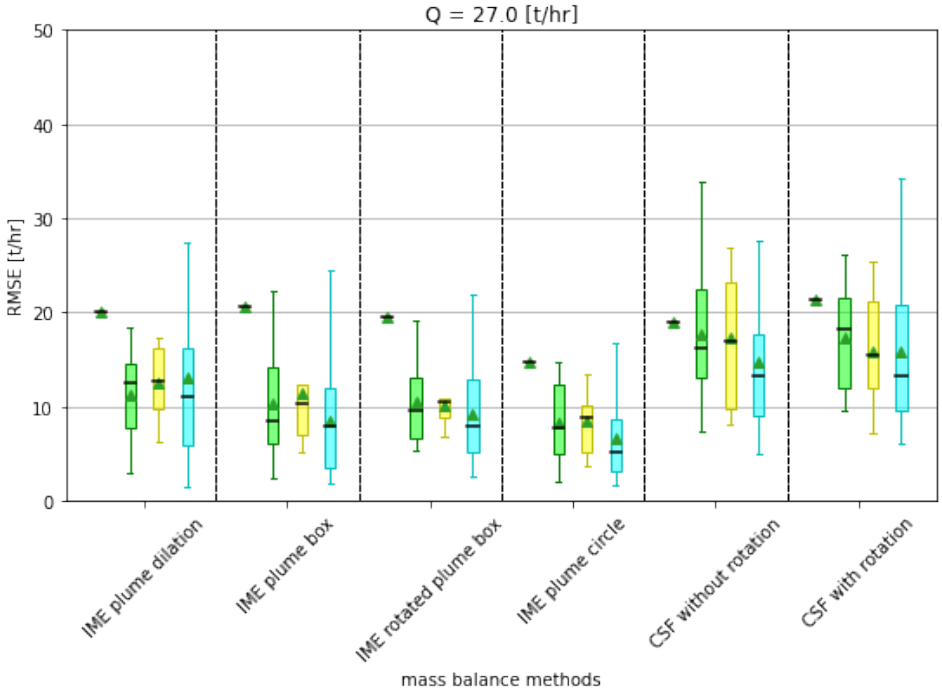


Figure A.8: This figure shows several box plots containing RMSE corresponding to mass balance quantification of plumes with known emission rate of 27 [t/hr] generated over Algeria region using the NCEP meteorological data. Each box plot represents biases calculated for plumes belonging to a particular class of plumes. Rest of the description of this figure is same as Figure A.4.

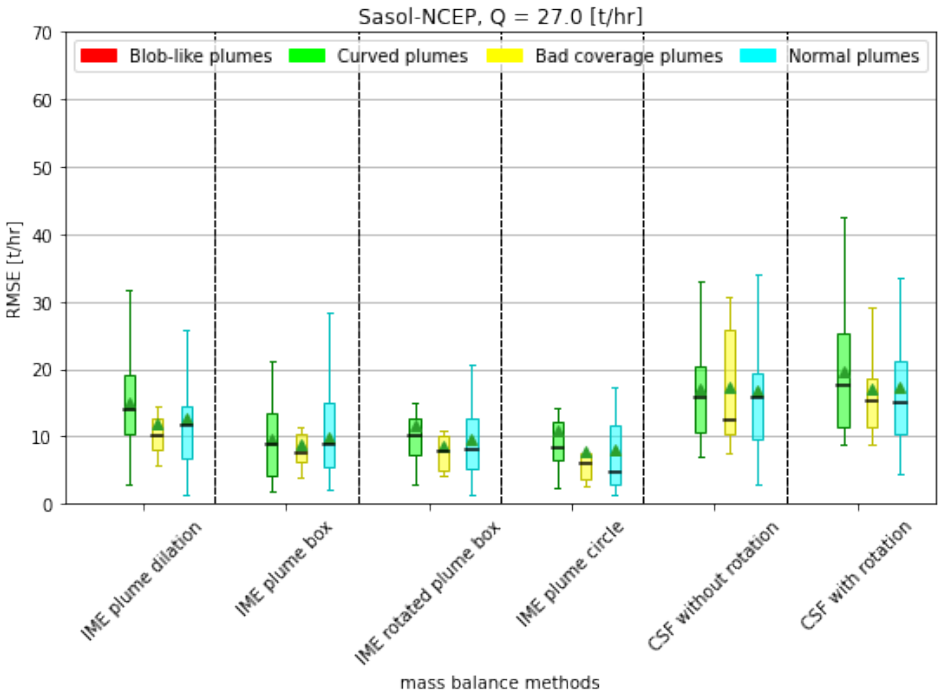


Figure A.9: This figure shows several box plots containing RMSE corresponding to mass balance quantification of plumes with known emission rate of 27 [t/hr] generated over Sasol region using the NCEP meteorological data. Each box plot represents biases calculated for plumes belonging to a particular class of plumes. Rest of the description of this figure is same as Figure A.4.