

A Spatiotemporal Analysis of Methane Emissions in South Africa Using Observations of Sentinel-5P's TROPospheric Monitoring Instrument

A Dissertation

Submitted in Partial Fulfilment of Requirements for the Degree of

Master of Research (M.Res.) in Geospatial Data Science

Keneuoe Maliehe

20395575

Supervision:

Prof Stuart Marsh

Dr Salim Alam



**University of
Nottingham**

UK | CHINA | MALAYSIA

CDT in Geospatial Systems

University of Nottingham

26 August 2022

Abstract

Methane is a potent greenhouse gas emitted into the atmosphere by anthropogenic (60%) and biological (40%) sources. Its growth is attributed to the industrial revolution, with sectors such as energy production, agriculture and waste treatment taking the lead as emitters. South Africa is committed to monitoring its growth to lessen the effects of climate change on the globe. The study identified methane (CH_4) emission hotspots over South Africa from space-based solar backscatter measurements using observations of Tropospheric Monitoring Instrument (TROPOMI) and compared observed concentrations to surface in-situ data and to an Emissions Database for Global Atmospheric Research (EDGAR) database. Even though no statistical correlation was found between the space-based observations and bottom-up inventory, correlations could be identified visually. Weak positive correlation exists between the space-based observations and surface observations. Monthly CH_4 total-averaged dry-air mole fraction (X_{CH_4}) were predicted for the year 2022 using seven statistical models based on a time series of three and half years. The predictions were compared to actual monthly X_{CH_4} for 2022 evaluated using root mean square error (RMSE) and mean absolute percentage error (MAPE) performance metrics. The Holt-Winters's additive (HWA) model performed best with a RMSE of 4.95 and MAPE of 25% due to its capability to capture both the trend and seasonality components of the data well. The study demonstrated the capability of TROPOMI for the estimation of CH_4 concentrations in the atmosphere and for identification of trend patterns along both spatial and temporal profiles.

Acknowledgements

First and foremost I would like to acknowledge my funders, the United Kingdom Research and Innovation Engineering and Physical Sciences Research Council and the University of Nottingham without whose financial contributions, this entire project would not have been possible.

The sincerest gratitude goes out to my supervisors, Stuart Marsh and Salim Alam for your guidance, time, insights and encouragement. Your knowledge and support was instrumental in the delivery of this project and I am grateful. I look forward to continuing my PhD research with you.

To my partner for your constant support and for being there with me in every way through the hardest thing I ever had to do. You watched me go through so much, accepted the new me and all the limitations that she came with, and you ensured me that I was still good enough. Nathi, for the inspiration. Just when I thought it was enough, you showed me that the sky is the limit. Your persistence everyday shows me that nothing is impossible. My life-size *popi*.

My family in Lesotho and South Africa. Ellen, my migration and this dissertation would not have been possible without your support, emotionally, physically and financially. From uplifting me when I had hit rock-bottom, letting me stay at your house indefinitely during my would-not-have-liked-to-be-so-frequent trips home during the course of the year, I cannot thank you enough. I have not known such kindness and selflessness. Nkhono Thabi for taking such good care of my child during the few months that I had to sacrifice him. thank you for loving him like your own. And Nkhono Tuntu for taking care of all the administrative affairs pertaining to “my stuff”. I don’t thank you enough. And for cheering me on this year, reminding me always of the commitment and the extraordinary in me, I am grateful. Shoeshoe, thank you for always having my back, for being there at the airport with us, the most memorable stopovers and for being Nathi’s favourite aunt. South Africa and OR Tambo would not be the same without you.

Larissa, thank you for showing up every time, no matter the task. I cannot believe this year is almost over and how far we have both come from those hikers sitting at the top of Faerie Glen Nature Reserve looking down at Pretoria and our lives. May we never look back. Thank you for giving me my own personal ArcGIS Pro classes on WhatsApp.

And finally, I am grateful to my colleagues and mentors in the CDT. A special mention to Jamie, thank you for taking care of literally everything.

Contents

List of Acronyms	10
1 Introduction	14
1.1 Research Problem Statement	17
1.2 Research Aim and Objectives	19
2 Theoretical Background	21
2.1 Introduction	21
2.2 The Atmosphere and its Constituents	21
2.3 The Greenhouse Effect	22
2.4 Climate Change	24
2.5 Methane	26
2.6 Satellite Observations of Atmospheric Methane	27
3 Literature Review	29
3.1 Introduction	29
3.2 Understanding the Global Methane Budget	29
3.3 Space-based Observations of Methane	30
3.4 Applications of Space-based Observations	32
3.5 Emission Inventories	34
3.6 Forecasting Models for Methane Emissions	35
3.7 Research Gaps	37
4 Research Design	39
4.1 Introduction	39
4.2 Description of Study Area	39
4.3 Resources and Tools	41
4.3.1 Satellite data	41
4.3.2 Bottom-up inventory	42
4.3.3 Surface measurements	43
4.3.4 Computing infrastructure	45
4.4 Method Design	45
4.4.1 Automated satellite data download and access	45
4.4.2 Image processing and spatial analysis	46
4.4.3 Temporal analysis	50
4.4.4 Validation	55

5	Results	56
5.1	Introduction	56
5.2	Spatial Analysis	57
5.3	Time Series Analysis and Forecasting	61
5.4	Validation	64
6	Discussion	74
7	Conclusions and Recommendations	76

List of Figures

1	Layers of the atmosphere, retrieved from McElroy (2012)	23
2	The Earth’s energy balance, retrieved from Seinfeld (2016)	24
3	Globally averaged atmospheric CH_4 in <i>ppb</i> and its annual growth rate, as estimated by Saunio et al. (2020). It shows the decline in emissions in the 1990s, the rapid growth in 2007 and other significant changes in the growth rate of CH_4	30
4	A summary of space-based instruments measuring CH_4 onboard completed and ongoing low earth orbiting (LEO) satellites with the length of missions. .	32
5	Map of South Africa showing the location of its nine administrative areas . .	39
6	Mean emissions of CH_4 over one year, 2018, reported in EDGAR v6.0, for all source types	44
7	Concentrations of atmospheric column-averaged dry-air mole fractions of CH_4 in South Africa from May 2018 to July 2019	65
8	Concentrations of atmospheric column-averaged dry-air mole fractions of CH_4 in South Africa from August 2019 to October 2020	66
9	Concentrations of atmospheric column-averaged dry-air mole fractions of CH_4 in South Africa from November 2020 to December 2021	67
10	The overlay process highlighting that the areas of higher building density and industrial activity are in areas of higher CH_4 emissions.	68
11	Monthly time-series analysis from May 2018 to July 2022 shows a steady increase of CH_4 emissions over the country in the four year period.	69
12	Seasonal decomposition of the data shows the data exhibits both trend and seasonality	70
13	Graphical representations of mean monthly X_{CH_4} forecast using the baseline model, the naive model, the simple moving average (SMA) and the simple exponential smoothing (SES) models show the poor performance of the models where the monthly X_{CH_4} have been severely underestimated by the models. .	71
14	Graphical representations of mean monthly X_{CH_4} forecast using the HWA, Holt-Winters’s multiplicative (HWM), auto-regressive integrated moving average (ARIMA) and seasonal auto-regressive integrated moving average (SARIMA) models show the better performance of the models.	72
15	Correlation between the EDGAR inventory (a) and TROPOMI (b) averaged observations for May 2018.	73

16	Line plot of TROPOMI measurements being compared to Global Atmosphere Watch (GAW) and National Oceanic and Atmospheric Administration (NOAA) surface measurements of CH_4	73
----	---	----

List of Tables

1	Mixing ratios of gases in the atmosphere	22
2	Residence times and global warming potential of Kyoto greenhouse gases in the atmosphere	25
3	Methane emissions source types reported by the bottom-up EDGAR anthropogenic inventory with their corresponding International Panel on Climate Change (IPCC) codes	43
4	Description of substrings in the TROPOMI product name	47
5	Tests Statistics for the Augmented Dickey-Fuller (ADF) and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) tests on the time-series data	62
6	Forecasting accuracy metrics for the evolution of X_{CH_4} in South Africa in 2022	63
7	X_{CH_4} emissions data forecast for 2022 using the HWA model. The forecasts are presented against actual values from the TROPOMI X_{CH_4} values obtained for January to July 2022	63

List of Scripts

1	Pre- and post-processing of the Sentinel-5 Precursor (S-5P) images using <code>harp</code> atmospheric toolbox	47
2	<code>bin_spatial</code> parameters	49
3	<code>bin_spatial</code> parameter calculations	49

List of Acronyms

CH_4 methane 1, 6, 7, 14–34, 36–44, 51, 56–61, 63–69, 73–77

CO carbon monoxide 15, 42

CO_2 carbon dioxide 14, 15, 18, 21–27, 35, 36, 41, 76

O_2 oxygen 21–23

X_{CH_4} CH_4 total-averaged dry-air mole fraction 1, 6, 8, 27, 48–52, 58, 61–63, 71, 72, 77

Ar argon 21, 22

H_2 hydrogen 22

He helium 22

Kr krypton 22

N_2 nitrogen 21–23

N_2O nitrous oxide 15, 22, 23, 25, 26

Ne neon 22

OH hydroxyl radical 22, 26, 38

O_3 ozone 22, 23

SF_6 sulphur hexafluoride 25, 26

ppb parts per billion 14

ppm parts per million 22

ADF Augmented Dickey-Fuller 8, 50, 51, 61, 62

AFOLU agriculture, forestry, and other land-use 40

AIRS Atmospheric Infrared Sounder 31

ARIMA auto-regressive integrated moving average 6, 35, 36, 53, 54, 63, 72

ARMA auto-regressive moving average 54, 62

C Celsius 23, 24

CarbonSat Carbon Monitoring Satellite 31

CEDS Community Emissions Data System 34

CFC chlorofluorocarbon 23, 25

COP Conference of the Parties 41

COVID-19 coronavirus disease 2019 44, 58

CrIS Cross-Track Infrared Sounder 31

CTM chemical transport model 27, 32, 33, 74

DEAT Department of Environmental Affairs and Tourism 16

DFFE Department of Forestry, Fisheries, and the Environment 15, 41

EDGAR Emissions Database for Global Atmospheric Research 1, 6, 8, 19, 33, 34, 41–43, 49, 56, 64, 73, 77

ENVISAT Environmental Satellite 30

ESA European Space Agency 30, 46

ESRL Earth System Research Laboratory 43, 44

EU European Union 17

G3E Geostationary Emission Explorer for Europe 31, 32, 37

GAUGE Green house gAs UK and Global Emission 38

GAW Global Atmosphere Watch 7, 38, 41, 44, 64, 73

GDP gross domestic product 40

GEO-CAPE Geostationary Coastal and Air Pollution Events 31, 32

GeoCARB Geostationary Carbon Observatory 31, 32

GeoFTS Geostationary Fourier Transform Spectrometer 31

GMD Global Monitoring Division 43

GOSAT Greenhouse Gases Observing Satellite 27, 30, 33, 34

HFC hydrofluorocarbon 25

HWA Holt-Winters’s additive 1, 6, 8, 36, 63, 72, 77

HWM Holt-Winters’s multiplicative 6, 36, 63, 72

IASI Infrared Atmospheric Sounding Interferometer 31

IMG Interferometer Monitor for Greenhouse Gases 31

IPCC International Panel on Climate Change 8, 15, 17, 18, 41, 43

K Kelvin 22

KPSS Kwiatkowski-Phillips-Schmidt-Shin 8, 50, 51, 61, 62

L1 level 1 42

L2 level 2 41

L3 level 3 50

LEO low earth orbiting 6, 31, 32

LiDAR Light Detection and Ranging 31

MAPE mean absolute percentage error 1, 36, 54, 55, 62, 63

MASE mean absolute scaled error 35

MERLIN MEthane Remote sensing Lidar missioN 31

NAEIS National Atmospheric Emissions Inventory System 15

NDC nationally determined contribution 18–20, 40, 41, 56

NNAR neural network auto-regression 35, 36

NOAA National Oceanic and Atmospheric Administration 7, 29, 30, 38, 41, 43, 44, 64, 73

OSM Open Street Map 41, 50

PFC perfluorocarbon 25

RMSE root mean square error 1, 35, 36, 54, 62, 63

S-5P Sentinel-5 Precursor 9, 16, 19, 21, 31, 41, 48, 56

SARIMA seasonal auto-regressive integrated moving average 6, 35, 37, 54, 63, 72

SAWS South African Weather Services 38, 43, 44

SCIAMACHY SCanning Imaging Absorption spectroMeter for Atmospheric CHartogra-
phY 27, 30, 33, 34

SES simple exponential smoothing 6, 37, 62, 71

SLCP short-lived climate pollutant 26

SMA simple moving average 6, 36, 52, 62, 71

SRON Netherlands Institute for Space Research 42

SWIR shortwave infrared 27, 32

TANSO-FTS Thermal And Near infrared Sensor for carbon Observation - Fourier Transform Spectrometer 30

TCCON total carbon column observing network 33

TES Tropospheric Emission Spectrometer 31

TIR thermal infrared 27, 31, 32

TROPOMI TROPOspheric Monitoring Instrument 1, 6–8, 16, 17, 19–21, 31–34, 37, 41, 42, 46–48, 52, 56, 57, 61, 63, 64, 73–77

UK United Kingdom 17, 38

UNEP United Nations Environment Programme 25

UNFCCC United Nations Framework Convention on Climate Change 14, 17, 18, 21, 25, 26, 41, 43

USA United States of America 17, 33, 34, 36, 74

WFM-DOAS weighting function modified differential optical absorption spectroscopy 42

WGS84 World Geodetic System 1984 50

WMO World Meteorological Organization 38, 44

WRF-Chem weather research and forecasting coupled with chemistry 34

1 Introduction

CH_4 is a potent greenhouse gas emitted by anthropogenic (60%) and natural (40%) sources. Its concentration in the atmosphere has risen from a pre-industrial value of 722 parts per billion (*ppb*) to a current global average of 1,850 *ppb* (de Gouw et al., 2020). Globally, its growth is attributed to the industrial revolution, with sector such as energy production, agriculture and waste treatment taking the lead as culprits. In South Africa, the key drivers to a growth in emissions have been identified as population growth and development activities (Shikwambana et al., 2021a). Due to its global warming potential which is 25 times that of carbon dioxide (CO_2) and its shorter lifetime in the atmosphere, measuring and monitoring it is crucial for mitigating the impacts of climate change.

CH_4 emissions in the atmosphere have both consequences for climate change and air quality. It is well known that exposure to high concentrations of CH_4 can lead to respiratory problems (Nesme et al., 2021). In many parts of the developed world such as Europe, north America and other highly industrialised parts of the world, climate change policies have become more important than air pollution policies. This is different in parts of the world with mega-cities with fast-growing economies where air pollution is extremely toxic and therefore the protection of human life through the enforcement of air quality policies is more urgent than the subject of climate change. Even though global policies for climate change and local or regional policies for air quality are viewed differently in different parts of the world depending on different perspectives, it is important to note that the sources of air pollutants and greenhouse gases are the same. As thus there is a need for combined strategies to address the challenges of both climate change and air pollution.

Therefore, South Africa has committed to reducing its national emissions and to take part in a global stocktaking of greenhouse gases exercises which is scheduled to take place every five years. The global stocktake has been mandated by the United Nations Framework Convention on Climate Change (UNFCCC) in efforts to support the Paris Agreement (Bergamaschi et al., 2018). The country has several bottom-up greenhouse gas inventories, which include CH_4 .

There are, however, several problems with them. Seymore et al. (2014), even though now dated, analysed and summarised South Africa’s greenhouse gas inventories at the time and found that the country had several, dating back to 1988. The inventories came from six sources and have been summarised below:

- The inventory of Scholes and van der Merwe (1996) for 1988 despite being significantly outdated was quite comprehensive, covering CO_2 , CH_4 , nitrous oxide (N_2O), NO_x and carbon monoxide (CO). Seymore et al. (2014) reported that this was a good inventory which provided a useful reference point for 1988, but could however not be useful in analysis as its activity classification was inconsistent with the structure of energy balances being used in 2014.
- The inventory of Howells and Solomon (2000) for 1994 was fairly outdated as well and had the same limitations as that of Scholes and van der Merwe.
- The Department of Environmental Affairs and Tourism inventories for 1990, 1994 and 2000 were inconsistent with one-another as they utilised different IPCC guidelines, The 1990 and 1994 inventories were compiled using IPCC 1996 guidelines and the 2000 inventory using the 2000 guidelines.
- The International Energy Agency’s inventories for 1971, 1975, 1980, 1985, 1990, 1995, 2000, 2005, 2007, 2008 and 2009 were the only source of a comprehensive time-series of emission inventory, however rather limited in terms of gas composition as they only reported CO_2 .
- The Department of Energy inventories for 2007 and 2008 had the limitation of assuming nuclear and renewable energy sectors emitted zero emissions.

Seymore et al. (2014) concluded all inventories to be inconsistent, and to have major discrepancies in terms of scope of sectors and gas components, as well as methods of estimation used. It is also noteworthy that none of the inventories used space-based estimation methods.

Since then South Africa’s Department of Forestry, Fisheries, and the Environment (DFFE) developed the National Atmospheric Emissions Inventory System (NAEIS) which was de-

signed for companies to report their greenhouse gas emissions. Its key objective was to enhance the data collection process by the department for a national greenhouse gas inventory. However, due to the complexity of bottom-up estimation calculations, it was unrealistic to expect every company to have an environment scientist who would report accurate emissions, or for every company to ensure the right kind of system in place for emission measurements. This is expected to be a source of error for the emissions inventory. There was also a latency of four years, at least, for publication of national greenhouse gas inventories. The greenhouse gas inventory by the environment agency then, the Department of Environmental Affairs and Tourism (DEAT), for the period spanning 1990 to 2000 was published in 2009 (Seymore et al., 2014), the inventory for the period spanning 2000 to 2010 was published in 2014, that of 2013 to 2015 in 2019, and lastly the inventory for 2017 was published in August 2021 (Stevens, 2021). It is not rare for greenhouse gas inventory estimates to have uncertainties in emission factors for most CH_4 sources, as has been evidenced in other parts of the world. As such, it is now common practice to have inventories that are an amalgamation of bottom-up and top-down approaches. It is within this context that it becomes necessary for South Africa to look to space-based measurements for improvements and to augment what is existing. The TROPOMI onboard the S-5P satellite provides the opportunity for this as well as for applications related to quick identification of hotspots and CH_4 plumes, tracking of large-scale CH_4 transport events in the atmosphere and estimation and prediction of CH_4 emission trends in the country (Jacob et al., 2016). The spatial and temporal resolution of TROPOMI in combination, are unrivalled compared to other instruments in orbit and the primary goal of the mission has been to accurately measure concentrations of CH_4 in the atmosphere.

Satellite observations of atmospheric CH_4 began in 1996 (Jacob et al., 2016) and have continued in recent years at global, regional as well as city scales. With improving capabilities of instruments in precision, spatial and temporal resolution, satellites are significantly enhancing the information that exists within the atmospheric chemistry domain. Therefore this research attempted to quantify the concentration of CH_4 in the atmosphere in South Africa

during the years 2018 to the end of 2021 using space-based observations from TROPOMI. Further, using a time series of the observations, predicted emission concentrations for 2022. Finally, it compared the space-based observations to an existing high quality bottom-up inventory and surface flask measurements taken from South Africa's only CH_4 monitoring station in Cape Point.

1.1 Research Problem Statement

In order to properly formulate responses to control air and greenhouse pollutants, knowledge of both emission levels and their respective sources is a prerequisite. Understanding emissions is at the core of managing environmental pollution and responding to climate change. Globally, the effort to limit climate change and control emissions of greenhouse gases is mainly coordinated through the UNFCCC, whose long-term objective is to stabilise atmospheric greenhouse gas concentrations at a level that would prevent dangerous anthropogenic interference with the climate system. Through the UNFCCC, the Kyoto Protocol sets targets for many industrialised countries and limits the emission increases of Annex 1 parties. There are currently non-annex 1 countries that are producing more greenhouse gases than annex-1 countries, such as China and India. China is the single largest emitter of greenhouse gases globally, followed by the United States of America (USA), European Union (EU), India and Russia. However, the developing world sees the urgency around global warming and climate change differently and argue that “social and economic development and poverty eradication are the first and overriding priorities” (UNFCCC Secretariat, 2012) of their countries. The IPCC and G7 group of economies (Canada, France, Germany, Italy, Japan, United Kingdom (UK), and the USA) both have a goal of keeping accumulated greenhouse gas emissions below levels that will generate a maximum increase in average global temperatures of 2° above pre-industrial levels. This will require approximately a 50% reduction in global emissions to be made by the middle of the 21st century (Pleijes, 2009).

In the immediate term, the enhanced transparency framework of the Paris Agreement

requires countries to submit national sector-resolved emissions by November 2024 as a basis for setting their nationally determined contributions (NDCs). Outside of the Paris Agreement, there is a Global Methane Pledge of 2021, which has been signed by over 110 countries who are committing to reduce their collective CH_4 emissions by 30% of 2020 levels by 2030 (Jacob et al., 2022).

The reason for the recent focus on CH_4 emissions is that they have a very important effect on global radiative forcing. CH_4 's potency as a greenhouse gas and shorter lifetime relative to CO_2 and other long-living greenhouse gases (Pleijel et al., 2009) makes it a low-hanging fruit for climate change mitigation. Moreover, CH_4 has an impact on air quality and subsequently human health, and quantitative analysis of its concentration through space and time becomes critical for projecting and mitigating changes to the climate and air quality (Kort et al., 2014).

South Africa signed the Paris Agreement in 2015, which legally bound it to efforts to limit temperature rise by accounting for and reducing its greenhouse gas emissions and contributions. It has several national bottom-up greenhouse gas inventories conducted using IPCC guidelines where emissions from individual sectors are collected using source-dependent activity data and emission factors (Jacob et al., 2016). Bottom-up inventory estimates of CH_4 emissions of most countries, including South Africa, have large uncertainties in measurements not being rigorously made or tested sufficiently across different sites and discrepancies in emission factors for most CH_4 source sectors occur (Worden et al., 2021). Therefore, to complement bottom-up methods, it has been demonstrated that atmospheric transport modelling or inverse modelling can be used to check for consistency of bottom-up inventories. The scientific and technical advice body of the UNFCCC stated in 2017 its notice of the increasing capability of satellite observations to monitor greenhouse gas concentrations and emissions and therefore their relevance in supporting the objectives of the Paris Agreement (Bergamaschi et al., 2018).

Despite all the efforts, observations and monitoring techniques, forecasts indicate that the world's highest emitters will increase their greenhouse gas emissions by 3.67% by 2030 Tudor

and Sova (2021). This indicates that countries will not meet the targets their governments set out in their NDCs. This highlights the need for more impactful policies, and measures to ensure countries will reach their targets ahead of time. The production of accurate estimates of emissions and their evolution through time is critical for robust policy-making including monitoring and evaluation processes.

To support South Africa’s climate policy and its political will to contribute to fighting climate change, the study aimed to identify CH_4 emission hotspots over the country from space-based solar backscatter measurements, with the expectation that these hotspots are over areas of intense industrial activity or areas with high population densities. The hotspots were identified using daily observations of the TROPOMI onboard the S-5P satellite and averaged to weekly and monthly time intervals. A time series was also performed to investigate the presence of seasonal trends and patterns over the course of three years, from TROPOMI’s launch date to the end of 2021. And using the time series, concentrations for the first seven months of 2022 were predicted and validated using real measurements from TROPOMI. Lastly, as it is accepted that the EDGAR inventory is a valuable tool for countrywide to global estimation of greenhouse gas emissions, validation of the TROPOMI measurements was done using an EDGAR inventory, in addition to surface measurements.

1.2 Research Aim and Objectives

The aim of the research is thus to evaluate the capability of TROPOMI as a tool to aid South Africa in its climate policy and in supporting the Paris Agreement. It is to ascertain whether it can accurately measure CH_4 concentrations in its atmosphere and facilitate predictions of CH_4 concentration in the future in order to serve as a warning system for the country not reaching its emission reduction targets.

The aim is supported by the following objectives:

1. To spatially analyse trends and patterns in CH_4 emissions using satellite retrievals in order to answer the question of where the highest concentrations of CH_4 are located

within the country and to ascertain any relationships between them and events on the surface. The expectation is that high concentrations of emissions are related to a number of anthropogenic source sectors such as livestock rearing, oil and gas systems, coal mining, landfills and wastewater treatment.

2. To monitor emissions during the period 2018 to 2021 using monthly averaged data of TROPOMI. This will comprise of a time-series analysis of the average CH_4 emissions across the study area in order to test for seasonality and for the identification of a temporal trend during the period of observation. This objective will help in answering the question of how have the CH_4 emissions behaved over time. Ascertaining whether they are increasing in line with observed CH_4 emission estimates globally? And further, to determine if the rate of emissions can be characterised by annual variables and if the patterns are influenced by climatic conditions or surface conditions such as vegetation cycles.
3. To produce forecasts of CH_4 emissions by employing time-series based prediction (or forecasting) models to the space-based observations of CH_4 . This can contribute to the timely monitoring and evaluation of South Africa's progress towards the target it set out in the Paris Agreement and its NDC, and act as an early-warning system.
4. To test agreement with an existing bottom-up inventory. Even though it has been noted that discrepancies exist between bottom-up and top-down approaches of CH_4 inventories, it is expected that at the very least a correlation exists between estimates obtained by the different approaches. This objective will therefore test how the TROPOMI observations of CH_4 agree or disagree with those reported by a high quality bottom-up inventory by the European Commission.

2 Theoretical Background

2.1 Introduction

This section provides a theoretical background of the concepts that have been addressed throughout this dissertation as well as explains briefly how they are interlinked. As atmospheric emissions and concentrations of CH_4 continued to rise, CH_4 came to be the most important human-influenced (fuelled by the rise in industrial and economic activity of the planet) greenhouse gas in terms of climate forcing, after CO_2 . Terms such as the earth's energy balance, radiative forcing and greenhouse effect are introduced in attempts to illustrate how the increase of the gases and CH_4 concentrations in the atmosphere contribute to climate change. The structure of the atmosphere and its layers are introduced to show how CH_4 interacts with the atmosphere at different levels and how those affect the way that CH_4 can be measured from space using remote sensing techniques. And most specifically why the TROPOMI instrument onboard the S-5P is a superior instrument for short- to medium-term monitoring of different components of the atmosphere to aid the objective of the UNFCCC of keeping emissions low in order to reverse the effects of climate change.

2.2 The Atmosphere and its Constituents

The word atmosphere has Greek origins, '*atmos*' and '*sphaira*' which can be translated into English, '*vapour*' and '*ball*'. Indeed, atmospheres are gaseous envelopes around the planets. The atmospheres of the different planets have different compositions and constituents. Different from those of other planets, the atmosphere of planet Earth supports life and its composition is determined by the biological processes that are essential for supporting life on Earth (Wayne, 2000).

Its composition is primarily made up of gases, with nitrogen (N_2), oxygen (O_2) and argon (Ar) accounting for 78%, 21% and 0.93%, respectively. The remaining 0.07% is made up of trace gases and despite their extremely low concentrations, play a crucial role in Earth's

radiation balance in the greenhouse effect, the ozone layer and other environmental issues (Seinfeld, 2016). Table 1 summarises the concentrations of the major atmospheric gases. Concentrations of constituents of the atmosphere are given as mixing ratios, by volume, and are also referred to as mole fraction, in parts per million (*ppm*). Figure 1 present the different layers of the atmosphere with the yellow line indicating the altitude and temperatures in the respective layers. The layers of the atmosphere are important in this context as the sinks of CH_4 are defined by different oxidation processes of the hydroxyl radical (OH) in different parts of the atmosphere.

Table 1. Mixing ratios of gases in the atmosphere

Gas	Mixing ratio (mol/mol)
N_2	0.78
O_2	0.21
Ar	0.0093
CO_2	365×10^{-6}
neon (Ne)	18×10^{-6}
ozone (O_3)	$(0.01-10) \times 10^{-6}$
helium (He)	5.2×10^{-6}
CH_4	1.7×10^{-6}
krypton (Kr)	1.1×10^{-6}
hydrogen (H_2)	500×10^{-9}
N_2O	310×10^{-9}

2.3 The Greenhouse Effect

The greenhouse effect refers to the absorption and re-radiation of energy by atmospheric gases, resulting in a downward flux of infrared radiation from the atmosphere to the surface. Figure 2 shows incoming and outgoing energy fluxes from Earth on an annual-average basis. At equilibrium, the total rate at which energy leaves the Earth is 102 W/m^2 of reflected sunlight + 239 W/m^2 of infrared which equates to 341 W/m^2 , the incoming solar radiation at the top of the atmosphere, establishing the equilibrium temperature of the Earth (Jacob, 2011).

The global average surface temperature of the Earth is approximately 288 Kelvin (K)

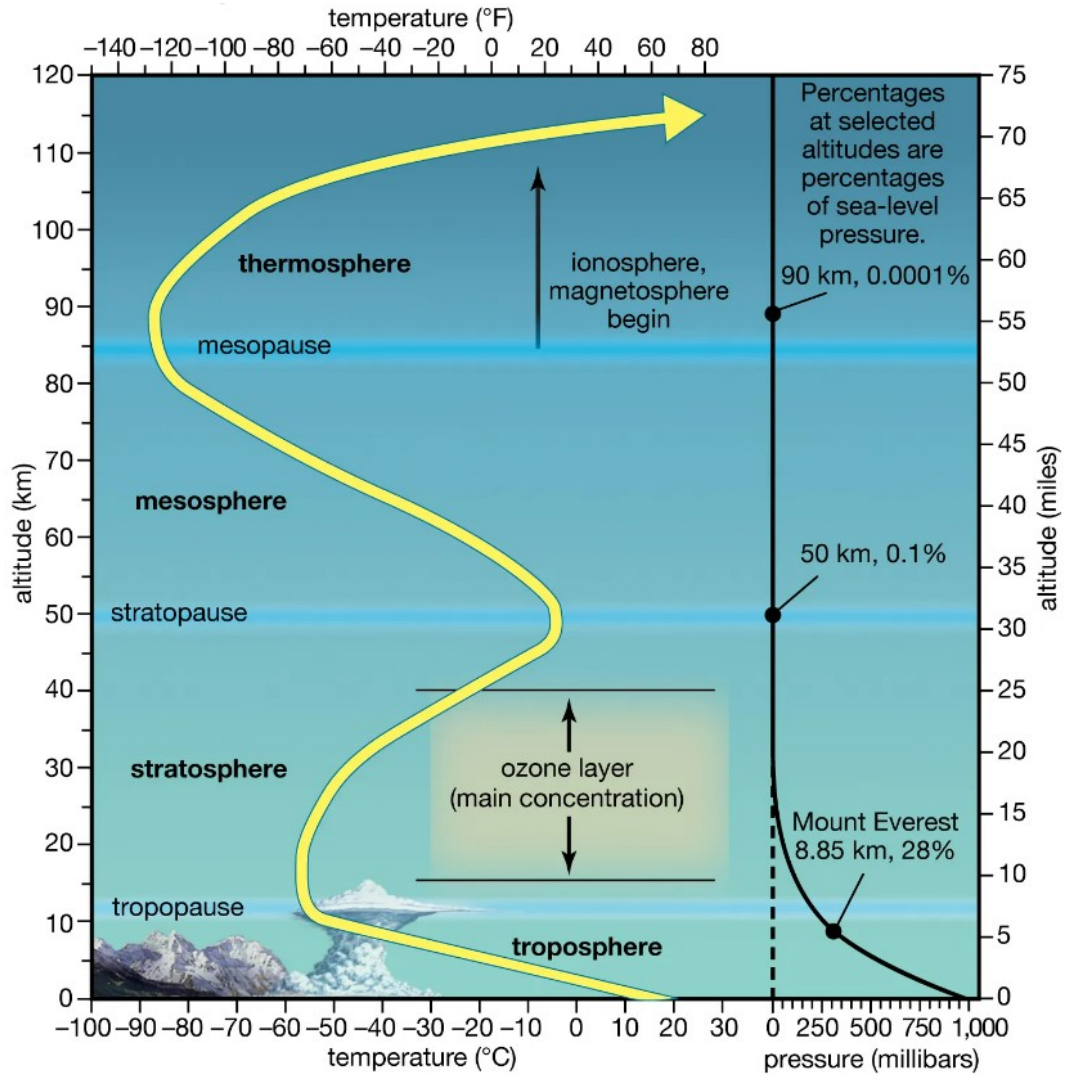


Figure 1. Layers of the atmosphere, retrieved from McElroy (2012)

or 15° Celsius (C) for which the corresponding blackbody irradiance is 396 W/m^2 . However, much of the infrared energy emitted at the surface is absorbed by molecules in the atmosphere such as CO_2 , water vapour, CH_4 , N_2O , chlorofluorocarbons (CFCs), and O_3 and re-emitted in both upwards and downward directions, further heating Earth's surface and maintaining the atmospheric temperature gradient. The Earth's energy balance is summarised graphically in Figure 2. N_2 and O_2 which account for 99% of the atmosphere's composition are transparent to infrared radiation (Wayne, 2000). It is essentially the 1% of

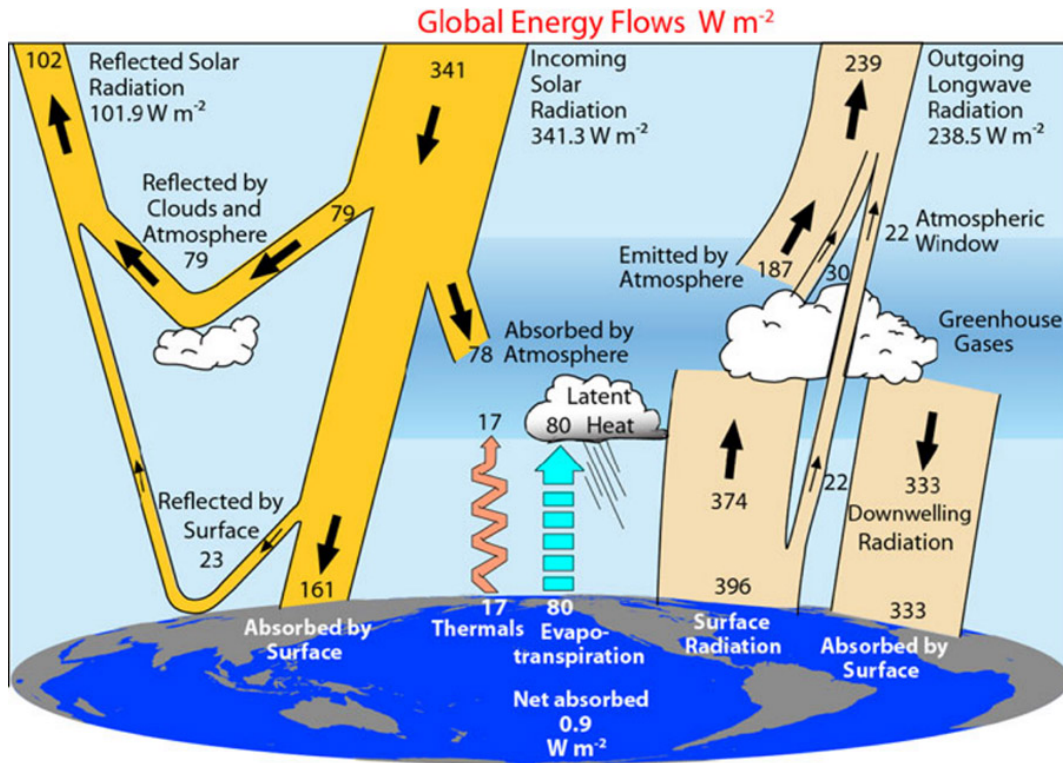


Figure 2. The Earth's energy balance, retrieved from Seinfeld (2016)

trace gases that are responsible for the absorption and remittance of infrared radiation. At equilibrium, emitted longwave flux at the top of the atmosphere is $239 W/m^2$, significantly less than emitted at the surface ($396 W/m^2$). This aids the Earth in keeping a habitable temperature of $15^\circ C$ instead of $-18^\circ C$. This natural phenomenon is known as the greenhouse effect and it is made possible by water vapor, CO_2 , CH_4 and other gases which have come to be known as the greenhouse gases. Despite their small amounts, they profoundly affect the Earth's temperature (Seinfeld, 2016).

2.4 Climate Change

Pleijel et al. (2009) states that for a long period in the establishment of atmospheric particles and aerosols as air pollutants, no or little attention was paid to greenhouse gases and climate change and very few attempts were made to control greenhouse gases until the 1990s. The

detection of the widespread destruction of ozone in the stratosphere over Antarctica is what prompted the control of ozone-depleting substances under the Montreal Protocol in the 1990s. Climate change policies over the past thirty years benefited considerably from the control of ozone-depleting substances such as CFCs and other halocarbons, hydrofluorocarbons (HFCs) and perfluorocarbons (PFCs) (Neelin, 2011).

Major steps in controlling greenhouse gases were taken in 1992 when the UNFCCC was signed and subsequently when the Kyoto Protocol was signed in 1997 (Pleijel et al., 2009). Steps were then taken to establish binding targets for emission reductions for the main long-lived greenhouse gases including CO_2 , CH_4 , N_2O and others as detailed in Table 2. The gases are nick-named the Kyoto greenhouse gases, and also included in the table is the potential of each gas to increase surface temperatures in one hundred years. A number of additional pollutants have global-warming properties, including many ozone-depleting substances but they are regulated under a different protocol, the United Nations Environment Programme (UNEP) Montreal Protocol on Substances that Deplete the Ozone Layer and they are typically controlled under air quality regulation.

Table 2. Residence times and global warming potential of Kyoto greenhouse gases in the atmosphere

Greenhouse Gas	Residence time in the atmosphere	Global warming potential
CO_2	100 years	1
CH_4	10 years	25 times
N_2O	114 years	280 times
HFCs	260 years	11 700 times
PFCs	>50 000 years	6 630 times
sulphur hexafluoride (SF_6)	3 200 years	23 900 times

Climate change induced by greenhouse gases not only influences the behaviour of the atmosphere and weather systems over land and sea. It also brings with it modifications of a range of physical, chemical and biological processes within the terrestrial and marine ecosystems. Many of these process-changes directly or indirectly affect the composition of the atmosphere (Pleijel et al., 2009).

2.5 Methane

CH_4 is a short-lived climate pollutant (SLCP) and together with other SLCPs contribute as much as one third of the current total greenhouse forcing. This relates to a strong impact on climate change. Due to CH_4 short residence time in the atmosphere relative to other longer-lived greenhouse gases (CO_2 , SF_6 , N_2O , etc.) it is argued that strategies for reducing SLCP emissions should be prioritised as their reduction will lead to nearer-term climate benefits (Worden et al., 2021; Seinfeld, 2016).

Atmospheric concentrations of CO_2 , CH_4 and N_2O were recorded in 2011 to be 391 ppm, 1,803 ppb and 324 ppb respectively, which exceeded pre-industrial levels by approximately 40%, 150% and 20% respectively. Saunio et al. (2020) reported that concentrations have now exceeded the highest concentrations recorded in ice cores in the last 800,000 years. Globally averaged CH_4 was estimated to be 722 $\pm$ 25 ppb (IPCC 2013) and in 2011 it grew to 1,803 $\pm$ 2 ppb. Measurements of global CH_4 levels began in 1978 and in 1980 through to 2006, a decline was identified in its growth where global emissions approached a steady state at approximately 550 $TgCH_4yr^{-1}$ (Seinfeld, 2016).

Understanding and quantifying the global CH_4 budget is important for assessing pathways to mitigating climate change. The relative importance of CH_4 compared to CO_2 , other than its shorter atmospheric lifetime and stronger global warming potential is its significant growth rate over the past ten years (Saunio et al., 2020). Bottom-up methods suggest almost 30% larger global emissions (737 $TgCH_4yr^{-1}$) of CH_4 than top-down inversion methods, reckons Saunio et al. (2020), adding that atmospheric constraints on top-down estimates suggest that at least some of the bottom-up emissions are overestimated. However, Bergamaschi et al. (2018) offers that the difference could be explained by the significant contribution of biogenic sources (40%) which is not included in national inventories reported to UNFCCC.

The main sink of atmospheric CH_4 , according to Seinfeld (2016), is oxidation with the OH radicals in the troposphere, which accounts for at least 87% of the total sinks of CH_4 . Other sinks include photochemical reactions in the stratosphere, reactions with chlorine radicals in

the marine boundary layer and soil (Bergamaschi et al., 2018).

CH_4 is expressed as either fluxes or atmospheric concentrations. Flux refers to the difference between the amount of gas added to the atmosphere by emissions from various sources and the amount taken up by surface sinks that remove the gas from the atmosphere, while concentration refers to the amount of gas contained in a volume of air. Units for expressing CH_4 measurements are teragrams of CH_4 ($TgCH_4yr^{-1}$) for fluxes and *ppm* or *ppb* for atmospheric concentrations.

The general approach for inferring CH_4 emissions from observed atmospheric concentrations is to use a three-dimensional chemical transport model (CTM) describing the sensitivity of concentrations to emissions. Due to errors in measurements and in model transport, the best that can be achieved is an error-weighted statistical fit of emissions to the observations. Two different methods have been used for Greenhouse Gases Observing Satellite (GOSAT) and SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY (SCIAMACHY) retrievals, namely the CO_2 proxy method and the full physics method (Jacob et al., 2016).

2.6 Satellite Observations of Atmospheric Methane

Shortwave infrared (SWIR) satellite instruments observe solar backscatter from the Earth surface between the spectral ranges of 1.63-1.70 μm and at 2.2-2.4 μm (Jacob et al., 2022). Satellites that measure CH_4 in the thermal infrared (TIR) are not sensitive to CH_4 near surface but provide measurements of background CH_4 in the upper troposphere (Jacob et al., 2022). Satellites that observe in the SWIR also do not directly measure gases. Instead they obtain information on the column-averaged dry-air mixing ratios of the gases in the atmosphere, the average amount of a particular gas in a vertical column above a specified location excluding water vapour. These column-averaged dry-air mixing ratio values are denoted as X_{CH_4} with the X to indicate that they are from satellite measurements and therefore different from in-situ ground based or near-surface measurement. An X_{CH_4} concentration of 1,800 *ppb* means

that the atmosphere about a location contains on average 1,800 CH_4 molecules per 1 billion air molecules.

3 Literature Review

3.1 Introduction

The objective of this chapter is to provide an overview of the current state of knowledge in space-based observations of CH_4 , amongst other greenhouse gases, in support of the Paris Agreement and climate change mitigation strategies.

Section 3.2 presents the global CH_4 budget to provide context and to show why it is important for South Africa or any country on the planet. Section 3.3 gives an overview of satellite missions targeting CH_4 , what they have achieved in the past and currently, and what those proposed for the future will offer. Section 3.4 provides an account of current themes in terms of applications of space-based measurements of CH_4 , and section 3.5 an overview of emission inventories and which are used in different parts of the world. Finally, section 3.6 provides a review of time series forecasting or prediction models, especially those that have been applied in greenhouse gas estimations for climate policy.

3.2 Understanding the Global Methane Budget

Understanding the global budget is important for South Africa as every country needs to know how its emissions contribute to the overall global warming potential. Emissions of greenhouse gases from man-made activities are the most dominant cause of global warming. Figure 3 summarises the global growth trajectory of CH_4 between 1980 and 2020. The upper level shows the globally-averaged, monthly mean atmospheric concentrations determined from NOAA surface sites (Bergamaschi et al., 2018). Following a rapid decrease in CH_4 emissions in the 1990's, a steady state occurred until 2006 (Kort et al., 2014). The atmospheric burden of CH_4 grew again in 2007 (Kort et al., 2014). In 2017, global CH_4 emissions were estimated by atmospheric inversions to be $576 \text{ TgCH}_4\text{yr}^{-1}$, of which 60% ($359 \text{ TgCH}_4\text{yr}^{-1}$) was attributed to anthropogenic sources (Saunio et al., 2020). From NOAA's observations, global CH_4 emissions were estimated in 2021 to be 15% higher than in the 1984 - 2006

period (World Meteorological Organization, 2021). 2021 was the second year in a row where a record increase in atmospheric levels of CH_4 was reported. Findings by NOAA showed the increase was 17 *ppb*, the largest annual increase recorded since systematic measurements began in 1983 (World Meteorological Organization, 2021). The increase during 2020 was 15.3 *ppb*, with atmospheric levels averaging to 1,895.7 *ppb* during 2021, or around 162% greater than pre-industrial levels. Multiple studies (Bergamaschi et al., 2015; de Gouw et al., 2020; Jacob et al., 2016; Krautwurst et al., 2017; Turner et al., 2017) have attempted to uncover the cause of this renewed growth.

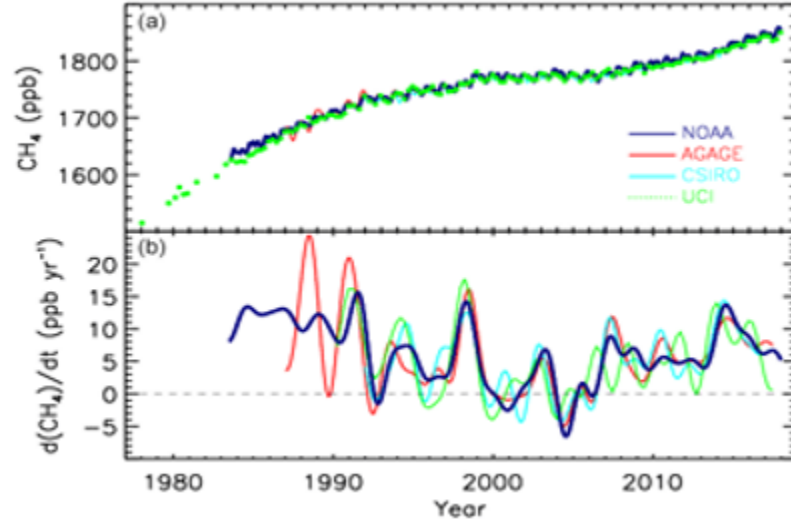


Figure 3. Globally averaged atmospheric CH_4 in *ppb* and its annual growth rate, as estimated by Saunio et al. (2020). It shows the decline in emissions in the 1990s, the rapid growth in 2007 and other significant changes in the growth rate of CH_4

3.3 Space-based Observations of Methane

Measurements of total CH_4 column by solar backscatter began in 2002, with SCIAMACHY onboard European Space Agency (ESA)’s Environmental Satellite (ENVISAT), until it went out of orbit in 2012. GOSAT which was launched in 2009 continues today to measure CH_4 with the Thermal And Near infrared Sensor for carbon Observation - Fourier Transform Spectrometer (TANSO-FTS) instrument (Alexe et al., 2015; Jacob et al., 2016). The GHGSat

instrument was launched on a microsatellite in June 2016 by GHGSat Inc with a high spatial resolution of $50 \times 50 \text{ m}^2$, over targeted viewing domains for the detection of point sources. It has a swatch of $12 \times 12 \text{ km}^2$ but with a precision of 1-5%. Carbon Monitoring Satellite (CarbonSat), also known as TanSat, was launched in 2016 and designed to measure CH_4 globally with an unprecedented combination of fine pixel resolution ($2 \times 2 \text{ km}^2$) and high precision of 0.4% (Jacob et al., 2016). The TROPOMI onboard the S-5P satellite, launched in October 2017, observes CH_4 with a $5.5 \times 7 \text{ km}^2$ spatial resolution and a precision of less than 1% (Veefkind et al., 2012).

The first Light Detection and Ranging (LiDAR) mission, MEthane Remote sensing Lidar mission (MERLIN), is a joint German and French space mission planned for launch in 2024. It has the scientific objective of delivering CH_4 atmospheric columns for all latitudes, with emphasis on poorly accessible regions in the tropics and at high latitudes (Ehret et al., 2017; Bousquet et al., 2018). The active sensor will make possible the constraining of CH_4 emissions irrespective of season, weather, or latitude and will operate both day and night (Bousquet et al., 2018).

CH_4 measurements in the TIR region of the electromagnetic spectrum began with the Interferometer Monitor for Greenhouse Gases (IMG) in 1996-1997 and continued from 2002 with the launch of Atmospheric Infrared Sounder (AIRS) and Tropospheric Emission Spectrometer (TES) in 2004. Infrared Atmospheric Sounding Interferometer (IASI), launched in 2006, even though it offers very low sensitivity in the boundary layer, it provides statistically consistent CH_4 emissions. Cross-Track Infrared Sounder (CrIS) is planned to launch on 30 Sep 2022. All instruments listed above have been onboard sun-synchronous LEO satellites and their mission launch dates and completion have been summarised in Figure 4.

Geostationary satellites, Geostationary Fourier Transform Spectrometer (GeoFTS), Geostationary Carbon Observatory (GeoCARB), Geostationary Coastal and Air Pollution Events (GEO-CAPE) and Geostationary Emission Explorer for Europe (G3E) have been planned and proposed for the future to map CH_4 on continental scales with a 2-5 km resolution (Jacob et al., 2016). Sheng et al. (2018) conducted simulation experiments to test the ability

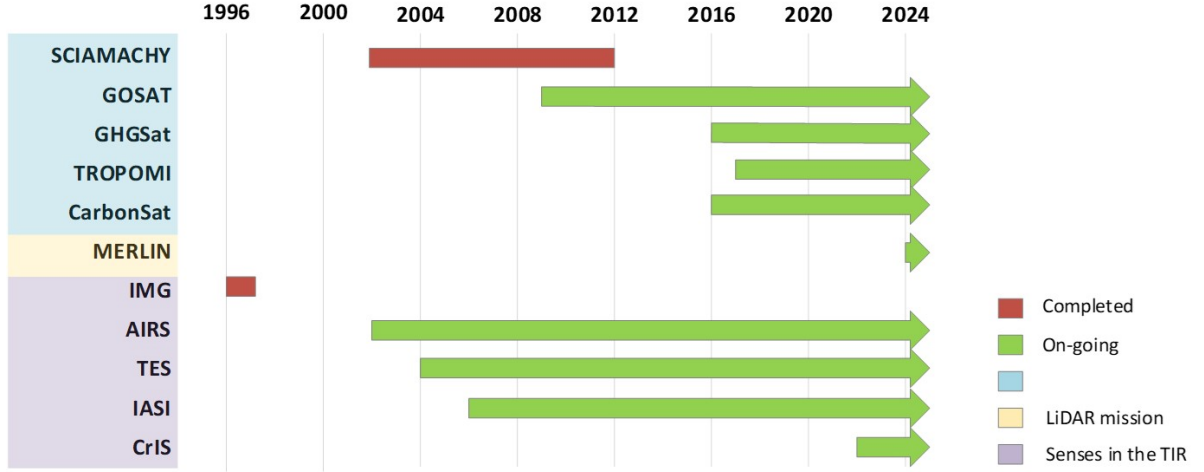


Figure 4. A summary of space-based instruments measuring CH_4 onboard completed and ongoing LEO satellites with the length of missions.

of these future satellites for constraining CH_4 emissions. After an inverse analysis using a GEOS-Chem CTM in a one-week simulation, Sheng et al. (2018) concluded that Geo-CARB and GEO-CAPE will perform a much better job than all the LEO satellites, including TROPOMI. The geostationary satellites are less sensitive to cloud cover and are better capable of constraining CH_4 emissions due to their finer pixel resolution and more frequent observations. Bousserez et al. (2016) conducted similar experiments to compare the efficacy of SWIR, TIR and multispectral SWIR and TIR measurements in geostationary satellites from LEO satellites concluding that geostationary satellites would constrain twice as many independent flux patterns and would greatly mitigate the impact of inherent uncertainties in source inversion methods. A further advantage of the geostationary satellites, particularly G3E, is going to be the ability to monitor point-source emissions of CH_4 due to its planned higher spatial resolution of $2 \times 3 \text{ km}$ (Butz et al., 2015).

3.4 Applications of Space-based Observations

The applications of space-based observations for emission hotspot detection and plume identification exist in literature and have been applied in different parts of world including the

USA (Buchwitz et al., 2017; Turner et al., 2015; Kort et al., 2014), Turkmenistan (Schneising et al., 2020), Argentina (Puliafito et al., 2020), and South Africa (Shikwambana et al., 2021a, 2022). Pandey et al. (2019) and Schneising et al. (2020) made use of TROPOMI observations to detect CH_4 leakages and monitor their dispersion in the atmosphere through time, demonstrating the effectiveness of routine satellite measurements in detecting and quantifying emissions from unpredictable events, such as gas and oil leaks which release dangerous amounts of CH_4 in the atmosphere. A common occurrence has also been the use of an EDGAR inventory for validation and verification or for testing satellite observations for agreement (Kort et al., 2014).

Estimation of CH_4 emissions on regional scales and using global inverse analyses for retrievals on GOSAT and SCIAMACHY have been the most predominant. Examples can be found in the works of Alexe et al. (2015); Lunt et al. (2019); Turner et al. (2015) and Wecht et al. (2014). Comparison of retrievals between the instruments have been noted (Alexe et al., 2015). Retrievals using TROPOMI are growing in significance, as in Qu et al. (2021) and Hu et al. (2018) who compared retrievals from GOSAT and TROPOMI.

Inversions for relating surface emissions of CH_4 to atmospheric mole fractions have typically been by a GOES-Chem CTM as conducted by Lunt et al. (2019); Turner et al. (2015) and Wecht et al. (2014) for global analysis, and nested versions became the alternative when studies were on a regional scale. Alexe et al. (2015) utilised a TM5-4DVAR model, which is a customized Bayesian optimisation CTM. Validations of the retrievals made against a total carbon column observing network (TCCON) or independent shipboard and aircraft observations and were found to be agreeable with some minor uncertainties and errors declared. Wecht et al. (2014) validated their results against an EDGAR v4.2 inventory.

Tracking of large-scale transport events was also identified, for example Pandey et al. (2019) and Schneising et al. (2020) used TROPOMI observations to detect CH_4 leakages and monitor their dispersion in the atmosphere through time, demonstrating the effectiveness of routine satellite measurements in detecting and quantifying greenhouse gas emissions from unpredictable events.

Additional recurring themes were estimation of CH_4 emission trends, where de Gouw et al. (2020) and Turner et al. (2015) used satellite retrievals to analyse trends and patterns in emissions, and Buchwitz et al. (2017) and Kort et al. (2014) detected CH_4 emission hotspots in the USA. Estimation methods included a fast data-driven method (Buchwitz et al., 2017) applied to GOSAT and SCIAMACHY retrievals and a weather research and forecasting coupled with chemistry (WRF-Chem) model was applied by Kort et al. (2014). Validation and agreement were tested using EDGAR v4.2.

3.5 Emission Inventories

A popular practice in space-based CH_4 studies is the use of bottom-up inventories either for validation/ verification of the space-based observations or as prior emission estimates in inversion analyses. It is common to see a national inventory of the country whose observations are being mapped or the global EDGAR being compared to the derived space-based estimations. For example Puliafito et al. (2020) used a high resolution Argentinian inventory, GEAA, for comparison with TROPOMI measurements and Thi Quynh Nguyen et al. (2021) made use of the Regional Emission inventory for Asia for validation when mapping emissions in Vietnam. In Varon et al. (2022)’s integrated CH_4 inversion toolbox, several bottom-up inventories make up prior estimates for different sections of the world, namely; (i) the US EPA Inventory of US Greenhouse Gas Emissions and Sinks is used for the USA, (ii) the National Institute of Ecology and Climate Change (INECC) inventory for Mexico, (iii) the Environment and Climate Change Canada National Inventory Report (ECCC NIR) for Canada, and (iv) the Global Fire Emissions Database and (v) EDGAR for the rest of the world. The Community Emissions Data System (CEDS) bottom-up inventory has also come up in comparison studies. Tan et al. (2022) developed a model using GEOS-Chem and two inventories EDGAR v4.3.2 and CEDS version 2017-05-18 where the inventories showed considerable differences reflecting large uncertainties in estimates of Chinese CH_4 emissions.

3.6 Forecasting Models for Methane Emissions

Forecasting of greenhouse gases are seen as essential for the monitoring and evaluating of global climate change policy (Tudor and Sova, 2021). At a local South Africa level, Durbach et al. (2017) were part of a project forecasting greenhouse gas emissions for the country for 2050 and subsequently developed a tool to assess uncertainty in the forecast greenhouse gas emissions. Tudor and Sova (2021) argue that the aim of time series forecasting is to identify the best predictive model within a pool of candidates and employ it to produce forecasts for the series of interest. They, therefore, compared seven predictive models, namely; (i) the exponential smoothing state-space model, (ii) the Holt-Winters (HW) model, (iii) the exponential smoothing state space model with Box-Cox transformation, ARMA errors, trend and seasonal components (TBATS) model, (iv) the ARIMA model, (v) the structural time series model, (vi) neural network auto-regression (NNAR) model, and (vii) the naive model on 49 countries to test if they would meet their 2030 targets for the global climate policy. They reported that NNAR presented the best performance for greenhouse gas emissions citing the lowest RMSE and mean absolute scaled error (MASE). In 2022, Tudor and Sova (2022) continued with the research on greenhouse gas emissions forecasting and predicted the aggregate greenhouse gas emissions in nine Central and Eastern European countries for the 2019 – 2050 horizon using the same set of six statistical methods and one machine-learning method.

Bhatti et al. (2021) predicted future concentrations of particulate matter using SARIMA for Pakistan, focusing mostly on the seasonal analysis of variation in the time series for the dataset. (Bhatti et al., 2021)’s model is well explained and the choice well defended.

More studies have predicted the evolution of CO_2 than any other greenhouse gas or air pollutant. Tudor (2016) predicted CO_2 emissions levels for Bahrain at the end of 2020 and 2021, again attesting to the supremacy of NNAR. Qader et al. (2021) are in agreement with regards to the supremacy of the NNAR model which they used to forecast the evolution of CO_2 for Bahrain with high accuracy. Malik et al. (2020) predicted CO_2 emissions for

Pakistan using an ARIMA univariate model with historical data for Pakistan up till 2020, forecasting yearly to 2030. What was interesting in this paper was that the forecasting was carried out under two different scenarios, (i) business as usual, and (ii) where the assumption was that the country adopted mitigation strategies to curb emissions. The model performed satisfactorily when evaluated using RMSE, MAPE and an additional mean absolute error metric. It was however not known if other models would have performed better, the study had no justification for the model and performance metrics chosen. Hopali and Cakmak (2020) predicted daily energy-related CO_2 emissions by a factory in the USA using ARIMA, HWA and HWM forecasting models. They found ARIMA to be the most appropriate model based on a MAPE score of 10.4%. However, this study only evaluated three models and there was no motivation on why only those three were selected for the time series forecasting. It also failed to provide a defense on robustness and it was not clear if any tests were done for stationarity of the auto-regressive model.

A study predicting coalbed CH_4 was also reviewed (Linghu et al., 2020). Even though the study was not based on space-based CH_4 observations but mainly for the rational allocation of resources and energy conversation, emission reduction was still an objective and was therefore relevant. Further it affirmed that the core of time series prediction is to find the law from the data and use it to make predictions about future data, as denoted by Box et al. (2016).

Time series and forecasting, even though common in air quality and greenhouse gas emissions literature, are not as wide for CH_4 emissions. Lastly most of the time series forecasting models reviewed were statistical in nature, with the exception of Tudor and Sova (2021), Qader et al. (2021) and Tudor and Sova (2022)'s best performing model which was a combination of a machine learning and auto-regressive model, the NNAR. Despite the popularity and successes of the NNAR model, it is omitted from this research as it employs complex and sophisticated machine learning techniques that are beyond the scope of the research. ARIMA and the Holt Winters' models have also been used successfully in the prediction of CO_2 emissions (Hopali and Cakmak, 2020; Malik et al., 2020). Therefore they have been employed in this research alongside others, namely the (i) naive, (ii) SMA, (iii)

SES, and (iv) SARIMA forecasting models.

3.7 Research Gaps

CH_4 emissions in South Africa have been studied on a qualitative basis using a combination of data sources such as indicators from the World Bank, census office and satellite data (Shikwambana et al., 2022). Studies that exclusively use satellite data for quantitative analysis or conducting chemical transport models currently do not exist. However, a study by Hakkarainen et al. (2021) quantitatively analysed other greenhouse gases and air pollutants using TROPOMI.

Additional research is needed to identify the cause of discrepancies between top-down and bottom-up results for many CH_4 emission sectors. There is a lack of confidence in the available CH_4 emission estimates as studies are still showing discrepancies between bottom-up inventories and top-down inversions (Worden et al., 2021; Turner et al., 2015; Saunio et al., 2020). Of course, the problem with bottom-up approaches involves spatial resolution as sampling using in-situ methods can only be performed at sparse locations or with a coarse spatial resolution leading to extrapolation of few samples or small areas to large areas, or relying on the use of proxy data such as energy consumption data and emissions ratios applied to species. These are the main sources of error.

With top-down approaches using space observations, even though achieving global coverage with frequencies of up to one day, the limitation is being sun-synchronous and therefore covering an area at the same local time each day. The temporal under-sampling affects the ability to accurately quantify fluxes. Seemingly geostationary satellites will mitigate some of these challenges, however not without the limitation that spatial coverage will be regional. For example, G3E will only be targeting central Europe (Butz et al., 2015). So far no geostationary satellites have been planned for covering South Africa or the continent of Africa.

The causes of the recent growth in concentrations of CH_4 in the atmosphere. Two main challenges in reducing the uncertainties in this atmospheric growth of CH_4 are (i) the

variety of geographically overlapping CH_4 sources on the planet and (ii) the destruction of CH_4 by short-leaved OH . Saunio et al. (2020) have reported that the most important source of uncertainty in the CH_4 budget is attributable to natural emissions, especially those from wetlands and other inland waters. They made several suggestions for improving CH_4 measurements one of them being the intensification of CH_4 observations at local scales and urban-scale monitoring to constrain bottom-up land surface models, and at regional scales (surface networks and satellites) to constrain atmospheric inversions. A lot of funding is required for the development and implementation of these networks and it is encouraging that some Annex 1 countries are well on their way such as with the UK and Ireland's Green house gAs UK and Global Emission (GAUGE) project, more of which can be read in Palmer et al. (2018). NOAA as well as the GAW programme of the World Meteorological Organization (WMO) also maintain wide networks of stations across the world for monitoring greenhouse gases at global and local scales. However, South Africa has only one observation station at Cape Point. It is the author's view that this physical constraint has to be addressed first by the construction of more observation stations for a more homogeneous coverage before this gap can be addressed. It is also interesting that the CH_4 concentrations at the Cape Point observation station published by the GAW program are different from those published by NOAA, even though captured by one provider, the South African Weather Services (SAWS).

It is also noted that significant uncertainties exist in the identification and quantification of individual CH_4 sources. One of the ways of addressing this is by the development of a 3D variational inversion system using isotopic and/or co-emitted species such as ethane to improve source partitioning. Addressing this gap requires a thorough understanding of atmospheric chemistry and atmospheric science engineering. This is an interesting area that can be addressed by future works by the author in addition to looking at improvements to transport models and the representation of photochemical sinks in top-down inversions.

4 Research Design

4.1 Introduction

This section outlines the methodology utilised for achieving the aim of analysing the spatial and temporal distributions of CH_4 and making predictions of emissions for 2022. The study area and datasets utilised are also presented.

4.2 Description of Study Area

The study was constrained to South Africa (Figure 5), a country on the southern tip of the African continent between the latitudes of 22° and 35° S, with a landmass covering an area of approximately 1.2 million km^2 . Most of South Africa experiences cold and dry weather conditions during the seasons of autumn and winter, and in turn the hot and rainy weather is experienced in spring and summer. It's weather is also influenced significantly by the El Niño southern oscillation.



Figure 5. Map of South Africa showing the location of its nine administrative areas

South Africa’s population is increasing at a rapid rate, reported to be 57,8 million in 2018 and growing at a rate of 29.8%. The World Bank reported its gross domestic product (GDP) to be 419.95 billion in 2021. The country is urbanising rapidly and with that, adding pressure to electricity demand as more houses and businesses are being established (Shikwambana et al., 2021b). The following paragraphs summarise industrial activity in South Africa and its profile with respects to CH_4 emissions and climate change mitigation.

38% of the country’s land mass is mainly occupied by cattle, sheep and goats, which are the main drivers of the livestock industry in the country. The agricultural sector as a whole contributed 2.3% to South Africa’s GDP in 2020 (Casey, 2021), and 2016 records indicate that animal production contributed 46% to the agricultural gross value. This sector contributes significantly to the CH_4 emissions in South Africa and Stevens (2021) reported that the agriculture, forestry, and other land-use (AFOLU) livestock category contributed 50.4% to the country’s CH_4 emissions in 2017. Emissions from animal husbandry are mainly due to enteric fermentation and manure management. The waste sector was the second biggest contributor to the CH_4 emissions of 2017, contributing 41.0% (Stevens, 2021). Car ownership levels were reported to be at 210 vehicles per 1,000 people in 2016. In many of its cities, motorised trips can be split equally between private and public transport trips, daily, however, with 20% of the population walking. Around 600,000 internal combustion engine vehicles are sold yearly (Brink and Stander, 2016). Most of South Africa’s freight is transported by road between its cities, with the remaining 20% being transported by rail. The above are contributing factors to the country’s increasing greenhouse gas emissions, which is currently the leading emitter of greenhouse gases, including CH_4 , on the African continent (Crippa et al., 2019).

In 2004, South Africa put in place legislation that mandates big industrial emitters to report their greenhouse gases (Shikwambana et al., 2020). The country has taken a strong stance on climate change on the international stage and has developed a significant and forward-looking domestic strategy that lays out a roadmap for mitigation and adaptation in line with the country’s NDC to the Paris Agreement (Averchenkova et al., 2019).

The DFFE in South Africa introduced a climate change Bill [B9-2022] to parliament on 18 February 2022. If the Bill becomes law it will form the first legal framework for a response to the impacts of climate change in South Africa. It acknowledges the urgent threat that climate change presents and emphasizes the need for an effective, progressive and incremental response and one of the most important aspects of the Bill is the DFFE minister’s obligation to determine a pathway for the reduction of greenhouse gas emissions progressively over time through a “national greenhouse gas emissions trajectory” which is consistent with global obligations and efforts. The Bill provides that within one year of the Bill coming into operation, the minister must publish a list of greenhouse gas emitting sectors subject to sectoral emissions targets with the thresholds expressed in CO_2 equivalents.

Further to this new Bill and the NDC, it has been a signatory to a number of international conventions for mitigating the impacts of climate change, including UNFCCC, IPCC and Conference of the Parties (COP) for a number of years, again demonstrating the country’s political will to reduce greenhouse gas emissions and play a positive role in the reduction of global warming. Despite its political will, bottom-up inventories for 2018 estimated its CH_4 emissions to be 45,140 kt of CO_2 equivalent.

4.3 Resources and Tools

The dataset being analysed were satellite data from the TROPOMI instrument onboard the S-5P satellite. Secondary datasets which aided the analysis were the EDGAR high quality bottom-up greenhouse gas inventory, surface flask measurements of atmospheric CH_4 from NOAA and the GAW program and Open Street Map (OSM) data.

4.3.1 Satellite data

TROPOMI level 2 (L2) daily observations from 30 April 2018 (when the TROPOMI operational record began) to 31 July 2022, is the primary collection of data utilised in the study. The data from May 2018 to December 2021 is utilised for the spatial analysis and training of

the time series forecast model while the data from 2022 is held out for testing the forecast.

TROPOMI retrieves atmospheric CH_4 columns from backscatter sunlight in the $2.3\mu m$ CH_4 absorption band, with daily global coverage at $5.5 \times 7\text{ km}^2$ nadir pixel resolution (Varon et al., 2022). For observations before August 2019, the pixel resolution was $7 \times 7\text{ km}^2$ and measurements are made at approximately 13h30 local solar time (Veefkind et al., 2012). The CH_4 retrieval, produced by the Netherlands Institute for Space Research (SRON), is based on the RemoTeC full-physics algorithm as column-average dry-air mixing ratios in *ppm* or *ppb* along with scattering properties of the atmosphere (Varon et al., 2022; Veefkind et al., 2012). Raw level 1 (L1) products are also available from the manufacturer and Schneising et al. (2019, 2020) have used them for developing a weighting function modified differential optical absorption spectroscopy (WFM-DOAS) algorithm optimized to retrieve CH_4 and CO simultaneously.

4.3.2 Bottom-up inventory

The study utilised the EDGAR inventory as one of the validation datasets. Agreement of TROPOMI observations to CH_4 fluxes reported by EDGAR were tested using EDGAR v6.0 of 2018 (Figure 6), constrained to the extent of South Africa. EDGAR is a global database of emissions of greenhouse gases and air pollutants compiled by the Joint Research Council for the European Commission. It provides gridded emissions data from atmospheric inversions and land surface models for wetland or biomass burning provided at monthly scale. The emissions it reports are yearly and monthly estimates of fluxes re-gridded to $0.1^\circ \times 0.1^\circ$. They are then converted into teragrams of CH_4 ($TgCH_4$) per grid cell. The gridded inventory reports emissions from various anthropogenic source categories that are mandated by climate policy such as fossil-fuel related activities, agriculture, waste management, transport and many more. Table 3 provides a complete list, and for each code the EDGAR dataset has a separate grid mapping the average fluxes for the source type. Its temporal resolution ranges from monthly to yearly from 1970 to 2018.

This independently compiled inventory is often used to compare emission estimates to

Table 3. Methane emissions source types reported by the bottom-up EDGAR anthropogenic inventory with their corresponding IPCC codes

IPPC Code	Description
1A1a	Power industry
1A1b	Oil refineries and Transformation industry
1A2	Combustion for manufacturing
1A3a_CDS	Aviation climbing and descent
1A3a_CRS	Aviation cruise
1A3a_LTO	Aviation landing and takeoff
1A3b_noRES	Road transportation no resuspension
1A3c+ 1A3e	Railways, pipelines, off-road transport
1A3d	Shipping
1A4+1A5	Energy for buildings
1B1a	Fuel exploitation COAL
1B1a+1B2aiii2+1B2aiii3+1B2bi+1B2bii	Fuel exploitation
1B2aiii2+1B2aiii3	Fuel exploitation OIL
1B2bi+1B2bii	Fuel exploitation GAS
2B	Chemical processes
2C1+2C2	Iron and steel production
3A1	Enteric fermentation
3A2	Manure management
3C1b	Agricultural waste burning
3C2+3C3+3C4+3C7	Agricultural soils
4A+4B	Solid waste landfills
4C	Solid waste incineration
4D	Waste water handling
5B	Fossil Fuel Fires

those reported by member states of the UNFCCC. It also features frequently in CH_4 satellite retrievals research (Wecht et al., 2014; Puliafito et al., 2020; Buchwitz et al., 2017; Liu et al., 2021) either as a priori for chemical transport modelling or for validation of calculated emission estimates.

4.3.3 Surface measurements

Surface measurements from the NOAA Global Monitoring Division (GMD) of the Earth System Research Laboratory (ESRL) collected by SAWS was the first set of surface flask CH_4 measurements for comparison with the space-based estimates (Lan et al., 2022). A second group of surface measurements has also been obtained from Mkololo et al. (2022)

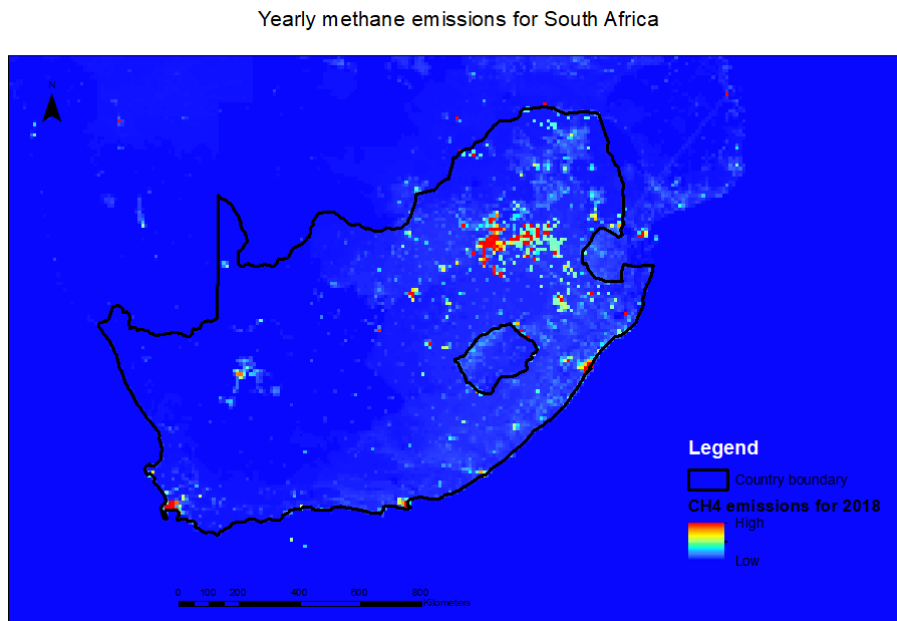


Figure 6. Mean emissions of CH_4 over one year, 2018, reported in EDGAR v6.0, for all source types from the GAW program of the WMO. The data is also collected by SAWS from South Africa's only CH_4 monitoring station at Cape Point.

The Cape Point station is located in a nature reserve at the southern end of the Cape Peninsula in South Africa, approximately 60 *km* from the Cape Town city centre. Its coordinates are 34°21' S and 18°29' E. Situated at 230 *m* above sea level on the Cape Peninsula, it is primarily exposed to maritime air from the South Atlantic.

In 2020, the coronavirus disease 2019 (COVID-19) pandemic impacted flask shipping and sampling frequency to the NOAA ESRL at many sites causing isolated gaps and some delayed processing into the second half of 2021. South Africa and the Cape Point measurement site was impacted by this and there was no data from March 2020 to December 2020. Flask shipping and sampling continued as normal from January 2021 and there is data for the whole year in 2021. The GAW data was also impacted in 2021 and has missing records from January to June in 2021.

4.3.4 Computing infrastructure

The project and research in its entirety was carried out on a desktop computer, a Lenovo ThinkPad with an 11th generation Intel Core i7 processor running at 2.8 gigahertz. The RAM of the computer was 16 gigabytes. It ran a Windows 10 Pro 64-bit operating system.

A number of software packages were made use of during the course of the project in the execution of the research. The biggest components included acquiring the data, which fits in the category of big data, and processing it. This was done in Visual Studio Professional 2022 using Python 3.9. Second most important was ArcGIS Pro 2.9 for the geo-processing of vector data and the production of all the maps. Quantum GIS 3.16 was used for some data conversions, Microsoft One Drive for storage of all data inputs and outputs, and finally GitHub was used for code back up and version control.

4.4 Method Design

4.4.1 Automated satellite data download and access

The download of satellite data from the Copernicus open access hub took place using the `sentinelSAT` Python package (Upadhyay et al., 2022). The package provides both a command line interface and a Python library for the search, download and metadata retrieval for all Sentinel products. It has at its disposal three functions, namely `SentinelAPI`, `read_geojson` and `geojson_to_wkt` which can be imported individually or in unison by not specifying the function names when importing the library. All the functions were imported for this work. `SentinelAPI` facilitated the connection to the Copernicus open access hub over the world wide web and it took as arguments, username and password credentials for authentication, as well as the website address to the open access hub. The login credentials are to a generic account and are the same for all users. They are published on the Copernicus website, the login is `s5pguest` and password is also `s5pguest`.

The `read_geojson` function takes the directory of a file that will be used as a footprint or the area to be downloaded. A shapefile of the South African country boundary was

converted into a geojson using Quantum GIS and its directory supplied to `read_geojson`. The `geojson_to_wkt` function then converted the given geojson file to a well-known-text (*.wkt*) format and allocated it to a variable, *footprint*. The footprint specifies the geographical search area and the algorithm only searches for products whose footprint intersects or is included in the geography. The footprint could also be specified by typing the coordinates of a simple polygon in *.wkt* format into the variable space. Even though there isn't a limitation on the number of points that can be specified for the bounding box, there is an 8,000 character limit related to the number of characters that can go into the next function, `api.query`. The footprint was then used as an input parameter for the `api.query` function together with other query keywords which specified the data to be downloaded. A number of keywords were specified, namely the start and end date, the product type and platform name. Table 4 was useful for the selection of keywords as the `api.query` function works by sifting through the TROPOMI product names and matches the search query to different parts of the product name sub-string. Table 4 also offers an explanation of all the elements in the sub-string. Lastly, the `api.download_all` function took as input the result of the `api.query` and the directory path where the data files were to be download. On execution, the products began to download to the chosen location in the background and depending on internet speed and start and end date specified, the data products of about 62 megabytes each, took from two minutes to eight hours to download.

In order to cover South Africa, two products were downloaded for each day, and on rare occasions, three products were required. For the period spanning 30 April 2018 to 31 July 2022, 3,417 products were downloaded.

4.4.2 Image processing and spatial analysis

Once downloaded, the products were processed using the `harp` (Zhou et al., 2020), a Python toolkit for reading, processing and inter-comparing satellite remote sensing data, model data, in-situ data and ground-based remote sensing data. It is part of an atmospheric toolbox developed for ESA by S&T Corp in partnership with European Organisation for the Exploitation

Table 4. Description of substrings in the TROPOMI product name

Substring	Meaning
S5P	Mission name
NRTI, OFFL or RPRO	Processing stream, near real-time, offline or reprocessing
L2_CH4	Product identifier
YYYYMMDDTHHMMSS	Start of granule in UTC, T is a fixed character
YYYYMMDDTHHMMSS	End of the granule in UTC, T is a fixed character
10028	Orbit number
01	Collection number
010302	Processor version number as “MMmmp”, with “MM” the major version number, “mm” the minor version number, and “pp” the patch level.
YYYYMMDDTHHMMSS	The time of processing for this granule in UTC
.nc	The file name extension. All Sentinel 5 precursor files are netCDF-4 files

of Meteorological Satellites and the Royal Belgian Institute for Space Aeronomy. The Python interface of the `harp` library provides a set of functions to import and export `harp` products. Its `import_product` function takes as input parameters, the file or files being imported, operations, options and post operations. A list of operations is provided as a single string where each individual operation is separated by a semi-colon, as in lines 3 to 7 of Script 1.

```

1  ## Apply quality filter, select variables, bin the time to desired
   period
2  period = harp.import_product(files_array, operations = ";".join([
3      "CH4_column_volume_mixing_ratio_dry_air_validity>50",
4      "keep(latitude_bounds,longitude_bounds,
   CH4_column_volume_mixing_ratio_dry_air)",
5      "bin_spatial(151, -35, 0.1, 251, 10, 0.1)",
6      "derive(latitude {latitude})",
7      "derive(longitude {longitude})",
8  ]), options="ch4=bias_corrected", post_operations="bin();squash(time,
   (latitude,longitude))")
9
10 ## Execute join operation
11 product_bin_bc = harp.execute_operations(period, "", "bin()")
12
13 ## Export joined product

```

```
14 harp.export_product(product_bin_bc, r"Outputs\\period.nc")
```

Script 1. Pre- and post-processing of the S-5P images using **harp** atmospheric toolbox

The operations used in the image pre-processing and appearing in Script 1 are described here and in the following paragraphs. `CH4_column_volume_mixing_ratio_dry_air_validity > 50` in line 3 was a comparison filter which searches all the data values in the *qa_value* file of the TROPOMI product. Items for which the value of the provided variable does not match the expression get excluded. Therefore all pixels for which the value was less than 50 were excluded.

The `keep(latitude.bounds,longitude.bounds,CH4_column_volume_mixing...)` in line 4 specified the variable or variables to be included in the ingested product. Exclusion of the `keep` function in the list of operations keeps all the TROPOMI variables, and these were kept only where additional geophysical parameters such as elevation, albedo or surface pressure needed to be analysed. For the pre-processing, only the X_{CH4} , latitude and longitude bounds were kept.

In line 5, `bin_spatial(151, -35, 0.1, 251, 10, 0.1)` mapped the image products onto a regular spatial latitude and longitude grid for all the variables in the TROPOMI product and the parameters of the function are listed in Script 2 where:

- `lat_steps` was the total number of steps or pixels the resultant variable maps have in the north - south direction, which is similar to the number of rows.
- `lat_edge` was the lower limit of the latitude bound.
- `lat_res` was the resolution in the north - south direction.
- `lon_steps` was the total number of steps or pixels the resultant variable maps have in the east - west direction, which is identical to the number of columns.
- `lon_edge` was the lower limit of the longitude bound.
- `lon_res` was the resolution in the east - west direction.

Script 3 details how the calculations of each parameter were carried out. The first step was to determine the desired resolution, which was $0.1^\circ \times 0.1^\circ$, similar to the EDGAR inventory. The value of 0.1 was therefore allocated for both `lat_res` and `lon_res`. Thereafter was to determine the latitude and longitude bounds. South Africa is between the latitudes of 22° and 35° South and between the longitudes of 16° and 33° East. Therefore, `lat_min` was allocated -35 and `lon_min` was allocated 10 to accommodate a buffer area on both sides. For the desired `lat_max` to be -20, at a resolution of 0.1, the number of steps to be accounted for was 150. For the desired `lon_max` to be 35, similarly with a resolution of 0.1, the number of steps to be allocated was 250. It was necessary to add 1 to each of the steps as the grid is specified in terms of cell corners and not the cell center positions.

In lines 6 and 7, `derive(latitude latitude)` and `derive(longitude longitude)` left it to `harp` to derive the latitude and longitude bounds of the variables after the `bin_spatial` operation had been executed.

In line 8, `post_operations="bin();squash(time,(latitude,longitude))"` removed the time dimension from the latitude and longitude variables and `options = "ch4=bias_corrected"` specified whether to ingest the 'normal' X_{CH_4} column (default) or the bias corrected X_{CH_4} column. For ingestion of the normal X_{CH_4} column, the `options` could be omitted from the parameters of the `import_product` function as it is an optional parameter.

```
1 bin_spatial(lat_steps,lat_edge,lat_res,lon_steps,lon_edge,lon_res)
```

Script 2. `bin_spatial` parameters

```
1 bin_spatial((lat_max-lat_min)/lat_resolution + 1, lat_min,
    lat_resolution, (lon_max-lon_min)/lon_resolution + 1, lon_min,
    lon_resolution)
```

Script 3. `bin_spatial` parameter calculations

An image of a single day does not reflect the X_{CH_4} values of X_{CH_4} across the entire country. Further, the conditions of the atmosphere with respects to weather, rain or clouds

on a particular day are capable of invalidating the obtained measurements of that day. Application of the quality filter also stripped the dataset of a significant amount of pixels. It was then more realistic to work with averaged values over a number of days by temporal averaging to a resolution that is representative enough of the geophysical properties being mapped. This was realised by the `bin()` function in line 11 of Script 1. The function performed an averaging in the time dimension such that all samples in the same bin got averaged for all variables in the product. The data was temporally averaged to weekly and monthly intervals.

After image processing with `harp`, the images were exported into the World Geodetic System 1984 (WGS84) coordinate system as level 3 (L3) NetCDF files for further analysis. `Harp` exports products as either NetCDF, HDF4 or HDF5. The downloaded products were imported into `harp` using the `import_product` function.

The monthly NetCDF files were imported into ArcGIS Pro and each converted to a raster for further analysis. A simple overlay process was employed where a band math algorithm was used to add the cumulative X_{CH4} pixels values to each other. This allowed the creation of a heat map where areas that consistently had high emissions resulted in higher values and those that consistently had low emissions resulted in lower values. The heat map raster was then overlaid with data from OSM of the following categories, (i) polygons where the land use was industrial, (ii) point locations of buildings, (iii) point locations of landfills, and (iv) point locations of wastewater treatment sites. The purpose was to investigate if the locations of the aforementioned items would coincide with high emissions as the items have been identified as high emitters.

4.4.3 Temporal analysis

The time series and model prediction were undertaken in the following steps:

1. Data preparation and decomposition of the time series data;
2. Testing for stationarity using the ADF and KPSS tests;
3. Box-Cox transformation and differencing to make the series stationary;

4. Building the time series forecast; and
5. Performance evaluation.

Data preparation and decomposition of series data

The CH_4 amount fractions across the study area were averaged. This was executed in ArcGIS Pro where the L3 NetCDF files were converted to GeoTIFFs in order to fit raster requirements in ArcGIS Pro. The GeoTIFF was converted into a point shapefile where the centroid of each grid was converted into a point with the CH_4 value of the pixel. The unweighted mean of the pixels within the country boundary were then calculated for each time interval. Monthly averages were collated in a text file alongside their dates of occurrence as an index. The data was plotted using the `seaborn` package (Waskom, 2021) in Python.

Testing for stationarity using the ADF and KPSS tests

A prerequisite to fitting a time series model is stability of the series (Bhatti et al., 2021). When a time-series exhibits non-stationarity, stabilisation techniques are required before the data can be processed further. ADF and KPSS tests were employed to check the stationarity of the X_{CH_4} time series. This determined whether the mean, variance and co-variance of the series were constant. Using the ADF test, the null hypothesis ($p - value > 0.05$) is that the series is not stationary, and following the KPSS test, the null hypothesis is that the series is stationary. Stationarity of the series was tested using the ADF and KPSS statistic against critical values in a t -distribution under a 95% confidence interval. The null hypothesis was also evaluated.

Box-Cox transformation and differencing

For a time series with poor stability, the research employed a differential Box-Cox transformation to achieve stability. Box-Cox applies a transformation to an abnormally distributed series to the best approximation of a normal distribution, making the variance constant. The transformation follows Equation 1. Box-Cox differencing was applied to remove the trend

component from the series.

$$y(\lambda) = \begin{cases} \frac{y^\lambda - 1}{\lambda} & : \lambda \neq 0 \\ \log y & : \lambda = 0 \end{cases} \quad (1)$$

Building the time series forecast

For building the time series forecast and deploying the models, a hold-out technique for the test data was applied by keeping the 2022 monthly dataset for testing while the complete X_{CH4} TROPOMI averages were used to train the time series. The models used in the study and implemented in a Python environment are summarised in the following paragraphs.

Naive forecasting is one of the simplest forecasting techniques in which the last period's actual value is used as the current period's forecast (Ostermann et al., 2022), without any adjustments or any attempts to establish causal factors. It follows Equation 2 and it serves as the baseline model for the forecasting. The performance of all models following the baseline were expected to be superior to the naive forecasting model to be considered appropriate.

$$Y_t = Y_{t-t_\Delta} \quad (2)$$

SMA is also a simple forecast method, commonly used with time series data to smooth out short-term fluctuations and highlight longer-term trends, according to Zeiml et al. (2019). It creates a series of averages of different subsets of the full series. The first element of the moving average is obtained by averaging the initial fixed subset of the number series. Then the subset is modified by moving forward and excluding the first number of the series and including the next value in the subset. This is presented in Equation 3, where Y_t is data in the current period and m is the length of the moving period. The equation has been adopted from Swari et al. (2021).

$$Y_{t-1} = Y_t + Y_{t-1} + \dots + Y_{t-m+1} \quad (3)$$

The basic idea pertaining to Holt Winters' models (additive and multiplicative) is said to be to make smooth estimates of long-term trends, trend increments and seasonal changes for a time series with line trends, seasonal changes and random changes, and then to make a prediction model and extrapolate the predicted values (Linghu et al., 2020). Holt's forecasting (or prediction) models smooth the time series with three exponential smoothing formulae that are applied to the mean, trend and each seasonal sub-series. The exponential component (E_t) is given by Equation 4, the trend given by Equation 5 and finally the k^{th} step is given by Equation 6. The equations have been adopted from Tudor and Sova (2022).

$$E_t = \omega T_t + (1 - \omega)(E_{t-1} + T_{t-1}) ; 0 < \omega < 1 \quad (4)$$

$$T_t = v(E_t - E_{t-1}) + (1 - v)T_{t-1} ; 0 < v < 1 \quad (5)$$

$$F_{t+k|t} = E_t + kT_t \quad (6)$$

One the time series was deemed stationary, the auto-regressive moving average model was established using Equation 7, adopted from Linghu et al. (2020) and Bhatti et al. (2021). This model is denoted ARIMA (p, d, q) and is also known as the auto-regressive integrated moving average model. Its construction is based on two sub-models, the auto-regressive model and the moving average model. In simple terms, the model says that "a data series formed by the predicted object over time can be regarded as a random series probably including volatility, periodic trends and seasonal factors, and after technically processing, a certain mathematical model can be used to describe this series approximately" - Bhatti et al. (2021).

$$Y_t = \varphi_0 + \varphi_1 X_{t-1} + \dots + \varphi_p X_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \quad (7)$$

In Equation 7, Y_t is the linear combination of the observed values at previous p periods, ε_t is the random error, a series of independent and identically distributed random variables with the homogeneity of variance satisfied (Bhatti et al., 2021). The terms p and q are the order of

the auto-regressive model and the order of the the moving average model, respectively. The terms, $\varphi_i(i = 1, 2, \dots, p)$ and $\theta_j(j = 1, 2, \dots, q)$ are the undetermined coefficients of the model.

The model was further adapted for a time-series with significant seasonal effects in the form of the SARIMA model. The SARIMA model is based on a short-term correlation through a low-order auto-regressive moving average (ARMA) (p, q) model and seasonal correlation is extracted through an ARMA (P, Q) model with a period step S , changing the model structure to ARIMA $(p, d, q) \times (P, D, Q)$, with a multiplicative relationship between the two aspects.

The goodness of fit of the candidate models were assessed using RMSE and MAPE performance metrics. RMSE is a common metric for evaluating the performance of time series forecasting and prediction models and is usually recommended above most metrics (Tudor and Sova, 2021). It represents the square root of the mean square error and is computed by Equation 8, adopted from Ostermann et al. (2022) and Tudor and Sova (2021). In simpler terms it is the standard deviation of the forecast errors. MAPE, on the other hand, measures the difference of forecast errors and divides it by the actual observed value (Equation 9).

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (8)$$

$$MAPE = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (9)$$

The notable difference between RMSE and MAPE and why performance is evaluated using both metrics is that the former is scale dependant, while the latter is independent of scale. This means that RMSE cannot be used to measure out-of-sample forecast accuracy when multiple series of different measurement units are being analysed. This brought the need for an accompanying scale-independent metric, which MAPE satisfies. MAPE measures forecasting error as a percentage and is the most common measure as it scales different

variables units to percentage and therefore can be used to evaluate accuracy out-of-sample. The use of MAPE was identified in the studies of Bhatti et al. (2021) and Linghu et al. (2020).

The univariate time series analysis for this research was carried out in Visual Studio Professional 2022 in the Python language using the following packages: (i) `scipy.stats`, (ii) `statsmodels`, (iii) `pylab`, (iv) `sklearn`, and (v) `seaborn`.

4.4.4 Validation

A correlation analysis was performed in Python using the `xarray` package by Hoyer and Hamman (2017) to determine the existence of any relationships between the satellite observations and bottom-up estimates. Before this could be achieved, a geographical subset was created with a mask applied over the data pairs and to get the datasets to be co-located. And thereafter, a Pearson correlation statistics was calculated for the pair using Equation 10

$$r = \frac{\sum (x_i - \hat{x})(y_i - \hat{y})}{\sqrt{\sum (x_i - \hat{x})^2 \sum (y_i - \hat{y})^2}} \quad (10)$$

5 Results

5.1 Introduction

The aim of the research was to evaluate the capability of TROPOMI as a tool to aid South Africa in its climate policy and in meeting its targets as set out in its NDC. This will be achieved if TROPOMI can accurately measure CH_4 concentrations in the atmosphere above South Africa and in facilitating predictions of CH_4 concentrations for the country into the future. The success of this aim will see south Africa moving to new frontiers in climate change mitigation as it will serve as a warning system for the country's risk of not reaching its emission reduction targets.

The aim was meant to be achieved by conducting a number processes as specified under the method design in section 4.4. Firstly, it involved a spatial analysis of TROPOMI CH_4 observations for identification of hotspots that are related to a number of anthropogenic source sectors such as livestock rearing, oil and gas systems, coal mining, landfills and wastewater treatment. Secondly, a time-series analysis of the average CH_4 emissions across the country in order to test for seasonality, trends of any components that can aid in providing intelligence pertaining to greenhouse gases and climate change. To also answer the question of whether the evolution of CH_4 emissions in South Africa are in line with observed CH_4 emission estimates globally? Thirdly, production of forecasts of CH_4 emissions through the employment of univariate time-series-based prediction models. Lastly it involved validating the data obtained using an EDGAR inventory for the establishment of any relationship between estimates of CH_4 obtained using bottom-up methods and those retrieved using space-based methods. Surface flask measurements during the corresponding years were also investigated.

This chapter presents the results obtained for the spatio-temporal analysis of CH_4 emissions in South Africa using observations of TROPOMI onboard S-5P. The first sub-section 5.2 addresses the spatial analysis component, followed by section 5.3 which is the temporal analysis and lastly the cross-validation with EDGAR and surface data in section 5.4.

5.2 Spatial Analysis

Figures 7 to 9 are visual representations of how the column-averaged dry-air mole fractions of CH_4 evolved across South Africa over the course of three and half years, as per the space-based CH_4 observations from TROPOMI since it began its operational record in May 2018. TROPOMI's daily observations from 01 May 2018 to 31 December 2021 were retrieved and processed to quantify the CH_4 concentrations in the atmosphere above South Africa on a monthly basis. Overall, the maps show the gradual increase of concentrations from May 2018 to December 2021 at a quick glance. The country icons represented using the inferno colour bar start off lighter, increasing in intensity until completely dark at the end of the timeline. The mean concentrations started at 1,800 *ppb* in May 2018 and gradually rose to 1,836 *ppb* by the end of 2021. This is an increase of 36 *ppb* observed for the country in a period of three and a half years. However, the lowest concentrations were in November 2018 at 1,792 *ppb* and the highest recorded are in September 2021 at 1,838 *ppb*, with an absolute range of 46 *ppb*. The evolution across different parts of the country on a month to month basis is described in detail in the following paragraphs.

May 2018 to December 2018

The monthly average for May exhibits moderate CH_4 levels in the north and north-eastern parts of the country. For June, this region shows a significant increase in the CH_4 monthly average. The increase can be traced to the Eastern Cape Province and KwaZulu Natal province. The central parts of the country including the Free State and Northwest province show a slight decrease in CH_4 concentrations, from May to June. There are no major changes in monthly averages over the western parts of South Africa in the same months. During the month of July, the concentration levels did not change significantly from those observed in June. However, the western parts show some increases, specifically over Cape Town and the Cape Point area. August showed a decrease in atmospheric CH_4 levels across the country, with an exception over the extreme north in the Limpopo province where there is a significant increase. September exhibited some enhancements over the east coast, which

persisted throughout October including parts of the Mpumalanga Highveld coal belt and towards the Namibian border. December is characterised by an overall decrease in the CH_4 concentrations, especially along the east coast in KwaZulu Natal and extending to the Eastern Cape.

January 2019 to December 2019

The monthly average for January showed higher concentrations in the eastern and central parts of South Africa, with the higher concentrations moving towards the west of the country in February. An increase in concentrations was also noted over the greater Gauteng, Limpopo and Mpumalanga Highveld coal belt in February. In March, a belt of low concentration levels that was present over the North West and Free State disappeared, with concentrations in excess of 1,790 *ppb* over the country except for a small area over the Western Cape and parts of the Eastern Cape which were still below 1,970 *ppb*. In April the CH_4 concentrations declined across the country except for a portion over Gauteng, which followed the same pattern in May. However, May saw an increase with a belt of higher emissions on the east coast that remained through-out the year and into 2020. The enhancement of CH_4 over the east coast looks like it started in March, however pixels are missing over the east coast in April and making it difficult to say if the enhancement was present in April. An enhancement of emissions was also present in the central interior covering areas such as the Free State, Gauteng and Mpumalanga in March, the Northern Cape and Mpumalanga in June, Mpumalanga and parts of the Western Cape towards Cape Point in July and greater Gauteng, and Mpumalanga in December. These enhancements were between the ranges of 1,825 and 1,843 *ppb*. The Western Cape and Eastern Cape showed sharp declines in the range of 1,753 and 1,771 *ppb* in November and December. The mean X_{CH_4} decreased for the whole country in the last two months of 2019, predating the COVID-19 pandemic. The highest CH_4 concentration was recorded in September 2019 at 2,069 *ppb* on the west coast in KwaZulu Natal.

January 2020 to December 2020

The year started with a number of missing pixels in January to April. The number of valid pixels began to stabilise in May providing a stable picture for the year, before dropping again in December. Despite the missing pixels, there was an increase in CH_4 concentrations from the lower levels in the preceding months. Concentrations of between 1,825 and 1,844 *ppb* persisted over the east coast. Most of the country had concentrations in the region of 1,807 and 1,824 *ppb* except for more enhancements over the area north of Lesotho going all the way up to Gauteng in March, the north-east of the country encompassing Limpopo and the eastern parts of Mpumalanga in April, and Gauteng in May. June saw the concentrations between 1,825 and 1,844 *ppb* spread across the country with more areas covered almost homogeneously in June and July. Concentrations dropped in August and September with only a few areas to the west of the country still showing the enhanced values between 1,825 and 1,844 *ppb*. Concentrations of CH_4 increased from October till the end of year, however a decline was experienced in November over the Western Cape, parts of the Karoo and the Eastern Cape. December has a magnitude of missing pixels over the east of the country.

January 2021 to December 2021

The year of 2021 saw emission concentrations above 1,825 *ppb*, increasing as high as 1,959 *ppb* in November. The Western Cape, the Karoo in the Northern Cape and Eastern Cape still had lower concentrations of CH_4 in the atmosphere than the rest of the country, however increasing slightly to between 1,790 and 1,824 *ppb* in January till March. The rest of the country measured at more than 1,825 *ppb*. Concentrations of over 1,862 started to appear in February originating from Namibia and spread across the north east of the country. The north-east of Northern Cape towards the Botswana border, the North West, Free State, Gauteng and Mpumalanga were all affected. There are missing pixels over Limpopo and KwaZulu Natal. In March, as more pixels were available for the country, it was possible to see that Limpopo and KwaZulu Natal were also affected by the widespread CH_4 enhancement, which intensified towards the north of the country. The CH_4 enhancement over the north dropped slightly in April with a few parts over the Northern Cape towards the Namibian

border affected. The North West province had a number of spots of enhanced values, similar to parts of Gauteng, Mpumalanga, KwaZulu Natal and the north-east of Limpopo towards Thohoyandou. The lower atmospheric concentrations in the Western Cape, the Karoo in the Northern Cape and Eastern Cape increased slightly from what they were before March, and remained between 1,807 and 1,843 *ppb* until September. In October the south west of the country experienced a decline in atmospheric concentrations of CH_4 with values between 1,790 and 1,806 *ppb*. However, enhancements over 1,844 *ppb* persisted over the likely suspects Mpumalanga, Gauteng, Limpopo and parts of KwaZulu Natal until November. December has a significant number of missing pixels over the eastern parts of the country.

Overall, three particular regions were identified as high emitters across consecutive months, namely:

1. The Mpumalanga Highveld coal belt characterised by the country's coal mines and the coal-powered Eskom power stations which have chosen to be in those areas to be in close proximity to the coal mines.
2. Gauteng and greater Johannesburg region characterised by the gold mines and the Johannesburg Stock Exchange. A top of industrial activities in manufacturing and production of goods can be found in the East Rand and Isando areas of Gauteng.
3. Richards Bay industries and surrounding suburbs. Richards Bay is a growing metropolitan with several industries in relative proximity to residential areas. In addition to being in close proximity to the largest coal export terminal in the world and the second-largest port in South Africa, Richards Bay hosts several commercial, light, and heavy industrial activities such as paper, fertiliser and sugar production (Okello et al., 2020).
4. Limpopo is not associated with heavy industry. However, there is a thriving manufacturing sector in the area and it hosts South Africa's richest and largest agricultural areas.

These are areas of significant CH_4 emitting activities. The results of the overlay process in Figure 10 show that these areas coincide with higher building densities, landfills and

wastewater treatment sites which have been regarded as proxies for high-emission activities. The Limpopo, Mpumalanga, Gauteng, KwaZulu Natal and Eastern Cape provinces have the most buildings and industrial activity. There is also a significant amount of the contributing factors in the Western Cape towards the Cape Town area, however due to missing data, it is not possible to see if there are high emissions in the area.

5.3 Time Series Analysis and Forecasting

The results of the time series analysis and forecasting are presented in this sub-section.

The temporal evolution of X_{CH_4} observed by TROPOMI in the atmosphere over South Africa is displayed in Figure 11. The time series indicates that atmospheric concentrations of CH_4 have steadily increased since May 2018. Figure 12 demonstrates the decomposition of the time series data. It showcases that the data exhibits both trend and seasonality and therefore a model that can incorporate both into the forecast would be more desirable.

The test statistics for the ADF and KPSS tests are presented in Table 5. For a sample of 51 observations the critical value for a series with a trend is -3.50. The calculated ADF statistic is 0.062, a positive value and therefore greater than the critical value. Therefore at a 95% confidence interval, the null hypothesis is accepted. This is conformed by the p-value of 0.963 which greater than 0.05. Therefore the series is not stationary. The KPSS statistic is 1.049, which is greater than the critical value of 0.46 at a 95% confidence interval. This is sufficient to reject the null hypothesis. A p-value of 0.01 which is less than 0.05 also causes a rejection of the null hypothesis and therefore the the series is not-stationary. Table 5 indicates the new test statistics after applying a Box-Cox transformation to make the variance constant, and a Box-Cox differencing to remove the trend component of the series.

It can be confirmed visually from Figures 13 and 14 that among the seven models, Holt Winters' models outperformed the rest, by observing the yellow and green lines. The yellow lines represent the test data, therefore the actual X_{CH_4} values obtained from the satellite. The green line represent the X_{CH_4} values fabricated by the models. In Figure 13, the naive,

Table 5. Tests Statistics for the ADF and KPSS tests on the time-series data

	Original Series	Post-transformation
ADF Statistics	0.062	-7.154
Critical-values (at 5%)	-2.92	-2.92
p-values	0.963	0.000
KPSS Statistics	1.049	0.274
Critical-values (at 5%)	0.46	0.46
p-values	0.010	0.100

SMA and SES models forecast values of 1,835 to 1,836 *ppb* for all the months in the year. The naive model forecast 1,836 *ppb* for all the months which is equivalent to the last X_{CH_4} value in the series, December 2021's measurement. The SMA model forecast 1,835 *ppb*, which is equivalent to the average of the last 6 months in the series, corresponding to the m that was specified in the model. The SES model also predicted 1,836 *ppb*, similar to the naive model. The plots show the distance between the forecast values and the actual values which do not intersect at any point.

In contrast, the Holt Winters models and the auto-regressive moving average models in Figure 14 appear to be more in-sync with the actual data showing at least one point of intersection between the test data and predicted values. The ARMA has approximated the trend without any seasonal components. The Holt Winters models incorporated both trend and seasonality and their test and predicted values intersect at two occurrences

Using the RMSE and MAPE evaluation metrics, the lowest MAPE is desired as it indicates the percentage of error in the prediction. The smallest RMSE is also desired as it indicates the size of the error based on the scale of the sample. Further the models are rated based on the performance of the baseline and the expectation is that they will have a RMSE which is less than 9.5 and a MAPE score below 48%. The SMA and SES models performed worse than the baseline, with RMSEs of 11.07 and 9.8 respectively. Their MAPE were also higher than 48%, with the SMA scoring 57% and the SES scoring 50%. They were immediately discarded.

The ARIMA has a RMSE of 6.97 and MAPE score of 33% which indicates that it has performed better than the baseline. The SARIMA also performed better than the baseline with a RMSE of 6.58 and MAPE score of 28%. The HWA and HWM performed almost similarly, with the best MAPE scores at 25%. Due to a lower RMSE for the HWA at 4.95 in comparison to the HWM at 5.05, the HWA performed the best of all the models. The scores have been summarised in Table 6.

Table 6. Forecasting accuracy metrics for the evolution of X_{CH_4} in South Africa in 2022

Model	RMSE	MAPE
Naïve	9.5	0.48
SMA	11.07	0.57
SES	9.8	0.5
HWA	4.95	0.25
HWM	5.02	0.25
ARIMA	6.97	0.33
SARIMA	6.58	0.28

The X_{CH_4} values predicted by the best performing forecast model, the HWA model are presented in Table 7. It has predicted higher CH_4 emissions for 2022, exceeding the highest recorded measurement in September 2021 of 1,838 *ppb*. The model also predicted that the highest X_{CH_4} would be 1,852 *ppb* in June 2022. The lowest predicted value is 1,840 *ppb* with an average of 1,847.433 *ppb*. In contrast the lowest actual X_{CH_4} value is 1,839 *ppb* and the highest values occurred in February and again in July at 1,849 *ppb*. The actual mean of observations for the first half of 2022 is 1,844.9 *ppb*.

Table 7. X_{CH_4} emissions data forecast for 2022 using the HWA model. The forecasts are presented against actual values from the TROPOMI X_{CH_4} values obtained for January to July 2022

Month	Actual Values	HWA Forecasts
01-2022	1844	1840.688
02-2022	1849	1845.355
03-2022	1842	1849.355
04-2022	1839	1845.688
05-2022	1844	1848.565
06-2022	1847	1852.065
07-2022	1849	1850.315

5.4 Validation

Figure 15 shows the EDGAR v6.0 dataset compared alongside the TROPOMI averaged data for the same period. Even though some similarity can be identified visually, a correlation coefficient of 3.45×10^{-19} was achieved. Similarities are noticeable in KwaZulu Natal and in the Guauteng and Mpumalanga regions where high emissions reported on the EDGAR inventory (a) are observable on the map produced from TROPOMI data (b).

Figure 16 is a graph showing the relationship between surface measurements of CH_4 and observations of TROPOMI on corresponding dates. It shows the discrepancy between the datasets. The datasets tend to agree on the trend but the seasonal fluctuations are different for all of them with the NOAA measurements having a smooth curve and sharp spikes are noticed in June every year with the GAW data.

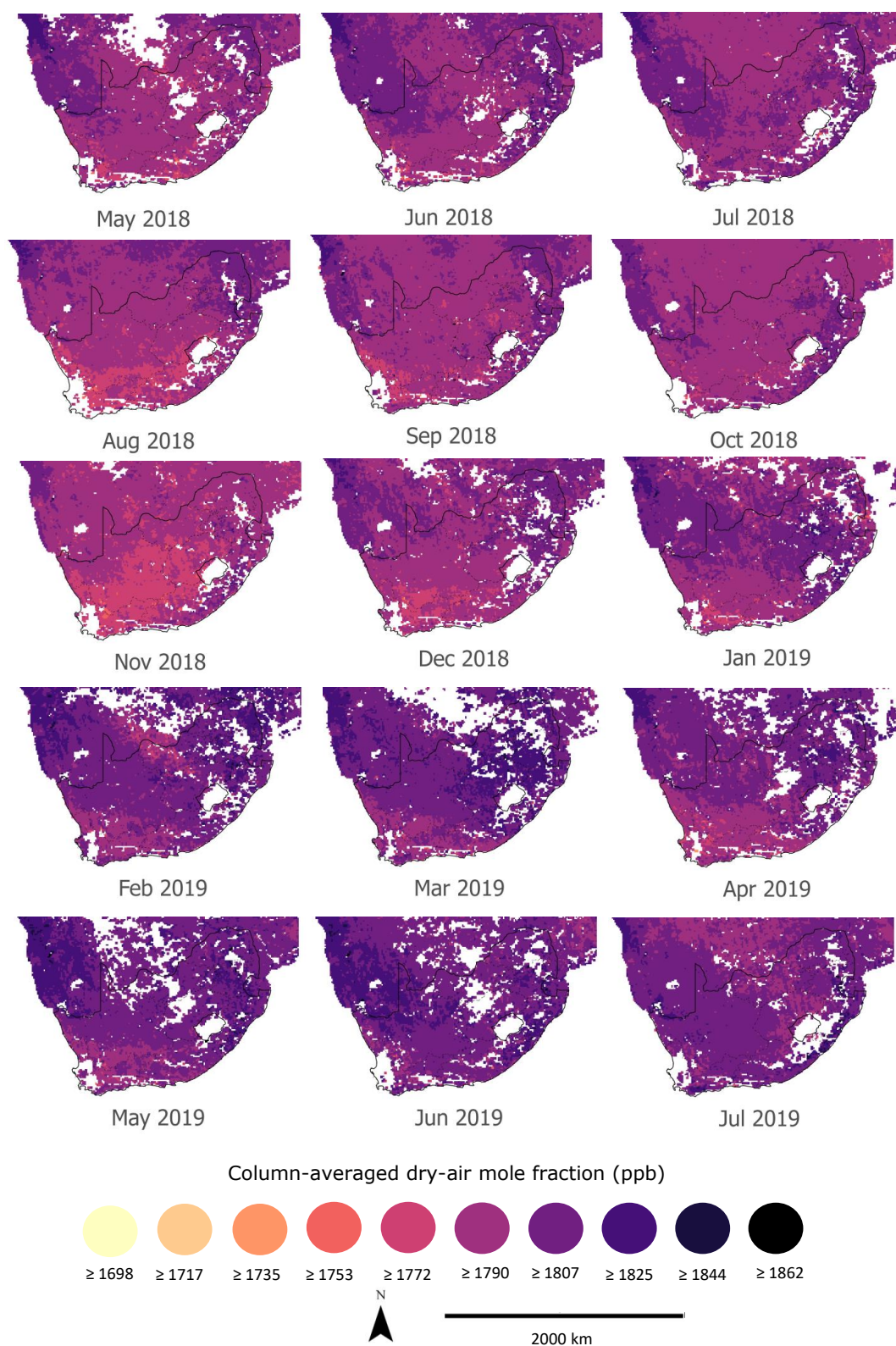


Figure 7. Concentrations of atmospheric column-averaged dry-air mole fractions of CH_4 in South Africa from May 2018 to July 2019

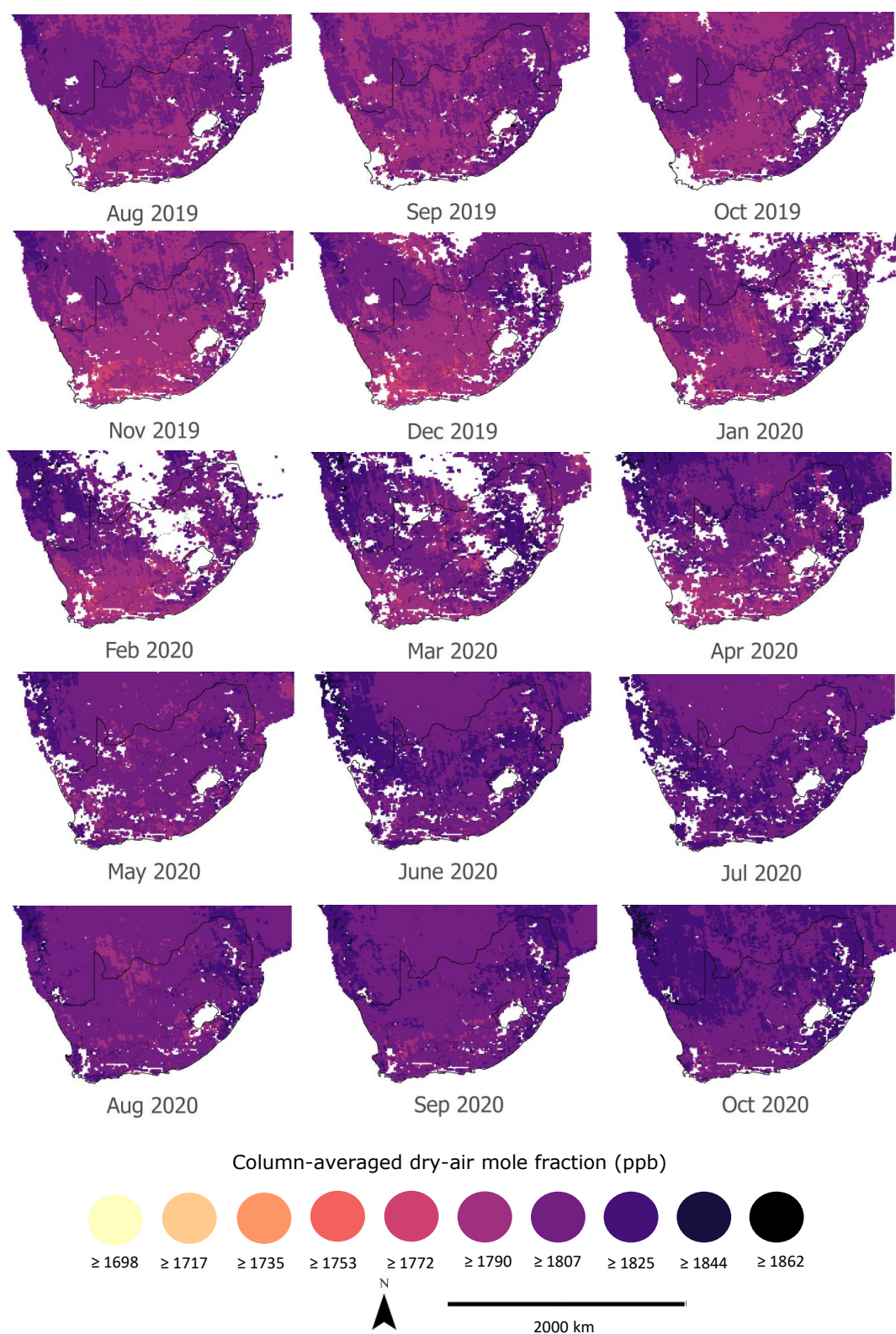


Figure 8. Concentrations of atmospheric column-averaged dry-air mole fractions of CH_4 in South Africa from August 2019 to October 2020

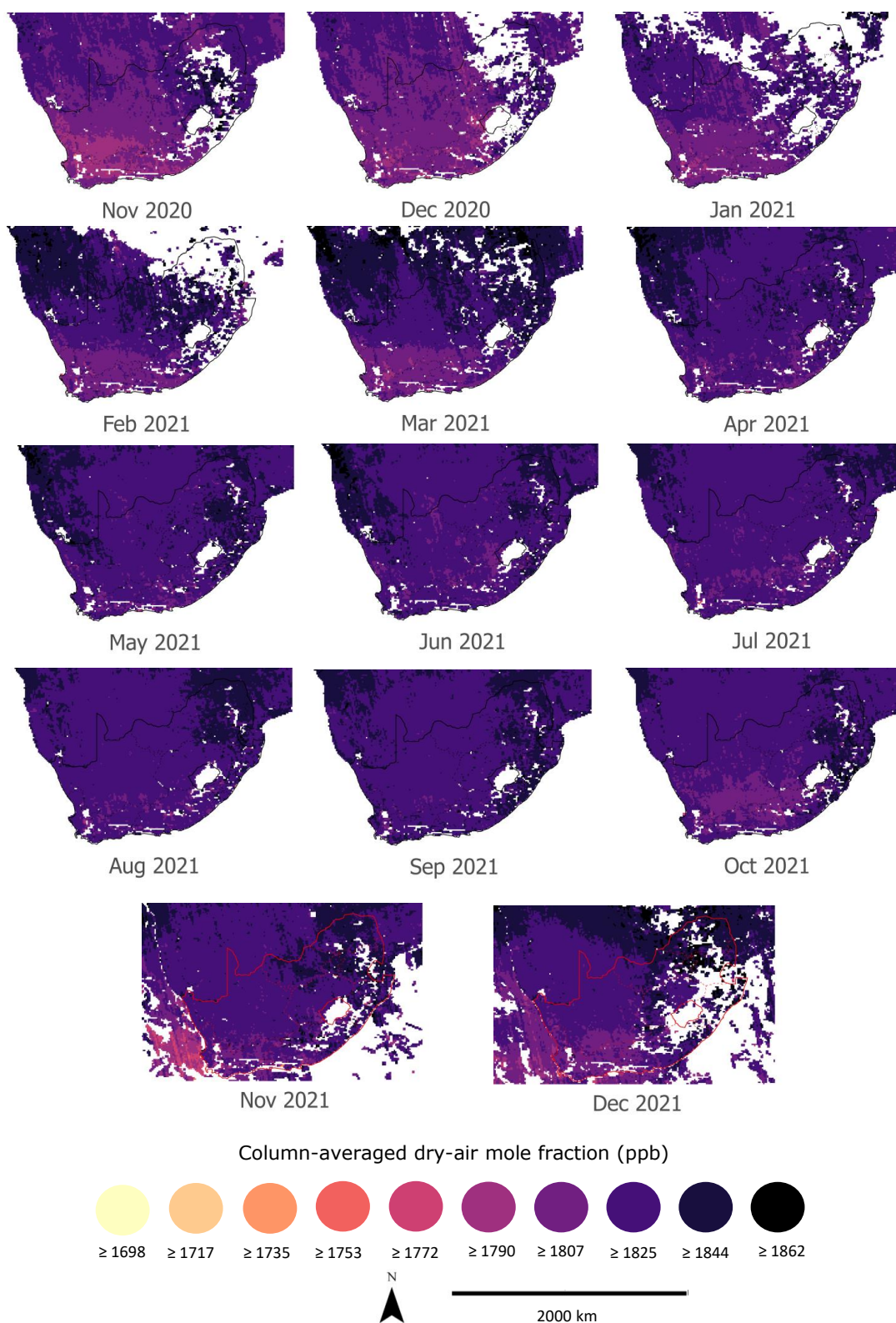


Figure 9. Concentrations of atmospheric column-averaged dry-air mole fractions of CH_4 in South Africa from November 2020 to December 2021

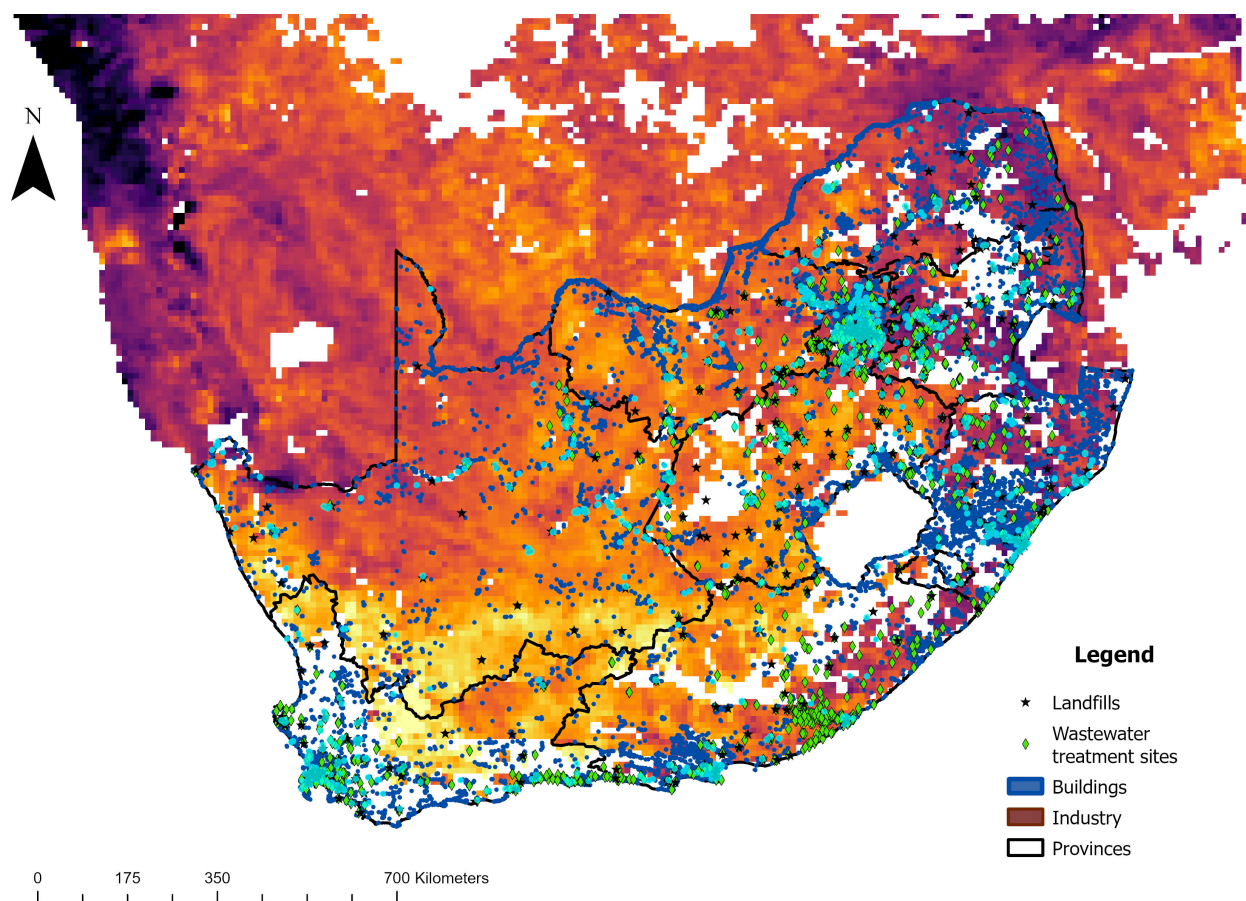


Figure 10. The overlay process highlighting that the areas of higher building density and industrial activity are in areas of higher CH_4 emissions.

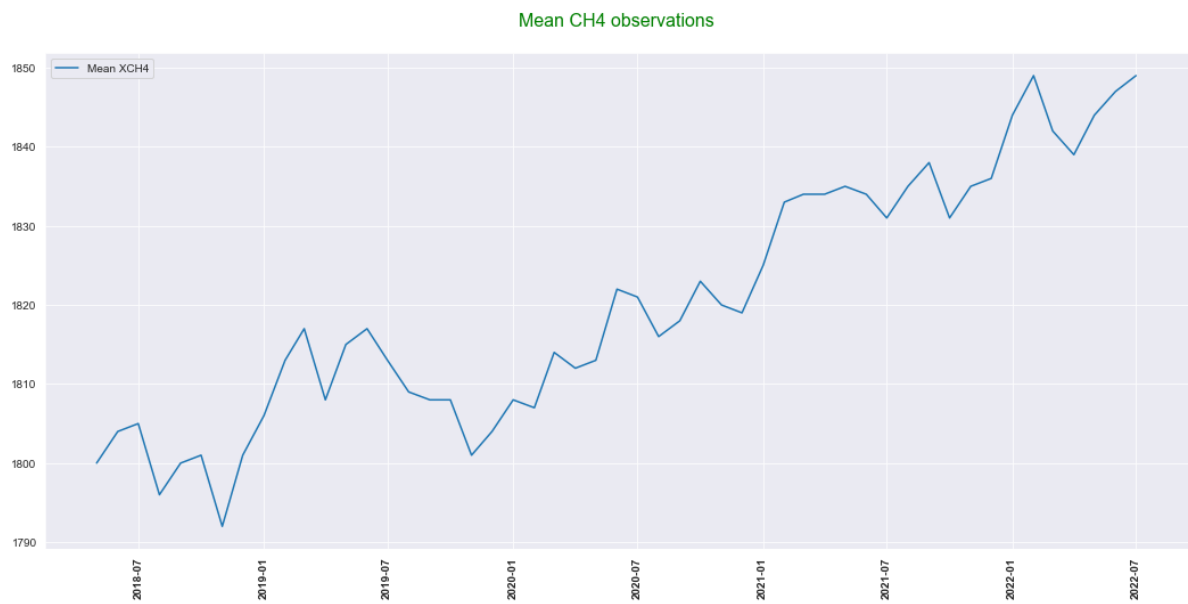


Figure 11. Monthly time-series analysis from May 2018 to July 2022 shows a steady increase of CH_4 emissions over the country in the four year period.

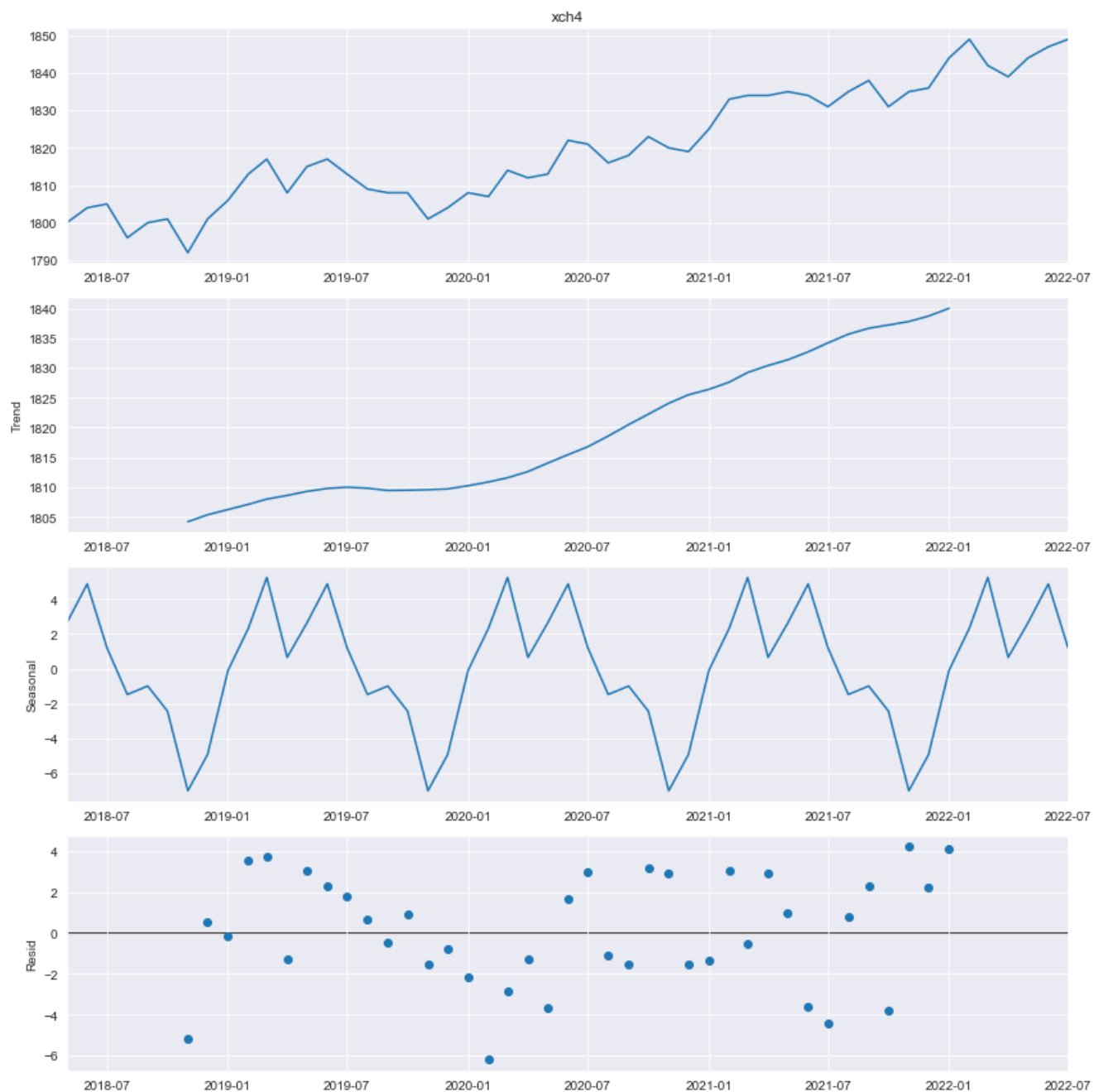


Figure 12. Seasonal decomposition of the data shows the data exhibits both trend and seasonality

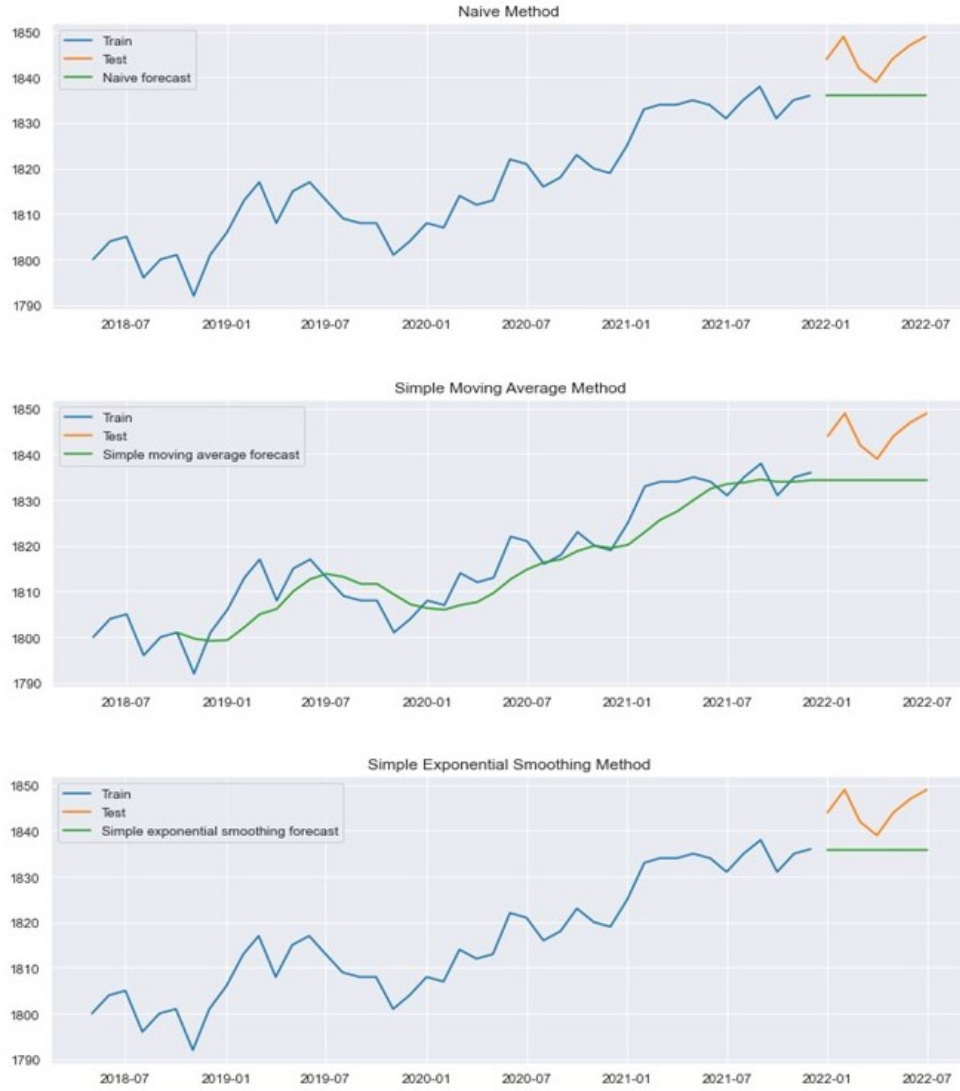


Figure 13. Graphical representations of mean monthly X_{CH_4} forecast using the baseline model, the naive model, the SMA and the SES models show the poor performance of the models where the monthly X_{CH_4} have been severely underestimated by the models.

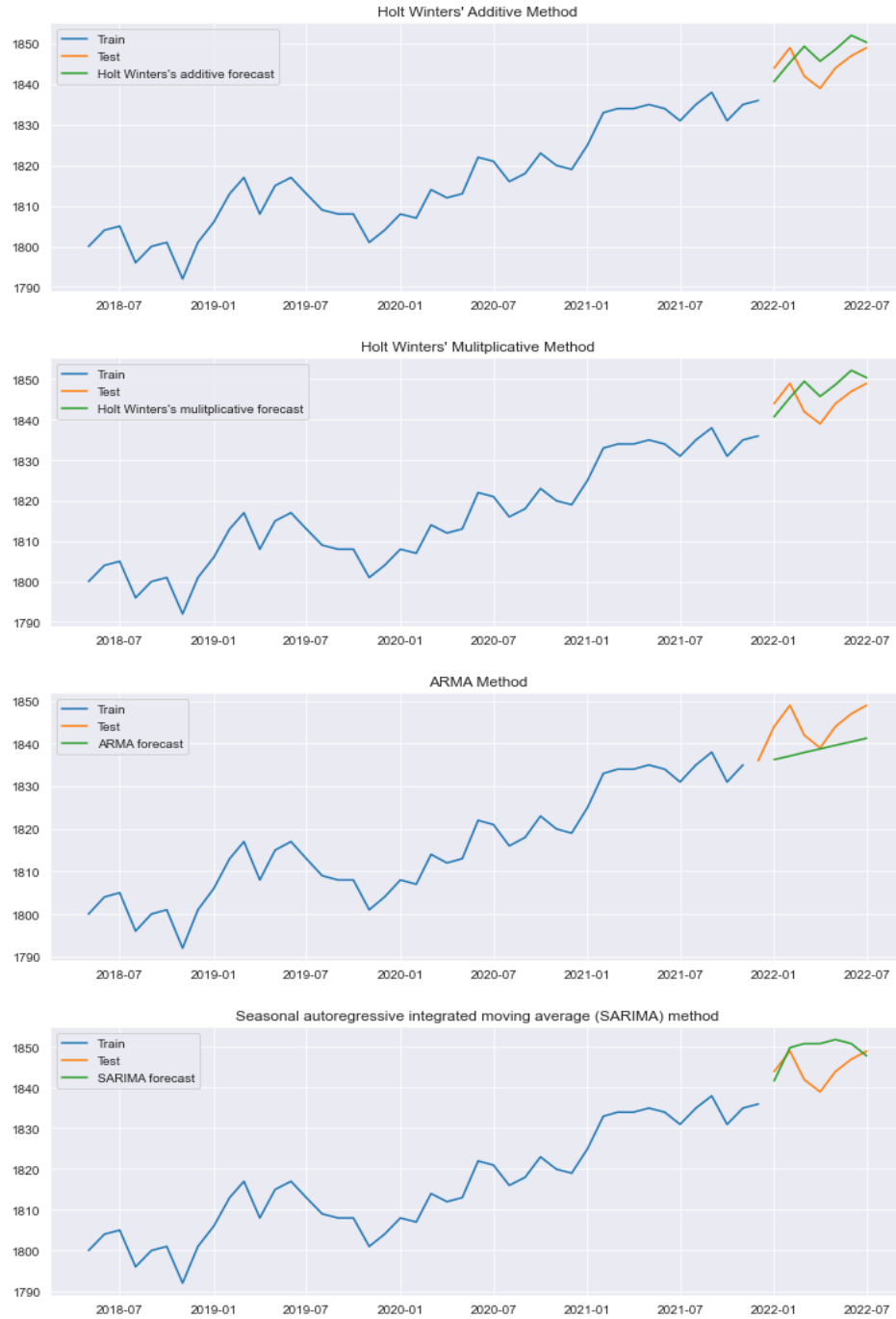


Figure 14. Graphical representations of mean monthly X_{CH_4} forecast using the HWA, HWM, ARIMA and SARIMA models show the better performance of the models.

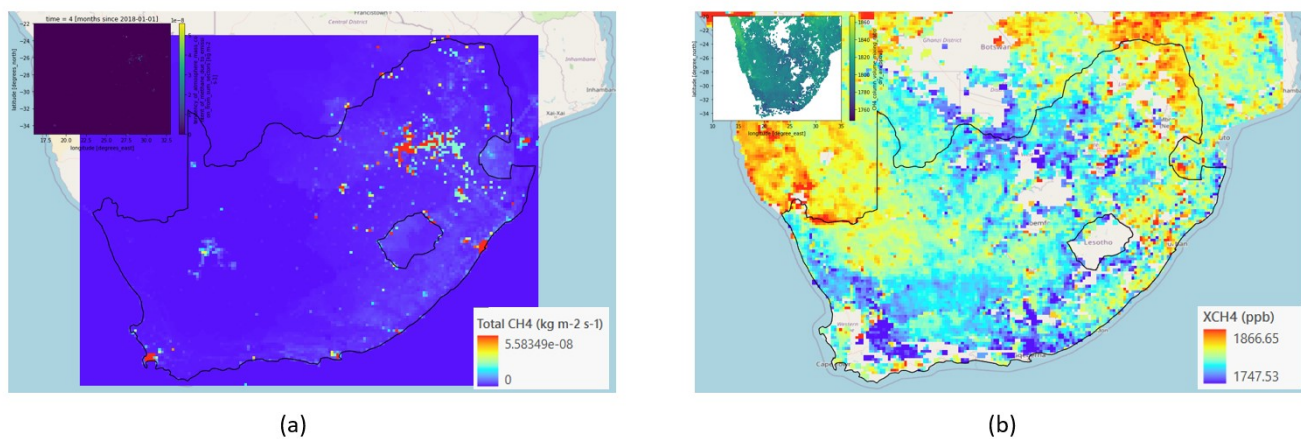


Figure 15. Correlation between the EDGAR inventory (a) and TROPOMI (b) averaged observations for May 2018.

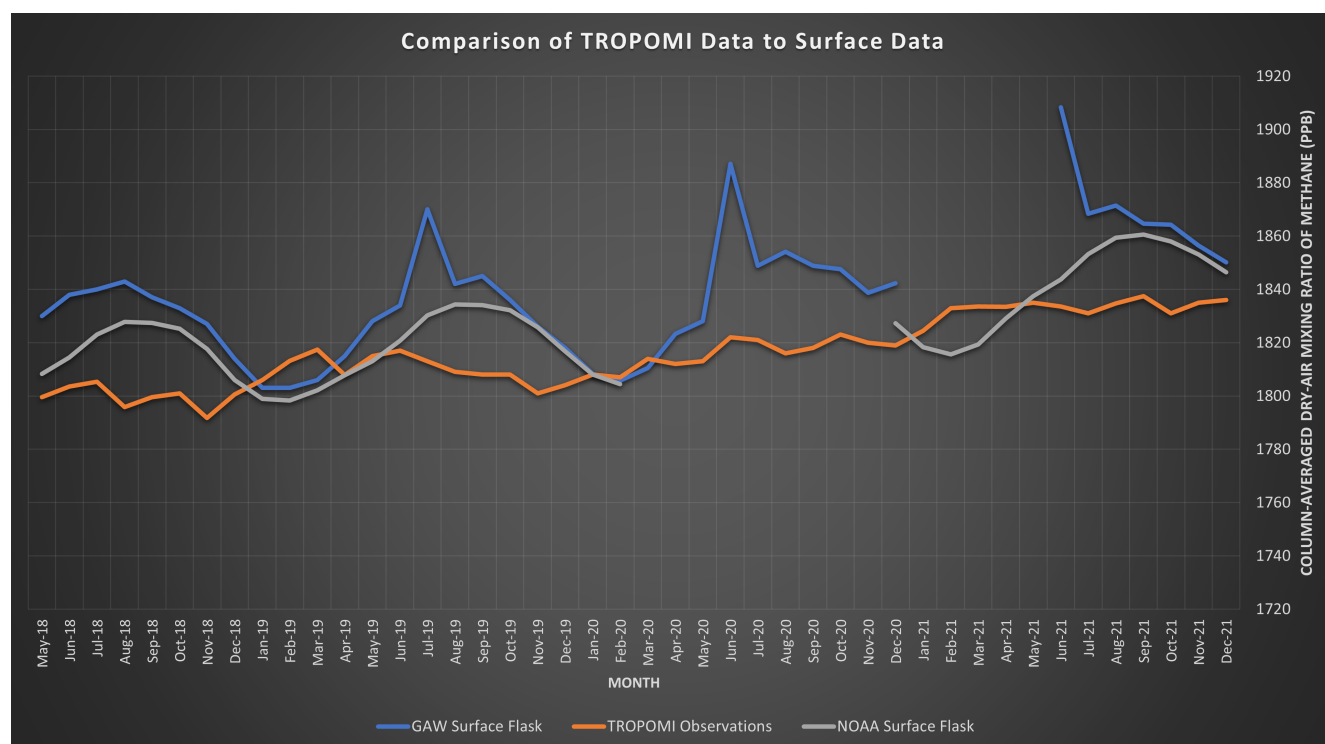


Figure 16. Line plot of TROPOMI measurements being compared to GAW and NOAA surface measurements of CH_4 .

6 Discussion

Both the spatial and the temporal analysis confirmed the rise in CH_4 emissions. However, there was also an expectation to see a decline in CH_4 emissions during the pandemic. Contrarily, emissions continued to rise steadily according to previous trends and seasonal fluctuations. In consideration of this, transportation was one of the only sources of emissions that were reduced significantly during the pandemic. However, we could not expect reduced levels of CH_4 from reduced transportation activity in South Africa. Car ownership levels in South Africa, as mentioned in section 4.2, are very low in comparison to the USA for example, whose proportion of vehicles was 800 vehicles per 1,000 people in the same year. Therefore the impact of reduced mobility could not have been enough to reduce emissions, as the biggest contributors of CH_4 emission, agriculture, landfills and wastewater treatment continued to operate under the guise of "essential goods and services"

There was also an expectation to notice higher emissions estimates observed in spring, in line with a study by Shikwambana et al. (2021a) that identified the months of September, October and November as having the most significant recorded measurements of CH_4 in the country. This has not been identified by this research, as the seasonal decomposition of the time-series showed instead the highest emissions occurring in the autumn months every year. The seasonal component showed the lowest measurements are experienced in October followed by a sharp increase November. This continues until February and decline again every year towards June, continuing to drop until November.

It remains to be addressed in future works if measurements from a CTM would yield result which are closer to Shikwambana et al. (2021a) or to this research. Especially as there are several variables that affect continuous monitoring of emissions and their concentrations, cloud cover and precipitation being unequal throughout a particular day for examples. It has also been made clear that the success of a TROPOMI retrieval is highly influenced by cloud cover. And even though a posterior correction accounts for most of the bias in the albedo, there is still a need for further investigation as it affects the accuracy of observed

emissions. It is therefore important to note that this research did not incorporate the above in the estimations.

The forecasting accuracy of 75% could also be improved. The forecasting models researched were all a-priori. In light of that, it is not possible to say if a model dependant on social, economic, etc, variables would have been more effective. Further other studies (e.g (Tudor and Sova, 2021)) trained their time-series models on data aging back to 1970, whereas this study trained the models with only 44 observations. Even though the forecast were deemed satisfactory, it can be expected that observations of a longer time span may increase the model accuracy.

One of the goals of TROPOMI is to provide consistent records of atmospheric concentrations on CH_4 with the accuracy and temporal resolution that enable inverse modelling of CH_4 sources on weekly scales. Until an inversion has been performed, only estimations can be achieved as effects of chemical transport in the atmosphere have to be accounted for before accurate measurements of what's on the ground or at the bottom of the atmosphere can be achieved. However, some important limitations with the TROPOMI data have to be addressed. The data gaps that result from the quality filter and the effects of the albedo on the observed CH_4 values are problematic and do not inspire confidence when working with the product.

7 Conclusions and Recommendations

The research for this dissertation analysed 3,417 satellite images, each weighing between 65 and 87 megabytes and including the outputs and maps produced totalling 296 gigabytes. The workflow which comprised of a combination of automation tools, models and batch processes in ArcGIS Pro, and script writing in Python to enabled insights to be drawn from simple cartographic techniques. Python was the workhorse of the project, the following packages having carried out most of the work; `sentinelsat`, `harp`, `xarray`, `pandas` and `statsmodels`.

This dissertation makes two important contributions to existing literature on greenhouse gas emissions and climate policy support. The first contribution is the focus on CH_4 for predictive modelling within the current discourse on the renewed growth of CH_4 emissions in the atmosphere. It took off from norm where predictions were made for CO_2 and sums of greenhouse gases in their entirety, and continued to extend the statistical methods employed for time series of other gases to CH_4 . The second contribution is on the use of Python packages to carry out the time series forecasting, where R and Matlab are mostly utilized in literature.

The dissertation had the broad aim of assessing TROPOMI's capability in supporting South Africa's climate policy by accurately mapping CH_4 emission quantities and their evolution through time as well as for identifying hotspots for quick identification of dangerous plumes over the country. The research has shown that TROPOMI can constrain the emissions of CH_4 and is useful in the identification of hotspots. But the aim of whether the estimations provided by the satellite are accurate have not been met. This remains a challenge until an inverse analysis of the emissions have been performed to account for climate, wind and other weather conditions that result in the mixing of CH_4 particles in the air. In addition to the great disparity between surface flask measurements, there is the challenge of spatial coverage with South Africa being represented by only one station in the global monitoring network. Three out of the four objectives were met with satisfactory results. Both the spatial analysis and the temporal analysis showed the growth of CH_4 , at a rate of 12 *ppb* per year, which is in

line with global measurements. Further, high emission activities have corresponded to areas of high emissions across the country. And lastly, the HWA time series model has predicted the future observations of TROPOMI for South Africa with an adequate of accuracy of 75%, which could potentially be improved by machine learning models.

A zero correlation was found between the TROPOMI observations and the EDGAR inventory of the same time period, where a positive correlation was expected. The objective of validation of the TROPOMI against an EDGAR was also met, however with unsatisfactory results and therefore necessitating the conversion of TROPOMI X_{CH_4} to fluxes. The limitation for the conversion was access to South African wind data which could not be delivered within the project timeline. There is still scope for this exercise to be carried out in future work.

As it has now been established that time-series based models for forecasting of CH_4 work satisfactorily, future work can look at creating a warning system for South Africa and use forecasts to see if the country can meet its targets based on the current state of affairs. Machine learning based time-series forecasting models are also an interest area for the future.

References

- Alexe, M., Bergamaschi, P., Segers, A., Detmers, R., Butz, A., Hasekamp, O., Guerlet, S., Parker, R., Boesch, H., Frankenberg, C., Scheepmaker, R. A., Dlugokencky, E., Sweeney, C., Wofsy, S. C., and Kort, E. A. (2015). Inverse modelling of CH₄ emissions for 2010-2011 using different satellite retrieval products from GOSAT and SCIAMACHY. *Atmospheric Chemistry and Physics*, 15(1):113–133.
- Averchenkova, A., Gannon, K. E., and Curran, P. (2019). Governance of climate change policy: A case study of South Africa. Technical Report November, Grantham Research Institute on Climate Change and the Environment and Centre for Climate Change Economics and Policy, London School of Economics and Political Science, London.
- Bergamaschi, P., Corazza, M., Karstens, U., Athanassiadou, M., Thompson, R. L., Pison, I., Manning, A. J., Bousquet, P., Segers, A., Vermeulen, A. T., Janssens-Maenhout, G., Schmidt, M., Ramonet, M., Meinhardt, F., Aalto, T., Haszpra, L., Moncrieff, J., Popa, M. E., Lowry, D., Steinbacher, M., Jordan, A., O’Doherty, S., Piacentino, S., and Dlugokencky, E. (2015). Top-down estimates of European CH₄ and N₂O emissions based on four different inverse models. *Atmospheric Chemistry and Physics*, 15(2):715–736.
- Bergamaschi, P., Danila, A., Weiss, R. F., Ciais, P., Thompson, R. L., Brunner, D., Levin, I., Meijer, Y., Chevallier, F., Bovensmann, H., Crisp, D., Basu, S., Dlugokencky, E., Engelen, R., Gerbig, C., Günther, D., Hammer, S., Henne, S., Houweling, S., Peylin, P., Pinty, B., Ramonet, M., Reimann, S., Röckmann, T., Schmidt, M., Strogies, M., Sussams, J., Tarasova, O., van Aardenne, J., Vermeulen, A. T., and Vogel, F. (2018). Atmospheric monitoring and inverse modelling for verification of greenhouse gas inventories. Technical Report August, European Commission Joint Research Centre, Ispra.
- Bhatti, U. A., Yan, Y., Zhou, M., Ali, S., Hussain, A., Qingsong, H., Yu, Z., and Yuan, L. (2021). Time series analysis and forecasting of air pollution particulate matter (pm_{2.5}): An sarima and factor analysis approach. *IEEE access*, 9:41019–41031.
- Bousquet, P., Pierangelo, C., Bacour, C., Marshall, J., Peylin, P., Ayar, P. V., Ehret, G., Bréon, F., Chevallier, F., Crevoisier, C., Gibert, F., Rairoux, P., Kiemle, C., Armante, R., Bès, C., Cassé, V., Chinaud, J., Chomette, O., Delahaye, T., Edouart, D., Estève, F., Fix, A., Friker, A., Klonecki, A., Wirth, M., Alpers, M., and Millet, B. (2018). Error budget of the methane remote

- lidar mission and its impact on the uncertainties of the global methane budget. 123(20):11,766–11,785.
- Bousserez, N., Henze, D., Rooney, B., Perkins, A., J Wecht, K., J Turner, A., Natraj, V., and Worden, J. (2016). Constraints on methane emissions in north america from future geostationary remote-sensing measurements. *Atmospheric chemistry and physics*, 16(10):6175–6190.
- Box, G. E. P., Jenkins, G. M., Reinsel, G. C., and Ljung, G. M. (2016). *Time series analysis: forecasting and control*. Wiley series in probability and statistics. Wiley, New York, 5th ed edition.
- Brink, J. and Stander, H. (2016). Perspectives on the future of personal transport in south africa : transportation engineering. *Civil engineering (Johannesburg, South Africa)*, 24(8):18–24.
- Buchwitz, M., Schneising, O., Reuter, M., Heymann, J., Krautwurst, S., Bovensmann, H., Burrows, J. P., Boesch, H., Parker, R. J., Somkuti, P., Detmers, R. G., Hasekamp, O. P., Aben, I., Butz, A., Frankenberg, C., and Turner, A. J. (2017). Satellite-derived methane hotspot emission estimates using a fast data-driven method. *Atmospheric Chemistry and Physics*, 17(9):5751–5774.
- Butz, A., Orphal, J., Checa-Garcia, R., Friedl-Vallon, F., Von Clarmann, T., Bovensmann, H., Hasekamp, O., Landgraf, J., Knigge, T., Weise, D., Sqalli-Houssini, O., and Kemper, D. (2015). Geostationary emission explorer for europe (g3e): Mission concept and initial performance assessment. *Atmospheric measurement techniques*, 8(11):4719–4734.
- Casey, N. H. (2021). A profile of South African sustainable animal production and greenhouse gas emissions. *Animal Frontiers*, 11(4):7–16.
- Crippa, M., Oregionni, G., Guizzard, D., Muntean, M., Schaaf, E., Lo Vullo, E., Solazzo, E., Monforti-Ferrario, F., and Olivier, J. (2019). *Fossil CO₂ and GHG emissions of all world countries - 2019 Report*. Publications Office of the European Union.
- de Gouw, J. A., Veeffkind, J. P., Roosenbrand, E., Dix, B., Lin, J. C., Landgraf, J., and Levelt, P. F. (2020). Daily Satellite Observations of Methane from Oil and Gas Production Regions in the United States. *Scientific Reports*, 10(1):1–10.
- Durbach, I., Merven, B., and McCall, B. (2017). Expert elicitation of autocorrelated time series with application to e3 (energy-environment-economic) forecasting models. *Environmental modelling and software : with environment data news*, 88:93–105.

- Ehret, G., Bousquet, P., Pierangelo, C., Alpers, M., Millet, B., Abshire, J. B., Bovensmann, H., Burrows, J. P., Chevallier, F., Ciais, P., Crevoisier, C., Fix, A., Flamant, P., Frankenberg, C., Gibert, F., Heim, B., Heimann, M., Houweling, S., Hubberten, H. W., Jöckel, P., Law, K., Löw, A., Marshall, J., Agustí-Panareda, A., Payan, S., Prigent, C., Rairoux, P., Sachs, T., Scholze, M., and Wirth, M. (2017). MERLIN: A French-German space LiDAR mission dedicated to atmospheric methane. *Remote Sensing*, 9(10):1–29.
- Hakkarainen, J., Szelağ, M. E., Ialongo, I., Retscher, C., Oda, T., and Crisp, D. (2021). Analyzing nitrogen oxides to carbon dioxide emission ratios from space: A case study of Matimba Power Station in South Africa. *Atmospheric Environment: X*, 10(March).
- Hopali, E. and Cakmak, A. (2020). Prediction of daily co2 emissions of a factory using arima and holt-winters seasonal methods. *International journal of information, business and management*, 12(3):51–63.
- Howells, M. and Solomon, M. (2000). South african greenhouse gas energy emissions inventory for the year 1994. *Journal of Energy in Southern Africa*, 11(4):272–278.
- Hoyer, S. and Hamman, J. (2017). xarray: N-D labeled arrays and datasets in Python. *Journal of Open Research Software*, 5(1).
- Hu, H., Landgraf, J., Detmers, R., Borsdorff, T., Aan de Brugh, J., Aben, I., Butz, A., and Hasekamp, O. (2018). Toward Global Mapping of Methane With TROPOMI: First Results and Intersatellite Comparison to GOSAT. *Geophysical Research Letters*, 45(8):3682–3689.
- Jacob, D. J. (2011). *Introduction to Atmospheric Chemistry*. Princeton University Press.
- Jacob, D. J., Turner, A. J., Maasakkers, J. D., Sheng, J., Sun, K., Liu, X., Chance, K., Aben, I., McKeever, J., and Frankenberg, C. (2016). Satellite observations of atmospheric methane and their value for quantifying methane emissions. *Atmospheric Chemistry and Physics*, 16(22):14371–14396.
- Jacob, D. J., Varon, D. J., Cusworth, D. H., Dennison, P. E., Frankenberg, C., Gautam, R., Guanter, L., Kelley, J., McKeever, J., Ott, L. E., Poulter, B., Qu, Z., Thorpe, A. K., Worden, J. R., and Duren, R. M. (2022). Quantifying methane emissions from the global scale down to point sources using satellite observations of atmospheric methane. *Atmospheric Chemistry and Physics*, pages 9617–9646.

- Kort, E. A., Frankenberg, C., Costigan, K. R., Lindenmaier, R., Dubey, M. K., and Wunch, D. (2014). Four corners: The largest US methane anomaly viewed from space. *Geophysical Research Letters*, 41(19):6898–6903.
- Krautwurst, S., Gerilowski, K., Jonsson, H. H., Thompson, D. R., Kolyer, R. W., Iraci, L. T., Thorpe, A. K., Horstjann, M., Eastwood, M., Leifer, I., Vigil, S. A., Krings, T., Borchardt, J., Buchwitz, M., Fladeland, M. M., Burrows, J. P., and Bovensmann, H. (2017). Methane emissions from a Californian landfill, determined from airborne remote sensing and in situ measurements. *Atmospheric Measurement Techniques*, 10(9):3429–3452.
- Lan, X., Dlugokencky, E., Mund, J., Crotwell, A., Crotwell, M., Moglia, E., Madronich, M., Neff, D., and Thoning, K. (2022). Atmospheric Methane Dry Air Mole Fractions from the NOAA GML Carbon Cycle Cooperative Global Air Sampling Network, 1983-2021.
- Linghu, J., Chen, J., and Yan, Z. (2020). Research on forecasting coal bed methane demand and resource allocation system based on time series. *Energy exploration and exploitation*, 38(5):1467–1483.
- Liu, M., van der A, R., van Weele, M., Eskes, H., Lu, X., Veefkind, P., de Laat, J., Kong, H., Wang, J., Sun, J., Ding, J., Zhao, Y., and Weng, H. (2021). A New Divergence Method to Quantify Methane Emissions Using Observations of Sentinel-5P TROPOMI. *Geophysical Research Letters*, 48(18):1–10.
- Lunt, M. M., Palmer, P. P., Feng, L., Taylor, C. C., Boesch, H., and Parker, R. R. (2019). An increase in methane emissions from tropical Africa between 2010 and 2016 inferred from satellite data. *Atmospheric Chemistry and Physics*, 19(23):14721–14740.
- Malik, A., Hussain, E., Baig, S., and Khokhar, M. F. (2020). Forecasting CO₂ emissions from energy consumption in Pakistan under different scenarios: The China–Pakistan Economic Corridor. *Greenhouse Gases: Science and Technology*, 10(2):380–389.
- McElroy, M. B. (2012). ionosphere and magnetosphere. publisher: Britannica.
- Mkololo, T., Labuschagne, C., and Joubert, W. (2022). Atmospheric CH₄ at Cape Point by South African Weather Service, dataset published as CH₄_{CPT_ssurface – insitu}_{SAWS}_{data}_{atWDCGG}.
- Neelin, J. D. (2011). *Climate Change and Climate Modeling*. Cambridge University Press, Cambridge.

- Nesme, N., Foucher, P. Y., Doz, S., Lezeaux, O., and Camy-Peyret, C. (2021). Comparison of estimation methods to quantify methane plume concentration at high spatial resolution from hyperspectral images. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 43(B3-2021):411–417.
- Okello, N. O., Okello, T. W., and Zunckel, M. (2020). Public perceptions of air quality status and suggestions for improvement: The case of Richards Bay and its surroundings, uMhlathuze Local Municipality, South Africa. *Clean Air Journal*, 30(1):1–10.
- Ostermann, A., Fabel, Y., Ouan, K., and Koo, H. (2022). Forecasting charging point occupancy using supervised learning algorithms. *Energies (Basel)*, 15(9):3409.
- Palmer, P. I., O’Doherty, S., Allen, G., Bower, K., Bösch, H., Chipperfield, M. P., Connors, S., Dhomse, S., Feng, L., Finch, D. P., Gallagher, M. W., Gloor, E., Gonzi, S., Harris, N. R. P., Helfter, C., Humpage, N., Kerridge, B., Knappett, D., Jones, R. L., Le Breton, M., Lunt, M. F., Manning, A. J., Matthiesen, S., Muller, J. B. A., Mullinger, N., Nemitz, E., O’Shea, S., Parker, R. J., Percival, C. J., Pitt, J., Riddick, S. N., Rigby, M., Sembhi, H., Siddans, R., Skelton, R. L., Smith, P., Sonderfeld, H., Stanley, K., Stavert, A. R., Wenger, A., White, E., Wilson, C., and Young, D. (2018). A measurement-based verification framework for UK greenhouse gas emissions: an overview of the Greenhouse gAs Uk and Global Emissions (GAUGE) project. *Atmospheric Chemistry and Physics*, 18(16):11753–11777.
- Pandey, S., Gautam, R., Houweling, S., Van Der Gon, H. D., Sadavarte, P., Borsdorff, T., Hasekamp, O., Landgraf, J., Tol, P., Van Kempen, T., Hoogeveen, R., Van Hees, R., Hamburg, S. P., Maasakkers, J. D., and Aben, I. (2019). Satellite observations reveal extreme methane leakage from a natural gas well blowout. *Proceedings of the National Academy of Sciences of the United States of America*, 116(52):26376–26381.
- Pleijel, H., Karlsson, P. E., and Simpson, D. (2009). Ozone and Methane - Climate and Environment Connected. In Pleijel, H., editor, *Air Pollution and Climate change. Two Sides of the Same Coin?*, chapter 4, page 168. Swedish Environmental Protection Agency.
- Puliafito, S. E., Berná, L., Lopez-Noreña, A., Pascual, R., and Bolaño-Ortiz, T. (2020). Atmospheric Methane Emissions for Argentina. Comparison With TROPOMI Satellite Measurements. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, IV-3/W2-2020(March):107–112.

- Qader, M. R., Khan, S., Kamal, M., Usman, M., and Haseeb, M. (2021). Forecasting carbon emissions due to electricity power generation in bahrain. *Environmental science and pollution research international*, 29(12):17346–17357.
- Qu, Z., Jacob, D. J., Shen, L., Lu, X., Zhang, Y., Scarpelli, T. R., Nesser, H., Sulprizio, M. P., Maasakkers, J. D., Bloom, A. A., Worden, J. R., Parker, R. J., and Delgado, A. L. (2021). Global distribution of methane emissions: A comparative inverse analysis of observations from the TROPOMI and GOSAT satellite instruments. *Atmospheric Chemistry and Physics*, 21(18):14159–14175.
- Saunois, M., Stavert, A. R., Poulter, B., Bousquet, P., Canadell, J. G., Jackson, R. B., Raymond, P. A., Dlugokencky, E. J., Houweling, S., Patra, P. K., Ciais, P., Arora, V. K., Bastviken, D., Bergamaschi, P., Blake, D. R., Brailsford, G., Bruhwiler, L., Carlson, K. M., Carrol, M., Castaldi, S., Chandra, N., Crevoisier, C., Crill, P. M., Covey, K., Curry, C. L., Etiope, G., Frankenberg, C., Gedney, N., Hegglin, M. I., Höglund-Isaksson, L., Hugelius, G., Ishizawa, M., Ito, A., Janssens-Maenhout, G., Jensen, K. M., Joos, F., Kleinen, T., Krummel, P. B., Langenfelds, R. L., Laruelle, G. G., Liu, L., Machida, T., Maksyutov, S., McDonald, K. C., McNorton, J., Miller, P. A., Melton, J. R., Morino, I., Müller, J., Murguía-Flores, F., Naik, V., Niwa, Y., Noce, S., O’Doherty, S., Parker, R. J., Peng, C., Peng, S., Peters, G. P., Prigent, C., Prinn, R., Ramonet, M., Regnier, P., Riley, W. J., Rosentreter, J. A., Segers, A., Simpson, I. J., Shi, H., Smith, S. J., Steele, L. P., Thornton, B. F., Tian, H., Tohjima, Y., Tubiello, F. N., Tsuruta, A., Viovy, N., Voulgarakis, A., Weber, T. S., van Weele, M., van der Werf, G. R., Weiss, R. F., Worthy, D., Wunch, D., Yin, Y., Yoshida, Y., Zhang, W., Zhang, Z., Zhao, Y., Zheng, B., Zhu, Q., Zhu, Q., and Zhuang, Q. (2020). The global methane budget 2000–2017. *Earth System Science Data*, 12(3):1561–1623.
- Schneising, O., Buchwitz, M., Reuter, M., Bovensmann, H., Burrows, J. P., Borsdorff, T., Deutscher, N. M., Feist, D. G., Griffith, D. W., Hase, F., Hermans, C., Iraci, L. T., Kivi, R., Landgraf, J., Morino, I., Notholt, J., Petri, C., Pollard, D. F., Roche, S., Shiomi, K., Strong, K., Sussmann, R., Velasco, V. A., Warneke, T., and Wunch, D. (2019). A scientific algorithm to simultaneously retrieve carbon monoxide and methane from TROPOMI onboard Sentinel-5 Precursor. *Atmospheric Measurement Techniques*, 12(12):6771–6802.
- Schneising, O., Buchwitz, M., Reuter, M., Vanselow, S., Bovensmann, H., and P. Burrows, J. (2020).

- Remote sensing of methane leakage from natural gas and petroleum systems revisited. *Atmospheric Chemistry and Physics*, 20(15):9169–9182.
- Scholes, R. J. and van der Merwe, M. R. (1996). Greenhouse gas emissions from south africa. *South African Journal of Science*, 92:220–222.
- Seinfeld, J. H. (2016). *Atmospheric Chemistry and Physics : From Air Pollution to Climate Change*. John Wiley and Sons, Incorporated, third edition.
- Seymore, R., Inglesi-Lotz, R., and Blignaut, J. (2014). A greenhouse gas emissions inventory for South Africa: A comparative analysis. *Renewable and Sustainable Energy Reviews*, 34:371–379.
- Sheng, J. X., Jacob, D. J., Maasakkers, J. D., Zhang, Y., and Sulprizio, M. P. (2018). Comparative analysis of low-Earth orbit (TROPOMI) and geostationary (GeoCARB, GEO-CAPE) satellite instruments for constraining methane emissions on fine regional scales: Application to the Southeast US. *Atmospheric Measurement Techniques*, 11(12):6379–6388.
- Shikwambana, L., Mhangara, P., and Kganyago, M. (2021a). Assessing the relationship between economic growth and emissions levels in south africa between 1994 and 2019. *Sustainability (Switzerland)*, 13(5):1–15.
- Shikwambana, L., Mhangara, P., and Mbatha, N. (2020). Trend analysis and first time observations of sulphur dioxide and nitrogen dioxide in South Africa using TROPOMI/Sentinel-5 P data. *International Journal of Applied Earth Observation and Geoinformation*, 91(May):102130.
- Shikwambana, L., Mokgoja, B., and Mhangara, P. (2022). A Qualitative Assessment of the Trends, Distribution and Sources of Methane in South Africa. *Sustainability*, 14(6):3528.
- Shikwambana, L., Ncipha, X., Sangeetha, S. K., Sivakumar, V., and Mhangara, P. (2021b). Qualitative study on the observations of emissions, transport and the influence of climatic factors from sugarcane burning: A south african perspective. *International Journal of Environmental Research and Public Health*, 18(14).
- Stevens, L. (2021). The 7th National Greenhouse Gas Inventory Report. Technical report, Department of Forestry, Fisheries and the Environment, Pretoria.
- Swari, M., Qusyairi, M., Mandyartha, E., and Wahanani, H. (2021). Business intelligence system using simple moving average method (case study : Sales medical equipment at pt. semangat sejahtera bersama). *Journal of Physics: Conference Series*, 1899(1):12121.

- Tan, H., Zhang, L., Lu, X., Zhao, Y., Yao, B., Parker, R. J., and Boesch, H. (2022). An integrated analysis of contemporary methane emissions and concentration trends over China using in situ and satellite observations and model simulations. *Atmospheric Chemistry and Physics*, 22(2):1229–1249.
- Thi Quynh Nguyen, T., Takeuchi, W., Misra, P., and Sachiko, H. (2021). Technical note: Emission mapping of key sectors in Ho Chi Minh City, Vietnam, using satellite-derived urban land use data. *Atmospheric Chemistry and Physics*, 21(4):2795–2818.
- Tudor, C. (2016). Predicting the evolution of co2 emissions in bahrain with automated forecasting methods. *Sustainability*, 8(9).
- Tudor, C. and Sova, R. (2021). Benchmarking ghg emissions forecasting models for global climate policy. *Electronics (Basel)*, 10(24):3149.
- Tudor, C. and Sova, R. (2022). Eu net-zero policy achievement assessment in selected members through automated forecasting algorithms. *ISPRS international journal of geo-information*, 11(4):232.
- Turner, A. J., Frankenberg, C., Wennberg, P. O., and Jacob, D. J. (2017). Ambiguity in the causes for decadal trends in atmospheric methane and hydroxyl. *Proceedings of the National Academy of Sciences of the United States of America*, 114(21):5367–5372.
- Turner, A. J., Jacob, D. J., Wecht, K. J., Maasackers, J. D., Lundgren, E., Andrews, A. E., Biraud, S. C., Boesch, H., Bowman, K. W., Deutscher, N. M., Dubey, M. K., Griffith, D. W., Hase, F., Kuze, A., Notholt, J., Ohyama, H., Parker, R., Payne, V. H., Sussmann, R., Sweeney, C., Velasco, V. A., Warneke, T., Wennberg, P. O., and Wunch, D. (2015). Estimating global and North American methane emissions with high spatial resolution using GOSAT satellite data. *Atmospheric Chemistry and Physics*, 15(12):7049–7069.
- UNFCCC Secretariat (2012). Report of the Conference of the Parties on its Seventeenth Session. In *Part One: Proceedings*, number March, Durban, South Africa. United Nations Framework Convention on Climate Change.
- Upadhyay, P., Czerkawski, M., Davison, C., Cardona, J., Macdonald, M., Andonovic, I., Michie, C., Atkinson, R., Papadopoulou, N., Nikas, K., and Tachtatzis, C. (2022). A flexible multi-temporal and multi-modal framework for sentinel-1 and sentinel-2 analysis ready data. *Remote sensing (Basel, Switzerland)*, 14(5):1120.

- Varon, D. J., Jacob, D. J., Sulprizio, M., Estrada, L., Downs, W. B., Shen, L., Hancock, S. E., Nesser, H., Qu, Z., Penn, E., Chen, Z., Lu, X., Lorente, A., Tewari, A., and Randles, C. A. (2022). Integrated methane inversion (imi 1.0): A user-friendly, cloud-based facility for inferring high-resolution methane emissions from tropomi satellite observations. *Geoscientific Model Development Discussions*, 2022:1–31.
- Veefkind, J. P., Aben, I., McMullan, K., Förster, H., de Vries, J., Otter, G., Claas, J., Eskes, H. J., de Haan, J. F., Kleipool, Q., van Weele, M., Hasekamp, O., Hoogeveen, R., Landgraf, J., Snel, R., Tol, P., Ingmann, P., Voors, R., Kruizinga, B., Vink, R., Visser, H., and Levelt, P. F. (2012). 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.
- Waskom, M. L. (2021). seaborn: statistical data visualization. *Journal of Open Source Software*, 6(60):3021.
- Wayne, R. P. (2000). *Chemistry of Atmospheres : An Introduction to the Chemistry of the Atmosphere of Earth, the Planets, and their Satellites*. Oxford University Press, Oxford, 3rd edition.
- Wecht, K. J., Jacob, D. J., Frankenberg, C., Jiang, Z., and Blake, D. R. (2014). Mapping of North American methane emissions with high spatial resolution by inversion of SCIAMACHY satellite data. *Journal of Geophysical Research: Atmospheres*, 119(12):7741–7756.
- Worden, J. R., Cusworth, D., Qu, Z., Yin, Y., Yuzhong Zhang, Bloom, A. A., Ma, S., Byrne, B., Scarpelli, T., Maasakkers, J. D., Crisp, D., Duren, R., and Jacob, D. J. (2021). The 2019 Methane Budget And Uncertainties At 1 Degree Resolution And Each Country 4 Through Bayesian Integration Of GOSAT Total Column Methane Data And A Priori 5 Inventory Estimates. *Atmospheric Chemistry and Physics Discussions*.
- World Meteorological Organization (2021). State of the Global Climate 2021 WMO Provisional report. Technical report, World Meteorological Organisation.
- Zeiml, S., Altendorfer, K., Felberbauer, T., and Nurgazina, J. (2019). Simulation based forecast data generation and evaluation of forecast error measures. In *2019 Winter Simulation Conference (WSC)*, volume 2019- of *WSC '19*, pages 2119–2130. IEEE Press.
- Zhou, W., Zhao, Y., Zhang, G., and Shen, X. (2020). Harp: Holistic analysis for refactoring python-based analytics programs. In *Proceedings of the ACM/IEEE 42nd International Conference on*

Software Engineering, ICSE '20, page 506–517, New York, NY, USA. Association for Computing Machinery.