Aviation H_2O and NO_x climate cost functions based on local weather

MSc thesis

J. van Manen

Faculty of Aerospace Engineering
Department of Control and Operations
Section Aircraft Noise and Climate Effects

TU Delft

Aviation H_2O and NO_x climate cost functions based on local weather

MSc thesis

by

J. van Manen

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Tuesday April 25, 2017 at 15:00.

Student number:	4001729	
Project duration:	June 15, 2016 – April 25, 2017	
Thesis committee:	Prof. Dr. V. Grewe,	TU Delft (ANCE), DLR, supervisor
	Dr. M. Snellen,	TU Delft (ANCE)
	Dr. Ir. A. Sahai,	TU Delft (ANCE)
	Dr. Ir. W. J. C. Verhagen,	TU Delft (ATO)

Cover image by C. Baron (`https://unsplash.com/@crisbaron`).
Colours for plots by C. A. Brewer (`http://colorbrewer2.org`).
An electronic version of this thesis is available at `http://repository.tudelft.nl/`.

TU Delft

Preface

This thesis forms the final part of my MSc Aerospace Engineering programme at TU Delft, and is worth 42 ECTS credits. A separate literature review on the same topic was previously completed for 12 ECTS credits. My MSc track was Flight Performance and Propulsion (FPP) with a focus on gas turbines, which I combined with Honours courses on the environmental impact of aviation. I am grateful that I was allowed to work on my thesis within the Aircraft Noise and Climate Effects (ANCE) group. Perhaps this work is better seen as a culmination of my Honours track than my MSc.

In fact this document could be placed in a much wider frame. My BSc minor focussed on sustainable design and engineering, and around the same time my father began working at Clean Sky. If the ANCE group had existed when I started my MSc three years ago, I would probably have enrolled in it over FPP, but as it is I used the Honours curriculum to fill in the gaps. For Honours credits I performed a two-month internship at DLR for Dr Grewe, whom I had approached during his guest lectures for Prof. Simons' new Aircraft Emissions and Climate Effects course. When I started this thesis he was a Visiting Professor, but by now he is a Full Professor in the ANCE department.

This thesis has been written as a stand-alone document, but the scope of the work was defined within the EU project ATM4E (`http://atm4e.eu/`) and the two cannot entirely be separated. Work Package 1 of ATM4E comprises formulation of novel noise and local air quality Environmental Change Functions (ECFs), formulation of algorithmic climate ECFs from REACT4C (`http://www.react4c.eu/`) data, and integration into a common format. Investigating algorithmic ECFs for water vapour and NO_x emissions is precisely the scope of this MSc thesis, while contrails were concurrently researched by the University of Reading. At key stages of this thesis, knowledge of the implementation of results almost certainly influenced the direction taken. Thus a little scientific value will have been lost to gain a lot of practical relevance. Note that the term Climate Cost Function (CCF) from REACT4C was generalised to Environmental Change Function in ATM4E, as cost suggests monetary impact and the scope is expanded to local air quality and noise impacts. For clarity, this thesis will use the old terminology.

I would like to thank Prof. Grewe for his supervision over my internship and thesis. If our paths had not crossed I would feel more like an engineer than a scientist, and would potentially have a different career path. Thanks Simon for being in the unique position of listening to my frustrations and understanding the subject material. Thanks to my friends, mother, father, brother, girlfriend, officemates, housemates, and anyone else who had to witness one of my rants. Finally, good luck to the entire ATM4E project team!

J. van Manen
Den Haag, 7 April 2017

Abstract

Aviation contributes significantly to anthropogenic global warming, and one promising possibility for mitigation is climate-optimised routing. For the REACT4C project a novel approach was used to simulate the variation of aviation water vapour and NO_x emission climate impact with location and weather patterns, but this is too computationally expensive to apply beyond initial research. Results showed about 10% climate impact reduction from a 1% cost increase. For implementation of climate-optimised routing, algorithms are needed which will allow climate impact to be estimated in real-time from weather predictions. This research focuses on formulating algorithmic approximations of aviation water vapour and NO_x emission climate impact based on local weather data by systematically examining correlations between climate impact data and weather data at the time of emission in the REACT4C dataset.

The methods and models used for generating the REACT4C data are assessed in detail down to their first publications and potential errors and omissions are identified. The analysis is split into direct water vapour, short-lived ozone from NO_x, and methane from NO_x climate impact. Long-lived ozone and stratospheric water vapour from methane effects are neglected. The water vapour and NO_x ozone and methane Climate Cost Function (CCF) results from REACT4C are reverse-engineered to the original grid they were emitted from to prevent inflation of statistical power. Weather and chemistry data at the time of emission are interpolated to the same grid for regression analysis. Literature reviews are used to identify causal predictors and derived variables. A variety of statistical tools are applied to assess variability of the CCFs and search for the best predictors. Four algorithms are developed for each species, using zero-dimensional instantaneous regression analysis. A tailored trade-off framework is applied to choose the best algorithm for application.

The chosen algorithmic CCF for water vapour emissions is linear with potential vorticity and has an adjusted R^2 of 0.59. Both the mean and the variance of the water vapour climate impact appear strongly determined by the altitude of an emission relative to the tropopause. The relationship between water vapour CCF results and emission altitude is investigated to critically reflect and expand on results from a previous publication.

The chosen algorithmic CCF for ozone is bilinear with geopotential and temperature plus their interaction and has an adjusted R^2 of 0.42. Ozone climate impact appears moderately determined by altitude and temperature of the emission location. The relationship between ozone CCF results, background NO_x concentration and latitude during emission is investigated to critically reflect and expand on results from a previous publication.

The chosen algorithmic CCF for methane is bilinear with geopotential and the solar incidence, and has an adjusted R^2 of 0.17. Methane climate impact has low variability and is relatively independent of weather at the time and location of emission. The relationship between methane CCF results and background NO_x concentrations during emission is investigated to critically reflect and expand on results from a previous publication. Methane climate impact can be more accurately predicted by using simulated ozone climate impact, but the variance left unexplained by the ozone algorithm would lead to worse results in application. The correlation between methane and ozone is weaker than in previous studies. Chemical concentrations, lightning frequency, and lightning NO_x production at the time and location of emission do not predict aviation NO_x climate impact beyond the extent of basic meteorology unless a large amount of predictors are included in the regression.

The chain of models and assumptions from basic climate science to algorithmic CCFs is assessed to identify relative effects on uncertainty of the results. Several steps are identified that should be revisited and several opportunities for future data analyses to increase understanding and certainty of algorithmic CCFs. Future steps for research into and application of algorithmic CCFs depend on upcoming verification activities for the results presented here.

Contents

List of Figures

List of Tables

Acronyms

AGTP Absolute Global Temperature Change Potential

AGW Anthropogenic Global Warming

AGWP Absolute Global Warming Potential

ATM Air Traffic Management

ATM4E Air Traffic Management for the Environment

ATR Average Temperature Response

ATTILA Atmospheric Tracer Transport in a LAgrangian model

CCF Climate Cost Function

CCM Climate-Chemistry Model

COSMO Consortium of Small-scale MOdels

CWF Chemical Weather Forecasting

DLR Deutsches Zentrum für Luft- und Raumfahrt

E39C ECHAM4/L39 + CHEM

EA East Atlantic

ECHAM5 ECMWF-Hamburg v5

ECMWF European Centre for Medium-range Weather Forecasts

EI Emission Index

EMAC ECHAM/MESSy Atmospheric Chemistry

ETS Emissions Trading Scheme

GCM General Circulation Model

GHG Greenhouse Gas

GTP Global Temperature Potential

GWP Global Warming Potential

ICAO International Civil Aviation Organisation

IFS Integrated Forecasting System

IPA Institut für Physik der Atmosphäre

IPCC Intergovernmental Panel on Climate Change

IQR Interquartile Range

ISA International Standard Atmosphere

KDE Kernel Density Estimation

LOSU Level of Scientific Understanding

LTO Landing and Take-Off

MECCA Module Efficiently Calculating the Chemistry of the Atmosphere

MESSy Modular Earth Submodel System

NAFC North Atlantic Flight Corridor

NAO North Atlantic Oscillation

NMHC Non-Methane Hydrocarbon

NOAA National Oceanic and Atmospheric Administration

NWP Numerical Weather Prediction

PMO Primary-Mode Ozone

PV Potential Vorticity

REACT4C Reducing Emissions from Aviation by Changing Trajectories for the benefit of Climate

RF Radiative Forcing

RRTM Rapid and accurate Radiative Transfer Model

RSE Residual Standard Error

SWV Stratospheric Water Vapour

UNFCCC United Nations Framework Convention on Climate Change

UTLS Upper Troposphere/Lower Stratosphere

VOC Volatile Organic Compounds

WMO World Meteorological Organisation

WP Weather Pattern

Chemical symbols

CH_4 Methane

CO Carbon monoxide

CO_2 Carbon dioxide

H_2O Water vapour

HNO_3 Nitric acid

HO_2 Hydroperoxy radical

N_2O Nitrous oxide

NO Nitric oxide

NO_2 Nitrogen dioxide

NO_3 Nitrate

NO_x Nitrogen oxides (NO and NO_2)

NO_y All reactive species containing nitrogen and oxygen

$O(^1D)$ Atomic oxygen in excited state

$O(^3P)$ Atomic oxygen in ground state

O_2 Molecular oxygen

O_3 Ozone

OH Hydroxyl radical

PAN Peroxy acetyl nitrate

RH Alkanes

SO_2 Sulphur dioxide

1 Introduction

1.1. Background

The popular debate continues, but the scientific consensus on Anthropogenic Global Warming (AGW) is well-established (e.g. Cook et al., 2013). An estimated 3.5% of anthropogenic Radiative Forcing (RF) relative to pre-industrial times was caused by aviation in 2005 (Lee et al., 2009, 2010). Including the relatively uncertain contribution of contrail-induced cirrus the estimate increases to 4.9%. Simple forecasts indicate that the share of aviation could grow to 4–4.7% of anthropogenic RF excluding contrail-induced cirrus by 2050 (Lee et al., 2009).

Depending on the emission studied, various mitigation options for aviation climate impact are being pursued, e.g. biofuels, taxes, and combustion technologies. For short-lived species such as water vapour, NO_x, and contrails one promising development is climate-optimised Air Traffic Management (ATM) (Lee et al., 2010) as the impact of these emissions shows strong spatiotemporal variation.

The European research project Reducing Emissions from Aviation by Changing Trajectories for the benefit of Climate (REACT4C) (`http://www.react4c.eu/`) investigated the feasibility and mitigation gain from such measures. A novel modelling chain was introduced that coupled a climate model, an emissions model, and an ATM model (Grewe et al., 2014b). Eight distinct weather patterns were identified for the case study in the North Atlantic Flight Corridor (NAFC) (Irvine et al., 2013), for which the 3-D (4-D originally) spatial variation in globally-averaged climate impact per mass emitted (i.e. Climate Cost Functions) was calculated for carbon dioxide, water vapour, and NO_x emissions, and per distance flown for contrail formation (Grewe et al., 2014b). Grewe et al. (2014a) presented results of the optimisation for one of these weather patterns, showing that up to 25% climate impact reduction is possible for only 0.5% higher operating costs. Roughly two million CPU-hours were needed (`http://delta.tudelft.nl/article/reduce-climate-impact-by-flying-more/30948`). The other seven REACT4C patterns were subsequently simulated. Grewe et al. (2017b) presented results from optimising air traffic for all eight patterns with the climate metric used in this thesis, and find that 10% reduction can be achieved for 1% higher operating costs, or lower operating costs if market-based measures for emission reduction are introduced. One separate day was simulated using the same methodology within the Deutsches Zentrum für Luft- und Raumfahrt (DLR) project WeCare (Grewe et al., 2017a), with an enlarged domain and finer resolution.

1.2. Objective

Besides the results of the ATM optimisation for the first REACT4C weather pattern, Grewe et al. (2014a) presented intermediate results in the form of CCF contours, shown here for a NO_x emission in Figure 1.1. For the purpose of this thesis these are the final results of the method in Grewe et al. (2014b). The ozone CCF results show high values in the south, coinciding with the jet stream, and low values around 30°W, 60°N, coinciding with the low-pressure system (visible from the trough, the southwards bend in the geopotential height). This suggests that the variability of the ozone CCF can be partially predicted from the weather situation at the time of emission. The colour scales are the same in both plots, which indicates that the variability of the cooling effect from methane is significantly lower than

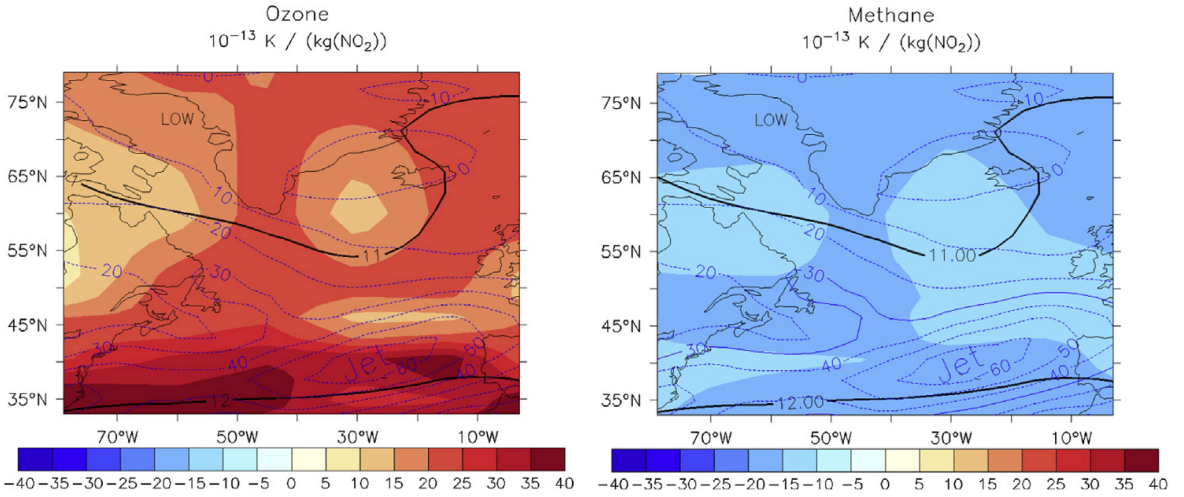


Figure 1.1: Climate Cost Function (CCF) results for ozone (left) and methane (right) from a NO_x emission at 12 UTC, 200 hPa, Weather Pattern (WP)1. Black isolines show the geopotential height in [km], blue dotted isolines show the wind velocity in [m s^{-1}]. The climate metric is F-ATR20 in [10^{-13} K $kg(NO_2)^{-1}$], the global-mean temperature response over 20 years to local aviation given a strategic re-routing decision and growth of air traffic (see Section 3.7). The image is adapted from Grewe et al. (2014a).

that of the warming effect from ozone. With this colour scale no pattern is visible with the jet stream or geopotential for the methane CCF data.

If climate-optimised ATM is to become feasible, CCFs must be calculated real-time and the computational costs reduced drastically. The full set of REACT4C results presents an opportunity to investigate and derive approximations from model weather data, that introduce some extra uncertainty to the process but facilitate real-time calculation in a Numerical Weather Prediction (NWP) routine. Positive results bring real-life application of climate-optimised ATM one step closer.

The research objective is to formulate algorithmic approximations of aviation water vapour and NO_x emission climate impact based on local weather data by systematically analysing variations in climate impact and local weather data at the time of emission across eight simulated datasets for distinct weather patterns over the North Atlantic from the REACT4C project.

In the REACT4C approach a NO_x emission leads to short-lived ozone, methane, and Primary-Mode Ozone (PMO) (long-lived ozone) climate effects (Grewe et al., 2014b). PMO results are linearly scaled from methane results, thus they are not meaningful to analyse. Ozone and methane climate effects vary considerably in their mechanisms and time-scales (Lee et al., 2010), so for this thesis they may be seen as separate objectives that are simply recombined. Several metrics (see Section 3.7) are calculated in Grewe et al. (2014b) but currently only F-ATR20 is available. The three sub-objectives for this research are thus:

- Develop a water vapour CCF algorithm
- Develop an ozone CCF algorithm
- Develop a methane CCF algorithm

1.3. Structure

Chapter 2 gives a more detailed introduction to radiation, climate and the impact of aviation. Details on how water vapour and NO_x CCFs were calculated in REACT4C (Grewe et al., 2014b) are provided in Chapter 3. The methodology of this thesis, i.e. data preparation and statistics, is described in Chapter 4. Results for water vapour algorithmic CCFs are presented in Chapter 5, including a discussion on the final algorithm. As ozone and methane effects result from the same NO_x emission, results for the ozone and methane algorithmic CCFs are both presented in Chapter 6, including discussions on the final algorithms and the relationship between both effects. Chapter 7 discusses the methodology and results on a higher level. Chapter 8 presents conclusions on the research and recommendations for further analysis.

2

Climate change and aviation

This chapter introduces the concepts of Radiative Forcing (RF) and Anthropogenic Global Warming (AGW) (Section 2.1), provides details on the various emissions and climate impacts of aviation (Section 2.2), and discusses climate mitigation strategies for aviation (Section 2.3).

2.1. Climate change

2.1.1. Radiation

Radiation is one the three fundamental modes of heat transfer, and the one responsible for the Earth being suitable for life. Due to its high temperature from nuclear fusion, the Sun emits a large amount of shortwave (ultraviolet) radiation. The Earth absorbs a large amount of the shortwave radiation incident on it (at day-time), which forms an equilibrium with the smaller amount of longwave (infrared) radiation the Earth itself emits due to its much lower temperature (at both day and night-time). The gases and aerosols that make up the Earth's atmosphere tip the balance to a significantly higher (hence habitable) surface temperature by reflecting, absorbing and emitting specific wavelengths of radiation, popularly called the greenhouse effect. Note that this is a false analogy: greenhouses primarily block cooling through convection (Wood, 1909).

Any increase or decrease in the concentrations of Greenhouse Gases (GHGs) and aerosols (including clouds) of the Earth's atmosphere thus leads to a radiative imbalance and this is commonly called Radiative Forcing (RF). Note that all changes to the radiative balance of the system cause RF, e.g. changes in incident solar radiation and albedo changes from ice caps melting. Emissions, anthropogenic or not, that cause RF directly or indirectly through formation of other GHGs all have an atmospheric lifetime, which can vary by several orders of magnitude and the RF decreases proportionally to the amount of the GHG remaining. Physically, RF depends on the characteristics of the radiation that the emitted or formed GHG absorbs - e.g. the concentrations of other GHGs and their radiation spectra. Thus a doubling of GHG concentrations does not necessarily cause a doubling of the greenhouse effect.

2.1.2. Surface temperature change

If the radiation balances changes, it takes time for the Earth system to find a new balance at a different temperature where it again emits as much radiation as it absorbs. The concept of relating temperature to RF is symbolised by the climate sensitivity parameter λ, which relates global-mean RF from a sustained change in GHG concentration to (long-term) equilibrium surface temperature change:

$$\Delta T_s^{eq} \approx \lambda \cdot RF \tag{2.1}$$

The parameter λ was originally thought to be independent of the species being studied, but this only holds for (meridionally) homogeneous RF (Lee et al., 2010). For inhomogeneous sustained concentration changes from aviation, the climate sensitivity parameter depends on the species and its spatial distribution, but is the same order of magnitude as for carbon dioxide (Ponater et al., 2006). If the GHG concentration change from an emission is not sustained but temporary, or the time horizon is

too short, the Earth system will not reach a new equilibrium temperature and transient effects must be considered.

There are multiple ways to define RF. Instantaneous RF keeps all atmospheric temperatures constant as they tend to change to correct the radiation imbalance, thus it represents the instantaneous result of an emission. Adjusted RF allows the stratospheric temperature to change and is more representative of surface temperature changes, at least for carbon dioxide and ozone (Myhre et al., 2013). The sign and size of λ are much more dependent on the perturbation if instantaneous RF is used (Hansen et al., 1997), thus using adjusted RF facilitates calculation of temperature change. Whenever the term RF is used without any qualifier in this thesis, adjusted RF is meant.

To compare the climate impact of non-CO_2 emissions, the Intergovernmental Panel on Climate Change (IPCC) (e.g. Myhre et al., 2013) prescribes usage of GWP100, the Global Warming Potential over 100 years. This is calculated from the integrated RF over 100 years of the emission, divided by the integrated RF over 100 years of an equal mass of carbon dioxide. This is commonly used to sum all GHG emissions into CO_2-equivalent emissions. The Global Temperature Potential (GTP) is an alternative metric that relates temperature change from an emission to the temperature change from an equal mass of carbon dioxide.

The total accumulated anthropogenic RF in 2011, relative to 1750, was 2.29 [1.13–3.33] W m^{-2} (Myhre et al., 2013), of which 1.68 [1.33–2.03] W m^{-2} is estimated to be from carbon dioxide emissions. About 16% of the current 100-year integrated RF from anthropogenic emissions is caused by the transport sector (Fuglestvedt et al., 2008). An estimated 3.5% of anthropogenic RF relative to pre-industrial times was caused by aviation in 2005 (Lee et al., 2009, 2010), 4.9% including the relatively uncertain contribution of contrail-induced cirrus. The most recent IPCC report (Boucher et al., 2013; Myhre et al., 2013) found that the Level of Scientific Understanding (LOSU) for contrail-induced cirrus has increased and gives a similar estimate to Lee et al. (2009). Simple forecasts indicate that the share of aviation could grow to 4–4.7% of anthropogenic RF excluding contrail-induced cirrus by 2050 (Lee et al., 2009). Of 197 parties to the convention, 132 have ratified the United Nations Framework Convention on Climate Change (UNFCCC) Paris Agreement (`http://unfccc.int/paris_agreement/items/9485.php`) as of February 2017. Its central aim is to limit AGW to 2 °C relative to pre-industrial times and pursue efforts to limit the increase to 1.5 °C.

The scientific consensus on AGW is undeniable (Cook et al., 2013), the societal importance of staying under 2 °C is shared by the majority of governments, and aviation and its alternatives are responsible for a significant and growing share of climate change. There is a clear need to better understand and reduce the climate impact of aviation.

2.2. Aviation emissions and effects

The first major international assessment of mechanisms and magnitude of aviation climate impact was the IPCC Special Report on Aviation (Penner et al., 1999). Lee et al. (2010) provided a comprehensive update of the IPCC report, that serves as a current reference for aviation climate effects. Figure 2.1 gives a visual overview of the relative contributions of the major families of aircraft emissions, which are discussed in the following sections.

2.2.1. Carbon dioxide

Carbon dioxide is the main chemical product of fossil fuel combustion, and the most important historical cause of AGW relative to pre-industrial times in general (Myhre et al., 2013). Typically, burning 1 kg of kerosene produces 3160 ± 60 g of CO_2 (Lee et al., 2010), expressed as an Emission Index (EI) of 3160 g kg^{-1}. Carbon dioxide is a very stable GHG, that typically stays in the atmosphere for several hundreds of years (Myhre et al., 2013), and as such its climate impact may be assumed to be independent of time and location of emission.

2.2.2. NO_x

NO_x, or nitrogen oxides (NO and NO_2), is a chemical product of combustion of fossil fuels with air. The ambient nitrogen reacts with ambient oxygen once all the fuel has combusted due to the high temperatures needed for efficient combustion of that fuel. The most important determinant for NO_x formation is the combustor flame temperature, which is highest for stoichiometric combustion (Saravanamuttoo et al., 2009). The EI is 12–17 g kg^{-1} (Lee et al., 2010), but NO_x emissions have insignificant direct

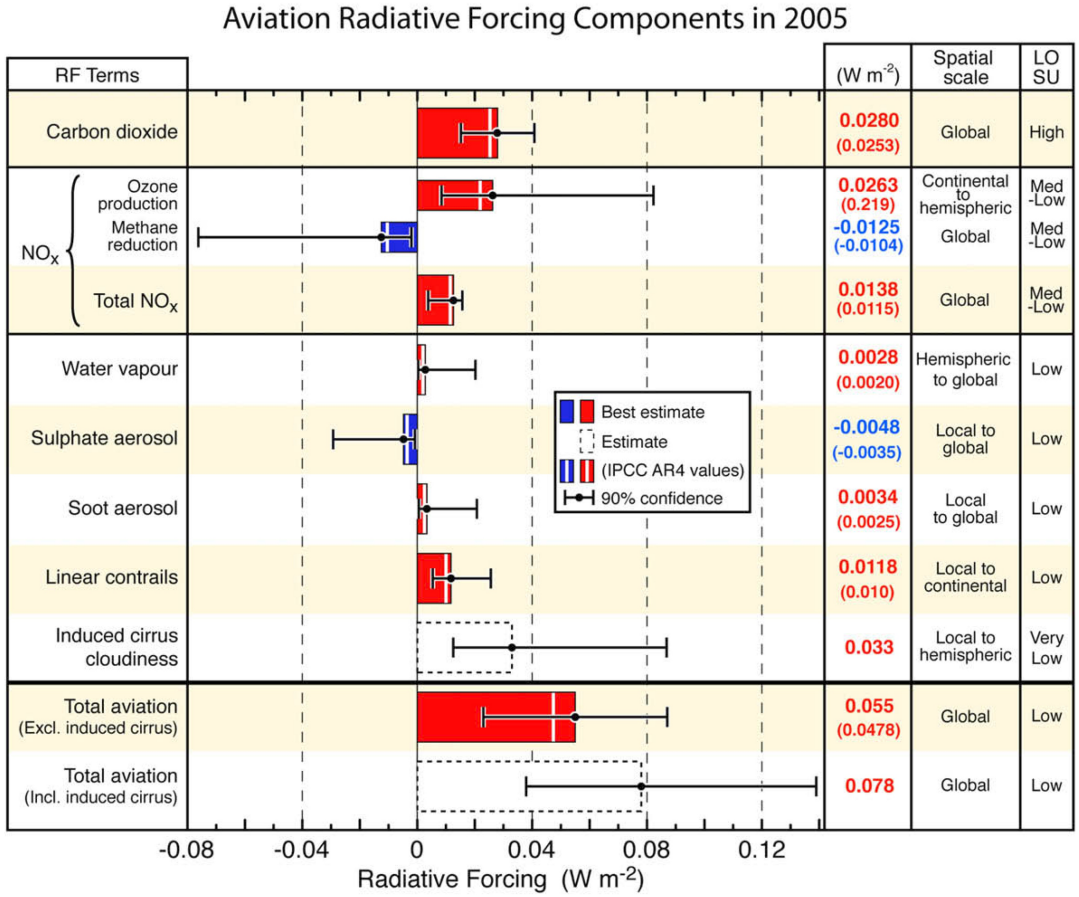


Figure 2.1: Aviation RF in 2005 per emission type from emissions between 1850 and 2005. IPCC AR4 values for white line and values in brackets are from Forster et al. (2007). Median values and error bars for total NO_x and total aviation were based on Monte Carlo simulations in Lee et al. (2009). For total NO_x, ozone and methane impacts were assumed 100% correlated. Level of Scientific Understanding (LOSU) on the right is based on judgement in Lee et al. (2009). The image is from Lee et al. (2009).

climate impact. Their climate impact lies in a complex series of photochemical reactions that depend on background mixing ratios of the reacting species, temperature, pressure, and availability of sunlight (Lee et al., 2010).

In the short term, an atmospheric NO_x emission leads to an increased rate of ozone (O_3) production. Ozone is an effective short-lived GHG so this causes a short-term warming effect. In the longer term, NO_x emissions lead to increased mixing ratios of OH, which enhances oxidation of methane (CH_4). Methane is a long-lived GHG so this causes a long-term cooling effect. The enhanced oxidation of methane also leads to decreased concentrations of ozone and stratospheric water vapour (Myhre et al., 2013), thus the actual cooling effect is amplified.

Figure 2.2 shows the main reaction pathways for a NO_x emission. An increase in atmospheric NO_x (from an aircraft emission) shifts the chemical balance towards increased NO_2 and OH. The NO_2 reacts to form ozone (if light is available), and the OH reacts to oxidise methane among others (CO, O_2, Non-Methane Hydrocarbons (NMHCs)). The atmospheric lifetime of the NO_x emission is determined by formation of HNO_3, which can be washed out. PAN is important for long-range transport of nitrogen compounds. More details on the relevant reactions are given in Sections 3.4.3 and 3.4.4.

The integrated impacts of the short and long-term components of NO_x climate impact depend on different factors and, by definition, timescales. The net effect of a NO_x emission thus not only depends on the location and time of the emission but also on the metric chosen to evaluate the climate impact. Globally, the net effect of aviation NO_x emissions is significantly warming (see Figure 2.1) and comparable to that of carbon dioxide.

Note that in Figure 2.1 the confidence bounds for combined NO_x climate impact are an order of magnitude smaller than for ozone or methane. Lee et al. (2009) assumed that ozone and methane

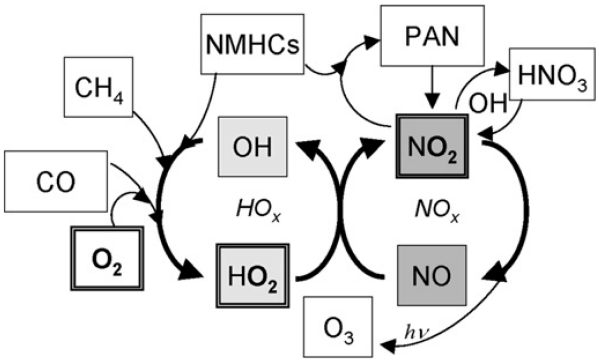


Figure 2.2: Schematic of the main reaction pathways between a NO_x emission and ozone and methane. The image is from Lee et al. (2010).

impacts are 100% correlated, thus their uncertainties are 100% anticorrelated, but noted that this is an approximation. Holmes et al. (2011) assessed uncertainties in aviation NO_x ozone and methane climate impact with an ensemble of published models and found an $R^2 = 0.79$ correlation.

Only about 1% of NO_x emissions originate from aviation (Berntsen and Isaksen, 1999). The major sources are surface fossil fuels, lightning, soils, and biomass burning. In the Upper Troposphere/Lower Stratosphere (UTLS) the only net sources are aviation and lightning (Berntsen and Isaksen, 1999) and they are of the same order of magnitude - here aviation is overrepresented compared to its total emissions due to their altitude. Lightning is the main source of NO_x at flight altitudes (Berntsen and Isaksen, 1999) and through non-linearity in NO_x chemistry it is a strong driver of aviation NO_x climate impact (Grewe et al., 2002).

NO_x emissions in the Landing and Take-Off (LTO) phase negatively affect local air quality directly and through the enhanced production of ozone (EPA, 1999).

2.2.3. Water vapour

Water vapour is the other main chemical product of fossil fuel combustion. Naturally abundant water vapour causes two-thirds of the Earth's natural greenhouse effect (Myhre et al., 2013). Typically, the EI from kerosene combustion is around 1240 g kg^{-1} (Lee et al., 2010). Water vapour is a mostly inert GHG, and its atmospheric lifetime strongly depends on the rate of removal through the hydrological cycle (e.g. rain). Variations in water vapour climate impact are thus driven by background humidity and precipitation rates. The current climate impact of aviation water vapour emissions is expected to be low, but future developments such as supersonic aviation (at higher altitudes) and hydrogen-powered aviation could have a significant impact through water vapour emissions (Lee et al., 2010). A new study by Wilcox et al. (2012b) simulated aviation water vapour impact in more detail than those included in Figure 2.1 and found 0.86 [0.34–1.34] mW m^{-2} annual-mean global-mean adjusted RF from 2006 aviation water vapour emissions. Not only is the mean estimate significantly lower than that in Figure 2.1, but the confidence bounds are an order of magnitude smaller.

2.2.4. Sulphate aerosols

Sulphate aerosols form shortly after emission from SO_2 emissions. The EI for gaseous SO_2 is in the range of 0.6–1.0 g kg^{-1} (Lee et al., 2010), which is consistent with kerosene sulphur contents. The direct greenhouse effect of sulphate aerosols is slightly cooling, but they have an uncertain role in the radiative properties of contrails and clouds (Lee et al., 2010). Sulphate aerosols in the LTO phase negatively affect local air quality (EPA, 1999).

2.2.5. Soot aerosols

Soot particles are a combination of organic carbon compounds and elemental carbon (C). Particles are typically 30–60 nm in size and are emitted with an EI of 0.01–0.05 g kg^{-1} (Lee et al., 2010). The direct greenhouse effect of soot aerosols is slightly warming, but they have an uncertain role in the radiative properties of contrails and clouds (Lee et al., 2010). Soot aerosols in the LTO phase negatively affect local air quality (EPA, 1999).

2.2.6. Contrails

Contrails are line-shaped cirrus clouds, formed by the mixing of hot, humid aircraft exhaust with the cold ambient air. The commonly accepted method (Lee et al., 2010) for determining whether contrails will form is the Schmidt-Appleman criterion (Schmidt, 1941; Appleman, 1953; Schumann, 1996). This determines, based purely on physical properties including the water vapour EI, whether or not a contrail will form. The precise composition and properties of a contrail depend, among others, on aerosol mixing ratios and the time-of-day. Generally speaking, during daylight contrails reflect incoming solar radiation and thus have a mostly cooling effect. During night-time contrails absorb the Earth's outgoing longwave radiation and thus have a mostly warming effect (Stuber et al., 2006). Globally, the net effect of contrails is significantly warming.

So-called linear persistent contrails are relatively well understood using the Schmidt-Appleman criterion, and their modelling can be verified by identifying contrails on satellite imagery (Mannstein et al., 2012). Some contrails, however, spread and become cirrus clouds. The calculation of this impact is very uncertain, but may be a stronger warming than all other aircraft climate impacts combined (see Figure 2.1). Several new studies since Lee et al. (2009) have increased the understanding of contrail-induced cirrus, leading the IPCC (Boucher et al., 2013; Myhre et al., 2013) to state that the LOSU for contrail-induced cirrus has increased. Burkhardt and Kärcher (2011) found 37.5 mW m^{-2} accumulated RF globally in 2002 from linear and cirrus effects. Bock and Burkhardt (2016) found similar results for 2002, and provide an updated estimate of 56 mW m^{-2} accumulated RF globally in 2006.

2.2.7. Other species

Kerosene-powered aircraft emit a multitude of other species in trace amounts, but these have negligible climate impact. Especially at low power settings (i.e. in parts of the LTO phase), aircraft engines can emit significant amounts of carbon monoxide (CO) and Volatile Organic Compounds (VOC), which negatively affect local air quality directly and through enhanced production of ozone (EPA, 1999).

2.3. Mitigation options for aviation

Mitigation options for reducing the climate impact of aviation are usually categorised as technological versus operational, but here different categories are used along the chain from emission to climate metric.

The most basic mitigation option is to reduce air traffic. Rail and road transport are more readily powered electrically by renewable sources and are viable alternative to short-haul aviation, and the societal benefit of aviation over surface transport is controversial. As aviation is exempt from the UNFCCC and the Paris Agreement, introducing environmental legislation is difficult. The European Emissions Trading Scheme (ETS) is a market-based measure that now includes aviation, but the resulting CO_2 price is currently so low that it has no meaningful impact. Following the Paris Agreement, 66 International Civil Aviation Organisation (ICAO) member states (`http://www.icao.int/environmental-protection/Pages/market-based-measures.aspx`) agreed upon a global market-based measure for aviation.

The second option is to reduce the amount of GHG emissions per passenger per distance flown. As carbon dioxide and water vapour are direct products of combustion, their emission can be reduced by improving fuel efficiency. Options in this area include improved aerodynamics and higher density cabins. The financial motive to reduce fuel consumption through more efficient combustion in the past decades has led to higher temperature combustion, which causes higher NO_x emissions and more contrail formation. Alternative fuels can be used, such as hydrogen which produces no carbon dioxide when emitted, or biofuels which cause zero net carbon dioxide emission due to biomass absorbing ambient carbon dioxide. Hydrogen fuels would require an extensive re-design of passenger aircraft and could increase water vapour and NO_x emissions. Biofuels present problems with large-scale availability, and in some forms compete with agriculture. Hydrogen and biofuel combustion produces (almost) no sulphur aerosols (Lee et al., 2010), but the effect this would have on contrail formation is thought to be negligible (Grewe et al., 2017b). Low NO_x kerosene combustion technology is being pursued, mainly by moderating the air-to-fuel ratio and temperature in the combustor.

The final option is to reduce the resultant climate impact for constant emissions or predictably varying emissions per flight condition. Previous work has often focussed (Lee et al., 2010) on destroying contrails or changing their ambient properties via technological means or fuel additives. One promising

research endeavour is to optimise air traffic not for cost, time, or emissions but net climate impact from all emissions. This is the focus of the REACT4C project, and this thesis focuses on the spatiotemporal variation of the climate impact per emission (CCFs) as a basis for this optimisation.

3

REACT4C simulation data

REACT4C (http://www.react4c.eu/) was a European Union project that ran from 2010-2014 and was coordinated by the DLR Institut für Physik der Atmosphäre (IPA). A novel modelling chain was introduced, that enabled ATM optimisation for minimum total climate impact for an air traffic routing decision as opposed to fuel burn or emissions. A test case was set up for one winter day in the NAFC, which was representative for one of eight typical weather patterns previously identified by Irvine et al. (2013). The method used was published by Grewe et al. (2014b) and the results were published separately by Grewe et al. (2014a). The other seven weather patterns were simulated later with some nuances to the set-up, which will be explained in this chapter. Grewe et al. (2017b) present results from optimising all eight patterns with the climate metric used in this thesis.

The chapter will summarise the method used for water vapour and NO_x CCFs calculation, as contrails, carbon dioxide and ATM implementation are beyond the scope of this thesis. The chapter is structured as follows: weather patterns (Section 3.1), base model set-up (Section 3.2), REACT4C-specific model set-up (Section 3.3), atmospheric processes (Section 3.4), sensitivity studies (Section 3.5), radiative forcing calculation (Section 3.6), climate metric calculation (CCFs) (Section 3.7), and verification (Section 3.8). For more details the reader is generally referred to Grewe et al. (2014b). Though PMO effects are not analysed in this thesis, Section 3.6.4 explains the scaling used in REACT4C.

3.1. Weather patterns

To facilitate calculation of the REACT4C CCFs, Irvine et al. (2013) presented a method for classifying distinct weather patterns in the NAFC. Several million CPU-hours were needed (Section 1.1) for the REACT4C research, and the total budget for the project was more than four million euros (http://cordis.europa.eu/project/rcn/93557_en.html, including other research). Irvine et al. (2013) state that the reason for identifying classifications is the computational cost of calculating the climate-optimised routes, which is partially due to the CCF calculation and partially due to the ATM optimisation. Though market-based measures can be applied to mitigate the cost increase from climate-optimised routing (Grewe et al., 2017b), an approach that requires millions of CPU-hours for every day would need to be very well parallelised for real-time calculation and would add significant cost.

Their inputs were 21 years of European Centre for Medium-range Weather Forecasts (ECMWF) ERA-Interim re-analysis data. The zonal wind in the NAFC from this dataset is averaged at 250 hPa (about 10 km), and the latitude with the highest zonal average is the latitude of the jet stream. Synoptic typing of weather patterns is achieved by comparing the geopotential height anomaly at 250 hPa to the North Atlantic Oscillation (NAO) and East Atlantic (EA) teleconnection patterns. The teleconnection pattern data are constructed from monthly-mean values and National Oceanic and Atmospheric Administration (NOAA) index values, and the seasonal cycle is removed from the daily data using Fourier filtering. Taking the inner product of the daily data and the pattern data yields values ranging from -1 to 1, where 1 is daily weather completely matching the index (i.e. NAO or EA), -1 is the opposite and 0 is orthogonal. The results for all 21 years of winter and summer days were analysed and split up into five winter and three summer patterns. Figure 3.1 demonstrates the split.

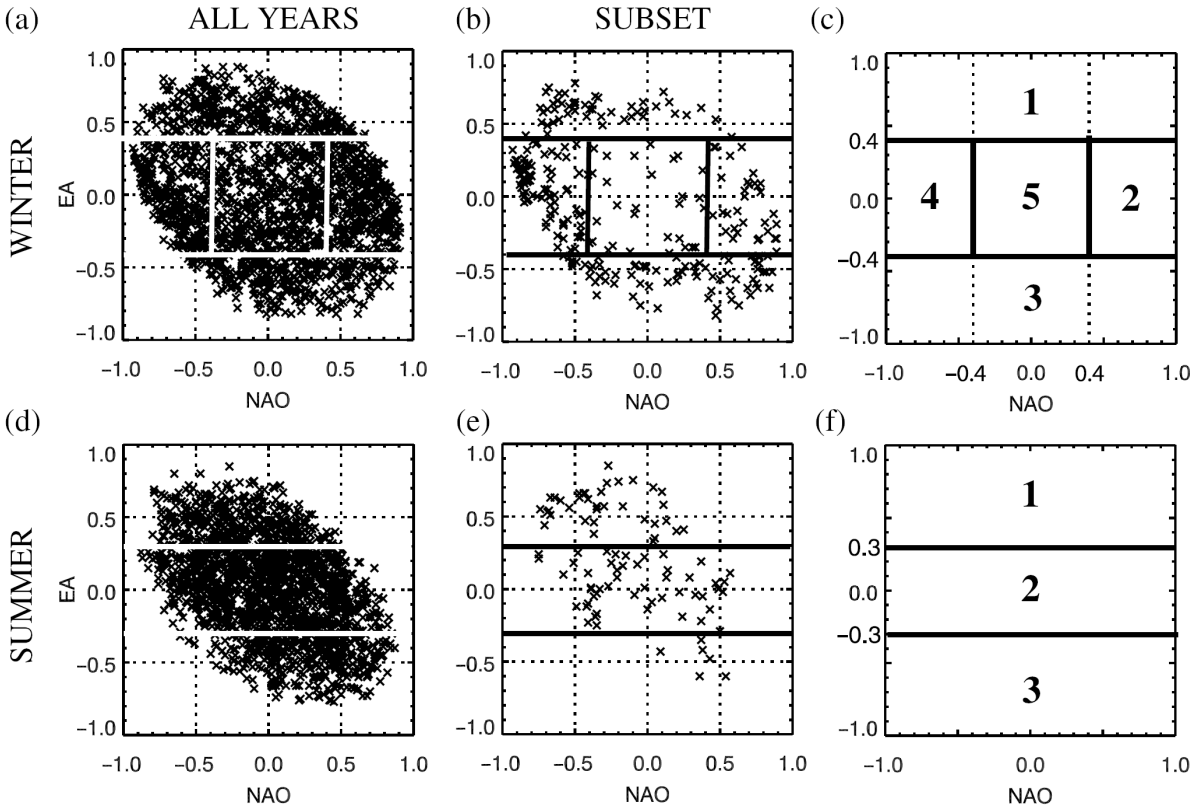


Figure 3.1: Projection coefficients of daily 250 hPa geopotential height anomalies in the NAFC on to NAO and EA. Left plots show all included days, middle plots are specific years for which optimal routing data were available, right plots show pattern demarcation. The image is from Irvine et al. (2013).

Irvine et al. (2013) assessed sensitivity of the patterns to a 0.1 shift in the thresholds; results were similar. No specific clarifications are given for the choice of five patterns for winter and three for summer. In the left and middle of Figure 3.1 the extent of the spread of data on the EA axis appears similar, but the summer data are more densely grouped towards the middle. This may explain the lower thresholds chosen for summer (0.3 versus 0.4). The summer data are also grouped more tightly around the diagonal, which could explain the choice not to differentiate patterns along the NAO axis. The likely assumption is that the location on the NAO axis can be used to determine the location on the EA axis (or vice versa). Irvine et al. (2013) state that the teleconnection patterns are weaker in summer and thus increasing the amount of types would mean they are less distinct.

The method employed leads to all summer and winter days in the 21-year time period being represented by one of the eight patterns. No mention is made by Irvine et al. (2013) of spring and autumn weather patterns or applicability of results from these patterns to spring and autumn weather.

Irvine et al. (2013) build upon an earlier study by Woollings et al. (2010), which is not fully comparable as it uses low-level wind and may miss additional upper tropospheric features. Woollings et al. (2010) found three preferred latitudinal locations of the jet stream in winter, and applied the same approach to different regions and seasons in their appendix. This leads to one/two modes for summer, one for spring, and four for autumn. These results are compiled in Figure 3.3. Qualitatively, the low-level jet stream locations in autumn and spring resemble some intermediate stage between winter and summer but in distinct ways. Similar conclusions can be drawn by visually inspecting the NOAA teleconnection pattern maps for the NAO and EA (via `http://www.cpc.ncep.noaa.gov/data/teledoc/telecontents.shtml`).

Irvine extended this methodology to investigate the applicability to spring and autumn patterns. If two five-month seasons are used, i.e. NDJFM (November–March) and MJJAS (May–September), the methodology returns the same five winter and three summer patterns as for the DJF and JJA seasons. Because there are only three distinct summer weather patterns, and in spring and autumn the weather transitions between the winter and summer patterns as visible in Figure 3.3 and the NOAA maps,

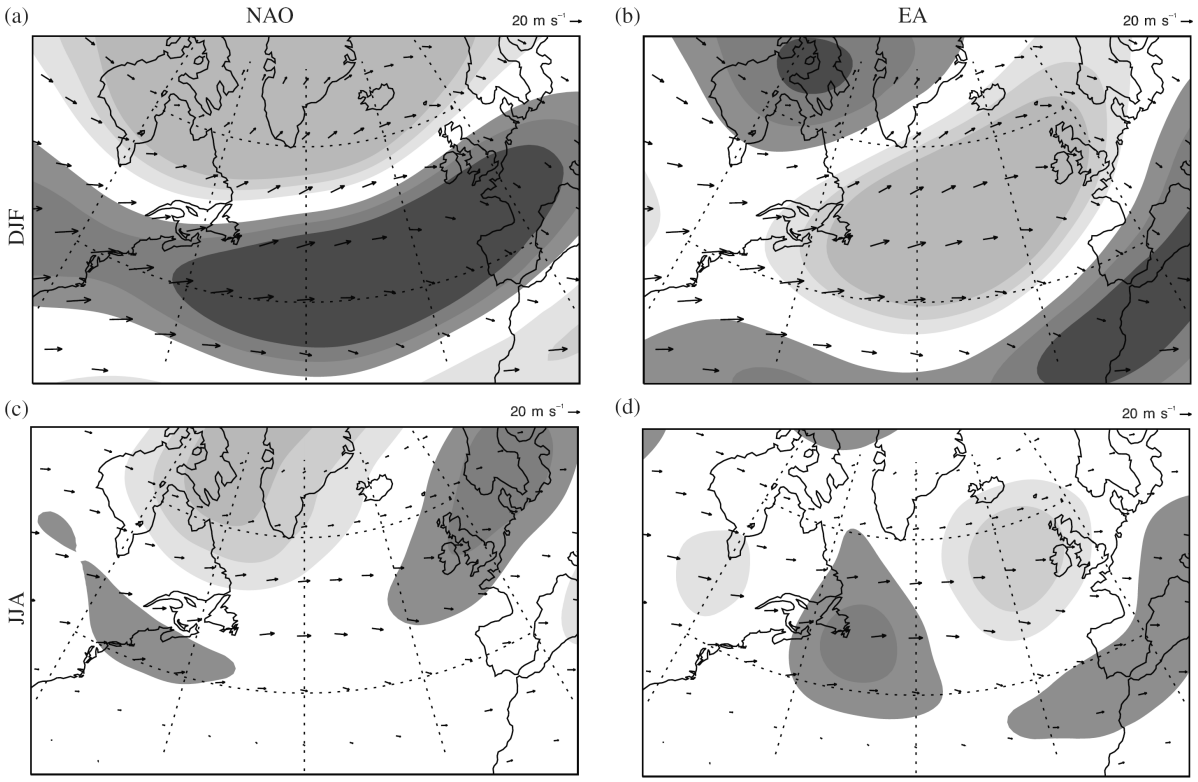


Figure 3.2: Teleconnection patterns in the NAFC for winter and summer, shown with their positive phases. Dark and light grey represent positive, respectively negative geopotential height anomalies at 250 hPa. The image is from Irvine et al. (2013).

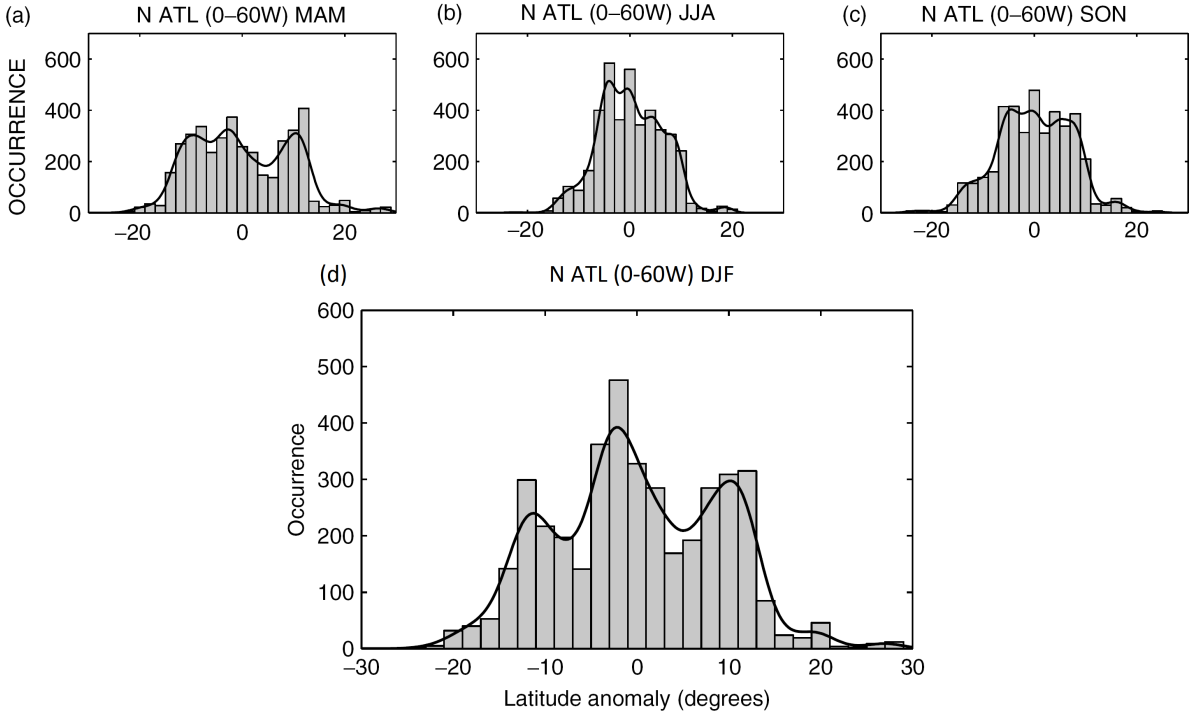


Figure 3.3: Histograms and non-normalised Kernel Density Estimations (KDEs) of low-level (925–700 hPa) jet stream latitude relative to mean in the NAFC for (a) spring, (b) summer, (c) autumn, (d) winter. The analysis used 44 years of ERA-40 data and the jet stream latitude was calculated similarly to Irvine et al. (2013). The image is adapted from Woollings et al. (2010).

clearly defined spring and autumn weather patterns were not identified. Overall, the five winter and three summer patterns from REACT4C cannot be applied to represent spring and autumn (E. A. Irvine, personal communication, 27 March, 2017).

To generate CCF data, the model described in the following sections was initialised up to a simulated day that is representative for the respective weather patterns in Figure 3.1, at which point the emissions were released. The published REACT4C study (Grewe et al., 2014a,b) examines a simulated winter day that represents WP1. Grewe et al. (2017b) state that the chosen simulation days best represent the jet stream location of each weather pattern.

3.2. Base model set-up

The base model used by Grewe et al. (2014b) is ECHAM/MESSy Atmospheric Chemistry (EMAC) version 2.42 (Jöckel et al., 2006, 2010), a modular Climate-Chemistry Model (CCM) based on the ECMWF-Hamburg v5 (ECHAM5) (Roeckner et al., 2003) General Circulation Model (GCM) and the Module Efficiently Calculating the Chemistry of the Atmosphere (MECCA) (Sander et al., 2011) chemistry sub-model. The model is run in T42L41 resolution, i.e. 41 pressure levels and 42nd level truncation in spectral space (or 2.8° × 2.8° latitude-longitude resolution at the equator). The modular set-up of EMAC facilitates development of specific sub-models for research.

3.3. REACT4C-specific set-up

The Lagrangian sub-model Atmospheric Tracer Transport in a LAgrangian model (ATTILA) (Reithmeier and Sausen, 2002) was implemented to track transport of Lagrangian air parcels. The new sub-model AIRTRAC (Grewe et al., 2014b, suppl) was developed specifically for the REACT4C study and keeps track of the concentration changes from emissions using a tagging approach (Grewe, 2013). Though not relevant for this literature review, the sub-model CONTRAIL (Grewe et al., 2014b, suppl) was also developed for this study and performs similar tasks to AIRTRAC with parametrisations specific to contrail development. The sub-models MECCA, TREXP, CLOUD, SCAV, and RAD4ALL are mentioned in Grewe et al. (2014b) or its supplements and discussed in this chapter. A complete list of sub-models employed is not published, but it may be assumed that Jöckel et al. (2016) used the same sub-models (V. Grewe, personal communication, 12 April, 2016). A notable omission from Grewe et al. (2014b) is the choice of background emissions inventories, as concentrations affect non-linear chemistry and radiation.

3.3.1. Lagrangian method

In fluid mechanics, the Lagrangian point of view is contrasted to the Eulerian point of view. Eulerian methods have a fixed point of view, e.g. a control volume and keep track of the fluid properties inside that volume as a function of time. Lagrangian methods follow individual fluid parcels through space and keep track of their properties as a function of time. The Lagrangian sub-model ATTILA has previously been applied for water vapour in the AirClim study (Grewe and Stenke, 2008; Fichter, 2009). Stenke et al. (2009) applied ATTILA to chemically active species as well and concluded that the main improvements come through water vapour representation. The main reason for using a Lagrangian approach for all species in the REACT4C study (Grewe et al., 2014b) is to facilitate a tagging approach.

3.3.2. Tagging approach

Grewe (2013) introduced a general tagging method which is contrasted with the more established perturbation method, used in e.g. AirClim (Grewe and Stenke, 2008; Fichter, 2009) or Stevenson and Derwent (2009). In a perturbation approach the base model (EMAC in this case) is run once for a perturbation in each time-region grid point and once without any perturbations. In a tagging approach many if not all emissions are introduced simultaneously in one model run. Grewe et al. (2012) showed the advantage of tagging approaches in terms of non-linear compensation effects in chemistry. All of the emissions interact with the model identically, and they are assumed small enough that they do not affect background concentrations. The trade-off is accuracy versus computational efficiency. The AIRTRAC sub-model keeps track of the respective H_2O, NO_x, O_3, HNO_3, OH, HO_2, and depleted CH_4 from each NO_x time-region grid point.

Dimension	Number	Unit	Values
Longitude	6	°W	75, 60, 45, 30, 15, 0
Latitude	7	°N	30, 35, 40, 45, 50, 60, 80
Pressure	4	hPa	400, 300, 250, 200
Time	3	UTC	6, **12**, 18

Table 3.1: Time-region grid points for REACT4C CCFs. Subsequent weather patterns were only analysed for 12 UTC, printed in bold. Data are from Grewe et al. (2014b).

3.3.3. Emissions grid

The time-region grid for releasing emissions was defined by six longitudes, seven latitudes, four pressure altitudes and three times-of-day, shown in Table 3.1. Due to computational costs and low variability within WP1 (V. Grewe, personal communication, April 2017), the CCFs for other weather patterns were only calculated for 12 UTC so the time dimension is not relevant for analysis across weather patterns. In the published REACT4C method (Grewe et al., 2014b) this means 504 time-region grid points in total, and 168 for the other weather patterns. The grid points correspond roughly to the NAFC at altitudes representative for cruising aircraft and at distinct times of day. At each time-region grid point of the model 5×10^5 kg of NO (2.33×10^5 kg of N) and 1.25×10^7 kg of water vapour were released. The emissions were released over a 15 minute model time-step across the EMAC grid box that each time-region grid point was located in (see Figure 3.4). The time-region grid points at 0°W and 45°W are located exactly on the border of two EMAC grid boxes. Emissions at 0°W are released into the grid box west of this longitude, and emissions at 45°W are released into the grid box east of this longitude (S. H. Rosanka, personal communication, 22 March, 2017).

3.3.4. Trajectories

For each time-region grid point (red triangles in Figure 3.4), the release mass was divided into 50 air parcels that were randomly distributed across the EMAC gridbox (black grid in Figure 3.4) that the time-region grid point is located in, using the TREXP sub-model (Jöckel et al., 2010). The ATTILA sub-model manages transport of the parcels and the AIRTRAC sub-model manages the concentration changes of species in the air parcels.

3.3.5. Model outputs

The REACT4C-specific EMAC model was run for 90 days (V. Grewe, personal communication, April 2016). Chemical and physical processes were calculated within the model, as explained in the next sections. Depending on the species, the model output is the 90-day perturbation and RF history. After the model integration was completed, the 90-day histories were averaged over the 50 tracers for each time-region grid point. RF for some species and all long-term climate impacts were calculated from the perturbation history during post-processing.

3.4. Atmospheric processes

3.4.1. Water vapour physics

For the Lagrangian water vapour parcels Grewe et al. (2014b) took atmospheric loss processes into account with the following differential equation:

$$\frac{d\mathrm{H_2O}^e}{dt} = -\frac{pr}{\mathrm{H_2O}^b}\mathrm{H_2O}^e \tag{3.1}$$

Here $\mathrm{H_2O}^e$ refers to the mixing ratio in [mol mol^{-1}] in the Lagrangian parcel, $\mathrm{H_2O}^b$ refers to the total water vapour mixing ratio in the respective Eulerian EMAC grid box in [mol mol^{-1}] (i.e. the background mixing ratio), and pr (precipitation) is the water vapour loss rate in [mol mol^{-1} s^{-1}] in the respective grid box due to rainfall and snowfall. This rate is calculated by the CLOUD sub-model, which re-implements the ECHAM5 cloud cover scheme (Roeckner et al., 2003, 2006) into the Modular Earth Submodel System (MESSy) standard (Jöckel et al., 2010). Essentially, the Lagrangian water vapour loss rate is proportional to that of the ECHAM5 background model running on T42L41 resolution. In each timestep the background rates on the EMAC grid are mapped to each Lagrangian air parcel within that cell and used to calculate Equation (3.1) within AIRTRAC.

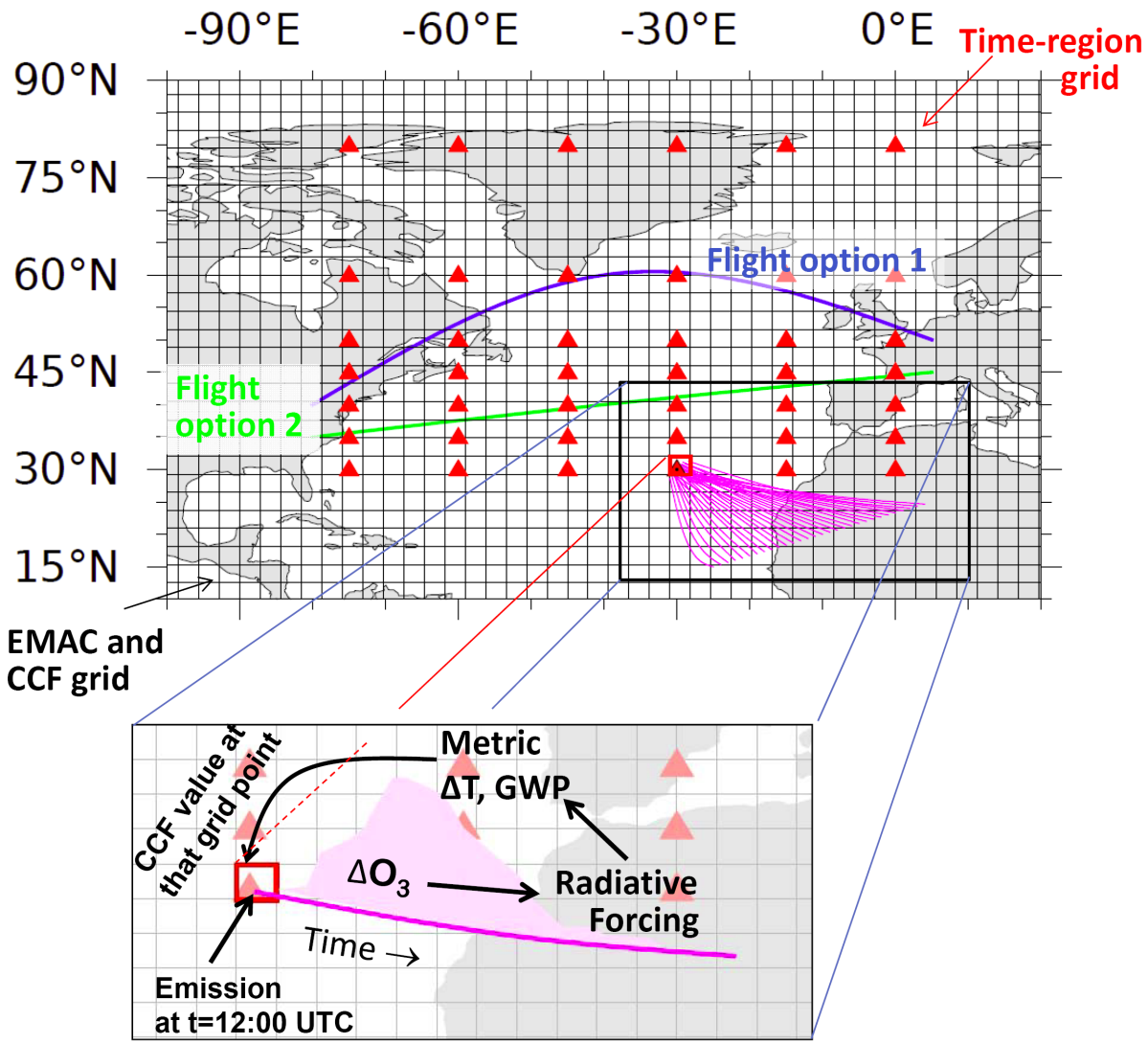


Figure 3.4: Visual representation of REACT4C CCF method. The black grid represents the background EMAC resolution. For each time-region grid point 50 trajectories (magenta) are released in the corresponding EMAC gridbox. The insert shows how the trajectory evolves over time, RF is calculated, and the corresponding metric is mapped back to the time-region grid. The green and blue lines show two aircraft route options. The image is from Grewe et al. (2014b).

3.4.2. NO_x chemistry

The NO_x chemistry tagging used by Grewe et al. (2014b) follows the tagging method presented by Grewe et al. (2010). Detailed background chemistry is part of the MECCA sub-model on the EMAC grid. For computational efficiency, simplified chemistry is used for the emissions tagging built into the AIRTRAC sub-model. Background (b) and tagged emissions (e) mixing ratios [mol mol^{-1}] are separated. Production (P) and loss (L) rates are given in [mol mol^{-1} s^{-1}]. As with water vapour, the background processes are calculated on the EMAC grid while the tagged processes (via AIRTRAC) are calculated for Lagrangian air parcels, requiring conversions within the model.

3.4.3. Ozone chemistry

The ozone-forming reaction for NO and ozone-destroying reaction for NO_2 are taken into account (Grewe et al., 2014b):

$$HO_2 + NO \longrightarrow OH + NO_2 \tag{3.2}$$

$$NO_2 + O_3 \longrightarrow NO + 2\,O_2 \tag{3.3}$$

With the tagging approach (Grewe et al., 2010), this gives the following production and loss terms (Grewe et al., 2014b):

$$P^e_{\mathrm{O_3}} = P^b_{\mathrm{O_3}} \cdot \frac{1}{2}\left(\frac{\mathrm{HO_2^e}}{\mathrm{HO_2^b}} + \frac{\mathrm{NO}^e}{\mathrm{NO}^b}\right) \tag{3.4}$$

$$L^e_{\mathrm{O_3}} = L^b_{\mathrm{O_3}} \cdot \frac{1}{2}\left(\frac{\mathrm{NO_2^e}}{\mathrm{NO_2^b}} + \frac{\mathrm{O_3^e}}{\mathrm{O_3^b}}\right) \tag{3.5}$$

Note that Reaction (3.2) does not directly form ozone. It is known to be the rate-limiting step in the formation of ozone (e.g. Jaeglé et al., 1999), and is followed by (Lee et al., 2010):

$$\mathrm{NO_2} \xrightarrow{\mathrm{hv}} \mathrm{NO} + \mathrm{O(^3P)} \tag{3.6}$$

$$\mathrm{O(^3P)} + \mathrm{O_2} + \mathrm{M} \longrightarrow \mathrm{O_3} + \mathrm{M} \tag{3.7}$$

As a simplification, all background nitrogen species are grouped into one family (except NO_x and HNO_3), and it is assumed that the emissions are small enough not to affect the specific reaction rates. Re-writing Reaction (3.4) for the latter assumption gives (Grewe et al., 2014b):

$$\frac{P^b_{\mathrm{O_3}}}{\mathrm{HO_2^b} \cdot \mathrm{NO}^b} = \frac{P^b_{\mathrm{O_3}} + P^e_{\mathrm{O_3}}}{\left(\mathrm{HO_2^b} + \mathrm{HO_2^e}\right)\left(\mathrm{NO}^b + \mathrm{NO}^e\right)} \tag{3.8}$$

As a further simplification, the contribution of aviation to the production of ozone is assumed to depend only on NO_x. Destruction of ozone is split into two terms, one for NO_x and one for all other loss processes. The ratio of NO to NO_2 is assumed to be equal between emission and background, so the combined NO_x mixing ratio can be used. The differential equation for the contribution of aviation to the ozone mixing ratio is then (Grewe et al., 2014b):

$$\frac{d\mathrm{O_3^e}}{dt} = P^b_{\mathrm{O_3}} \frac{\mathrm{NO_x^e}}{\mathrm{NO_x^b}} - \frac{1}{2} D^b_{\mathrm{O_3},1}\left(\frac{\mathrm{NO_x^e}}{\mathrm{NO_x^b}} + \frac{\mathrm{O_3^e}}{\mathrm{O_3^b}}\right) - D^b_{\mathrm{O_3},2}\frac{\mathrm{O_3^e}}{\mathrm{O_3^b}} \tag{3.9}$$

Here, $D^b_{\mathrm{O_3},1}$ in [mol mol^{-1} s^{-1}] is the background destruction rate through NO_x and $D^b_{\mathrm{O_3},2}$ is the background destruction rate through all other loss processes. Grewe et al. (2014b) state that an exchange of $\mathrm{NO_x^e}$ with $\mathrm{HNO_3^e}$ is taken into account and that the latter is washed out. Based on literature (Penner et al., 1999; Lee et al., 2010) the reaction for NO_x loss is almost certainly:

$$\mathrm{OH} + \mathrm{NO_2} + \mathrm{M} \longrightarrow \mathrm{HNO_3} + \mathrm{M} \tag{3.10}$$

3.4.4. Methane chemistry

For methane a similar procedure is followed with more chemical reactions. Methane oxidation requires OH through Reaction (3.17), thus the majority of the tagged reactions account for OH production and loss. The reactions taken into account for OH production are (Grewe et al., 2014b):

$$\mathrm{H_2O} + \mathrm{O(^1D)} \longrightarrow 2\,\mathrm{OH} \tag{3.11}$$

$$\mathrm{HO_2} + \mathrm{O_3} \longrightarrow \mathrm{OH} + 2\,\mathrm{O_2} \tag{3.12}$$

$$\mathrm{HO_2} + \mathrm{NO} \longrightarrow \mathrm{OH} + \mathrm{NO_2} \tag{3.13}$$

The reactions taken into account for OH loss are (Grewe et al., 2014b):

$$\mathrm{OH} + \mathrm{O_3} \longrightarrow \mathrm{HO_2} + \mathrm{O_2} \tag{3.14}$$

$$\mathrm{OH} + \mathrm{CO} \xrightarrow{\mathrm{O_2}} \mathrm{HO_2} + \mathrm{CO_2} \tag{3.15}$$

$$\mathrm{OH} + \mathrm{RH} \xrightarrow{\mathrm{O_2}} \mathrm{RO_2} + \mathrm{H_2O} \tag{3.16}$$

$$\mathrm{OH} + \mathrm{CH_4} \xrightarrow{\mathrm{O_2}} \mathrm{CH_3O_2} + \mathrm{H_2O} \tag{3.17}$$

$$\mathrm{OH} + \mathrm{HO_2} \longrightarrow \mathrm{H_2O} + \mathrm{O_2} \tag{3.18}$$

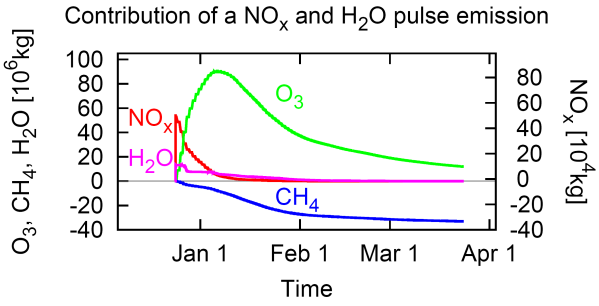


Figure 3.5: Temporal evolution of NO_x (red), ozone (green), methane (blue) and water vapour (magenta) global atmospheric mass after an emission (see Section 3.3.3) at 75°W, 30°N, 200 hPa, 23 December (WP1), 6 UTC. The image is from Grewe et al. (2014b).

An extra reaction is taken into account for production of HO_2 (Grewe et al., 2014b):

$$RO_2 + NO \xrightarrow{O_2} HO_2 + R'CHO + NO_2 \tag{3.19}$$

Two extra reactions are taken into account for loss of HO_2 (Grewe et al., 2014b):

$$RO_2 + HO_2 \longrightarrow ROOH + O_2 \tag{3.20}$$
$$HO_2 + HO_2 \longrightarrow H_2O_2 + O_2 \tag{3.21}$$

Aircraft emission contributions to H_2O, CO, RH, and CH_4 mixing ratios are considered negligible in terms of their effects on OH. Production and loss terms for OH and HO_2 can be formed analogous to ozone and combined to produce differential equations. The methane loss from aviation NO_x emissions then depends on OH through Reaction (3.17). The differential equations contain too many terms to list and can be found in Grewe et al. (2014b).

3.4.5. Typical example of temporal evolution

Figure 3.5 shows a typical example of the temporal evolution of several species after an emission. The complexity of the system of reactions is clearly visible. Methane decreases at first due to the increased OH from Reaction (3.2). Once NO_x is removed from the atmosphere by scavenging of HNO_3, the increased ozone leads to a stronger methane reduction (Grewe et al., 2014b).

3.5. Sensitivity studies

A number of sensitivity studies were performed to qualitatively balance accuracy and computational costs for NO_x and ozone masses. An extensive sensitivity study was performed on the ATTILA Lagrangian parcels. From 24 time-region grid points 50 parcels were released and the mean NO_x mixing ratio over the first month after release was calculated. Then the amount of parcels was varied between 2 and 48 to assess convergence. The mean standard deviation per grid point for NO_x mixing ratio is almost 20% if only 2 parcels are used, and 12% for ozone. This drops for increasing numbers of parcels and both the mean standard deviation and the extremes are below 10% for 20 parcels. Grewe et al. (2014b) conclude that the error almost converges when 50 parcels are used.

Adding one time point led to $\pm$40% and $\pm$25% variation in masses of NO_x and ozone respectively. Releasing emissions in the mid-points of the current time-region grid (Figure 3.4), thus adding resolution in both horizontal dimensions simultaneously, led to $\pm$50% and $\pm$35% variation in masses respectively. Grewe et al. (2014b) conclude that horizontal resolution is more important than temporal resolution, and that the variation of climate impact from NO_x emissions itself is larger, thus the resolution is appropriate.

As a first-order estimation, using the mid-point release grid would mean twice as many grid points in two dimensions i.e. four times the grid points, thus also four times the trajectories to calculate to explain $\pm$50% NO_x variation. Reducing the trajectories released from 50 to 20 (for example) means a 60% decrease in trajectories while the variation is increased by at most 10% for NO_x. It is not known how the computational cost and alternative refinements (e.g. to the amount of tagged chemicals) compare to the number of trajectories. Note that in the separate day simulated for the WeCare project a finer time-region grid was applied (Grewe et al., 2017a).

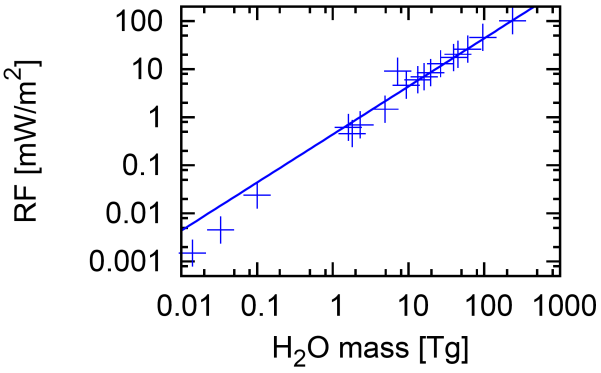


Figure 3.6: Resultant adjusted radiative forcing and atmospheric water vapour mass change for a uniform emission at varying latitude-altitude locations from Grewe and Stenke (2008), with a linear fit at $\mathrm{RF}_{\mathrm{adj}} = 4.28 \times 10^{-13}$ W m^{-2} kg^{-1}. The image is from Grewe et al. (2014b).

3.6. Radiative forcing

The EMAC model in the REACT4C study (Grewe et al., 2014b) was run for 90 model days. The AIRTRAC sub-model was set up to keep track of tagged species, but to quantify climate impact the concentration changes must be translated to radiative forcing.

3.6.1. Water vapour adjusted RF

For adjusted RF from a water vapour emission a new correlation was developed, based on Grewe and Stenke (2008). The AirClim study presented by Grewe and Stenke (2008) had a comparable approach to REACT4C but used perturbations instead of tagging. Depending on its location in their latitude-altitude grid, the same uniform water vapour emission led to a different change in atmospheric water vapour. Grewe et al. (2014b) compared this resulting change in atmospheric water vapour to the resulting radiative forcing (see Figure 3.6), and found an almost linear correlation. The results from Grewe and Stenke (2008) cannot be used to relate an emission to adjusted RF, but due to the REACT4C set-up the resulting atmospheric mass change is already available and the relation can be used. The linear fit shown in Figure 3.6 is thus applied to the 90-day time history of the water vapour mass from an emission, yielding 90-day adjusted RF time history.

3.6.2. Ozone RF

Instantaneous RF for ozone from NO_x emissions is calculated by the RAD4ALL EMAC sub-model (Jöckel et al., 2006; Roeckner et al., 2006), which re-implements the ECHAM5 radiation scheme (Roeckner et al., 2003) into the MESSy interface. A separate relation was set up by Grewe et al. (2014b) to relate ozone instantaneous RF to adjusted RF (see Section 2.1.2 for definitions).

Ozone instantaneous RF

The ECHAM5 radiation code (and by extension the RAD4ALL code) has a larger time-step than the model dynamics for computational efficiency, typically 2 hours (Roeckner et al., 2003). The solar irradiation in a grid cell varies with Sun-Earth distance and effective solar zenith angle, which accounts for curvature of the atmosphere. For both shortwave and longwave radiation, the ECHAM5 radiation scheme follows the methods implemented in the ECMWF Integrated Forecasting System (IFS) model at the time of development (Cy23r1), i.e. the Fouquart and Bonnel (1980) method for shortwave radiation and the Rapid and accurate Radiative Transfer Model (RRTM) (Mlawer et al., 1997) method for longwave radiation.

The shortwave scheme uses the Eddington approximation for the spherical angles and the delta-Eddington approximation for reflectivity. Four shortwave spectral bands are used, one for visible and ultraviolet light, three for near infrared. Ozone absorption is accounted for in the visible and ultraviolet band (0.25-0.69 μm) and the farthest infrared band (2.38–4.00 μm).

The main difference in the longwave scheme is that there are emission sources within the atmosphere itself and the Earth is a diffuse source compared to the solar beam Roeckner et al. (2006). Sixteen longwave spectral bands are used, of which ozone absorption is accounted for in the 630–700, 700–820, 980–1080, and 1080–1180 cm^{-1} bands. Each band is split up into a number of intervals, 50 in total for the bands that are relevant for ozone absorption.

Forster et al. (2011) assessed the performance of a large number of CCM and GCM radiation codes compared to line-by-line codes. For the stratospheric ozone test the ECHAM5 code has a shortwave

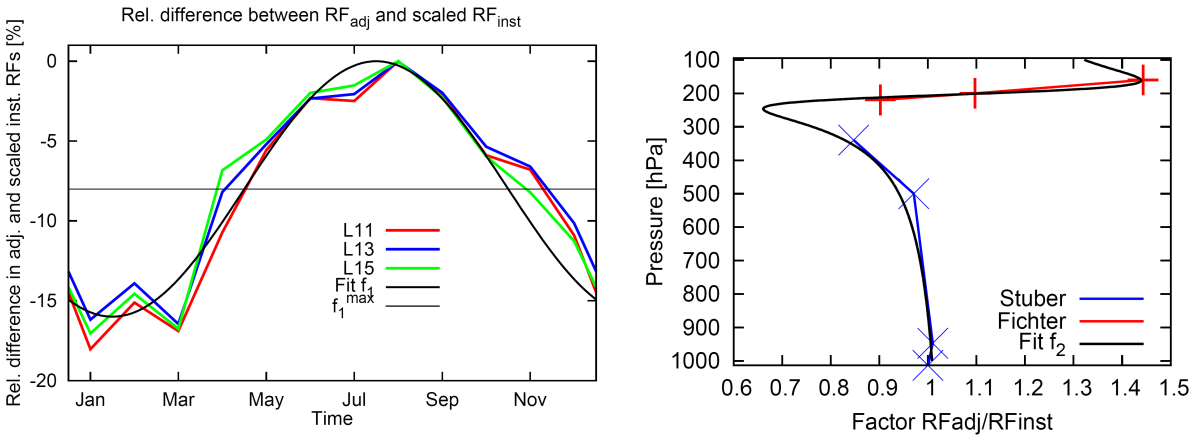


Figure 3.7: Results from Fichter (2009) and Stuber (2003) used to identify ozone adjusted RF time and altitude dependence in Grewe et al. (2014b). The image is adapted from Grewe et al. (2014b).

accuracy of around 10% and a longwave accuracy of almost 30%. For the tropospheric ozone test, solar radiation plays only a minor role and the longwave accuracy is within 10%. Sukhodolov et al. (2014) assessed the shortwave performance of the ECHAM family of GCMs specifically. The main weakness found in the four-band scheme used in ECHAM5 is in modelling responses to variations in solar irradiance, which was not part of the REACT4C study.

The ECHAM5 radiation code is not the most accurate one available, but there is an important trade-off to make with computational efficiency and it is more detailed than the methods employed for water vapour, methane, and PMO RF calculation.

Ozone adjusted RF

Grewe et al. (2014b) developed an analytical formula for converting ozone instantaneous RF to adjusted RF:

$$\mathrm{RF}_{\mathrm{adj}} = f_1(t) \times f_2(p) \times \mathrm{RF}_{\mathrm{inst}} \tag{3.22}$$

Monthly AirClim results from Fichter (2009) were adopted and used to identify time dependence. Exactly what results were adopted is not mentioned, but EMAC simulations were run for each monthly perturbation of three latitude-altitude cells, including spin-up time to allow stratospheric temperature adjustment. Analysing the time variance of instantaneous and adjusted RF of these cells reveals a sinusoidal correlation - Equation (3.23) - after scaling by the yearly maximum value and shifting to the ratio of instantaneous versus adjusted RF (Figure 3.7 left).

For the altitude dependence, the four data points from Fichter (2009) are combined with five data points from a similar study by Stuber (2003). The fit found, Equation (3.24), closely reproduces the data used (Figure 3.7 right, one point is out of range). A gap is visible in the source data in the 200 to 400 hPa range of calculated CCFs, where the fit produces a minimum. Grewe et al. (2014b) describe in text and their Figure A4 that the factor between adjusted and instantaneous RF should become negative in the lower stratosphere, but this is not reproduced by the fit, though one data point at 50 hPa from Stuber (2003) was taken into account.

$$f_1(t) = 0.08\left[\sin\left(2\pi\frac{t-7.5}{12} + \frac{\pi}{2}\right) - 1\right] \tag{3.23}$$

$$f_2(p) = 1.05 - 1.1\frac{\frac{1}{60}p - 3.4}{2.0\left(\frac{1}{60}p - 3.4\right)^2 + 1} \tag{3.24}$$

3.6.3. Methane adjusted RF

Grewe et al. (2014b) use the so-called IPCC formula to calculate adjusted RF for methane from NO_x emissions, proposed in the IPCC First Assessment Report (Shine et al., 1990). The IPCC report states

that the equation's functional form is from Wigley (1987), the coefficient was derived from Hansen et al. (1988), and the CH_4-N_2O overlap term was taken from Hansen et al. (1988). Hansen's results are derived from a 1-D model given by Lacis et al. (1981). Wigley (1987) is not publicly available, but Wigley is one of the contributors to Shine et al. (1990). The following form was thus applied in REACT4C (Shine et al., 1990):

$$\mathrm{RF}_{\mathrm{adj}} = 0.036\left(\sqrt{M} - \sqrt{M_0}\right) - \left[f\left(M, N_0\right) - f\left(M_0, N_0\right)\right] \tag{3.25}$$

$$f(M,N) = 0.47 \ln\left[1 + 2.01 \times 10^{-5} (MN)^{0.75} + 5.31 \times 10^{-15} M (MN)^{1.52}\right] \tag{3.26}$$

Here M and N are the methane and N_2O volumetric concentrations in [ppbv] and the adjusted RF is calculated from a concentration change relative to M_0 and N_0. The term $f(M,N)$ is the CH_4-N_2O overlap term from Hansen et al. (1988), which represents the overlap in radiation spectra of the two GHGs. N_2O is not tagged by the AIRTRAC sub-model: the background N_2O concentration is taken based on current measured values (V. Grewe, personal communication, 12 April, 2016).

3.6.4. PMO adjusted RF

PMO adjusted RF is calculated using results from Dahlmann (2012):

$$\mathrm{RF}_{\mathrm{adj}}\left(\mathrm{PMO}\right) = -0.29 \mathrm{RF}_{\mathrm{adj}}\left(\mathrm{CH_4}\right) \tag{3.27}$$

Dahlmann (2012) states that the coefficient for PMO RF above was calculated from the model ECHAM4/L39 + CHEM (E39C) and AirClim results for O_3^{net} (sum of direct ozone and PMO), direct ozone and methane RF, but no further details are given on the calculation. The value is also compared to other results from literature (0.23–0.58) and is in the low range of results (Dahlmann, 2012). Grewe et al. (2014b) do not provide details on their reason to choose this specific value. Note that Dahlmann (2012) used the same IPCC formula (3.6.3) to calculate the methane RF that this scaling is based on. Because of this linear scaling from methane results there is no point in analysing the variability PMO results or developing algorithms.

3.7. Climate metrics

To avoid misinterpretation of results in climate science, it is important to pose the adequate climate question before devising metrics to answer that question (Grewe and Dahlmann, 2015). Grewe et al. (2014b) identified three political questions regarding climate impact of aviation and formulated preferred metrics for each question:

- Q1: "What is the *short-term* climate impact of the REACT4C re-routing strategy?"
- Q2: "What is the *long-term* climate impact of the REACT4C re-routing strategy?"
- Q3: "What is the medium-range climate impact of a *present-day* REACT4C re-routing decision?"

In literature (Shine et al., 2005; Boucher and Reddy, 2008; Fuglestvedt et al., 2010) a distinction has been made between pulse (P), sustained (S), and future (F) emissions when presenting climate impact metrics. The three underlying metrics discussed by Grewe et al. (2014b) are Absolute Global Warming Potential (AGWP), Absolute Global Temperature Change Potential (AGTP), and Average Temperature Response (ATR). The three climate metrics finally calculated for the single published weather pattern (WP1) are F-ATR20 (Q1), P-AGWP20 (Q1), and P-AGWP100 (Q2) (Grewe et al., 2014b), where the number represents the time horizon in [years]. Results for the ATM optimisation Pareto front in Grewe et al. (2014a) show that F-ATR20 presents the lowest mitigation gain potential of the three metrics, and P-AGWP100 presents the highest mitigation gain.

The F-ATR with 20 or 100 year time horizon is identified as the best metric for answering Q1 and Q2 in Grewe et al. (2014b), as temperature change better represents climate impact than does integrated RF, and a future emissions scenario (or sustained scenario) represents the effect of a strategic re-routing decision instead of a present-day decision. The P-AGWP metrics are presented as the furthest abstraction from this answer (Grewe et al., 2014b), but have the benefit of simplicity and allowing comparisons to previous research, as these metrics relative to carbon dioxide AGWP have been widely

Table 3.2: Values of the constants for Equation (3.30), from Boucher and Reddy (2008).

		$i = 1$	$i = 2$
c_j	K (W m^{-2})$^{-1}$	0.631	0.429
d_j	years	8.4	409.5

used (see Section 2.1.2). The only results available for the other seven REACT4C weather patterns are F-ATR20, as used for the ATM optimisation in Grewe et al. (2017b). As F-ATR20 calculation follows from the other metrics, they are also explained here.

3.7.1. Method for P-AGWP and P-AGTP

The REACT4C simulation set-up (Grewe et al., 2014b) with its 15-minute emissions represents a pulse set-up (before application of a future emissions scenario) whereas some recent studies have used sustained emissions. In comparison, Grewe and Stenke (2008) ran their model for 5 years with constant emissions. The REACT4C approach (Grewe et al., 2014b, suppl) follows Fuglestvedt et al. (2010) for P-AGWP and P-AGTP. Absolute Global Warming Potential (AGWP) and Absolute Global Temperature Change Potential (AGTP) are defined as follows (Grewe et al., 2014b, suppl):

$$\mathrm{AGWP}\, H_x = \int_0^H F_x(t)\, \mathrm{d}t \tag{3.28}$$

$$\mathrm{AGTP}\, H_x = \int_0^H F_x(t) R(H-t)\, \mathrm{d}t \tag{3.29}$$

$$R(t) = \sum_{j=1}^{2} \frac{c_j}{d_j} \exp\left(\frac{-t}{d_j}\right) \tag{3.30}$$

Here $F_x(t)$ is the time-dependent adjusted RF of the emission. The response function, Equation (3.30), represents climate response (i.e. surface temperature change) to an RF change on both the short and long timescale as given by Fuglestvedt et al. (2010) and first proposed by Boucher and Reddy (2008) based on 1000 simulated years of CO_2 response in the HadCM3 GCM. The values of the constants are given in Table 3.2. Adding up the two terms of c_j gives an equilibrium climate sensitivity of 1.05 K (W m^{-2})$^{-1}$. Fuglestvedt et al. (2010) give the tentative interpretation that the first term of the constants is the ocean mixed-layer response and the second term is the deep ocean response.

Depending on the species, either the concentration, instantaneous RF, or adjusted RF are output from the EMAC simulation and during post-processing the time-history of adjusted RF is calculated. The peak RF in the 90-day adjusted RF history was determined, and an exponential decay was fit from this point up to one year after emission (V. Grewe, personal communication, 18 April, 2016). The average of this extrapolated history is then the year-mean adjusted RF: $\overline{F_x}$. This is used to estimate an equivalent pulse A_x given an assumed exponential lifetime τ_x (Grewe et al., 2014b, suppl):

$$\overline{F_x} = \int_0^1 A_x \exp\left(-\frac{t}{\tau_x}\right) \mathrm{d}t = A_x \tau_x \left[1 - \exp\left(-\frac{1}{\tau_x}\right)\right] \tag{3.31}$$

Specific assumptions about the evolution of RF as a function of time are made for each species, detailed in the next sections. These can then be used directly to calculate AGWP or combined with Equation (3.30) to calculate AGTP.

3.7.2. Water vapour

The exponential lifetime τ for water vapour is assumed to be 0.05 years (Grewe et al., 2014b, suppl). Applying Equation (3.31) and approximating:

$$A_{\mathrm{H_2O}} = \frac{\overline{F_{\mathrm{H_2O}}}}{\tau_{\mathrm{H_2O}} \left[1 - \exp\left(-\frac{t}{\tau_{\mathrm{H_2O}}}\right)\right]} \approx \frac{\overline{F_{\mathrm{H_2O}}}}{\tau_{\mathrm{H_2O}}} \tag{3.32}$$

For water vapour RF a pulse with exponential decay is assumed (Grewe et al., 2014b, suppl):

$$F_{\mathrm{H_2O}}(t) = A_{\mathrm{H_2O}} \exp\left(-\frac{t}{\tau_{\mathrm{H_2O}}}\right) \tag{3.33}$$

Applying Equation (3.28) and solving the integral gives (Grewe et al., 2014b, suppl):

$$\text{P-AGWP}\, H_{\mathrm{H_2O}} = A_{\mathrm{H_2O}} \tau_{\mathrm{H_2O}} \left[1 - \exp\left(-\frac{H}{\tau_{\mathrm{H_2O}}}\right)\right] \tag{3.34}$$

Applying Equations (3.29) and (3.30) and solving the integral gives (Grewe et al., 2014b, suppl):

$$\text{P-AGTP}\, H_{\mathrm{H_2O}} = \sum_{j=1}^{2} \frac{A_{\mathrm{H_2O}} \tau_{\mathrm{H_2O}} c_j}{\tau_{\mathrm{H_2O}} - d_j} \left[\exp\left(-\frac{H}{\tau_{\mathrm{H_2O}}}\right) - \exp\left(-\frac{H}{d_j}\right)\right] \tag{3.35}$$

3.7.3. Ozone

Two methods for ozone climate response are presented, one where an instant peak $A_{\mathrm{O_3}}$ decays exponentially, and another with an exponential increase followed by a decay. The first, more simple, method is evaluated against the second and deviates only 1–3% for a 1–3 month period of exponential increase. The simple method was used for CCF calculation and is explained here. During extrapolation of the 90-day adjusted RF time history to year-mean adjusted RF, the exponential lifetime τ was derived for each separate time-region pulse and this was used to calculate the long-term response (V. Grewe, personal communication, 18 April, 2016). Analogous to water vapour (Grewe et al., 2014b, suppl):

$$A_{\mathrm{O_3}} = \frac{\overline{F_{\mathrm{O_3}}}}{\tau_{\mathrm{O_3}} \left[1 - \exp\left(-\frac{t}{\tau_{\mathrm{O_3}}}\right)\right]} \approx \frac{\overline{F_{\mathrm{O_3}}}}{\tau_{\mathrm{O_3}}} \tag{3.36}$$

$$\text{P-AGWP}\, H_{\mathrm{O_3}} = A_{\mathrm{O_3}} \tau_{\mathrm{O_3}} \left[1 - \exp\left(-\frac{H}{\tau_{\mathrm{O_3}}}\right)\right] \tag{3.37}$$

$$\text{P-AGTP}\, H_{\mathrm{O_3}} = \sum_{j=1}^{2} \frac{A_{\mathrm{O_3}} \tau_{\mathrm{O_3}} c_j}{\tau_{\mathrm{O_3}} - d_j} \left[\exp\left(-\frac{H}{\tau_{\mathrm{O_3}}}\right) - \exp\left(-\frac{H}{d_j}\right)\right] \tag{3.38}$$

3.7.4. Methane

The perturbation lifetime for methane is in the order of 12 years (Lee et al., 2010), so the assumption of a pulse with exponential decay would not be accurate. Instead the RF response to a NO_x emission is assumed to decrease linearly to a negative maximum F_m at a time H_0, followed by an exponential decay from that point on (Grewe et al., 2014b, suppl):

$$\begin{aligned} F_{\mathrm{CH_4}}(t) &= \frac{t}{H_0} F_m \quad \text{for } 0 \le t \le H_0 \\ F_{\mathrm{CH_4}}(t) &= F_m \exp\left[\frac{-(t - H_0)}{\tau_{\mathrm{CH_4}}}\right] \quad \text{for } H_0 < t \end{aligned} \tag{3.39}$$

The value of H_0 is set to 1 year to approximate the period during which a NO_x emission induces a methane perturbation through OH. Applying Equation (3.28) and solving the integrals gives (Grewe et al., 2014b, suppl):

$$\text{P-AGWP}\, H_{\mathrm{CH_4}} = \frac{H_0}{2} F_m + \tau_{\mathrm{CH_4}} F_m \left[1 - \exp\left(\frac{-(H - H_0)}{\tau_{\mathrm{CH_4}}}\right)\right] \tag{3.40}$$

Applying Equations (3.29) and (3.30), re-arranging and solving the integrals gives (Grewe et al., 2014b, suppl):

$$\text{P-AGTP}\,H_{\text{CH}_4} = \sum_{j=1}^{2} \frac{c_j F_m}{H_0} \left(\left[H_0 - d_j \right] \exp\left[\frac{-(H - H_0)}{d_j} \right] + d_j \exp\left[\frac{-H}{d_j} \right] \right) + \sum_{j=1}^{2} \frac{c_j F_m \tau_{\text{CH}_4}}{\tau_{\text{CH}_4} - d_j} \left(\exp\left(\frac{-(H - H_0)}{\tau_{\text{CH}_4}} \right) - \exp\left(\frac{-(H - H_0)}{d_j} \right) \right) \tag{3.41}$$

3.7.5. Method for F-ATR

The REACT4C approach (Grewe et al., 2014b, suppl) follows Fuglestvedt et al. (2010) for P-AGWP and P-AGTP but provides no details on F-ATR20 calculation. The Average Temperature Response (ATR) metric was proposed by Schwartz Dallara et al. (2011) for quantifying climate impact mitigation gain of aviation technology advances, as it represents the average temperature change from a technological change over the average life-time of an aircraft. In the REACT4C method, F-ATR20 is used to represent the short-term climate impact of a structural re-routing decision given a business-as-usual geographical spread and growth of the aviation sector (Grewe et al., 2014b, 2017b). ATR is calculated by the following equation (Grewe et al., 2014b):

$$\text{ATR}\,H = \frac{1}{H} \int_{t_0}^{t_0+T} \mathrm{d}T(t)\,\mathrm{d}t \tag{3.42}$$

In REACT4C, to find F-ATR for a timescale H the F-AGTP was calculated for each year from 0 to H, and Equation (3.42) was subsequently applied to derive the average temperature change. Shine et al. (2005) introduced the concept of AGTP and showed how to derive it for both pulse and sustained emissions. They mention in deriving S-AGTP how F-AGTP could be derived.

The future scenario is implemented by taking into account the IPCC Fa1 scenario (Penner et al., 1999) for emissions. This scenario is based on the earlier IS92a IPCC growth scenario (Leggett et al., 1992), taking into account expected technology improvements for fuel burn and NO_x emission reductions. For REACT4C the net growth from scenario Fa1, including economic growth and technological improvement, was factored into the temporal development of the F-AGTP from 0 to H (V. Grewe, personal communication, 18 April, 2016).

According to Shine et al. (2005), the concentration change $\Delta\chi$ from a change in emissions ΔS with a time constant α is:

$$\frac{\mathrm{d}\Delta\chi(t)}{\mathrm{d}t} = \Delta S(t) - \frac{\Delta\chi(t)}{\alpha} \tag{3.43}$$

For sustained emissions ΔS is constant, thus (Shine et al., 2005):

$$\Delta\chi(t) = \alpha \Delta S \left[1 - exp\left(-\frac{t}{\alpha} \right) \right] \tag{3.44}$$

Then by representing RF as a factor times the concentration change and assuming ΔS equals unity, the S-AGTP is calculated. Shine et al. (2005) uses a single timescale response as opposed to the dual-timescale response from Boucher and Reddy (2008) used by Grewe et al. (2014b). Converting the variables and assuming the perturbation lifetime τ does not equal either of the climate response lifetimes d_j:

$$\text{S-AGTP}\,H_x = \sum_{j=1}^{2} \frac{A_x \tau_x c_j}{d_j} d_j \left[1 - \exp\left(-\frac{H}{d_j} \right) \right] - \frac{1}{d_j^{-1} - \tau_x^{-1}} \left[\exp\left(-\frac{H}{\tau_x} \right) - \exp\left(-\frac{H}{d_j} \right) \right] \tag{3.45}$$

The form of Equation 3.45 resembles that of the water vapour and ozone P-AGTP formulas after assuming an exponential decay. Shine et al. (2005) do not mention an exponential decay explicitly but it is contained in the assumption that RF is concentration change times a factor. This is confirmed by the fact that the Shine et al. (2005) formula for P-AGTP is identical to Equation (3.35) for water

vapour and Equation (3.38) for ozone. A more general approach to P-AGWP and S-AGWP is given by Berntsen et al. (2005).

For a future emissions scenario, ΔS in Equation (3.44) is not constant but equal to the assumed trend of the future emissions growth. It is outside of the scope of this thesis, but a similar approach as above for Equation 3.45 combined with the Fa1 emissions scenario (Penner et al., 1999) could be used to derive the form of the F-AGTP. Discretising Equation (3.42) for one AGTP per year:

$$\text{F-ATR}\,H_x = \frac{1}{H}\sum_{k=1}^{H}\text{F-AGTP}\,k_x \tag{3.46}$$

Figure 1.1 shows two examples of the calculated F-ATR20 data, that were used as an input for the ATM optimisation (Grewe et al., 2014a,b).

3.8. Verification

Validating the global climate response to a local aircraft emission is not feasible, as this would require tracking the emitted species for the duration of their atmospheric lifetimes and accurately accounting for the climate response due to all other natural and anthropogenic perturbations. Grewe et al. (2014b) use results from other studies to achieve a limited verification of the REACT4C results, insofar as is possible given the novelty of this research.

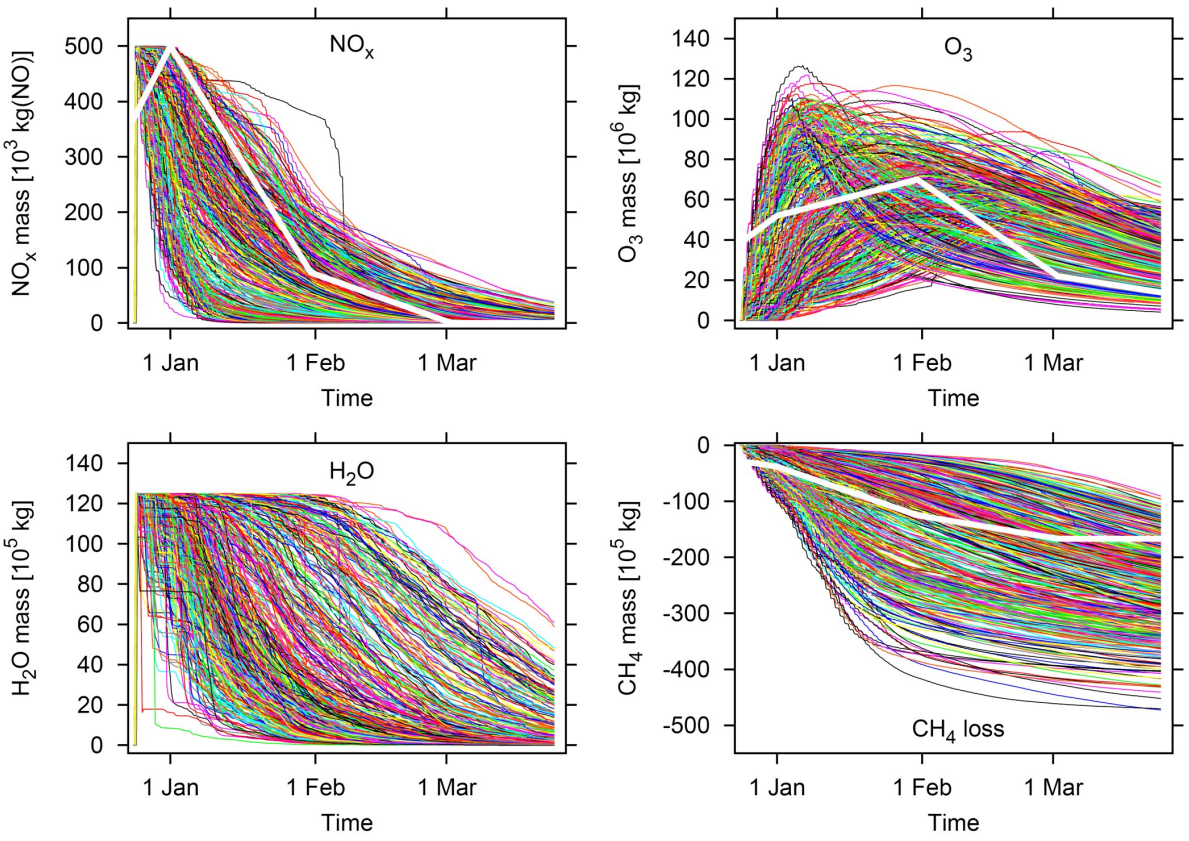


Figure 3.8: Temporal evolution of NO_x, ozone, water vapour and methane perturbation masses after emissions of NO_x and water vapour into the time-region grid (see Section 3.3.3). Each line represents one time-region grid point. The white overlaid lines are re-scaled results from Stevenson et al. (2004). The image is from Grewe et al. (2014b).

Figure 3.8 shows the time history of the perturbations for the 90-day model integration after summing the 50 trajectories for each time-region grid point, with re-scaled January values from Stevenson et al. (2004) as white lines. The overall trend is clearly agreeable. Other verification comparisons given by Grewe et al. (2014b) yield similar results: similar order of magnitude and all variations can be explained by differences between conditions in the NAFC and global averages. Using the REACT4C actual ATM

emissions as inputs for AirClim (Grewe and Stenke, 2008; Fichter, 2009) shows very little difference in NO_x climate impact and some difference in water vapour climate impact, but the latter is very small from both models. GWP and GTP (i.e. AGWP relative to CO_2) results for ozone and methane were calculated and compared favourably to Fuglestvedt et al. (2010). These comparisons to literature are all agreeable, but this thesis will investigate the significant variability between time-region grid points as visible in Figure 3.8. This variability could not be verified by Grewe et al. (2014b) as no comparable data are available for the local variability of the normalised climate impact within a region.

4

Methodology for preparing algorithmic CCFs

This chapter documents the methodology applied in this thesis. Figure 4.1 is a graphical representation of the steps and methods employed in this thesis per emission species, insofar as they directly lead to the research objective (Section 1.2). The structure of this chapter roughly follows Figure 4.1 and the chronological order in which methods were applied. Section 4.1 defines the terminology used from this point on in this thesis, Section 4.2 explains how the REACT4C data was prepared for analysis, Section 4.3 explains how the variability of the CCF itself is analysed and visualised. Section 4.4 gives the requirements that were devised as a basis for the methodology behind algorithm development in this thesis. Section 4.5 explains the methods used for formulating algorithms, and Section 4.6 details the trade-off framework developed from the requirements. Finally Section 4.7 explains how the original data and final algorithm results are visualised for reference.

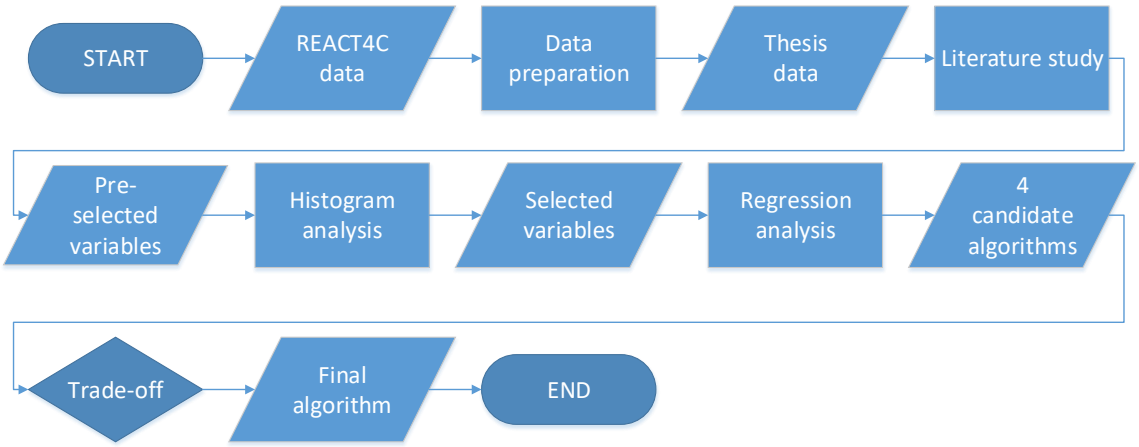


Figure 4.1: Flowchart of the steps used for preparing algorithmic CCFs from REACT4C data. Steps that do not feed directly into the results are omitted.

4.1. Terminology

For the sake of this thesis, Climate Cost Function (CCF) refers to the global-mean F-ATR20 (see Section 3.7) from concentration changes of a species from a local emission, normalised to 1 kg of emission. The five winter and three summer patterns from Irvine et al. (2013) (see Section 3.1) are sequentially numbered WP1-WP8 here, so WP6-WP8 are the summer patterns. For example, Figure 1.1 contains ozone and methane CCF results for 200 hPa, WP1.

4.2. Preparation of data for analysis

The REACT4C CCF data was presented as needed for the ATM optimisation, on a $28 \times 18 \times 6$ grid over the North Atlantic, with WP1 also including three time coordinates. These results were calculated on a $6 \times 7 \times 4$ grid, as explained in Section 3.3.3, and interpolated afterwards. The two extra altitudes in the dataset are copied from and extend above and below 200 hPa and 400 hPa respectively, to prevent the ATM optimisation having exceeding the simulated altitudes. The interpolated data points increase the amount of computation needed for analysis, though this is not expected to be significant. Using the interpolated data is likely to inflate the statistical power of any analysis done. No more information is contained than in the original data, and due to interpolation all the added points will follow the same trend. If regression analysis is then done, the results will change little when compared to regression on the original data, but the accuracy of the fit found would be overestimated. For this reason the REACT4C CCF data is reverse-engineered to its original resolution (Section 4.2.1).

All other data from REACT4C, e.g. temperature or geopotential, were calculated and presented on the higher-resolution EMAC grid. This is a $128 \times 64 \times 41$ global longitude-latitude-level grid, of which 28×18 in the NAFC conform the interpolated CCF data. The vertical coordinate is a hybrid pressure coordinate system instead of pressure altitude, and the data is available in 15-minute time steps. For application of regression analysis, the EMAC data is interpolated to the same $6 \times 7 \times 4$ grid as the CCF data (see Section 4.2.2). For accurate 2-D plotting of weather patterns, basic weather variables such as wind vectors and geopotential are converted to a $28 \times 18 \times 4$ grid that matches the interpolated CCF data.

Figure 3.4 shows an overlay of both grids. The scope of the analysis for this thesis is the geographical extent of the CCF time-region grid. In total, around 13.3 GiB of data are made available from REACT4C. After selecting the time-step and geographical extent of the emissions and interpolating, around 6.1 MiB of data is used for analysis.

4.2.1. CCF reverse engineering

Visual inspection of the provided CCF results along one latitude or longitude for one level show unmistakeable bilinear interpolation used to convert from the CCF grid to the EMAC grid without conservation of the original data points. Simply re-interpolating the data would introduce unnecessary errors by rounding off the peaks present in the original data. Here the data are extrapolated from two nearby points instead of interpolated. Due to the bilinear interpolation initially used, this approach returns exactly the original CCF data on the time-region grid.

4.2.2. EMAC data interpolation

Correctly interpolating the EMAC data to the time-region grid is not trivial. However, the global extent of EMAC model data and the interpolation to a more course grid means that no extrapolation is needed and the introduced interpolation error is limited. Both bilinear and bicubic interpolation per sigma-pressure level were initially applied. Bilinear interpolation is finally applied here due to its simplicity and possible overcorrection from application of the bicubic scheme.

Certain variables, e.g. wind direction and relative humidity, are not continuous and interpolating them directly produces varying amounts of error. Interpolating two wind directions that vary slightly, e.g. 359° and 1° (azimuth, so 0° is north) leads to the interpolated result of 180°, which is the opposite direction. Wind data is thus interpolated in cartesian coordinates and then transformed to polar coordinates. Relative humidity varies from 0% to about 120%, and 120% is not at all similar to 0%. For relative humidity, direct interpolation is expected to produce negligible error compared to interpolating the underlying data and recalculating the metric.

The EMAC model uses sigma-pressure levels for vertical discretisation under the hydrostatic assumption (Roeckner et al., 2003), that gradually move from following the Earth's surface at low altitudes to fixed pressure levels at high altitudes. As altitude is commonly given in pressure [Pa] instead of geometric height, there is no convenient way to directly apply trilinear interpolation without applying weighing factor or converting to geometric height. Here instead the data are first interpolated bilinearly per sigma-pressure level. Then the data at each longitude-latitude coordinate are linearly interpolated to the fixed-pressure levels of the time-region grid (see Section 3.3.3). Roeckner et al. (2003) gives the following relations for the sigma-pressure levels:

$$p_{k+1/2} = A_{k+1/2} + B_{k+1/2} p_s \tag{4.1}$$

$$p_k = \frac{1}{2}(p_{k+1/2} + p_{k-1/2}) \tag{4.2}$$

Here k is the level, indermediate indices represent grid-cell boundaries, and p_s in [Pa] is the surface pressure. At the lowest sigma-pressure level ($k = n_{lev}$), $A = 0$ and $B = 1$ so the surface pressure is obtained. At the highest level ($k = 0$), $A = 0$ and $B = 0$ so the pressure is 0 Pa, with A reaching a maximum roughly at the median pressure level and B monotonously decreasing with altitude. This produces levels that entirely depend on surface pressure at low altitude, and pressure levels entirely independent of surface pressure at high altitude with a gradual transition. Grewe et al. (2014b) used 41 levels, for which the A and B constants are not published, but are encoded in the data made available. It is assumed that the zonal variation of longitude distances is negligible within each 2.8° × 2.8° EMAC grid cell, so the interpolation is performed in [°] and not converted to cartesian distances.

Much literature exists on various and complex methods of interpolating weather and climate data. Except for nearest-neighbour, bilinear interpolation is generally the simplest method available. Given the large amount of uncertainty already present in the data, researching more complex interpolation methods was out of the scope of this thesis.

4.3. Assessing the variability of CCF data

To assess general spatial and seasonal tendencies of the REACT4C CCFs, sets of box plots are generated dividing the entire data set up into longitudes, latitudes, altitudes, and weather patterns, i.e. keeping all but one dimension constant. Box plots allow for compact visualisation of the variability as opposed to e.g. 2-D heat maps for each altitude and weather pattern. If an interaction between two dimensions is suspected, sets of box plots can be generated per factor along one dimension, i.e. keeping all but two dimensions constant. This is most readily done with the season (two values) or level (four values), as otherwise the compactness of box plots is lost. Tukey box plots are used for visualisation with the addition of the mean. The lines of the box represents the first, second (i.e. the median), and third quartiles. The whiskers extend either to minimum and maximum or to 1.5 times the length of the box, i.e. the Interquartile Range (IQR), whichever is shortest. Figure 4.2 shows an example for the ozone CCF data per WP. From this example it is clear that the ozone CCF results are higher for summer and that there is variability within the seasons.

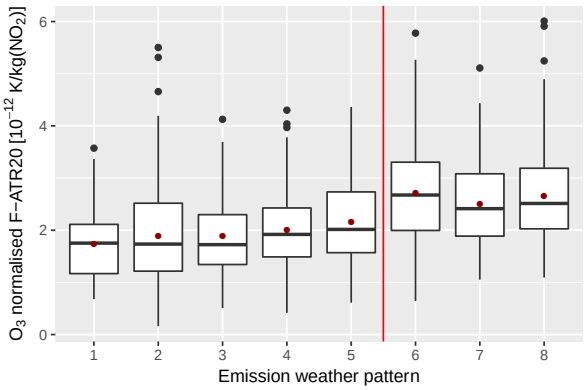


Figure 4.2: Box plot of REACT4C ozone CCF data per WP. The red line separates winter and summer WPs, the dark red dots represent the mean. Other details are standard for Tukey box plots.

Outliers can be readily identified from the box plots. Careful judgement is needed on whether to include these outliers in the analysis, as removing out-lying data points that represent the actual variability of the problem will lead to incorrect results. Key information in determining the validity of an outlier include the magnitude of its variance from the other data, its location and our understanding of the expected spatial variability, and whether the data adjacent to the outlier show a similar trend or it is entirely isolated.

Ozone and methane climate impacts are caused by the same local NO_x emission mass, and partially depend on the same reactions (see Sections 3.4.3 and 3.4.4). An attempt is made to correlate methane

impact to ozone impact to compare results with literature, and potentially avoid the formulation of a separate algorithm for methane (see Section 6.1.3).

4.4. Requirements for an algorithmic CCF

As the algorithms are intended for implementation in a NWP routine, a number of constraints can be devised to guide the analysis. Given the intention to implement in a different software package than that employed for this thesis, an algorithm should be portable. In practice this means a limited number of steps in the calculation, and closed-form expressions with a limited number of terms. The algorithm should not be difficult to present or to discretise and implement in a different routine. Instantaneous, zero-dimensional expressions are preferred as integration in space and/or time means the algorithm is sensitive to the set-up of the model it is implemented in.

The algorithm should be externally valid to an extent, so it has merit outside of the geographical extent of the NAFC and the specific weather situations and seasons of the eight weather patterns from REACT4C. Deriving algorithms from the REACT4C data that are valid at e.g. surface pressure over the South Pole is infeasible. On the other hand, algorithms that perfectly predict the REACT4C data yet are invalid for a slight deviation in geographical extent or season have no scientific merit. The season and weather pattern should thus not be used directly in the algorithm, and caution should be used when considering longitude, latitude, altitude or derivatives thereof as variables. Limitations in the applicable domain of the algorithm are acceptable but should be clearly described. Future verification activities, that are out of scope for this thesis, are needed before any claims can be made about the external validity and veracity of an algorithm.

Simplicity is preferred so the algorithm and its derivation can still visually conveyed and supported with theory. If possible only one step should be used that consists of a closed-form expression with one or two independent variables, so 2-D/3-D scatter plots can be used. The final goal in this field of research is for the algorithmic CCF to be implemented and used for climate-optimised ATM, thus simplicity will be key in convincing various stakeholders in the future.

Causality is not expected to be a major issue in this research. The EMAC model equations are known, the REACT4C-specific set-up is known, the investigated variables are based on literature reviews, and the climate impact of an emission physically depends on meteorology. Note that causality within the REACT4C simulated data is not necessarily physical causality as not all physical processes are modelled, and processes that are not accurately represented at the $2.8° \times 2.8°$ resolution are parametrised.

If an algorithm is to be implemented in NWP, variables must be used that are available in these routines. Essentially this includes all physical variables of the atmosphere, e.g. primitive variables such as wind and temperature, precipitation, the location of the tropopause, cloud physics, lightning incidence, and solar irradiance. NWP routines do not normally incorporate chemistry, and not all routines are designed for specific details such as lightning. Some NWP routines are already optimised for real-time prediction of e.g. the wind field for calculating wind-optimal aircraft trajectories. Algorithmic CCFs that only make use of variables currently available in NWP or derivatives thereof are preferred, and variables that describe large-scale variations of the atmosphere are preferred once more to minimise the computational cost of NWP implementation. Chemical Weather Forecasting (CWF) is a promising new field of research that combines NWP with chemical reactions and transport, but research has focussed on surface-level air pollution (Kukkonen et al., 2012). There is thus some potential for future real-time CCF calculation based on chemical variables, and they should be included in the scope of the analysis.

Reproducing the spatial pattern of the CCFs is more important than absolute accuracy. The REACT4C data is already quite uncertain (Grewe et al., 2014b), but an incorrect pattern from an algorithmic approximation leads to incorrect re-routing decisions.

4.5. Formulation of algorithmic CCFs

The research objective of this thesis (Section 1.2) is to formulate algorithmic approximations, but this must be operationalised to formulate a methodology. Given the requirements for the algorithm discussed in Section 4.4 and the timeframe of this thesis, the scope was defined as 0-D instantaneous regression analysis on the entire REACT4C CCF dataset for 12 UTC emissions. The distinction between regression and algorithm in this thesis is only made when suggesting extra calculation steps beyond the regression equation to e.g. prevent negative CCF results.

4.5.1. Data selection for regression

As a modular CCM, EMAC tracks a multitude of variables during a simulation and many more are calculated in post-processing. The approach agreed for this thesis was to request specific datasets based upon a literature review per emission species. An initial dataset is provided with the CCF data and basic weather variables: geopotential altitude, relative humidity, surface pressure, temperature, and wind in polar coordinates.

A literature review is performed per species, to identify driving variables behind the CCF variations. A distinct part of this process is to review causal links in the EMAC background model by examining the model equations and the driving variables. The most valuable source of this information outside of Grewe et al. (2014b) is the ECHAM5 model description report (Roeckner et al., 2003). Relevant datasets are made available based on the results from these literature reviews, as opposed to specific variables.

The hypothesis of this thesis (Section 1.2) and the expectation expressed in Frömming et al. (2017) is that the CCFs can be accurately represented by local weather data. As discussed in Section 4.4, inclusion of non-weather variables in an algorithm limits the potential for application of algorithmic CCFs. The method applied here is to include non-weather variables in the analysis, but keep a clear distinction in formulation, selection, and discussion of algorithms between those that are compatible with NWP application and those that are not (e.g. background chemical concentrations).

The list of variables in each dataset is output to a text file, and the descriptions of the variables are scanned manually for data that matches those identified in literature and otherwise potentially relevant variables. The same variables may be included in multiple datasets, and these are plotted to determine if and why there are differences. Some variables identified in literature will not be available directly in the dataset and are derived. At a later stage more variables may be derived to facilitate specific interactions between basic variables.

Histograms and KDEs are output for every selected variable, for the full set of $6 \times 7 \times 4 \times 8$, to assess whether the variable has any meaningful variation and what distribution the data has. Variables with no variation, very little variation, or exact copies of another included variable are excluded from further analysis. Many weather variables are inherently sparse, e.g. snow or lightning, and some may have been unused in the REACT4C model set-up. By excluding variables with little variation, there is some risk of excluding information that explains outliers in the CCF data. Generally the sparse data and their outliers line up across multiple variables (e.g. all snow data, all lightning data). This risk is then mitigated by keeping one of the variables in the analysis. Tables 5.1 and 6.1 contain the variables before and after histogram selection for water vapour and NO_x analysis respectively.

4.5.2. Regression analysis

The first step made in formulating algorithm candidates is to generate a scatter plot matrix of the CCF and all the selected independent variables mutually. This visualises the scatterplot of each pair of variables with a simple univariate fit, the histogram/KDE of each variable, and the correlation coefficient of each pair in one overview. The Spearman rank coefficient is employed here as opposed to the Pearson product-moment correlation coefficient. The Spearman coefficient can return high results for all monotonous relationships, whereas the Pearson coefficient purely looks at linear correlation and we anticipate non-linear relationships. There is no convenient coefficient for non-monotonous relationships so these should be visually inspected in scatter plots. The scatter plot matrix is used to identify the strongest univariate relationships between the CCF data and weather data, and to identify potential collinearity among the weather variables. Generally, two medium strong predictors that have no collinearity are better combined in a bivariate regression than two strong predictors that have full collinearity, as here the second variable would add no new information. Figure 4.3 shows a scatter plot matrix of the water vapour, ozone and methane CCF data as an example. Both the Spearman coefficient and the scatter plot show that the strongest correlation is between ozone and methane. If ozone and methane CCF data would be used to predict the water vapour CCF, the collinearity would make their combined predictive power significantly lower than the coefficients of 0.11 and -0.03 suggest.

More detailed univariate scatter plots of the CCF over each independent variable are generated, and shape- and colour-coded per dimension based on results from the boxplot analysis. This allows more detailed exploration of the relationships and formulation of options to pursue. For the strongest predictors, box plots are generated for all four dimensions to assess whether the variability shows similar trends to the CCF. If one or more dimensions in the CCF box plot analysis are shown to be

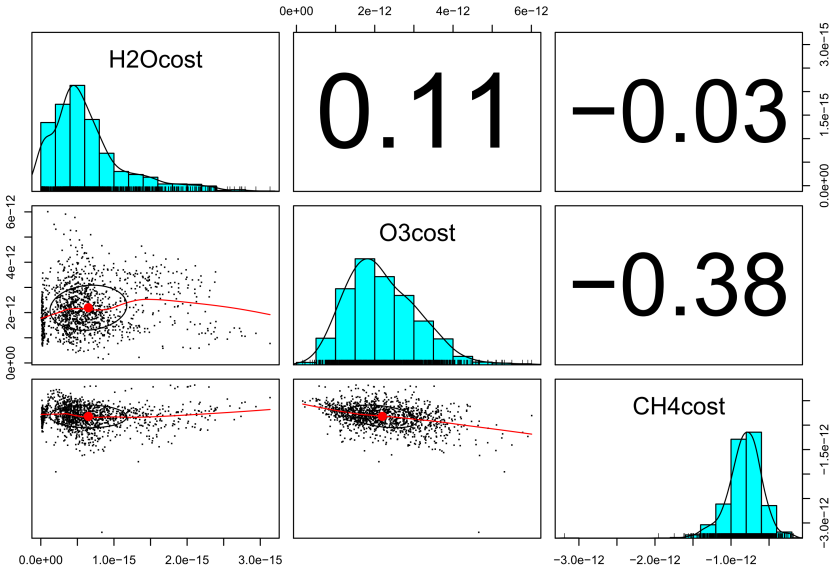


Figure 4.3: Scatterplot matrix of water vapour (H2Ocost in [K kg(fuel)$^{-1}$]), ozone (O3cost in [K kg(NO_2)$^{-1}$]), and methane (CH4cost in [K kg(NO_2)$^{-1}$]) CCF data from REACT4C. Histograms and KDEs are shown on the diagonal. Spearman rank coefficients are shown above the diagonal. Scatter plots with LOESS non-parametric fit are shown below the diagonal.

leading, a box plot for that dimension is generated for all predictors.

By performing 0-D regression, we are committing what is termed the atomic fallacy. Each data point is taken as an independent measurement (such as a person or an atom) for regression, while both measured and simulated meteorological data show strong spatial patterns. Spatial autocorrelation is taken into account for residuals in the algorithm trade-off. Methods for spatial regression were investigated, but are out of scope for this thesis due to time constraints.

As the REACT4C data contains five winter and three summer patterns, and their ability to represent all seasonal variation is yet unclear (Section 3.1), this atomic fallacy may also affect the ability of an algorithm to capture seasonal trends. If the seasonal trend is not similar in the predictor(s) and season is not added as a separate category, the balance of weather patterns will skew results towards winter values. Adding season to the algorithm presents a separate problem of using the small sample of eight patterns to determine a constant factor between summer and winter CCFs without any information for spring or autumn.

4.5.3. Presentation of regression results

The general form of a simple linear regression is:

$$y^j = \beta_0 + \beta_1 \times x_1^j + \epsilon^j \tag{4.3}$$

With (j=1,...,n) where n is the number of observations. The two β coefficients are the parameters that are optimised via least-squares to best predict the dependent variable y from the dependent variable x. The residuals ϵ are the part of y that cannot correctly be predicted by this regression which are minimised via least-squared in the form $\sum_{j=1}^{n} \left(\epsilon^j\right)^2$. In this thesis the notation $y = \beta_0 + \beta_1 \times x_1 + \epsilon$ is used with the understanding that we are working with vectors. If more predictors are used it is no longer simple regression, and there is a set of parameters β_i with (i=0,...,m) where m is the number of predictors in the model.

For univariate regression, or bivariate regression where at least one of the predictors is categorical, the fit can be conveniently visualised in a scatter plot of CCF data versus the independent variable with the algorithm presented as a fitted curve or multiple fitted curves.

For bivariate and higher regressions, the data is visualised in a set of partial residual scatter plots (also called component+residual plots). The x-axis of each sub-plot is original data of each independent variable or interaction term between variables in the regression. The y-axis for predictor i is formulated as $\beta_i \times x_i + \epsilon$, i.e. the residuals of the full regression model for each value of the x-axis plus the regression coefficient for that independent variable times the value of the variable. This allows interpretation of the relationship between the independent variable, the dependent variable, and the regression given that other variable are included.

Linear regression models are presented as formulas with the β_i coefficients for each term. The least-squares optimised estimates for β_i, the standard error, the t-statistic, and the two-sided p-value of the t-statistic are provided in tables. The t-statistic is β_i divided by its standard error (the null hypothesis is that β_i is zero), and the p-value gives the probability of this result being due to chance using the Student's t-distribution. The maximum p-value allowed for statistical significance is often 0.05 across many disciplines, but is set at 0.001 for this analysis to limit the amount of algorithm candidates. With the large dataset made available for this thesis, a p-value of 0.001 is easily attainable and reducing the chance of a correlation being due to random variability is worthwhile if the results of this thesis should have any merit for external application. Non-linear regressions are presented similarly, but the coefficients are not called β_i to avoid confusion with linear regression.

To test the accuracy of the algorithm across the four dimensions, each is added separately to the regression to see how much the adjusted R^2 (see Section 4.6.2) improves and whether the resulting β_i coefficients are statistically significant.

4.6. Algorithm trade-off

Based on the algorithm requirements in Section 4.4, a set of criteria for selecting the final algorithm were devised. In the process of investigating and formulating algorithms, many regressions are considered with and without additional calculation steps. To manage time and effort, four arbitrary candidates per species are worked out in detail and presented in this thesis. These candidates are based on the maximum achievable adjusted R^2 for a given amount and type of variables, and intended to present meaningfully different results to trade off. The set of criteria presented below is used to select a final algorithm per species.

4.6.1. Pattern of residuals

The most important criterion for a non-perfect algorithmic CCF is that deviation from the original CCF does not lead to an incorrect routing decision when implemented. In theory errors are acceptable, as long as they are constant per weather pattern. As a further approximation, errors between emission levels are ignored when quantifying the residual pattern.

The first method used is to create Tukey box plots for the fit, i.e. the algorithm, and the residuals for all four dimensions. Visual inspection of the residual box plots can reveal remaining trends in the residuals, and does not neglect inter-level or weather pattern variation.

The second method used is calculation of the range of the residuals from the algorithm within each longitude-latitude field. The range of these ranges is represented by the mean, which gives a single number to report. For options with equal general accuracy, this allows a consideration of whether the errors are conveniently distributed for this application.

The third method used is a spatial autocorrelation test of the residuals per longitude-latitude field. Several options exist, but here Moran's I (Moran, 1950) is applied. This metric varies from -1 to 1, with -1 representing negative spatial autocorrelation i.e. a checkerboard pattern, 1 representing perfect positive spatial autocorrelation i.e. a landscape with the horizon in the middle, and 0 representing no spatial autocorrelation. The distances between data points in the grid are converted to Cartesian distances on the globe, and the weighting function used is reciprocal distance. The p-value used is 0.05, and only significant results with a negative I are considered problematic. If all errors in a longitude-latitude field are perfectly grouped, the likelihood of an incorrect re-routing is smaller than if high and low errors are scattered around.

4.6.2. General residual error

Residual versus fit scatter plots are generated for every algorithm candidate. These allow inspection of normality of the residuals, remaining trends between residuals and fit i.e. variability that could be explained by higher order terms, and heteroskedasticity of the residuals i.e. whether the variance is roughly constant for all fitted values or varies. Heteroskedasticity is mainly a problem for accurately calculating confidence intervals.

The Residual Standard Error (RSE) is used to compare the algorithms on accuracy across the entire dataset. It is calculated from the square root of the mean of the squared residuals. For linear regression models, the adjusted R^2 is also presented, which is R^2 adjusted by the numbers of parameters and observations. The R^2 metric represents the fraction of the variance in the dependent variable that is

explained by the regression and gives a convenient metric for the accuracy of the algorithm. The R^2 is not defined for non-linear regression, constant regression, or linear regression with no intercept and thus not reported in these cases.

4.6.3. External veracity

Literature is used to estimate ranges for the independent variables of the regression in other seasons and geographical extents, and the resulting range of the algorithmic CCF is compared to what is known about regional differences in climate impact from literature. As explained in Section 4.4, veracity in entirely different spatial domains is not expected, but veracity for moderate shifts in spatial domain, for different weather situations, or for spring/autumn is critical for application of algorithmic CCFs.

4.6.4. Mechanistic background

Expectations on the nature of the relationship between the dependent and independent variables are revisited, based on literature and reasoning. To avoid fitting noise instead of meaningful variation, it is important to identify causal relationships and avoid complex fits such as higher-order polynomials if they have not been identified in previous research.

4.6.5. Parsimony

All else being equal, a univariate linear regression is preferable. Statistical tests and metrics are more readily available for linear regression, and the less variables an algorithm employs the easier it is to visualise and the lower the risk of fitting noise.

4.6.6. Weighting of categories

No weights are assigned to the categories (i.e. Sections 4.6.1–4.6.5) in any attempt to make an objective trade-off. The final algorithm candidate is based on a subjective weighting and the other candidates are presented in equal detail to allow future work to apply a different algorithm based on different subjective criteria.

4.7. Presentation of 2-D contour plots

As has been explained, in REACT4C the CCF results were bilinearly interpolated to the EMAC grid. Grewe et al. (2014b) do not provide details how further interpolation progressed during the ATM optimisation, i.e. what CCF value is used for the section of a route between two grid-points. The most likely options are bilinear or nearest-neighbour interpolation, depending on how the ATM routine discretises airspace. Papers on the REACT4C results (Grewe et al., 2014a; Frömming et al., 2017) choose to show the CCF results using interpolated contours instead of a grid (i.e. Figure 1.1). The same is true for AirClim results (Grewe and Stenke, 2008; Dahlmann et al., 2016). For these two reasons, i.e. representing actual usage of the results and allowing comparisons to literature, the CCF data and algorithmic results are shown on 2-D longitude/latitude fields with bilinearly interpolated contours. Bilinearly interpolated isolines of the geopotential are overlaid, as is the wind vector field. The drawback is that 32 plots are needed per species and seasonal and spatial trends, especially for altitude, are not as obvious as in box plots. Frömming et al. (2017) present similar results of the REACT4C data for all weather patterns but only at 250 hPa. All contour plots are included in the appendices with the original and algorithmic plot for each level and weather pattern side-by-side to allow visual comparison.

5

Water vapour algorithmic CCF results

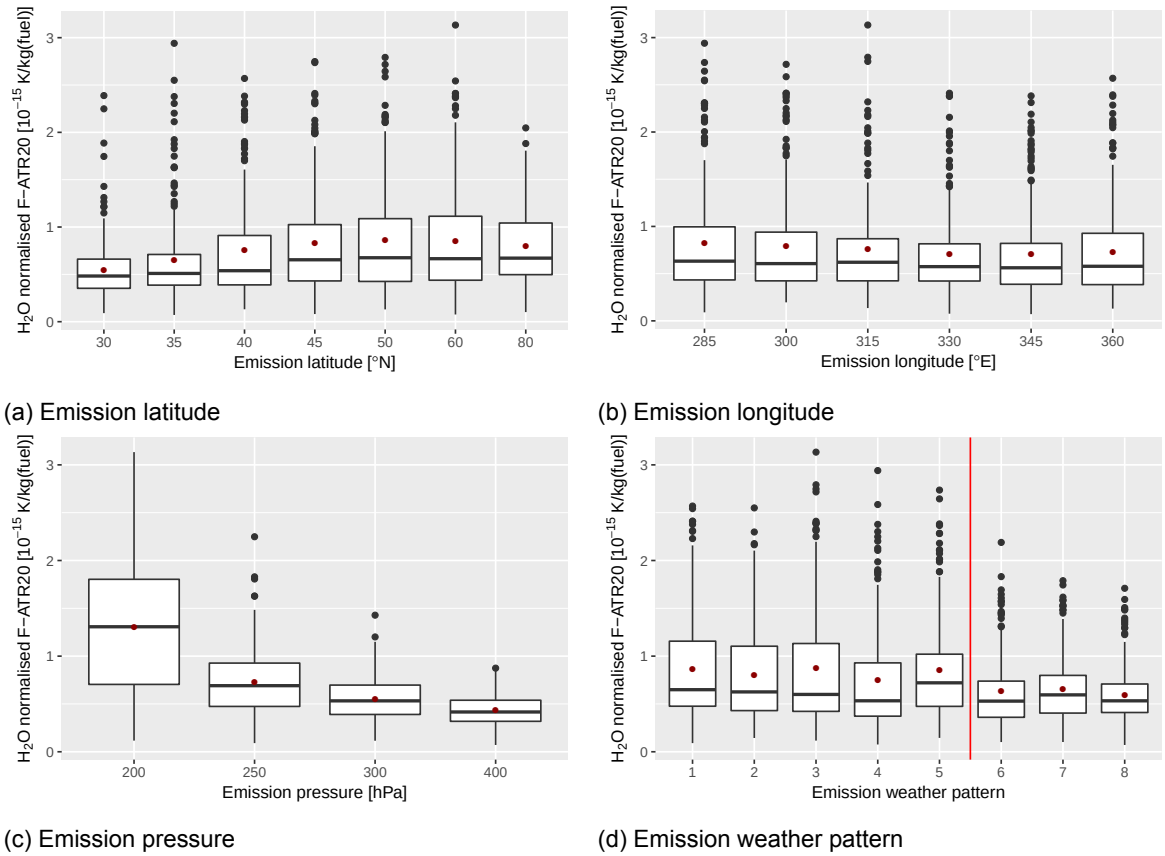
Following the methodology presented in Chapter 4, here REACT4C water vapour CCF results are described and analysed (Section 5.1), the literature review and selection of regression variables is presented (Section 5.2), four algorithms are discussed (Section 5.3), the trade-off is explained and the final algorithm is chosen (Section 5.4), and finally a discussion on all results for water vapour is presented (Section 5.5).

Note that the water vapour CCF data in [K kg(fuel)$^{-1}$] for WP1 were multiplied by 24 to correct an unidentified error. The RF data for WP1 show no discrepancies with the other weather patterns, and a multiplication of around 24 brings the CCF data in line with the other weather patterns, so it is assumed that an error occurred in converting [hours] to [days] in the RF extrapolation for calculating the climate metrics (see Section 3.7).

5.1. Dimensional variability of REACT4C CCFs

Figure 5.1 shows box plots of the CCF data for each of the four dimensions. There is some meridional variability, with higher latitudes representing higher CCF results. This may be caused by the emission pressure levels being constant while the tropopause is naturally lower at higher latitudes. This is also suggested by Gauss et al. (2003), to explain the sensitivity to polar re-routing in their results. CCF data for emissions at 80°N are slightly lower and have less variance than for emissions at 60°N, so a linear trend cannot be assumed. There is low zonal variability. Mid-Atlantic longitudes show slightly lower variability but this is negligible compared to the variability along other dimensions and no fundamental variability along longitudes is expected due to the prevailing westerly winds (see e.g. Figure 1.1).

There is significant altitudinal variability, with non-linear increase for decreasing pressure. Heteroskedasticity is present, with the variance increasing proportionally to the mean. The monotonic increase for increasing altitude is in overall agreement with other studies, and the most likely cause is the increasing proximity to the stratosphere and hence longer atmospheric lifetime of a water vapour emission. The heteroskedasticity may be caused by emissions at lower altitude (400 hPa) always staying in the troposphere, while emissions at a higher altitude (200 hPa) may be emitted into and spend the majority of their lifetime in either the troposphere or the stratosphere. An extra simulated altitude in the stratosphere (100–150 hPa) would aid in assessing this. Wilcox et al. (2012b) found an almost linear relationship between water vapour RF from January emissions between 200 and 290 hPa, but this is normalised from all aviation emissions at that altitude in the Northern Hemisphere. Considering, in the troposphere, the negligible effect of aviation emissions on humidity, non-linearity in the physical processes is not expected to be relevant. Figure 5.1 shows 336 (concurrent) simulations per altitude however, while Wilcox et al. (2012b) ran one simulation per altitude. Results from Fichter (2009), more readily available in Dahlmann et al. (2016), are difficult to interpret due to a non-linear colour scale, but when reverse-engineered they show a stronger global-mean year-mean non-linear RF altitude trend for sustained emissions than these results, within the same latitude-altitude domain. Gauss et al. (2003) previously found that a 1 km increase of global subsonic cruise altitude causes a doubling of the water vapour perturbation. Fichter (2009) also investigated the effect of global altitude shifts of aviation on the global climate impact, and found that the difference due to water vapour emissions is of similar



(a) Emission latitude

(b) Emission longitude

(c) Emission pressure

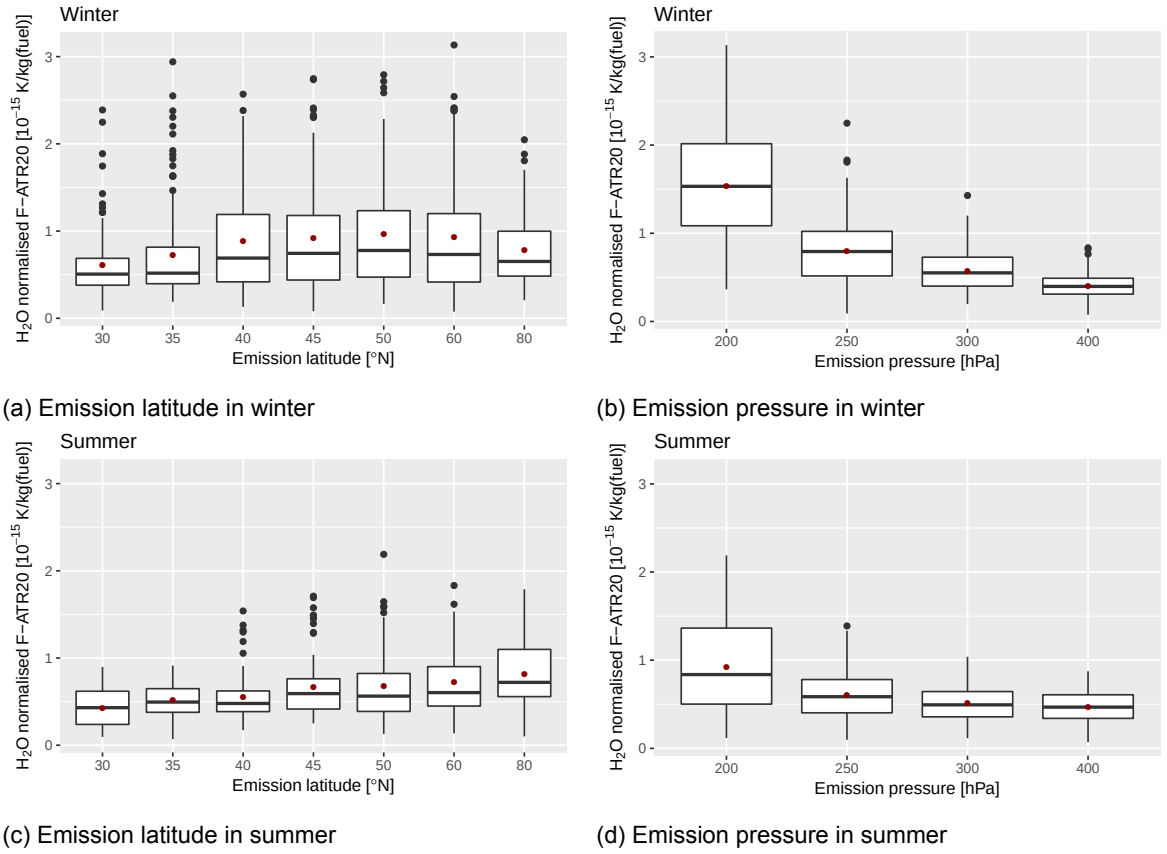
(d) Emission weather pattern

Figure 5.1: Box plots of REACT4C water vapour CCF data for each dimension. The red line represents the distinction between winter (1–5) and summer (6–8) weather patterns.

magnitude to other species. Water vapour is known to be a relatively small contributor to the climate impact of aviation (Section 2.2.3), thus results from Fichter (2009) show that the altitude trend in water vapour emission climate impact is very strong compared to its overall impact.

There is some seasonal variability, with lower mean CCFs and lower variance in summer. This is expected due to higher summer tropopause altitudes and higher background humidity, thus lower atmospheric lifetimes for a water vapour emission. There is some variation between the constituent weather patterns per season, with WP4 showing lower results. This suggests that the weather pattern distinctions devised by Irvine et al. (2013) (Section 3.1) are not of great importance for the variability of water vapour emission climate impact. The difference in variance between summer and winter data appears stronger than the difference in mean/median values, thus the skewness of the distributions is strongly determined by the season. Wilcox et al. (2012b) found a sinusoidal relationship between water vapour RF and month for year-round emissions with a peak in March, but this is not readily comparable due to the year-round emissions modelled with some seasonal variation in the emissions. Considering typical lifetimes of water vapour emissions in the REACT4C simulation domain of several weeks (Grewe and Stenke, 2008), a peak in March agrees with winter emissions leading to higher CCFs. Gauss et al. (2003) also found the largest RF in spring, mid-northern latitudes for year-round cryoplane emissions.

Figure 5.2 shows box plots of the water vapour CCF data for latitude and level after splitting into summer and winter data. Summer emissions show a monotonic trend for increasing latitude, while winter emissions show a maximum for moderate latitudes, which explains the non-monotonic trend when viewing all data. Winter emissions show more heteroskedasticity with latitude, with moderate latitudes showing the largest variance. A potential explanation for the difference between 60°N and 80°N is the transition from Ferrel cell to polar cell and location of the polar jet around this latitude. This could lead to 80°N emissions being transported northwards instead of southwards like the others, and raining out sooner. The difference between summer and winter emissions could then be explained by seasonal precipitation difference at the North Pole.



(a) Emission latitude in winter

(b) Emission pressure in winter

(c) Emission latitude in summer

(d) Emission pressure in summer

Figure 5.2: Box plots of REACT4C water vapour CCF data per level and latitude split into season of emission.

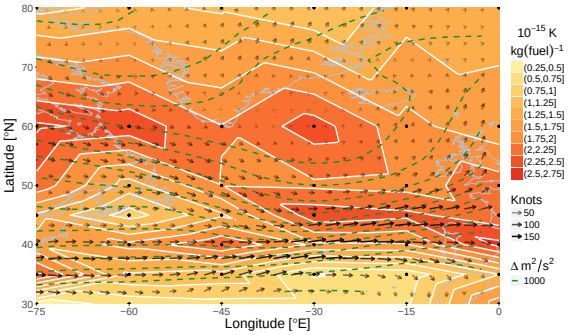
The non-linear trend with altitude is stronger for winter emissions than summer emissions. Emissions at 300 hPa and 400 hPa show very little seasonal trend in their CCF results, suggesting that the differences in mean and variance at higher altitudes are due to seasonal differences in tropopause altitude. Winter emissions at 200 hPa show a higher mean and higher variance here, strengthening the hypothesis that the CCF is largely determined by the tropopause.

Figure 5.3 (left) shows bilinearly interpolated contour plots of the water vapour CCFs data for 200 hPa, WP1 as this is the case shown in Figure 1.1 and Grewe et al. (2014a), though no water vapour results are presented there. The lowest values occur at 30°N, and the highest values occur at 40–60°N. There is no directly discernible correlation with the wind or geopotential field, and there are many complex details in the pattern of the water vapour CCF data for this altitude and weather pattern. Appendix A (left side) contains these visualisations for all 32 combinations of altitude and weather pattern.

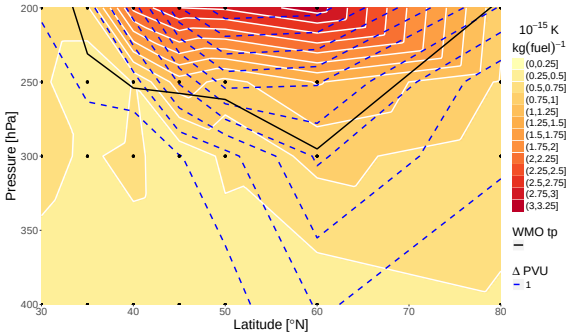
Figure 5.3 (right) shows bilinearly interpolated contour plots of the water vapour CCFs data for 315°E, WP3 as this case contains the single highest value according to Figure 5.1. The altitude trend identified in Figure 5.1 is present, but is either convolved with a non-monotonic latitude trend as was also previously identified, or better explained as a trend relative to the tropopause. The highest value occurs at the lowest pressure relative to the tropopause. One low value (40°N, 250 hPa) occurs for an emission released into the stratosphere, but most low values occur significantly below the tropopause. Appendix B (left side) contains these visualisations for all 48 combinations of longitude and weather pattern.

5.2. Regression variable selection

In the REACT4C study (Grewe et al., 2014b) a linear relationship between water vapour mass and RF was derived (Section 3.6.1), and a constant exponential lifetime for climate metrics (Section 3.7.2) so the CCF variation is entirely driven by the removal rate of the water vapour emission mass, which is



(a) 12 UTC, 200 hPa, WP1. The dashed lines are geopotential isolines in [1000 m^2 s^{-2}], and the arrows are wind vectors.



(b) 12 UTC, 315°E, WP3. The solid black line represents the World Meteorological Organisation (WMO) thermal tropopause, and the dashed lines are Potential Vorticity (PV) (see Section 5.3.2) isolines in [PVU].

Figure 5.3: Water vapour CCF bilinear contour plots in longitude-latitude field (left) and latitude-altitude field (right). The black points are the emissions locations i.e. the actual data that are interpolated.

taken as proportional to the background precipitation rate (Section 3.4.1). Emissions are not added to the background and cannot affect precipitation rates.

Section 5.1 mentioned several studies into the spatial and seasonal variation of aviation water vapour climate impact, but no studies were found that specifically investigated driving parameters behind the variation. The zonal-mean year-mean normalised RF results in Fichter (2009) visually resemble a non-linear relationship with pressure relative to tropopause in the troposphere, and from 150 hPa upwards a weaker relationship to pressure altitude. The latter results should not be taken into account however, as an incorrect upper boundary condition was used in this study (V. Grewe, personal communication, March 2017). Similar visual results from the prior study (Grewe and Stenke, 2008), with less simulations but more altitudes, do show a similar non-linear trend with emission altitude and tropopause altitude. Their results show striking similarity between water vapour perturbation lifetimes and RF, which led to the correlation applied in Grewe et al. (2014b) (Section 3.6.1). Note again that the REACT4C methodology uses results from Grewe and Stenke (2008) so no conclusions can be drawn from similar results. It is well established that removal processes for water vapour are much weaker in the stratosphere, leading emissions above the tropopause to have higher impact on average (Lee et al., 2010). Therefore, all data relating to the tropopause are taken into account.

The underlying hydrological physics of the ECHAM5 model are very likely to be correlated to the CCF results. Roeckner et al. (2003) describe the model set-up of ECHAM5. The EMAC mechanisms differ only in that they have been implemented in a modular set-up. Equation 10.1 in Roeckner et al. (2003) gives a differential equation for the mass ratio of water vapour with all the source and sink turns. They state that apart from convective detrainment, condensation and deposition are the most important cloud generation processes. Below -35 °C only deposition occurs, and the 200–400 hPa release altitude means that deposition is a stronger driver than condensation as most emissions occur at ambient temperatures below -35 °C. The total water mass (vapour, liquid, and ice) ratio is represented by a bell-shaped beta distribution within the model cells, which is used to calculate condensation and deposition terms. The beta distribution is determined by several constants which are not known, but should not vary internally. Thus the liquid and ice water mass ratios may be represented overall by the water vapour mass ratio. As the lifetime of a water vapour is by definition (Grewe et al., 2014b) proportional to precipitation rates, and precipitation rates depend on water vapour, liquid and ice masses, all data relating to precipitation and humidity are taken into account.

The three types of variables pursued for developing a water vapour algorithm are thus:

- general meteorology
- tropopause
- precipitation and humidity

Table 5.1 summarises all variables that were read in from the supplied datasets and down-selected by viewing in histograms.

Table 5.1: Variable overview for water vapour algorithm analysis. The 3rd and 4th dimensions in 3-D/4-D are always time followed by altitude, and bold-printed rows are those selected for analysis via histograms. Note that the time dimension is not used in this analysis and that weather pattern is not counted as a dimension for this table.

Dataset	Name	Description	Units	Dimension
CCF	**H2Ocost**	**Cost function for H_2O**	**K kg(fuel)$^{-1}$**	**4-D**
tropo	**pblh**	**Planetary boundary layer height**	**m**	**3-D**
tropo	**tp_clim**	**Climatological tropopause pressure**	**Pa**	**3-D**
tropo	**tp_PV**	**PV tropopause pressure**	**Pa**	**3-D**
tropo	**tp_WMO**	**WMO tropopause pressure**	**Pa**	**3-D**
tropo	**PV**	**Potential vorticity**	**PVU**	**4-D**
cloud	**aclc**	**Large scale cloud cover**	**0–1**	**4-D**
cloud	condensation	Condensate in cloud covered part of gridboxn	kg kg^{-1}	4-D
cloud	iwc	Large scale cloud snow/ice content	kg kg^{-1}	4-D
cloud	lwc	Large scale cloud liquid water content	kg kg^{-1}	4-D
cloud	mimelt	Large scale frozen precipitation melting	kg m^{-2} s^{-1}	4-D
cloud	misedi	Large scale ice sedimentation	kg kg^{-1}	4-D
cloud	**prec_cover**	**Large scale precipitation cloud cover**	**-**	**4-D**
cloud	rain_evap	Large scale rain evaporation	kg kg^{-1}	4-D
cloud	rain_form	Large scale rain formation inside cloud	kg kg^{-1}	4-D
cloud	rainflux	Large scale rain precipitation flux	kg m^{-2} s^{-1}	4-D
cloud	rhc	Critical relative humidity for natural clouds	%	4-D
cloud	snow_form	Large scale snow formation inside cloud	kg kg^{-1}	4-D
cloud	snow_subl	Large scale snow sublimation	kg kg^{-1}	4-D
cloud	snowflux	Large scale snow precipitation flux	kg m^{-2} s^{-1}	4-D
convect	**CAPE**	**Convective available potential energy**	**m^2 s^{-2}**	**3-D**
convect	**cth**	**Convective cloud top height**	**m**	**3-D**
convect	cv_cldwater	Convective cloud water content	kg kg^{-1}	4-D
convect	cv_lwc	Convective cloud water content (3-D in cloud)	kg kg^{-1}	4-D
convect	cv_precflx	Convective precipitation flux	kg m^{-2} s^{-1}	4-D
convect	cv_snowflx	Convective snow precipitation flux	kg m^{-2} s^{-1}	4-D
convect	massfd	Downward mass flux	kg m^{-2} s^{-1}	4-D
convect	cv_precnew	Freshly formed precipitation flux	kg m^{-2} s^{-1}	4-D
convect	cv_snownew	Freshly formed snow flux	kg m^{-2} s^{-1}	4-D
convect	conv_tte	Convective temperature tendency	K s^{-1}	4-D
convect	conv_qte	Convective humidity tendency	s^{-1}	4-D
convect	cv_rform	Convective precipitation formation (water, in cloud value)	kg kg^{-1}	4-D
convect	cv_sform	Convective precipitation formation (snow, in cloud value)	kg kg^{-1}	4-D
convect	cv_cover	Estimated convective cloud cover	0–1	4-D
convect	conv_bot	Bottom level of convection / convective cloud base	levels	3-D
convect	conv_top	Top level of convection	levels	3-D
ECHAM5	**geopot**	**Geopotential**	**m^2 s^{-2}**	**4-D**
ECHAM5	**geosp**	**Surface geopotential**	**m^2 s^{-2}**	**3-D**
ECHAM5	**tm1**	**Dry air temperature**	**K**	**4-D**
ECHAM5	**tpot**	**Potential temperature**	**K**	**4-D**
ECHAM5	**tvirt**	**Virtual temperature**	**°C**	**4-D**

ECHAM5	**rhum**	**Relative humidity**	**%**	**4-D**
ECHAM5	**aps**	**Surface pressure**	**Pa**	**3-D**
ECHAM5	**um1**	**Zonal wind times cosine of latitude**	**m s^{-1}**	**4-D**
ECHAM5	**vm1**	**Meridional wind times cosine of latitude**	**m s^{-1}**	**4-D**
ECHAM5	qtec	Convective detrained humidity	kg kg^{-1}	4-D
ECHAM5	xtecl	Convective detrained liquid	kg kg^{-1}	4-D
ECHAM5	xteci	Convective detrained ice	kg kg^{-1}	4-D
ECHAM5	**tte**	**Dry air temperature tendency**	**K s^{-1}**	**4-D**
ECHAM5	qte	Specific humidity tendency	kg kg^{-1} s^{-1}	4-D
ECHAM5	xlm1	Cloud water	kg kg^{-1}	4-D
ECHAM5	xim1	Cloud ice	kg kg^{-1}	4-D
ECHAM5	xite	Cloud ice tendency	kg kg^{-1} s^{-1}	4-D
ECHAM5	xlte	Cloud water tendency	kg kg^{-1} s^{-1}	4-D
g3b	**aprl**	**Large scale precipitation**	**kg m^{-2} s^{-1}**	**3-D**
g3b	**aprc**	**Convective precipitation**	**kg m^{-2} s^{-1}**	**3-D**
g3b	**aprs**	**Snowfall**	**kg m^{-2} s^{-1}**	**3-D**
g3b	tke	Turbulent kinetic energy	m^2 s^{-2}	4-D
g3b	**vervel**	**Vertical velocity**	**m s^{-1}**	**4-D**
g3b	**q**	**Specific humidity**	**kg kg^{-1}**	**4-D**
g3b	**xl**	**Cloud water**	**kg kg^{-1}**	**4-D**
g3b	**xi**	**Cloud ice**	**kg kg^{-1}**	**4-D**
g3b	**xvar**	**Variance of total water amount q+xl+xi**	**kg kg^{-1}**	**4-D**
g3b	**qvi**	**Vertically integrated water vapor**	**kg m^{-2}**	**3-D**
g3b	**xlvi**	**Vertically integrated cloud water**	**kg m^{-2}**	**3-D**
g3b	**xivi**	**Vertically integrated cloud ice**	**kg m^{-2}**	**3-D**

Three definitions of the tropopause pressure are included: climatological, PV, and WMO. The first is determined from previous simulations of the WMO tropopause (V. Grewe, personal communication, March 15, 2017). The PV or dynamic tropopause altitude is defined at a specific (unknown) level of PV and is contrasted with the classical WMO or thermal tropopause. This is defined as the lowest altitude at which the atmospheric temperature lapse rate (per unit altitude) decreases below the 2 °C km^{-1} threshold, provided the average lapse rate between this altitude and any other altitude within 2000 m higher is also below 2 °C km^{-1}. In the International Standard Atmosphere (ISA) the troposphere has a positive lapse rate of 6.5 °C km^{-1} and the stratosphere has a negative lapse rate, with the tropopause being an isothermal layer in between, but in reality the boundary between the two layers is not isothermal or so unambiguous as demonstrated by the difference between the PV and WMO tropopause.

Several variables were derived from the REACT4C data to facilitate analysis. First, pressure relative to WMO/PV tropopause pressure. This rescales the pressure altitude of an emission to that of the tropopause as literature suggested this distinction would be significant. Second, dry/potential/virtual temperature relative to the mean of the 2-D longitude/latitude field. This is meant to capture the other spatial dimensions when working with the relative tropopause pressure. Third, wind direction and speed in polar coordinates as this is easier to visualise and understand. Note that the REACT4C variables $um1$ and $vm1$ are scaled by the cosine of the latitude. Finally the season (winter/summer) of the weather pattern was encoded as a category. In regression analysis it might be desirable to add a switch for seasonal behaviour of the data and we do not want to explicitly use the weather patterns.

During analysis, it was concluded that the variable $geopot$ was actually geopotential relative to surface geopotential, as the geopotential field for a certain pressure did decrease with increasing latitude, but showed strong patterns with the topography of Greenland that was not visible in literature. The absolute geopotential was calculated through addition and used for all analysis, and all references to geopotential in the results refer to absolute geopotential. The resulting absolute geopotential data still show some pattern with the topography of Greenland but the isolines line up with literature on REACT4C (Grewe et al., 2014a; Frömming et al., 2017).

The virtual temperature data were found to be erroneous as they linearly depend on temperature

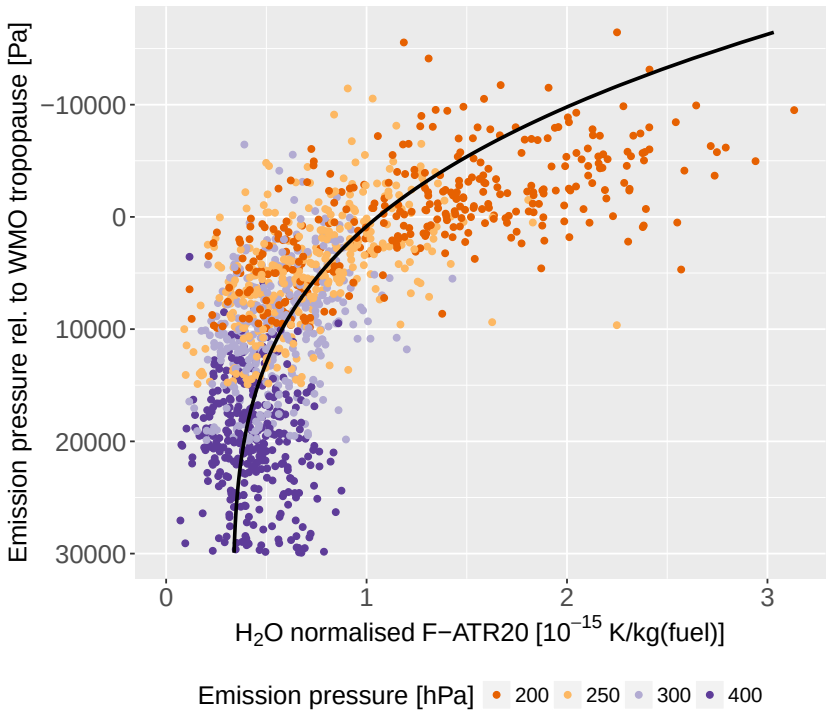


Figure 5.4: Scatterplot of water vapour CCF data versus emission pressure relative to tropopause pressure coloured by level, with the non-linear regression found shown as a black curve.

and altitude, and do not depend on humidity. The majority of the precipitation data was empty or very sparse, which is to be expected given the spatiotemporal domain of the data and the effect of interpolation to the time-region grid.

5.3. Algorithm candidates

Scatter plot matrices and regression techniques were used to identify the strongest (combinations of) predictors from the 30 variables in Table 5.1 plus dimensions and derived variables. Due to time constraints for detailed analysis and space constraints for presenting results here, four algorithm candidates are discussed in detail. A subjective selection was made from preliminary results based on the highest adjusted R^2 per type of regression or variable.

Several requirements for the algorithms were devised to guide the regression. They should have monotonous behaviour, i.e. always increase for an increase in relative altitude. Above the tropopause the gradient should be higher than below the tropopause due to stratospheric lifetimes. The CCF should never become negative as this is physically impossible for a water vapour emission.

5.3.1. Quartic dependence on relative tropopause pressure

Figure 5.4 shows the pattern of the water vapour CCF data versus relative tropopause pressure data. There is a clear trend of increasing CCF data for increasing altitude relative to the tropopause, but it is non-linear and non-trivial to model. None of the reviewed literature serves as a guide, though if reverse-engineered as in Section 5.1, results from Fichter (2009) do suggest a non-linear relationship with pressure relative to the tropopause. The following equation showed the best middle-ground between accuracy and complexity:

$$\text{F-ATR}\,20_{\text{H}_2\text{O}} = A + B \times (lev - tp_{WMO} - C)^4 + \epsilon \tag{5.1}$$

The parameter values from non-linear least squares optimisation are summarised in Table 5.2.

The regression is shown as the black curve in Figure 5.4. It appears to underfit the 400 hPa data and overfit the 300 hPa and 250 hPa data. The CCF values for 200 hPa have much higher variance and a correct fit is more difficult to identify. The curve significantly overestimates CCF values for the emissions highest above the tropopause because the regression by definition covers the domain of the independent variable. Though there are few data points, the CCF values appear to decrease with

Table 5.2: Parameter overview for water vapour quartic relative tropopause pressure algorithm.

	Units	Estimate	Std. Error	t value	$Pr\,(X > \lvert t \rvert)$
A	K kg(fuel)$^{-1}$	3.34×10^{-16}	2.81×10^{-17}	1.19×10^{1}	5.73×10^{-31}
B	K kg(fuel)$^{-1}$ Pa^{-4}	4.21×10^{4}	2.50×10^{3}	1.69×10^{1}	4.36×10^{-58}
C	Pa	2.29×10^{-34}	4.49×10^{-35}	5.11×10^{0}	3.79×10^{-7}

altitude at this end. At a release pressure more than 100 hPa above the tropopause pressure (emissions far below the tropopause), the water vapour climate impact appears to be practically constant.

The regression is clearly non-linear due to the constant within the power term. Linear versions were tested with different constants applied, but the result was arbitrary and sensitive to the constant chosen, which physically represents the relative altitude at which the CCF results is minimum. Adding other weather variables to the regression did not explain the remaining variation, and due to the non-linear regression this easily leads to convergence problems. Due to the nature of non-linear regression, addition of dimension variables to the regression generally leads to convergence problems and no meaningful results for interpretation.

Physically, water vapour climate impact is known to depend on its atmospheric lifetime, which is much larger in the stratosphere due to its stability. The relative altitude at the time of emission is thought to be a proxy for the eventual altitude of the tracer during its 90-day lifetime, with emissions close to the tropopause having a highly uncertain destination. Additional data at an extra altitude above 200 hPa would help to identify whether more distinctly stratospheric emissions show less variance in the CCF results due to the likelihood of ending up in the troposphere decreasing.

The year-mean tropopause in Fichter (2009) varies between 100–250 hPa depending on latitude, while it varies between 100–350 hPa in this dataset because of seasonal variation. At a latitude of 80°N in winter should have one of the lowest tropopause altitudes possible, so on the negative end no problems are expected. On the high end, at the equator, for low emissions there is a risk of the relative pressure exceeding the value of constant C in the regression, and the trend with altitude reversing. This can be mitigated by setting the algorithm to constant results below the value of C or even below the extent of the training data used here. If the algorithm is used for emissions above 200 hPa in winter at high latitudes there is a risk of overestimating the resulting CCF. This should be mitigated by careful application. Application above 200 hPa is generally unwise given that we have no data to model stratospheric effects.

5.3.2. Linear dependence on potential vorticity

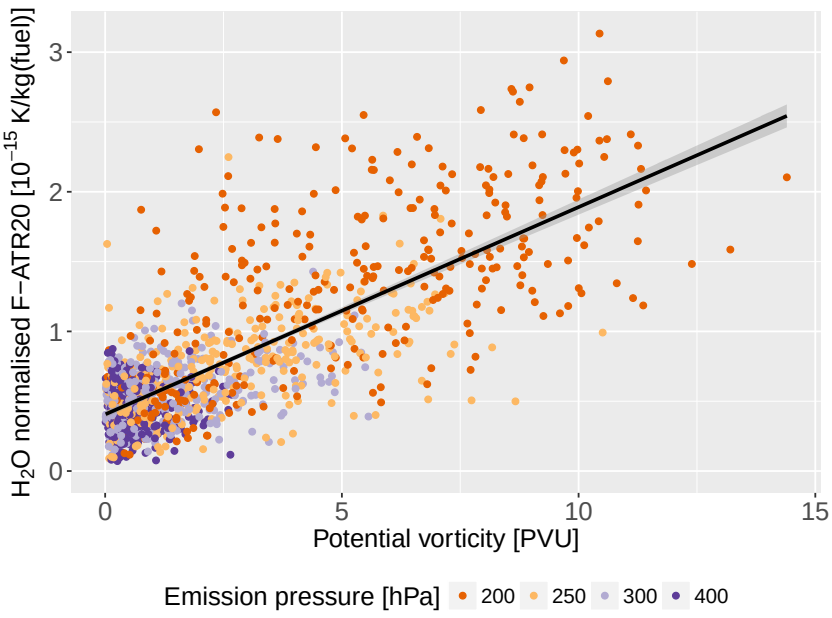


Figure 5.5: Scatterplot of water vapour CCF data versus PV coloured by emission pressure, with the linear regression found shown as a black line with 95% confidence interval for the regression parameters.

Figure 5.5 shows the pattern of the water vapour CCF data versus Potential Vorticity (PV). There is no discernible pattern beyond linear so simple linear regression was used:

$$\text{F-ATR}\,20_{\text{H}_2\text{O}} = \beta_0 + \beta_1 \times PV + \epsilon \tag{5.2}$$

The parameter values from least squares optimisation are summarised in Table 5.3.

Table 5.3: Parameter overview for water vapour linear potential vorticity algorithm.

	Units	Estimate	Std. Error	t value	$Pr\,(X > \lvert t \rvert)$
β_0	K kg(fuel)$^{-1}$	4.05×10^{-16}	1.20×10^{-17}	3.37×10^{1}	7.90×10^{-181}
β_1	K kg(fuel)$^{-1}$ PVU^{-1}	1.48×10^{-16}	3.40×10^{-18}	4.37×10^{1}	5.63×10^{-260}

The regression is shown as the black line in Figure 5.5 with 95% confidence intervals. Potential vorticity is a well-understood concept in meteorology but not a primitive variable. It combines vorticity and stratification in a way that only diabatic and frictional processes affect a parcel's PV. Stratospheric air has higher PV than tropospheric air, and when a parcel crosses the tropopause it takes significant time for the parcel to reach the same PV as its surroundings. Thus this concept is widely used to track cross-tropopause transport in meteorology (Grewe and Dameris, 1996), and very important for studying rapid cyclogenesis (Davis and Emanuel, 1991). The definition of PV is:

$$PV = -g\,(\zeta_\Theta + f)\,\frac{\partial \Theta}{\partial p} \tag{5.3}$$

Where the left term in brackets represents absolute vorticity on an isentropic surface (relative vorticity on an isentropic surface ζ_Θ and Coriolis parameter f), and the right term represents the static stability (vertical gradient of potential temperature Θ). Here Potential Vorticity Units (PVU) are used, defined as 10^{-6} m^2 s^{-1} K kg^{-1} .

As can be seen in Figure 5.5, heteroskedasticity is absent whereas the boxplots for altitudes (Figure 5.2) showed it. PV seems to represent vertical variation in the CCF results more smoothly than altitude itself. Grewe and Dameris (1996) present a figure a figure with zonally-averaged PV isolines and the WMO tropopause for 1 January. In the Southern Hemisphere PV is by definition negative, so the absolute value should be taken as a prior step in the algorithm to avoid negative CCFs. For the 200–400 hPa cruise altitude, the PV decreases strongly towards the equator and this algorithm would give results close to the intercept. Water vapour CCFs being lower in the tropics as suggested by this algorithm is sensible as the background humidity is higher (Tian et al., 2013), and the tropopause altitude is higher.

The PV isolines in Grewe and Dameris (1996) roughly follow the contours of the thermal tropopause and suggest that PV could be interpreted as a proxy for the relative altitude to the tropopause. Results from Fichter (2009) show a non-linear behaviour with altitude that follows the tropopause altitude meridionally, which visually lines up well with the PV contours in Grewe and Dameris (1996).

Addition of level to the regression strongly improves the accuracy through a negative direct term and/or a negative interaction term with PV (i.e. $\beta_2 \times lev \times PV$). Addition of latitude strongly improves the accuracy through a negative interaction term with PV. Addition of season slightly improves the accuracy through an interaction term with PV, i.e. an increased slope between the CCF and PV in winter.

Values of PV that represent the dynamic or PV tropopause vary in literature but are in the order of 2–4 PVU (Grewe and Dameris, 1996; Wilcox et al., 2012a; Irvine et al., 2013). The thermal or WMO tropopause and the PV tropopause by definition give different results, but optically the behaviour of the CCF in Figure 5.5 when considering 3 PVU to represent the tropopause, lines up with the behaviour in Figure 5.4.

5.3.3. Reciprocal dependence on specific humidity

Figure 5.6 shows the pattern of the water vapour CCF data versus specific humidity. Clearly humidity decreases with altitude, and the CCF increases. The shape of the data suggests a reciprocal relationship. For simplicity a linear regression was used such that the algorithmic CCF has a vertical asymptote at zero humidity, and a horizontal asymptote for high humidity:

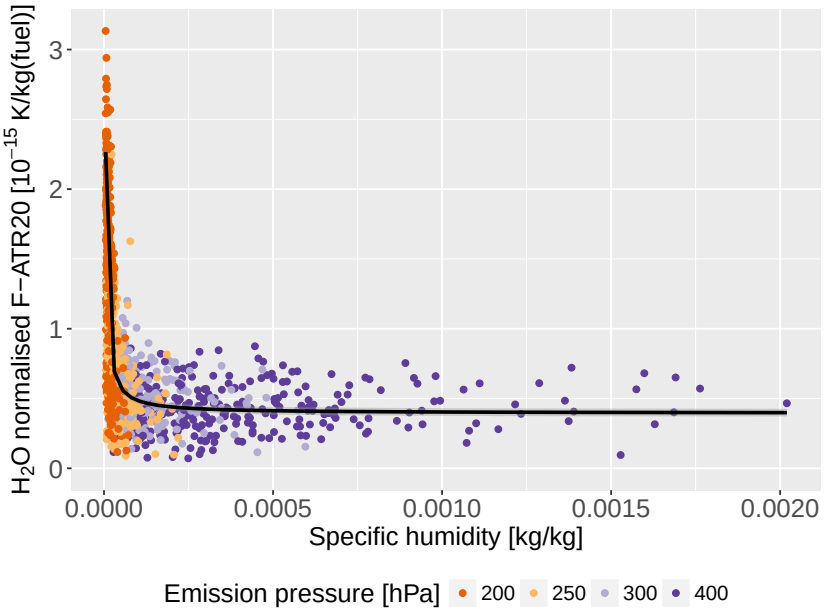


Figure 5.6: Scatterplot of water vapour CCF data versus specific humidity coloured by emission pressure, with the reciprocal regression found shown as a black curve with 95% confidence interval for the regression parameters.

$$\text{F-ATR}\,20_{\text{H}_2\text{O}} = \beta_0 + \frac{\beta_1}{q} + \epsilon \tag{5.4}$$

The parameter values from least squares optimisation are summarised in Table 5.4.

Table 5.4: Parameter overview for water vapour reciprocal specific humidity algorithm.

	Units	Estimate	Std. Error	t value	$Pr\,(X > \lvert t \rvert)$
β_0	K kg(fuel)$^{-1}$	3.94×10^{-16}	1.43×10^{-17}	2.76×10^{1}	4.32×10^{-133}
β_1	K kg(fuel)$^{-1}$ (kg kg^{-1})$^{-1}$	9.09×10^{-21}	2.56×10^{-22}	3.55×10^{1}	7.68×10^{-195}

The regression is shown as the black line in Figure 5.6. The result is visually satisfying but bears further inspection. As the algorithm approaches infinity for zero humidity, the data points with low humidity and low CCF have a high error. The result is negative heteroskedasticity. For very high humidity, the algorithm asymptotically approaches the intercept (β_0, for simple linear regression the y-intercept) and hence the algorithm is stable for extrapolation to higher values of q.

Specific humidity is a primitive variable that is present in the most basic atmospheric models. It is simply the mass of water vapour in an air parcel over the mass of the wet air parcel. In the REACT4C set-up (see Section 3.4.1) the emitted water vapour is not added to the background and does not influence precipitation rates. All else being equal, higher specific humidity means higher likelihood of achieving critical humidity and precipitation occurring, which is fed into the Lagrangian loss rate. An alternative explanation is that specific humidity is a proxy for altitude (relative to the tropopause) and thus the destination of the tracer, as Figure 5.6 clearly shows that humidity decreases with altitude. With only instantaneous data the distinction cannot be tested, but it may be assumed that both effects are present.

Addition of level to the regression strongly improves accuracy through a negative interaction term with reciprocal specific humidity. Addition of latitude barely improves the accuracy through a negative interaction term. Addition of season barely improves the accuracy through a direct term, that lowers the y-intercept of the regression in winter.

The lowest value of q, for which the maximum algorithm result of about 2.3 × 10^{-15} K kg(fuel)$^{-1}$ is achieved, is about 0.005 g kg^{-1}. The lowest value for 300 hPa in Tian et al. (2013) is about 0.03 g kg^{-1} and this is the highest altitude they provide results for. Humidity decreases with altitude, but as this dataset contains winter values at 200 hPa, 80°N it is unlikely that even lower humidity values would occur as long as application of this algorithm is limited to 200 hPa and 80°N. Application to the Southern Hemisphere should give similar results.

5.3.4. Linear dependence on pressure and squared tropopause

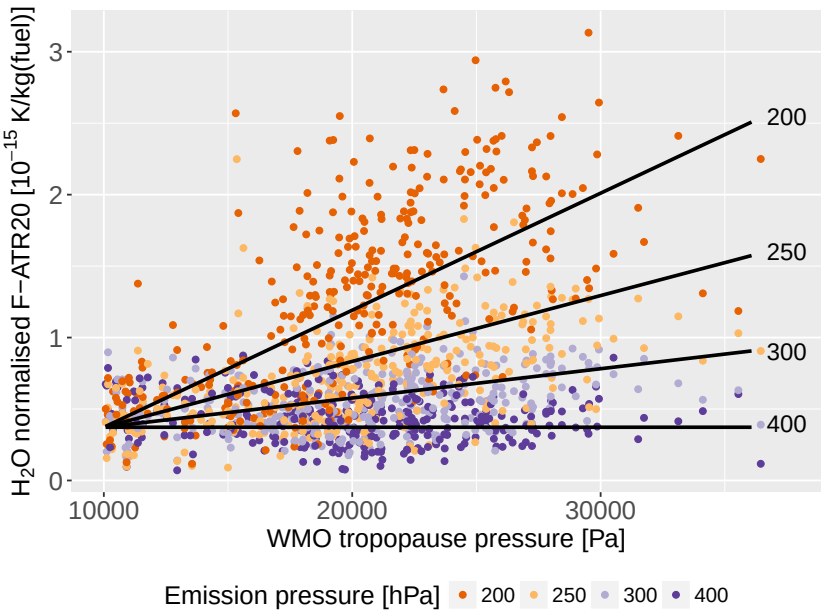


Figure 5.7: Scatterplot of water vapour CCF data versus the pressure altitude of the thermal tropopause coloured by emission pressure, with the best regression found shown as separate black lines per emission pressure (200–400 hPa).

Figure 6.9 shows the pattern of the water vapour CCF data versus tropopause pressure. As opposed to the non-linear relative tropopause pressure algorithm, the approach here is to vary the algorithm for both tropopause pressure and release pressure. For interactive analysis 3-D plots were used but these are not as readily interpreted in print. The data suggest that for a high tropopause, the CCF is independent of the release altitude. For the lowest release altitude, 400 hPa, the CCFs appear constant. It seems that for progressively higher altitude emissions, the slope of the CCF-tropopause pressure relationship changes non-linearly. For simplicity, the shape of this relationship was assumed linear. Performing linear regression for a linear relationship with tropopause altitude starting at 100 hPa and a quadratic relationship with altitude starting at 400 hPa gives:

$$\text{F-ATR}\,20_{\text{H}_2\text{O}} = \beta_0 + \beta_1 \times (tp_{WMO} - 10000) \times (40000 - lev)^2 + \epsilon \tag{5.5}$$

The parameter values from least squares optimisation are summarised in Table 5.5.

Table 5.5: Parameter overview for water vapour linear pressure and squared tropopause algorithm.

	Units	Estimate	Std. Error	t value	$Pr\,(X > \lvert t \rvert)$
β_0	K kg(fuel)$^{-1}$	3.72×10^{-16}	1.21×10^{-17}	3.09×10^{1}	1.47×10^{-158}
β_1	K kg(fuel)$^{-1}$ Pa^{-3}	2.05×10^{-28}	4.46×10^{-30}	4.59×10^{1}	1.59×10^{-277}

The regression is shown as the four black lines in Figure 5.7 - one line per altitude. The algorithm is continuous for altitude but only these four are shown for clarity and for comparisons to the original data. The constraints for minimum tropopause pressure and maximum release pressure appear to work well, but the quadratic and linear terms are arbitrary and fixed due to the extent of the dataset.

This algorithm demonstrates that the non-linear algorithm for relative tropopause pressure (Section 5.3.1) does not completely capture the nature of the trend of the CCF data. Literature (Grewe and Stenke, 2008; Fichter, 2009; Lee et al., 2010; Wilcox et al., 2012b) is clear on longer water vapour lifetimes in the stratosphere leading to higher climate impact for stratospheric emissions, which suggests the tropopause at the time of emission should be distinctive for water vapour climate impact. This algorithm shows that the CCF can modelled with similar accuracy with (Section 5.3.1) and without this distinction. Here a higher tropopause only leads to high CCF results for emissions above 400 hPa, and the effect is always linear. The slope does not depend on the distinction between tropospheric and stratospheric emissions. Emissions at 400 hPa having constant approximate CCF values in this algorithm can be explained by these emissions always being significantly below the tropopause. The

increasing slope with altitude may represent the uncertainty of a tracer ending up in the stratosphere depending on its emission above or below the tropopause. There is a clear trend of increasing CCF for increasing emission altitude relative to the tropopause but no distinction in terms of emissions below or above the tropopause.

Addition of level to the regression improves accuracy but only through a negative interaction term. Addition of latitude improves accuracy more strongly, either directly through a negative term or through a negative interaction term. Season improve the accuracy less through an interaction term, i.e. an increase of the slope between the CCF and the combination of level and tropopause in winter.

This algorithm should not be applied below 400 hPa because it would give a negative trend for tropopause altitude. If applied above 200 hPa the trend would be steeper than supported by data. Given the extent of the REACT4C domain, application for lower tropopause pressures than in this dataset is not likely. In the tropics a higher tropopause is expected, and values lower than 100 hPa could easily lead to negative CCFs for 200 hPa emissions. A minimum equal to the intercept should be applied as a further step in the algorithm.

5.4. Algorithm trade-off

Based on the algorithm requirements in Section 4.4, a set of trade-off criteria were devised in Section 4.6. Table 5.6 provides an overview of the results in this section.

Table 5.6: Overview of results of the subjective water vapour algorithm trade-off, per algorithm and category. If three or more results are distinct, they are ranked into +, ± and −. If only two distinct categories are identified, only + and − are used.

Algorithm	Pattern	Accuracy	External veracity	Background	Parsimony
$(lev - tp)^4$	−	−	±	±	−
PV	+	+	+	±	+
$\frac{1}{q}$	−	±	±	+	±
$tp \times lev^2$	+	±	−	−	−

5.4.1. Pattern of residuals

The mean values of the spread of residuals per level and WP are:

- $(lev - tp)^4$: 1.33×10^{-15}
- PV : 1.22×10^{-15}
- $\frac{1}{q}$: 1.33×10^{-15}
- $tp \times lev^2$: 1.20×10^{-15}

The PV and bivariate algorithms offer about 10% improvement over the other two.

Boxplots of the residuals (not shown) reveal that all algorithms significantly underpredict 200 hPa CCF data and slightly overpredict other altitudes, and the heteroskedasticity from the CCF data is still present in the residuals. All algorithms show little remaining trend per weather pattern and season. All algorithms overpredict 80°N data.

The 2-D autocorrelation test (Moran's I) is inconclusive. The amount of level/WP sets that are statistically significant varies per algorithm, but none give a negative I.

Overall there is a distinction between relatively accurate pattern (PV and bivariate algorithms) and relatively inaccurate pattern (quartic relative tropopause and reciprocal humidity algorithms).

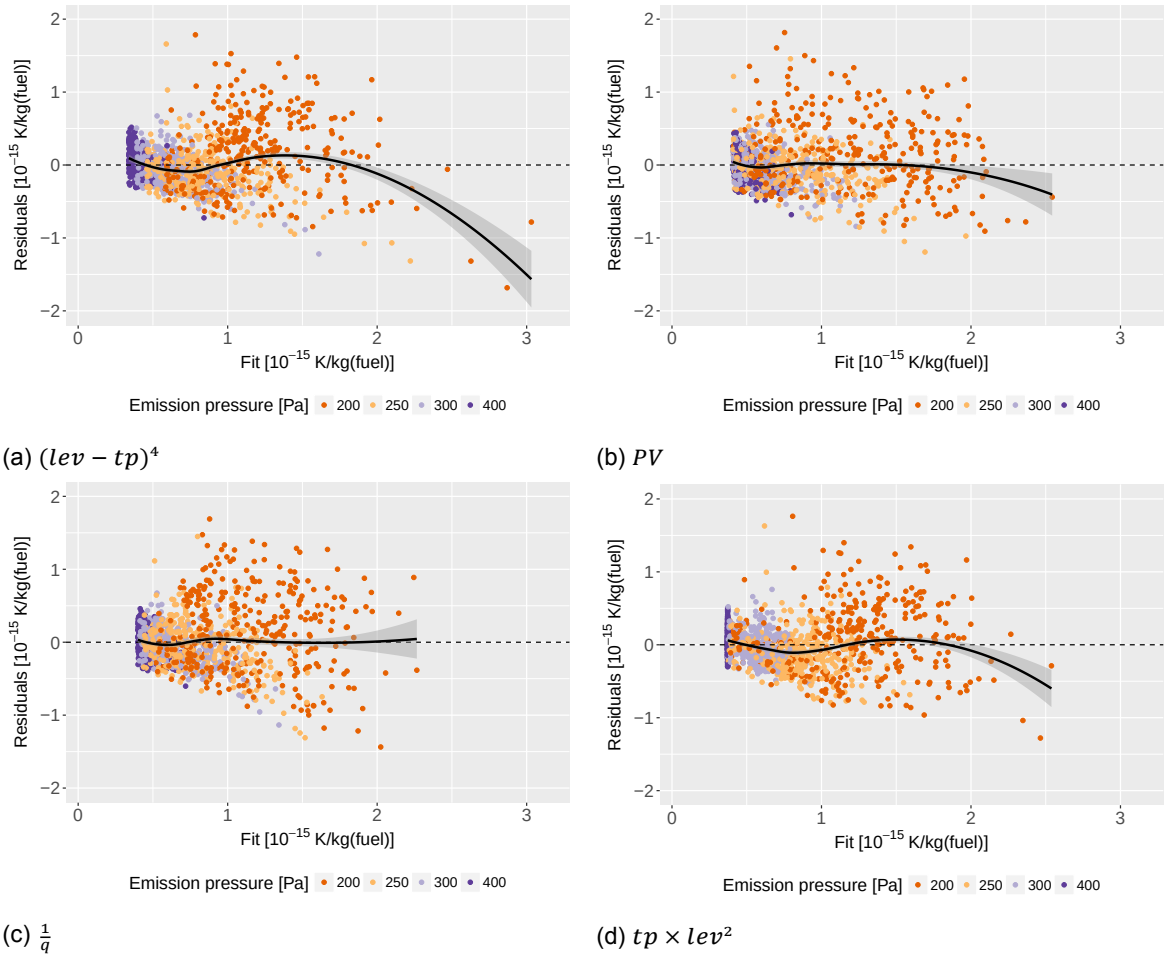
5.4.2. General accuracy

The Residual Standard Error (RSE) (and adjusted R^2) for each algorithm within this dataset are:

- $(lev - tp)^4$: 3.5×10^{-16}
- PV : 3.3×10^{-16} (0.59)
- $\frac{1}{q}$: 3.7×10^{-16} (0.48)

- $tp \times lev^2$: 3.2×10^{-16} (0.61)

Again the PV and bivariate algorithms offer about 10% improvement over the other two candidates.



(a) $(lev - tp)^4$ (b) PV

(c) $\frac{1}{q}$ (d) $tp \times lev^2$

Figure 5.8: Scatterplots of residuals versus fit for the four water vapour algorithms. Residuals are original values minus fitted values. The black curves are LOESS regressions to identify the trend of the residuals over the algorithm values, with 95% confidence bands.

Figure 5.8 shows residual plots for the four algorithms. The humidity algorithm shows the least remaining pattern. Both tropopause algorithms seem to have a remaining trend in residuals versus fit for 250–400 hPa data, and have large negative residuals for high values of fit. This means that the highest values of the algorithm are large overpredictions of the original CCF data. The PV algorithm appears to have less heteroskedasticity in its residuals, and the trend looks almost as good as for humidity.

In terms of accuracy, the PV and bivariate algorithms are best, the specific humidity algorithm is worst and the quartic algorithm is moderate. In terms of residual plots, the PV and humidity algorithms are best and the tropopause algorithms are worst. Overall the PV algorithm is best, the quartic algorithm is worst, and the other two are moderate.

5.4.3. External veracity

Results here are summarised from Section 5.3.

- $(lev - tp)^4$: a high tropopause or low altitude emission risks exceeding the minimum of the parabola and reversing the altitude trend. A low tropopause or high altitude emission is likely to overpredict the water vapour climate impact.
- PV: safe for low values (in tropics) due to intercept (β_0), stable for high values of PV due to

conservative trend assumed. If absolute value is applied, the Southern Hemisphere should not be a problem.

- $\frac{1}{q}$: returns infinite results for dry air and a hard limit would add uncertainty. At lower latitudes the humidity increases beyond this dataset and the CCF is expected to decrease so the asymptote will lead to overprediction of the water vapour climate impact for low latitudes.
- $tp \times lev^2$: returns a negative trend with tropopause altitude for emissions below 400 hPa, an uncertain high trend for emissions above 200 hPa, and potentially negative CCF results for a tropopause above 100 hPa in the tropics.

The bivariate algorithm is the most uncertain in terms of external application due to the two hard-coded constants. The PV algorithm is the most certain because of its simplicity.

5.4.4. Mechanistic background

Results here are summarised from Section 5.3.

- $(lev - tp)^4$: the underlying variable is the most obvious driver from literature, but the fourth order non-linear regression is suggested by data instead of theory.
- PV: this is a well-known descriptive variable in meteorology as a flow tracer but not a direct physical driver (i.e. a primitive variable), known to correlate with altitude relative to tropopause.
- $\frac{1}{q}$: the underlying variable represents both basic weather and likelihood of precipitation, the reciprocal function is simple enough to support with literature. No arbitrary constant is applied.
- $tp \times lev^2$: represents what is not explained by the simple pressure relative to the tropopause which is exactly why it is not well understood. The constants and the power term were suggested by data.

The humidity algorithm represents the best combination of an obvious driver with a simple functional form derived. The two tropopause algorithms lack theoretical support.

5.4.5. Parsimony

Non-linear regression in the quartic tropopause algorithm represents unnecessary complexity. Combining two variables with one quadratic term and two hard-coded constants is also complex. Simple linear regression for PV is the simplest option. Univariate linear regression for inverse specific humidity represents a middle way.

5.4.6. Choice

The PV algorithm is the only one that scores well across all criteria in Table 5.6 - low spatial and absolute error, stable for external application, and simple. Its drawback compared to the other candidates is that PV is not as clear of a driving variable from what literature has been published on the topic of aviation water vapour climate impact.

Taking into account the recommendation of making the regression applicable for the Southern Hemisphere, the final form of the algorithmic CCF is:

$$\mathrm{aCCF}_{\mathrm{H_2O}} = 4.05 \times 10^{-16} + 1.48 \times 10^{-16} \times |PV| \approx \text{F-ATR}\,20_{\mathrm{H_2O}} \tag{5.6}$$

Figure 5.9 shows a comparison between the original CCF data and the final algorithm (aCCF) results, where points on the black line indicate perfect agreement. The minimum for the aCCF, i.e. the y-intercept or β_0 in Section 5.3.2, is clearly visible and overestimates the lowest CCF values. The 400 hPa data appears to be correctly predicted in terms of the mean, but this y-intercept leads to less variability than is present in the original data. The algorithm appears optimised for 300 hPa and 250 hPa data. A remaining trend appears visible for 200 hPa data, as the algorithm predicts low CCF values for this altitude correctly but underpredicts the highest CCF values. Difficulties with properly representing the climate impact of emissions at 200 hPa were expected from Figure 5.1. Figure 5.9 suggests that a higher-order PV term would improve the algorithm's predictions, at least for 200 hPa, but the adjusted R^2 (so overall accuracy at all altitudes) improves negligibly for addition of any number of higher-order terms to the original regression as presented in Section 5.3.2.

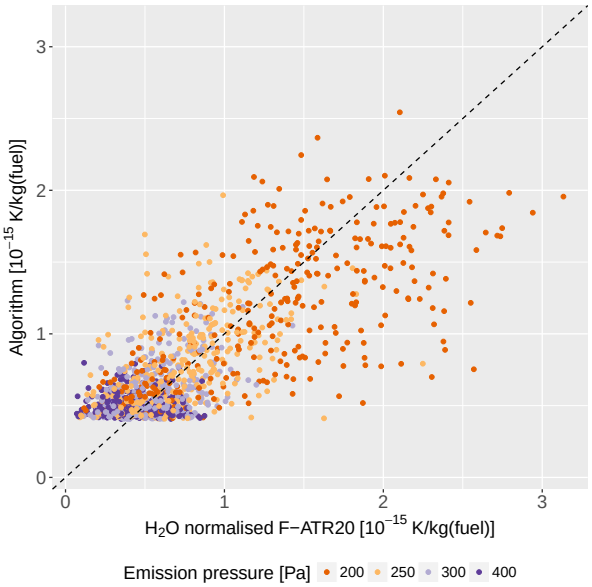


Figure 5.9: Scatterplot of final water vapour |PV| algorithmic CCF results versus REACT4C data coloured by level, with the dotted line indicating perfect agreement.

5.5. Discussion

From the results in the previous sections and others not shown in this thesis, a number of discussions and conclusions arise. Here discussions on various topics are given as far as they relate directly to the results of the water vapour CCF data, the weather data, and the algorithm formulation. General discussions regarding the methodology are given in Chapter 7.

5.5.1. Relationship with emission altitude

When reverse-engineered and plotted alongside the REACT4C water vapour CCFs versus emission pressure (not shown), the results from Fichter (2009) show much better agreement than those from Wilcox et al. (2012b). Note that Fichter (2009) results are interpolated contours, but there are 15 emission points within the REACT4C latitude-altitude domain in their methodology. The data presented here are partially calculated from results of Grewe and Stenke (2008). Fichter (2009) expanded upon the methodology of Grewe and Stenke (2008) so some commonality is to be expected, but used different emission domains. There is no reason to doubt the E39C perturbation studies performed by Fichter (2009), beyond that the set-up omits zonal variation and no seasonal results are presented. Note that monthly results using the update Dahlmann (2012) methodology are currently calculated (Grewe et al., 2017a) but not available in literature.

The algorithms presented here, and other attempts not presented here, show complex behaviour of the water vapour CCF with altitude and tropopause altitude. Notably, the PV isolines presented in Grewe and Dameris (1996) appear similar to the water vapour RF contours in Fichter (2009). The agreement of both the original data and the analysis presented here with the more extensive methodology of Fichter (2009), suggests that the linear relationship between normalised water vapour RF and altitude in Wilcox et al. (2012b) does not adequately represent the nature of the variability of water vapour emission climate impact.

Results from Grewe and Stenke (2008) demonstrate that within the REACT4C altitude range, the mid-northern latitudes of the REACT4C domain are most sensitive to water vapour emissions and lower normalised RF and thus climate impact should be expected towards the equator. Their results also show that the variation at this altitude is negligible when compared to stratospheric water vapour emissions. None of the algorithms presented here are expected to be accurate for application above 200 hPa.

5.5.2. Maximum attainable R^2

When all the selected variables from Table 5.1 are combined in a linear regression model with longitude, latitude, level and season, the resulting adjusted R^2 is 0.68 for 34 variables. Introducing interaction terms of course increases the accuracy, but first-order interactions already require 595 parameters so

the result is meaningless. Though imperfect, this gives an estimate of the amount of variation in the CCF data that can be explained by the weather data.

5.5.3. Ensemble-mean algorithm

Application of the mean of an ensemble of simulations is commonly used in meteorology to mitigate the effect of chaos on the results of a prediction. When the ensemble-mean is calculated from the four algorithms presented in Section 5.3, the RSE is 3.2×10^{-16} K kg(fuel)$^{-1}$. As the non-linear regression is included in the mean, (adjusted) R^2 is undefined. This is approximately the same result for absolute accuracy as the bilinear tropopause and altitude algorithm, and a slight improvement upon the |PV| algorithm (Section 5.4.2).

Figure 5.10: Scatterplot of water vapour ensemble (mean of four) algorithmic CCF results versus REACT4C data coloured by level, with the dotted line indicating perfect agreement.

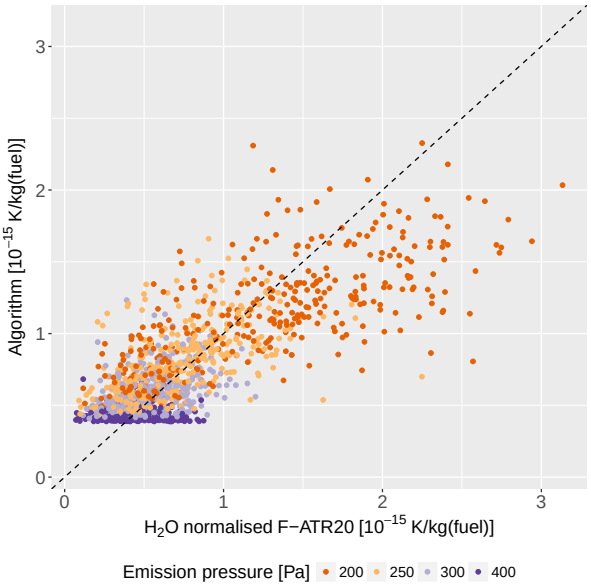


Figure 5.10 shows the resulting ensemble-mean algorithm results versus the original REACT4C data, for comparison to Figure 5.9. The results are difficult to distinguish from those of the PV algorithm. Comparisons to the other three algorithms (not shown) yield similar results, as the four algorithms all produce similar results (Section 5.4). Application of the ensemble-mean to these four algorithms does not appear to meaningfully improve the accuracy for the complexity it introduces.

5.5.4. Final algorithm

Figure 5.11 shows the same CCF-PV scatterplot as Figure 5.5 but coloured and regressed per season. Clearly the intercept is very stable, but the slope is prone to change. The main differences are in 200 hPa data. The latitude of emission is not indicated in the plot, but almost all data points with low CCF values for PV values above around 2 PVU are emitted at 80°N.

Several interesting options remain to improve upon this algorithm. The remaining latitude and season variation may be accounted for by adding these variables to the regression, though this would limit external veracity. The components of PV could be analysed separately in the REACT4C data to assess exactly what causes PV to be the most accurate (sole) predictor of water vapour CCF results. Diabatic heating rates and cross-tropopause fluxes can be investigated to expand understanding of the relationship between PV, tropopause and water vapour emission climate impact in this dataset.

Addition of a second parameter from Table 5.1 to the regression was within the scope of this research, and offered no significant improvements. Addition of dimensions as in Section 5.3.2 is problematic for external veracity, and because the dimensions are not continuous.

The adjusted R^2 of 0.59 means about 60% of variance in the CCF is explained by the algorithm. The maximum obtained with 35 variables is 0.68. Assuming the latter is a good estimate of the maximum attainable R^2, the PV algorithm is a good result and room for improvement within zero-dimensional instantaneous regression analysis appears marginal.

Figure 5.12 (left) shows bilinearly interpolated contour plots of the water vapour PV algorithm results

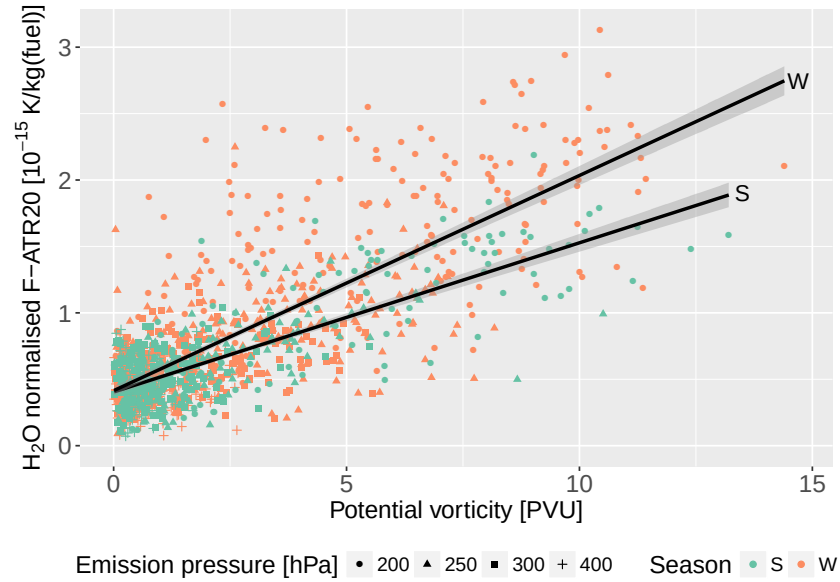
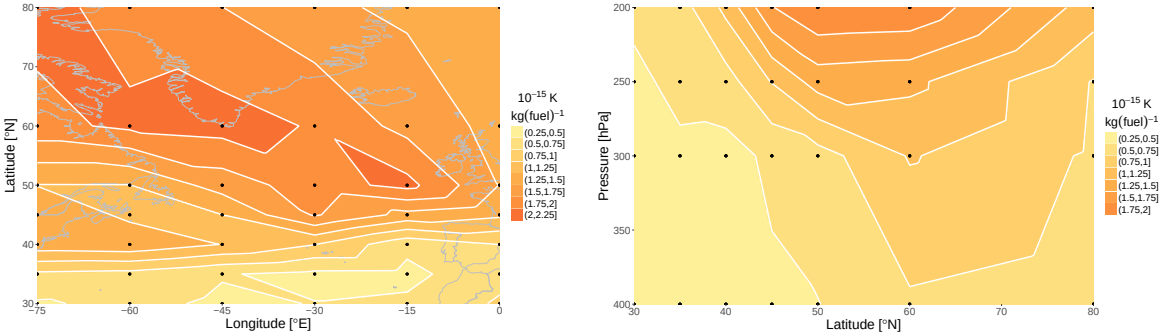


Figure 5.11: Scatterplot of water vapour CCF data versus PV coloured by season and shaped per emission pressure, with seasonal regression and 95% confidence intervals for regression parameters.



(a) 12 UTC, 200 hPa, WP1

(b) 12 UTC, 315°E, WP3

Figure 5.12: Water vapour final $|PV|$ algorithmic CCF bilinear contour plots in longitude-latitude field (left) and latitude-altitude field (right). The black points are the emissions locations i.e. the actual data that are interpolated.

for 200 hPa, WP1 to compare to Figure 5.3 (left). The algorithm produces a similar level of small-scale details as are present in the original data, but the pattern is not entirely reproduced. The lowest values still occur at low latitudes, but the high values are shifted north. Especially 285°E, 80°N appears overpredicted by the algorithm. The low value at 300°E, 45°N is captured by the algorithm, but it appears to smooth out the adjacent high values. Overall, only the overprediction at 285°E, 80°N is expected to potentially lead to incorrect climate-optimised routing and this is a location with relatively little air traffic. Appendix A (right side) contains these visualisations for all 32 combinations of altitude and weather pattern, for comparison to the original CCF data on the left side.

Figure 5.12 (right) shows bilinearly interpolated contour plots of the water vapour PV algorithm results for 315°E, WP3 to compare to Figure 5.3 (right). The dependence on the algorithm on PV is very clear, and the pattern of the CCF data relative to the tropopause is reproduced. The high values at 40–60°N, 200 hPa are underestimated. This suggests that, for this latitude-altitude field, a higher-order dependence on PV may improve the prediction. Appendix B (right side) contains these visualisations for all 48 combinations of longitude and weather pattern, for comparison to the original CCF data on the left side.

6

NO_x algorithmic CCF results

Following the methodology presented in Chapter 4, here REACT4C ozone and methane CCF results are described, analysed and compared (Section 6.1), the literature review and selection of regression variables is presented (Section 6.2), four ozone algorithms are discussed (Section 6.4), the trade-off is explained and the final ozone algorithm is chosen (Section 6.5), and finally a discussion on all results for ozone is presented (Section 6.6). The same structure is repeated for methane in Sections 6.7, 6.8 and 6.9.

6.1. Variability of REACT4C NO_x CCF data

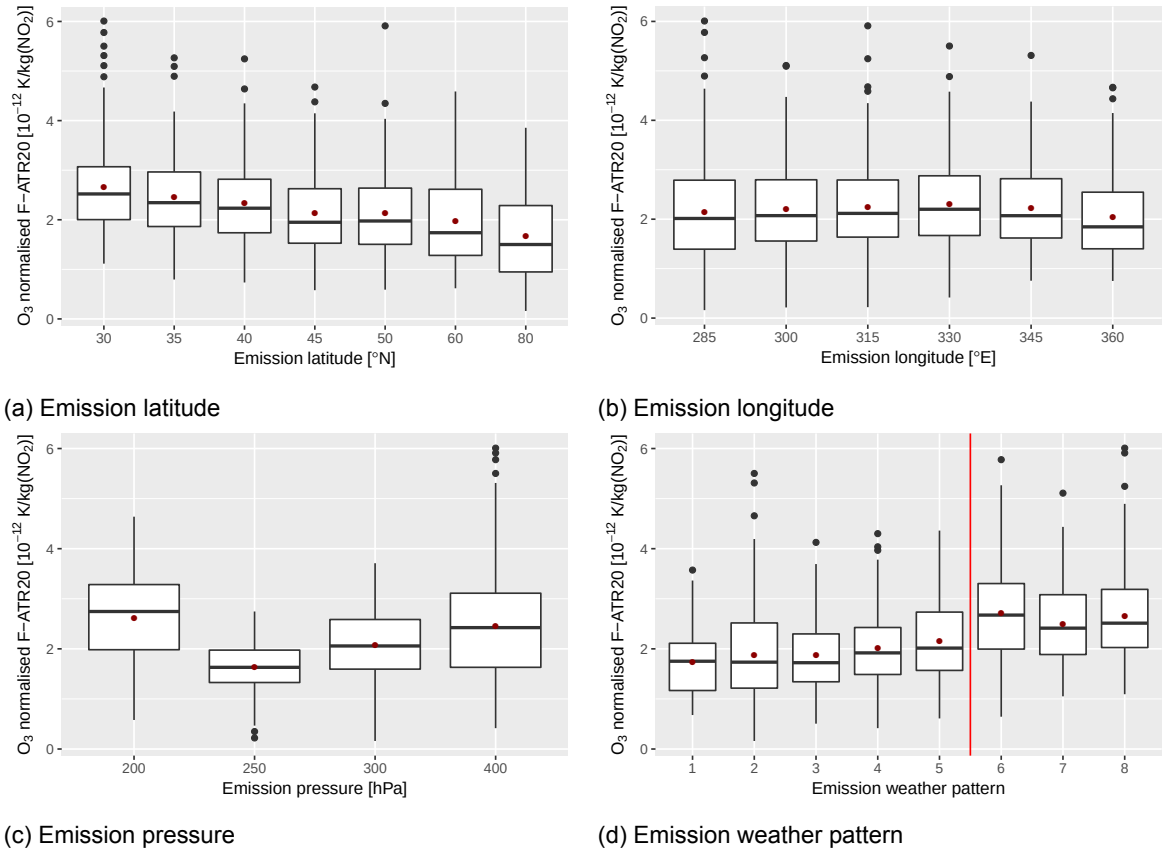
Before attempting to find an algorithmic approximation, the REACT4C NO_x CCF results for ozone and methane are described and analysed, as is the relationship between ozone and methane results.

6.1.1. Variability of ozone CCF data

Figure 6.1 shows box plots of the CCF data for each of the four dimensions. There is some meridional variability, with a clear decreasing linear trend for an increase in emission latitude. This is likely due to higher latitudes corresponding to residence at higher latitudes for the rest of the simulation. Köhler et al. (2013) also found a monotonic decreasing trend for ozone RF resulting from a constant amount of NO_x emitted globally in different latitude bands (2–3 bands in the REACT4C domain). Results from Grewe and Stenke (2008) and Fichter (2009), more readily available in Dahlmann et al. (2016), show a monotonic year-mean zonal-mean decrease in sustained ozone RF with latitude within the REACT4C domain. Results from Gilmore et al. (2013) agree on a year-mean decreasing trend of ozone production per mass of NO_x emitted. There is low zonal variability in the box plot and no fundamental variability along longitudes is expected due to the prevailing westerly winds (see e.g. Figure 1.1).

There is significant altitudinal variability, with a non-monotonic trend with a minimum at 250 hPa. Detailed analysis is necessary to explain this behaviour, though such a marked difference between 200 and 250 hPa may correspond to the tropopause altitude. Köhler et al. (2008) also found a non-monotonic relationship between ĩ0–14 km altitude for ozone RF resulting from a constant amount of NO_x emitted globally at 16 different altitudes. This departure from the linearity seen for lower altitudes, however, coincides with a change in NO_x pulse sizes and may have been caused by chemical non-linearity. Results from Grewe and Stenke (2008) and Fichter (2009) show a very slight increase in ozone RF with altitude within the REACT4C domain that does not appear to match these results.

There is visible seasonal variability, with summer emissions leading to higher climate impact. This is assumed to be due to temperature and solar radiation. Stevenson et al. (2004) found a minor seasonal relationship for ozone RF resulting from a constant amount of NO_x emitted globally in four different months. In their results, summer emissions also lead to higher climate impact but the difference is smaller than here. Berntsen et al. (2003) previously found disagreement between five models on the month of the maximum ozone perturbation from global aviation NO_x emissions, with one model showing a peak in July and the others in late winter or spring. Gilmore et al. (2013) used an adjoint sensitivity approach to model regional and temporal differences in NO_x direct impact on ozone. Those results show peak ozone production efficiency in summer. There is also intra-seasonal variability, with WP5



(a) Emission latitude

(b) Emission longitude

(c) Emission pressure

(d) Emission weather pattern

Figure 6.1: Box plots of REACT4C ozone CCF data for each dimension. The red line represents the distinction between winter (1–5) and summer (6–8) weather patterns.

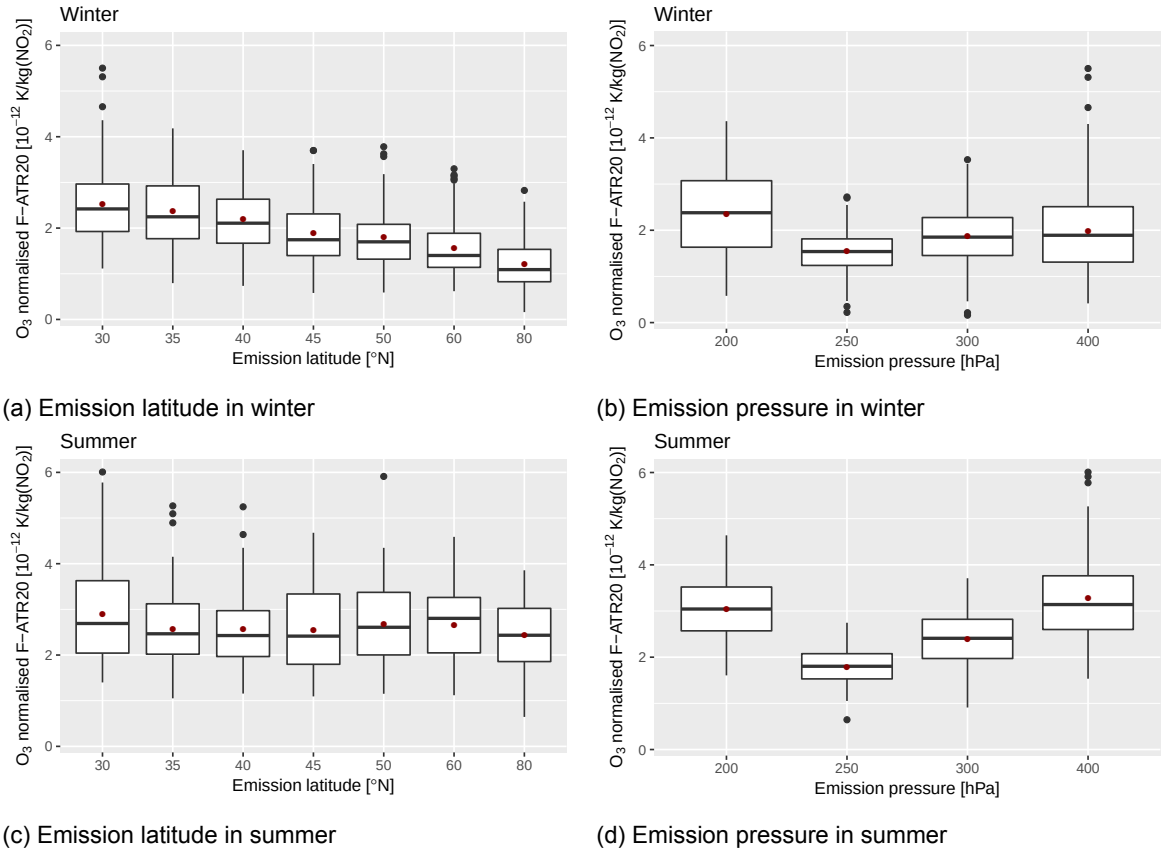
and WP6 producing relatively high CCF values. This suggests that the weather pattern distinctions devised by Irvine et al. (2013) (Section 3.1) are of importance for the variability of ozone climate impact.

Figure 6.2 shows box plots of the ozone CCF data for latitude and level after splitting into summer and winter data. There is a strong linear trend with latitude for all winter emissions, and practically no trend for summer emissions. This is likely caused by the availability of sunlight for NO_x chemistry, as in winter the effect of latitude on sunlight is stronger. The minimum ozone CCF values occur at 250 hPa for both summer and winter. If this is driven by tropopause altitude we expect more marked differences. The parabolic nature appears stronger for summer data. The highest values appear to occur for 30°N, summer, 400 hPa.

Figure 6.3 shows bilinearly interpolated contour plots of the ozone CCFs data for 200 hPa, WP1 to compare to Figure 1.1. The results in the two figures are in agreement, but this colour scale and the use of wind vectors over wind contours paints a less strong visual correlation with the weather situation. High values of the ozone CCF data do coincide with the jet stream, but are also present due south where there is little wind. There is a small peak value at 330°E, 80°N that is not visible in Figure 1.1 as it cuts off the original data on the borders of the REACT4C domain. Appendix C (left side) contains these visualisations for all 32 combinations of altitude and weather pattern.

6.1.2. Variability of methane CCF data

Figure 6.4 shows box plots of the methane CCF data for each of the four dimensions. There is almost no meridional variability. Köhler et al. (2013) found a monotonic decreasing trend with latitude for methane RF. Many factors could cause a discrepancy between results for eight weather patterns in the NAFC and year-mean zonal-mean results for latitude bands. Results from Grewe and Stenke (2008) and Fichter (2009) appear to agree with Köhler et al. (2013) on meridional variability. There is almost no zonal variability and no fundamental variability along longitudes is expected due to the prevailing westerly winds (see e.g. Figure 1.1).



(a) Emission latitude in winter

(b) Emission pressure in winter

(c) Emission latitude in summer

(d) Emission pressure in summer

Figure 6.2: Box plots of REACT4C ozone CCF data per level and latitude split into seasons.

There is some altitudinal variability, with apparently linearly increasing cooling effect for increasing emission altitude. Köhler et al. (2008) found almost constant altitude behaviour for methane RF from a constant amount of NO_x emitted globally at different altitudes. As mentioned in Section 6.1.1 these results may be affected by chemical non-linearity. Results from Grewe and Stenke (2008) and Fichter (2009) show little altitudinal variability.

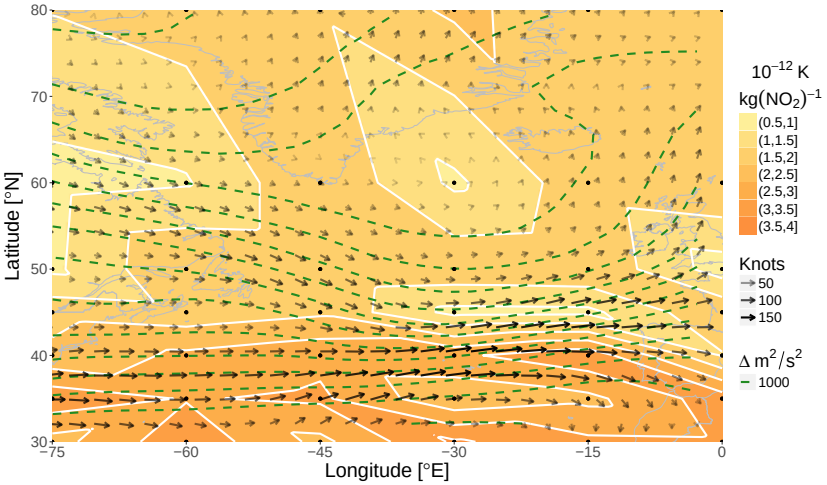
There is some seasonal variability, but the intra-seasonal variation is larger than the inter-seasonal variation. This suggests that the weather pattern distinctions devised by Irvine et al. (2013) (Section 3.1) are of importance for the variability of methane climate impact. Stevenson et al. (2004) found a minor seasonal relationship for methane RF, where summer emissions lead to stronger cooling.

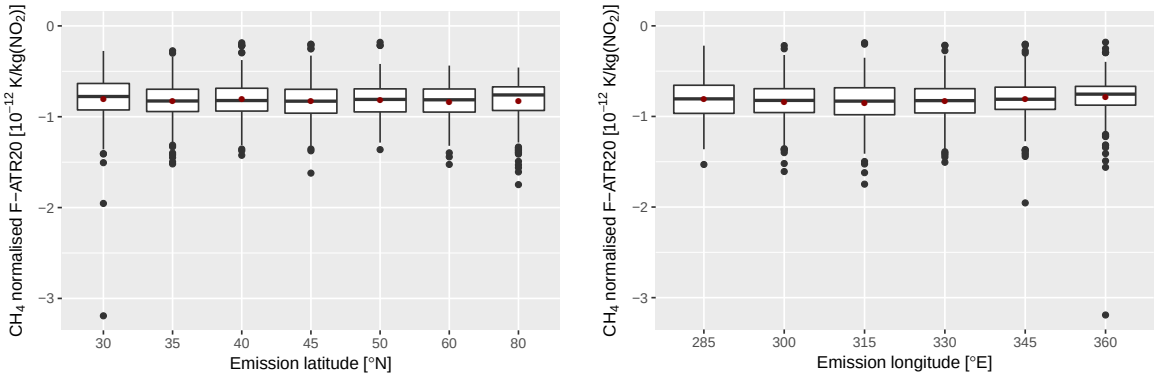
Figure 6.5 shows box plots of the methane CCF data for latitude and level after splitting into summer and winter data. The lack of meridional trend overall appears to come from combining a decreasing trend for summer data and an increasing trend for winter data. The summer trend thus agrees with literature, while the winter trend does not. The slight linear altitude trend overall appears to come from increasing but differently shaped trends for summer and winter - in summer the largest gradient is between 300 hPa and 400 hPa, while in winter it is between 200 hPa and 250 hPa.

Figure 6.6 shows bilinearly interpolated contour plots of the ozone CCFs data for 200 hPa, WP1 to compare to Figure 1.1, and for 400 hPa, WP2 to identify an outlier. On first sight the left figure does no agree with Figure 1.1, but the interpolation applied leads to different contour shapes, and the bins used for colouring are half the size. Investigation of the original and reverse-engineered methane CCF shows that both plots are correct in terms of the pattern, but all values in both sets of data are between -10 and -5 $\times 10^{13}$ K kg(NO_2)$^{-1}$. The contours in Figure 1.1 may be incorrect or the data shown there may include the climate impact of PMO. Both figures show little visual correlation with the weather situation but lower values at 80°N.

There is one clear outlier at 360°E, 30°N, 400 hPa, WP2 which is also present before interpolation. The CCF data at 330°E and 345°E are also clearly higher than the surrounding data in Figure 6.6 (right). This is likely to be a feature of model dynamics at the time of emission and in the proceeding 90 days.

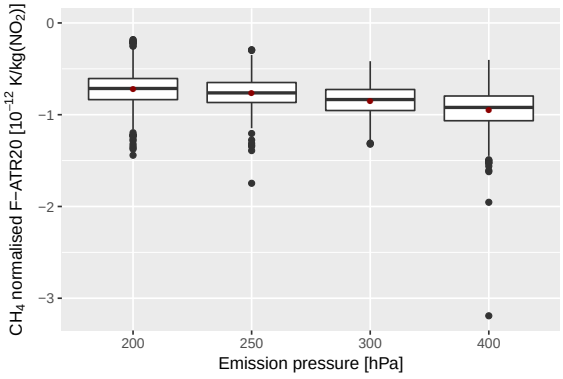
Figure 6.3: Ozone CCF bilinear contour plot, 12 UTC, 200 hPa, WP1. Dashed lines are geopotential isolines, arrows are wind vectors, black points are the emissions locations i.e. the actual data that are interpolated.



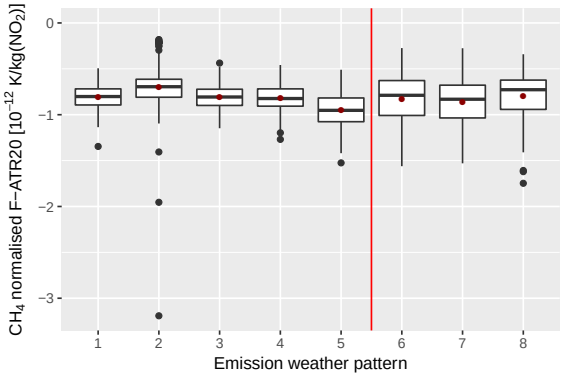


(a) Emission latitude

(b) Emission longitude

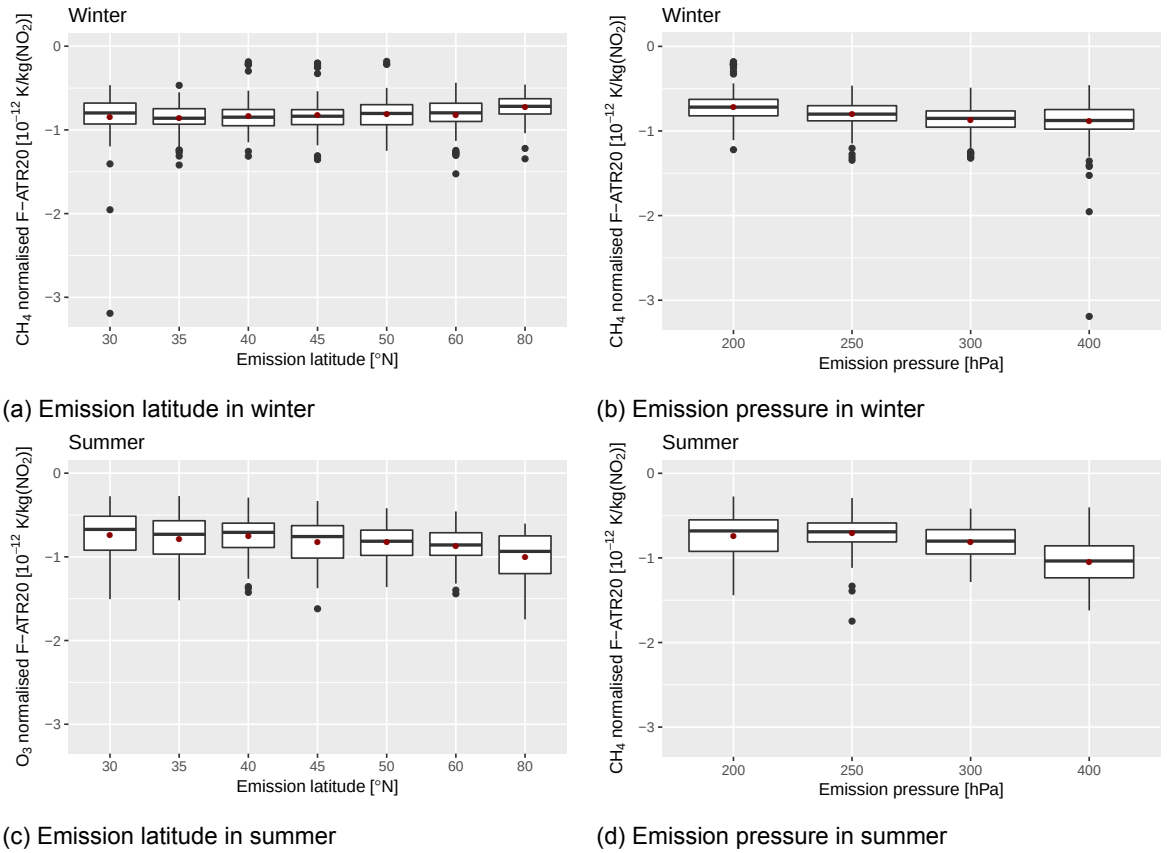


(c) Emission pressure



(d) Emission weather pattern

Figure 6.4: Box plots of REACT4C methane CCF data for each dimension. The red line represents the distinction between winter (1–5) and summer (6–8) weather patterns.



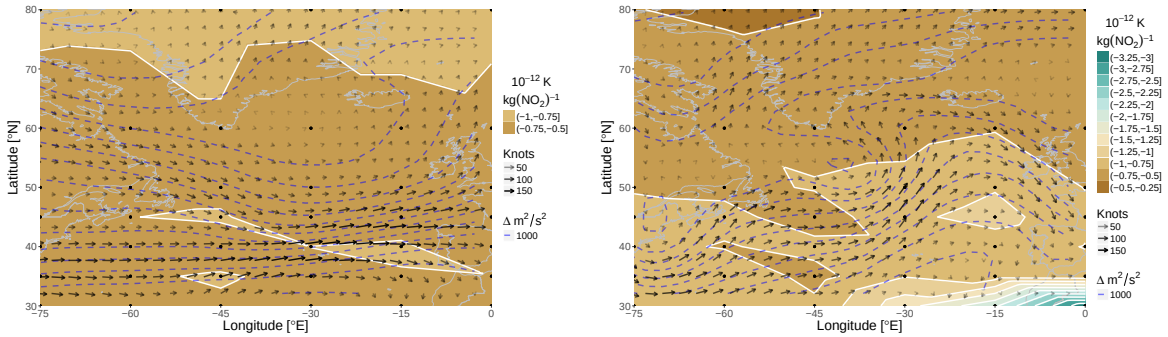
(a) Emission latitude in winter

(b) Emission pressure in winter

(c) Emission latitude in summer

(d) Emission pressure in summer

Figure 6.5: Box plots of REACT4C methane CCF data per level and latitude split into seasons.



(a) 12 UTC, 200 hPa, WP1

(b) 12 UTC, 400 hPa, WP2

Figure 6.6: Methane CCF bilinear contour plots, left to compare with Figure 1.1, right to identify outlier. The outlier is visible bottom right, dashed lines are geopotential isolines, arrows are wind vectors, black points are the emissions locations i.e. the actual data that are interpolated.

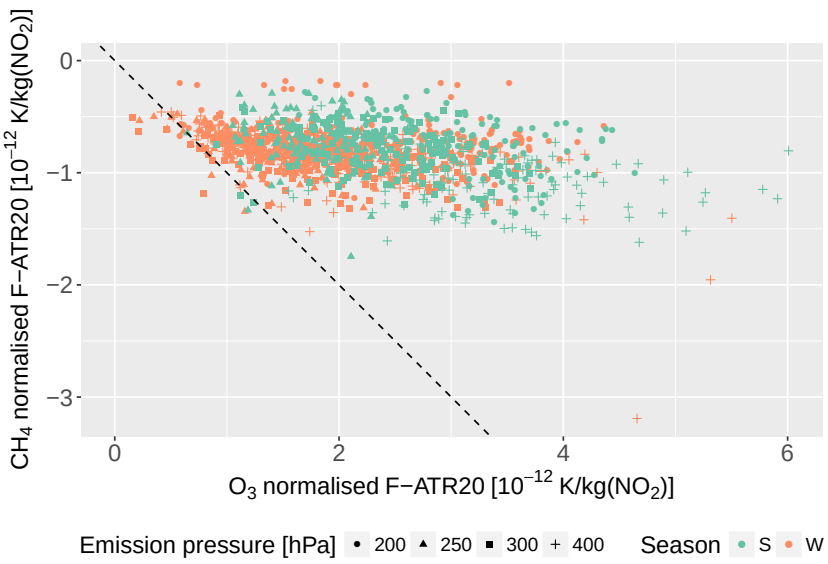


Figure 6.7: Scatterplot of methane CCF data versus ozone CCF coloured by emission season, with the dashed line representing net zero cumulative effect. Note that PMO has been omitted and that a similar figure is presented in Frömming et al. (2017).

A possibility remains that the outlier is caused by an error in the simulation, but the error affecting three neighbouring locations and not the rest of the concurrently simulated locations is unlikely. As it is not an unphysically low value the outlier is left in the dataset. Appendix D (left side) contains these visualisations for all 32 combinations of altitude and weather pattern.

6.1.3. Relationship between ozone and methane

Figure 6.7 shows the methane and ozone CCF results together. Emissions in both summer and winter show a decreasing trend of methane CCF for increasing ozone CCF, but the trend is not very strong and the data are noisy. No regression is shown in the plot for clarity and because multiple (linear) options are available for predicting the methane CCF data:

- constant
  - $RSE = 2.33 \times 10^{-13}$
- linear, forced through origin, no seasonal slope
  - $RSE = 3.14 \times 10^{-13}$,
  - $\partial\,\text{F-ATR}\,20_{\text{CH}_4}\,/\,\partial\,\text{F-ATR}\,20_{\text{O}_3} = -0.33$
- linear, forced through origin, seasonal slope
  - $RSE = 3.02 \times 10^{-13}$
  - $\partial\,\text{F-ATR}\,20_{\text{CH}_4}\,/\,\partial\,\text{F-ATR}\,20_{\text{O}_3} = -0.37/-0.30$ (winter/summer)
- linear, y-intercept, no seasonality
  - $RSE = 2.16 \times 10^{-13}$, adjusted $R^2 = 0.14$
  - $\partial\,\text{F-ATR}\,20_{\text{CH}_4}\,/\,\partial\,\text{F-ATR}\,20_{\text{O}_3} = -0.098$
- linear, y-intercept, full seasonality
  - $RSE = 2.13 \times 10^{-13}$, adjusted $R^2 = 0.16$
  - $\partial\,\text{F-ATR}\,20_{\text{CH}_4}\,/\,\partial\,\text{F-ATR}\,20_{\text{O}_3} = -0.091/-0.14$ (winter/summer)

Forcing a linear regression through the origin means the variance is at best 30% worse than simply taking the mean, i.e. assuming the methane CCF does not vary. The most elaborate option presented only achieves an adjusted $R^2 = 0.16$ so predicting the methane CCF from the instantaneous, zero-dimensional ozone CCF data is not very accurate. Note that R^2 is not defined when a regression does not include an intercept, and that options that do not include an intercept have higher residuals than just assuming the mean of the methane CCF data. The effect of the ozone CCF on the methane CCF varies between -0.37 and -0.091 depending on the season and intercept. Holmes et al. (2011) presented results of ozone versus methane year-mean RF from an ensemble of models and found an $R^2 = 0.79$ correlation with what appears to be a methane/ozone trend of -0.6. At least ten of the models in Holmes et al. (2011), namely those from Hoor et al. (2009) and Myhre et al. (2011), calculate methane RF from a further simplification of the IPCC formula (Section 3.6.3). It is thus unlikely that the difference in methane/ozone trend is caused by application of the IPCC formula in REACT4C, and likely that the difference is caused by local versus global-scale emissions and application of the F-ATR20 metric. Note that the studies used by Holmes et al. (2011) examine accumulated RF from historical aviation emissions in a certain year (see Figure 2.1), whereas the data here was first calculated as the integrated RF over a timeframe (i.e. AGWP) of a pulse emission and transformed to the Average Temperature Response (ATR) from a strategic re-routing decision given a future emission scenario (see Section 3.7).

Whether or not the linear regression should cross the origin depends on what is emitted and how the reactions are partitioned. Assessing limiting cases of the reaction pathways for ozone and methane to theoretically establish whether either ozone or methane climate impact from a NO_x emission can be zero when the other is non-zero is out of scope for this thesis, and all data in Figure 6.7 are for the same pulse of 5×10^5 kg of NO so limiting cases would be largely irrelevant.

Brasseur et al. (1998) find that increasing ambient NO_x mixing ratios (e.g. from aviation emissions), lead to increased ozone production rates and increased OH concentrations (thus increased methane oxidation) up to 0.1–0.2 ppbv, above which increasing NO_x leads to decreasing rates of ozone production and decreased OH (their Figure 3). Beyond this value for the background NO_x mixing ratio, a NO_x emission will lead to a decreased rate of ozone production and a decreased OH concentration. Grewe et al. (2012) present a similar figure for ozone production rates (no OH) at 50°N in Europe for summer conditions, which gives a NO_x mixing ratio of around 2–3 ppbv for the maximum rate of ozone production and around 20 ppbv for net-zero ozone production rate. The results from Brasseur et al. (1998) suggest similar behaviour for OH and thus methane. It is thus possible for an aviation NO_x emission to lead to net-zero ozone production given high enough ambient NO_x mixing ratios, but undetermined whether this would also lead to net-zero methane oxidation.

At high stratospheric altitudes, NO_x emissions are known to lead to a reduction in ozone (Lee et al., 2010) and thus negative ozone CCF values (Grewe and Stenke, 2008). Though this is outside of the REACT4C simulation domain, and assessing whether the assumptions in Section 3.4 are valid in the upper stratosphere is outside of our scope here, this casts extra doubt on whether a regression must necessarily cross the origin. If a y-intercept is unphysical, a non-linear fit may be warranted but there are not enough low-end ozone CCF data to reliably suggest one.

6.2. Regression variable selection

Berntsen and Isaksen (1999) found strong effects of reduced lightning and convection on aviation ozone changes from the same NO_x emission. Grewe et al. (2002) similarly conclude that lightning representation is of influence on aviation NO_x impact. Stevenson et al. (2004) found that NO_x ozone climate impact mostly depends on local NO_x chemistry, i.e. sunlight, temperature and background NO_x concentrations. NO_x methane climate impact is more dependent on transport. Gauss et al. (2006) identify sunlight, tropopause, HNO_3 washout, latitude and altitude as drivers of NO_x climate impact. Köhler et al. (2008) analysed altitude variability, as discussed in Section 6.1.1, which warrants including other metrics of altitude than just pressure. Results from Stevenson and Derwent (2009) suggest asymptotic behaviour of NO_x ozone and methane climate impact for background NO_x concentrations during emission, and identify sunlight as a driver.

Background concentrations (at the time and location of emission) of all chemical species described in Sections 2.2.2 and 3.4.1 are deemed relevant for analysis due to the complexity of the chemical dynamics and the possibility that a background concentration of a chemical that is not tagged in the

Grewe et al. (2014b) approach still correlates to the climate impact of an emission.

The four types of variables pursued for developing ozone and methane algorithms are thus:

- general meteorology (as water vapour analysis)
- solar radiation
- background chemistry
- lightning

Table 6.1 summarises all variables that were read in from the supplied datasets and down-selected by viewing in histograms. Production and loss terms of the tagged chemical species from Grewe et al. (2014b) are available in the data. These were ignored to avoid introducing unnecessary complexity, because of the preference for weather variables, and because of potential issues with representing the same terms in a different model set-up.

Table 6.1: Variable overview for NO_x ozone and methane algorithm analysis. The 3rd and 4th dimensions in 3-D/4-D are always time followed by altitude, and bold-printed rows are those selected for analysis via histograms. Note that the time dimension is not used in this analysis and that weather pattern is not counted as a dimension for this table. IC is intra-cloud, CG is cloud-to-ground.

Dataset	Name	Description	Units	Dimension
CCF	**O3cost**	**Cost function for short-lived O_3**	**K kg$(NO_2)^{-1}$**	**4-D**
CCF	**CH4cost**	**Cost function for CH_4**	**K kg$(NO_2)^{-1}$**	**4-D**
tropo	**tp_PV**	**PV tropopause pressure**	**Pa**	**3-D**
tropo	**tp_WMO**	**WMO tropopause pressure**	**Pa**	**3-D**
tropo	**PV**	**Potential vorticity**	**PVU**	**4-D**
ECHAM5	**geopot**	**Geopotential**	**m^2 s^{-2}**	**4-D**
ECHAM5	**geosp**	**Surface geopotential**	**m^2 s^{-2}**	**3-D**
ECHAM5	**tm1**	**Dry air temperature**	**K**	**4-D**
ECHAM5	**tpot**	**Potential temperature**	**K**	**4-D**
ECHAM5	**tvirt**	**Virtual temperature**	**°C**	**4-D**
ECHAM5	**rhum**	**Relative humidity**	**%**	**4-D**
ECHAM5	**aps**	**Surface pressure**	**Pa**	**3-D**
ECHAM5	**um1**	**Zonal wind times cosine of latitude**	**m s^{-1}**	**4-D**
ECHAM5	**vm1**	**Meridional wind times cosine of latitude**	**m s^{-1}**	**4-D**
ECHAM5	**cossza**	**Cosine of solar zenith angle**	**-**	**3-D**
ECHAM5	**zi0**	**Solar irradiance**	**W m^{-2}**	**3-D**
g3b	tke	Turbulent kinetic energy	m^2 s^{-2}	4-D
g3b	**vervel**	**Vertical velocity**	**m s^{-1}**	**4-D**
g3b	**q**	**Specific humidity**	**kg kg^{-1}**	**4-D**
lnox	xnox_ave	Lightning NO_x emission	kg(N) s^{-1} m^{-3}	4-D
lnox	fpscg_ave	CG flash frequency	s^{-1}	3-D
lnox	fpsic_ave	IC flash frequency	s^{-1}	3-D
lnox	fpsm2cg_ave	CG flash frequency	s^{-1} m^{-2}	3-D
lnox	fpsm2ic_ave	IC flash frequency	s^{-1} m^{-2}	3-D
lnox	**npcanz_ave**	**Number of lightning events**	**-**	**3-D**
lnox	NOxcg_ave	CG NO_x lightning emission	kg(N)	3-D
lnox	NOxic_ave	IC NO_x lightning emission	kg(N)	3-D
lnox	telnox_ave	Lightning NO_x emission tendency	mol mol^{-1} s^{-1}	4-D
tracer	**N2O**	**N_2O mixing ratio**	**mol mol^{-1}**	**4-D**
tracer	**PAN**	**PAN mixing ratio**	**mol mol^{-1}**	**4-D**
tracer	**CO**	**CO mixing ratio**	**mol mol^{-1}**	**4-D**
tracer	**HNO3**	**HNO_3 mixing ratio**	**mol mol^{-1}**	**4-D**
tracer	**CH4**	**CH_4 mixing ratio**	**mol mol^{-1}**	**4-D**
tracer	**O3**	**O_3 mixing ratio**	**mol mol^{-1}**	**4-D**
tracer	**NO**	**NO mixing ratio**	**mol mol^{-1}**	**4-D**

tracer	**NO2**	**NO_2 mixing ratio**	**mol mol^{-1}**	**4-D**
tracer	**NO3**	**NO_3 mixing ratio**	**mol mol^{-1}**	**4-D**
tracer	**NOX**	**NO_x mixing ratio**	**mol mol^{-1}**	**4-D**
tracer	**HO2**	**HO_2 mixing ratio**	**mol mol^{-1}**	**4-D**
tracer	**O3P**	**$O(^3P)$ mixing ratio**	**mol mol^{-1}**	**4-D**
tracer	**OH**	**OH mixing ratio**	**mol mol^{-1}**	**4-D**

In addition, several variables were derived from the REACT4C data. First, pressure relative to WMO/PV tropopause pressure. This rescales the pressure altitude of an emission to that of the tropopause as the water vapour analysis showed this was valuable and the ozone CCF shows interesting altitude trends in Figure 6.1. Second, latitude of the jet stream for a specific longitude, i.e. the latitude of the largest wind speed. There seems to be very little zonal variation, so the meridional distance relative to the jet stream may numerically express the observed trends of higher ozone CCFs in the jet stream. Third, wind direction and speed converted into in polar coordinates, as this is easier to visualise and understand. Note that the REACT4C variables $um1$ and $vm1$ are scaled by the cosine of the latitude. Finally the season (winter or summer) of the weather pattern was encoded as a category. It might be desirable to add a factor for seasonal behaviour of the data and if using the weather patterns directly presents issues for generalising any results found.

As for water vapour, the absolute geopotential was calculated from the relative geopotential and the surface geopotential. When $geopot$ is used with a qualifier in these results, absolute geopotential is meant. The virtual temperature again appears erroneous.

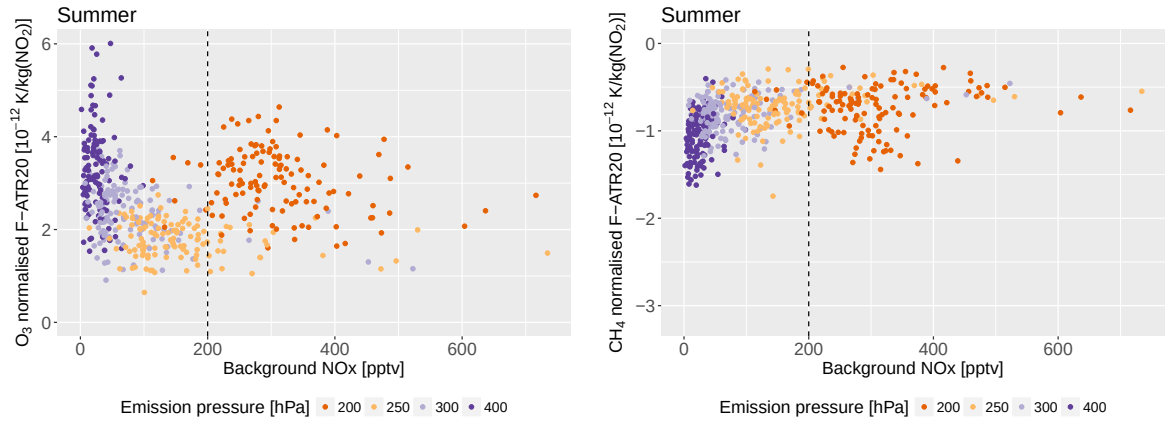
6.3. Relationship of ozone and methane with background NO_x

Stevenson and Derwent (2009) presented results from 111 NO_x perturbation studies distributed across the globe, and find an exponential fit relating ozone AGWP100 to latitude and background NO_x at 250 hPa, and a separate fit relating methane AGWP100 to background NO_x at 250 hPa. All emissions occur in the month July between 200–300 hPa and most locations are in the Northern Hemisphere. The locations with low background NO_x, high ozone impact, and high (negative) methane impact are mostly in the Southern Hemisphere or the North Pacific. Their background NO_x mixing ratios vary between roughly 20–210 pptv, and their pulse size is 10 Gg(NO_2) per location. Established theories on the relationship between ambient NO_x mixing ratios and ozone production rates and OH concentrations were mentioned in Section 6.1.3.

Figure 6.8 shows both ozone and methane CCF data split per altitude and season and related to background NO_x in [pptv] at the time and location of emission. Exponential fits from Stevenson and Derwent (2009) are not super-imposed due to the difficulty of scaling the climate metrics, and because latitude is included in their ozone fit. Latitude of emission is not visualised in the figure because it is far less distinctive than altitude. Results from Stevenson and Derwent (2009) show a much stronger methane response relative to ozone than ours, which is partially caused by the choice of metric as the methane RF pulse decays much slower than the ozone pulse. The pulse size in the REACT4C data is 0.15 Gg(NO_2) so chemical non-linearity may explain part of the difference.

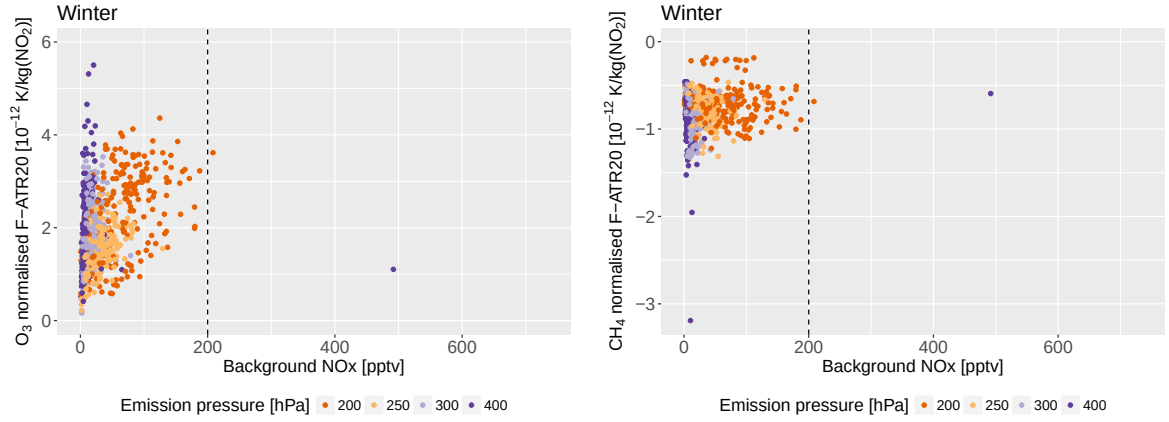
The data for emissions in summer compare well with the exponential fit results from Stevenson and Derwent (2009), except for 200 hPa data. The distinction between 200 hPa and 250 hPa data here is at roughly 200 pptv, also the maximum in the dataset used by Stevenson and Derwent (2009). However, for both ozone and methane climate impact from NO_x emissions in summer, the plots appear to show that 250 hPa emissions with higher background mixing ratios do follow the exponential trend and 200 hPa emissions with background mixing ratios below 200 pptv do not follow the exponential trend. This is based on a relatively small amount of data points though.

The cut-off of roughly 200 pptv background NO_x mixing ratio for the results in Stevenson and Derwent (2009) and for the agreement of these results with those in Stevenson and Derwent (2009), appears to coincide with the 0.1–0.2 ppbv maximum ozone production rate and OH concentration found by Brasseur et al. (1998). If this distinction at 200 pptv is caused by saturation effects in NO_x chemistry, however, the 200 hPa CCF data should be significantly lower instead of higher. Grewe et al. (2012) find a much higher ambient NO_x mixing ratio for maximum ozone production rate (about 2 ppbv). In both results, the relationship between NO_x mixing ratio and ozone production rate is linear for low NO_x mixing ratios, decreases below linear up to the maximum ozone production rate, and beyond that point



(a) Ozone summer

(b) Methane summer



(c) Ozone winter

(d) Methane winter

Figure 6.8: Scatter plots of ozone and methane CCF data versus background NO_x concentration coloured by level and split per season. The vertical line at 200 pptv shows the extent of the data used in Stevenson and Derwent (2009).

a higher NO_x mixing ratio leads to a decreased rate of ozone production. A higher background NO_x mixing ratio should thus never lead to a higher climate impact from the same mass of emitted NO_x.

A more likely explanation for the distinction at 2 pptv NO_x or 200 hPa emissions is the destination of the tracer in the next 90 days of simulation. Only a limited amount of information on the climate impact of a NO_x emission is likely to be contained in the instantaneous data used here. NO_x emitted at 200 hPa may have a longer lifetime or may be transported to regions with lower background NO_x, leading to higher ozone production and methane destruction in the tracer compared to emissions at 250 hPa.

The data for emissions in winter do not compare well to results from Stevenson and Derwent (2009). In winter only one data point significantly exceeds 200 pptv background NO_x mixing ratio. These results give reason to doubt the applicability of findings from Stevenson and Derwent (2009) above 200 pptv background NO_x mixing ratios, at 200 hPa, and in winter for the Northern Hemisphere. They also reinforce the findings from Stevenson and Derwent (2009) within the domain of their dataset. Further analysis of the seasonal variability of NO_x climate impact is warranted. Note that Stevenson and Derwent (2009) state that their results are not expected to be valid beyond July emissions.

These plots may aid in explaining the altitude behaviour of the ozone CCF data (Figure 6.1). Both in summer and winter an increase in background NO_x and a decrease in ozone CCF are observed for a pressure decrease from 400 hPa to 250 hPa. At 200 hPa the background NO_x is higher again but the ozone CCF is higher too, which is a departure from the exponential trend. The analysis here shows that the likely cause is the lifetime of a higher altitude NO_x emission and not non-linear chemistry (at the time and location of emission).

6.4. Ozone algorithm candidates

Scatter plot matrices and regression techniques were used to identify the strongest (combinations of) predictors from the 30 variables in Table 6.1 plus dimensions and derived variables. Due to time constraints for detailed analysis and space constraints for presenting results here, four algorithm candidates are discussed in detail.

A subjective selection was made from preliminary results based on the highest adjusted R^2 per amount of predictors in the regression. Early results revealed that high adjusted R^2 values were not attainable for univariate regression, in contrast to the water vapour CCF analysis in Section 5.3. The specific requirement devised for an ozone algorithm is that the results should not become negative, as negative ozone climate impact is plausible but outside the domain of the dataset used here.

6.4.1. Linear dependence on specific humidity

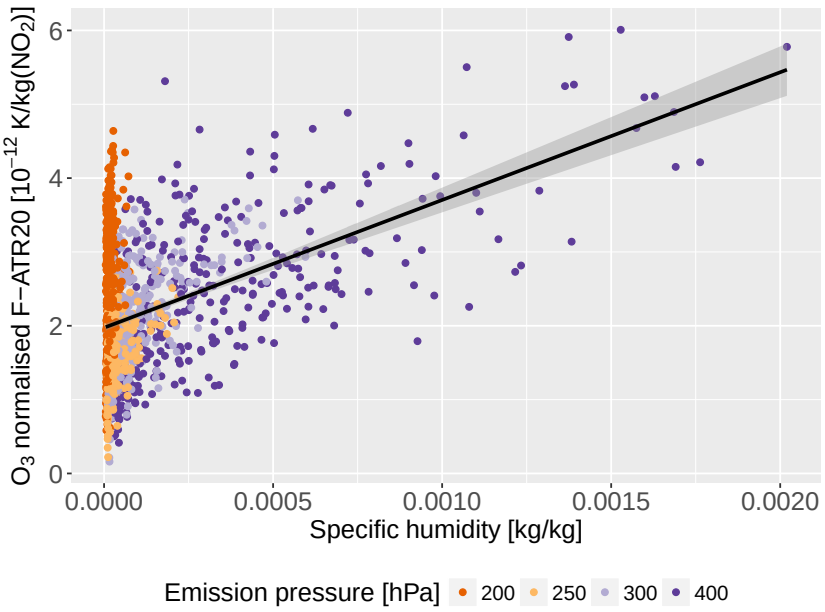


Figure 6.9: Scatterplot of ozone CCF data versus specific humidity coloured by level, with the linear regression found shown as a black line with 95% confidence interval for the regression coefficients in Table 6.2.

Figure 6.9 shows the ozone REACT4C CCF data versus the specific humidity q at the point of emission. The overall pattern is linearly increasing, and applying simple linear regression gives the following least-squares fit:

$$\text{F-ATR}\,20_{\text{O}_3} = \beta_0 + \beta_1 \times q + \epsilon \tag{6.1}$$

The parameter values from least squares optimisation are summarised in Table 6.2.

Table 6.2: Parameter overview for ozone linear specific humidity algorithm. The p-value for β_0 is beyond double-precision floating point range.

	Units	Estimate	Std. Error	t value	$Pr\,(X > \lvert t \rvert)$
β_0	K kg(NO_2)$^{-1}$	1.97×10^{-12}	2.50×10^{-14}	7.91×10^{1}	$< 2.23 \times 10^{-308}$
β_1	K kg(NO_2)$^{-1}$ (kg kg^{-1})$^{-1}$	1.73×10^{-9}	9.44×10^{-11}	1.83×10^{1}	4.57×10^{-67}

Visual inspection of Figure 6.9 reveals that the linear regression only properly captures the 400 hPa data, and that even for 400 hPa the variance around the regression is large. The relationship between ozone CCF and humidity seems to depend on emission altitude. The data for 200 hPa is practically independent of q. The lower the altitude, the clearer the trend. Specific humidity increases in summer and decreases with latitude, thus acting as a proxy for season and latitude at 400 hPa.

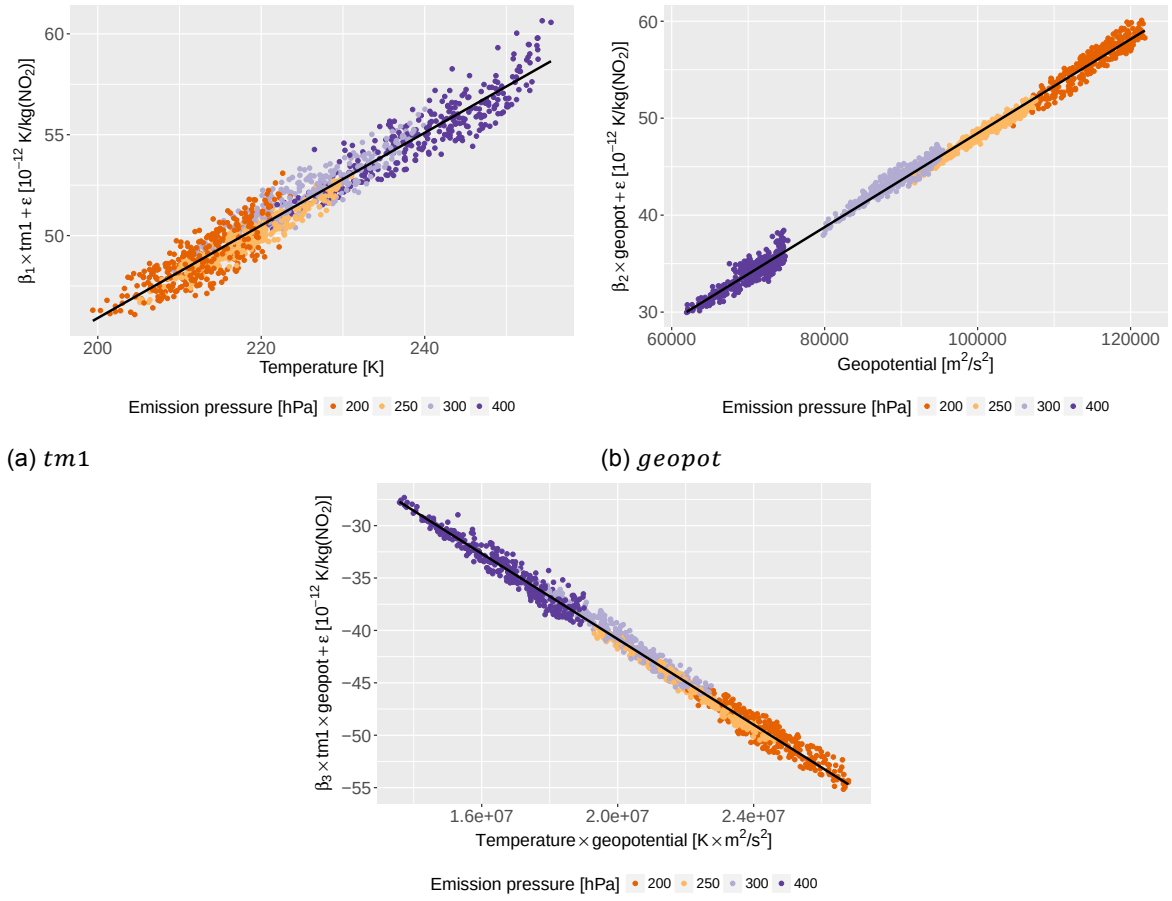
Physically, specific humidity is known to control OH production through Reaction (3.11) (Fuglestvedt et al., 1999). Higher humidity at the time of emission leads to a higher background OH production rate, which the Lagrangian OH production rate is proportional to (Grewe et al., 2014b). The Lagrangian

OH mixing ratio then drives the HNO_3 production rate through Reaction (3.10). Higher washout rates through higher HNO_3 production should lead to lower CCFs. Without verification by examining the production rates over a timespan, it is prudent to assume the correlation is simply caused by seasonal and zonal proxy.

Addition of season to the regression gives a non-significant interaction term (the slope does not change), but the y-intercept spreads ± 0.3 with summer higher (similar to the WP boxplot in Figure 6.1). Thus specific humidity does not accurately represent the seasonal trend on its own. Addition of latitude to the regression makes the q term superfluous - latitude is a stronger predictor than specific humidity but this confirms that q contains latitude information. Addition of level to the regression gives a non-significant interaction, but the direct effect of level adds accuracy. Thus specific humidity does not accurately represent the altitude trend on its own.

Climatological means of tropospheric specific humidity presented by Tian et al. (2013) do not exceed 1.5 g kg^{-1} for 400 hPa and 3.0 g kg^{-1} for 500 hPa pressure altitudes. This algorithm is unlikely to be valid at flight altitudes below 400 hPa, as 3.0 g kg^{-1} would lead to a CCF value 30% higher than the maximum of this dataset. Their maximum for 300 hPa is 0.45 g kg^{-1}. Due to the positive y-intercept, lower values of specific humidity are not a problem. The daily maximum specific humidity at 400 hPa is not considered but may present problems.

6.4.2. Bilinear dependence on temperature and geopotential



(a) $tm1$

(b) $geopot$

(c) $tm1 \times geopot$

Figure 6.10: Component plus residual scatter plots for the bivariate regression in Equation (6.2) and parameters in Table 6.3 coloured by emission pressure. Each x-axis is simply the values of the predictor, while each y-axis is the sum of the total residuals ϵ from the regression and the product of the regression coefficient β_i and the data on the x-axis. The black line shows the product of the regression coefficient β_i and the data on the x-axis, so without the residuals.

Figure 6.10 shows the ozone CCF data versus the geopotential and temperature at the point of

emission. The overall pattern is roughly linearly increasing, which leads to the following fit including interaction:

$$\text{F-ATR}\,20_{\mathrm{O_3}} = \beta_0 + \beta_1 \times tm1 + \beta_2 \times geopot + \beta_3 \times tm1 \times geopot + \epsilon \tag{6.2}$$

The parameter values from least squares optimisation are summarised in Table 6.3.

Table 6.3: Parameter overview for ozone bilinear temperature and geopotential algorithm.

	Units	Estimate	Std. Error	t value	$Pr\,(X > \lvert t \rvert)$
β_0	K kg(NO_2)$^{-1}$	-5.20×10^{-11}	2.34×10^{-12}	-2.22×10^{1}	2.36×10^{-93}
β_1	K kg(NO_2)$^{-1}$ K^{-1}	2.30×10^{-13}	1.04×10^{-14}	2.20×10^{1}	3.65×10^{-92}
β_2	K kg(NO_2)$^{-1}$ (m^2 s^{-2})$^{-1}$	4.85×10^{-16}	2.66×10^{-17}	1.82×10^{1}	2.46×10^{-66}
β_3	K kg(NO_2)$^{-1}$ (K m^2 s^{-2})$^{-1}$	-2.04×10^{-18}	1.20×10^{-19}	-1.70×10^{1}	1.03×10^{-58}

Temperature was observed as a weak predictor on its own but strong when combined with pressure altitude, especially when 200 hPa data is left out. As geopotential is more continuous and results in a stronger regression than pressure altitude it is used here.

Geopotential and temperature both increase in summer and decrease with latitude. Geopotential increases with pressure altitude while temperature decreases, potentially modelling the non-linear altitude trend seen in the ozone CCF data. Phsyically, geopotential is a representation of altitude and some altitude variation was identified in the CCF data and in literature. The definition of geopotential is:

$$\Phi\,(h) = \int_0^h g dz \tag{6.3}$$

Where Φ is the geopotential which is thus the vertically integrated gravitational acceleration from sea-level up to the altitude of interest. Commonly, geopotential height is used, which is $\frac{\Phi}{g_0}$ but as this is scaled by a constant the geopotential is used directly here.

Temperature directly feeds into all reactions on the background and foreground tagged models through the Arrhenius equation, but the impact of temperature at the time of emission on the 90-day ozone time history is likely to be small through this path. In Figure 3.5 and Figure 3.8 the majority of the NO_x and ozone mass changes occur several days or weeks after emission. For a given pressure level, temperature and geopotential are correlated, as is visible from the colour coding in Figure 6.10. This is sensible: in the troposphere buoyancy causes hot air to rise, thus for the same pressure hotter air is located at a higher altitude, and for the same pressure and latitude a high/low geopotential signifies a ridge/trough and hot/cold air influx from the south/north. At 200 hPa where the correlation between temperature and geopotential is the weakest (not shown), there is a split between emissions released into the stratosphere or troposphere. Temperature is approximately constant in the tropopause region and increases with altitude in the stratosphere, thus the correlation between geopotential and temperature is not expected to hold for the stratosphere.

Adding seasonality to the regression separately, which increases the other parameters and subtracts a constant for winter data, adds significant accuracy to the regression. The seasonality is not properly modelled. Addition of pressure level to the regression adds somewhat less accuracy than season. The non-linear altitude behaviour is not captured properly.

The y-intercept (β_0) is negative, though an ambient temperature of (close to) 0 K is unphysical for the UTLS. Climatological means of tropospheric temperature from Tian et al. (2013) do not exceed below 205 K for 300 hPa and the minimum in this dataset in 200 K. At this temperature, the algorithm gives zero CCF for a geopotential of 78000 m^2 s^{-2}, roughly 8 km. This is roughly the altitude of the 400 hPa data, but the temperature is around 220 K at this altitude both in Tian et al. (2013) and this dataset. Within the 200–400 hPa domain this regression is unlikely to give negative ozone CCF results.

6.4.3. Trilinear dependence on temperature, geopotential and specific humidity

Figure 6.11 shows the ozone CCF data versus the geopotential, temperature, interaction of the two, and specific humidity at emission. The overall pattern is roughly linearly increasing for each variable, which leads to the following trivariate fit including interaction:

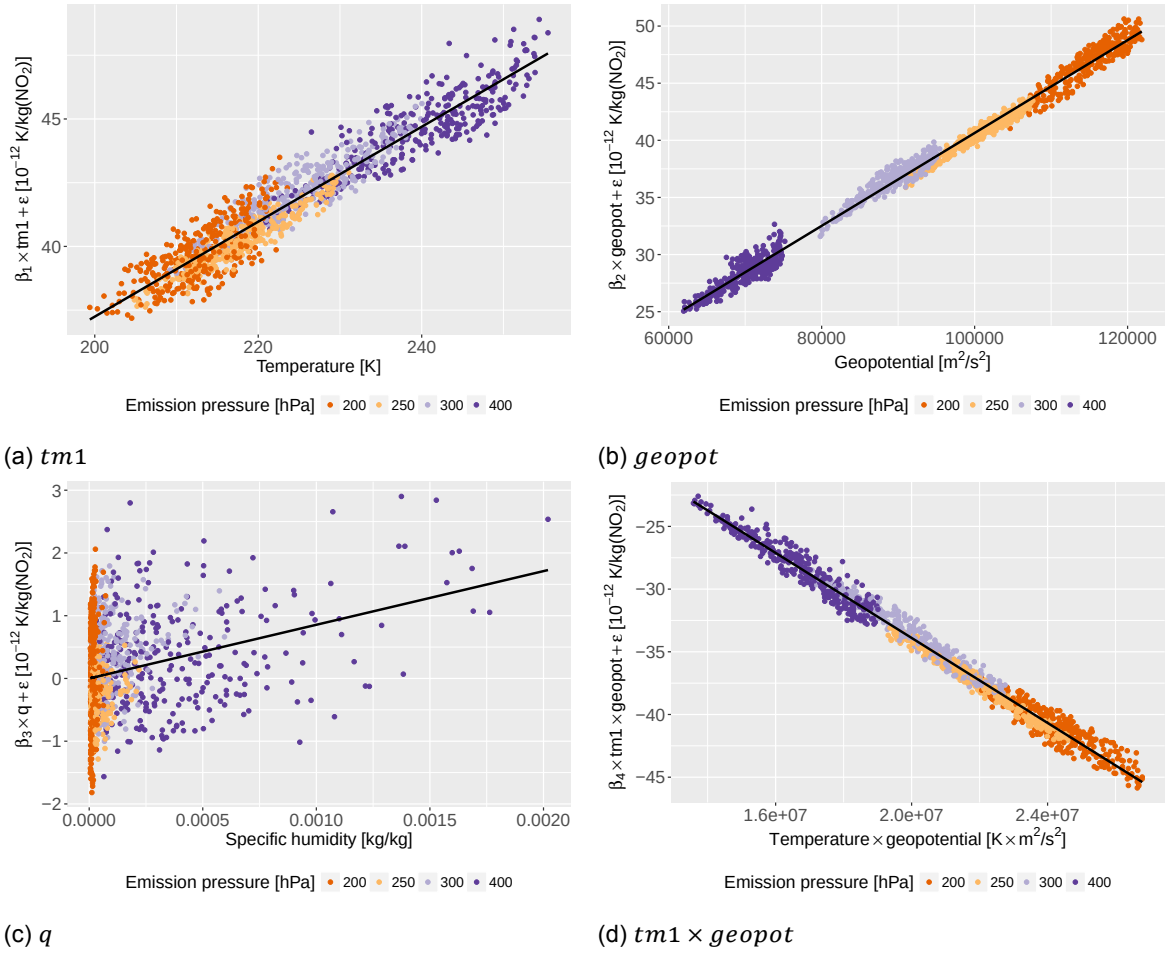


Figure 6.11: Component plus residual scatter plots for the trivariate regression in Equation (6.4) and parameters in Table 6.4 coloured by emission pressure. Each x-axis is simply the values of the predictor, while each y-axis is the sum of the total residuals ϵ from the regression and the product of the regression coefficient β_i and the data on the x-axis. The black line shows the product of the regression coefficient β_i and the data on the x-axis, so without the residuals.

$$\text{F-ATR}\,20_{\text{O}_3} = \beta_0 + \beta_1 \times tm1 + \beta_2 \times geopot + \beta_3 \times q + \beta_4 \times tm1 \times geopot + \epsilon \tag{6.4}$$

The parameter values from least squares optimisation are summarised in Table 6.4.

Table 6.4: Parameter overview for ozone trilinear temperature, geopotential and specific humidity algorithm.

	Units	Estimate	Std. Error	t value	$Pr\,(X > \lvert t \rvert)$
β_0	K kg$(NO_2)^{-1}$	-4.24×10^{-11}	2.75×10^{-12}	-1.54×10^{1}	2.33×10^{-49}
β_1	K kg$(NO_2)^{-1}$ K^{-1}	1.86×10^{-13}	1.23×10^{-14}	1.52×10^{1}	3.26×10^{-48}
β_2	K kg$(NO_2)^{-1}$ $(m^2\ s^{-2})^{-1}$	4.06×10^{-16}	2.89×10^{-17}	1.41×10^{1}	4.70×10^{-42}
β_3	K kg$(NO_2)^{-1}$ $(kg\ kg^{-1})^{-1}$	8.55×10^{-10}	1.32×10^{-10}	6.46×10^{0}	1.48×10^{-10}
β_4	K kg$(NO_2)^{-1}$ $(K\ m^2\ s^{-2})^{-1}$	-1.69×10^{-18}	1.30×10^{-19}	-1.30×10^{1}	1.27×10^{-36}

This algorithm combines strengths and weaknesses of the univariate and bivariate options presented previously. The effect of adding season, latitude or level to the regression is similar to the bivariate algorithm. The latitude trend is modelled slightly better, and addition of level now adds more accuracy than season. Addition of interactions with the pressure level makes specific humidity redundant. This implies that the seasonal trend is better captured with the addition of q and that q contains some information about the altitude dependence of the temperature and geopotential dependences.

The addition of an interaction between specific humidity and geopotential is statistically significant but adds little accuracy for its complexity. Specific humidity varies with geopotential roughly exponen-

tially, with the slope decreasing with pressure level. Versus temperature the specific humidity is noisy but bounded by an exponentially increasing maximum (critical humidity). The highest values of the ozone CCF occur for almost the highest specific humidity at 400 hPa so q seems to be useful mainly for predicting some of the positive outliers.

The y-intercept is negative again. If specific humidity is zero and 200 K temperature is again considered, negative results occur at around 70000 m^2 s^{-2}.

6.4.4. Quadrilinear dependence on temperature, geopotential, specific humidity and solar irradiance

Figure 6.12 shows the ozone CCF data versus the geopotential, temperature, interaction of the two, specific humidity, and solar irradiance at emission. The overall pattern is roughly linearly increasing for each variable, which leads to the quadrivariate fit including interaction given in Equation (6.5).

$$\text{F-ATR}\,20_{\text{O}_3} = \beta_0 + \beta_1 \times tm1 + \beta_2 \times geopot + \beta_3 \times q + \beta_4 \times zi0 + \beta_5 \times tm1 \times geopot + \epsilon \qquad (6.5)$$

The parameter values from least squares optimisation are summarised in Table 6.5.

Table 6.5: Parameter overview for ozone quadrilinear temperature, geopotential, specific humidity and solar irradiance algorithm.

	Units	Estimate	Std. Error	t value	$Pr\,(X > \lvert t \rvert)$
β_0	K $kg(NO_2)^{-1}$	-3.97×10^{-11}	2.76×10^{-12}	-1.44×10^{1}	1.25×10^{-43}
β_1	K $kg(NO_2)^{-1}$ K^{-1}	1.76×10^{-13}	1.23×10^{-14}	1.44×10^{1}	1.45×10^{-43}
β_2	K $kg(NO_2)^{-1}$ $(m^2\ s^{-2})^{-1}$	4.01×10^{-16}	2.86×10^{-17}	1.40×10^{1}	8.16×10^{-42}
β_3	K $kg(NO_2)^{-1}$ $(kg\ kg^{-1})^{-1}$	9.24×10^{-10}	1.32×10^{-10}	7.03×10^{0}	3.39×10^{-12}
β_4	K $kg(NO_2)^{-1}$ $(W\ m^{-2})^{-1}$	3.26×10^{-16}	5.86×10^{-17}	5.57×10^{0}	3.11×10^{-8}
β_5	K $kg(NO_2)^{-1}$ $(K\ m^2\ s^{-2})^{-1}$	-1.70×10^{-18}	1.29×10^{-19}	-1.32×10^{1}	1.91×10^{-37}

This regression adds solar irradiance to the trivariate algorithm. Solar irradiance in this dataset is a 3-D variable that is uncorrelated to geopotential, temperature, and humidity. The cosine of the solar zenith angle, $cossza$, is a stronger predictor but its improvements compared to $zi0$ lie in the range of a solar zenith angle above 90° i.e. beyond the horizon. When there is no sunlight incident, $zi0$ is simply zero but $cossza$ is continuous.

Solar radiation was identified as a driver in literature (Section 6.1, mainly because Reaction (3.6) requires light below a specific wavelength to take place and this reaction is needed for a NO_x emission to produce ozone.

The effect of adding season, latitude or level to the regression is similar to the trivariate algorithm. Addition of interactions with season or level is statistically insignificant. Addition of season or level directly adds a small amount of accuracy and makes $zi0$ redundant, thus its addition in this algorithm improves the seasonal and altitudinal trend. The addition of an interaction between specific humidity and geopotential, solar irradiance and specific humidity or solar irradiance and temperature is statistically significant but adds little accuracy for its complexity.

For zero humidity and zero incoming sunlight, the algorithm result becomes zero at 76000 m^2 s^{-2}. The solar irradiance as encoded in this dataset is irradiance at the top of the atmosphere, which is why the data in Figure 6.12 range from 0-1350 W m^{-2}. This range applies to all locations on Earth and the effect of $zi0$ in the regression is relatively small, so no extra problems are anticipated with application of this algorithm compared to the trivariate candidate.

6.5. Ozone algorithm trade-off

Based on the algorithm requirements in Section 4.4, a set of trade-off criteria were devised in Section 4.6. Table 6.6 provides an overview of the results in this section.

6.5.1. Pattern of residuals

The mean values of the spread of residuals per pressure level and WP are:

- q : 2.49×10^{-12}

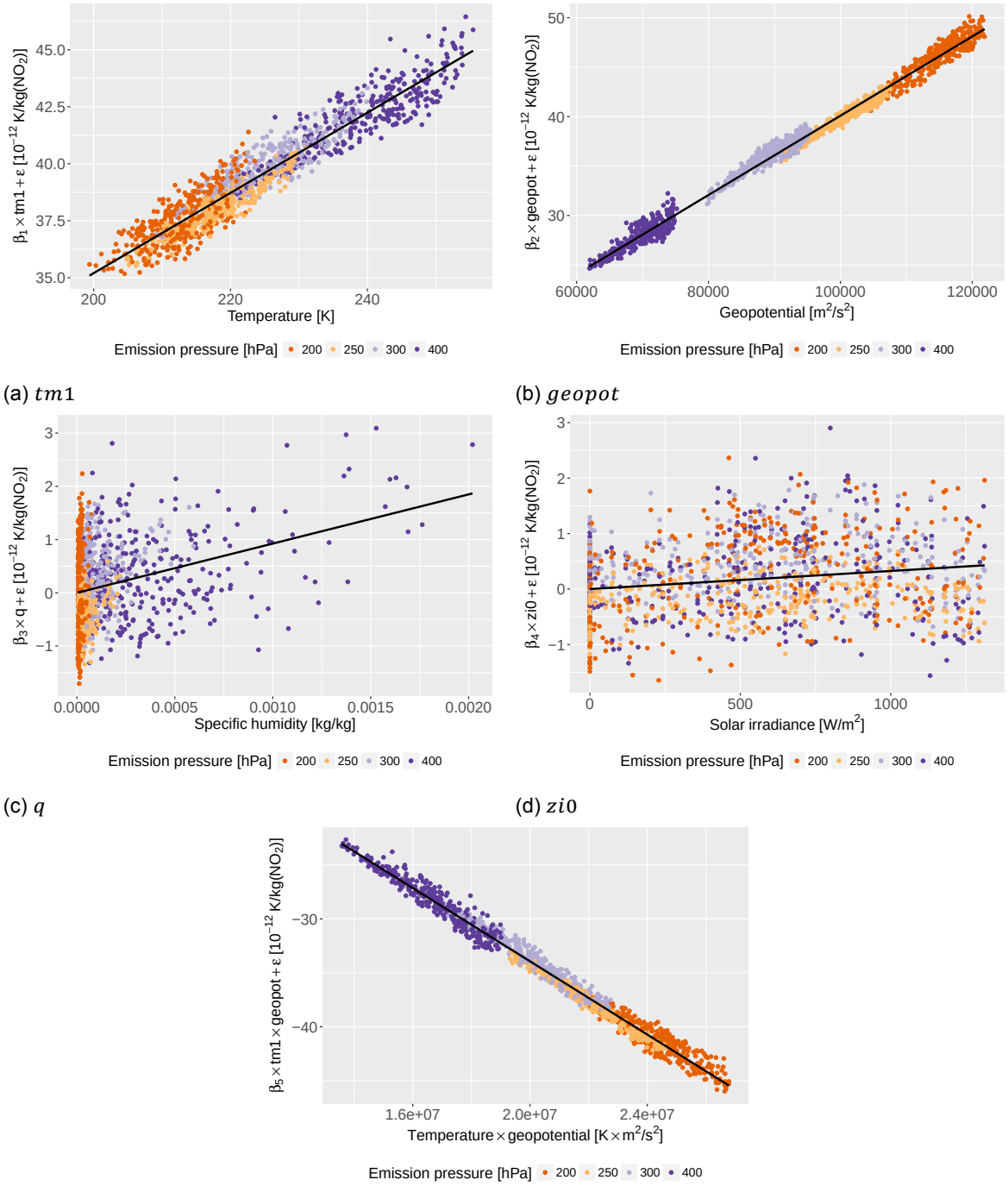


Figure 6.12: Component plus residual scatter plots for the quadrivariate regression in Equation (6.5) and parameters in Table 6.5 coloured by emission pressure. Each x-axis is simply the values of the predictor, while each y-axis is the sum of the total residuals ϵ from the regression and the product of the regression coefficient β_i and the data on the x-axis. The black line shows the product of the regression coefficient β_i and the data on the x-axis, so without the residuals.

- $tm1 \times geopot$: 2.39×10^{-12}
- $tm1 \times geopot + q$: 2.34×10^{-12}
- $tm1 \times geopot + q + zi0$: 2.35×10^{-12}

Interestingly, more variables is not always better. This is sensible though unexpected, as regular

Table 6.6: Overview of results of the subjective ozone algorithm trade-off, per algorithm and category. If three or more results are distinct, they are ranked into +, ± and −. If only two distinct categories are identified, only + and − are used.

Algorithm	Pattern	Accuracy	External veracity	Background	Parsimony
q	−	−	−	−	+
$tm1 \times geopot$	±	±	+	+	±
$tm1 \times geopot + q$	+	+	±	−	−
$tm1 \times geopot + q + zi0$	+	+	±	−	−

least-squares regression does not optimise for this statistic.

Analysing box plots of the residuals (not shown) per dimension for the univariate algorithm gives clear residual patterns for WP, latitude and altitude. The bivariate, trivariate and quadrivariate algorithms predict the latitude trend well and are somewhat better at predicting the WP and altitude trends. The trivariate and quadrivariate algorithms present little improvement over the bivariate option. All algorithms return minimum results for 300 hPa altitude instead of 250 hPa like the original data.

The 2-D autocorrelation test (Moran's I) is inconclusive. The amount of pressure level/WP sets that are statistically significant varies per algorithm, but none give a negative I. The univariate (q) algorithm clearly produces the worst pattern, the quadrivariate algorithm is slightly worse than the trivariate algorithm, and the bivariate algorithm is an intermediate option.

6.5.2. General accuracy

The RSE (and adjusted R^2) for each algorithm within this dataset are:

- q : 8.02×10^{-13} (0.20)
- $tm1 \times geopot$: 6.84×10^{-13} (0.42)
- $tm1 \times geopot + q$: 6.74×10^{-13} (0.43)
- $tm1 \times geopot + q + zi0$: 6.67×10^{-13} (0.45)

Again the bivariate option shows a marked improvement over the univariate option, with diminishing returns for the more complex algorithms.

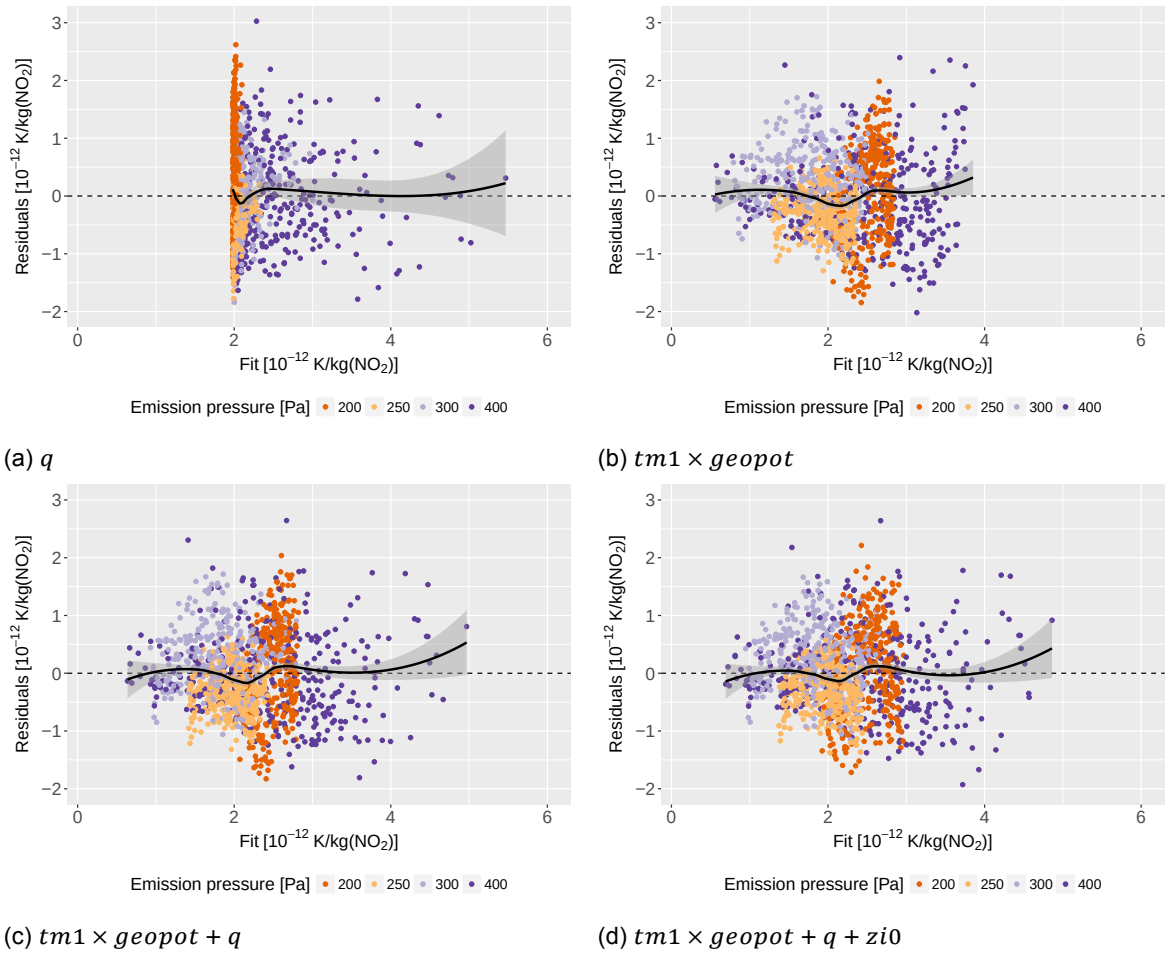
Figure 6.13 shows residual plots for the four algorithms. The univariate algorithm has a clear minimum fit value which is the y-intercept (β_0). The other three algorithms show little difference. There is no significant heteroskedasticity visible. There is a slight amount of trend left in all the residuals, but note that the LOESS trend is inaccurate on the extremes due to the lack of data. These results do not clearly favour any candidate.

The univariate option produces the worst results overall, with a marked improvement at the bivariate level, and more nuanced improvements for the trivariate and quadrivariate algorithms.

6.5.3. External veracity

Results here are summarised from Section 6.4.

- q: humidity can never be negative and the intercept is reasonable. It appears that the linear trend only captures 400 hPa data in Figure 6.9. High humidity at lower altitudes than 400 hPa limits the domain of application but this is limited to altitude shifts. It is unclear how high outliers in specific humidity could be in the 400 hPa range and whether application at lower latitudes is safe.
- $tm1 \times geopot$: the intercept is negative, while the ozone CCF should not become negative in this model set-up. This can be mitigated by adding a floor at zero. Only for very high temperature or geopotential (thus altitude), which is unlikely to occur at the same time, can the results of the algorithm exceed the maximum of the REACT4C data.
- $tm1 \times geopot + q$: similar to bivariate. The addition of q introduces the uncertain influence of high humidity values.
- $tm1 \times geopot + q + zi0$: similar to trivariate. The addition of $zi0$ should give no problems due to the range of the data used.



(a) q (b) $tm1 \times geopot$

(c) $tm1 \times geopot + q$ (d) $tm1 \times geopot + q + zi0$

Figure 6.13: Scatterplots of residuals versus fit for the four ozone algorithms. Residuals are original values minus fitted values. The black curves are LOESS regressions to identify the trend of the residuals over the algorithm values, with 95% confidence bands.

Overall, specific humidity is the most uncertain parameter in terms of external veracity and this causes the trivariate and quadrivariate algorithms to do worse than the bivariate algorithm. The extra parameter for the quadrivariate algorithm makes it no worse than the trivariate algorithm but also no better.

6.5.4. Mechanistic background

Results here are summarised from Section 6.4.

- q: humidity was not identified in literature as a driver, but it influences HNO_3 production through OH (Reaction 3.11) and varies with season and latitude.
- $tm1 \times geopot$: temperature is identified in literature as a driver, and geopotential is a way to express altitude. Both represent large-scale weather variation.
- $tm1 \times geopot + q$: as univariate and bivariate.
- $tm1 \times geopot + q + zi0$: in addition to the above, $zi0$ is identified in literature as a driver, has no altitudinal component, and varies predictably with longitude and latitude and day of the year.

Given that all data used are from the time of release, all variables are unlikely to directly affect the ozone CCF. They may be assumed to be proxies for the weather at a given location or proxies of the location itself, in both cases predicting the likely direction of transport of a tracer and its atmospheric lifetime. The differences between the algorithms are small for this category. Because q was not pre-identified as a driving variable, the bivariate algorithm gives the best results.

6.5.5. Parsimony

The less terms included in an algorithm, the more simple it is. Interaction terms are deemed less complex than an entirely separate variable for this thesis, especially for the bivariate option where the interaction term does not affect interactive 3-D plotting for presentations. This makes the univariate humidity algorithm the best in this category, while the trivariate and quadrivariate algorithms score poorly. The bivariate algorithm is a good compromise on simplicity.

6.5.6. Choice

The bivariate temperature and geopotential algorithm shows an overall strong improvement over the univariate algorithm. The trivariate and quadrivariate options are slightly more accurate but more complex and largely based on the same variables. Based on the scores assigned in Table 6.6, the bivariate algorithm presents a good compromise between accuracy and simplicity. The final form of the algorithmic CCF for ozone is:

$$\begin{aligned} \mathrm{a}\widetilde{\mathrm{CCF}}_{\mathrm{O_3}} = -5.20 \times 10^{-11} + 2.30 \times 10^{-13} \times tm1 + 4.85 \times 10^{-16} \times geopot \\ -2.04 \times 10^{-18} \times tm1 \times geopot \approx \text{F-ATR}\,20_{\mathrm{O_3}} \end{aligned} \tag{6.6}$$

$$\begin{aligned} \mathrm{aCCF}_{\mathrm{O_3}} &= \mathrm{a}\widetilde{\mathrm{CCF}}_{\mathrm{O_3}} \quad \text{for} \quad 0 \leq \mathrm{a}\widetilde{\mathrm{CCF}}_{\mathrm{O_3}} \\ \mathrm{aCCF}_{\mathrm{O_3}} &= 0 \quad \text{for} \quad \mathrm{a}\widetilde{\mathrm{CCF}}_{\mathrm{O_3}} < 0 \end{aligned} \tag{6.7}$$

Where the extra step is added to ensure no negative aCCF results can occur.

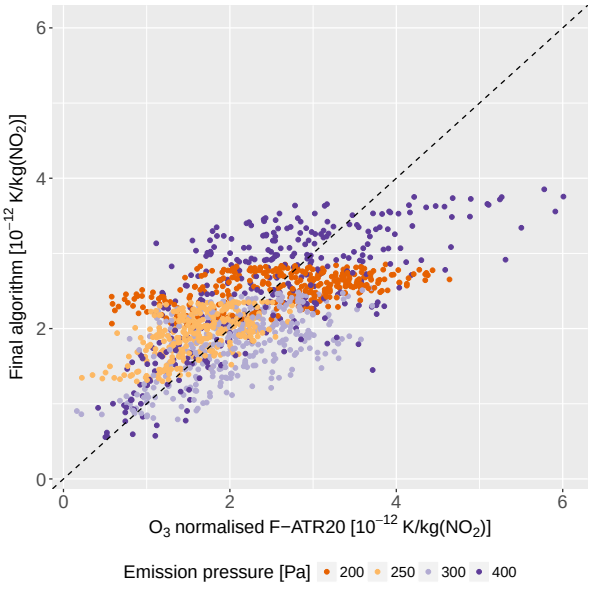


Figure 6.14: Scatterplot of ozone algorithmic CCF results versus REACT4C data coloured by level, with the black line indicating perfect agreement.

Figure 6.14 shows a comparison between the original CCF data and the final algorithm (aCCF) results, where points on the black line indicate perfect agreement. It is clear that the variability of the final algorithm results is lower than that of the original data. The algorithm appears to underpredict the highest ozone CCF values for emissions at 400 hPa, overpredict 250 hPa data, and underpredict 300 hPa data. No clear trend remains in Figure 6.14 but it is clear that the algorithm does not perfectly predict ozone CCF data.

6.6. Ozone discussion

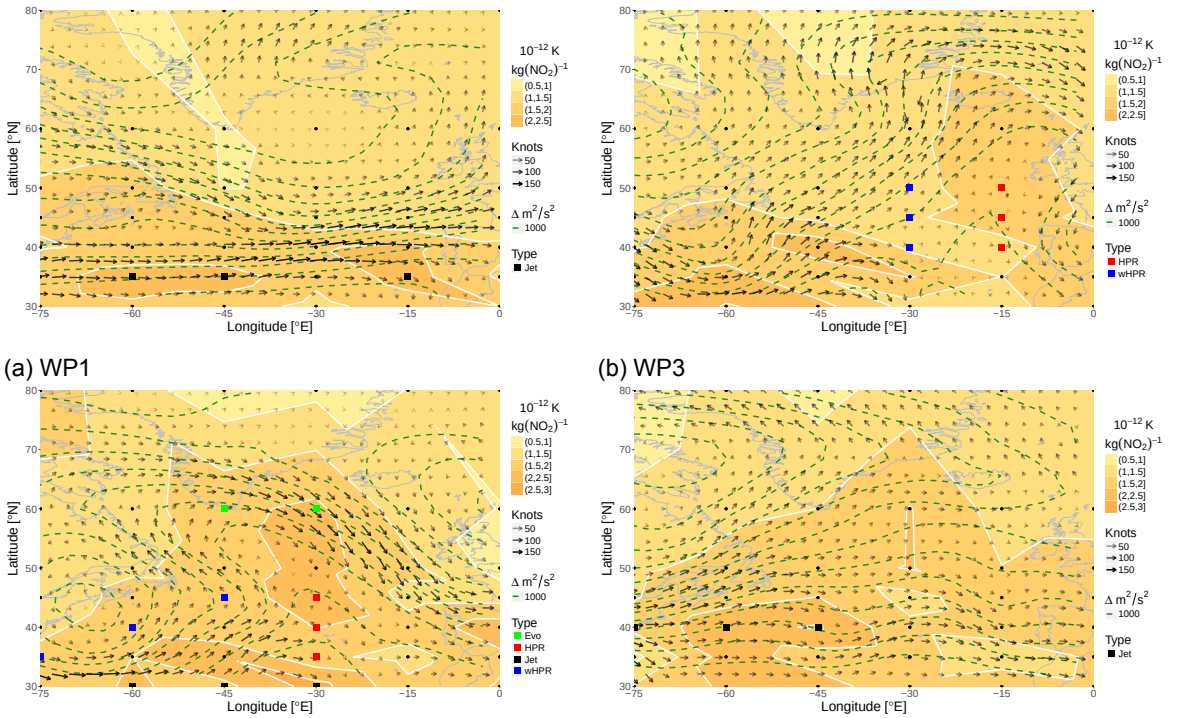
From the results in the previous sections and others not shown in this thesis, a number of discussions and conclusions arise. Here discussions on various topics are given as far as they relate directly to the results of the ozone CCF data, the weather data, and the algorithm formulation. General discussions regarding the methodology are given in Chapter 7.

6.6.1. Lightning and chemistry data

All lightning variables were rather sparse. The fact that it is the main source of background NO_x does not at all mean that instantaneous lightning data is useful. Lightning occurs infrequently so may lead to large spikes in NO_x production. Also, the contribution to background NO_x may occur predominantly in other latitudes, longitudes, altitudes, seasons, and times-of-day than included in this analysis.

Chemistry variables, especially OH and HO_2 for ozone and HNO_3 and HO_2 for methane had some predictive power in the regression analysis but did not add any information when meteorological variables were already included. NO_x was a weaker predictor than the aforementioned species. The intention was to include a more accurate algorithm that makes use of chemistry data in the trade-off.

6.6.2. Correlation with 2-D weather patterns



(a) WP1 (b) WP3

(c) WP4 (d) WP5

Figure 6.15: Ozone CCF bilinear contour plots, 12 UTC, 250 hPa, WPs 1, 3, 4, and 5. Square overlays represent locations used in analysis by Frömming et al. (2017), where black is jet stream, red is high-pressure ridge, blue is west of high-pressure ridge (Frömming et al. (2017) Table A.1), and green is portrayed in Frömming et al. (2017) Figure 7.

Figure 6.15 shows four 250 hPa ozone CCF contour plots, used here for comparing with Frömming et al. (2017) and a case study on the visual correlation with wind and geopotential. Frömming et al. (2017) examined the 250 hPa CCF results for all weather patterns from REACT4C. Based on visual likeness between the ozone CCF contours and wind and geopotential contours, they examined a subset of data from weather patterns 1, 3, 4, and 5. The fact that these weather patterns are chosen for analysis in Frömming et al. (2017) implies that they show the clearest visual correlation. They split the subset up into emissions inside the high pressure ridge (red points in Figure 6.15), emissions just west of the same high pressure ridge (blue points in Figure 6.15), and emissions in the jet stream (black points in Figure 6.15). Within this subset the peak ozone gain is found to occur earlier, at lower altitude, and at lower latitude for emissions inside the high pressure ridge (Frömming et al., 2017). This work expresses the same hypothesis that led to the research question for this thesis: the variation is largely determined by general meteorology at the time of emission, and is summarised in Grewe et al. (2017a). The two green points in Figure 6.15 are used to show the temporal development of NO_x and ozone in Frömming et al. (2017). Matthes et al. (2012) present similar temporal development plots in their Figure 44.5 for WP1, but these are based on preliminary data (V. Grewe, personal communication, 15 March, 2017).

The top-left plot in Figure 6.15 is similar to its 200 hPa version, that was previously used in Grewe

et al. (2014a) (i.e. Figure 1.1) to show the similarity of the ozone CCF pattern with wind and geopotential. The visual pattern is clear, though possibly biased by the interpolation, colour scheme and human error. If the pattern is truly significant it should be recognisable in scatter plot analysis too. General results across the full REACT4C dataset for wind speed and wind direction are among the worst of all variables in Table 6.1 - no meaningful correlation was found to the ozone CCF data.

The four longitude-latitude planes in Figure 6.15 are used here as a case study for a 2-D weather pattern algorithm basis. As they are all winter weather patterns and the pattern numbers are arbitrary, the allowed predictors for this case study are: wind speed, wind direction, geopotential, latitude, and longitude. Other primitive variables that could have been used are temperature and humidity, but these are out of scope for this case study as they are not visible in this type of weather map nor discussed in Frömming et al. (2017). Combining these predictors linearly with all interactions and eliminating statistically insignificant predictors leaves the following approximation, which is almost as accurate as the full model (for the reduced dataset) with much fewer variables:

$$\text{F-ATR}\,20_{\text{O}_3} = \beta_0 + \beta_1 \times lat + \beta_2 \times geopot + \epsilon \tag{6.8}$$

The parameter values from least squares optimisation are summarised in Table 6.7.

Table 6.7: Parameter overview for 2-D case study regression.

	Units	Estimate	Std. Error	t value	$Pr\,(X > \lvert t \rvert)$
β_0	K kg(NO_2)$^{-1}$	-2.06×10^{-12}	1.23×10^{-12}	-1.68×10^{0}	9.49×10^{-2}
β_1	K kg(NO_2)$^{-1}$ (°N)$^{-1}$	-8.33×10^{-15}	2.66×10^{-15}	-3.14×10^{0}	2.02×10^{-3}
β_2	K kg(NO_2)$^{-1}$ (m^2 s^{-2})$^{-1}$	4.06×10^{-17}	1.13×10^{-17}	3.58×10^{0}	4.46×10^{-4}

Geopotential is a stronger predictor within this subset than wind speed or direction, and the isolines have a similar pattern to the wind vectors in Figure 6.15. Adding latitude (lat) to the regression adds information about ridges and troughs in the geopotential field. The intercept (β_0) is non-physical but not a problem at this stage as this regression will not be applied at 0°N and 0 m^2/s^2 geopotential.

When applying this regression result to the subset of four longitude-latitude plans, the adjusted R^2 is 0.41. When applied to the total set of REACT4C data, even without forcing the same parameter values, the adjusted R^2 drops to 0.11. To continue this approach for all data would require separate regressions for all latitude-longitude planes, without a meaningful way of interpolating or extrapolating the results. A value of 0.41 is very close to the results from the bivariate algorithm and upwards presented here and 0.11 is less than for the univariate algorithm presented here (Section 6.4). This demonstrates that a visual correlation between CCF contours and contours of weather variables in a subset of the data does not necessarily translate to an actual correlation in the full dataset.

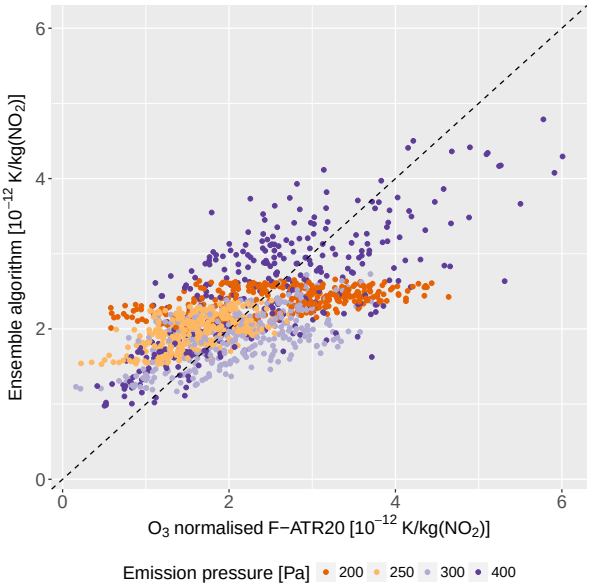
6.6.3. Ensemble-mean algorithm

Application of the mean of an ensemble of simulations is commonly used in meteorology to mitigate the effect of chaos on the results of a prediction. When the ensemble-mean is calculated from the four algorithms presented in Section 6.4, the RSE is 6.78×10^{-13} K kg(NO_2)$^{-1}$. The (adjusted) R^2 is 0.43. This is approximately the same result for absolute accuracy as the bilinear and trilinear algorithms (Section 6.5.2).

Figure 6.16 shows the resulting ensemble-mean algorithm results versus the original REACT4C data, for comparison to Figure 6.14. The results are difficult to distinguish from those of the bilinear temperature and geopotential algorithm. Comparisons to the trilinear and quadrilinear algorithms (not shown) yield similar results, as these algorithms include the same predictors and generally produce similar results (Section 6.5). The trend in 400 hPa data in Figure 6.16 does appear to improve over the bilinear algorithm, likely due to the influence of specific humidity on the other three algorithms.

Application of the ensemble-mean to these four algorithms does not appear to meaningfully improve the accuracy for the complexity it introduces. Considering that the trilinear algorithm includes the predictors of the linear and bilinear algorithms, and the quadrilinear algorithm adds one more predictor, the four algorithms are unlikely to predict different parts of the variability of the original CCF data. Taking the mean is likely to give similar results to the middle option in terms of complexity, as shown here.

Figure 6.16: Scatterplot of ozone ensemble (mean of four) algorithmic CCF results versus REACT4C data coloured by level, with the black line indicating perfect agreement.

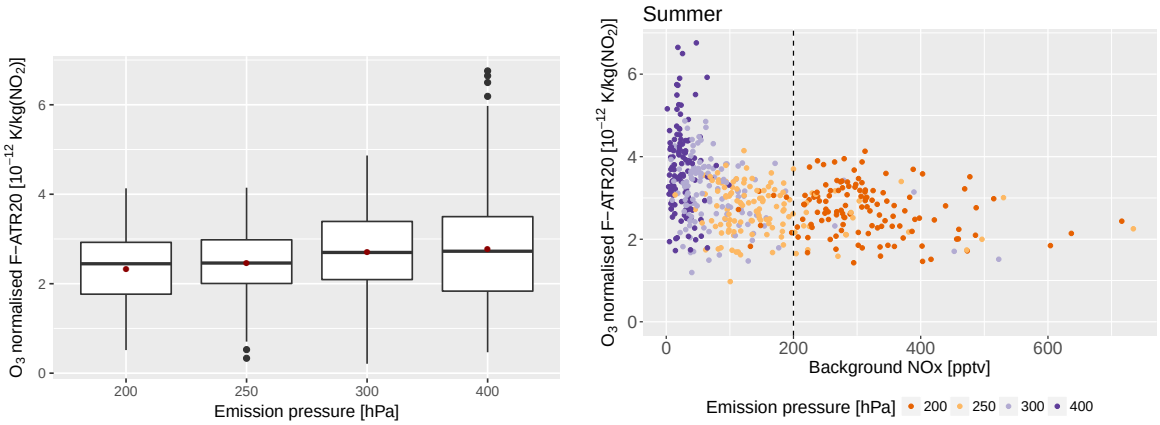


6.6.4. Maximum attainable adjusted R^2

When all the selected variables from Table 6.1 are combined in a linear regression model with longitude, latitude, level, season, and the latitude of the jet stream, the resulting adjusted R^2 is 0.57 for 36 variables. Introducing interaction terms of course increases the accuracy, but first-order interactions already require 666 parameters so the result is meaningless. Though imperfect, this gives an estimate of the amount of variation in the CCF data that can be explained by the combination of weather data, chemical data, and lightning data. If only the 19 weather variables are used, adjusted R^2 is 0.42, if only 13 chemistry variables are used, adjusted R^2 is 0.36, if only four dimensions are used, adjusted R^2 is 0.14.

6.6.5. Instantaneous RF scaling

The non-linear relationship with altitude for ozone CCF data in Figure 6.1 shows similarity to the instantaneous versus adjusted RF altitude scaling shown in Figure 3.7 (right). The closed-form equation for the fit used in this altitude scaling is given in Equation 3.24. This fit shows a minimum at 250 hPa while no data are used between about 220–330 hPa. This may (partially) explain the discrepancy between summer CCF data at 200 and 250 hPa in Figure 6.8. Using the expression in Equation 3.24, here the RF scaling is reversed to assess the sensitivity of the CCF data and the regression results to the application of this fit.



(a) Box plot per emission pressure.

(b) Scatter plot versus background NO_x mixing ratio.

Figure 6.17: Ozone CCF data after reversal of the REACT4C adjusted RF scaling.

Figure 6.17 (left) shows a box plot of the ozone CCF data versus altitude after the RF scaling was reversed, to compare to Figure 6.1. It appears that reversing the RF scaling removes almost all of the altitude trend from the REACT4C ozone CCF data. The remaining trend is more in terms of variance, i.e. the length of the boxes and whiskers, than in terms of the mean and median.

Figure 6.17 (right) shows a scatter plot of the ozone CCF data versus background NO_x mixing ratio after the RF scaling was reversed, to compare to Figure 6.8. When the RF scaling is removed, the summer ozone results compare well to Stevenson and Derwent (2009). The results for winter emissions (not shown) remain similar to those in Figure 6.8, so the observation that the relations in Stevenson and Derwent (2009) are not valid for winter is not sensitive to the RF scaling.

Table 6.8: Summary of Residual Standard Error (RSE) and adjusted R^2 for all four ozone algorithms before and after reversing the instantaneous/adjusted RF altitude scaling from Figure 3.7 (right).

Algorithm	RSE (adj R^2) before reversal	RSE (adj R^2) after reversal
q	8.02×10^{-13} (0.20)	8.19×10^{-13} (0.24)
$tm1 \times geopot$	6.84×10^{-13} (0.42)	7.79×10^{-13} (0.31)
$tm1 \times geopot + q$	6.74×10^{-13} (0.43)	7.54×10^{-13} (0.35)
$tm1 \times geopot + q + zi0$	6.67×10^{-13} (0.45)	7.38×10^{-13} (0.38)

The regressions from Section 6.4 were applied to the ozone CCF data, allowing the β coefficients to vary. The results are shown in Table 6.8. The Residual Standard Error (RSE) increases for all algorithms, and the adjusted R^2 decreases for all but the most simple algorithm. Analysis using scatter plot matrices (Section 4.5, not shown) reveals that after reversing the RF scaling, the Spearman coefficient of specific humidity q, temperature $tm1$, and solar irradiance $zi0$ increases. The coefficient for the absolute geopotential $geopot$ drops to almost zero. The variance of the CCF data itself increases by reversing the RF scaling which is why the absolute accuracy of the univariate algorithm in Table 6.8 decreases while the adjusted R^2 increases. If the maximum adjusted R^2 is calculated for an algorithm including all predictors after reversing the RF scaling as in Section 6.6.4, it drops slightly to 0.55.

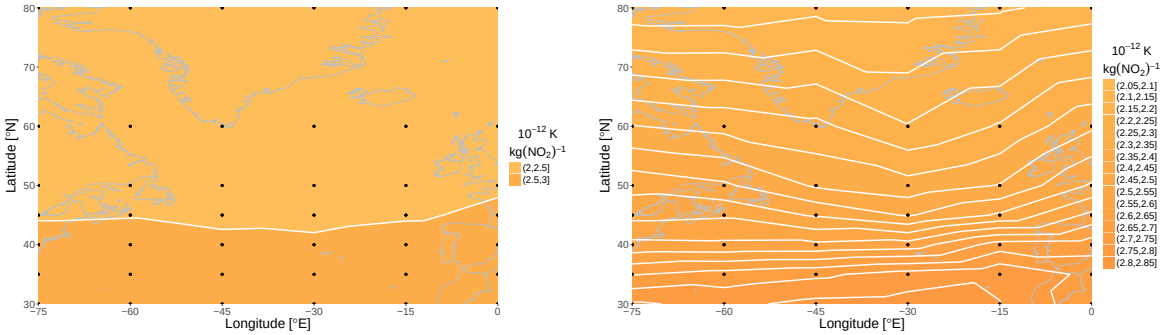
Neglecting the effect of altitude on the ratio between instantaneous and adjusted RF or inputting instantaneous RF into the climate metrics is incorrect. The fact that there is almost no altitude trend in Figure 6.17 (left), means that the shape of the fit in Figure 3.7 (right) largely determines the altitude trend of the ozone CCF data. Emissions at 200 and 400 hPa are within the two regions of the fit based on previous results, so regardless of what the fit should look like between 220–330 hPa, the underlying data determines that NO_x emissions at 200 hPa will have higher ozone climate impact than emissions at 400 hPa. It is unlikely that further simulations would show a local maximum between 220–330 hPa, and a (milder) local minimum around 250–300 hPa is expected. Thus the Spearman coefficient for geopotential would also be non-zero, and the final bivariate algorithm's accuracy will not decrease as much as in Table 6.8. Further discussions on the validity of this RF scaling are given in Section 7.1.4.

6.6.6. Final algorithm

The final bivariate algorithm for ozone CCF approximation explains around 40% of the variance (adjusted R^2 of 0.42) of the data, while a maximum of around 60% (adjusted R^2 of 0.57) is expected. However, Section 6.6.4 showed that with only weather variables the maximum attainable adjusted R^2 is also around 0.42. Chemical concentration data thus explain some of the variability in the CCF data that weather data do not, but in analysing algorithms with up to four predictors the concentration data were weaker predictors than specific humidity, temperature, geopotential and solar irradiance. If algorithms with more predictors are desired, chemical predictors should be added to the regression.

Section 6.4.2 discussed that the accuracy can be improved by adding pressure level or season to the regression but this is problematic for external veracity of the results and the improvement does not warrant the complexity it introduces. There are only four pressure levels in the data, and subtracting a constant from winter results complicates application to spring or autumn weather.

Figure 6.18 shows bilinearly interpolated contour plots of the ozone bilinear algorithm results for 200 hPa, WP1 to compare to Figures 1.1 and 6.3. The intervals and colours used for the CCF data in the left plot are identical to those in Figure 6.3 to facilitate comparison. The same results are shown in the right plot with intervals one order of magnitude smaller, as it was clear from Figure 6.14 that the algorithm has less variability than the original CCF data.



(a) Intervals of 0.5 × 10^{-13} K kg$(NO_2)^{-1}$ (b) Intervals of 0.05 × 10^{-13} K kg$(NO_2)^{-1}$

Figure 6.18: Ozone final $tm1 \times geopot$ algorithmic CCF bilinear contour plots with varying interval sizes, 12 UTC, 200 hPa, WP1. The black points are the emissions locations i.e. the actual data that are interpolated.

Reducing the interval size of the contours does not reveal significantly more detail. It is clear that some of the details in Figure 6.3 are not reproduced by the algorithm. Verification of this algorithm will depend on the level of detail of the ATM optimisation and on the CCFs of the other emissions taken into account. If the routing would lead an aircraft into the local minima in Figure 6.3, the algorithm will lead to false results. If the routing would only avoid the global maxima at low latitudes in Figure 6.3, the algorithm will lead to correct results.

Appendix C (right side) contains these visualisations (with 0.5 × 10^{-13} K kg$(NO_2)^{-1}$) intervals) for all 32 combinations of altitude and weather pattern, for comparison to the original CCF data on the left side.

6.7. Methane algorithm candidates

Scatterplot matrices and regression techniques were used to identify the strongest (combinations of) predictors from the 30 variables in Table 6.1 plus dimensions and derived variables. Due to time constraints for detailed analysis and space constraints for presenting results here, four algorithm candidates are discussed in detail.

A subjective selection was made from preliminary results based on the highest adjusted R^2 per amount of predictors in the regression. A constant fit is presented based on the low variability in Figures 6.4 and 6.7. Based on the trend in Figure 6.7 an extra bivariate option is presented that relates methane climate impact to ozone climate impact. The specific requirement devised for a methane algorithm is that the results should not become positive, as positive methane climate impact is plausible but outside the domain of the dataset used here.

6.7.1. Constant

Figure 6.19: Box plot of methane CCF data versus emission pressure, with the constant regression found (i.e. the mean) shown as a red line.

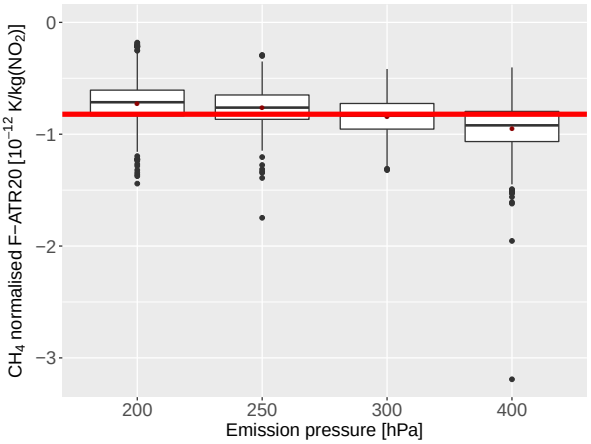


Figure 6.19 shows a box plot of the methane CCF data with the optimum constant regression for this

data as a red line as this dimension has the highest variability. The regression equation corresponds to the mean of the data:

$$\text{F-ATR}\,20_{\text{CH}_4} = \beta_0 + \epsilon \tag{6.9}$$

The parameter values from least squares optimisation are summarised in Table 6.9.

Table 6.9: Parameter overview for methane constant algorithm. The p-value for β_0 is beyond double-precision floating point range.

	Units	Estimate	Std. Error	t value	$Pr\,(X > \lvert t \rvert)$
β_0	K kg$(NO_2)^{-1}$	-8.21×10^{-13}	6.36×10^{-15}	-1.29×10^{2}	$< 2.23 \times 10^{-308}$

The reason for including a constant algorithm is that the variability of the methane CCF data itself is roughly three times smaller than the error (ϵ) from the most sophisticated ozone algorithm. Since both impacts result from the same NO_x emission, the variability of the methane impact can be said to be negligible compared to ozone, as seen in Figure 6.7. Early regression attempts revealed that the maximum attainable adjusted R^2 for methane algorithms (Section 6.9.3) is significantly lower than for water vapour and ozone.

This is an approximation that is not supported by the data in Figure 6.4 and the literature discussed in Section 6.1.2. Removing the outlier discussed in Section 6.1.2 negligibly changes the outcome. Addition of dimensions to the regression defeats the purpose of assuming constant CCF values.

6.7.2. Linear dependence on geopotential

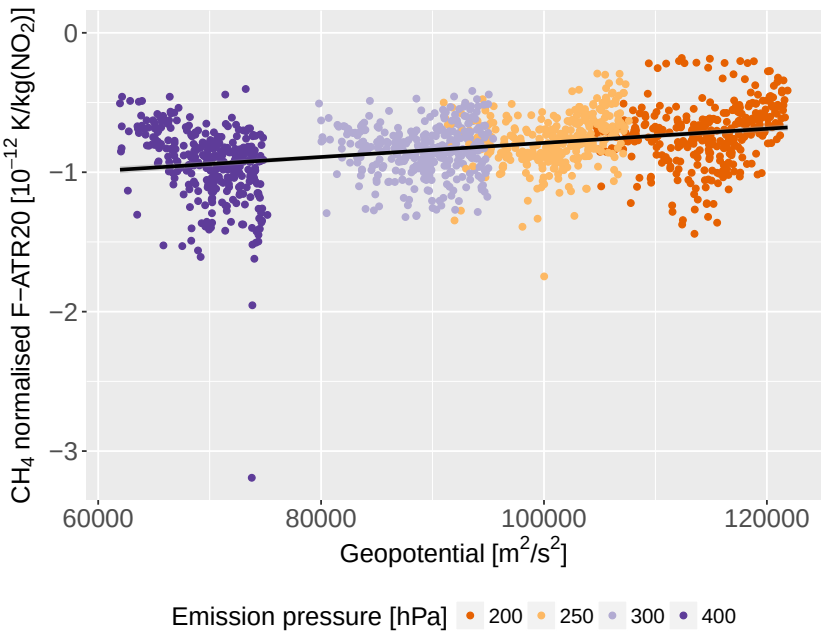


Figure 6.20: Scatterplot of methane CCF data versus geopotential coloured by level, with the linear regression found shown as a black line with 95% confidence interval for the regression coefficients.

Figure 6.20 shows the methane CCF data versus the geopotential at the point of emission. The overall pattern is weakly linearly increasing, which leads to the following fit:

$$\text{F-ATR}\,20_{\text{CH}_4} = \beta_0 + \beta_1 \times geopot + \epsilon \tag{6.10}$$

The parameter values from least squares optimisation are summarised in Table 6.10.

The tight confidence bounds in Figure 6.20 reveal that though much variation is left around the regression line, the overall trend is highly certain. Due to the correspondence of geopotential and pressure height the results are similar for a simple regression on to pressure. Geopotential is preferred because of the certainty provided by the range of data compared to only four pressure levels. In Section 6.1.2 the variability of the methane CCF data was shown to be strongest along the pressure

Table 6.10: Parameter overview for methane linear geopotential algorithm.

	Units	Estimate	Std. Error	t value	$Pr\,(X > \lvert t \rvert)$
β_0	K kg(NO_2)$^{-1}$	-1.30×10^{-12}	3.37×10^{-14}	-3.85×10^{1}	5.59×10^{-219}
β_1	K kg(NO_2)$^{-1}$ (m^2 s^{-2})$^{-1}$	5.06×10^{-18}	3.54×10^{-19}	1.43×10^{1}	2.60×10^{-43}

altitude dimension. The altitude of emission is expected to correlate to the climate impact through determining the destination of a tracer after emission and thus the background chemical concentrations, solar radiation, temperature etc. it encounters.

Addition of an interaction with latitude improves accuracy of the results by decreasing the absolute methane CCF with latitude and decreasing the positive effect of increasing geopotential with latitude. Addition of seasonality improves accuracy, by decreasing the absolute methane CCF result and the positive effect of geopotential for emissions in winter. Regression with only geopotential as a predictor does not explain all of the seasonal or meridional variability.

The intercept (β_0) is sensible as it gives plausible methane CCF values even at zero geopotential. Negative absolute geopotential is not possible and values approaching zero are not expected in the application of this algorithm, as the geopotential is measured from sea-level upwards. Application of this algorithm below 400 hPa would lead to lower geopotential values than included in this dataset, and thus lower methane CCF results than supported by this dataset. As geopotential decreases with latitude for constant pressure altitude, application above 80°N would lead to similar results as below 400 hPa but this is not a likely application. Application at lower latitudes or higher altitudes than the REACT4C data will produce lower absolute methane CCF results.

6.7.3. Bilinear dependence on geopotential and solar irradiance

Figure 6.21 shows the methane CCF data versus the geopotential and solar irradiance at the point of emission. The overall pattern is weakly linearly increasing with geopotential and decreasing with solar irradiance, which leads to the following fit:

$$\text{F-ATR}\,20_{\text{CH}_4} = \beta_0 + \beta_1 \times geopot + \beta_2 \times zi0 + \beta_3 \times geopot \times zi0 + \epsilon \tag{6.11}$$

The parameter values from least squares optimisation are summarised in Table 6.11.

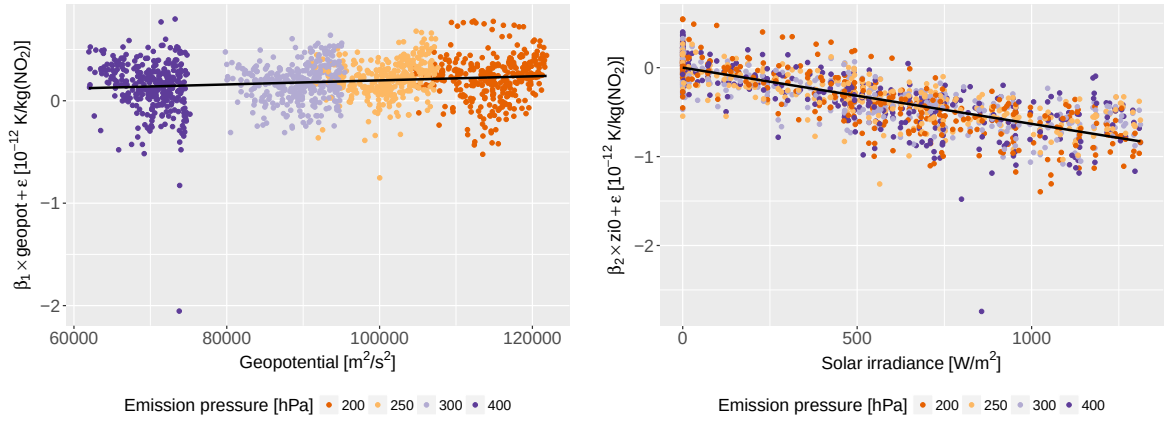
Table 6.11: Parameter overview for methane bilinear geopotential and solar irradiance algorithm.

	Units	Estimate	Std. Error	t value	$Pr\,(X > \lvert t \rvert)$
β_0	K kg(NO_2)$^{-1}$	-9.83×10^{-13}	5.41×10^{-14}	-1.82×10^{1}	3.18×10^{-66}
β_1	K kg(NO_2)$^{-1}$ (m^2 s^{-2})$^{-1}$	1.99×10^{-18}	5.85×10^{-19}	3.40×10^{0}	6.91×10^{-4}
β_2	K kg(NO_2)$^{-1}$ (W m^{-2})$^{-1}$	-6.32×10^{-16}	8.25×10^{-17}	-7.66×10^{0}	3.59×10^{-14}
β_3	K kg(NO_2)$^{-1}$ (W m^{-2} m^2 s^{-2})$^{-1}$	6.12×10^{-21}	8.68×10^{-22}	7.05×10^{0}	2.88×10^{-12}

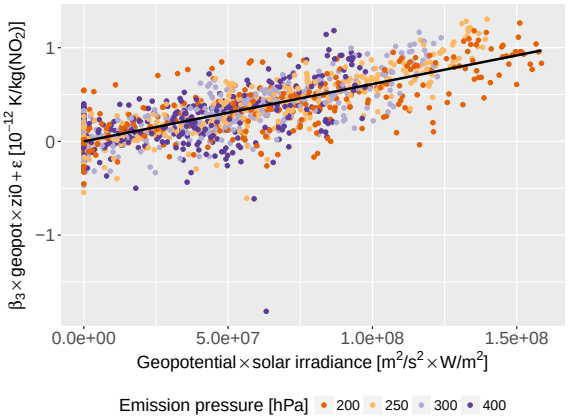
Solar radiation was identified as a driver in literature (Section 6.1. Stevenson et al. (2004) state that solar availability is overriding in terms of the seasonal trend in OH anomalies. Solar irradiance is the incoming solar radiation in [W m^{-2}] at the top of the atmosphere.

Addition of dimensions to the regression directly adds no statistically significant accuracy, i.e. the p-values of the parameters become insignificant though the adjusted R^2 improves. Pressure level, latitude, and season all improve accuracy if an interaction between the dimension and the interaction term of geopotential and solar irradiance is added to the regression. Addition of the pressure interaction adds a significant amount of accuracy, of the latitude interaction adds some accuracy, and of the seasonal term adds a negligible amount of accuracy. A complex relationship between geopotential, solar irradiance, and pressure altitude appears to be omitted in this algorithm.

The intercept is sensible as it gives plausible methane CCF values for zero geopotential and zero solar irradiance. The solar irradiance data used here covers the range of variability for Earth so no issues are expected in terms of external application. Assuming the maximum solar irradiance of around 1350 W m^{-2}, this algorithm produces a methane CCF value of zero for a geopotential of around 180000 m^2 s^{-2} or around 18 km altitude. Application of algorithmic CCFs is not recommended at this altitude, but just in case a ceiling should be added to the algorithm at zero.



(a) $geopot$ (b) $zi0$



(c) $geopot \times zi0$

Figure 6.21: Component plus residual scatter plots for the bivariate regression in Equation (6.11) and parameters in Table 6.11 coloured by emission pressure. Each x-axis is simply the values of the predictor, while each y-axis is the sum of the total residuals ϵ from the regression and the product of the regression coefficient β_i and the data on the x-axis. The black line shows the product of the regression coefficient β_i and the data on the x-axis, so without the residuals.

6.7.4. Bilinear dependence on ozone CCF and potential temperature

Figure 6.22 shows the methane CCF data versus the ozone CCF and potential temperature at the point of emission. The overall pattern is weakly linearly increasing, which leads to the following fit:

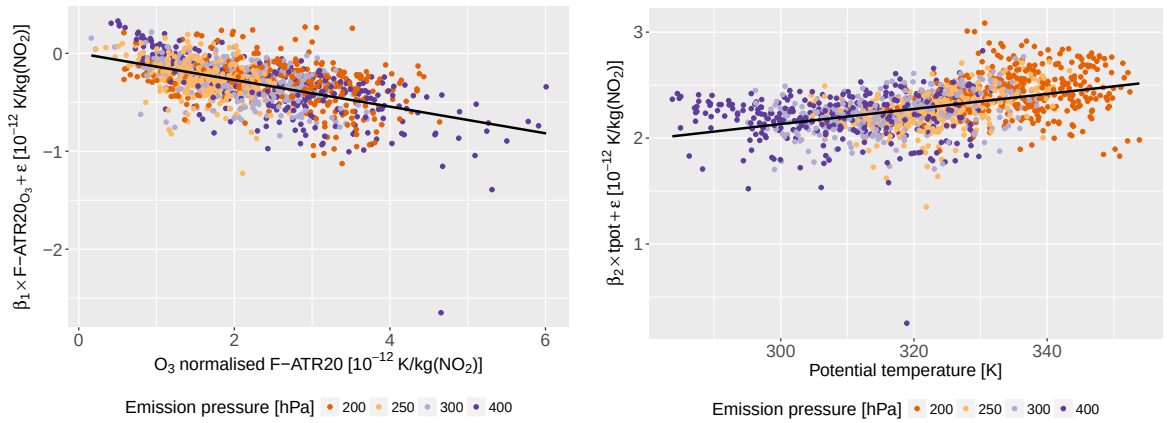
$$\text{F-ATR}\,20_{\text{CH}_4} = \beta_0 + \beta_1 \times \text{F-ATR}\,20_{\text{O}_3} + \beta_2 \times tpot + \epsilon \tag{6.12}$$

The parameter values from least squares optimisation are summarised in Table 6.12.

Table 6.12: Parameter overview for methane bilinear ozone and potential temperature algorithm.

	Units	Estimate	Std. Error	t value	$Pr\,(X > \lvert t \rvert)$
β_0	K kg(NO$_2$)$^{-1}$	-2.81×10^{-12}	1.31×10^{-13}	-2.15×10^{1}	3.31×10^{-88}
β_1	-	-1.36×10^{-1}	6.37×10^{-3}	-2.14×10^{1}	1.96×10^{-87}
β_2	K kg(NO$_2$)$^{-1}$ K^{-1}	7.11×10^{-15}	4.19×10^{-16}	1.69×10^{1}	1.57×10^{-58}

Potential temperature ($tpot$) of an air parcel is the temperature if the parcel would be adiabatically raised to 1000 hPa pressure, generally used because it increases with altitude unlike the actual temperature. It increases in summer, decreases with latitude, and increases with altitude. The actual temperature feeds into the chemical reactions in Section 3.4, but given the instantaneous scope of the regression analysis in this thesis is expected to act as a weather proxy. The potential temperature (for four fixed pressure levels) contains similar information but with a different altitude trend.



(a) $\text{F-ATR}\,20_{O_3}$ (b) $tpot$

Figure 6.22: Component plus residual scatter plots for the bivariate regression in Equation (6.12) and parameters in Table 6.12 coloured by emission pressure. Each x-axis is simply the values of the predictor, while each y-axis is the sum of the total residuals ϵ from the regression and the product of the regression coefficient β_i and the data on the x-axis. The black line shows the product of the regression coefficient β_i and the data on the x-axis, so without the residuals.

The variable $O3cost$ is the ozone CCF as presented in Section 6.1.1 and correlated to methane in Section 6.1.3. Its power as a predictor stems from both CCFs originating from the same NO_x emission and having a complex dependence on the lifetime and destination of the emission, and on being the only variable in the dataset that contains information beyond the time of emission. The highest adjusted R^2 of the ozone algorithms presented here is 0.45, which means that a large error is propagated into the methane results if they are calculated from the ozone algorithm results. It is plausible, if unlikely, that the algorithm could be applied to a dataset where the ozone CCF is simulated and the methane CCF is derived. A potential benefit of this algorithm is in reproducing the correlation between methane and ozone impacts from a NO_x emission, as seen in Figure 6.7.

Addition of latitude greatly improves accuracy through an interaction with potential temperature. This implies that the meridional trend is not properly modelled by these two predictors. Addition of pressure level to the regression greatly improves accuracy through an interaction with the ozone CCF data, but this requires removing the ozone CCF itself from the regression to keep statistical significance. This means that the relationship between the ozone CCF and altitude is a stronger predictor than the ozone CCF data itself.

The intercept is sensible, and the parameter for $O3cost$ compares well to the results in Section 6.1.3. The positive relationship with potential temperature leads to a positive trend with altitude and actual temperature, which is sensible given the altitude trend in the data and the relationship between temperature and reaction rates. Assuming a relatively high value of 350 K for the potential temperature, the algorithm produces zero methane CCF values only if provided with negative ozone CCF values so it is safe for external application.

6.8. Methane algorithm trade-off

Based on the algorithm requirements in Section 4.4, a set of trade-off criteria were devised in Section 4.6. Table 6.13 provides an overview of the results in this section.

Table 6.13: Overview of results of the subjective methane algorithm trade-off, per algorithm and category. If three or more results are distinct, they are ranked into +, ± and −. If only two distinct categories are identified, only + and − are used. *: strongly depends on application.

Algorithm	Pattern	Accuracy	External veracity	Background	Parsimony
$constant$	−	−	−	−	+
$geopot$	−	±	+	±	+
$geopot \times zi0$	±	±	±	±	−
$\text{F-ATR}\,20_{O_3} + tpot$	+	+/±*	−	+	±

6.8.1. Pattern of residuals

The mean values of the spread of residuals per pressure level and weather pattern are:

- $constant$: 8.10×10^{-13}
- $geopot$: 8.14×10^{-13}
- $geopot \times zi0$: 8.08×10^{-13}
- $\text{F-ATR}\,20_{\text{O}_3} + tpot$: 7.6×10^{-13}

Box plots of the residuals per dimension (not shown) reveal that only the $\text{F-ATR}\,20_{\text{O}_3} + tpot$ algorithm captures the complex weather pattern trend in Figure 6.4. The geopotential algorithm has an unexplained trend versus latitude. The $\text{F-ATR}\,20_{\text{O}_3} + tpot$ algorithm has an unexplained non-linear trend versus pressure level, likely caused by the behaviour of the ozone CCF data with pressure level (Figure 6.1). The constant algorithm of course contains no trends.

The 2-D autocorrelation test (Moran's I) is inconclusive. The amount of level/WP sets that are statistically significant varies per algorithm, but none give a negative I.

Overall the constant and geopotential algorithms perform poorly, the $\text{F-ATR}\,20_{\text{O}_3} + tpot$ algorithm performs well and the $geopot \times zi0$ algorithm presents a moderate option.

6.8.2. General accuracy

The RSE (and adjusted R^2) for each algorithm within this dataset are:

- $constant$: $2.33 \times 10^{-13} (0.00)$
- $geopot$: $2.17 \times 10^{-13} (0.13)$
- $geopot \times zi0$: $2.12 \times 10^{-13} (0.17)$
- $\text{F-ATR}\,20_{\text{O}_3} + tpot$: $1.96 \times 10^{-13} (0.29)$

The trend of improving accuracy is similar to that of the spread of the reisduals per pressure level and weather pattern. If the ozone bivariate algorithm (Section 6.4.2) results are used instead of simulated ozone CCF data, the RSE for the $\text{F-ATR}\,20_{\text{O}_3} + tpot$ algorithm becomes 2.17×10^{-13} K kg$(NO_2)^{-1}$, and the adjusted R^2 becomes 0.13 so it is as accurate as the $geopot$ algorithm.

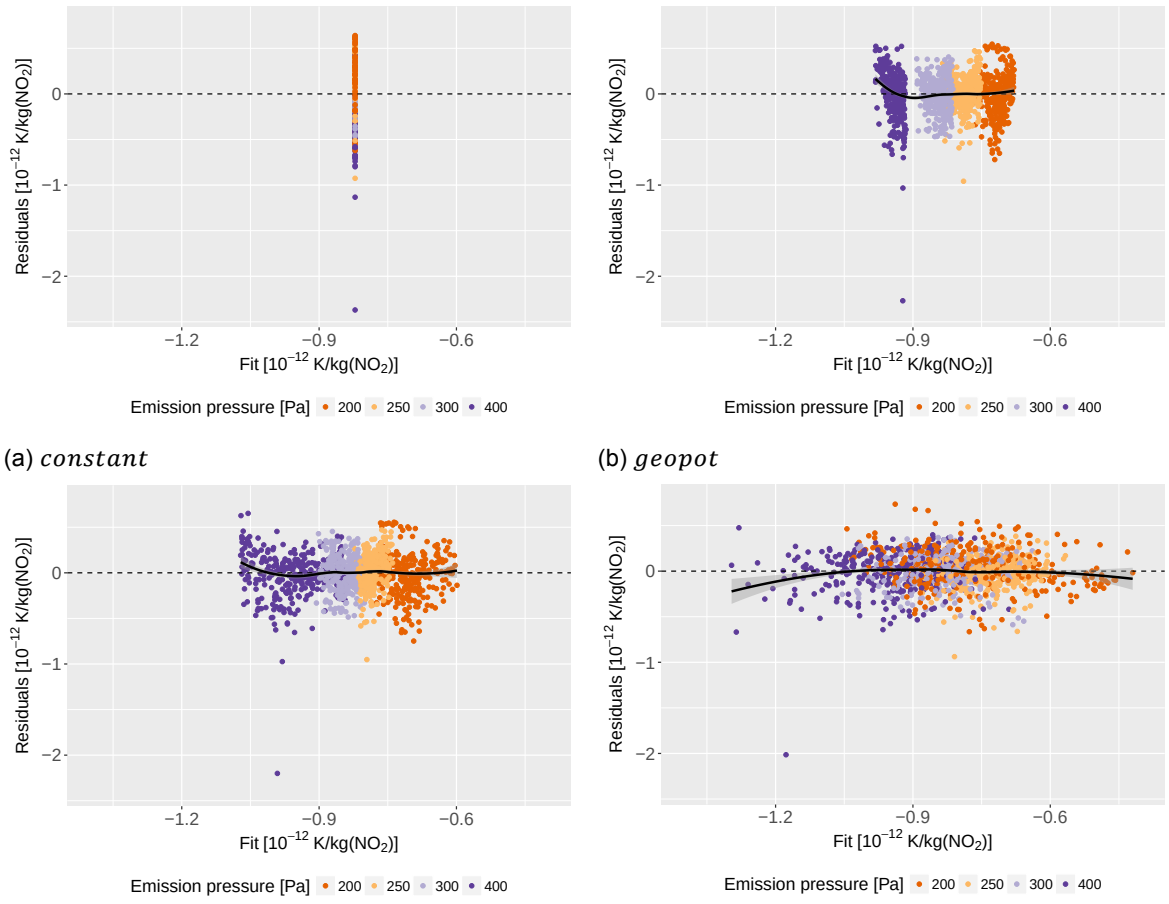
Figure 6.23 shows residual plots for the four algorithms. The results for the linear $geopot$ algorithm show some remaining trend at the ends due to the presence of a trend between geopotential and methane CCF data in each distinct pressure level. The two bilinear options show reasonable results, with little trend left. There is no visible heteroskedasticity.

If simulated ozone CCF data is used, the $\text{F-ATR}\,20_{\text{O}_3} + tpot$ algorithm gives the best results of the four candidates. If not, it is comparable to the $geopot$ algorithm in terms of accuracy and the $geopot \times zi0$ algorithm is best.

6.8.3. External veracity

Results here are summarised from Section 6.7.

- $constant$: literature showed clear spatial and seasonal trends, so a constant fit is unlikely to be correct. If there is no spatial variability of the methane CCF, climate-optimised ATM for this species becomes pointless. This option is likely to be more useful for general estimates of aviation climate impact than for re-routing.
- $geopot$: geopotential depends on pressure level, latitude, and the weather situation. The trend is relatively weak so there should not be problems applying the algorithm to other locations. Similar to the constant option, it is unlikely that the regional variation is properly captured, though for the same pressure level geopotential does decrease with latitude.
- $geopot \times zi0$: as $zi0$ is top-of-atmosphere incoming radiation and this dataset contains the full range of possible values, few problems are expected with external application. If $zi0$ is zero, this reduces to roughly the univariate $geopot$ algorithm. A ceiling at zero methane CCF is recommended.



(a) $constant$ (b) $geopot$

(c) $geopot \times zi0$ (d) $\text{F-ATR}\,20_{O_3} + tpot$

Figure 6.23: Scatterplots of residuals versus fit for the four methane algorithms. Residuals are original values minus fitted values. The black curves are LOESS regressions to identify the trend of the residuals over the algorithm values, with 95% confidence bands.

- $\text{F-ATR}\,20_{O_3} + tpot$: external veracity of this algorithm will depend on the ozone CCF input data used. It is not certain that the relationship between methane and ozone CCF in this dataset (Figure 6.7) will hold for different geographical domains, and there is some uncertainty regarding the accuracy of the altitudinal trend of the ozone CCF data used (Section 6.6.5). If algorithmic ozone CCF results are used as inputs, external veracity depends on the veracity of the ozone algorithm.

The $\text{F-ATR}\,20_{O_3} + tpot$ and constant algorithms are uncertain in terms of results produced outside of the REACT4C domain. The other two algorithms are more certain because of the stable and relatively weak relationship with geopotential. The $geopot \times zi0$ scores slightly lower because of the need for an extra calculation step to ensure negative methane CCF results.

6.8.4. Mechanistic background

Results here are summarised from Section 6.7.

- $constant$: the methane climate impact being constant disagrees with the REACT4C data (Figure 6.4) and literature (Section 6.1).
- $geopot$: all else being equal, higher altitude release of a NO_x emission leads to a different destination of the NO_x and ozone tracer as precursor for methane depletion, and some difference in chemistry at the time of release which is expected to be negligible.
- $geopot \times zi0$: similar to the $geopot$ algorithm, and in addition the solar irradiance may contain

information about reaction rates during the first hours after emission. It is likely to act as a proxy for latitude and season, which then influence the destination of a tracer after emission.

- $\text{F-ATR}\,20_{\text{O}_3} + tpot$: the (simulated) ozone CCF contains information about the chemical regime encountered by a NO_x emission and the tracer destination, which helps to determine the methane F-ATR20. The potential temperature contains actual temperature (thus reaction rates), altitude, and some location data which mainly determine the destination of the tracer.

The ozone and potential temperature algorithm has the best theoretical background, while the constant algorithm is clearly an approximation. The two geopotential algorithms are moderately well-understood, as the instantaneous character of the data used in this thesis does not allow confirmation of the causal relationship behind a correlation.

6.8.5. Parsimony

The less terms included in an algorithm, the more simple it is. Interaction terms are deemed less complex than an entirely separate variable for this thesis, especially for bivariate options where the interaction term does not affect interactive 3-D plotting for presentations. The constant option is clearly the most simple one, but this is not deemed a significant advantage over univariate geopotential. The most complex option presented here for methane is $geopot \times zi0$, which is still an acceptable level of complexity.

6.8.6. Choice

The ozone and potential temperature algorithm is clearly the most accurate one if simulated data is used, but it scores poorly on external veracity and for this trade-off it is assumed that algorithmic ozone CCF inputs are used. The bivariate geopotential and solar irradiance candidate presents marginal improvements in terms of accuracy compared to the constant and univariate geopotential options, but is within acceptable limits in terms of complexity and is chosen as the final algorithm based on the scores in Table 6.13. With the assumption of algorithmic ozone CCF inputs it is the most accurate option. The final form of the algorithmic CCF for methane is:

$$\begin{aligned} \text{aC}\widetilde{\text{CF}}_{\text{CH}_4} = -9.83 \times 10^{-13} + 1.99 \times 10^{-18} \times geopot - 6.32 \times 10^{-16} \times zi0 \\ +6.12 \times 10^{-21} \times geopot \times zi0 \approx \text{F-ATR}\,20_{\text{CH}_4} \end{aligned} \tag{6.13}$$

$$\begin{aligned} \text{aCCF}_{\text{CH}_4} &= \text{aC}\widetilde{\text{CF}}_{\text{CH}_4} \quad \text{for} \quad \text{aC}\widetilde{\text{CF}}_{\text{CH}_4} \leq 0 \\ \text{aCCF}_{\text{CH}_4} &= 0 \quad \text{for} \quad 0 < \text{aC}\widetilde{\text{CF}}_{\text{CH}_4} \end{aligned} \tag{6.14}$$

Where the extra step is added to ensure no positive aCCF results can occur.

Figure 6.24 shows a comparison between the original CCF data and the final algorithm (aCCF) results, where points on the black line indicate perfect agreement. It is clear that the variability of the final algorithm results is lower than that of the original data. The algorithm produces a clearer altitude trend than the original data, which may lead to incorrect vertical re-routing if applied for climate-optimised ATM.

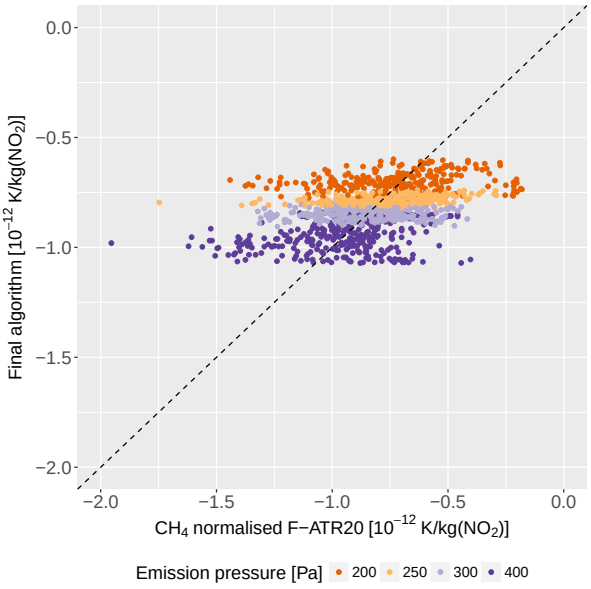
6.9. Methane discussion

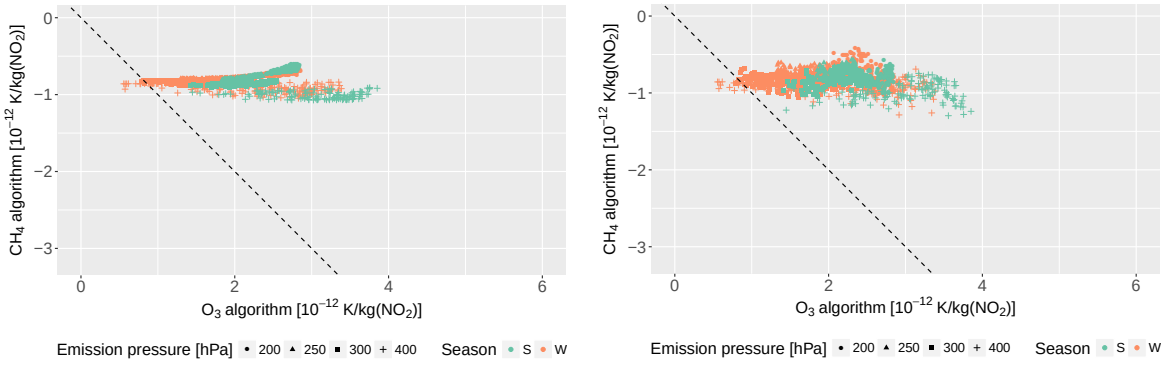
From the results in the previous sections and others not shown in this thesis, a number of discussions and conclusions arise. Here discussions on various topics are given as far as they relate directly to the results of the methane CCF data, its relation to ozone CCF data, the weather data, and the algorithm formulation. General discussions regarding the methodology are given in Chapter 7.

6.9.1. Outlier

None of the algorithms presented here, nor regressions with any of the other variables from Table 6.1 were able to explain the outlier in Figure 6.6 (right).

Figure 6.24: Scatterplot of methane algorithmic CCF results versus REACT4C data coloured by level, with the black line indicating perfect agreement. The outlier is outside of the range of the plot.





(a) $geopot \times zi0$

(b) $\text{F-ATR}\,20_{\text{O}_3} + tpot$, simulated ozone CCF

Figure 6.25: Scatterplots of methane algorithmic CCF results versus ozone algorithmic CCF results coloured by emission season, with the dashed line representing net zero cumulative effect. The left plot shows the final methane algorithm, the right plot shows the algorithm based on ozone CCF data using the simulated data. Note that PMO has been omitted.

6.9.2. Relationship with the ozone CCF

Figure 6.25 contains two scatter plots of the methane algorithm results versus the ozone algorithm results, for comparison to Figure 6.7. The left plots shows results from the final algorithm chosen in Section 6.8, while the right plot shows results from the $O3cost + tpot$ algorithm in the best case where the ozone CCF data is simulated.

Whereas the original data in Figure 6.7 show a clear decreasing trend of the methane CCF for increasing ozone CCF, here results for methane appear almost constant with some non-linear increase versus ozone at the lowest absolute values for methane. This coincides with 200 hPa release pressure. Both the ozone and methane algorithms chosen here linearly depend on geopotential. The non-linearity at 200 hPa is likely to be caused by the temperature versus geopotential correlation around the tropopause (Section 6.4.2). If the algorithmic methane results are regressed onto the ozone results, the RSE becomes lower than in Section 6.1.3, the negative trend becomes stronger, and the adjusted R^2 becomes almost zero.

Figure 6.25 (right) shows that the best-case alternative methane algorithm produces noisier results, which visually compare better to Figure 6.7. If there is a meaningful application of algorithmic NO_x CCFs where the ozone CCF can be simulated and the methane CCF is algorithmically estimated, the algorithm choice in Section 6.8 should be revisited.

6.9.3. Maximum attainable R^2

When all the selected variables from Table 6.1 are combined in a linear regression model with longitude, latitude, level, season, and the latitude of the jet stream, the resulting adjusted R^2 is 0.37 for 36 variables. Introducing interaction terms of course increases the accuracy, but first-order interactions already require 666 parameters so the result is meaningless. Though imperfect, this gives an estimate of the amount of variation in the CCF data that can be explained by the combination of weather data, chemical data, and lightning data. If only the 19 weather variables are used, adjusted R^2 is 0.26, if only 13 chemistry variables are used, adjusted R^2 is 0.30, if only four dimensions are used, adjusted R^2 is 0.14.

6.9.4. Ensemble-mean algorithm

Application of the mean of an ensemble of simulations is commonly used in meteorology to mitigate the effect of chaos on the results of a prediction. When the ensemble-mean is calculated from the four algorithms presented in Section 6.7, the RSE is 2.06×10^{-13} K kg$(NO_2)^{-1}$. The (adjusted) R^2 is 0.22. This is an intermediate result between the two bilinear algorithms presented here (Section 6.8.2). For the $\mathrm{F\text{-}ATR}\,20_{O_3} + tpot$ algorithm, simulated ozone CCF inputs were assumed.

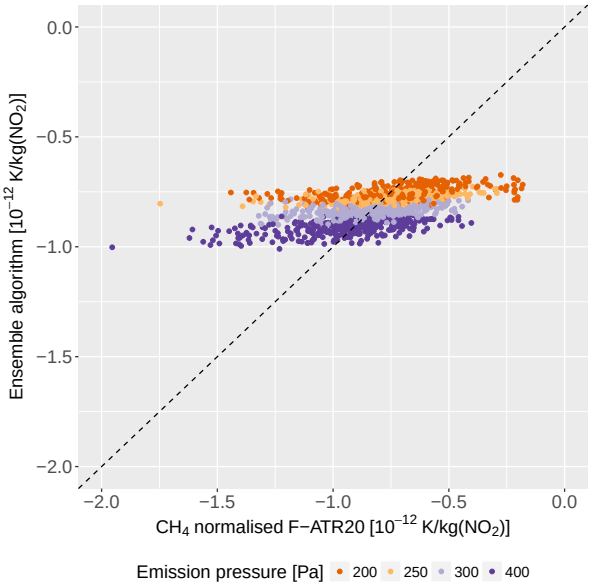


Figure 6.26: Scatterplot of methane ensemble (mean of four) algorithmic CCF results versus REACT4C data coloured by level, with the black line indicating perfect agreement. The outlier is outside of the range of the plot.

Figure 6.26 shows the resulting ensemble-mean algorithm results versus the original REACT4C data, for comparison to Figure 6.24. The results are not very distinctive from those of the geopotential with solar irradiance algorithm. Because of the addition of the (simulated) ozone and potential temperature algorith, the exaggerated altitude trend is somewhat reduced and the trend in the data is closer to the line of perfect agreement. Application of the ensemble-mean to these four algorithms does not appear to meaningfully improve the accuracy for the complexity it introduces.

6.9.5. Final algorithm

The final bivariate algorithm for methane CCF approximation explains around 20% of the variance (adjusted R^2 of 0.17) of the data, while a maximum of around 40% (adjusted R^2 of 0.37) is expected. However, Section 6.9.3 showed that with only weather variables the maximum attainable adjusted R^2 is around 0.26. Chemical concentration data thus explain some of the variability in the CCF data that weather data do not, but in analysing algorithms with up to four predictors the concentration data were weaker predictors than specific humidity, temperature, geopotential, solar irradiance and the ozone CCF data. If algorithms with more predictors are desired, chemical predictors should be added to the regression.

Figure 6.27 shows bilinearly interpolated contour plots of the methane bilinear algorithm results for 200 hPa, WP1 to compare to Figures 1.1 and 6.6 (left). The amount of contours and hence variability in the plot is the same as for the original data in Figure 6.6 (left), so reduced interval sizes are not

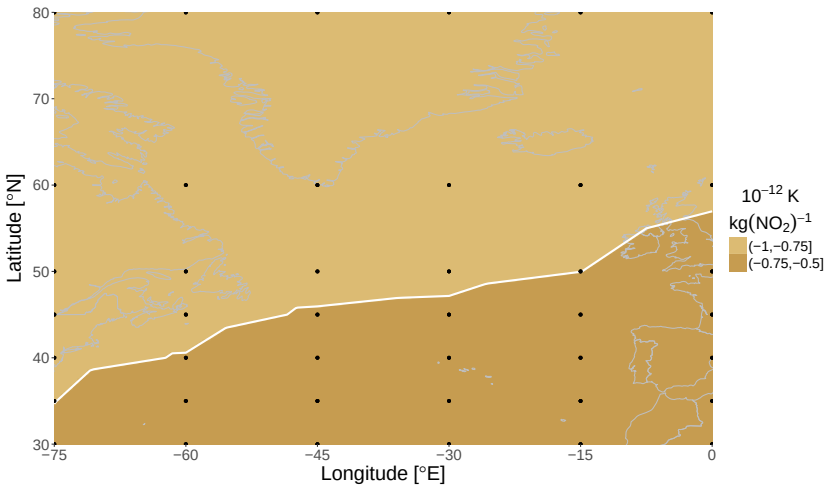


Figure 6.27: Methane final $geopot \times zi0$ algorithmic CCF bilinear contour plot, 12 UTC, 200 hPa, WP1. The black points are the emissions locations i.e. the actual data that are interpolated.

investigated. The influence of the solar irradiance on the methane aCCF results is clear from the shape of the contour. The pattern of the original CCF data is not reproduced at all, which agrees with the adjusted R^2 of 0.17. Depending on the other algorithms applied and the sensitivity of the ATM optimisation, it is likely that application of this methane algorithm will lead to some amount of incorrect re-routing.

Appendix D (right side) contains these visualisations for all 32 combinations of altitude and weather pattern, for comparison to the original CCF data on the left side.

7 Discussion

This chapter contains systematic discussions not pertaining to results for water vapour and NO_x, as these were given in Sections 5.5, 6.6 and 6.9 respectively. The structure is roughly chronological from the various steps of the REACT4C methodology, then the methodology of this thesis, then verification and validation of results, and finally potential for future developments in this area. Each of these topics is an input into the next topic, and potential issues with e.g. an assumption in REACT4C influence the validity of results of this thesis and hence the potential for application of these results. A qualitative assessment is given of the uncertainty added by each step between basic theory and algorithmic CCFs, and of the potential for improvements.

7.1. REACT4C methodology

7.1.1. Background model

There is fundamental uncertainty in all climate modelling, as models are not readily validated. Exponentially increasing computational power over the past decades has led to greatly enhanced model resolution, and scientific research has led to improved parametrisations of sub-grid processes. Some of the uncertainty in the REACT4C CCF data will be caused by the EMAC model, but the underlying ECHAM5 model has been verified and included in numerous model comparisons. Grewe et al. (2017a) used atmospheric measurements of ozone, NO, and NO_y to compare simulations in EMAC at $2.8° \times 2.8°$ resolution to the small-scale model Consortium of Small-scale MOdels (COSMO) at $0.44° \times 0.44°$, finding that differences in accuracy depend on the size of meteorological patterns at a time-of-day.

7.1.2. Weather patterns and geographical domain

The eight weather patterns selected by Irvine et al. (2013) represent a valid method for reducing the amount of simulations needed to simulate climate-optimised ATM in the NAFC. It has been demonstrated that these eight weather patterns cannot fully represent variability across all four seasons (Section 3.1). Within the NAFC, using results from these eight patterns to interpolate algorithmic CCFs adds significant uncertainty, as it is not known if the CCF results for these eight patterns sufficiently represent all yearly variability of the CCF results and how much bias is introduced by using unequal amounts of summer and winter patterns. Potentially, the method of determining weather patterns by NAO and EA index does not represent year-round CCF variability at all. Irvine et al. (2013) did use climate impact proxies to demonstrate that the eight patterns give distinguishable results for these proxies. Verification of the seasonal and weather pattern variability could be performed by simulating another day that has the same NAO and EA indices as the representative day used to generate one of the eight patterns, and analysing the intra-simulation variability of the CCF results.

The REACT4C methodology (Grewe et al., 2014b) was set up to use the NAFC as a test case. Using the results to derive algorithmic CCFs for external application adds great uncertainty. An effort was made in this thesis to argue which limitations will occur in predicting the CCFs in other geographic domains, but regression analysis and its results only give information about internal veracity of the relationship found. The algorithms in this thesis may be the result of fitting noise, or a physically valid

relationship found may show significant regional variations. The uncertainty caused by this limited simulation domain could be mitigated by applying the REACT4C methodology to other regions.

7.1.3. Tagging approach

The assumptions applied for application of a tagging approach in REACT4C add some inaccuracy compared to perturbation studies, as acknowledged by Grewe et al. (2014b). The emissions are not added to background concentrations so non-linearity in model physics and chemistry is not represented, and Lagrangian transport has various benefits and drawbacks compared to Eulerian schemes. The simplified physics and chemistry modelled within the Lagrangian tracer compared to the background (hydrological) physics and chemistry is not expected to introduce significant uncertainty. What is taken as a single longitude-latitude-pressure location for an CCF value for the purpose of this thesis is the sum of 50 smaller emissions that were randomly distributed in the EMAC gridbox. This adds uncertainty to the results that is expected to be relatively small. The sensitivity studies in Grewe et al. (2014b) (Section 3.5) present strong evidence for using 50 tracers per location, but significant variation if the time or space grid is refined. Thus the time-region grid specification adds significant uncertainty to the results. This casts into doubt the decision to simulate only one time coordinate for all weather patterns.

7.1.4. Radiative forcing parametrisations

The RF calculation for water vapour (Section3.6.1) is based on a strong correlation from Grewe and Stenke (2008) and adds a small amount of uncertainty to the water vapour results.

The instantaneous RF calculation for ozone within EMAC is not line-by-line but still expected to be accurate. The adjusted RF seasonal relationship for ozone (Section3.6.2) is based on a limited set of results from one publication, and adds a small amount of uncertainty to the ozone results. The adjusted RF altitude relationship for ozone is based on two previous publications and a complicated non-linear fit is presented through eight data points (Figure 3.7 right, one point is out of range), with a minimum at 250 hPa while no points are located between about 220–330 hPa. In Section 6.6.5 a sensitivity analysis was presented from entirely reversing the altitude relationship. After reversal, almost no altitude trend remains in the ozone CCF data and the accuracy of the algorithm options changes. The relationship with background NO_x better compares to Stevenson and Derwent (2009). Ignoring the difference between instantaneous and adjusted RF would be incorrect and a more accurate altitude scaling is expected to fall somewhere between that applied in REACT4C and the reversal applied here for sensitivity analysis. A minimum around 250–300 hPa is still expected. All algorithm candidates for ozone give a minimum at 300 hPa instead of 250 hPa for the original CCF data (not shown), so development of a more accurate altitude scaling may increase the accuracy of an algorithmic ozone CCF.

The minimum ozone CCF occurs for 250 hPa emissions for every weather pattern in this dataset, but the trend varies from pattern to pattern. In terms of the resulting flight routes, in the ISA 250 hPa should be converted to a flight altitude of 34,000 ft (FL340), whereas results in Grewe et al. (2014a) show very little altitude change for low optimisation and a shift to FL300 for full optimisation, which should mean 300 hPa. The uncertain effect on the CCF data may thus have negligible effect on the climate-optimised routing.

The methane RF relationship from the IPCC First Assessment Report is based on relatively old model results and a paper which was not peer-reviewed and is not publicly available at all (Section 3.6.3). This relationship adds significant uncertainty to the methane results, but as the relationship is proportional to the methane concentration the spatial error is likely to be small. However, many comparable studies, including Hoor et al. (2009) and Myhre et al. (2011), calculate methane RF from a further simplification of the IPCC formula.

7.1.5. Climate metrics

The REACT4C approach of simulating 90 days and then extrapolating climate metrics adds uncertainty to the results, but the chosen pulse extrapolation methods are grounded in previous research and the same method is applied to all locations so the spatial error is likely to be small. Considering the approximations used to calculate RF, the worst case is that the algorithms formulated in this thesis represent concentration changes more than they do climate impact. The calculation of F-ATR20 was not detailed in Grewe et al. (2014b) and a full derivation is beyond the scope of this thesis.

7.1.6. Verification

The limited verification activities in Grewe et al. (2014b) (Section 3.8) demonstrate that the general set-up of the REACT4C is correct and that the uncertainties introduced do not cause overall incorrect trends, but not whether the spatial pattern of the results is correct. No prior work exists on the variability of the climate impact of aircraft emissions in the NAFC or a domain of comparable size.

7.2. Thesis methodology

7.2.1. Interpolation

The approach of reverse engineering CCF data to their original grid and interpolating weather data to the same grid is warranted for the regression approach taken in this thesis, but adds some uncertainty to the weather data. More sophisticated interpolation methods should produce more accurate data but this effect is expected to be negligible compared to other uncertainties.

7.2.2. NO_x split-up

The investigation of the relationship between ozone and methane CCF data and the algorithm candidate relating methane to ozone with an adjusted R^2 of 0.29 show that combining ozone and methane in one algorithm is non-trivial and the split-up is warranted. It is not expected that more advanced methods would lead to a different understanding, given the different reaction pathways and atmospheric lifetimes.

For NO_x analysis an option is multivariate regression analysis, where multiple independent variables are regressed concurrently on the same dependent variables. This may aid in clarifying whether one or two algorithms should be developed for NO_x CCFs, but was not pursued in detail for this research.

7.2.3. Regression analysis

For the time, effort, and level of knowledge expected in an MSc thesis, the decision to focus on 0-D regression analysis is warranted. The first algorithm candidate investigated in detail was the non-linear quartic relative tropopause relationship which served as a case study for the drawbacks of non-linear regression without a strong observed relationship that is expected from prior literature, which is why linear regression was applied for all other candidates.

The case study for four 250 hPa datasets (Section 6.6.2) showed that visual patterns do not necessarily represent actual correlation between parameters and gave an estimate of the adjusted R^2 achievable for an ozone weather-based algorithm. Experimentation with including all selected independent variables in the regression (Sections 5.5.2, 6.6.4 and 6.9.3) gave estimates of best-case adjusted R^2 for each species. In each case the achieved results for the final algorithm was around 0.2 lower. Reasons for selecting only certain data for analysis and minimising the amount of predictors in the final algorithms were argued, and given the methodology followed, it is not expected that better results could be achieved without a significantly more complex approach. The best-case adjusted R^2 experiments can be interpreted as an upper limit for even the most sophisticated approach, which would likely extrapolate poorly and thus have little scientific value.

For all three species, the residuals Moran's I test in the algorithm trade-off produced varying amounts of statistically significant results for which the I value was always positive. In this methodology, positive I is deemed acceptable. The weighting used here was inverse Cartesian distance, and a different weighting would likely produce very different results of the test. Significant spatial autocorrelation in residuals either points to spatial autocorrelation in the independent variable that is not adequately modelled, or spurious spatial autocorrelation from the dependent variables. All algorithms presented here showed improved accuracy when season, latitude, or level were added to the regression, thus spatial autocorrelation in the CCF data is sensible. Given that no meaningful variation is expected with longitude and the tests were performed per longitude-latitude field, positive results from Moran's I for residuals suggest that the meridional variability of the CCFs is not adequately modelled. Moran's I tests on the CCF data are not presented here, but given the trends shown in the CCF box plots and the contour plots in the appendix, are very likely to show positive spatial autocorrelation through meridional variability. Literature, specifically in the field of ecology, (e.g. Dormann et al., 2007) suggests that in the presence of autocorrelation in the original data and in the residuals, spatial regression methods can produce more accurate results. This was observed relatively early on during research but spatial regression represents unnecessary complexity for an MSc thesis. The field is not as established and is focused around ecology, so little directly useful literature exists. There are also risks in terms of re-

sults being more difficult to interpret directly and internal veracity being improved at the cost of external veracity. If spatial regression is applied to this dataset in the future, the decision to interpolate all data to the CCF grid should be revisited.

7.3. Reflecting upon published research

Altitude behaviour for water vapour showed stronger agreement with Fichter (2009) than with Wilcox et al. (2012b). A caveat is that the results in REACT4C are based on a correlation from Grewe and Stenke (2008), upon which the methodology of Fichter (2009) is based. If there are unknown errors present in the Grewe and Stenke (2008) results, these errors are likely to propagate into the REACT4C results. If there are unknown errors in the Grewe and Stenke (2008) methodology, they are likely to be present in both the Fichter (2009) results and the REACT4C results. The linear altitude relationship shown in Wilcox et al. (2012b) is based on far fewer simulations than in AirClim (Grewe and Stenke, 2008; Fichter, 2009) or REACT4C (Grewe et al., 2014b), and the non-linear relationship with emission altitude and tropopause altitude shared by Fichter (2009) and this thesis is likely to be more accurate.

The comparison to exponential fits from Stevenson and Derwent (2009) for background NO_x concentrations provides an interesting reflection on the validity of their findings. Results for summer emissions agree well with their findings, while winter results do not agree at all. This comparison is not conclusive, however, given the suspected error in 250 hPa ozone CCF data. The comparison showed 250–400 hPa summer data in agreement with Stevenson and Derwent (2009) with 200 hPa deviating, but if 250 hPa results are underestimated in REACT4C the exponential relationship will not compare equally well. The difference between summer and winter will remain.

7.4. Verification and validation

The internal veracity of the presented algorithms across the instantaneous NAFC dataset is discussed Sections 5.5, 6.6, and 6.9 respectively. Here options for external verification and validation are discussed.

7.4.1. REACT4C dataset

For WP1 (Grewe et al., 2014b) two extra time coordinates were simulated. These results are not presented here, and not included in the main methodology. Preliminary analysis shows that the data do not vary as much within WP1 as they do between weather patterns, but to a similar degree as the algorithm results vary from the original data. Within six-hour time coordinates, changes in large-scale weather and hence the destination of an emitted tracer are relatively small. Considering the low variability with local time and longitude discussed in Sections 5.1 and 6.1, these differences are either caused by differences in weather or by the inherent randomness in this set-up. If the former is true, simulation of all three time coordinates for all weather patterns would have strengthened the algorithmic relationships investigated. If the latter is true, simulation of three time coordinates would have allowed investigation into how much of the variation can be explained by weather data. One option for this analysis was to test the algorithms on this extra data. As the data varies relatively little from 12 UTC and only represents WP1, results from testing the algorithms can not easily be extrapolated.

One option for this research was to split up the REACT4C data into training and testing data for algorithms. This option was neglected due to an assumption that the set of eight weather patterns is the minimum amount necessary to cover the variability of the CCFs and weather over the North Atlantic, thus using a subset for analysis would create bias through neglected variability and using the remainder for testing the results would give incorrectly low results. This route could be used to mitigate the bias introduced by an unequal number of summer/winter patterns by training the algorithms on two or three patterns per season, and using the remainder for testing. This can be combined with the 6 and 18 UTC approach. Another option that was outside the scope of this thesis is to assign relative weights to the weather and CCF data of each weather pattern based on the relative yearly frequency of each weather pattern as given by Irvine et al. (2013).

In future work, the 90-day time history per REACT4C emission can also be used to analyse what the variables in the algorithms presented here are acting as a proxy for, e.g. whether temperature correlates well to ozone CCF due to reaction rates or due to representing the meridional destination of the emission and corresponding chemistry. This would strengthen the causal relationship of the algorithms presented here.

7.4.2. WeCare dataset

Grewe et al. (2017a) present results from the DLR project WeCare, that used the REACT4C methodology for an extra day with a larger geographical extent and higher resolution. This dataset was not available on time for verification within this thesis but presents a valuable opportunity. If the data from this set that match the REACT4C extent are extracted and interpolated to the same grid, this dataset could be used to verify the weather pattern split-up approach by using synoptic typing to determine the pattern's location on the NAO and EA axis and comparing the CCF results to the closest REACT4C weather pattern.

7.4.3. AirTraf implementation

To test external veracity of the algorithm results, new simulations must be run to generate test data which is far beyond the scope of this thesis. As this thesis is framed within the Air Traffic Management for the Environment (ATM4E) scope (see Preface), external verification activities will be undertaken within the project by implementing the final algorithms in the new EMAC submodel AirTraf (Yamashita et al., 2016). A new test case will be simulated, where routes are optimised for minimal combined algorithmic CCF results. Concurrently, the ATM approach of the project will be verified. Details of the verification activities are not yet confirmed and are outside the scope of this thesis.

7.4.4. Validation

Beyond using historical data to nudge climate models, validation is not a trivial concept in climate science due to the timescales and lack of a control group. The input data from REACT4C were not validated, and even to validate only the atmospheric processes would require extensive satellite or aerial measurement campaigns without a control group to provide certainty.

7.5. Future development

Indirectly, this research gives some insight into which future developments could improve calculations of CCFs, beyond e.g. increased resolution of the simulation.

7.5.1. Chemical weather forecasting

For this research the focus was on parameters than can be implemented in NWP, but for scientific value chemistry and lightning data were included. CWF is a promising new field of research that combines NWP with chemical reactions and transport, but research has focussed on surface-level air pollution (Kukkonen et al., 2012). One caveat may be a lack of measured data at aviation cruise level to initialise the model, but satellite imagery (Miyazaki et al., 2012) and in-service aircraft measurements (Marenco et al., 1998; Clark et al., 2015) have potential. There is thus some potential for future real-time NO_x CCF calculation based on chemical variables, but this research showed no improvement in NO_x regression results compared to weather data. It is not expected that the development of CWF in combination with 0-D instantaneous algorithmic CCFs will improve upon the results from this thesis.

7.5.2. Medium-range weather forecasting

For this research the scope of application was now-casted NWP, thus instantaneous algorithms were researched that can be applied in a model that interpolates weather data based on measurement data. The experiments on maximum attainable R^2 per species confirmed that only limited information about the climate impact of an emission is available in instantaneous weather data. Weather models are known to be limited to about 7–14 day predictions due to chaotic effects, but there is potential for algorithms to be developed based on the first 7–14 days of weather data per REACT4C pattern. In Figure 3.8, the largest NO_x decrease and corresponding ozone peak and methane loss often occure around 15 days after emission. As this is at the uppermost range of NWP accuracy, a compromise could be made by using the first 7 days. Investigations into the 90-day time history of the REACT4C trajectories and local weather can demonstrate how well various timeframes of weather data can represent the concentration changes and thus climate impact (as intermediate steps are mostly linearised), and verify whether 7-day algorithms have more potential than instantaneous algorithms. This approach can be combined with CWF implementation if non-instantaneous chemistry data better predict CCFs than weather data do.

8

Conclusion and recommendations

8.1. Conclusion

The research objective of this MSc thesis was to formulate algorithmic approximations of aviation water vapour and NO_x emission climate impact based on local weather data by systematically analysing variations in climate impact and local weather data across eight simulated datasets for distinct weather patterns over the North Atlantic from the REACT4C project. The methodology selected is zero-dimensional instantaneous regression analysis on the full set of eight weather patterns. All data were interpolated to the original emissions grid to prevent inflation of statistical power. Regression variables were selected based on literature reviews, four algorithms were formulated per species, and a trade-off was made based on detailed requirements for algorithmic CCFs.

Water vapour climate impact is best explained by variables that represent the distinction between tropospheric and stratospheric emissions. The final algorithm (Equation (5.6)) is linear with PV and the underlying regression has an adjusted R^2 of 0.59. The most relevant previous publication found a linear relationship between RF and altitude, whereas this analysis shows a strongly non-linear relationship even if accounting for the tropopause altitude. The variability of water vapour climate impact increases strongly with altitude, which complicates accurate prediction of the impact of emissions near the tropopause.

Ozone climate impact is best explained by humidity, temperature, geopotential or solar irradiance. The final algorithm (Equation (6.7)) is linear with temperature, geopotential and their interaction and the underlying regression has an adjusted R^2 of 0.42. The most relevant previous publication found an exponentially decreasing relationship between ozone RF and background NO_x concentrations plus a linear decrease for increasing latitude. Analysis here confirms that this relationship only holds for emissions between 250 and 400 hPa in summer, with a background NO_x mixing ratio below 200 pptv. A potential flaw in the altitude scaling of the underlying REACT4C data was identified that risks invalidating the results for ozone. A sensitivity analysis was done to show the potential effect an entirely incorrect altitude scaling would have on the algorithms and on the background NO_x relationship.

Methane climate impact is best explained by geopotential, solar irradiance, potential temperature or the ozone climate impact. The final algorithm (Equation (6.14)) is linear with geopotential, solar irradiance and their interaction and the underlying regression has an adjusted R^2 of 0.17. The most relevant previous publication found an exponentially increasing relationship between methane RF and background NO_x concentrations. Analysis here confirms that this relationship only holds for emissions between 250 and 400 hPa in summer, with a background NO_x mixing ratio below 200 pptv. Analysis of the variation of methane and ozone climate impact for the same NO_x emission shows that using simulated ozone climate impact data to predict methane climate impact produces better results than using weather data. Using algorithmic CCF data as inputs produces worse results than weather data.

From the results of the algorithm formulation and experiments for maximum adjusted R^2 presented here, it is clear that water vapour, ozone, and methane normalised climate impact variations are decreasingly driven by the instantaneous weather situation in that order. This is sensible when considering the respective atmospheric lifetimes. These experiments show that chemical concentration data explain a part of the ozone and methane variability that weather data do not, but for the algorithms

presented here with relatively few predictors the chemistry data is less useful than weather data. A case study was done for ozone CCF data to demonstrate that a 2-D visual correlation to weather data does not necessarily apply to the underlying data.

The algorithms presented here are planned to be verified within the EMAC sub-model AirTraf, together with concurrent work on contrail algorithmic CCFs. Not all details on the procedure are known, and among others the accuracy of these algorithms will depend on the geographical domain and season chosen. The adjusted R^2 within the training data is assumed to be a reasonable estimate of external accuracy, and thus especially the methane algorithm appears unlikely to produce results accurate enough for application. All algorithms show improvement if geographical data are added, suggesting that spatial regression methods could find better algorithms. Using several days of weather data may improve predictions of tracer destination over instantaneous relationships suggested here.

8.2. Recommendations without new simulations

- Use 90-day REACT4C data to analyse how accurate algorithmic CCFs can be, and formulate 7-day algorithms for use with 7-day NWP or CWF.
- Re-size and interpolate the WeCare dataset, assess position on teleconnection pattern axes, and analyse variability of CCF results compared to most similar pattern(s).
- Test algorithms presented here on the WeCare dataset.
- Apply spatial regression techniques to this REACT4C dataset to verify whether zero-dimensional analysis is adequate.
- Re-consider choice of PMO scaling factor from methane and consider adding a Stratospheric Water Vapour (SWV) scaling factor from methane.
- Re-assess validity of the IPCC methane RF approximation and the impact on all REACT4C results including those presented here.
- Re-assess validity of the ozone instantaneous/adjusted RF altitude scaling approximation and the impact on all REACT4C results including those presented here, specifically for emissions at 250 and 300 hPa by addition of extra data from previous or new simulations.

8.3. Recommendations for new simulations

- Simulate higher altitude water vapour CCFs to allow an assessment of the nature of the increasing variability with altitude.
- Use a higher resolution longitude/latitude/level grid in future CCF simulations to allow more detailed analysis of the spatial variability of climate impact of aviation emissions.
- Use a regularly-spaced longitude/latitude/level grid and equal amounts of patterns per season in future CCF simulations if more algorithmic research is expected to be performed on the results, to prevent biases.
- Simulate CCFs for spring and/or autumn weather patterns for the REACT4C domain and verify that the winter/summer pattern typing approach represents yearly variability.
- Simulate CCFs for other geographical domains and seasons and apply methodology from this thesis to verify that zero-dimensional instantaneous algorithms are globally valid.

8.4. Final note

Much uncertainty remains, but this research represents another step in analysing and predicting variations in the climate impact of aviation emissions and in enabling mitigation strategies that do not require technological improvements.

Bibliography

Appleman, H.: The formation of exhaust condensation trails by jet aircraft, Bulletin of the American Meteorological Society, 34, 14–20, 1953.

Berntsen, T. K. and Isaksen, I. S. A.: Effects of lightning and convection on changes in tropospheric ozone due to NO_x emissions from aircraft, Tellus B, 51, 766–788, doi:10.1034/j.1600-0889.1999.t01-3-00003.x, 1999.

Berntsen, T. K., Gauss, M., Isaksen, I. S. A., Grewe, V., Sausen, R., Pitari, G., Mancini, E., Meijer, E., and Hauglustaine, D.: Source of NO_x at cruise altitudes: Implications for predictions of ozone and methane perturbations due to NO_x from aircraft, in: European Conference on Aviation, Atmosphere and Climate (AAC), edited by Sausen, R., Fichter, C., and Amanatidis, G., EUR 21051, pp. 190–196, European Commission, Friedrichshafen, Germany, URL http://www.pa.op.dlr.de/aac/proceedings/AAC-proceedings-chemistry.pdf, 2003.

Berntsen, T. K., Fuglestvedt, J. S., Joshi, M. M., Shine, K. P., Stuber, N., Ponater, M., Sausen, R., Hauglustaine, D. A., and Li, L.: Response of climate to regional emissions of ozone precursors: sensitivities and warming potentials, Tellus B, 57, 283–304, doi:10.1111/j.1600-0889.2005.00152.x, 2005.

Bock, L. and Burkhardt, U.: Reassessing properties and radiative forcing of contrail cirrus using a climate model, Journal of Geophysical Research: Atmospheres, 121, 9717–9736, doi:10.1002/2016JD025112, 2016.

Boucher, O. and Reddy, M.: Climate trade-off between black carbon and carbon dioxide emissions, Energy Policy, 36, 193–200, doi:10.1016/j.enpol.2007.08.039, 2008.

Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., Kerminen, V.-M., Kondo, Y., Liao, H., Lohmann, U., Rasch, P., Satheesh, S. K., Sherwood, S., Stevens, B., and Zhang, X.: Clouds and Aerosols, in: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., chap. 7, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Brasseur, G. P., Cox, R. A., Hauglustaine, D., Isaksen, I., Lelieveld, J., Lister, D. H., Sausen, R., Schumann, U., Wahner, A., and Wiesen, P.: European scientific assessment of the atmospheric effects of aircraft emissions, Atmospheric Environment, 32, 2329–2418, doi:10.1016/S1352-2310(97)00486-X, 1998.

Burkhardt, U. and Kärcher, B.: Global radiative forcing from contrail cirrus, Nature Climate Change, 1, 54–58, doi:10.1038/nclimate1068, 2011.

Clark, H., Sauvage, B., Thouret, V., Nédélec, P., Blot, R., Wang, K.-Y., Smit, H., Neis, P., Petzold, A., Athier, G., Boulanger, D., Cousin, J.-M., Beswick, K., Gallagher, M., Baumgardner, D., Kaiser, J., Flaud, J.-M., Wahner, A., Volz-Thomas, A., and Cammas, J.-P.: The first regular measurements of ozone, carbon monoxide and water vapour in the Pacific UTLS by IAGOS, Tellus B, 67, doi:10.3402/tellusb.v67.28385, 2015.

Cook, J., Nuccitelli, D., Green, S. A., Richardson, M., Winkler, B., Painting, R., Way, R., Jacobs, P., and Skuce, A.: Quantifying the consensus on anthropogenic global warming in the scientific literature, Environmental Research Letters, 8, 024 024, doi:10.1088/1748-9326/8/2/024024, 2013.

Dahlmann, K.: Eine Methode zur effizienten Bewertung von Massnahmen zur Klimaoptimierung des Luftverkehrs, Ph.D. thesis, Ludwig-Maximilians-Universität München, 2012.

Dahlmann, K., Grewe, V., Frömming, C., and Burkhardt, U.: Can we reliably assess climate mitigation options for air traffic scenarios despite large uncertainties in atmospheric processes?, Transportation Research Part D: Transport and Environment, 46, 40–55, doi:10.1016/j.trd.2016.03.006, 2016.

Davis, C. A. and Emanuel, K. A.: Potential Vorticity Diagnostics of Cyclogenesis, Monthly Weather Review, 119, 1929–1953, doi:10.1175/1520-0493(1991)119<1929:PVDOC>2.0.CO;2, 1991.

Dormann, C. F., McPherson, J. M., Araújo, M. B., Bivand, R., Bolliger, J., Carl, G., Davies, R. G., Hirzel, A., Jetz, W., Daniel Kissling, W., Kühn, I., Ohlemüller, R., Peres-Neto, P. R., Reineking, B., Schröder, B., M. Schurr, F., and Wilson, R.: Methods to account for spatial autocorrelation in the analysis of species distributional data: a review, Ecography, 30, 609–628, doi:10.1111/j.2007.0906-7590.05171.x, 2007.

EPA: Evaluation of air pollutant emissions from subsonic commercial jet aircraft, Tech. Rep. EPA420-R-99-013, United States Environmental Protection Agency, 1999.

Fichter, C.: Climate impact of air traffic emissions in dependency of the emission location and altitude, Ph.D. thesis, Manchester Metropolitan University, 2009.

Forster, P., Ramaswamy, V., Artaxo, P., Berntsen, T., Betts, R., Fahey, D. W., Haywood, J., Lean, J., Lowe, D. C., Myhre, G., Nganga, J., Prinn, R., Raga, G., Schulz, M., and Van Dorland, R.: Changes in Atmospheric Constituents and in Radiative Forcing, in: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, chap. 2, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2007.

Forster, P. M., Fomichev, V. I., Rozanov, E., Cagnazzo, C., Jonsson, A. I., Langematz, U., Fomin, B., Iacono, M. J., Mayer, B., Mlawer, E., Myhre, G., Portmann, R. W., Akiyoshi, H., Falaleeva, V., Gillett, N., Karpechko, A., Li, J., Lemennais, P., Morgenstern, O., Oberländer, S., Sigmond, M., and Shibata, K.: Evaluation of radiation scheme performance within chemistry climate models, Journal of Geophysical Research: Atmospheres, 116, doi:10.1029/2010JD015361, d10302, 2011.

Fouquart, Y. and Bonnel, B.: Computations of solar heating of the earth's atmosphere - A new parameterization, Beiträge zur Physik der Atmosphäre, 53, 35–62, 1980.

Frömming, C., Grewe, V., Matthes, S., Brinkop, S., Haslerud, A. S., Irvine, E. A., Rosanka, S., and van Manen, J.: Influence of weather situation on aviation emission effects: The REACT4C Climate Change Functions, Atmospheric Environment, in preparation, 2017.

Fuglestvedt, J., Berntsen, T., Myhre, G., Rypdal, K., and Skeie, R. B.: Climate forcing from the transport sectors, Proceedings of the National Academy of Sciences, 105, 454–458, doi:10.1073/pnas.0702958104, 2008.

Fuglestvedt, J. S., Berntsen, T. K., Isaksen, I. S., Mao, H., Liang, X.-Z., and Wang, W.-C.: Climatic forcing of nitrogen oxides through changes in tropospheric ozone and methane; global 3D model studies, Atmospheric Environment, 33, 961–977, doi:10.1016/S1352-2310(98)00217-9, 1999.

Fuglestvedt, J. S., Shine, K. P., Berntsen, T., Cook, J., Lee, D. S., Stenke, A., Skeie, R. B., Velders, G. J. M., and Waitz, I. A.: Transport impacts on atmosphere and climate: Metrics, Atmospheric Environment, 44, 4648–4677, doi:10.1016/j.atmosenv.2009.04.044, transport Impacts on Atmosphere and Climate: The ATTICA Assessment Report, 2010.

Gauss, M., Isaksen, I. S. A., Wong, S., and Wang, W.-C.: Impact of H_2O emissions from cryoplanes and kerosene aircraft on the atmosphere, Journal of Geophysical Research: Atmospheres, 108, doi:10.1029/2002JD002623, 4304, 2003.

Gauss, M., Isaksen, I. S. A., Lee, D. S., and Søvde, O. A.: Impact of aircraft NO_x emissions on the atmosphere - tradeoffs to reduce the impact, Atmospheric Chemistry and Physics, 6, 1529–1548, doi:10.5194/acp-6-1529-2006, 2006.

Gilmore, C. K., Barrett, S. R. H., Koo, J., and Wang, Q.: Temporal and spatial variability in the aviation NO_x-related O_3 impact, Environmental Research Letters, 8, 034 027, doi:10.1088/1748-9326/8/3/034027, 2013.

Grewe, V.: A generalized tagging method, Geoscientific Model Development, 6, 247–253, doi:10.5194/gmd-6-247-2013, 2013.

Grewe, V. and Dahlmann, K.: How ambiguous are climate metrics? And are we prepared to assess and compare the climate impact of new air traffic technologies?, Atmospheric Environment, 106, 373–374, doi:10.1016/j.atmosenv.2015.02.039, 2015.

Grewe, V. and Dameris, M.: Calculating the global mass exchange between stratosphere and troposphere, Annales Geophysicae, 14, 431–442, 1996.

Grewe, V. and Stenke, A.: AirClim: an efficient tool for climate evaluation of aircraft technology, Atmospheric Chemistry and Physics, 8, 4621–4639, doi:10.5194/acp-8-4621-2008, 2008.

Grewe, V., Dameris, M., Fichter, C., and Sausen, R.: Impact of aircraft NO_x emissions. Part 1: Interactively coupled climate-chemistry simulations and sensitivities to climate-chemistry feedback, lightning and model resolution, Meteorologische Zeitschrift, 11, 177–186, doi:10.1127/0941-2948/2002/0011-0177, 2002.

Grewe, V., Tsati, E., and Hoor, P.: On the attribution of contributions of atmospheric trace gases to emissions in atmospheric model applications, Geoscientific Model Development, 3, 487–499, doi:10.5194/gmd-3-487-2010, 2010.

Grewe, V., Dahlmann, K., Matthes, S., and Steinbrecht, W.: Attributing ozone to NO_x emissions: Implications for climate mitigation measures, Atmospheric Environment, 59, 102–107, doi:10.1016/j.atmosenv.2012.05.002, 2012.

Grewe, V., Champougny, T., Matthes, S., Frömming, C., Brinkop, S., Søvde, O. A., Irvine, E. A., and Halscheidt, L.: Reduction of the air traffic's contribution to climate change: A REACT4C case study, Atmospheric Environment, 94, 616–625, doi:10.1016/j.atmosenv.2014.05.059, 2014a.

Grewe, V., Frömming, C., Matthes, S., Brinkop, S., Ponater, M., Dietmüller, S., Jöckel, P., Garny, H., Tsati, E., Dahlmann, K., Søvde, O. A., Fuglestvedt, J., Berntsen, T. K., Shine, K. P., Irvine, E. A., Champougny, T., and Hullah, P.: Aircraft routing with minimal climate impact: the REACT4C climate cost function modelling approach (V1.0), Geoscientific Model Development, 7, 175–201, doi:10.5194/gmd-7-175-2014, 2014b.

Grewe, V., Dahlmann, K., Flink, J., Frömming, C., Ghosh, R., Gierens, K., Hendricks, J., Kaufmann, S., Kölker, K., Linke, F., Luchkova, T., Lührs, B., van Manen, J., Matthes, S., Minikin, A., Niklaß, M., Jöckel, P., Plohr, M., Righi, M., Rosanka, S., Schlage, R., Schmitt, A., Schumann, U., Terekhov, I., Unterstrasser, S., Vázquez-Navarro, M., Voigt, C., Wicke, K., Yamashita, H., Zahn, A., and Ziereis, H.: Mitigating the Climate Impact from Aviation: Achievements and Results of the DLR WeCare Project, Aerospace, in preparation, 2017a.

Grewe, V., Matthes, S., Frömming, C., Brinkop, S., Jöckel, P., Gierens, K., Champougny, T., Fuglestvedt, J., Haslerud, A., Irvine, E., and Shine, K.: Feasibility of climate-optimized air traffic routing for trans-Atlantic flights, Environmental Research Letters, 12, 034 003, doi:10.1088/1748-9326/aa5ba0, 2017b.

Hansen, J., Fung, I., Lacis, A., Rind, D., Lebedeff, S., Ruedy, R., Russell, G., and Stone, P.: Global climate changes as forecast by Goddard Institute for Space Studies 3-dimensional model, Journal of Geophysical Research, 93, 9341–9364, 1988.

Hansen, J., Sato, M., and Ruedy, R.: Radiative forcing and climate response, Journal of Geophysical Research: Atmospheres, 102, 6831–6864, doi:10.1029/96JD03436, 1997.

Holmes, C. D., Tang, Q., and Prather, M. J.: Uncertainties in climate assessment for the case of aviation NO, Proceedings of the National Academy of Sciences, 108, 10 997–11 002, doi:10.1073/pnas.1101458108, 2011.

Hoor, P., Borken-Kleefeld, J., Caro, D., Dessens, O., Endresen, O., Gauss, M., Grewe, V., Hauglustaine, D., Isaksen, I. S. A., Jöckel, P., Lelieveld, J., Myhre, G., Meijer, E., Olivie, D., Prather, M., Schnadt Poberaj, C., Shine, K. P., Staehelin, J., Tang, Q., van Aardenne, J., van Velthoven, P., and Sausen, R.: The impact of traffic emissions on atmospheric ozone and OH: results from QUANTIFY, Atmospheric Chemistry and Physics, 9, 3113–3136, doi:10.5194/acp-9-3113-2009, 2009.

Irvine, E. A., Hoskins, B. J., Shine, K. P., Lunnon, R. W., and Frömming, C.: Characterizing North Atlantic weather patterns for climate-optimal aircraft routing, Meteorological Applications, 20, 80–93, doi:10.1002/met.1291, 2013.

Jaeglé, L., Jacob, D. J., Brune, W. H., Faloona, I. C., Tan, D., Kondo, Y., Sachse, G. W., Anderson, B., Gregory, G. L., Vay, S., Singh, H. B., Blake, D. R., and Shetter, R.: Ozone production in the upper troposphere and the influence of aircraft during SONEX: approach of NO_x-saturated conditions, Geophysical Research Letters, 26, 3081–3084, doi:10.1029/1999GL900451, 1999.

Jöckel, P., Tost, H., Pozzer, A., Brühl, C., Buchholz, J., Ganzeveld, L., Hoor, P., Kerkweg, A., Lawrence, M. G., Sander, R., Steil, B., Stiller, G., Tanarhte, M., Taraborrelli, D., van Aardenne, J., and Lelieveld, J.: The atmospheric chemistry general circulation model ECHAM5/MESSy1: consistent simulation of ozone from the surface to the mesosphere, Atmospheric Chemistry and Physics, 6, 5067–5104, doi:10.5194/acp-6-5067-2006, 2006.

Jöckel, P., Kerkweg, A., Pozzer, A., Sander, R., Tost, H., Riede, H., Baumgaertner, A., Gromov, S., and Kern, B.: Development cycle 2 of the Modular Earth Submodel System (MESSy2), Geoscientific Model Development, 3, 717–752, doi:10.5194/gmd-3-717-2010, 2010.

Jöckel, P., Tost, H., Pozzer, A., Kunze, M., Kirner, O., Brenninkmeijer, C. A. M., Brinkop, S., Cai, D. S., Dyroff, C., Eckstein, J., Frank, F., Garny, H., Gottschaldt, K.-D., Graf, P., Grewe, V., Kerkweg, A., Kern, B., Matthes, S., Mertens, M., Meul, S., Neumaier, M., Nützel, M., Oberländer-Hayn, S., Ruhnke, R., Runde, T., Sander, R., Scharffe, D., and Zahn, A.: Earth System Chemistry integrated Modelling (ESCiMo) with the Modular Earth Submodel System (MESSy) version 2.51, Geoscientific Model Development, 9, 1153–1200, doi:10.5194/gmd-9-1153-2016, 2016.

Köhler, M. O., Rädel, G., Dessens, O., Shine, K. P., Rogers, H. L., Wild, O., and Pyle, J. A.: Impact of perturbations to nitrogen oxide emissions from global aviation, Journal of Geophysical Research: Atmospheres, 113, doi:10.1029/2007JD009140, d11305, 2008.

Köhler, M. O., Rädel, G., Shine, K. P., Rogers, H. L., and Pyle, J. A.: Latitudinal variation of the effect of aviation NO_x emissions on atmospheric ozone and methane and related climate metrics, Atmospheric Environment, 64, 1–9, doi:10.1016/j.atmosenv.2012.09.013, 2013.

Kukkonen, J., Olsson, T., Schultz, D. M., Baklanov, A., Klein, T., Miranda, A. I., Monteiro, A., Hirtl, M., Tarvainen, V., Boy, M., Peuch, V.-H., Poupkou, A., Kioutsioukis, I., Finardi, S., Sofiev, M., Sokhi, R., Lehtinen, K. E. J., Karatzas, K., San José, R., Astitha, M., Kallos, G., Schaap, M., Reimer, E., Jakobs, H., and Eben, K.: A review of operational, regional-scale, chemical weather forecasting models in Europe, Atmospheric Chemistry and Physics, 12, 1–87, doi:10.5194/acp-12-1-2012, 2012.

Lacis, A., Hansen, J., Lee, P., Mitchell, T., and Lebedeff, S.: Greenhouse effect of trace gases, 1970-1980, Geophysical Research Letters, 8, 1035–1038, 1981.

Lee, D. S., Fahey, D. W., Forster, P. M., Newton, P. J., Wit, R. C., Lim, L. L., Owen, B., and Sausen, R.: Aviation and global climate change in the 21st century, Atmospheric Environment, 43, 3520–3537, doi:10.1016/j.atmosenv.2009.04.024, 2009.

Lee, D. S., Pitari, G., Grewe, V., Gierens, K., Penner, J. E., Petzold, A., Prather, M., Schumann, U., Bais, A., Berntsen, T., Iachetti, D., Lim, L. L., and Sausen, R.: Transport impacts on atmosphere and climate: Aviation, Atmospheric Environment, 44, 4678–4734, doi:10.1016/j.atmosenv.2009.06.005, transport Impacts on Atmosphere and Climate: The ATTICA Assessment Report, 2010.

Leggett, J., Pepper, W., Swart, R., Edmonds, J., Meira Filho, L., Mintzer, I., Wang, M., and Watson, J.: Emissions Scenarios for the IPCC: an Update, in: Climate Change 1992: The Supplementary Report to The IPCC Scientific Assessment, chap. A3, pp. 68–95, Cambridge University Press, UK, 1992.

Mannstein, H., Vázquez-Navarro, M., Graf, K., Duda, D. P., and Schumann, U.: Contrail Detection in Satellite Images, pp. 433–447, Springer Berlin Heidelberg, Berlin, Heidelberg, doi:10.1007/978-3-642-30183-4_26, 2012.

Marenco, A., Thouret, V., Nédélec, P., Smit, H., Helten, M., Kley, D., Karcher, F., Simon, P., Law, K., Pyle, J., Poschmann, G., Von Wrede, R., Hume, C., and Cook, T.: Measurement of ozone and water vapor by Airbus in-service aircraft: The MOZAIC airborne program, an overview, Journal of Geophysical Research: Atmospheres, 103, 25 631–25 642, doi:10.1029/98JD00977, 1998.

Matthes, S., Schumann, U., Grewe, V., Froemming, C., Dahlmann, K., Koch, A., and Mannstein, H.: Atmospheric Physics, chap. Climate Optimized Air Transport, pp. 727–746, Springer, 2012.

Miyazaki, K., Eskes, H. J., Sudo, K., Takigawa, M., van Weele, M., and Boersma, K. F.: Simultaneous assimilation of satellite NO_2, O_3, CO, and HNO_3 data for the analysis of tropospheric chemical composition and emissions, Atmospheric Chemistry and Physics, 12, 9545–9579, doi:10.5194/acp-12-9545-2012, 2012.

Mlawer, E. J., Taubman, S. J., Brown, P. D., Iacono, M. J., and Clough, S. A.: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave, Journal of Geophysical Research: Atmospheres, 102, 16 663–16 682, doi:10.1029/97JD00237, 1997.

Moran, P. A. P.: Notes on Continuous Stochastic Phenomena, Biometrika, 37, 17–23, doi:10.2307/2332142, 1950.

Myhre, G., Shine, K., Rädel, G., Gauss, M., Isaksen, I., Tang, Q., Prather, M., Williams, J., van Velthoven, P., Dessens, O., Koffi, B., Szopa, S., Hoor, P., Grewe, V., Borken-Kleefeld, J., Berntsen, T., and Fuglestvedt, J.: Radiative forcing due to changes in ozone and methane caused by the transport sector, Atmospheric Environment, 45, 387–394, doi:10.1016/j.atmosenv.2010.10.001, 2011.

Myhre, G., Shindell, D., Bréon, F.-M., Collins, W., Fuglestvedt, J., Huang, J., Koch, D., Lamarque, J.-F., Lee, D., Mendoza, B., Nakajima, T., Robock, A., Stephens, G., Takemura, T., and Zhang, H.: Anthropogenic and Natural Radiative Forcing, in: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., chap. 8, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Penner, J. E., Lister, D. H., Griggs, D. J., Dokken, D. J., and McFarland, M.: Aviation and the global atmosphere: a special report of IPCC Working Groups I and III in collaboration with the Scientific Assessment Panel to the Montreal Protocol on Substances that Deplete the Ozone Layer, Cambridge University Press, UK, 1999.

Ponater, M., Pechtl, S., Sausen, R., Schumann, U., and Hüttig, G.: Potential of the cryoplane technology to reduce aircraft climate impact: A state-of-the-art assessment, Atmospheric Environment, 40, 6928–6944, doi:10.1016/j.atmosenv.2006.06.036, 2006.

Reithmeier, C. and Sausen, R.: ATTILA: atmospheric tracer transport in a Lagrangian model, Tellus B, 54, 278–299, doi:10.1034/j.1600-0889.2002.01236.x, 2002.

Roeckner, E., Bäuml, G., Bonaventura, L., Brokopf, R., Esch, M., Giorgetta, M., Hagemann, S., Kirchner, I., Kornblueh, L., Manzini, E., Rhodin, A., Schlese, U., Schulzweida, U., and Tompkins, A.: The atmospheric general circulation model ECHAM5. Part I: Model description, Tech. Rep. 354, Max-Planck-Institut für Meteorologie, 2003.

Roeckner, E., Brokopf, R., Esch, M., Giorgetta, M., Hagemann, S., Kornblueh, L., Manzini, E., Schlese, U., and Schulzweida, U.: Sensitivity of Simulated Climate to Horizontal and Vertical Resolution in the ECHAM5 Atmosphere Model, Journal of Climate, 19, 3771–3791, doi:10.1175/JCLI3824.1, 2006.

Sander, R., Baumgaertner, A., Gromov, S., Harder, H., Jöckel, P., Kerkweg, A., Kubistin, D., Regelin, E., Riede, H., Sandu, A., Taraborrelli, D., Tost, H., and Xie, Z.-Q.: The atmospheric chemistry box model CAABA/MECCA-3.0, Geoscientific Model Development, 4, 373–380, doi:10.5194/gmd-4-373-2011, 2011.

Saravanamuttoo, H. I. H., Rogers, G. F. C., Cohen, H., and V.Straznicky, P.: Gas turbine theory, Pearson Education, 6 edn., 2009.

Schmidt, E.: Die Entstehung von Eisnebel aus den Auspuffgasen von Flugmotoren, Schriften der Deutschen Akademie der Luftfahrtforschung, 44, 1–15, 1941.

Schumann, U.: On conditions for contrail formation from aircraft exhausts, Meteorologische Zeitschrift, 5, 4–23, 1996.

Schwartz Dallara, E., Kroo, I. M., and Waitz, I. A.: Metric for Comparing Lifetime Average Climate Impact of Aircraft, AIAA Journal, 49, 1600–1613, doi:10.2514/1.J050763, 2011.

Shine, K. P., Derwent, R. G., Wuebbles, D. J., and Morcrette, J.-J.: Radiative forcing of climate, in: Climate Change: The IPCC Scientific Assessment (1990), Report prepared for Intergovernmental Panel on Climate Change by Working Group I, edited by Houghton, J. T., Jenkins, G. J., and Ephraums, J. J., chap. 2, Cambridge University Press, Cambridge, Great Britain, New York, NY, USA and Melbourne, Australia, 1990.

Shine, K. P., Fuglestvedt, J. S., Hailemariam, K., and Stuber, N.: Alternatives to the Global Warming Potential for Comparing Climate Impacts of Emissions of Greenhouse Gases, Climatic Change, 68, 281–302, doi:10.1007/s10584-005-1146-9, 2005.

Stenke, A., Dameris, M., Grewe, V., and Garny, H.: Implications of Lagrangian transport for simulations with a coupled chemistry-climate model, Atmospheric Chemistry and Physics, 9, 5489–5504, doi:10.5194/acp-9-5489-2009, 2009.

Stevenson, D. S. and Derwent, R. G.: Does the location of aircraft nitrogen oxide emissions affect their climate impact?, Geophysical Research Letters, 36, doi:10.1029/2009GL039422, l17810, 2009.

Stevenson, D. S., Doherty, R. M., Sanderson, M. G., Collins, W. J., Johnson, C. E., and Derwent, R. G.: Radiative forcing from aircraft NO_x emissions: Mechanisms and seasonal dependence, Journal of Geophysical Research: Atmospheres, 109, doi:10.1029/2004JD004759, d17307, 2004.

Stuber, N.: Ursachen der Variabilität des Klimasensitivitätsparameters für räumlich inhomogene Ozonstörungen, Ph.D. thesis, Universität Hamburg, 2003.

Stuber, N., Forster, P., Rädel, G., and Shine, K.: The importance of the diurnal and annual cycle of air traffic for contrail radiative forcing, Nature, 441, 864–867, doi:10.1038/nature04877, 2006.

Sukhodolov, T., Rozanov, E., Shapiro, A. I., Anet, J., Cagnazzo, C., Peter, T., and Schmutz, W.: Evaluation of the ECHAM family radiation codes performance in the representation of the solar signal, Geoscientific Model Development, 7, 2859–2866, doi:10.5194/gmd-7-2859-2014, 2014.

Tian, B., Fetzer, E. J., Kahn, B. H., Teixeira, J., Manning, E., and Hearty, T.: Evaluating CMIP5 models using AIRS tropospheric air temperature and specific humidity climatology, Journal of Geophysical Research: Atmospheres, 118, 114–134, doi:10.1029/2012JD018607, 2013.

Wigley, T.: Relative contributions of different trace gases to the greenhouse effect., Climate Monitor, 16, 14–28, 1987.

Wilcox, L. J., Hoskins, B. J., and Shine, K. P.: A global blended tropopause based on ERA data. Part I: Climatology, Quarterly Journal of the Royal Meteorological Society, 138, 561–575, doi:10.1002/qj.951, 2012a.

Wilcox, L. J., Shine, K. P., and Hoskins, B. J.: Radiative forcing due to aviation water vapour emissions, Atmospheric Environment, 63, 1–13, doi:10.1016/j.atmosenv.2012.08.072, 2012b.

Wood, R. W.: Note on the theory of the greenhouse, Philosophical Magazine, 17, 319–320, doi:10.1080/14786440208636602, 1909.

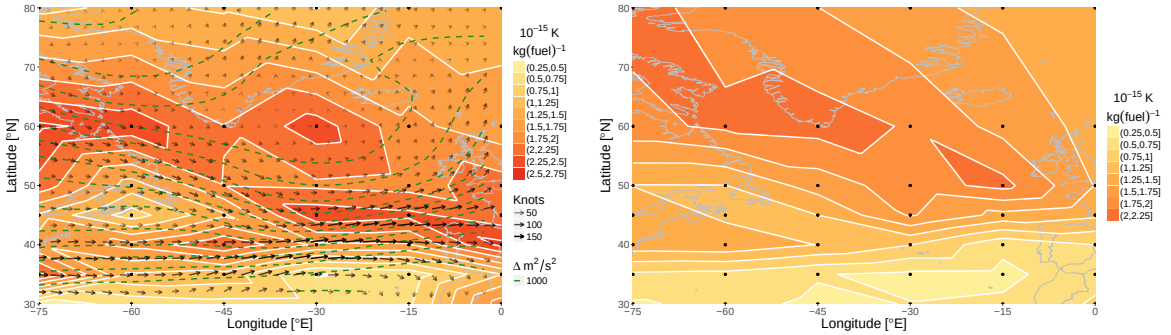
Woollings, T., Hannachi, A., and Hoskins, B.: Variability of the North Atlantic eddy-driven jet stream, Quarterly Journal of the Royal Meteorological Society, 136, 856–868, doi:10.1002/qj.625, 2010.

Yamashita, H., Grewe, V., Jöckel, P., Linke, F., Schaefer, M., and Sasaki, D.: Air traffic simulation in chemistry-climate model EMAC 2.41: AirTraf 1.0, Geoscientific Model Development, 9, 3363–3392, doi:10.5194/gmd-9-3363-2016, 2016.

A

Water vapour CCF pressure heatmaps

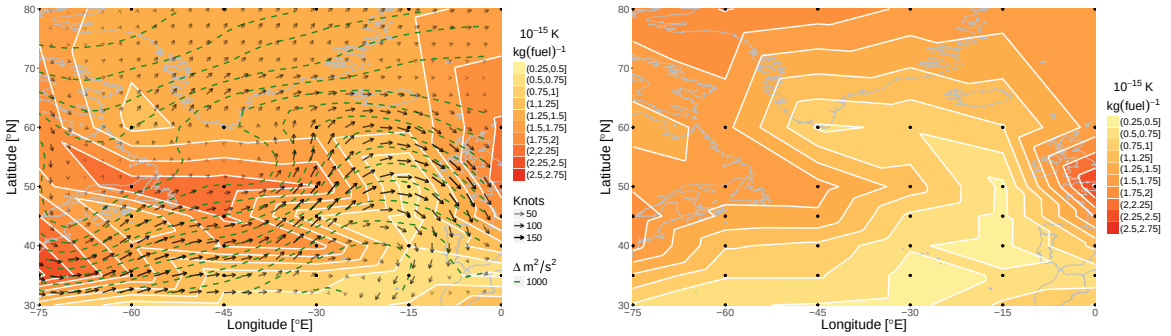
This appendix contains longitude/latitude heatmaps of the original and algorithmic water vapour CCFs. Note that the colour coding is the same for every plot, but the legends only show colours used in that specific plot. The black points are the initial REACT4C tracer emission locations. Bilinearly interpolated contours are shown because application of the CCFs would likely use bilinear interpolation. The dark green dotted lines are geopotential isolines in [1000 m^2 s^{-2}], and the magnitude of the wind vectors in [knots] is reflected by the length and opacity of the vectors.



(a) REACT4C data

(b) Final $|PV|$ algorithm

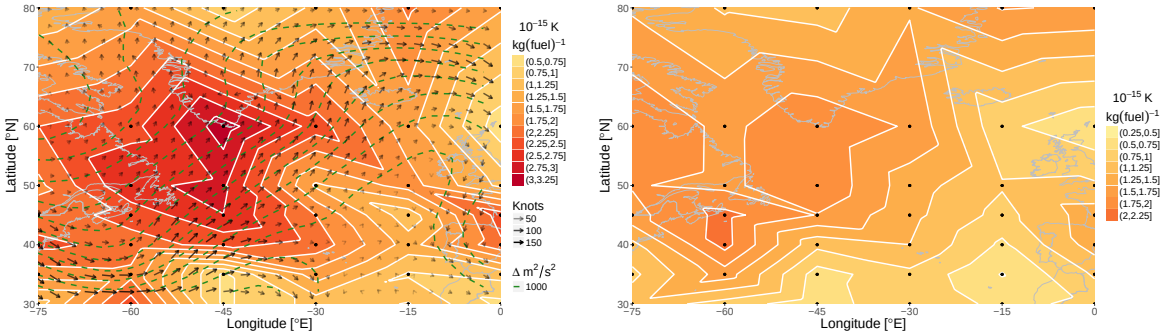
Figure A.1: Water vapour CCF/aCCF contour plots, 12 UTC, 200 hPa, WP 1



(a) REACT4C data

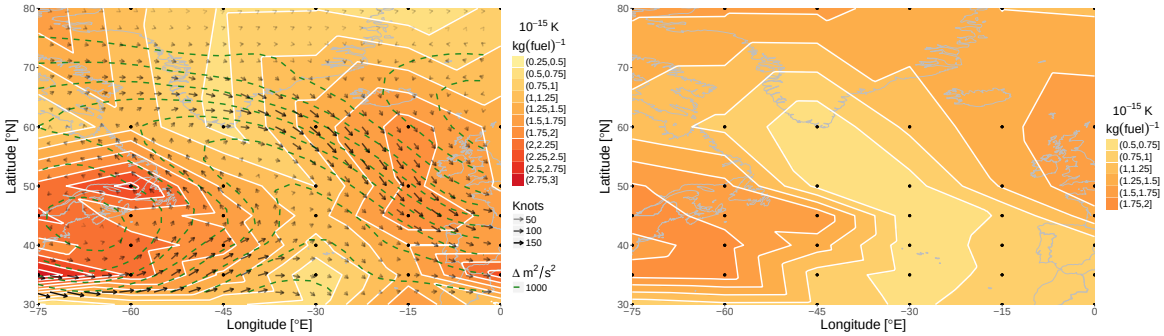
(b) Final $|PV|$ algorithm

Figure A.2: Water vapour CCF/aCCF contour plots, 12 UTC, 200 hPa, WP 2



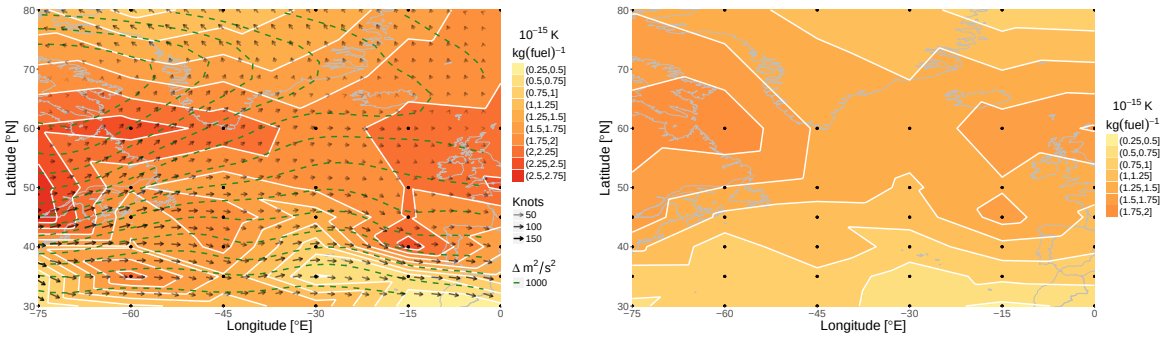
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure A.3: Water vapour CCF/aCCF contour plots, 12 UTC, 200 hPa, WP 3



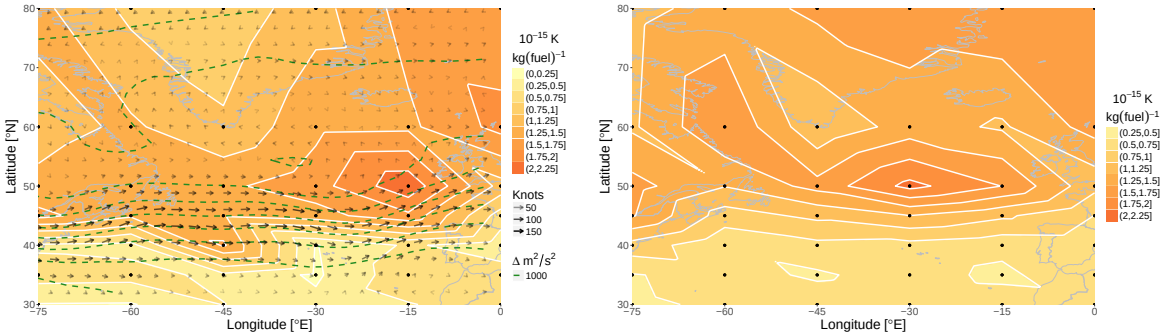
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure A.4: Water vapour CCF/aCCF contour plots, 12 UTC, 200 hPa, WP 4



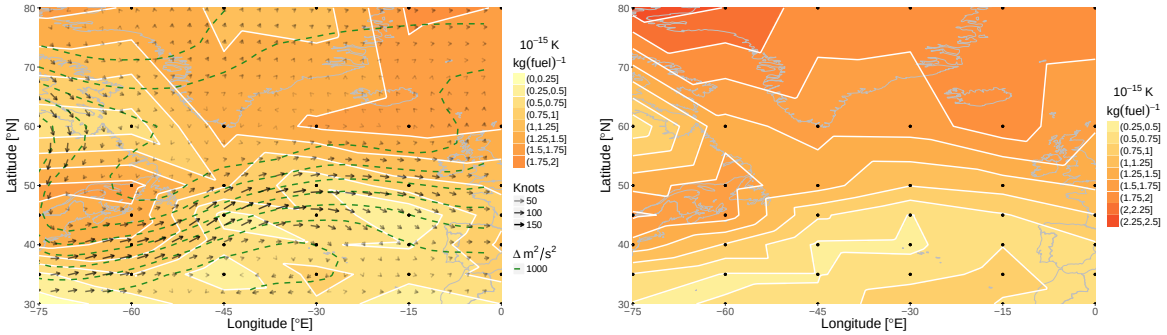
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure A.5: Water vapour CCF/aCCF contour plots, 12 UTC, 200 hPa, WP 5



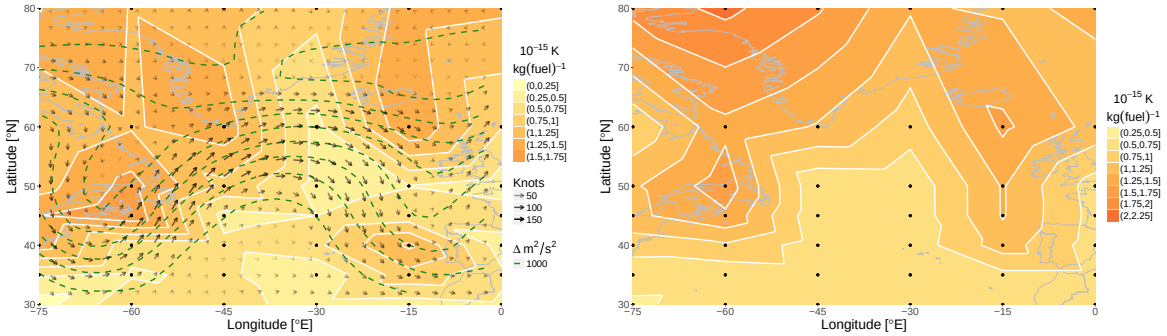
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure A.6: Water vapour CCF/aCCF contour plots, 12 UTC, 200 hPa, WP 6



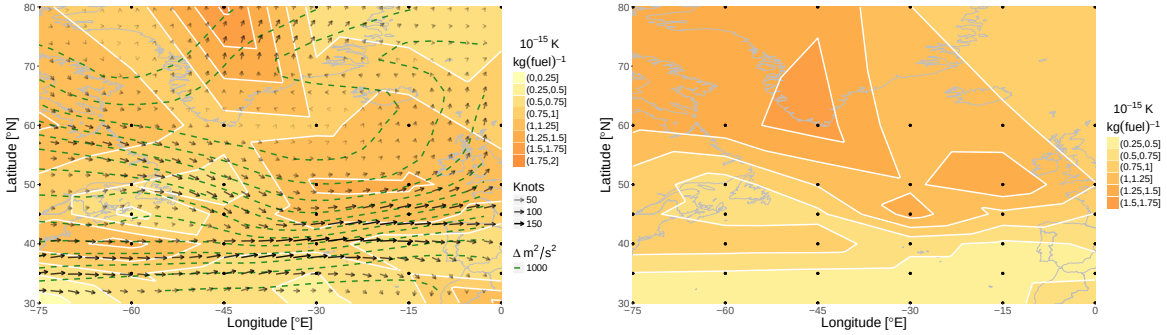
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure A.7: Water vapour CCF/aCCF contour plots, 12 UTC, 200 hPa, WP 7



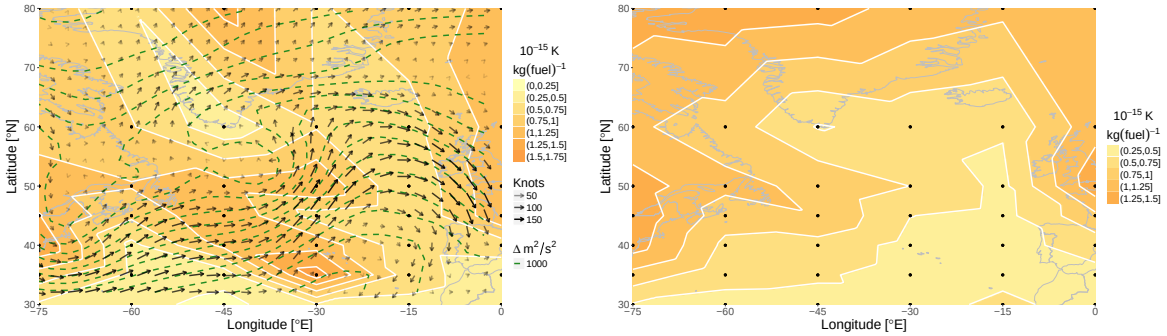
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure A.8: Water vapour CCF/aCCF contour plots, 12 UTC, 200 hPa, WP 8



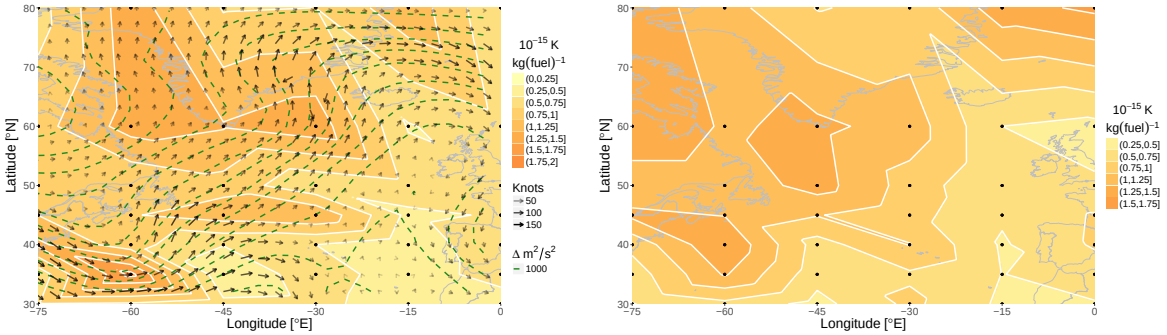
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure A.9: Water vapour CCF/aCCF contour plots, 12 UTC, 250 hPa, WP 1



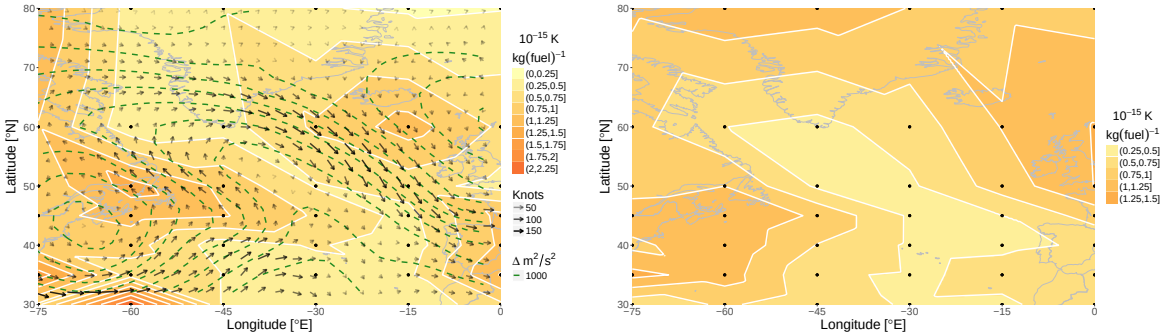
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure A.10: Water vapour CCF/aCCF contour plots, 12 UTC, 250 hPa, WP 2



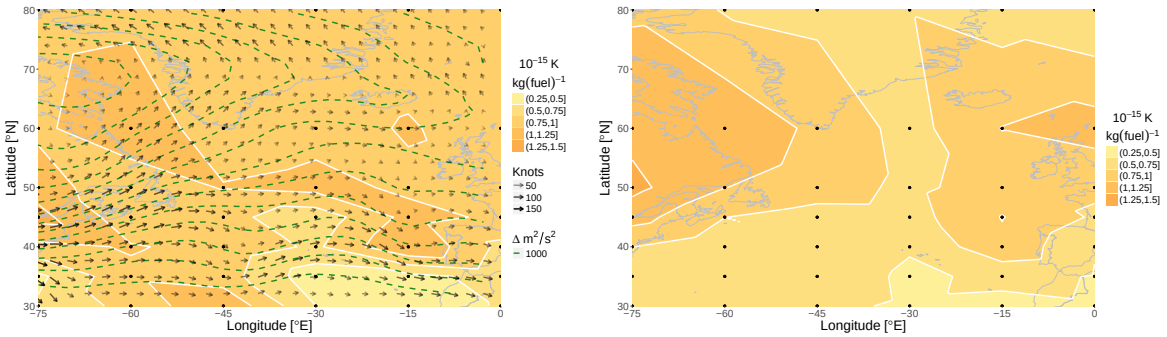
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure A.11: Water vapour CCF/aCCF contour plots, 12 UTC, 250 hPa, WP 3



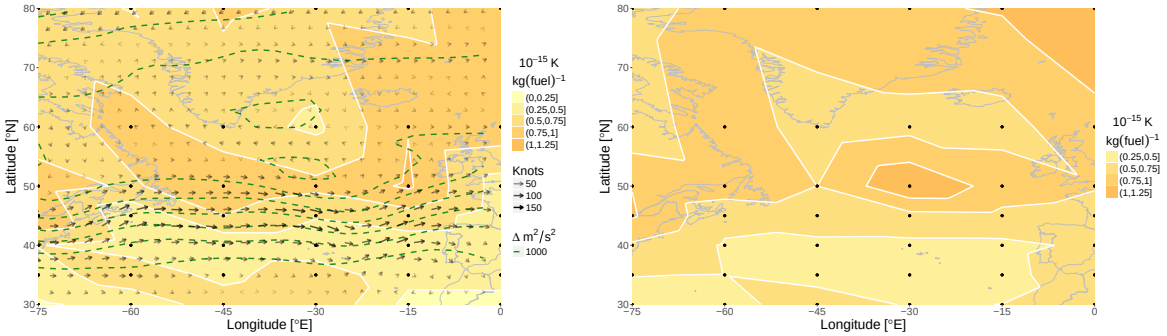
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure A.12: Water vapour CCF/aCCF contour plots, 12 UTC, 250 hPa, WP 4



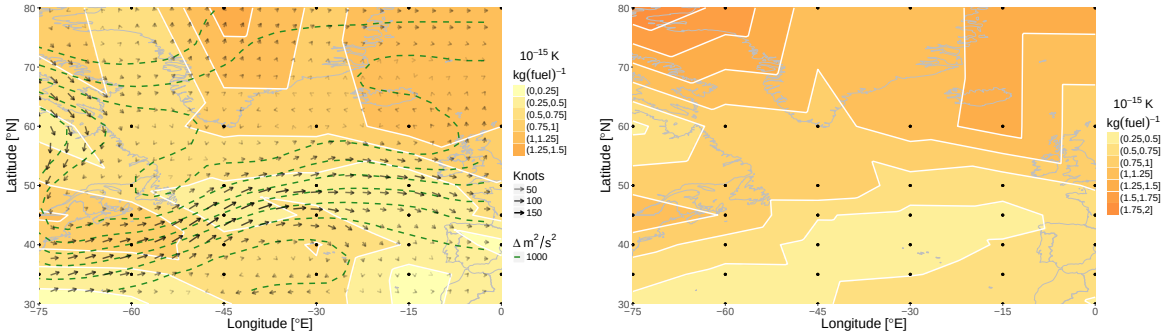
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure A.13: Water vapour CCF/aCCF contour plots, 12 UTC, 250 hPa, WP 5



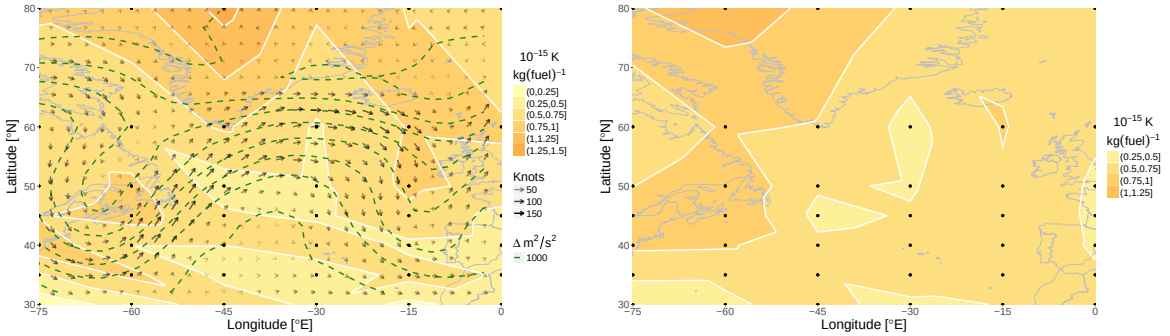
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure A.14: Water vapour CCF/aCCF contour plots, 12 UTC, 250 hPa, WP 6



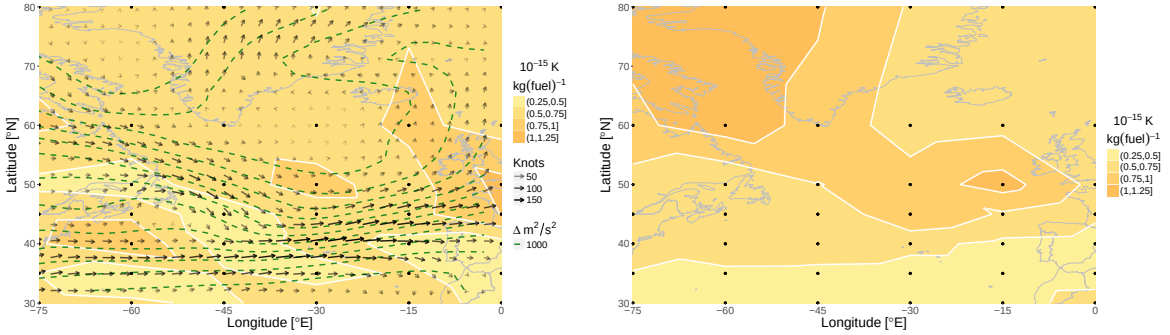
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure A.15: Water vapour CCF/aCCF contour plots, 12 UTC, 250 hPa, WP 7



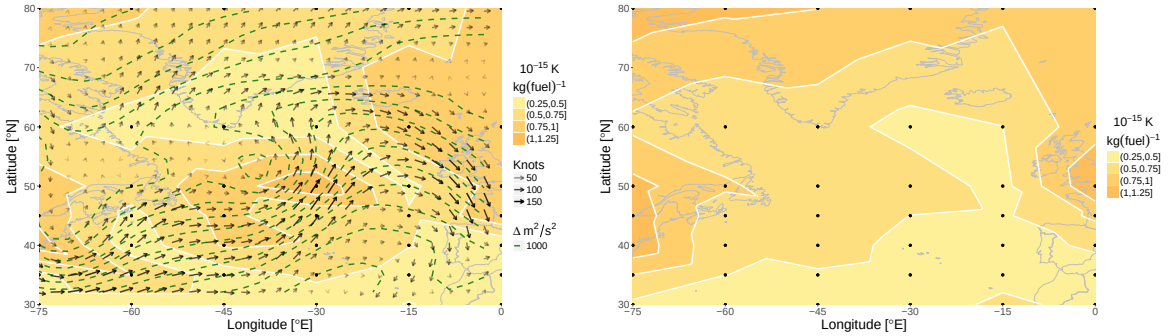
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure A.16: Water vapour CCF/aCCF contour plots, 12 UTC, 250 hPa, WP 8



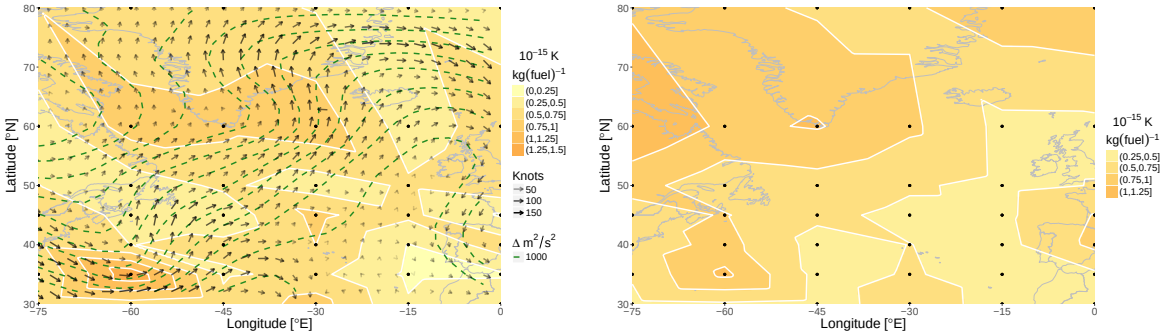
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure A.17: Water vapour CCF/aCCF contour plots, 12 UTC, 300 hPa, WP 1



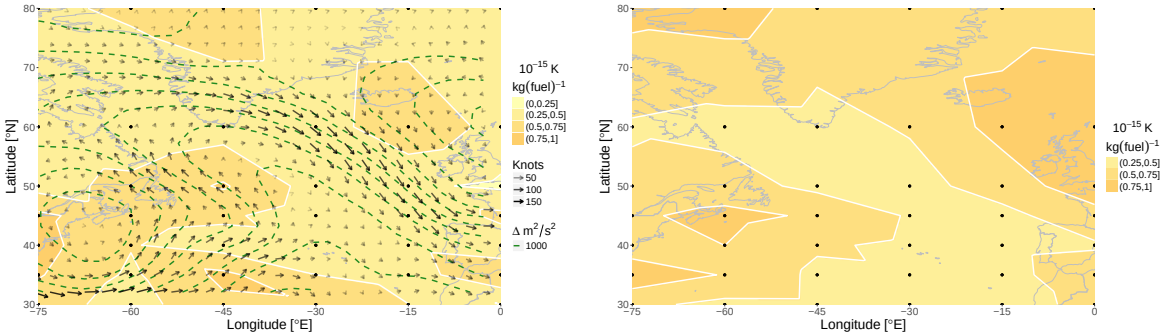
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure A.18: Water vapour CCF/aCCF contour plots, 12 UTC, 300 hPa, WP 2



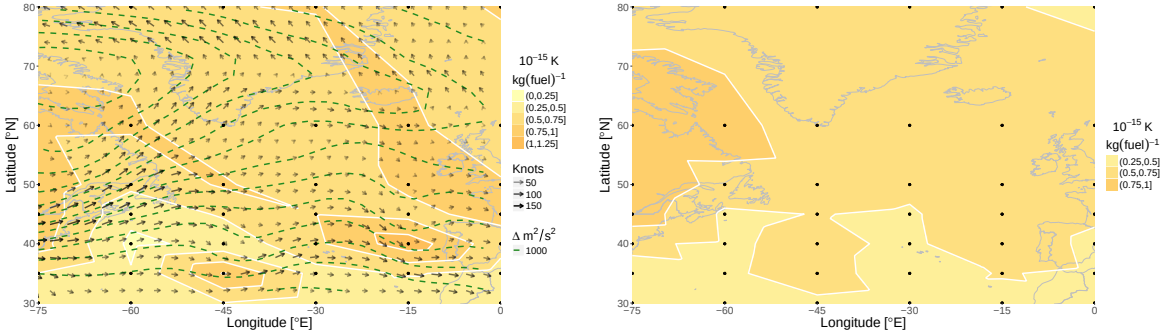
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure A.19: Water vapour CCF/aCCF contour plots, 12 UTC, 300 hPa, WP 3



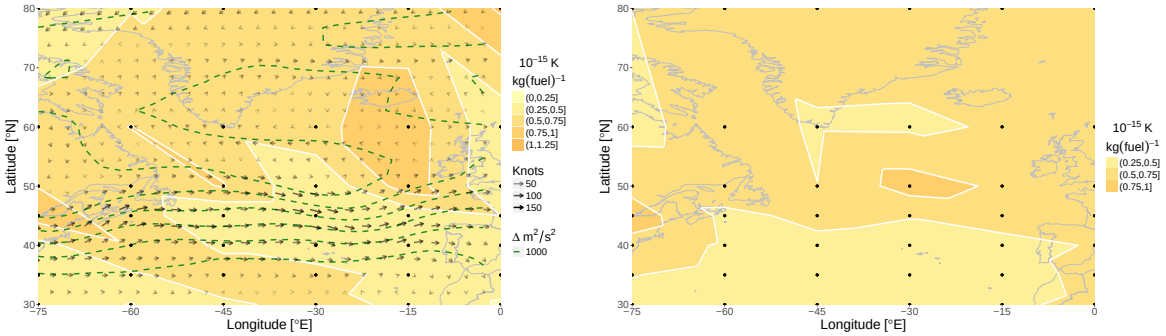
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure A.20: Water vapour CCF/aCCF contour plots, 12 UTC, 300 hPa, WP 4



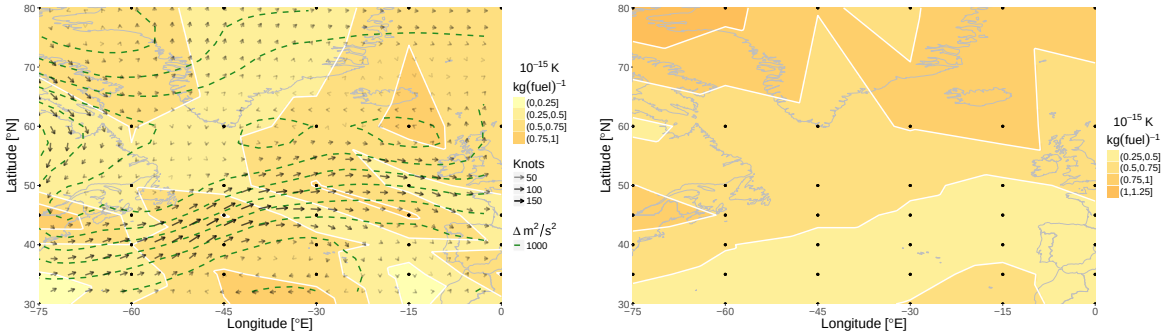
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure A.21: Water vapour CCF/aCCF contour plots, 12 UTC, 300 hPa, WP 5



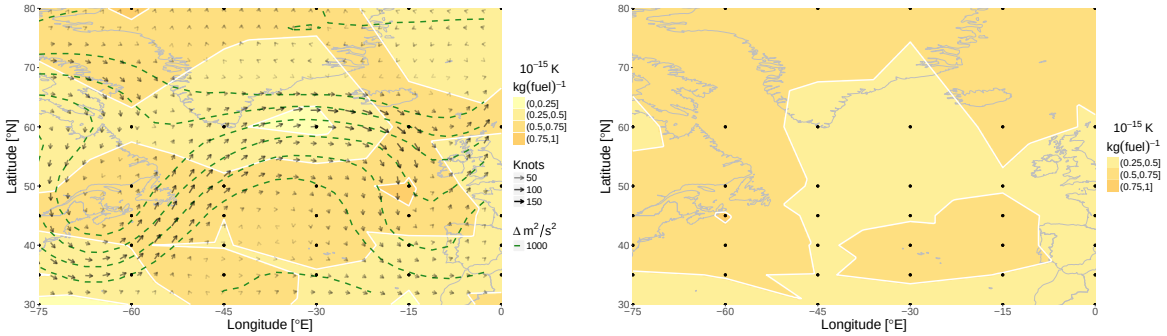
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure A.22: Water vapour CCF/aCCF contour plots, 12 UTC, 300 hPa, WP 6



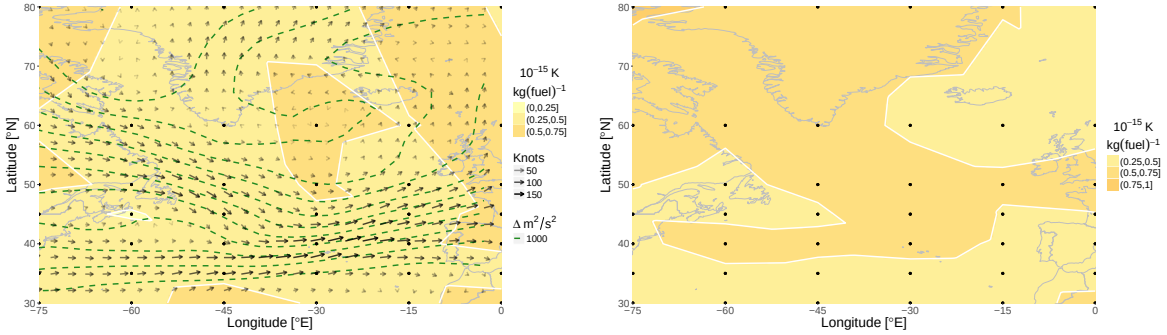
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure A.23: Water vapour CCF/aCCF contour plots, 12 UTC, 300 hPa, WP 7



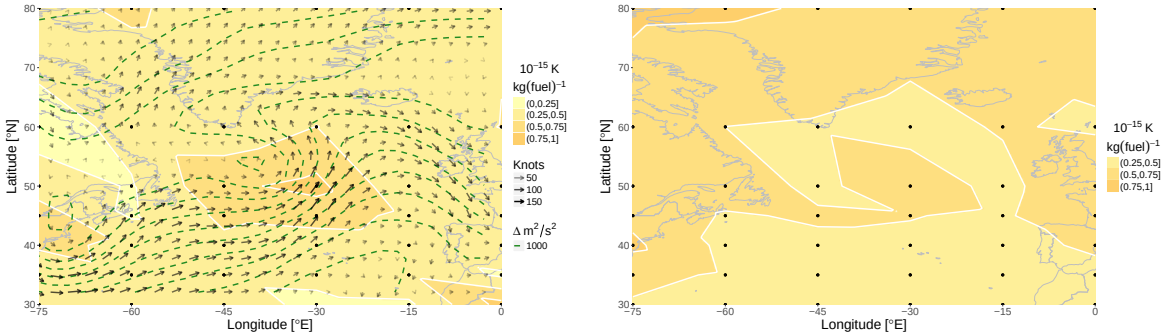
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure A.24: Water vapour CCF/aCCF contour plots, 12 UTC, 300 hPa, WP 8



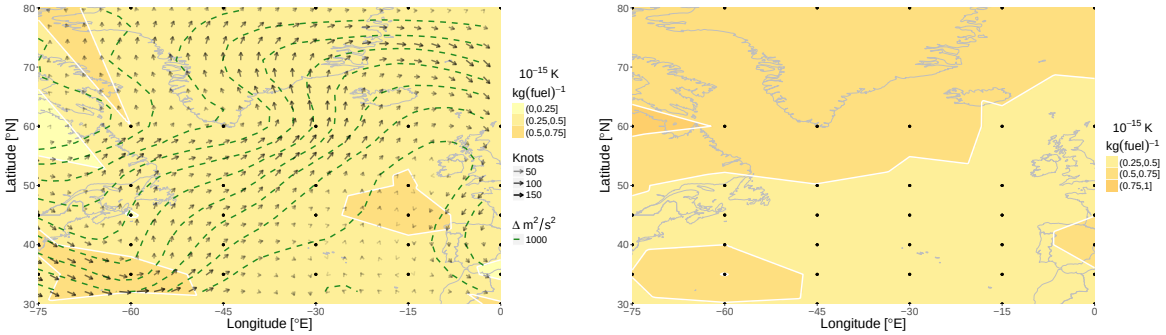
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure A.25: Water vapour CCF/aCCF contour plots, 12 UTC, 400 hPa, WP 1



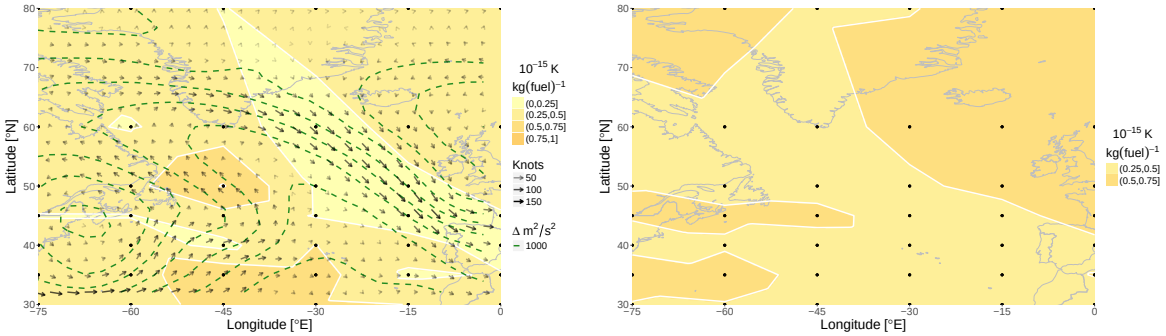
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure A.26: Water vapour CCF/aCCF contour plots, 12 UTC, 400 hPa, WP 2



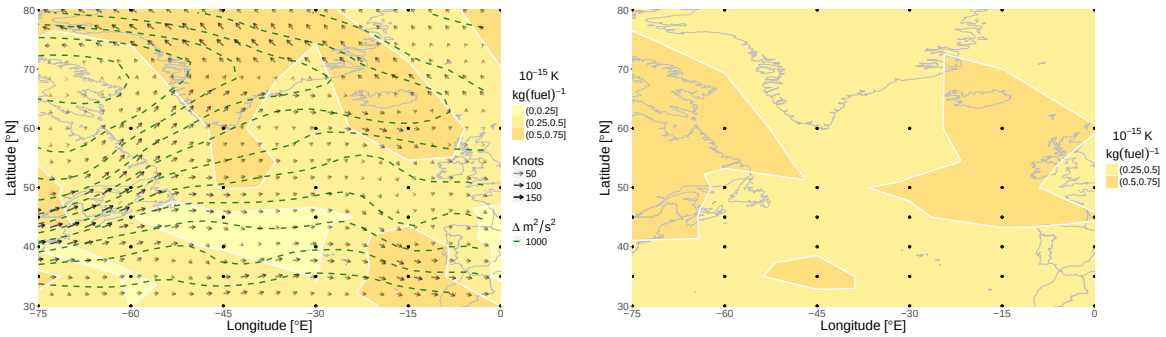
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure A.27: Water vapour CCF/aCCF contour plots, 12 UTC, 400 hPa, WP 3



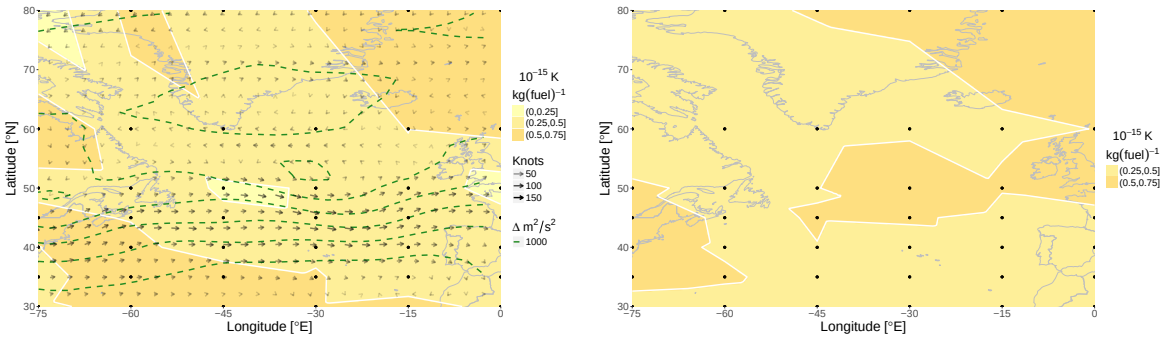
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure A.28: Water vapour CCF/aCCF contour plots, 12 UTC, 400 hPa, WP 4



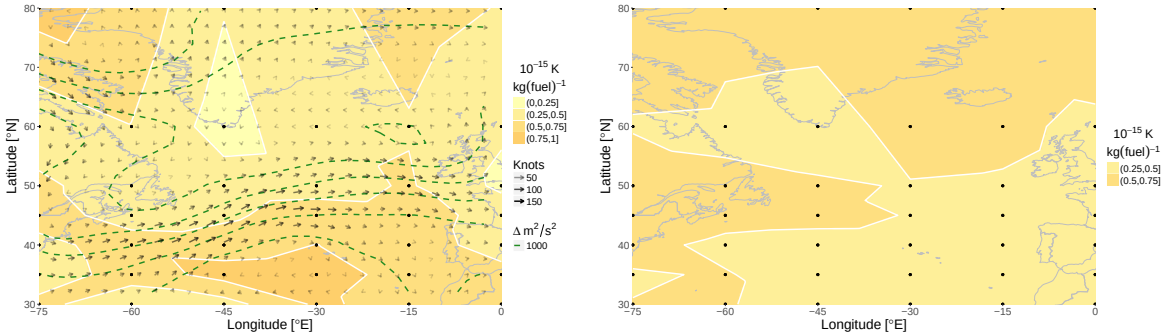
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure A.29: Water vapour CCF/aCCF contour plots, 12 UTC, 400 hPa, WP 5



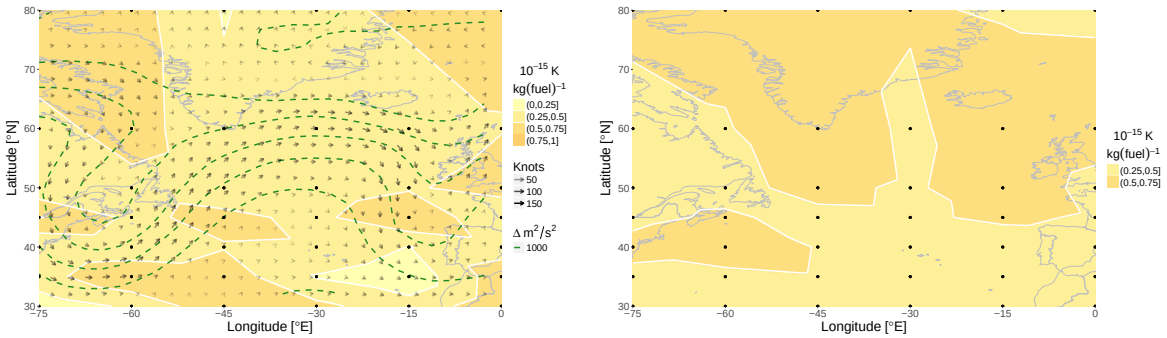
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure A.30: Water vapour CCF/aCCF contour plots, 12 UTC, 400 hPa, WP 6



(a) REACT4C data (b) Final $|PV|$ algorithm

Figure A.31: Water vapour CCF/aCCF contour plots, 12 UTC, 400 hPa, WP 7



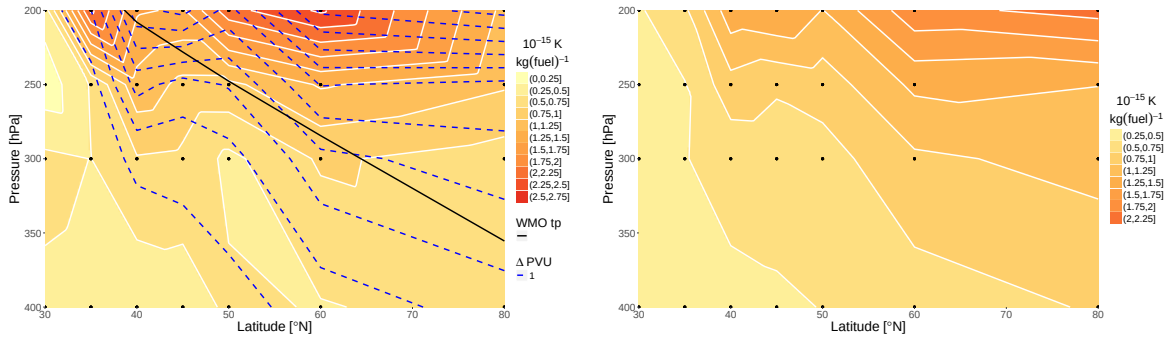
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure A.32: Water vapour CCF/aCCF contour plots, 12 UTC, 400 hPa, WP 8

B

Water vapour CCF zonal heatmaps

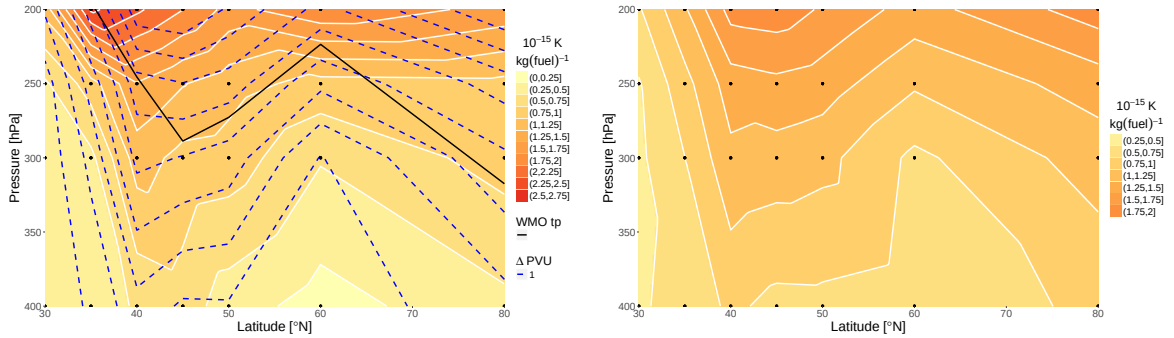
This appendix contains level/latitude heatmaps of the original and algorithmic water vapour CCFs. Note that the colour coding is the same for every plot, but the legends only show colours used in that specific plot. The black points are the initial REACT4C tracer emission locations. Bilinearly interpolated contours are shown because application of the CCFs would likely use bilinear interpolation. The dark blue dotted lines are potential vorticity isolines in [1 PVU], and the black line shows the location of the WMO thermal tropopause. If the tropopause is not visible at a specific latitude, it is always located above the plotted domain and never below it.



(a) REACT4C data

(b) Final $|PV|$ algorithm

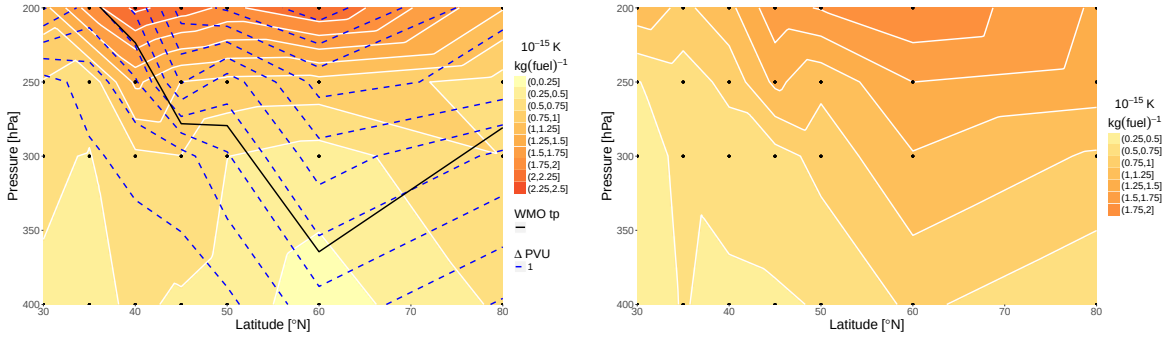
Figure B.1: Water vapour CCF/aCCF contour plots, 12 UTC, 285 °E, WP 1



(a) REACT4C data

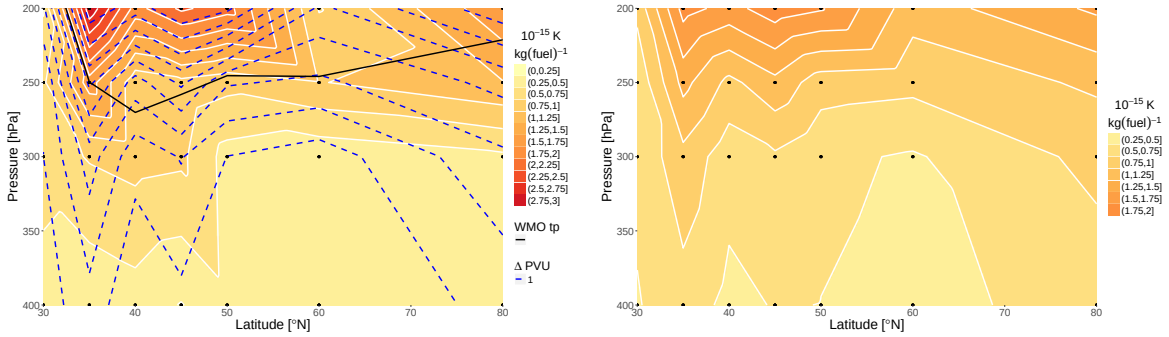
(b) Final $|PV|$ algorithm

Figure B.2: Water vapour CCF/aCCF contour plots, 12 UTC, 285 °E, WP 2



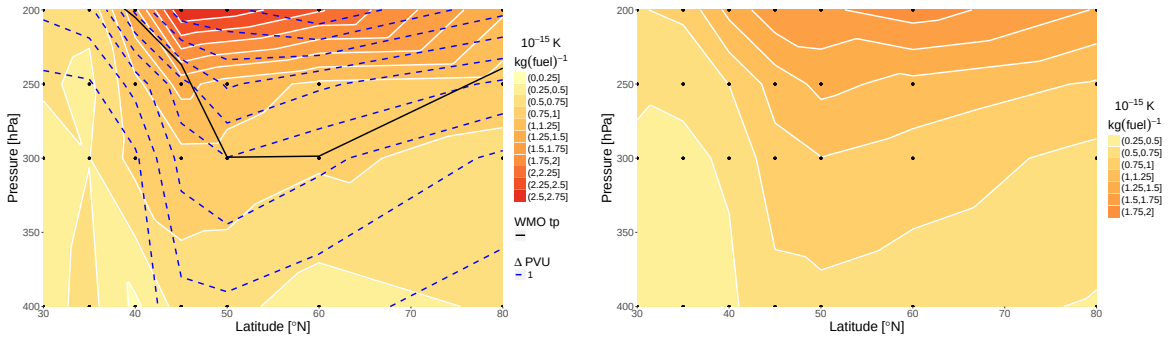
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.3: Water vapour CCF/aCCF contour plots, 12 UTC, 285 °E, WP 3



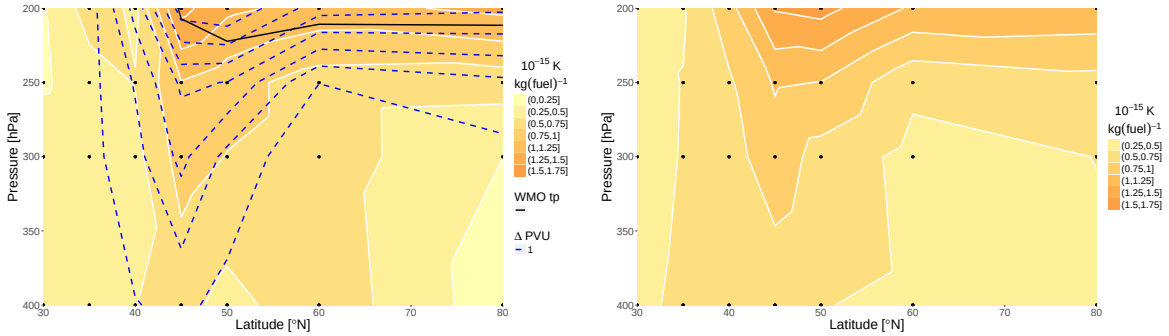
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.4: Water vapour CCF/aCCF contour plots, 12 UTC, 285 °E, WP 4



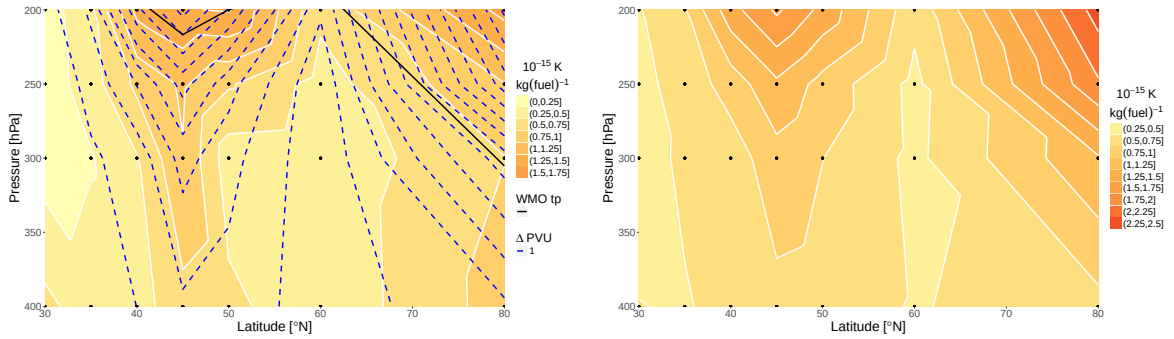
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.5: Water vapour CCF/aCCF contour plots, 12 UTC, 285 °E, WP 5



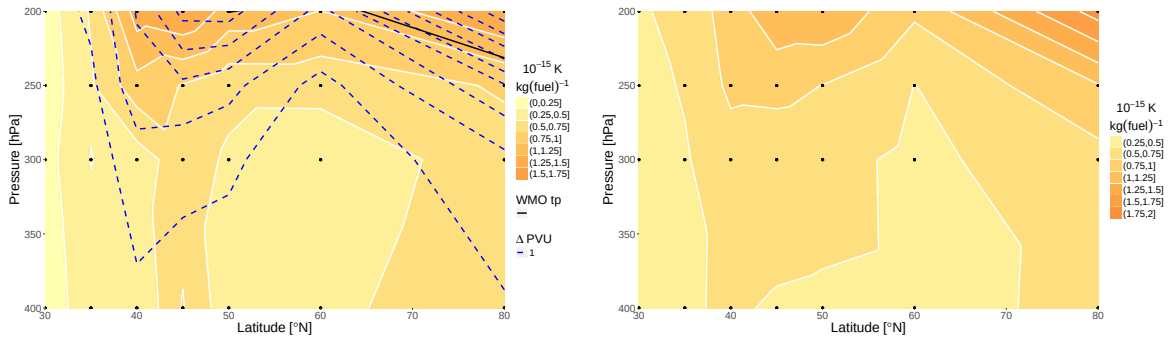
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.6: Water vapour CCF/aCCF contour plots, 12 UTC, 285 °E, WP 6



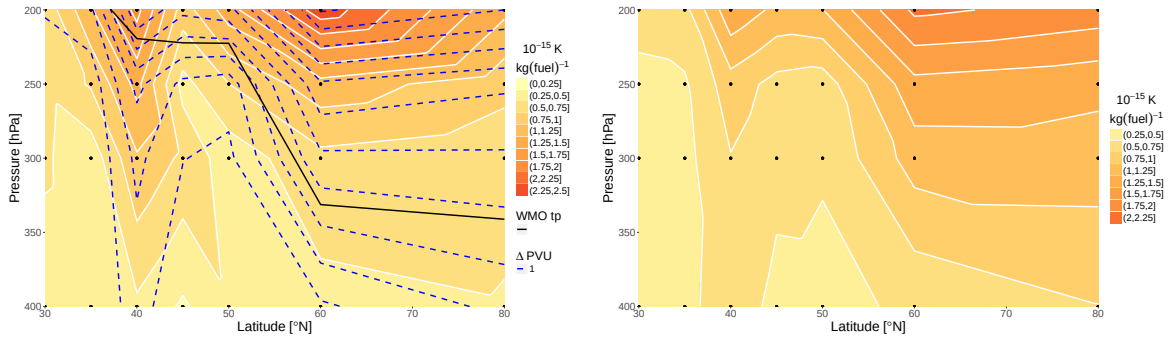
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.7: Water vapour CCF/aCCF contour plots, 12 UTC, 285 °E, WP 7



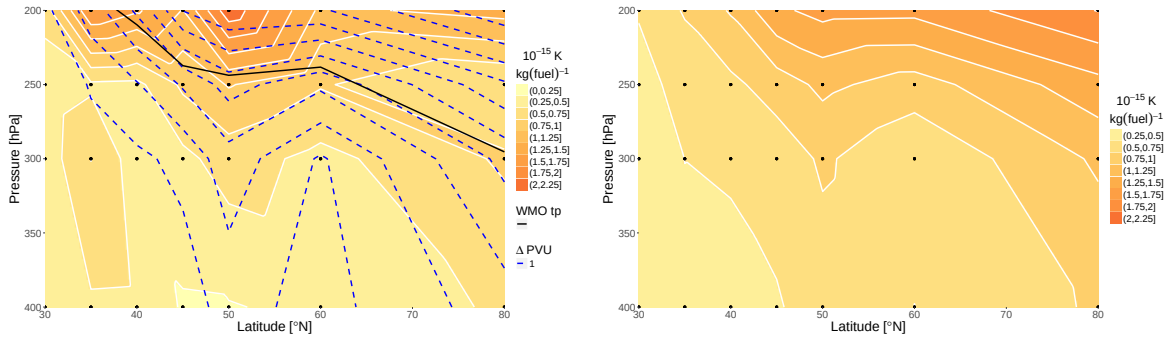
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.8: Water vapour CCF/aCCF contour plots, 12 UTC, 285 °E, WP 8



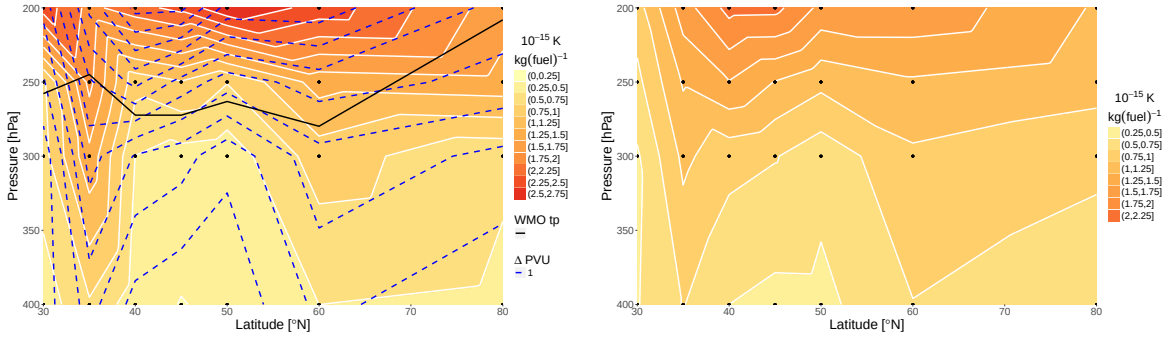
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.9: Water vapour CCF/aCCF contour plots, 12 UTC, 300 °E, WP 1



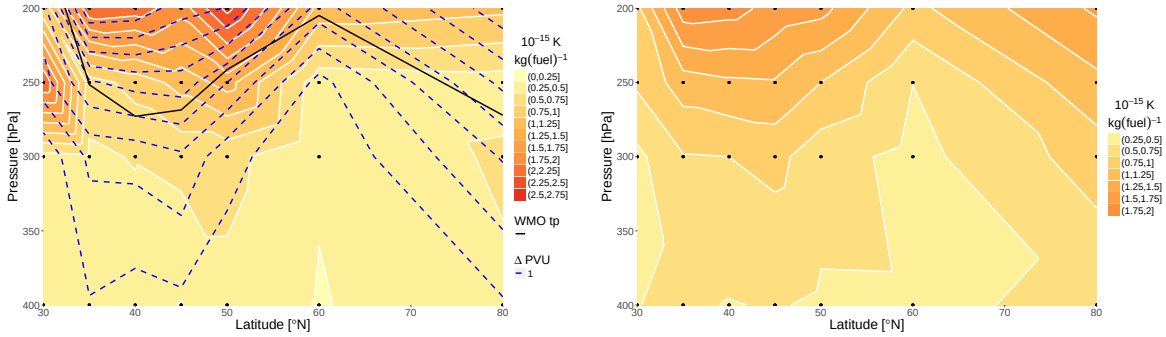
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.10: Water vapour CCF/aCCF contour plots, 12 UTC, 300 °E, WP 2



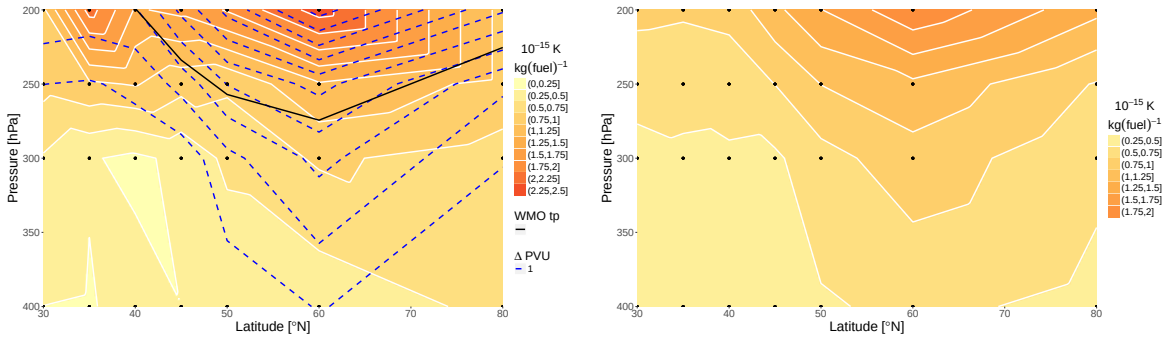
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.11: Water vapour CCF/aCCF contour plots, 12 UTC, 300 °E, WP 3



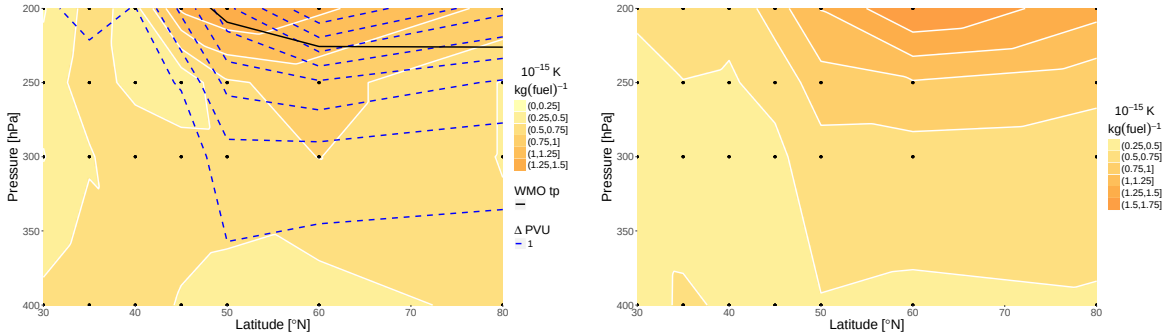
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.12: Water vapour CCF/aCCF contour plots, 12 UTC, 300 °E, WP 4



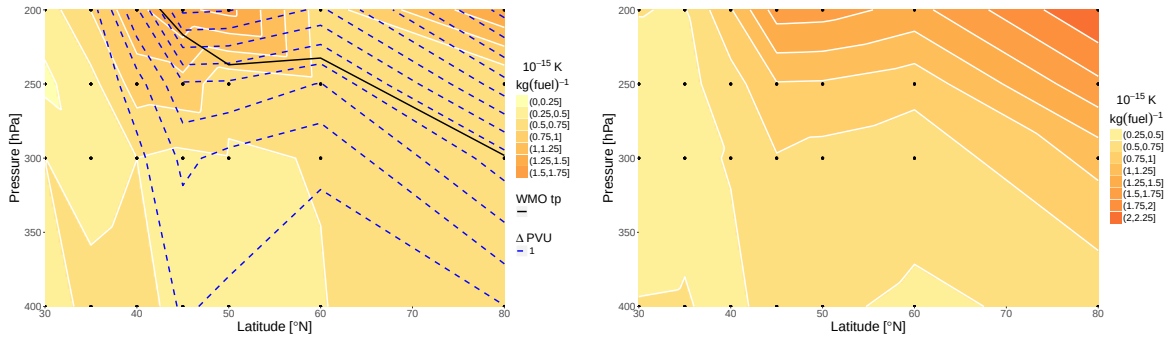
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.13: Water vapour CCF/aCCF contour plots, 12 UTC, 300 °E, WP 5



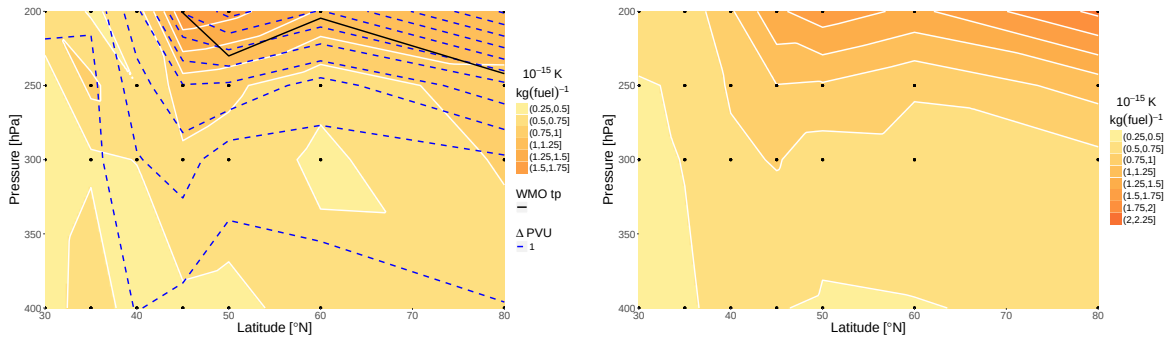
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.14: Water vapour CCF/aCCF contour plots, 12 UTC, 300 °E, WP 6



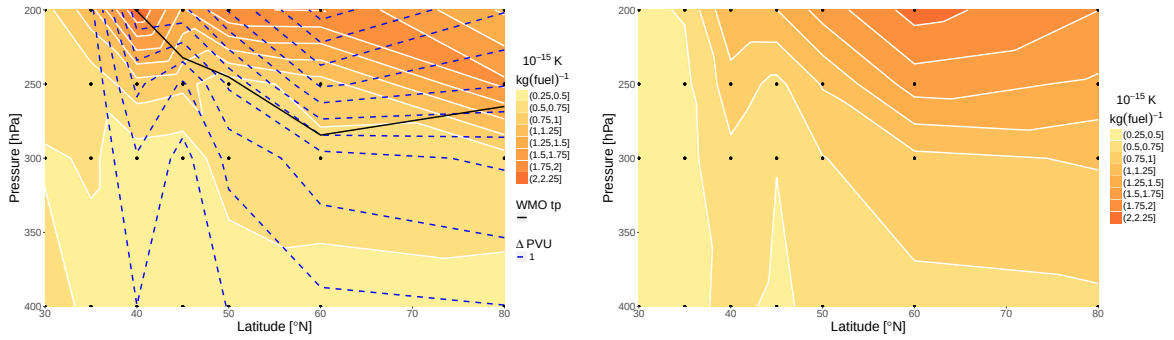
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.15: Water vapour CCF/aCCF contour plots, 12 UTC, 300 °E, WP 7



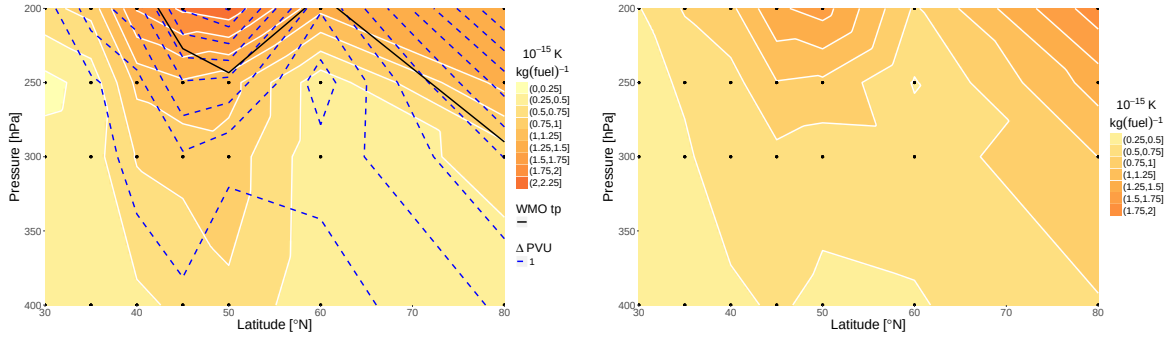
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.16: Water vapour CCF/aCCF contour plots, 12 UTC, 300 °E, WP 8



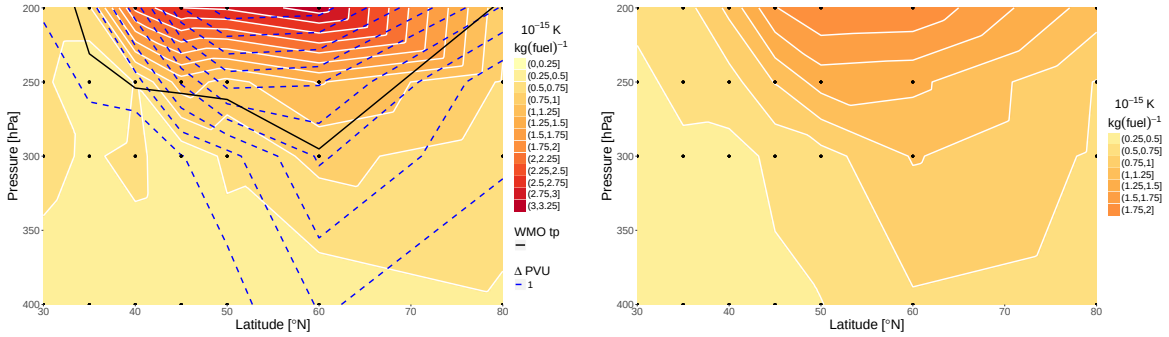
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.17: Water vapour CCF/aCCF contour plots, 12 UTC, 315 °E, WP 1



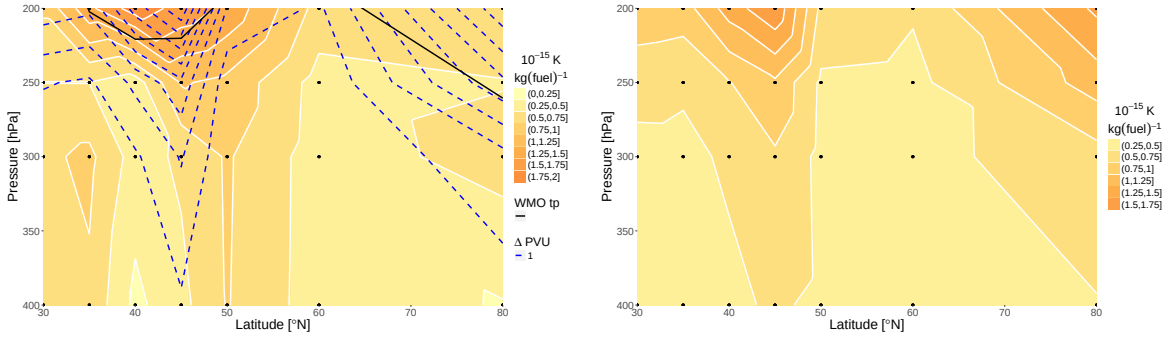
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.18: Water vapour CCF/aCCF contour plots, 12 UTC, 315 °E, WP 2



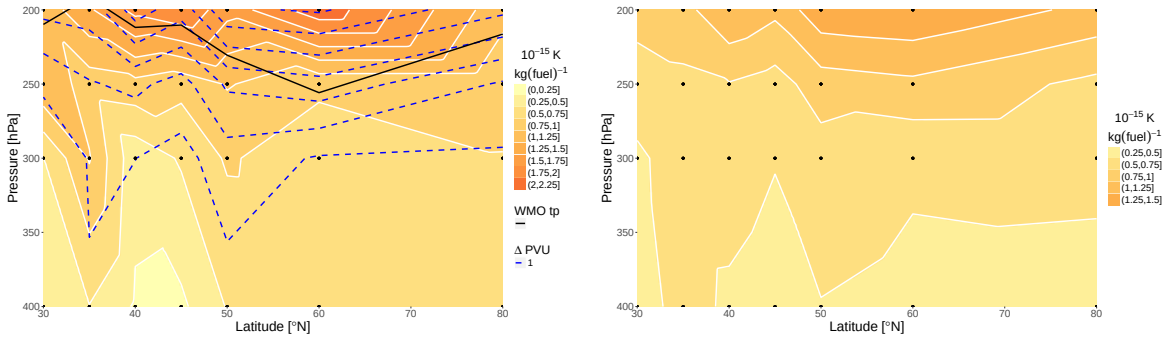
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.19: Water vapour CCF/aCCF contour plots, 12 UTC, 315 °E, WP 3



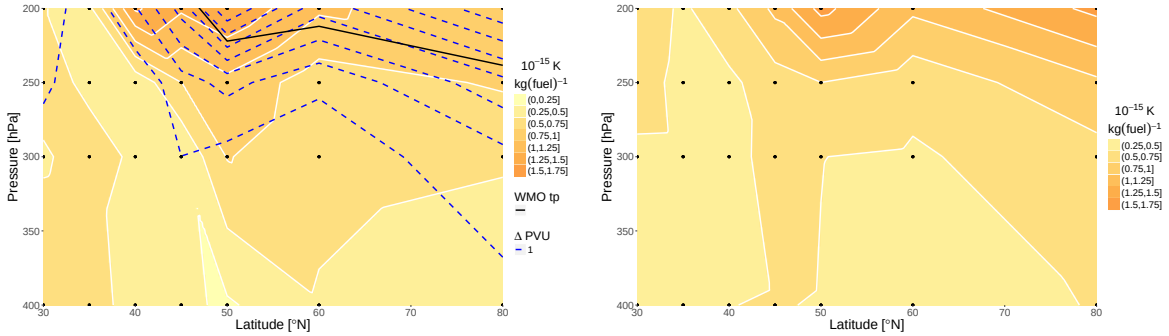
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.20: Water vapour CCF/aCCF contour plots, 12 UTC, 315 °E, WP 4



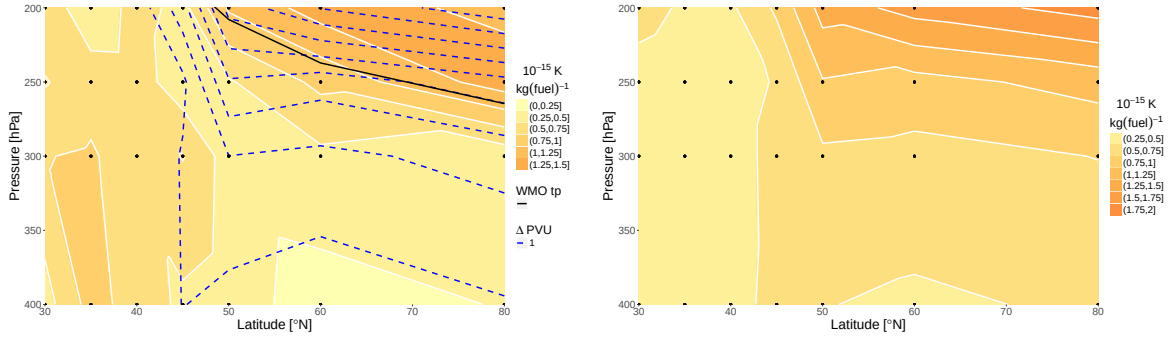
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.21: Water vapour CCF/aCCF contour plots, 12 UTC, 315 °E, WP 5



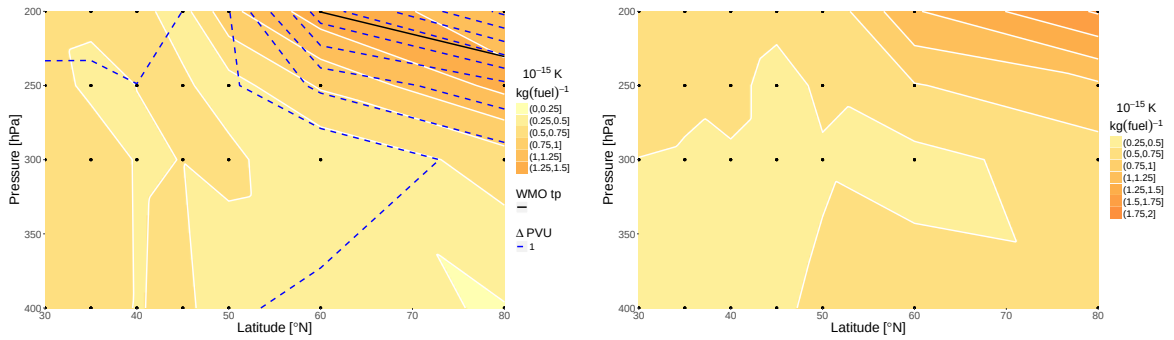
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.22: Water vapour CCF/aCCF contour plots, 12 UTC, 315 °E, WP 6



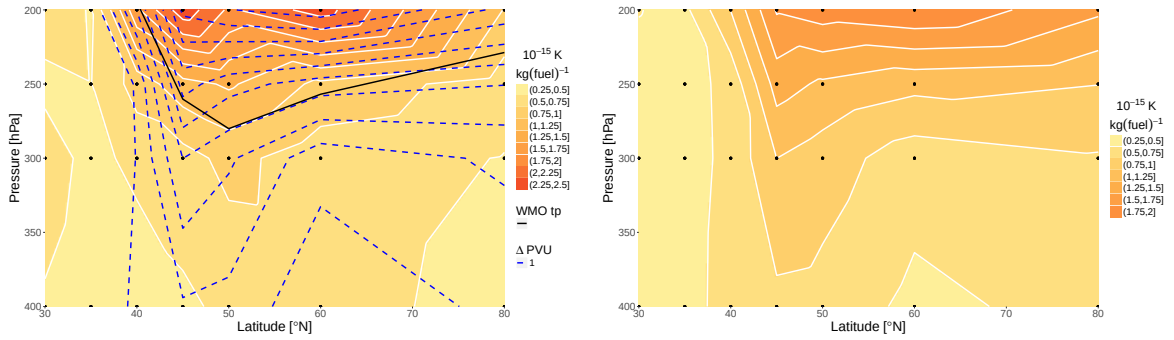
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.23: Water vapour CCF/aCCF contour plots, 12 UTC, 315 °E, WP 7



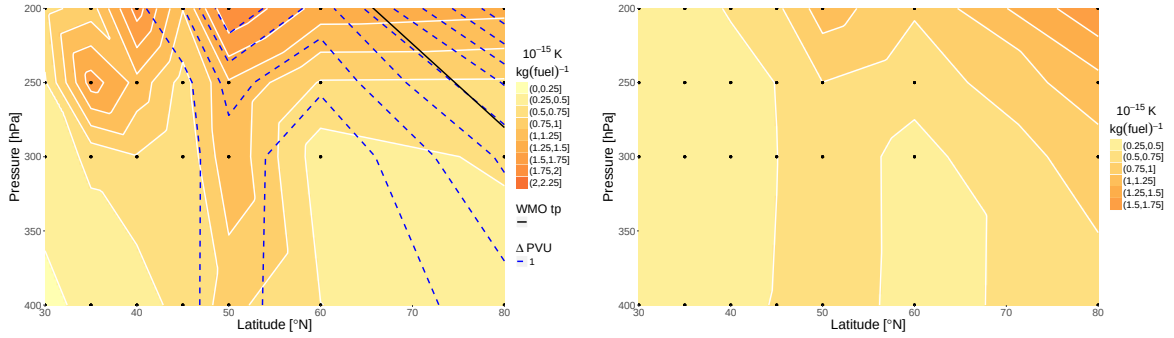
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.24: Water vapour CCF/aCCF contour plots, 12 UTC, 315 °E, WP 8



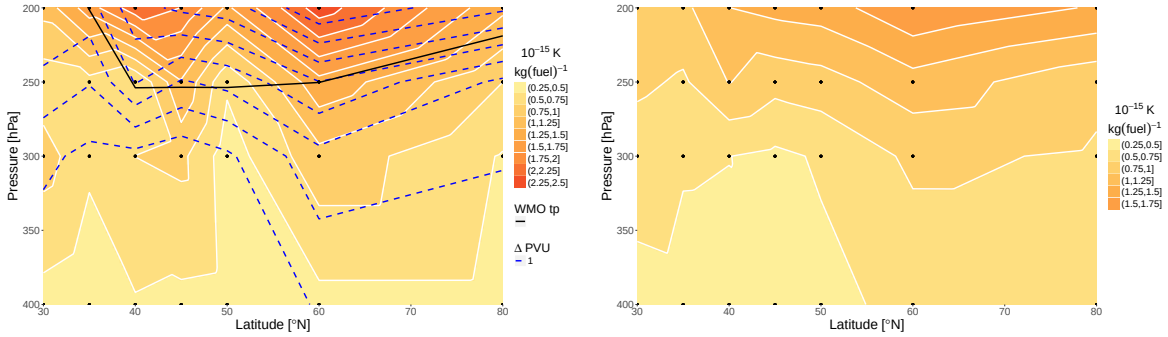
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.25: Water vapour CCF/aCCF contour plots, 12 UTC, 330 °E, WP 1



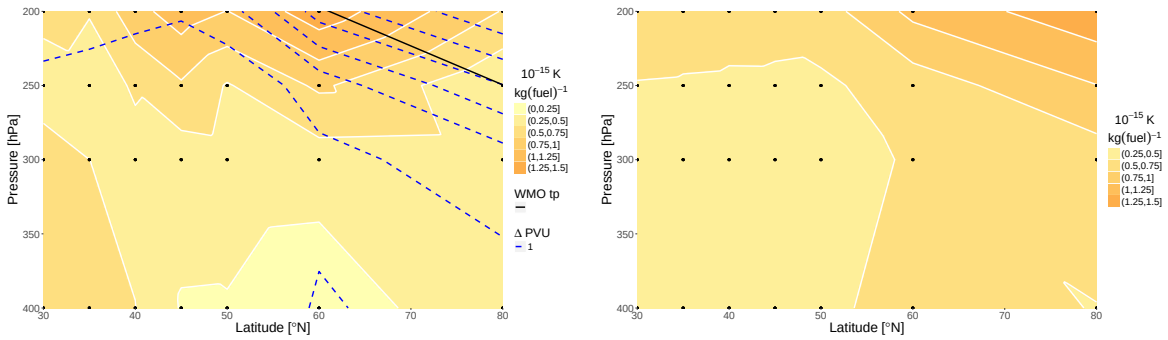
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.26: Water vapour CCF/aCCF contour plots, 12 UTC, 330 °E, WP 2



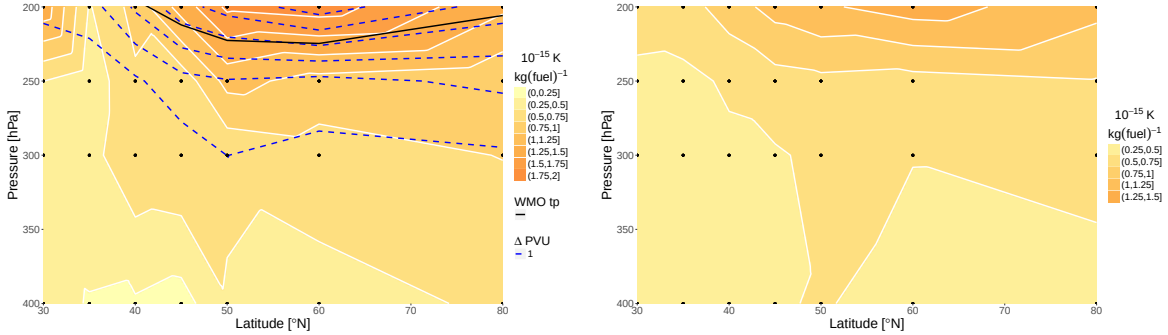
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.27: Water vapour CCF/aCCF contour plots, 12 UTC, 330 °E, WP 3



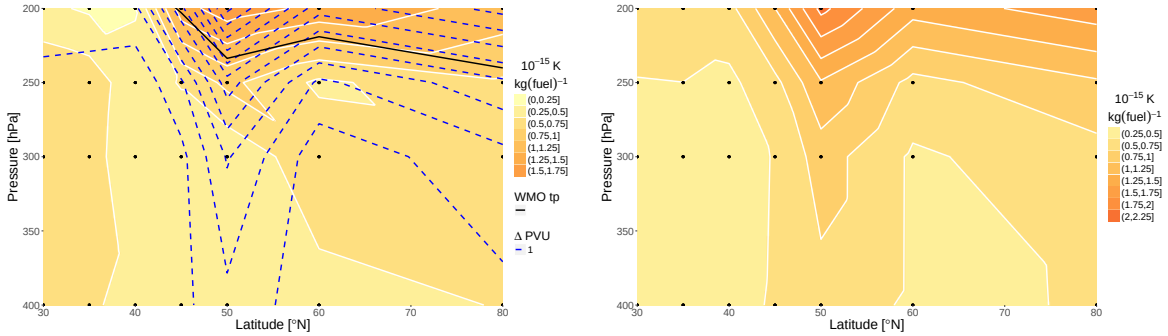
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.28: Water vapour CCF/aCCF contour plots, 12 UTC, 330 °E, WP 4



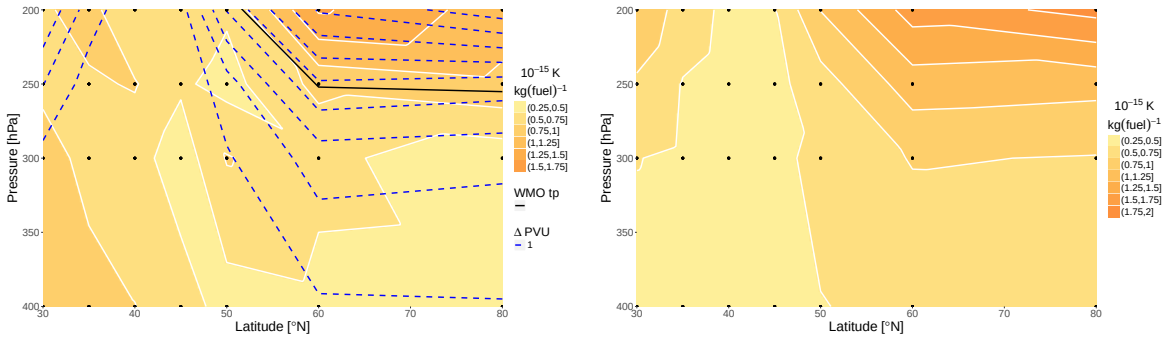
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.29: Water vapour CCF/aCCF contour plots, 12 UTC, 330 °E, WP 5



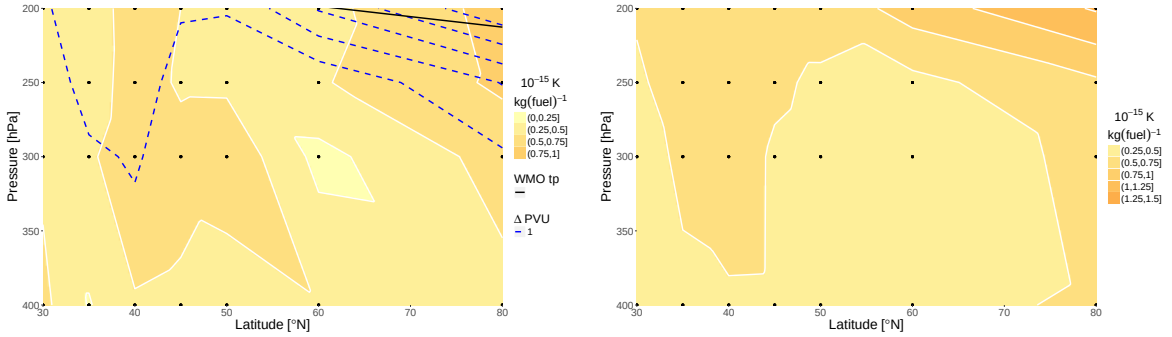
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.30: Water vapour CCF/aCCF contour plots, 12 UTC, 330 °E, WP 6



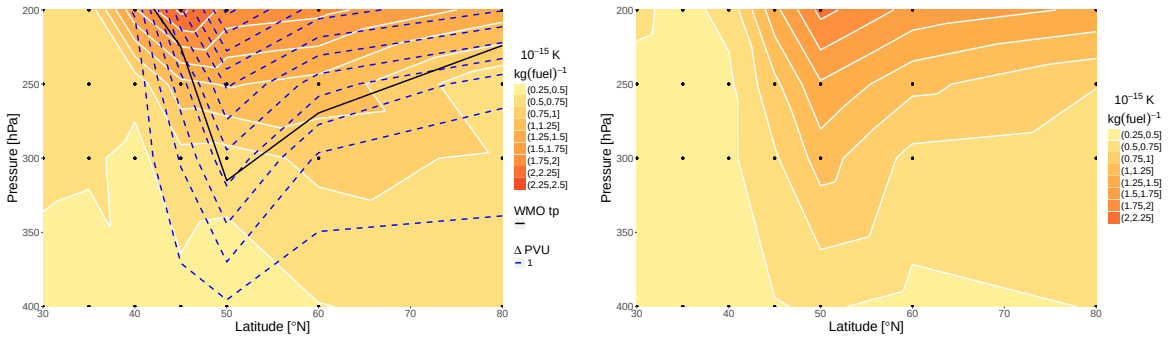
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.31: Water vapour CCF/aCCF contour plots, 12 UTC, 330 °E, WP 7



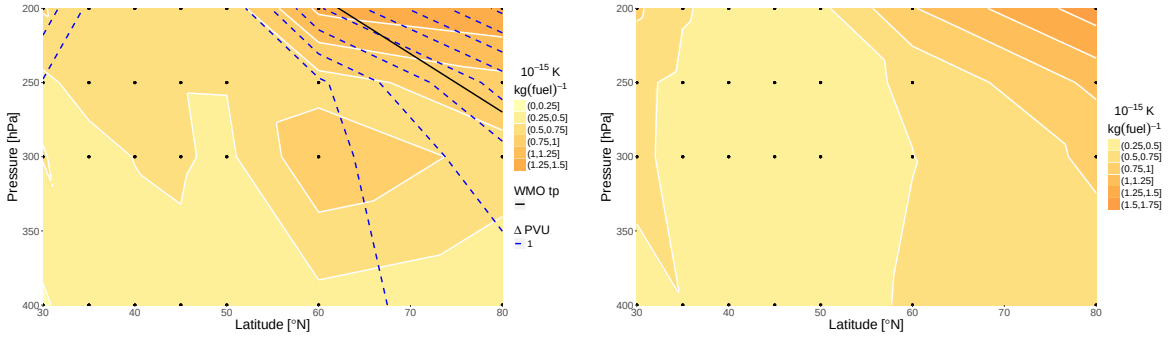
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.32: Water vapour CCF/aCCF contour plots, 12 UTC, 330 °E, WP 8



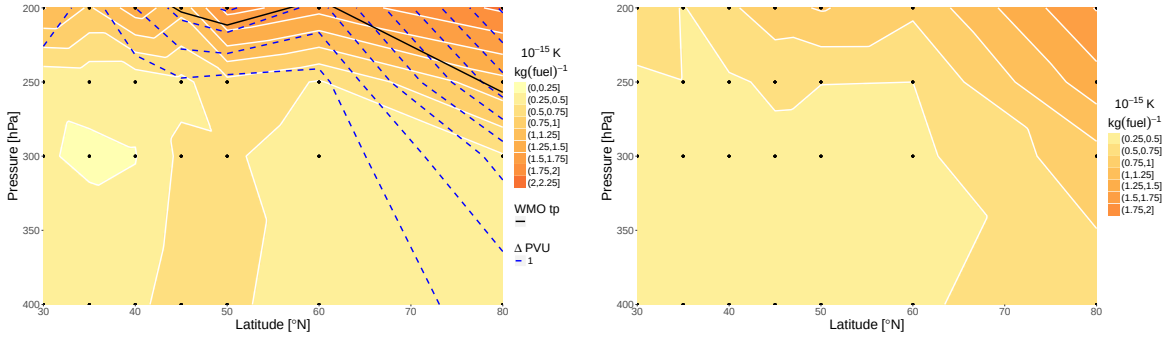
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.33: Water vapour CCF/aCCF contour plots, 12 UTC, 345 °E, WP 1



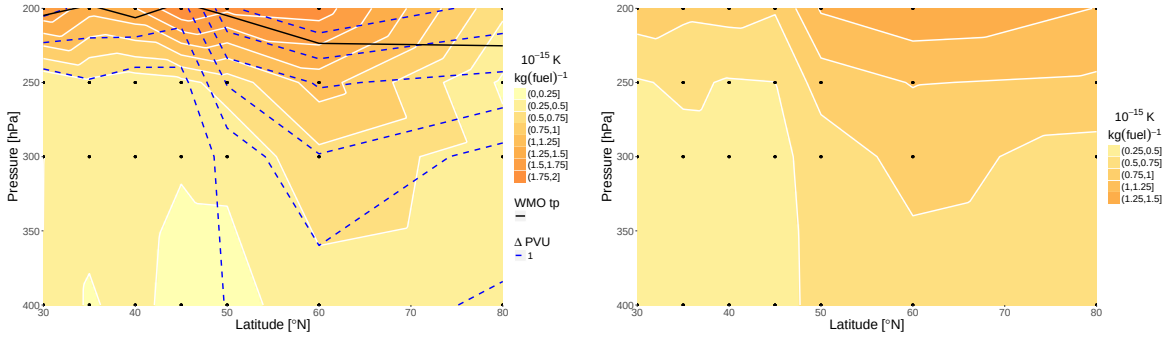
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.34: Water vapour CCF/aCCF contour plots, 12 UTC, 345 °E, WP 2



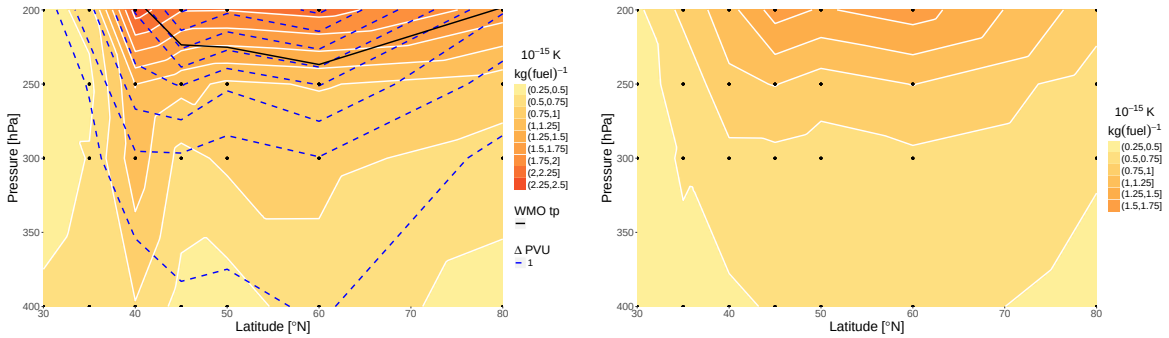
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.35: Water vapour CCF/aCCF contour plots, 12 UTC, 345 °E, WP 3



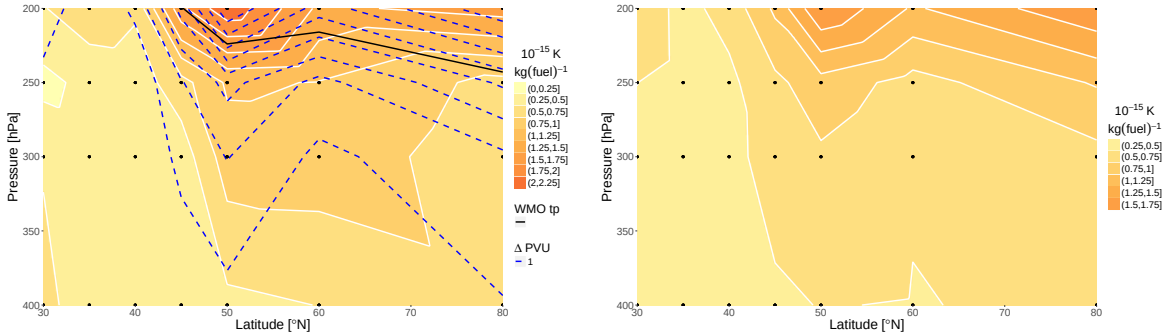
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.36: Water vapour CCF/aCCF contour plots, 12 UTC, 345 °E, WP 4



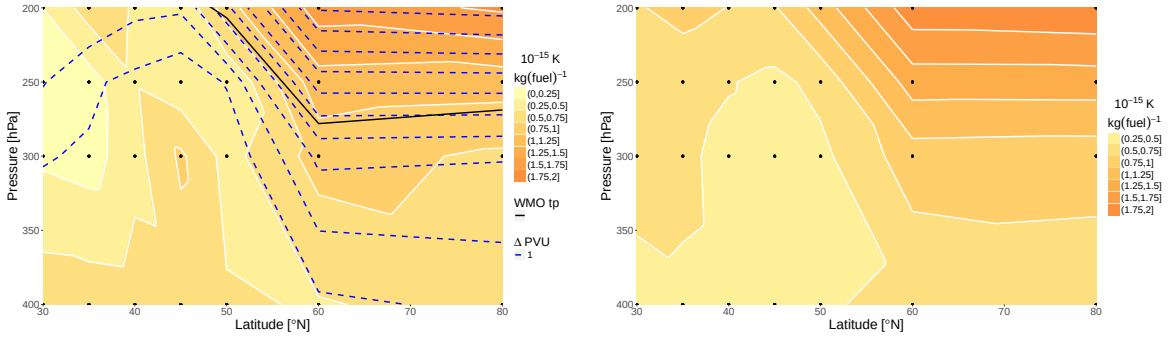
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.37: Water vapour CCF/aCCF contour plots, 12 UTC, 345 °E, WP 5



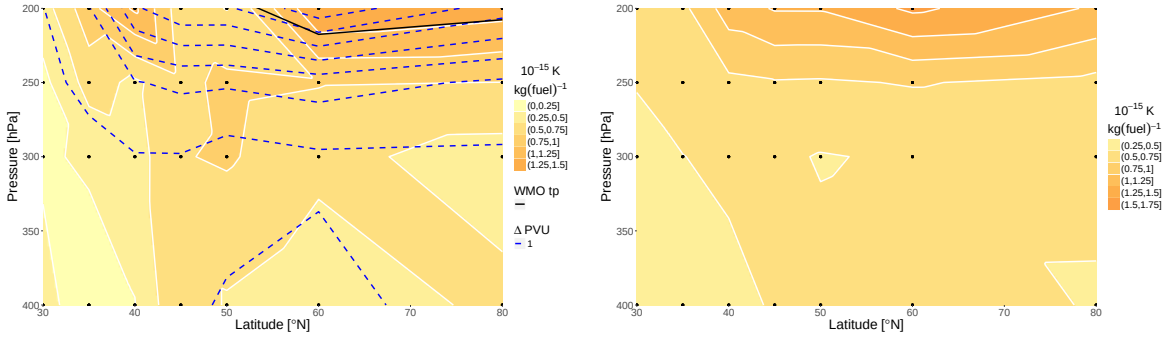
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.38: Water vapour CCF/aCCF contour plots, 12 UTC, 345 °E, WP 6



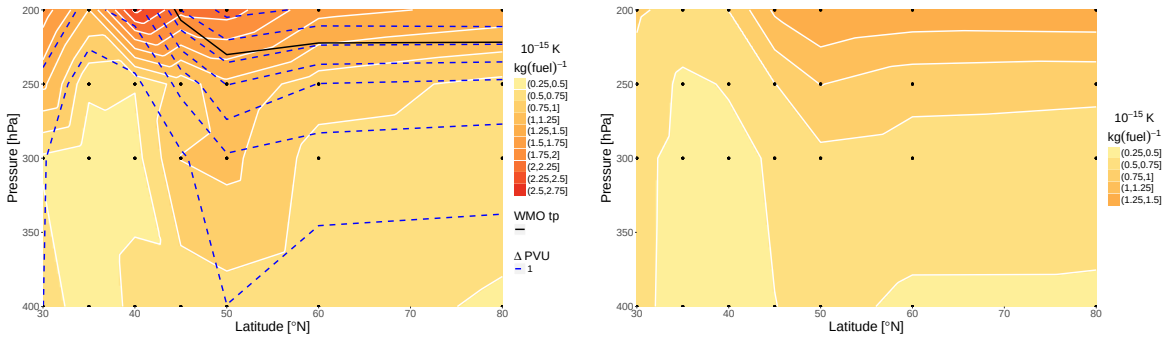
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.39: Water vapour CCF/aCCF contour plots, 12 UTC, 345 °E, WP 7



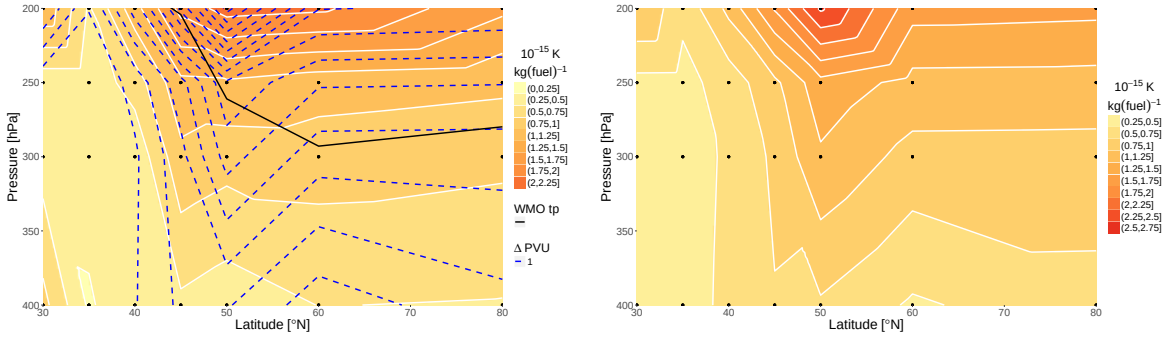
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.40: Water vapour CCF/aCCF contour plots, 12 UTC, 345 °E, WP 8



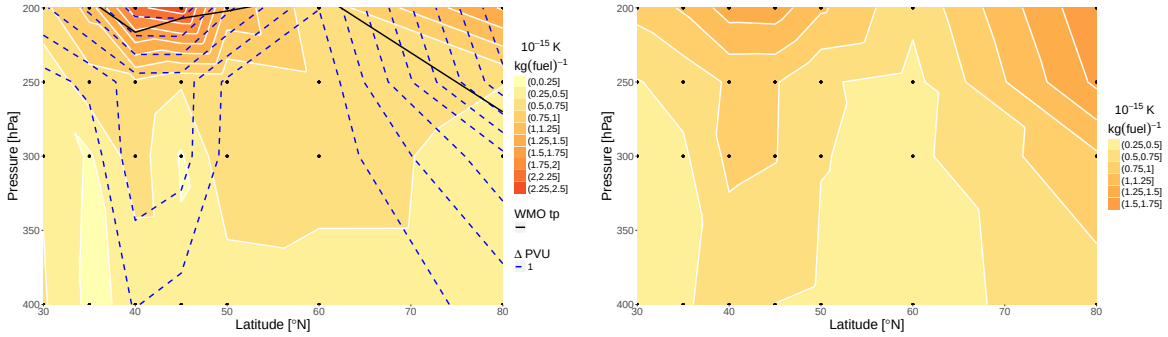
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.41: Water vapour CCF/aCCF contour plots, 12 UTC, 360 °E, WP 1



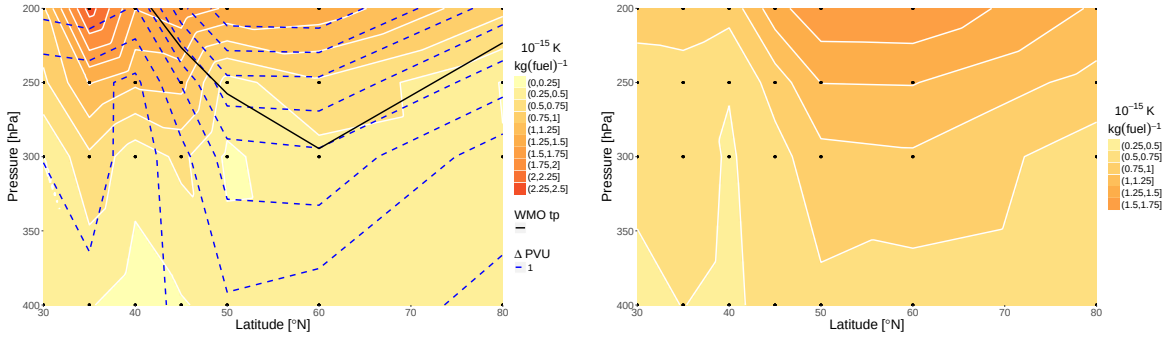
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.42: Water vapour CCF/aCCF contour plots, 12 UTC, 360 °E, WP 2



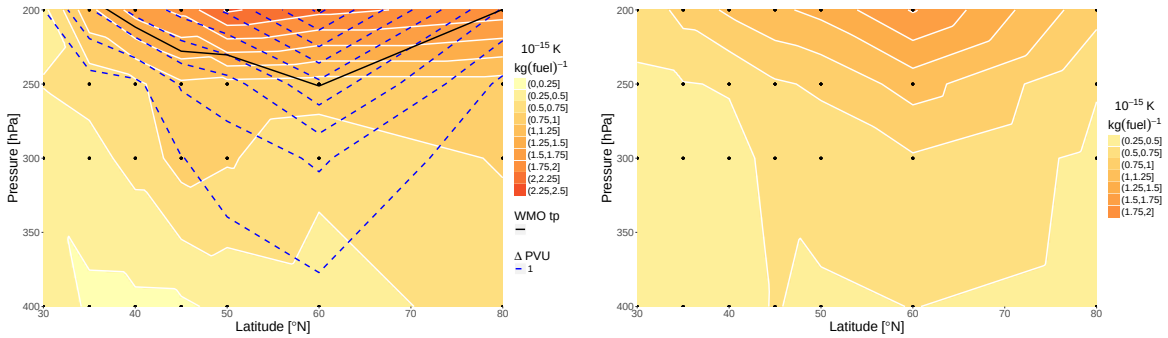
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.43: Water vapour CCF/aCCF contour plots, 12 UTC, 360 °E, WP 3



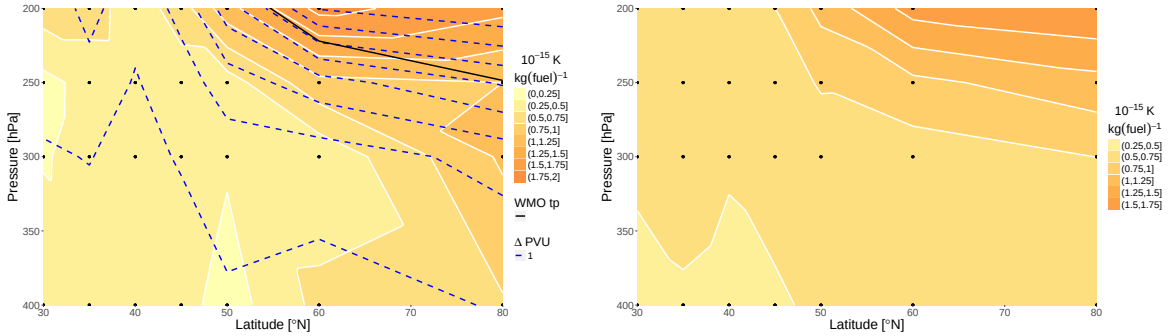
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.44: Water vapour CCF/aCCF contour plots, 12 UTC, 360 °E, WP 4



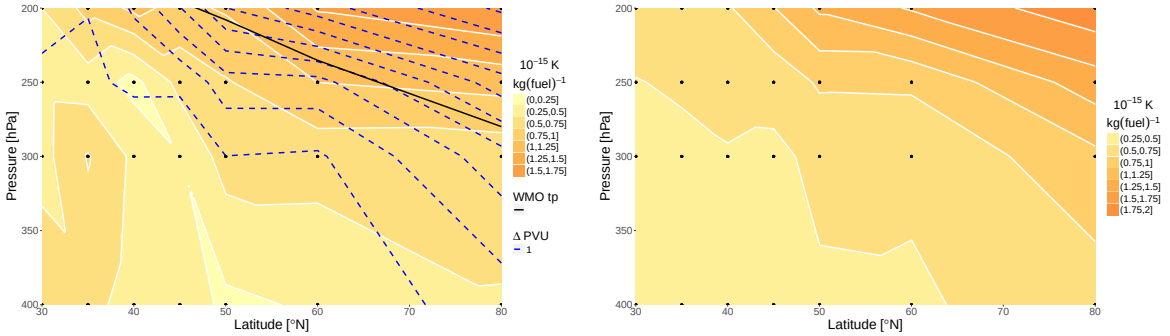
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.45: Water vapour CCF/aCCF contour plots, 12 UTC, 360 °E, WP 5



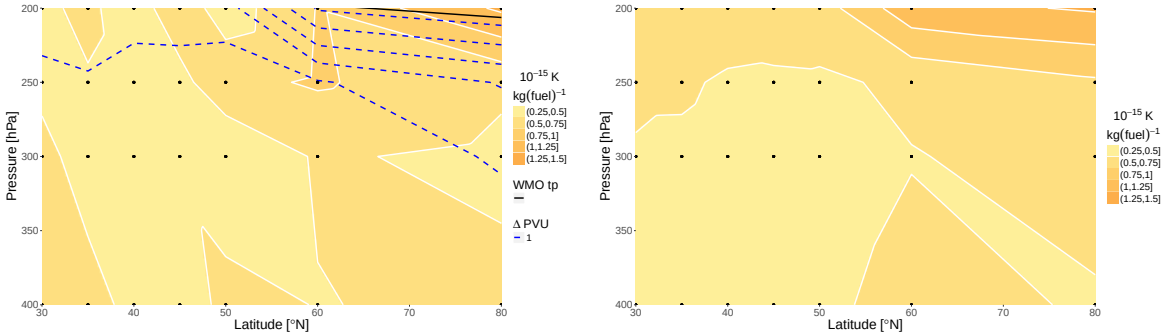
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.46: Water vapour CCF/aCCF contour plots, 12 UTC, 360 °E, WP 6



(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.47: Water vapour CCF/aCCF contour plots, 12 UTC, 360 °E, WP 7



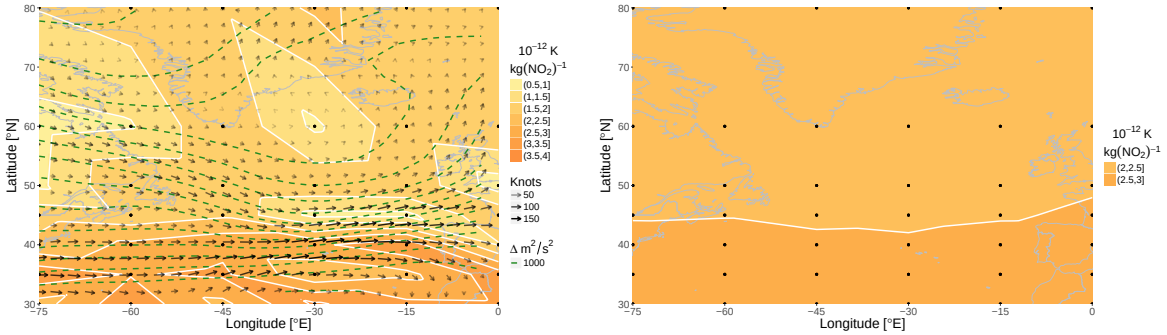
(a) REACT4C data (b) Final $|PV|$ algorithm

Figure B.48: Water vapour CCF/aCCF contour plots, 12 UTC, 360 °E, WP 8

C

Ozone CCF pressure heatmaps

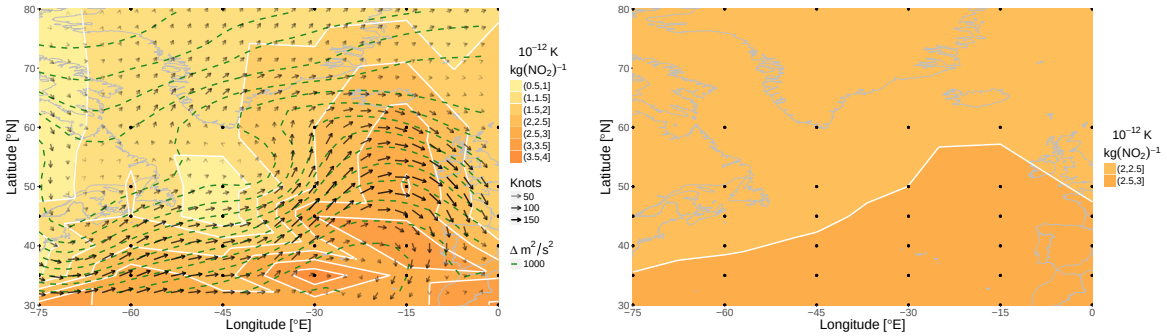
This appendix contains longitude/latitude heatmaps of the original and ozone CCFs. Note that the colour coding is the same for every plot, but the legends only show colours used in that specific plot. The black points are the initial REACT4C tracer emission locations. Bilinearly interpolated contours are shown because application of the CCFs would likely use bilinear interpolation. The dark green dotted lines are geopotential isolines in [1000 m^2 s^{-2}], and the magnitude of the wind vectors in [knots] is reflected by the length and opacity of the vectors.



(a) REACT4C data

(b) Final $geopot \times tm1$ algorithm

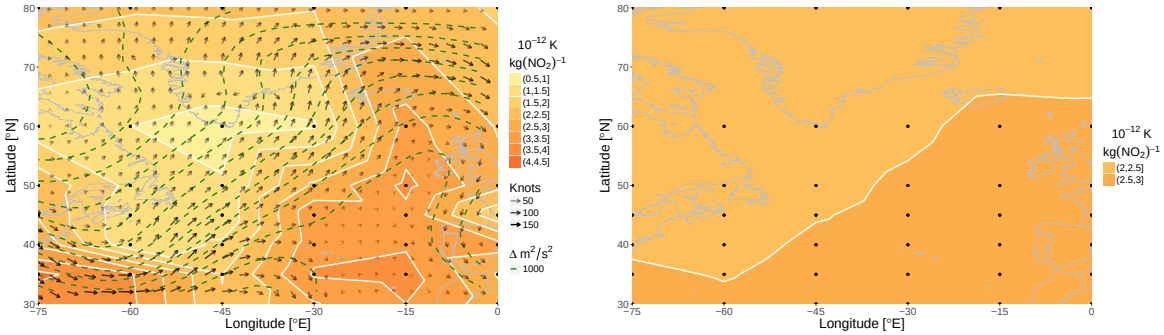
Figure C.1: Ozone CCF/aCCF contour plots, 12 UTC, 200 hPa, WP 1



(a) REACT4C data

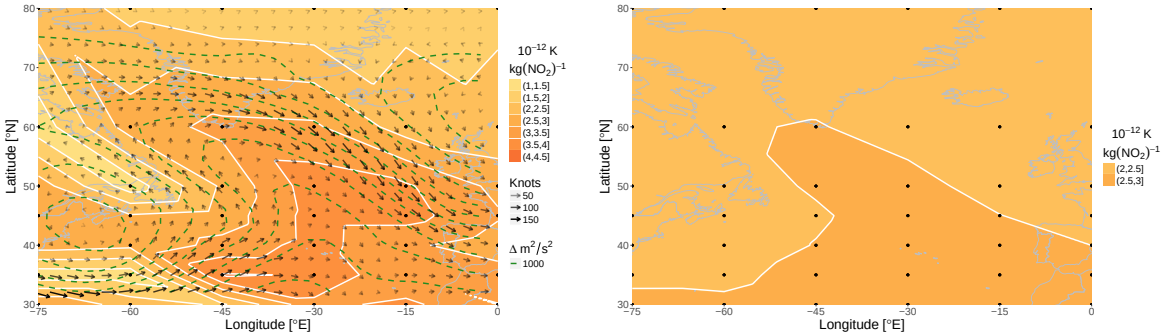
(b) Final $geopot \times tm1$ algorithm

Figure C.2: Ozone CCF/aCCF contour plots, 12 UTC, 200 hPa, WP 2



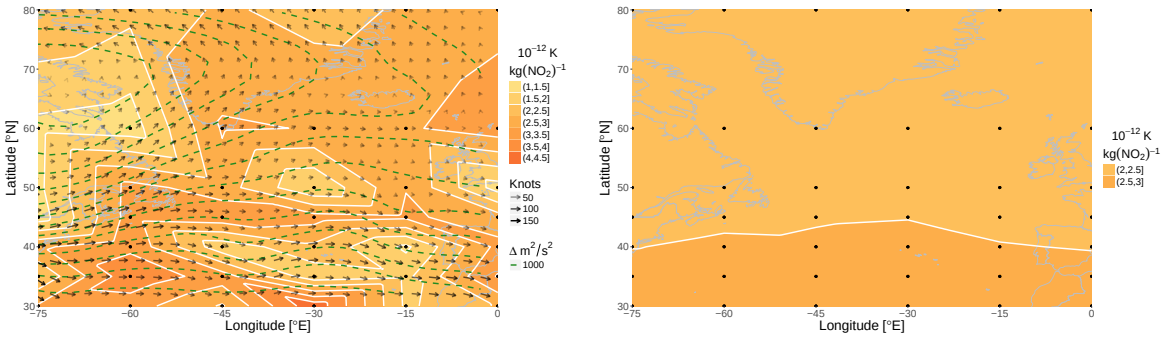
(a) REACT4C data (b) Final $geopot \times tm1$ algorithm

Figure C.3: Ozone CCF/aCCF contour plots, 12 UTC, 200 hPa, WP 3



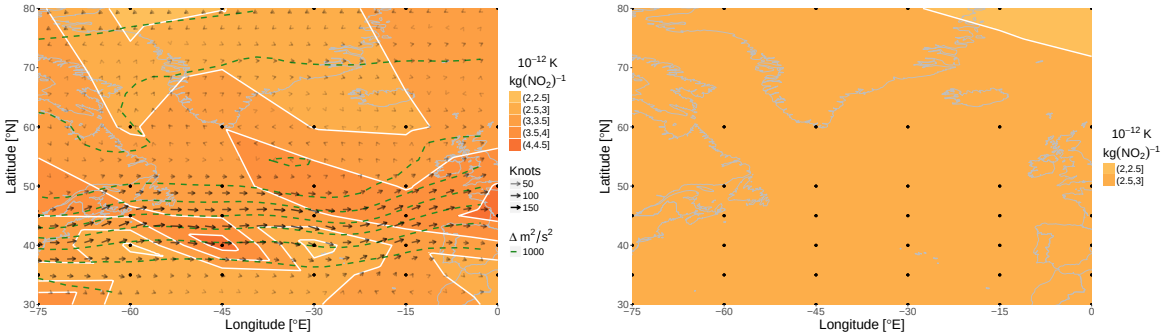
(a) REACT4C data (b) Final $geopot \times tm1$ algorithm

Figure C.4: Ozone CCF/aCCF contour plots, 12 UTC, 200 hPa, WP 4



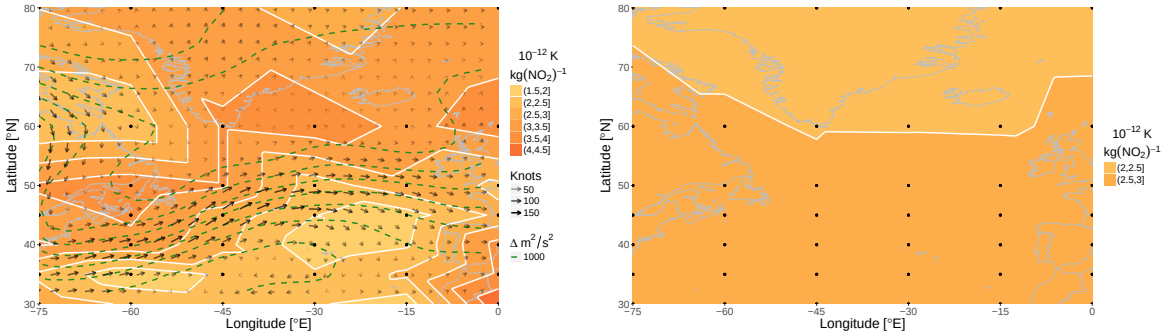
(a) REACT4C data (b) Final $geopot \times tm1$ algorithm

Figure C.5: Ozone CCF/aCCF contour plots, 12 UTC, 200 hPa, WP 5



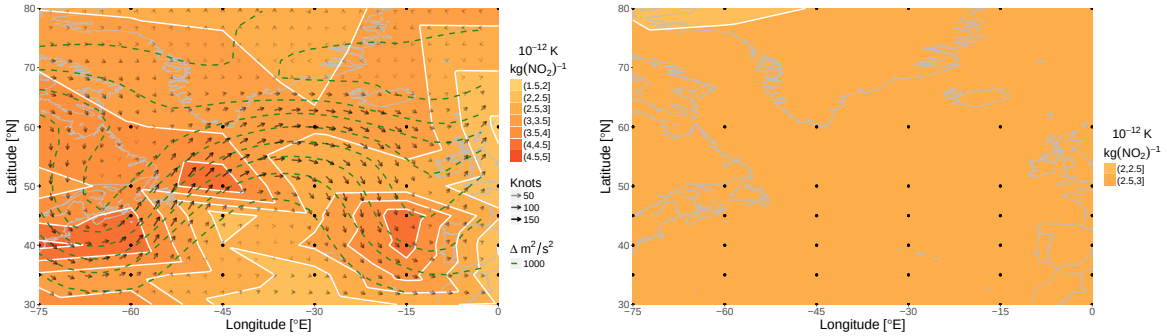
(a) REACT4C data (b) Final $geopot \times tm1$ algorithm

Figure C.6: Ozone CCF/aCCF contour plots, 12 UTC, 200 hPa, WP 6



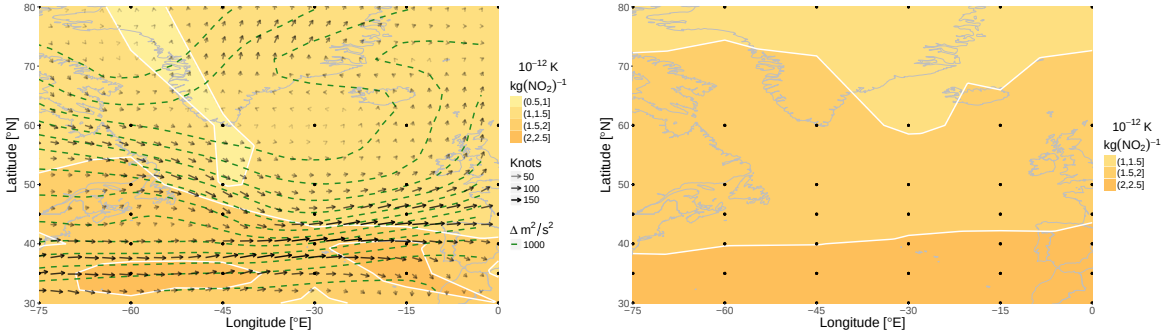
(a) REACT4C data (b) Final $geopot \times tm1$ algorithm

Figure C.7: Ozone CCF/aCCF contour plots, 12 UTC, 200 hPa, WP 7



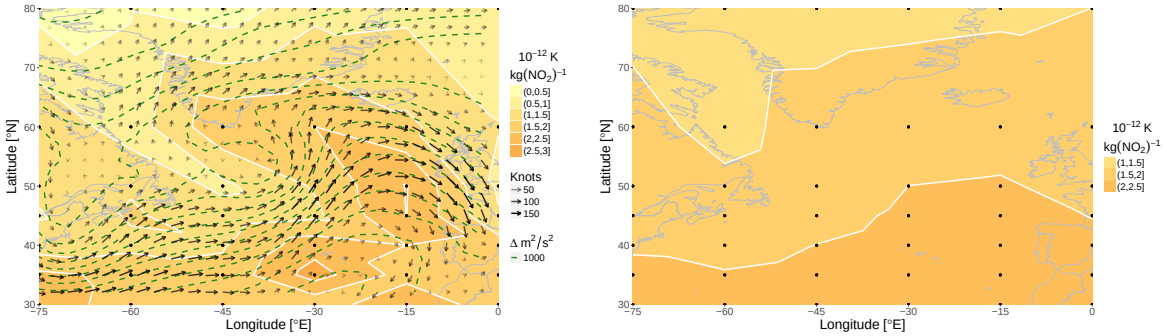
(a) REACT4C data (b) Final $geopot \times tm1$ algorithm

Figure C.8: Ozone CCF/aCCF contour plots, 12 UTC, 200 hPa, WP 8



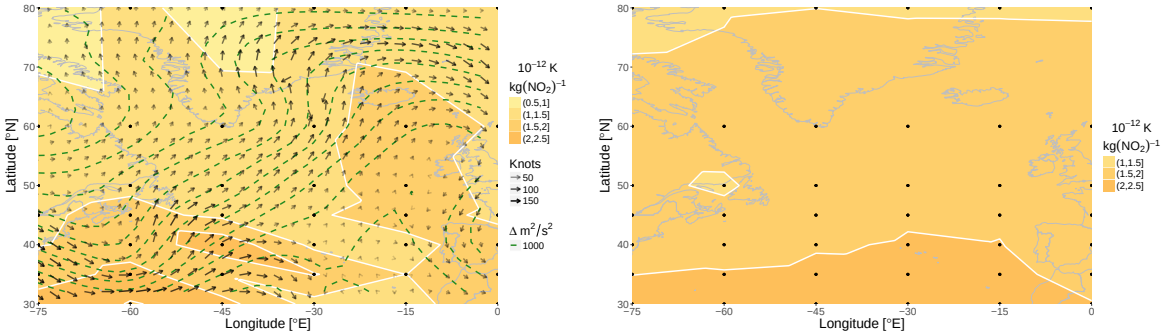
(a) REACT4C data (b) Final $geopot \times tm1$ algorithm

Figure C.9: Ozone CCF/aCCF contour plots, 12 UTC, 250 hPa, WP 1



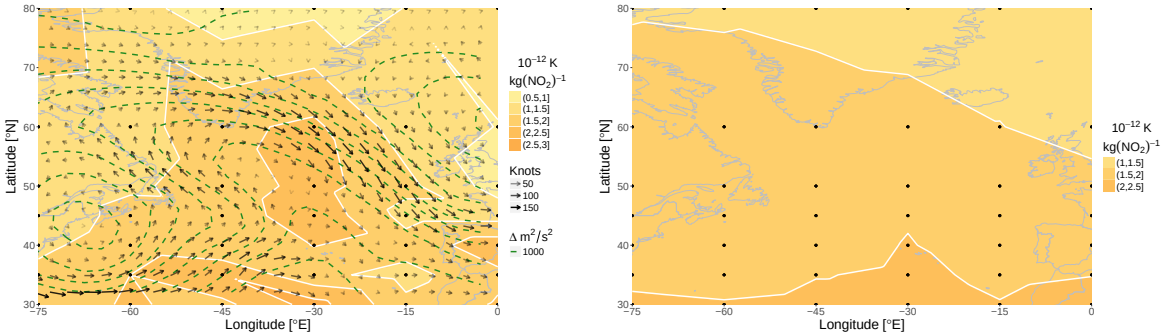
(a) REACT4C data (b) Final $geopot \times tm1$ algorithm

Figure C.10: Ozone CCF/aCCF contour plots, 12 UTC, 250 hPa, WP 2



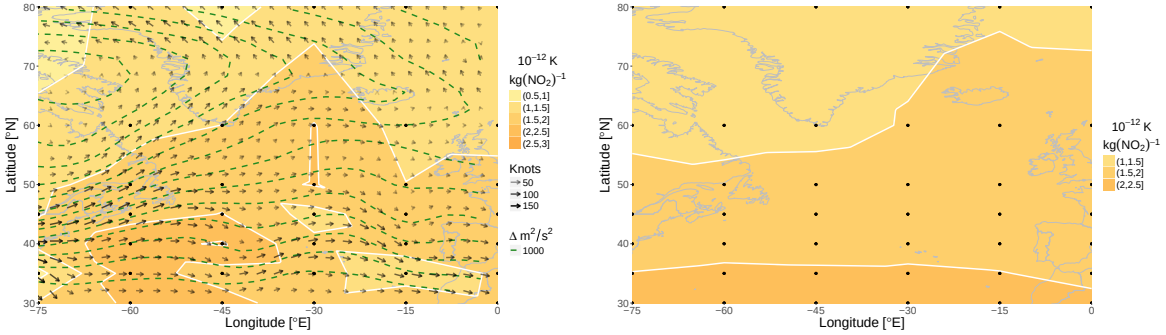
(a) REACT4C data (b) Final $geopot \times tm1$ algorithm

Figure C.11: Ozone CCF/aCCF contour plots, 12 UTC, 250 hPa, WP 3



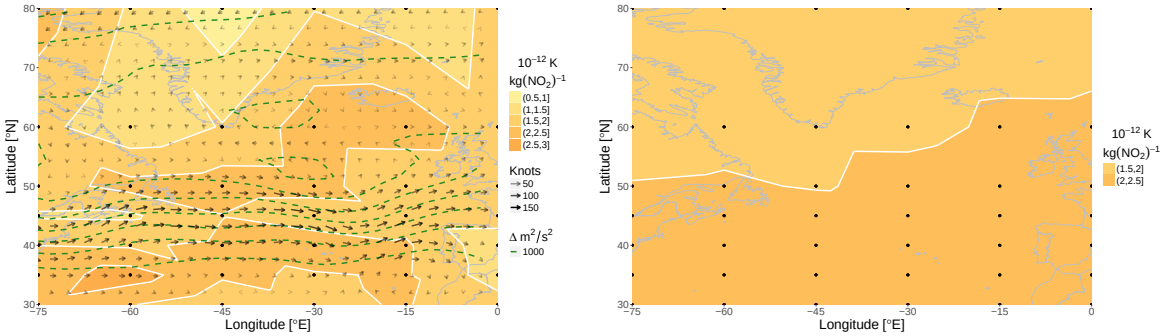
(a) REACT4C data (b) Final $geopot \times tm1$ algorithm

Figure C.12: Ozone CCF/aCCF contour plots, 12 UTC, 250 hPa, WP 4



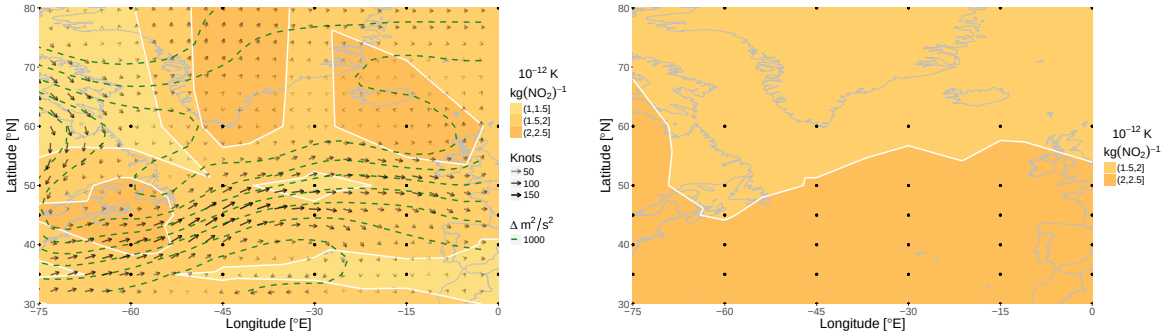
(a) REACT4C data (b) Final $geopot \times tm1$ algorithm

Figure C.13: Ozone CCF/aCCF contour plots, 12 UTC, 250 hPa, WP 5



(a) REACT4C data (b) Final $geopot \times tm1$ algorithm

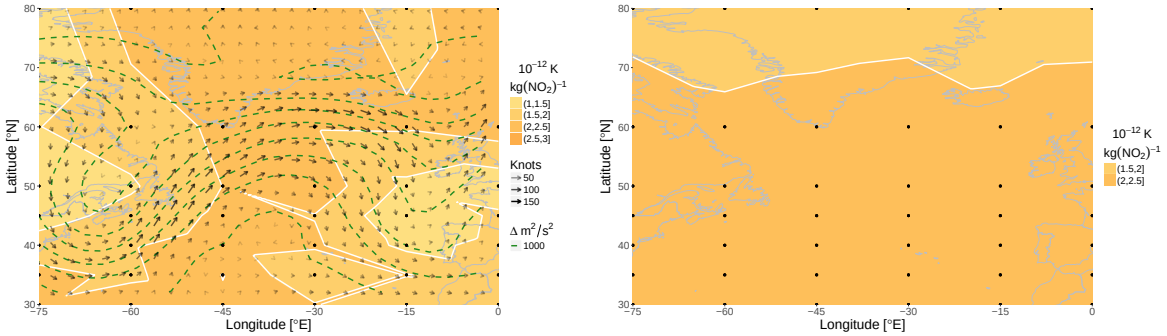
Figure C.14: Ozone CCF/aCCF contour plots, 12 UTC, 250 hPa, WP 6



(a) REACT4C data

(b) Final $geopot \times tm1$ algorithm

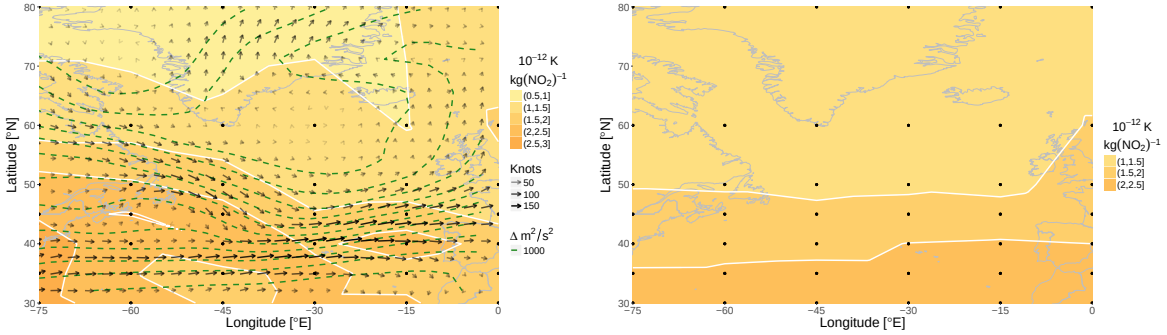
Figure C.15: Ozone CCF/aCCF contour plots, 12 UTC, 250 hPa, WP 7



(a) REACT4C data

(b) Final $geopot \times tm1$ algorithm

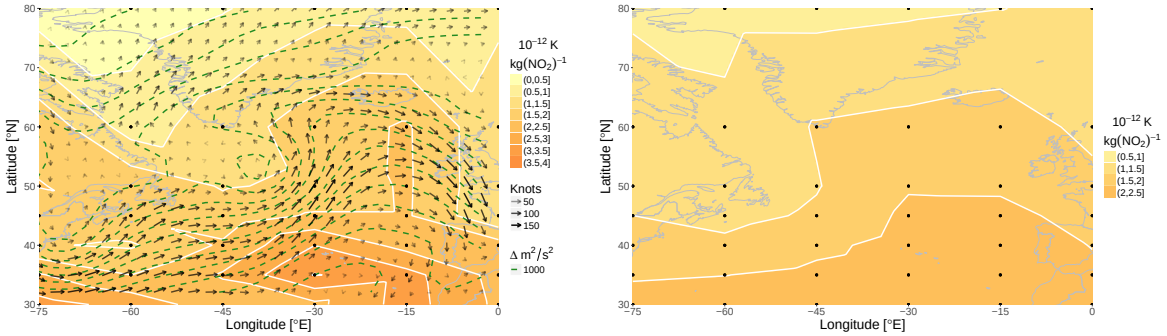
Figure C.16: Ozone CCF/aCCF contour plots, 12 UTC, 250 hPa, WP 8



(a) REACT4C data

(b) Final $geopot \times tm1$ algorithm

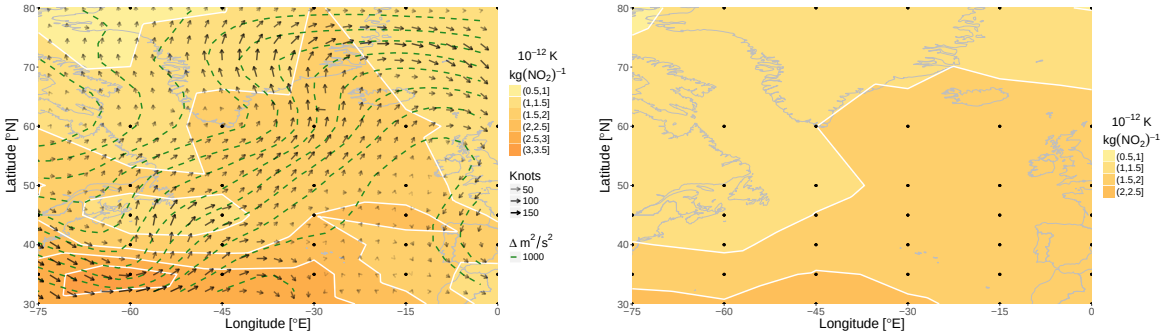
Figure C.17: Ozone CCF/aCCF contour plots, 12 UTC, 300 hPa, WP 1



(a) REACT4C data

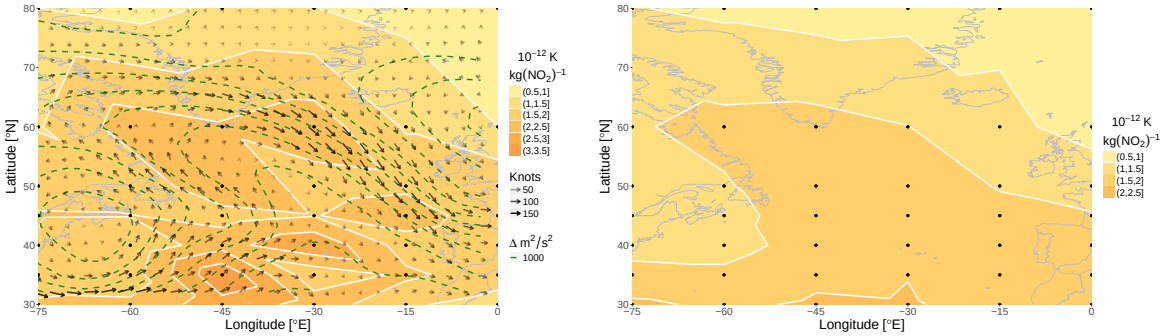
(b) Final $geopot \times tm1$ algorithm

Figure C.18: Ozone CCF/aCCF contour plots, 12 UTC, 300 hPa, WP 2



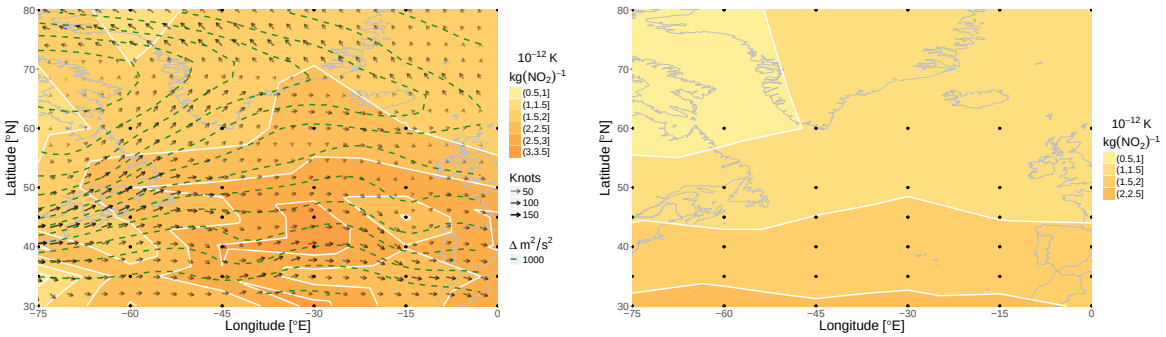
(a) REACT4C data (b) Final $geopot \times tm1$ algorithm

Figure C.19: Ozone CCF/aCCF contour plots, 12 UTC, 300 hPa, WP 3



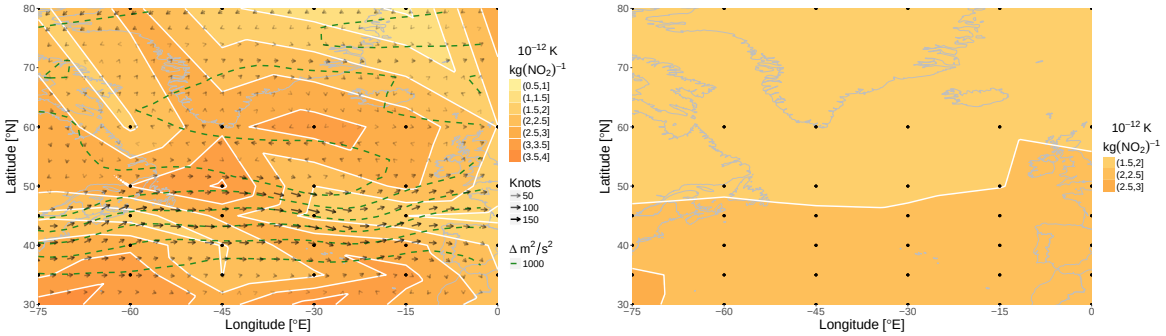
(a) REACT4C data (b) Final $geopot \times tm1$ algorithm

Figure C.20: Ozone CCF/aCCF contour plots, 12 UTC, 300 hPa, WP 4



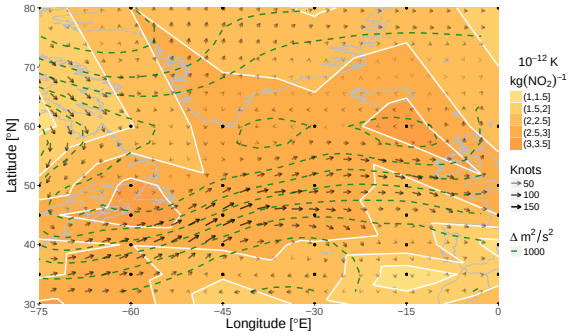
(a) REACT4C data (b) Final $geopot \times tm1$ algorithm

Figure C.21: Ozone CCF/aCCF contour plots, 12 UTC, 300 hPa, WP 5

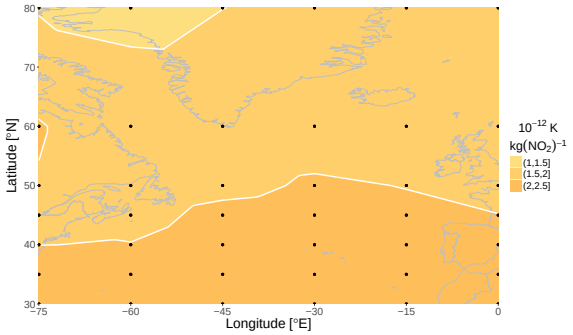


(a) REACT4C data (b) Final $geopot \times tm1$ algorithm

Figure C.22: Ozone CCF/aCCF contour plots, 12 UTC, 300 hPa, WP 6

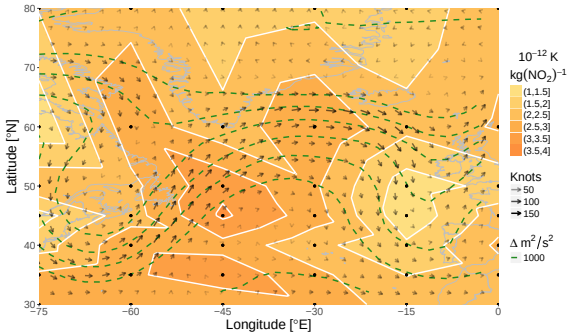


(a) REACT4C data

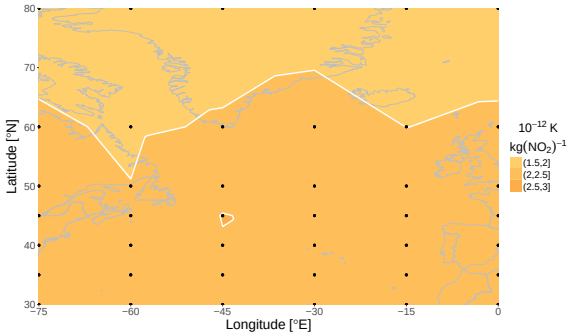


(b) Final $geopot \times tm1$ algorithm

Figure C.23: Ozone CCF/aCCF contour plots, 12 UTC, 300 hPa, WP 7

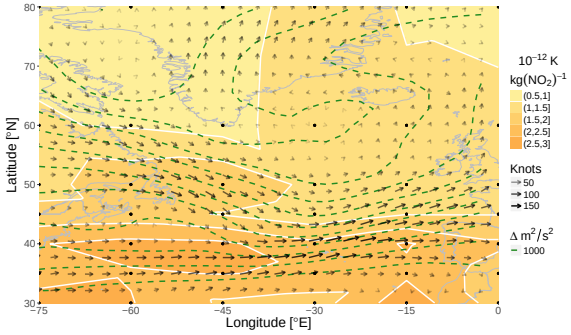


(a) REACT4C data

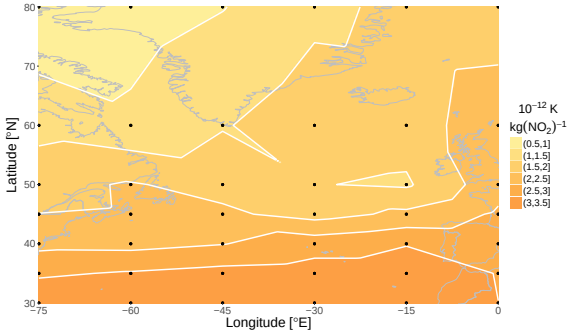


(b) Final $geopot \times tm1$ algorithm

Figure C.24: Ozone CCF/aCCF contour plots, 12 UTC, 300 hPa, WP 8

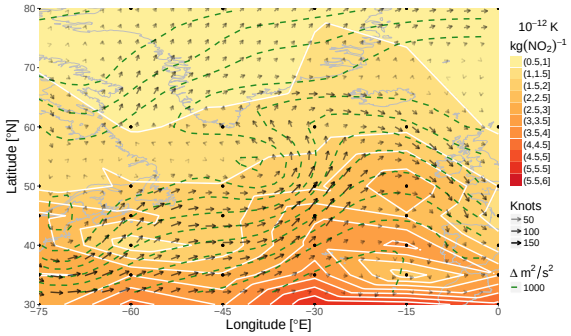


(a) REACT4C data

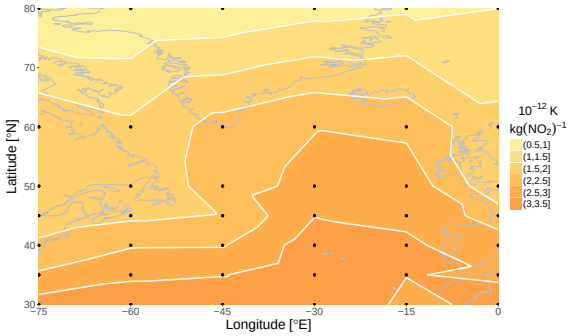


(b) Final $geopot \times tm1$ algorithm

Figure C.25: Ozone CCF/aCCF contour plots, 12 UTC, 400 hPa, WP 1

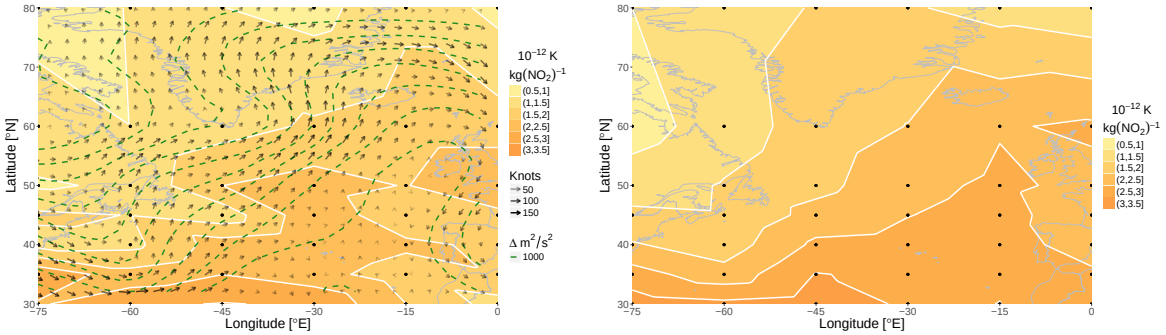


(a) REACT4C data



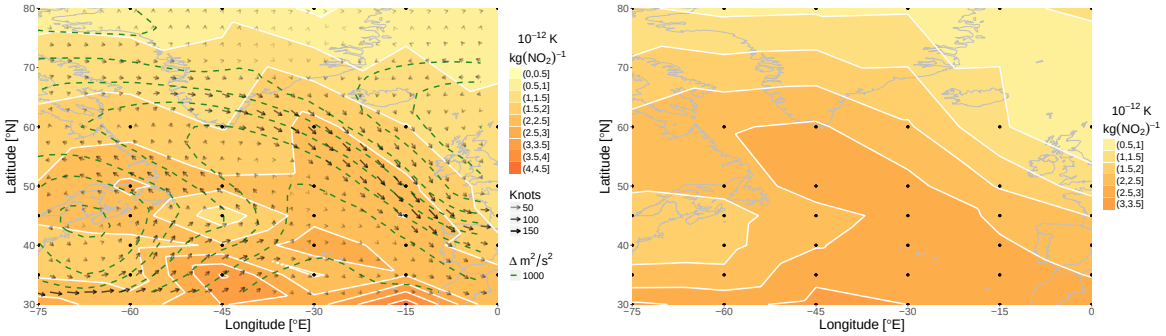
(b) Final $geopot \times tm1$ algorithm

Figure C.26: Ozone CCF/aCCF contour plots, 12 UTC, 400 hPa, WP 2



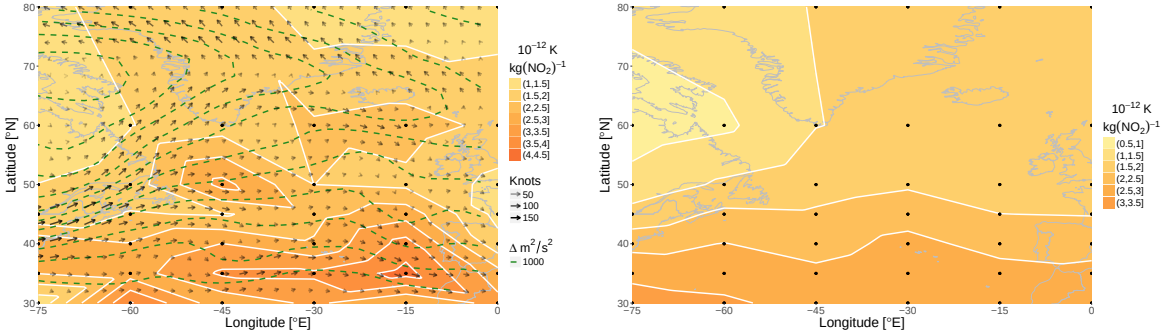
(a) REACT4C data (b) Final $geopot \times tm1$ algorithm

Figure C.27: Ozone CCF/aCCF contour plots, 12 UTC, 400 hPa, WP 3



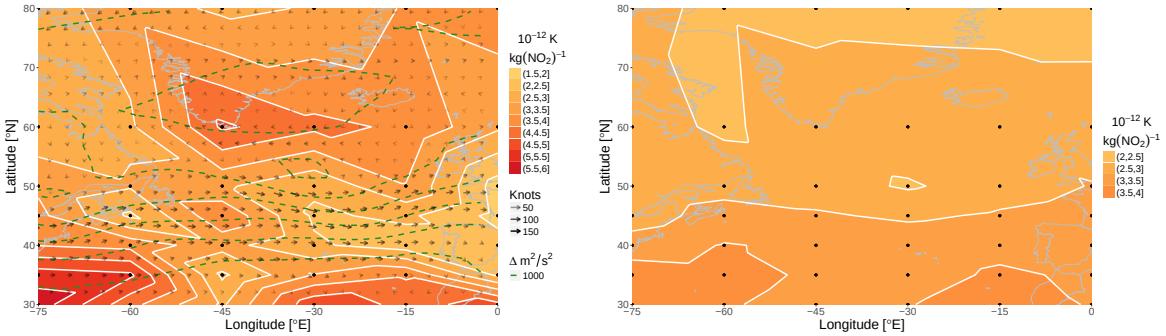
(a) REACT4C data (b) Final $geopot \times tm1$ algorithm

Figure C.28: Ozone CCF/aCCF contour plots, 12 UTC, 400 hPa, WP 4



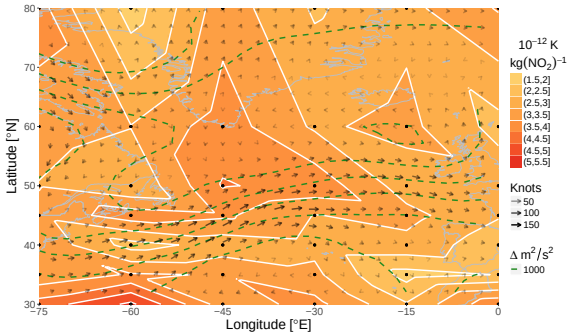
(a) REACT4C data (b) Final $geopot \times tm1$ algorithm

Figure C.29: Ozone CCF/aCCF contour plots, 12 UTC, 400 hPa, WP 5

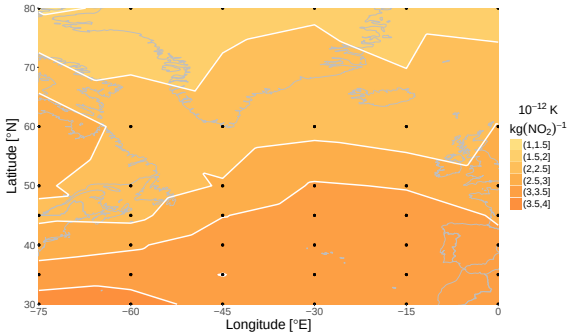


(a) REACT4C data (b) Final $geopot \times tm1$ algorithm

Figure C.30: Ozone CCF/aCCF contour plots, 12 UTC, 400 hPa, WP 6

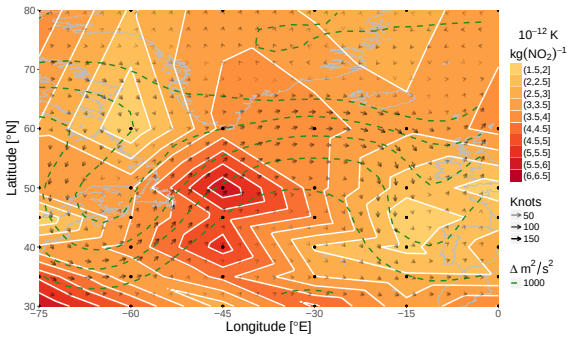


(a) REACT4C data

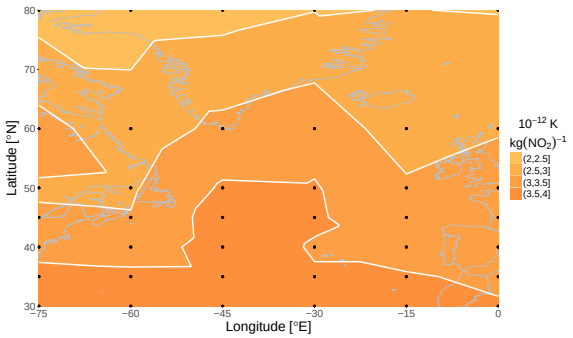


(b) Final $geopot \times tm1$ algorithm

Figure C.31: Ozone CCF/aCCF contour plots, 12 UTC, 400 hPa, WP 7



(a) REACT4C data



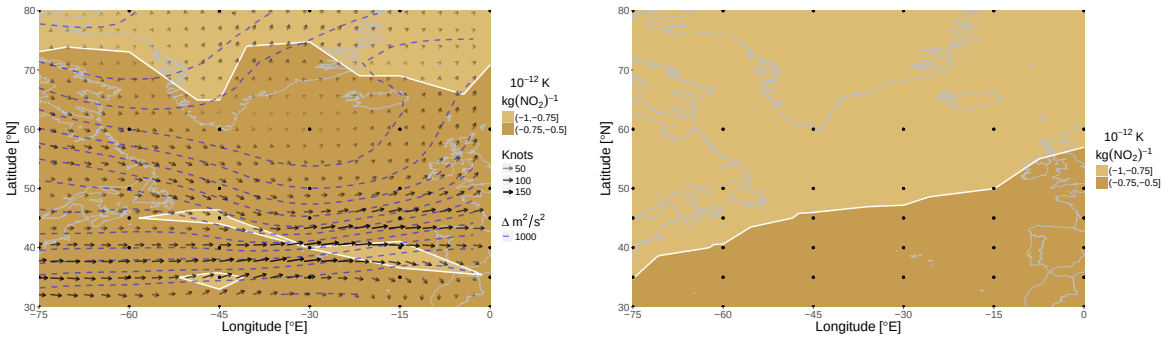
(b) Final $geopot \times tm1$ algorithm

Figure C.32: Ozone CCF/aCCF contour plots, 12 UTC, 400 hPa, WP 8

D

Methane CCF pressure heatmaps

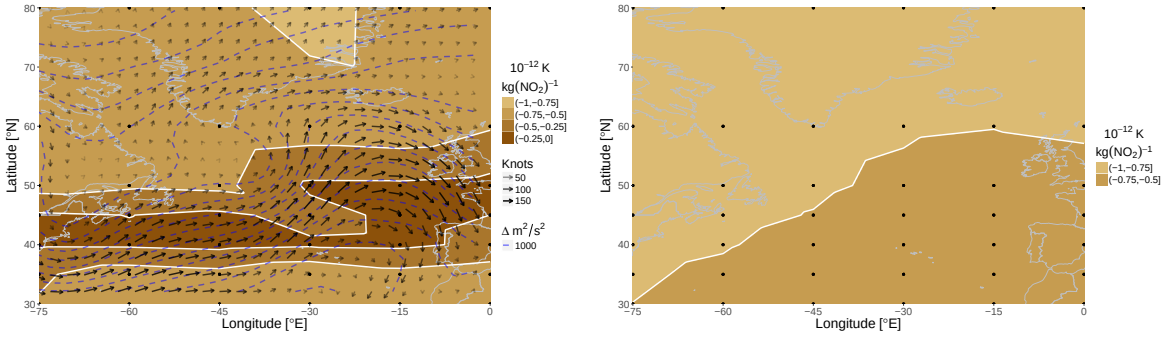
This appendix contains longitude/latitude heatmaps of the original and algorithmic methane CCFs. Note that the colour coding is the same for every plot, but the legends only show colours used in that specific plot. The black points are the initial REACT4C tracer emission locations. Bilinearly interpolated contours are shown because application of the CCFs would likely use bilinear interpolation. The dark blue dotted lines are geopotential isolines in [1000 m^2 s^{-2}], and the magnitude of the wind vectors in [knots] is reflected by the length and opacity of the vectors. A divergent colour scheme is applied to identify the outliers.



(a) REACT4C data

(b) Final $geopot \times zi0$ algorithm

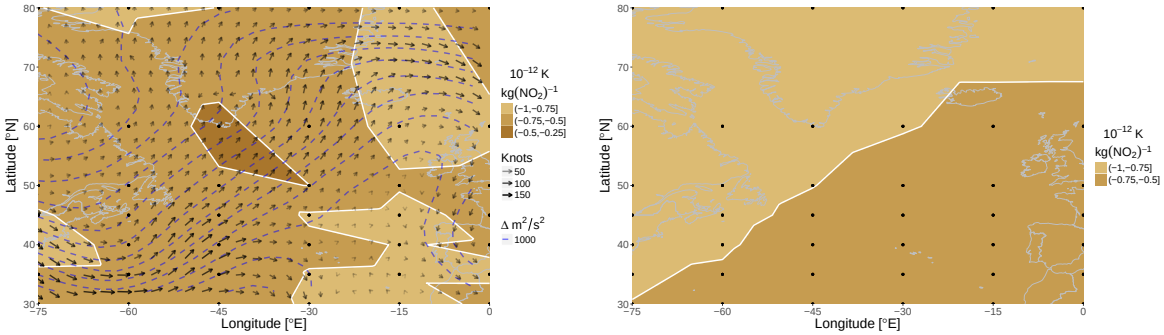
Figure D.1: Methane CCF/aCCF contour plots, 12 UTC, 200 hPa, WP 1



(a) REACT4C data

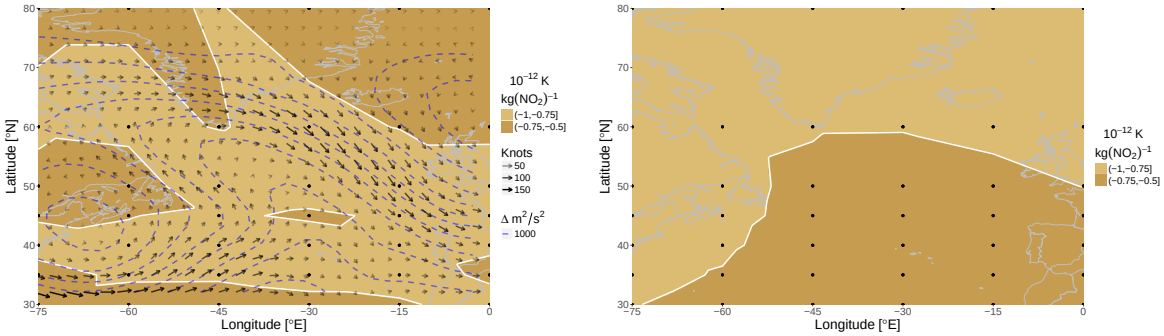
(b) Final $geopot \times zi0$ algorithm

Figure D.2: Methane CCF/aCCF contour plots, 12 UTC, 200 hPa, WP 2



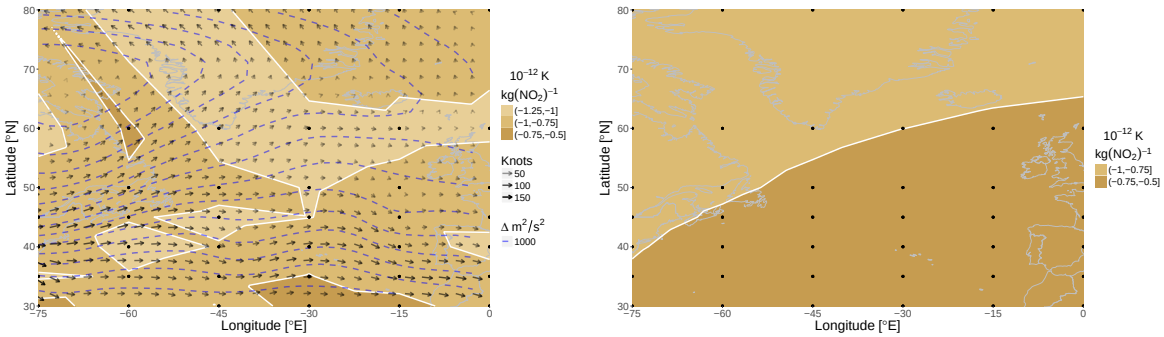
(a) REACT4C data (b) Final $geopot \times zi0$ algorithm

Figure D.3: Methane CCF/aCCF contour plots, 12 UTC, 200 hPa, WP 3



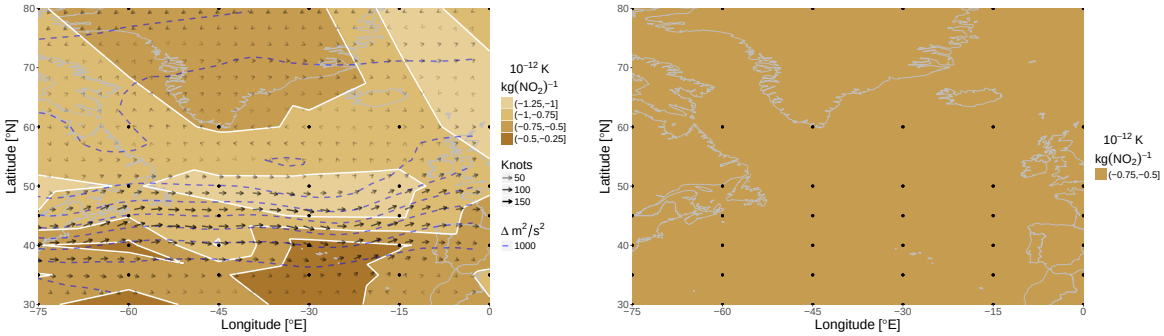
(a) REACT4C data (b) Final $geopot \times zi0$ algorithm

Figure D.4: Methane CCF/aCCF contour plots, 12 UTC, 200 hPa, WP 4



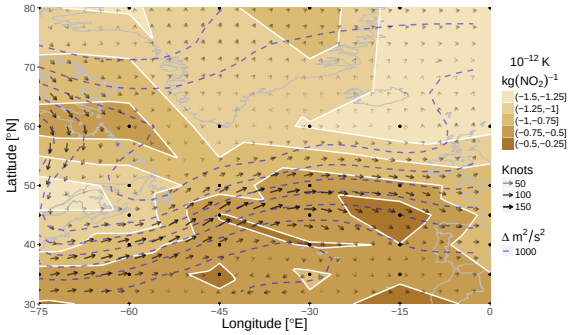
(a) REACT4C data (b) Final $geopot \times zi0$ algorithm

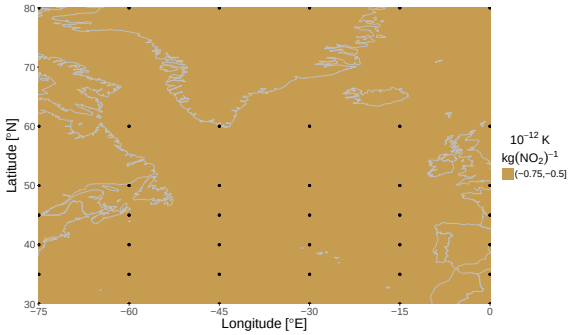
Figure D.5: Methane CCF/aCCF contour plots, 12 UTC, 200 hPa, WP 5



(a) REACT4C data (b) Final $geopot \times zi0$ algorithm

Figure D.6: Methane CCF/aCCF contour plots, 12 UTC, 200 hPa, WP 6

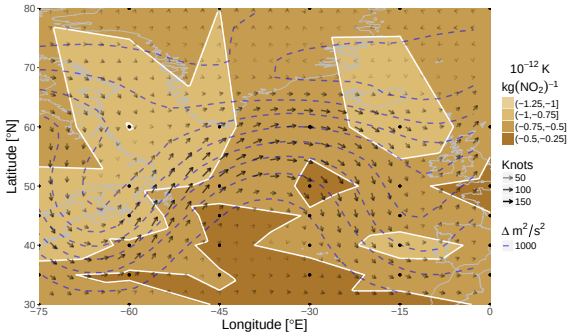


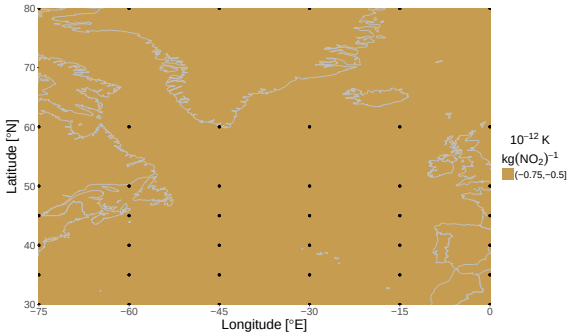


(a) REACT4C data

(b) Final $geopot \times zi0$ algorithm

Figure D.7: Methane CCF/aCCF contour plots, 12 UTC, 200 hPa, WP 7

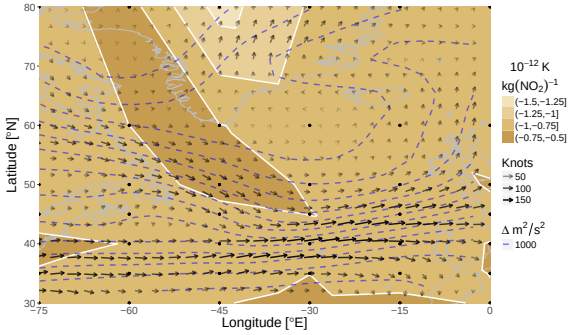


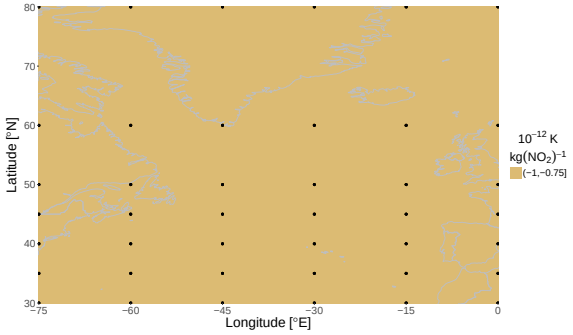


(a) REACT4C data

(b) Final $geopot \times zi0$ algorithm

Figure D.8: Methane CCF/aCCF contour plots, 12 UTC, 200 hPa, WP 8

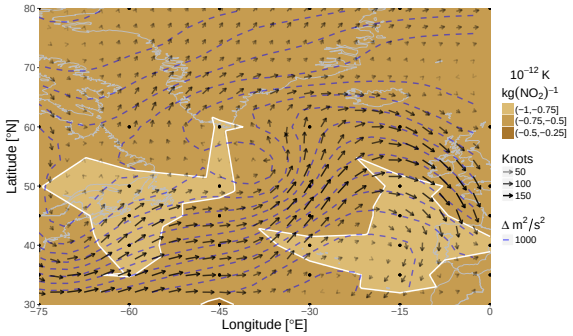


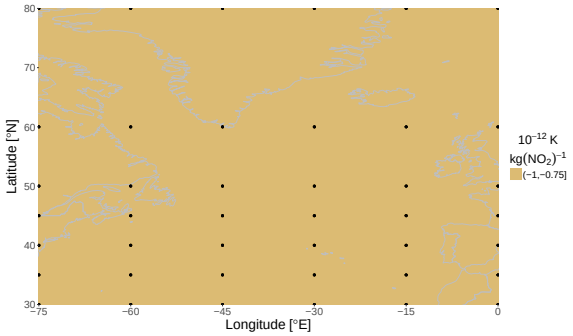


(a) REACT4C data

(b) Final $geopot \times zi0$ algorithm

Figure D.9: Methane CCF/aCCF contour plots, 12 UTC, 250 hPa, WP 1

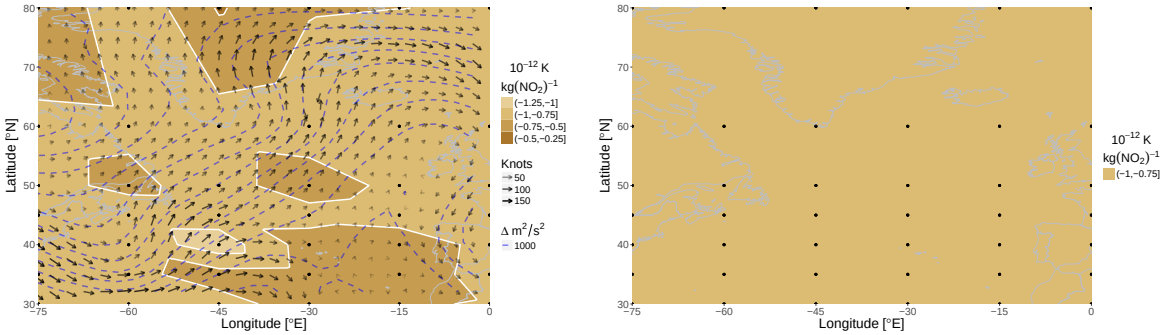




(a) REACT4C data

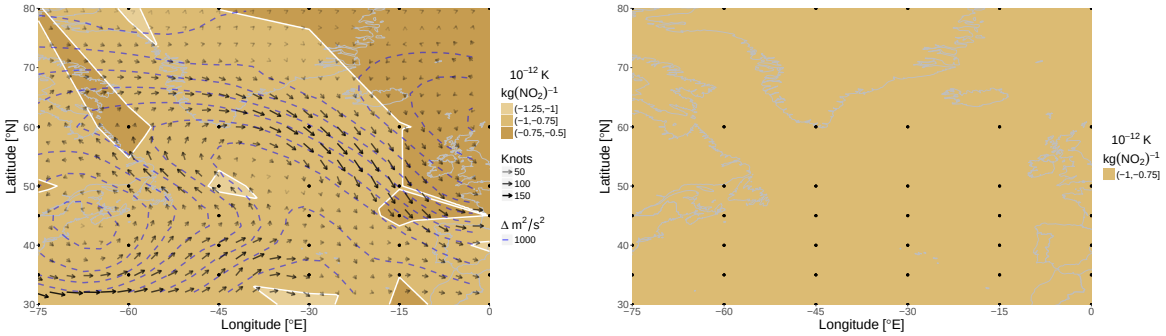
(b) Final $geopot \times zi0$ algorithm

Figure D.10: Methane CCF/aCCF contour plots, 12 UTC, 250 hPa, WP 2



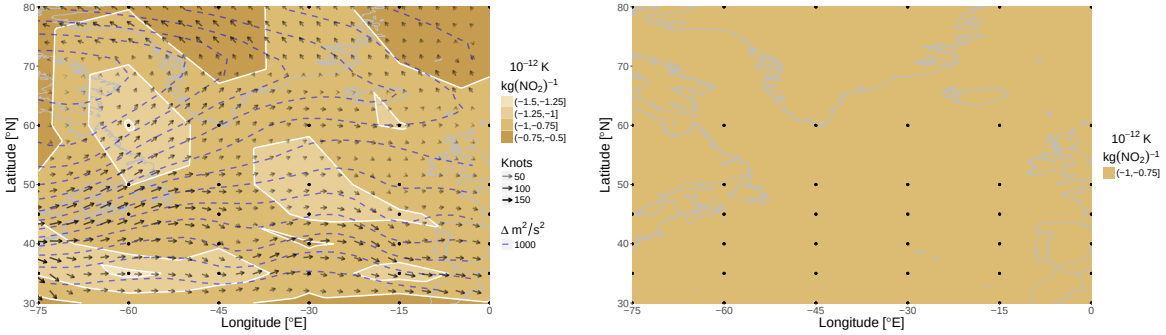
(a) REACT4C data (b) Final $geopot \times zi0$ algorithm

Figure D.11: Methane CCF/aCCF contour plots, 12 UTC, 250 hPa, WP 3



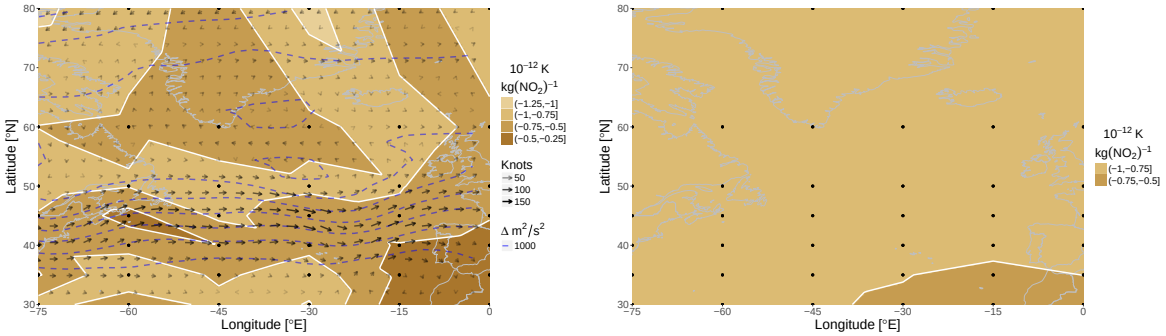
(a) REACT4C data (b) Final $geopot \times zi0$ algorithm

Figure D.12: Methane CCF/aCCF contour plots, 12 UTC, 250 hPa, WP 4



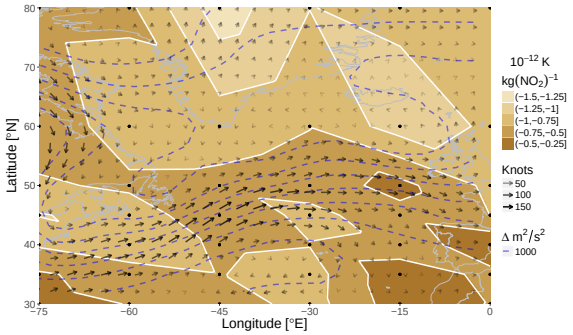
(a) REACT4C data (b) Final $geopot \times zi0$ algorithm

Figure D.13: Methane CCF/aCCF contour plots, 12 UTC, 250 hPa, WP 5

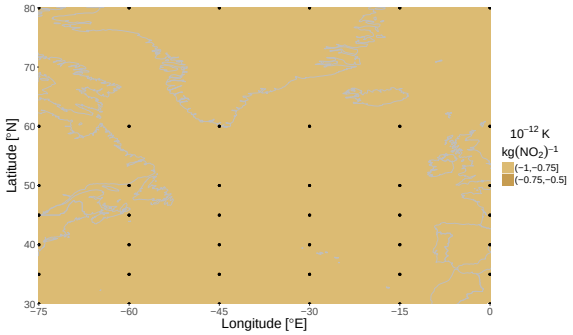


(a) REACT4C data (b) Final $geopot \times zi0$ algorithm

Figure D.14: Methane CCF/aCCF contour plots, 12 UTC, 250 hPa, WP 6

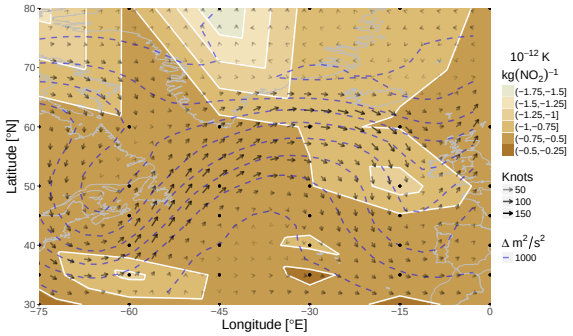


(a) REACT4C data

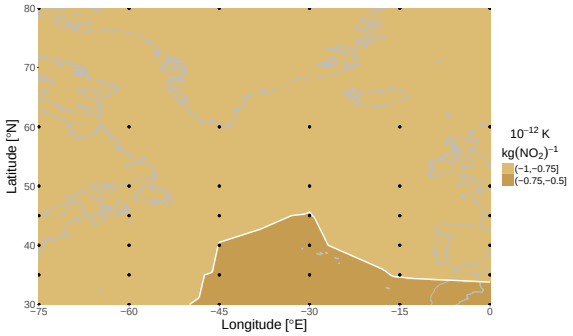


(b) Final $geopot \times zi0$ algorithm

Figure D.15: Methane CCF/aCCF contour plots, 12 UTC, 250 hPa, WP 7

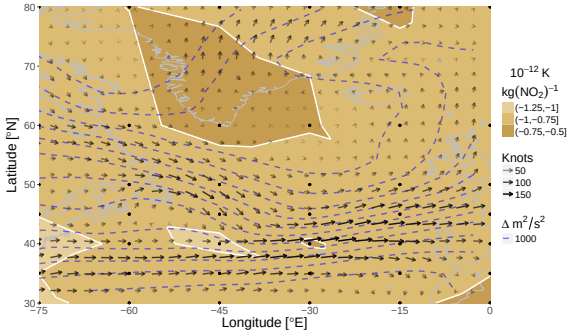


(a) REACT4C data

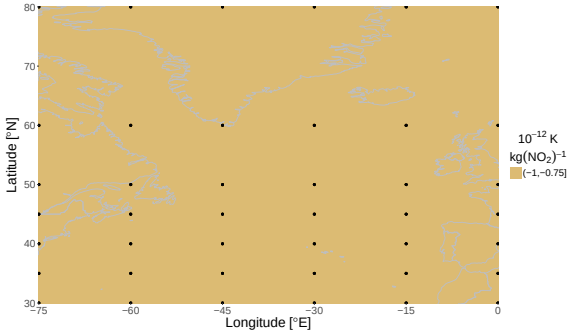


(b) Final $geopot \times zi0$ algorithm

Figure D.16: Methane CCF/aCCF contour plots, 12 UTC, 250 hPa, WP 8

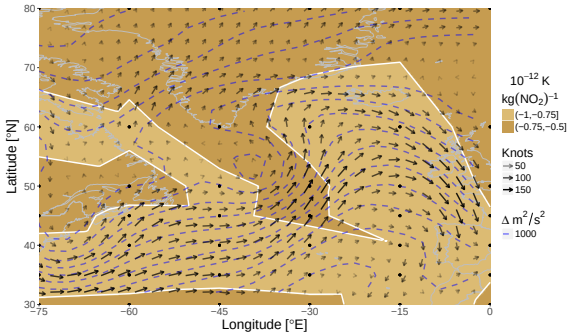


(a) REACT4C data

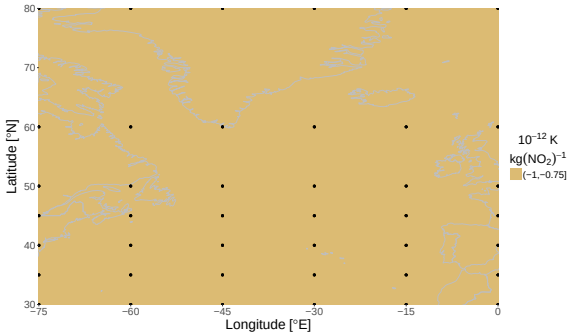


(b) Final $geopot \times zi0$ algorithm

Figure D.17: Methane CCF/aCCF contour plots, 12 UTC, 300 hPa, WP 1

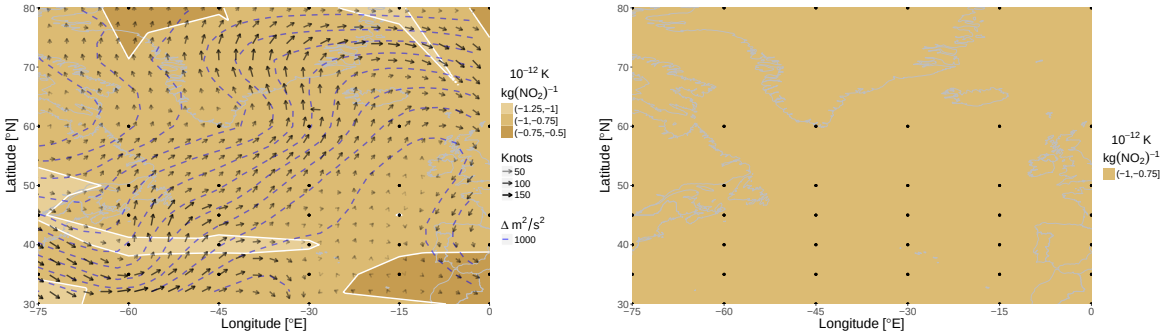


(a) REACT4C data



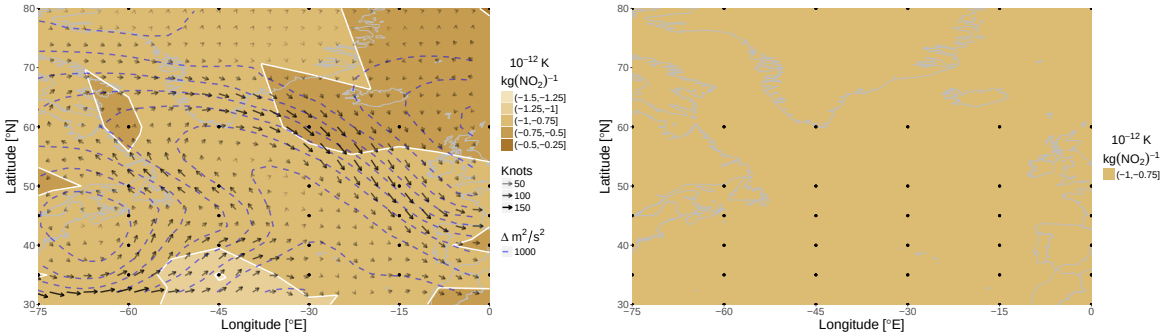
(b) Final $geopot \times zi0$ algorithm

Figure D.18: Methane CCF/aCCF contour plots, 12 UTC, 300 hPa, WP 2



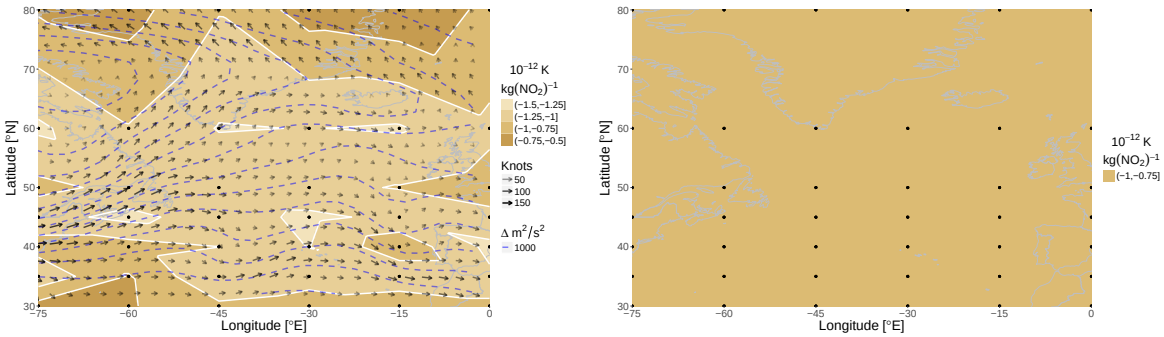
(a) REACT4C data (b) Final $geopot \times zi0$ algorithm

Figure D.19: Methane CCF/aCCF contour plots, 12 UTC, 300 hPa, WP 3



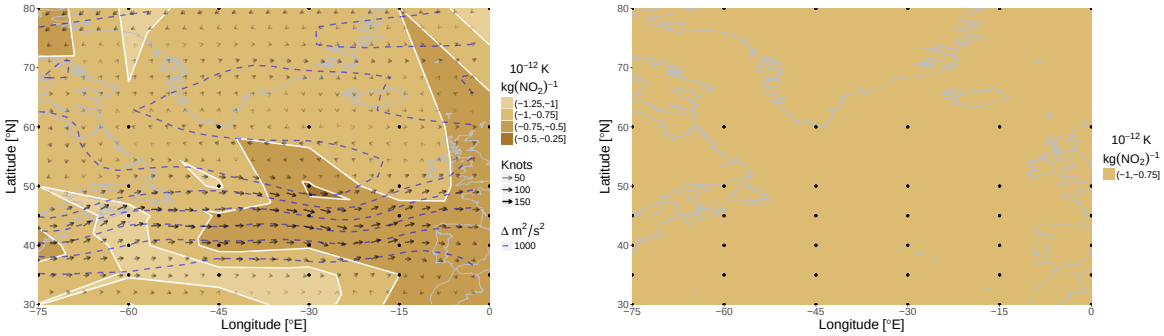
(a) REACT4C data (b) Final $geopot \times zi0$ algorithm

Figure D.20: Methane CCF/aCCF contour plots, 12 UTC, 300 hPa, WP 4



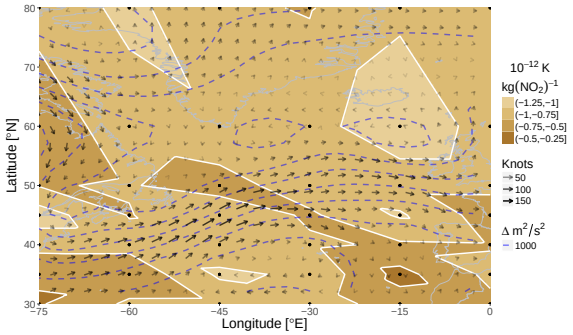
(a) REACT4C data (b) Final $geopot \times zi0$ algorithm

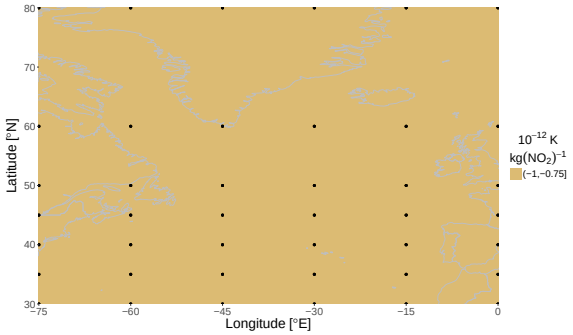
Figure D.21: Methane CCF/aCCF contour plots, 12 UTC, 300 hPa, WP 5



(a) REACT4C data (b) Final $geopot \times zi0$ algorithm

Figure D.22: Methane CCF/aCCF contour plots, 12 UTC, 300 hPa, WP 6

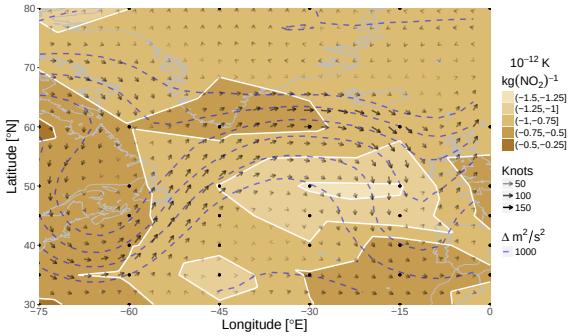


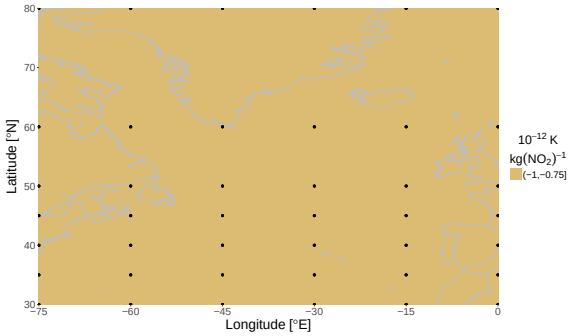


(a) REACT4C data

(b) Final $geopot \times zi0$ algorithm

Figure D.23: Methane CCF/aCCF contour plots, 12 UTC, 300 hPa, WP 7

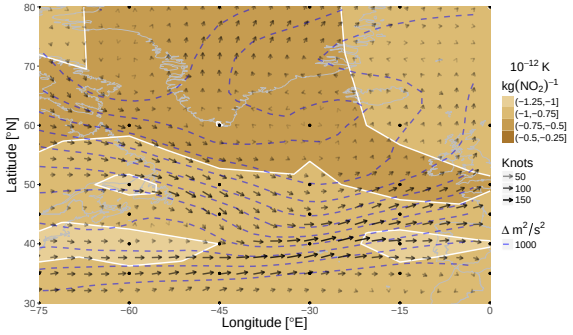


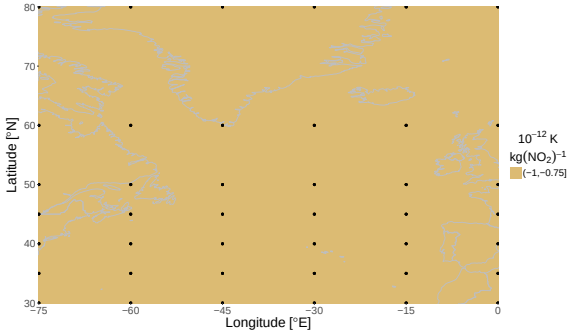


(a) REACT4C data

(b) Final $geopot \times zi0$ algorithm

Figure D.24: Methane CCF/aCCF contour plots, 12 UTC, 300 hPa, WP 8

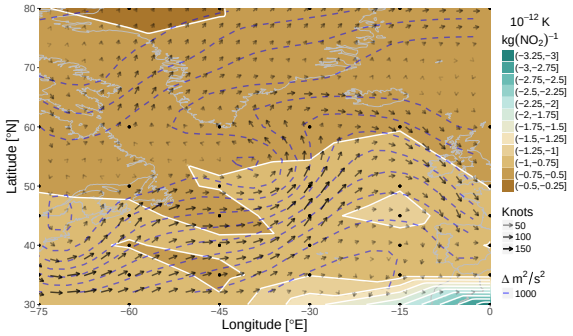


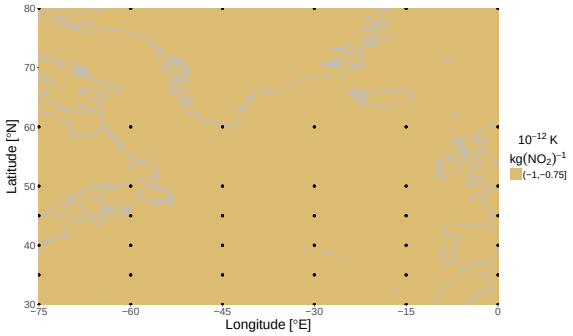


(a) REACT4C data

(b) Final $geopot \times zi0$ algorithm

Figure D.25: Methane CCF/aCCF contour plots, 12 UTC, 400 hPa, WP 1

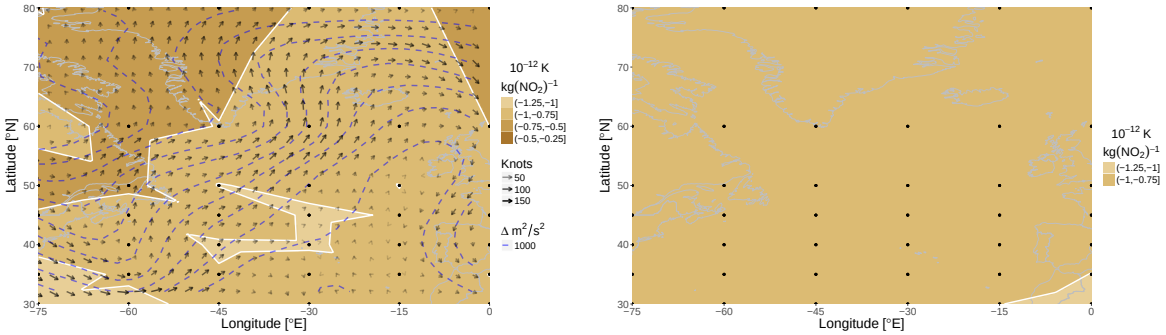




(a) REACT4C data

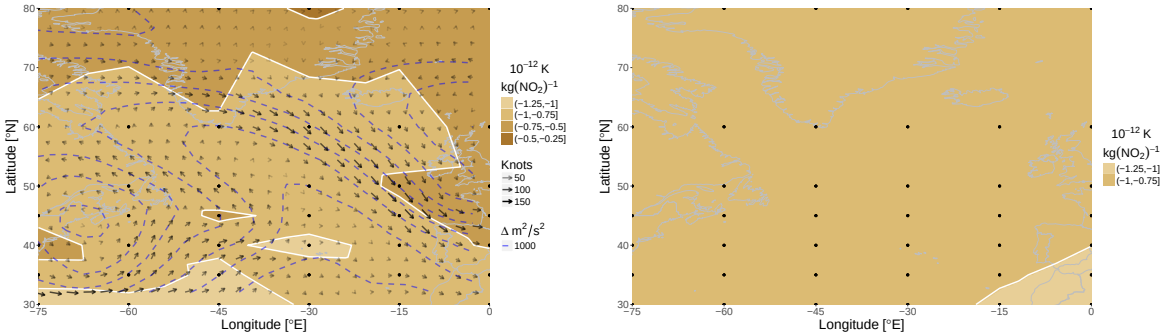
(b) Final $geopot \times zi0$ algorithm

Figure D.26: Methane CCF/aCCF contour plots, 12 UTC, 400 hPa, WP 2



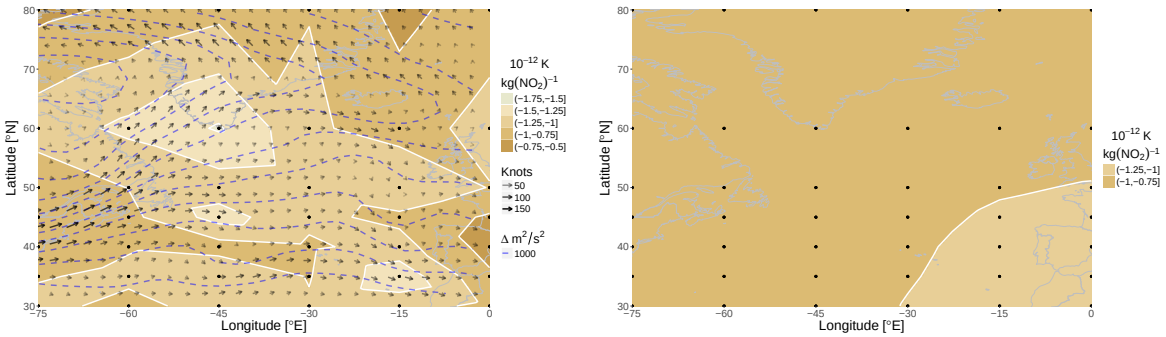
(a) REACT4C data (b) Final $geopot \times zi0$ algorithm

Figure D.27: Methane CCF/aCCF contour plots, 12 UTC, 400 hPa, WP 3



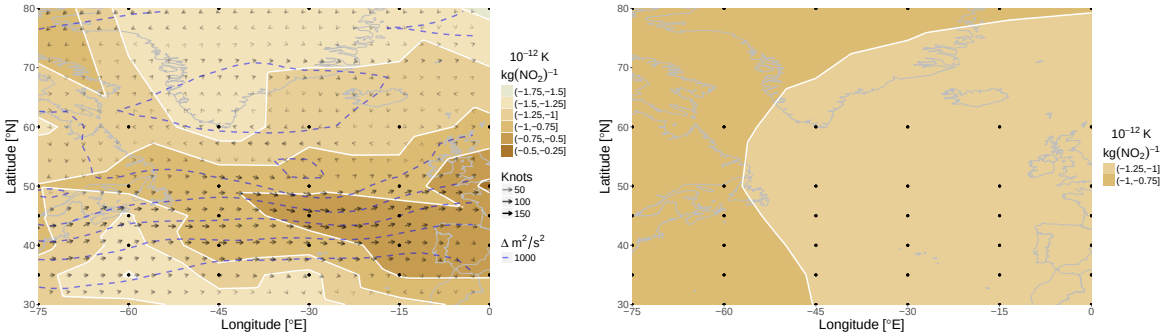
(a) REACT4C data (b) Final $geopot \times zi0$ algorithm

Figure D.28: Methane CCF/aCCF contour plots, 12 UTC, 400 hPa, WP 4



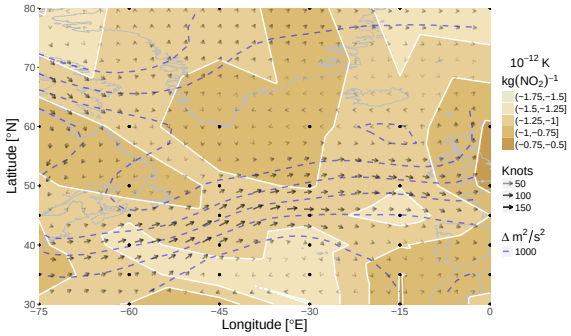
(a) REACT4C data (b) Final $geopot \times zi0$ algorithm

Figure D.29: Methane CCF/aCCF contour plots, 12 UTC, 400 hPa, WP 5

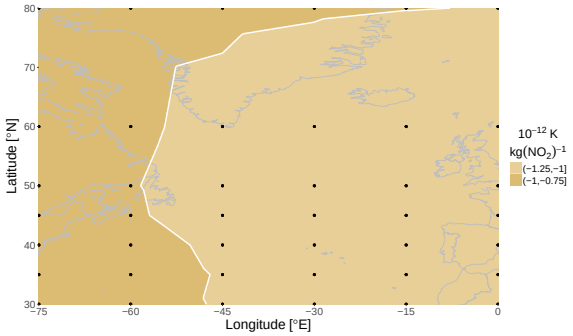


(a) REACT4C data (b) Final $geopot \times zi0$ algorithm

Figure D.30: Methane CCF/aCCF contour plots, 12 UTC, 400 hPa, WP 6

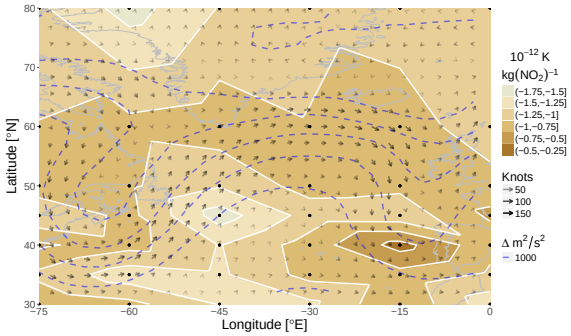


(a) REACT4C data

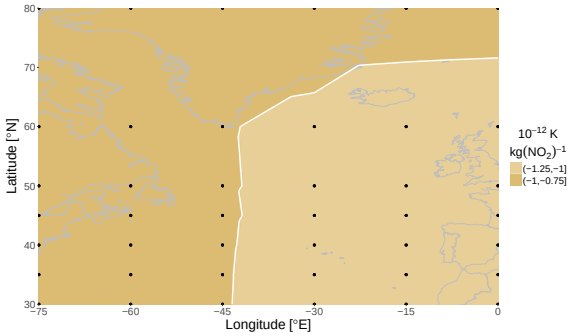


(b) Final $geopot \times zi0$ algorithm

Figure D.31: Methane CCF/aCCF contour plots, 12 UTC, 400 hPa, WP 7



(a) REACT4C data



(b) Final $geopot \times zi0$ algorithm

Figure D.32: Methane CCF/aCCF contour plots, 12 UTC, 400 hPa, WP 8