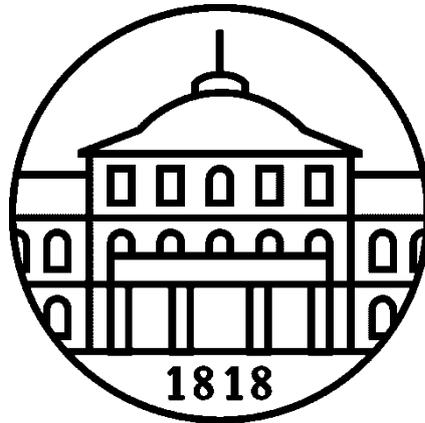


UNIVERSITY OF HOHENHEIM
Institute of Physics and Meteorology



**Meteorological conditions for strongly warming contrails and the statistics
of contrail's instantaneous radiative forcing**

Master's Thesis
M.Sc. Earth and Climate System Science

Handed in by
Lena Wilhelm



DLR

External institution:

**Deutsches Zentrum
für Luft- und Raumfahrt**
German Aerospace Center

Institute of Atmospheric Physics
Department Earth System Modelling

Head of Department: Prof. Dr. Volker Wulfmeyer (IPM)

Advisor: Dr. Klaus Gierens (DLR)

1. Examiner: Prof. Dr. Robert Sausen (DLR)

2. Examiner: Prof. Dr. Volker Wulfmeyer

Topic issued on: 12. July 2021

Date of submission: 07.01.2022

DECLARATION

Declaration*

I,

Lena Wilhelm

Matriculation number: 789141

declare that I have followed the Principles of Good Scientific Practice while writing the present

- | | | | |
|--------------------------|--------------------|-------------------------------------|-------------------|
| <input type="checkbox"/> | Bachelor's thesis. | <input checked="" type="checkbox"/> | Master's thesis, |
| <input type="checkbox"/> | seminar paper. | <input type="checkbox"/> | Diploma's thesis. |

I have written the paper/thesis independently and have used no other sources or aids than those given and have marked the passages taken from other works word-for-word or paraphrased.

Supervisor

Prof. Dr. Robert Sausen (DLR), Prof. Dr. Volker Wulfmeyer (Uni Hohenheim)

Topic of the paper/thesis

Meteorological conditions for strongly warming contrails and the statistics of contrails instantaneous radiative forcing

Semester : 5

I furthermore declare that the submitted unencrypted electronic document exactly and without exception corresponds to the contents and wording of the printed copy of the paper/thesis. I give my consent to this electronic version being checked for plagiarism with analytical software.

Stuttgart, 07.01.2022



Place, Date, Signature

* This declaration is to be included into the independently written paper/thesis as an annex. Papers/theses not including this declaration will not be accepted.

ABSTRACT

Persistent contrails and contrail cirrus are estimated to have a larger climate impact than all CO₂ emissions from global aviation since the introduction of jet engines. However, the measure for this impact, the radiative forcing (RF) or effective radiative forcing (ERF) comes with much larger uncertainties than those for CO₂. This study investigates one of the major causes for uncertainty, the natural variability. Specifically, the weather-induced variability is examined from a large dataset of instantaneous radiative forcing (iRF) values, produced from ten years of MOZAIC flights and ERA-5 reanalysis data.

Cdf's and pdf's of the iRF dataset show strong annual and interannual variations and a seasonal pattern. 80 % of the contrails have a small positive iRF of up to 20 Wm⁻², 10 % of all cases have an iRF \geq 19 Wm⁻², but these have a disproportionately large climate impact, and the remaining 10% have negative iRF. The distribution of iRF values declines exponentially at positive values and is heavily skewed. Monte Carlo experiments reveal the difficulty of determining a precise long-term mean from measurement campaign data. Depending on the chosen sample size, calculated means scatter considerably, which is caused exclusively by weather variability. This variability is the lower limit for uncertainty, which suggests, that there is a fundamental limit to the precision with which the RF and ERF of contrail cirrus can be determined.

The accurate local prediction of persistent contrails is still not possible because of errors in the humidity field in most weather prediction models. When the meteorological and dynamical conditions of persistent contrails and Big Hits (the strongest warming contrails) are known, they could be used as an addition to the SAC quantities to improve prediction possibilities. The data showed, that Big Hits favor small negative vertical velocities, small positive divergence, anticyclonic flow, low potential vorticities up to 4 PVU, large geopotential heights, and large lapse rates up to 10 K km⁻¹. The last four variables showed the strongest separation of the pdf's and are best suited for improving the prediction of persistent contrails. This was tested with a logistic regression model and with model output statistics for conditional probabilities. The results showed, that, 1) predicting Big Hits is quite reliable when it is already known that contrails will be persistent and 2) high probabilities for the persistence of contrails can be determined by introducing thresholds for dynamical variables and combining them with the SAC quantities. When such thresholds could be included in a weather prediction model like ECMWF's integrated forecast model (IFS), probabilities for contrail persistence could be produced before the flight planning period. Avoidance of persistent contrails or Big Hits would then become more reliable.

ACKNOWLEDGMENT

I want to express my deepest gratitude to my supervisor Dr. Klaus Gierens, who made this master thesis possible. He was the one who took me in as an intern in 2019, opened the world of research to me, and since then gave me so many opportunities, from writing scientific papers to being part of a working group and presenting in many seminars. With his (in my eyes) endless knowledge, he taught me to always question my results, have fun while redoing certain steps the fourth time, and not get intimidated by other experts in the field. Thanks, Klaus!

I also want to thank Prof. Wulfmeyer and Dr. Behrendt who made it possible to write my thesis externally and helped me with answering my (many) questions. I greatly appreciate Dr. Wulfmeyer's support with organizational issues.

Thank you to Prof. Dr. Robert Sausen who agreed to be my examiner albeit his full timetable and helped me with the preparation for the colloquium even on short notice. It was an honor to work alongside all the experts at DLR.

A big thanks to Flo, my special one, and to my friends and family who always believe in me.

Table of contents

Declaration	I
Abstract	II
Acknowledgment	III
List of figures	VI
List of tables	IX
Nomenclature	X
1 INTRODUCTION.....	1
1.1 Motivation and scientific background	1
1.1.1 Aviation in context.....	1
1.1.2 Contrail cirrus RF / ERF and their uncertainties.....	1
1.1.3 Weather variability induced uncertainty of contrail cirrus RF	3
1.1.4 Big Hits	4
1.2 Aims and objectives.....	5
1.3 Thesis outline	6
2 CONTRAILS AND AVIATION-INDUCED CIRRUS.....	7
2.1 Contrail formation and life cycle	7
2.1.1 Contrail Formation	7
2.1.2 Schmidt-Appleman criterion.....	9
2.1.3 Lifecycle and persistence	10
2.2 Contrail radiative forcing	12
2.2.1 Radiative forcing (RF) and instantaneous radiative forcing (iRF)	12
2.2.2 Energy forcing (EF)	13
2.2.3 Effective radiative forcing (ERF)	14
3 DATA	15
3.1 MOZAIC IAGOS	15
3.2 ERA-5 Reanalysis	17
4. METHODS.....	22
4.1 Statistics of iRF	22
4.2 Meteorological conditions of Big Hits.....	26
5 RESULTS.....	31

5.1 Statistics of iRF	31
5.1.1 Statistics	31
5.1.2 Monte-Carlo Experiments	37
5.1.3 Effect of Big Hits and maximum iRF	38
5.2 Meteorological conditions of Big Hits	41
5.2.1 Characterization – Seasonal and spatial patterns, length, iRF	41
5.2.2 Comparison of conditions for situations with and without strong contrails	44
5.2.2.1 Thermodynamic quantities	45
5.2.2.2 Dynamical characteristics	47
5.2.3 Prediction possibilities	51
 6 DISCUSSION	 58
6.1 Statistics of iRF	58
6.2 Meteorological conditions of Big Hits	61
 7 SUMMARY	 65
 8 REFERENCES	 69
 APPENDIX A	 80

- Figure 11:** Left: Seasonal variation of the mean iRF per month (black line) plus/minus one standard deviation (dotted lines) and the 5th – 95th inter-percentile range (dashed area) for the years 2000 to 2009. Right: Distribution functions (cdfs) for each season for the years 2000 to 2009 (colored) and the cdf for the whole dataset (black line).....**33**
- Figure 12:** Probability density functions for the iRF of years 2000 - 2009 (colored) and for the whole dataset (black). Left: Normal depiction. Right: Y-axis in logarithmic depiction...**35**
- Figure 13:** Black line: Pdf of the positive iRF values of years 2000 - 2009 plus its mathematical linear fit (red line). The straight-line fit indicates an exponential decline of the positive branch of the pdf.**36**
- Figure 14:** Pdf of the optical thickness of the contrails where iRF is larger than one (black) and its mathematical fit (der) in a half-logarithmic plot.**36**
- Figure 15:** Sample means (plus signs) of 1,000 samples of sample size 0.1 % (left) and 1 % (right) from the whole iRF dataset of persistent contrails. The blue line is the mean of all sample means and the red lines indicate the plus/minus one standard deviation range. Note the different scales on the x-axis.**37**
- Figure 16:** Pdf of the whole dataset (black line) with its first-order effect function (red) of the iRF (pdf x iRF). This shows the disproportionately large effect of high iRF values on the overall RF, despite their low probabilities.**39**
- Figure 17: Left:** Big Hits per flight, following a nearly exponential decline. **Right:** Big Hits per month.....**41**
- Figure 18:** All Big Hits coordinates drawn in red (13,995) on two different world maps. The red box in the left plot demarks the research area where MOZAIC flights were analyzed. The black arrow points to a gap in Big Hit cover in an area of**42**
- Figure 19:** Pdf of the length of Big Hits. The blue curve includes all Big Hits, the red one only Big Hits with a minimum length of 75 km and the blue for Big Hits with minimum 150 km length.**43**
- Figure 20:** Change of iRF along different flight paths. Flights of differing lengths were chosen randomly.....**44**
- Figure 21: Left:** Cdf of the iRF of all Big Hits (441,107 values). Most iRF have an iRF < 40 Wm^{-2} . **Right:** Pdf of all iRF of all Big Hits (black) with its mathematical fit (red).**44**
- Figure 22:** Temperature distribution in general situations (red), in situations that allow persistent contrails (black) and in situations with Big Hits (blue). The left panel shows the temperatures that ERA-5 reanalysis data predicts for the respective flight points from MOZAIC. The right panel shows MOZAIC’s measured temperatures. Note that the deciding factor for persistence and Big Hits (for both plots) are MOZAIC’s thermodynamic quantities.**45**

Figure 23: As figure 22 but for absolute humidity.....	46
Figure 24: IRF of all 441,107 Big Hit flight points plotted against the temperature. Highest iRF values occur at the highest temperatures.....	46
Figure 25: As figure 24, but for absolute humidity.....	46
Figure 26: As figure 22, but for RH_{max}	47
Figure 27: As figure 22, but for relative humidity with respect to ice.	47
Figure 28: Pdf's of dynamical characteristics for the three situations: Big Hit (red), persistent contrail (black), and not persistent (red). Upper left: Vertical velocity. Upper right: Divergence. Lower left: Vorticity. Lower right: Potential vorticity. The pdf's show large differences only for the vorticity and partly for potential vorticity.....	49
Figure 29: Left: Pdf for the normalized geopotential height for the three cases Big Hit (blue), persistent contrail (black) and not persistent (red). Big Hits occur predominantly at high geopotential heights. Right: Same as left plot, but for μ , the cosine of the solar zenith angle. Most Big Hits develop at night with a $\mu = 0$, or with very low or very high μ values.	50
Figure 30: Cdf and pdf of the lapse rate of Big Hits, persistent contrails, and for the general case. Persistent contrails and Big Hits develop particularly at high lapse rates.....	50
Figure 31: Correlation matrix of all variables in the iRF dataset (including 47,032 persistent or strongly warming contrails). Specifically, the correlation of variables with iRF is of interest (last row in each matrix). The correlations are illustrated as circles, with increasing size if a high positive (blue color) or negative correlation (red correlation) exists. The matrix includes all variables from ERA – 5.	52
Figure 32: Persistent contrails (black dots) and non-persistent contrails (red dots) in a Z – LR plot. The lines depict percentiles of each variable in 0.1 steps (10%). Those lines define 100 boxes of different sizes, where the probability for a persistent contrail can be counted for each box.....	54
Figure 33: Heatmap of conditional probabilities for contrail persistence in a two-dimensional space spanned from Z and LR percentiles (numbers on the axes). Highest probabilities occur with increasing geopotential height and increasing lapse rate.	56

LIST OF TABLES

Table 1: Description of used variables from the ERA-5 single level dataset (Hersbach et al., 2018b).	19
Table 2: Description of used variables from the ERA-5 pressure level dataset (Hersbach et al., 2018a).....	20
Table 3: Model parameters for various crystal habits. The parameter values assume that temperatures are given in K, fluxes are in Wm^{-2} , and radii in μm . For this study parameters for Myhre particles were used. Figure from (Schumann et al., 2012).	25
Table 4: Statistical key figures from the persistent contrails for the individual years and the whole dataset. iRFSW all years includes the instantaneous shortwave radiative forcing of all of the persistent contrails and the same for the instantaneous longwave radiative forcing at iRFLW. Note that the iRF referred to in the text is the net iRF, which is the sum of iRFSW and iRFLW.....	32
Table 5: Frequencies for persistent, cooling ($iRF < 0 Wm^{-2}$) and warming contrails ($iRF > 0 Wm^{-2}$) and Big Hits ($iRF > 19 Wm^{-2}$) per season and for the whole time series.....	34
Table 6: Number of Big Hits per season and in total with their percentages.	42
Table 7: Percentiles for different meteorological variables for the distribution of persistent contrails.	54
Table 8: As table 7, but for the distribution of Big Hits.	55
Table 9: For each Z percentile, the sum of the probability for persistent contrails of the LR percentiles and the sum of all of the probabilities of the two-dimensional probability matrix is given. Introducing additional conditions improves the total sum of probabilities. Note that the minimum count of persistent contrails per box reduces when adding further conditions.	55
Table 10: Maximum probabilities for the LR - Z matrix with the additional conditions from table 9. The highest probabilities can be seen at high lapse rates and high geopotential heights.	56

NOMENCLATURE

4D-Var	Variational Data Assimilation
A_{eff}	Effective albedo
AIC	Aviation Induced Cirrus
API	Application Programming Interface
BC	Black carbon
CARIBIC	Civil Aircraft for the Regular Investigation of the Atmosphere Based on an Instrument Container
cdf	Cumulative distribution function
CDS	Copernicus Data Service
CH ₄	Methane
CO ₂	Carbon dioxide
dam	Decameter
ECMWF	European Centre for Medium-Range Weather Forecast
EDA	Ensemble of Data Assimilations
EF	Energy Forcing
ERF	Effective Radiative Forcing
η	Propulsion efficiency
GCM	Global circulation model
HRES	High Resolution
IAGOS	In-service Aircraft for a Global Observing System
ICAO	International Civil Aviation Organization
IDL	Interactive Data Language
IFS	Integrated Forecast System
IPCC	Intergovernmental Panel on Climate Change
iRF	Instantaneous Radiative Forcing
ISSR	Ice supersaturated region
IWP	Ice water path
LES	Large-eddy simulations
LR	Lapse rate
LW	Longwave
MCH	MOZAIC Capacitive Hygrometer

MOZAIC	Measurement of Ozone and Water Vapour on Airbus In-service Aircraft
μ	Cosine of the solar zenith angle
N	Number of measurements
NO _x	Nitrogen oxide
O ₃	Ozone
OLR	Outgoing longwave radiation
pdf	Probability distribution function
PM	Particulate matter
PV	Potential vorticity
$r_{\text{eff cirrus}}$	Effective ice particle radius for contrail cirrus
r_{eff}	Effective ice particle radius for contrail
RF	Radiative Forcing
RF _{LW}	Longwave Radiative Forcing
RF _{SW}	Shortwave Radiative Forcing
RH	Relative humidity
RH _i	Relative humidity with respect to ice
RH _{max}	Relative humidity at maximum Schidt-Appleman temperature
RSR	Reflected solar radiation
SAC	Schmidt-Appleman criterion
SDR	Solar direct radiation
SW	Shortwave
T	Temperature
τ	Optical thickness
τ_c	Optical thickness of cirrus above contrail
TE	Total effect
T _{max}	Maximum Schmidt-Appleman temperature
TOA	Top of atmosphere
σ	Standard deviation
Z	Geopotential height

1 INTRODUCTION

1.1 Motivation and scientific background

1.1.1 Aviation in context

With the era of globalization and acceleration, aviation has become an important part of society over the past decades. The steady growth in flights, passenger numbers, and revenue bring immense economic and social benefits and are proof of the rising importance of this sector. According to the International Civil Aviation Organization (ICAO)(2019), the amount of carried passengers nearly doubled from 2.7 billion in 2010 to 4.5 billion in 2019 and carried freight tons increased from 188.5 billion in 2010 to 225.0 in 2019. This growth inevitably leads to an increase in negative externalities generated from global aviation. There is a large number of processes by which aviation contributes to climate change, overall combining to a net surface warming. This is caused mainly by aviation's large CO₂ emissions, which in 2018 accounted for 2.4 % of the overall anthropogenic CO₂ emissions, as well as by numerous non – CO₂ climate effects (Lee et al., 2021). Non – CO₂ effects include emissions of NO_x, water vapor, soot, and sulfur aerosols and the development of contrails and contrail cirrus, also called aviation induced cirrus (AIC). NO_x emissions lead to a short-term increase, but a long-term decrease of ozone (O₃), a decrease in methane (CH₄), and consequently a reduction of stratospheric water vapor (Lee et al., 2021). The impact of all of those effects on the climate is measured as effective radiative forcing (ERF) or radiative forcing (RF). Those metrics will be explained in chapter 2. In 2011 the net ERF for aviation contributed to 3.5 (5-95 % confidence interval: 4.0, 3.4) % of the overall anthropogenic ERF of 2290 (1130, 3330) mWm⁻². For 2018 non – CO₂ terms accounted for more than half of aviation's positive effective radiative forcing of 100.9 (55, 145) mWm⁻², namely 66.6 (21, 111) mWm⁻². The individual numbers for non – CO₂ effects can be found in figure 2. The largest effect, with 57.4 (17,98) mWm⁻², is ascribed to contrails cirrus in high humidity regions (Lee et al., 2021). With an ERF/RF ratio of 0.42 contrail cirrus are less effective in surface warming than CO₂ (ERF/RF = 1), but they still have a larger overall ERF.

1.1.2 Contrail cirrus RF / ERF and their uncertainties

The term contrail cirrus includes contrails plus cirrus cloudiness, that arises from the expansion of one or many contrails in ice supersaturated regions of the atmosphere. Contrails are line-shaped particle ice clouds (see figure 1), that are caused by the emission of water vapor,

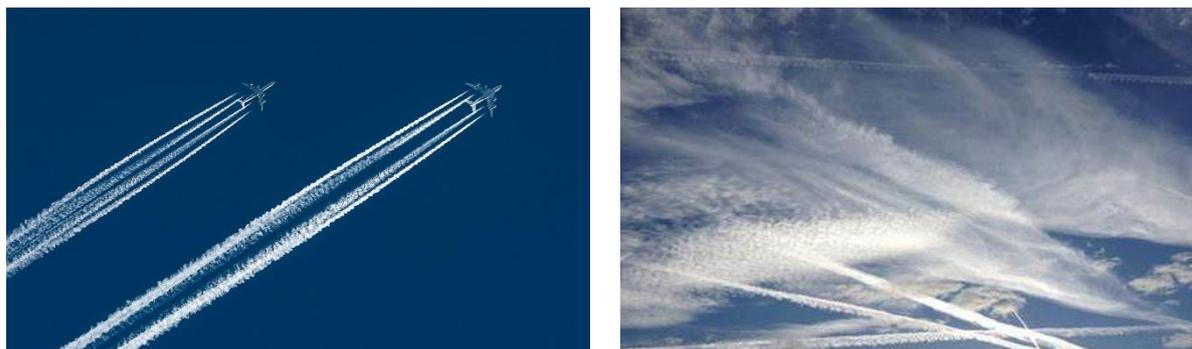


Figure 1: Left: Condensation trails forming behind the exhaust of an aircraft. **Right:** Aviation-induced cirrus. Source: <https://images6.alphacoders.com/983/983463.jpg>; <https://electricaircraft.mit.edu/wp-content/uploads/2>

aerosols, and particles from an aircraft exhaust, that mix with cold ambient air, which causes a local saturation, allowing condensation and freezing of water on aerosols (Appleman, 1953; Schmidt, 1941; Schumann, 1996).

Contrails impact the Earth's radiation-energy budget in two ways. They reflect solar radiation back to space and block thermal radiation from the Earth's surface and lower atmosphere. The shortwave radiative forcing (RF_{SW}) is absent during the night, while the longwave radiative forcing (RF_{LW}) occurs during day- and nighttime. The net RF, as their sum, adds up to the small warming effect listed in the IPCC charts.

The problem with the annual global RF or ERF for aviation is, that they come with very large uncertainties. Lee et al. (2021) give best estimates with 90 % uncertainty ranges (from the 5 % - 95 % confidence intervals), of which the non - CO_2 ERF terms contribute about 8 times more to the uncertainty of aviation net ERF than the CO_2 . Older estimates for contrail cirrus RF for 2011 (Boucher et al., 2013) and for aviation-induced cloudiness (which includes contrail cirrus) for 2005 (Lee et al., 2009) show smaller values than the new estimates from Lee et al. (2021), viz. 50 (10, 150) mWm^{-2} (old) vs. 86 (25, 146) mWm^{-2} (new) for 2011 and 33 (11, 78) mWm^{-2} (old) vs 82.6 (24, 140) mWm^{-2} (new) for 2005. The estimates from Lee et al. (2021) are based on results from Bickel et al. (2020), Bock and Burkhardt (2016), Burkhardt and Kärcher (2011), Chen and Gettelman (2013) and Schumann et al. (2015a).

What all estimates have in common is that they all have very large uncertainties and low confidence levels. Since the IPCC Special Report on Aviation and the Global Atmosphere in 1999 (Penner et al., 1999) the confidence level did not get any higher, despite of ongoing research. This is due to a variety of uncertainty sources. Lee et al. (2021, Appendix E) group them into two categories – A: Uncertainties in the radiative response to contrail cirrus, and B: Uncertainties from the upper-tropospheric water budget and the contrail cirrus scheme. Their calculated overall uncertainty for contrails cirrus RF is 70 % (Lee et al., 2021). So far, the determination of a precise global annual mean RF / ERF proves to be very difficult. This

however does not stem from a low level of scientific understanding (LOSU). For contrails, there has been considerable progress in scientific understanding during the last decades. Characteristics and the climate impact of contrails have been simulated with many different contrail models, namely Large-eddy simulations (LES) (eg. Lewellen and Lewellen, 2001; Unterstrasser, 2016), atmospheric global circulation models (GCM) (eg. Burkhardt and Kärcher, 2009; Ponater, 2002), and Lagrangian models (eg. Schumann, 2012; Unterstrasser and Sölch, 2010). New theories, eg. about the contrail formation for fuel cells (Gierens, 2021) and from hybrid-electric aircraft (Yin et al., 2020), have been introduced. The detection of contrails with satellite imagery (Mannstein et al., 1999) and the development of a contrail tracking algorithm for the development of contrails (Vazquez-Navarro et al., 2010) have been presented. These are only a few examples of the numerous studies about contrails and contrail cirrus with all of their connected complex processes. Unfortunately, all of these results did not contribute to a reduction of the uncertainties associated with contrail cirrus RF. Therefore, the many uncertainties do not limit the understanding of contrail cirrus forcing, but they limit the ability to determine a precise global mean.

1.1.3 Weather variability induced uncertainty of contrail cirrus RF

One major source for the large uncertainty is the natural variability, and specifically the weather-induced variability. Unfortunately, this variability only becomes visible in ensemble simulations or when simulations are repeated for the same year. Gettelman et al. (2021) presented such an ensemble study. They used air traffic data from 2006, scaled them to 2019 traffic amounts, and then calculated the ERF for the years 2019 and 2020, ignoring Covid-19 in 2020. The same amount of emissions is used for both years, only the meteorological conditions differed. The mean ERFs were $90 \pm 50 \text{ mWm}^{-2}$ for 2019 and $33 \pm 35 \text{ mWm}^{-2}$ for 2020, where the two standard deviations have been determined using the different results of the ten ensemble members. The mean of the two years is $62 \pm 59 \text{ mWm}^{-2}$ which is very similar to the ERF given by Lee et al. (2021) for 2018. What is remarkable is the huge difference between the two years. They are exclusively caused by changes in weather. The weather variability induces large uncertainties in the radiative forcing of contrail cirrus. This thesis examines exactly that: the weather variability induced uncertainty on contrail radiative forcing.

For this, the most basic measure for the radiative effect of a contrail, the instantaneous radiative forcing (iRF), and not the mean RF or ERF were used. The instantaneous RF of a contrail (iRF), also called cloud radiative effect, is the difference of the net radiation flux at top of the atmosphere in a situation with vs. without a contrail. It only considers the immediate radiative

effect at one single point in space and time. It can be three orders of magnitude larger than the RF metric (Gierens et al., 2020). IRF can represent strongly warming (endothermic) and strongly cooling (exothermic) cases. Forcing measures are only relevant for climate if contrails are persistent. That is, if they fulfill the thermodynamic Schmidt-Appleman criterion (Schumann, 1996) and are located in an ice-supersaturated region (ISSR) where the relative humidity with respect to ice RH_i exceeds 100 %. With those favorable conditions, ice crystals do not immediately sublime, instead, the total ice water content can grow by the uptake of ambient water, which leads to the formation of a persistent contrail and contrail cirrus. Persistent contrails can exist for several hours, with their appearance gradually approaching that of natural cirrus (see figure 1, right), when the contrail gets spread out by the local wind field and wind shear (Marquart, 2003).

1.1.4 Big Hits

On average, an aircraft spends about 15 % of its flight path in ISSR's (Gierens et al., 1999, p. 199). In theory, the climate impact of contrails can be eliminated by preventing the formation of persistent contrails (Teoh et al., 2020a). Since it is not efficient to try to prevent all persistent contrails in those 15 % of the flight distances, one should focus on contrails with the highest positive individual radiative forcings, so-called Big Hits (Royal Aeronautical Society, 2019). Big Hits are produced only on a small fraction of flight trajectories, meaning that a change in flight path could considerably prevent contrail formation. Using six weeks of flight data for the Japanese airspace, (Teoh et al., 2020a) demonstrated that only 2.2 % of flights contribute to 80% of contrail energy forcing (EF). The term EF will also be explained in chapter 2. Selectively diverting 1.7 % of those flights using a small-scale diversion strategy could reduce the EF by up to 59.3 % with an increase of fuel consumption and CO₂ emission of only 0.014 %. Even without additional fuel, a low-risk flight strategy could decrease the EF by 20 %. With such a minimally invasive optimized routing strategy, the negative warming effect of contrails could be reduced while keeping the additional aircraft operating costs and greenhouse gas emissions at a minimum. Such an avoidance strategy needs of course a reliable prediction of the conditions under which contrails can become such strong climate warmers (Big Hits). This subject is addressed in the second part of this study.

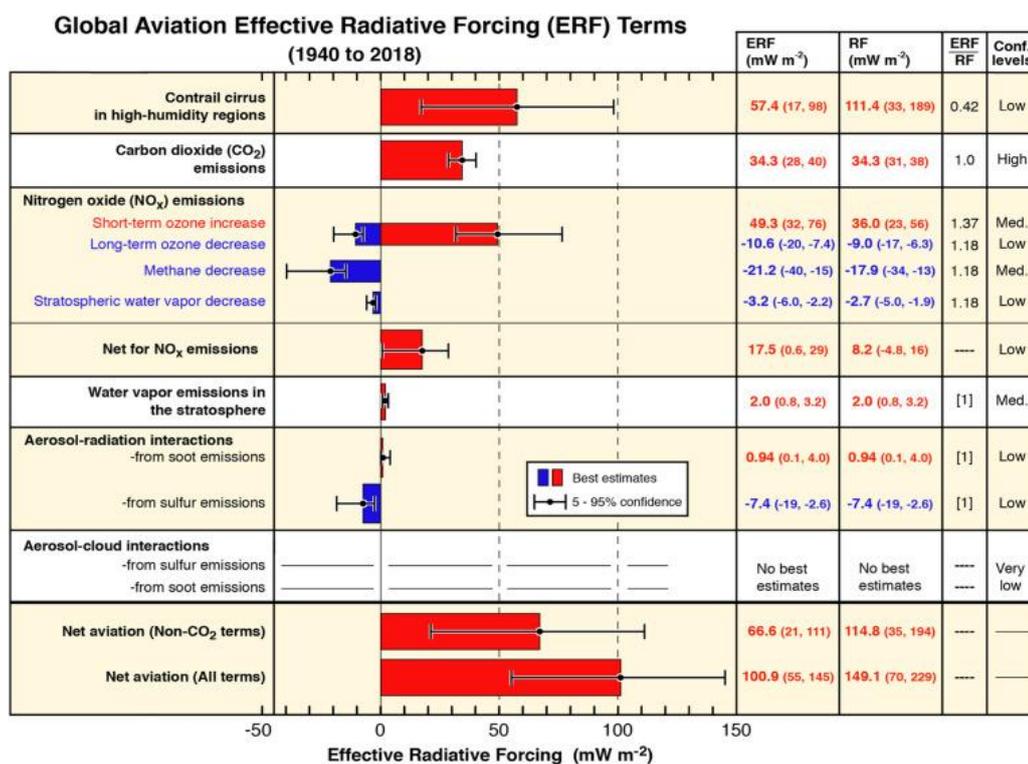


Figure 2: Best estimates for ERF, RF, and ERF/RF ratio including the confidence level from global aviation from 1940 to 2018. The bars and whiskers show best estimates of ERF with their 5–95 % confidence intervals. Red bars indicate a warming effect and blue bars indicate a cooling effect.

1.2 Aims and objectives

This thesis has two main goals. Firstly, this study aims to give a reason as to why we have not yet been able to significantly reduce the high uncertainty associated with contrail cirrus RF. For this, statistics of the instantaneous radiative forcing (iRF) of individual persistent contrails are produced from a combination of MOZAIC in-situ measurements from passenger aircraft and ERA – 5 reanalysis data. Ten years of data were used to achieve good statistics. The goal was to determine the variability in the data that is solely ascribable to weather variability. Therefore, all contrails in the data set are treated with the same methods, they are assumed to have fixed crystal size and depth, and thus only the weather varies, i.e. the amount of ice that can be created as well as the ambient radiation quantities. As the least number of degrees of freedom was used, this variability, which is a major source for uncertainty, is the lower limit for uncertainty.

Furthermore, this study intends to display the disproportionately large effect of strongly warming contrails. By avoiding them, a substantial part of aviation’s climate impact could be eliminated. Hence, the second objective of this thesis is to characterize Big Hits and analyze their meteorological conditions, to achieve better prediction possibilities. Different methods are used to try to produce probabilities for persistent contrails and Big Hits, based on their meteorological conditions. If regions where Big Hits would develop could be reliably predicted before the start of a flight, eg. by implementing those probabilities into global forecasts, the

flights could be rerouted to avoid areas of high probability. Such a strategy would require minimal additional fuel and emission of extra CO₂ since Big Hits occur only rarely. To see whether contrail prevention through flight route diversion is possible at all, the meteorological conditions and the iRF are also checked for their coherence.

1.3 Thesis outline

In chapter one, the scientific importance and motivation for the topic were presented. The next chapter gives a short introduction to contrail formation and radiative forcing metrics. In section 3 data sources are introduced, followed by section 4 that describes the methods used for the study and its objectives. Results can be found in chapter 5 and are discussed in chapter 6. The most important points are summarized in chapter 7.

2 CONTRAILS AND AVIATION-INDUCED CIRRUS

Condensation trails (hereon called contrails) are line-shaped, aviation-induced cirrus clouds that have a rather different formation mechanism than natural cirrus. Under favorable atmospheric conditions, they can become persistent and grow to a large cirrus cover, which can induce a considerable climate effect. This climate effect is usually measured in radiative forcing metrics. In the next subsection the formation mechanism and relevant thermodynamics, as well as the life cycle of contrails and contrail cirrus will be presented. To explain and differentiate the in the study used instantaneous radiative forcing (iRF), chapter 2.2 then lists the mainly used climate forcing metrics for contrail cirrus.

2.1 Contrail formation and life cycle

2.1.1 Contrail Formation

As mentioned before, contrails and contrail cirrus are ice clouds. There are two basic ice formation mechanisms, homogeneous and heterogeneous ice nucleation. Both mechanisms are schematically depicted in figure 3. Homogeneous ice nucleation is dominant for the formation of pure ice clouds, meaning natural cirrus, while heterogeneous ice nucleation is the decisive process for contrail cirrus. Homogeneous ice nucleation describes the process when supercooled liquid drops of water, sulfuric acid, or other substances freeze spontaneously without the presence of a particle nucleus (Bier, 2018; Koop, 2004). Depending on the drop radius of the liquid, it can be cooled way below its melting point T_m , down to the homogeneous nucleation temperature T_f . For pure water T_f is ~ 235 K (Koop, 2004), but the addition of solutes decreases the temperatures even further (Koop et al., 2000). Supercooled liquids can coexist with their crystalline phase, but, while the crystalline phase is stable, the supercooled liquid is in a metastable state (Koop, 2004). This means that there is a limit of supercooling, and when a critical vapor pressure or temperatures below T_f is reached, rapid nucleation happens in a matter of seconds (Pruppacher and Klett, 2010; Russo et al., 2014). Then, because of internal fluctuations, the small molecular clusters inside the bulk of the liquid drop will grow to a stable germ and once the critical size is reached, a macroscopic phase change occurs in the whole drop and it freezes to become an ice crystal (Pruppacher and Klett, 2010). In the upper cold troposphere ($T \leq -40$ °C) natural cirrus clouds form mainly by homogeneous freezing from liquid aerosol particles that consist of aqueous solutions of sulfuric acids and other substances. Freezing is only possible when the solution is highly diluted, which happens when the solution droplets gain water molecules from the ambient air. Therefore, the ambient relative humidity with respect to ice needs to exceed ~ 145 % (Gierens et al., 2012). When it comes to contrail

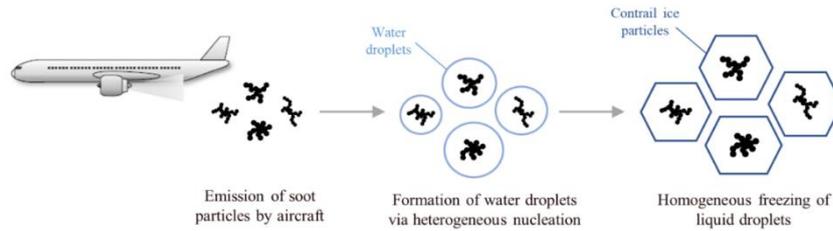


Figure 3: Formation of contrail ice particles by heterogeneous nucleation and subsequent freezing of the liquid droplet. Figure from (Teoh, 2020, p. 44).

formation, homogeneous nucleation is thought to be negligible (Kärcher and Yu, 2009). However, especially in the northern hemisphere, where the air is more polluted due to air traffic and more nuclei are available in the atmosphere, natural cirrus can also be formed by heterogeneous nucleation (Gierens, 2003). When the crystallization of supercooled liquids is initiated by surface contact with a solid particle nucleus, this is called heterogeneous nucleation. The addition of a foreign nucleus to a liquid increases the overall nucleation probability. Ice nuclei can be the many tiny aerosol particles in the atmosphere which have a concentration of $10^2 - 10^4 \text{ cm}^{-3}$ in the upper troposphere (Minikin et al., 2003; Schröder et al., 2002; Schröder and Ström, 1997), or when it comes to contrails, the particles from the exhaust plume of the aircraft that have a concentration of 10^9 cm^{-3} at the engine exit (Schröder et al., 1998). Aerosol particles in the upper troposphere are mainly supercooled solutions with sulfate-, nitrate- and ammonium-ions or organic materials (Bier, 2018). The exhaust gas volume of a conventional aircraft engine consists of 91.5 % natural components of the air, namely 75.2 % nitrogen and 16.2 % oxygen. The remaining volume is made up of 72 % carbon dioxide, 27.6 % water vapor, and 0.4 % pollutants, which consist of 84 % nitrogen oxides, 12 % carbon monoxide, and a small percentage of sulfur dioxide and particle emissions (Bier, 2018). Those particle emissions include soot particles and particulate matter (PM) like black carbon (BC), metallic compounds, and ultrafine organic volatile particles. Of all of those aircraft emissions, water vapor and particle emissions are most important for contrail formation (Bier, 2018; Teoh, 2020). However, particle emissions play only a secondary role, as contrails could even form when the aircraft would not emit any particles, eg. like in the exhaust of liquid hydrogen-fueled aircraft (Ström and Gierens, 2002). Therefore, contrail formation is mainly thermodynamically driven. Contrails form because of a local water saturation that develops during the isobaric mixing between the warm, moist exhaust gases from the aircraft and the cold, drier ambient air. Water molecules then condense at the surface of the aerosol particles from the exhaust or the surrounding air (mixed into the plume) and by the uptake of ambient water vapor, the liquid

droplets grow in size and then freeze homogeneously to become ice crystals (see figure 4). This all happens in less than a second.

Contrails form only in very humid and sufficiently cold regions, with temperatures below $-40\text{ }^{\circ}\text{C}$, which is usually in cruise levels of the upper troposphere and lower stratosphere. In such cold conditions, a direct condensation of ice on the particles would in principle be possible, but the particles would need to have special surface properties, like a crystalline lattice structure similar to that of ice, which is quite rare (Gierens, 2010; Pruppacher and Klett, 2010). Most of the ice crystals develop on coated and activated soot particles (Kärcher and Yu, 2009), which are produced by the incomplete combustion of hydrocarbons in the kerosene (Bockhorn, 1994).

2.1.2 Schmidt-Appleman criterion

As previously mentioned, the formation of contrails depends mainly on thermodynamic conditions during the mixing of the warm exhaust gases with the cold ambient air. In 1996 Schumann presented a criterion based on the findings from Schmidt and Appleman (Appleman, 1953; Schmidt, 1941), the so-called Schmidt-Appleman criterion (SAC) (Schumann, 1996), that can exactly decide whether contrail formation is possible.

Contrails develop only when, during the plume expansion process, local water saturation is achieved. This means, that the ambient atmospheric temperature has to be below a certain threshold temperature T_{max} and RH_{max} , which is the relative humidity that the mixture of exhaust and ambient air attains in the moment when its temperature is T_{max} , has to be above 100 % (Gierens et al., 2020). Equations for both T_{max} and RH_{max} are given in chapter 4.1.

The mixing process of the exhaust gas and the ambient air takes place isobarically, such that it follows a straight line on a $T - e$ phase diagram. Contrail formation is only possible when the mixing line intersects with the water saturation curve. In figure 5 two isobaric mixing lines for two possible phase trajectories are drawn, both starting at higher exhaust temperatures and cooling down to the ambient temperature during the mixing process. The two curves show the water vapor saturation partial pressure with respect to liquid water (upper one) and ice (lower one). The right line, which is the tangent to the water saturation curve, demarks the highest ambient temperatures, where contrail formation is still possible, namely the temperatures at the tangential point. For this trajectory T_{max} is at $\sim -40\text{ }^{\circ}\text{C}$ with a water vapor partial pressure $e \approx 18\text{ Pa}$. The left straight line shows a trajectory for contrail formation below the threshold T_{max} , where a large supersaturation is reached during the mixing process. The light grey area depicts conditions where short-lived contrails form, while the dark-grey area stands for the formation of persistent contrails, which happens only in ice supersaturated regions (ISSR`s).

The slope of the phase trajectory can be calculated with the equation

$$G = \frac{de}{dT} = \left(p * \frac{c_p}{\epsilon} \right) * \left[\frac{EI_{H_2O}}{Q*(1-\eta)} \right]. \quad (2.1)$$

It depends on the air pressure of the flight level p , the isobaric heat capacity of air $c_p = 1005 \text{ J kg}^{-1}\text{K}^{-1}$, the molar mass ratio of water and dry air $\epsilon = 0.622$, the emission index of water vapor $EI_{H_2O} = 1.25 \text{ kg per kg kerosene burnt}$, the combustion heat $Q = 43 \text{ MJ per kg of kerosene}$, and the overall propulsion efficiency of the aircraft $\eta = 0.35$ (for modern airlines) (Gierens, 2010). The overall propulsion efficiency is a function of the thrust of the engine F , the speed of the aircraft V and the amount of chemical energy provided by the combustion heat Q at flow rate m_f :

$$\eta = FV/(m_f Q). \quad (2.2)$$

Hence G depends on the atmospheric situation, as well as on the aircraft-, engine-, and fuel-combination (Schumann, 2005).

2.1.3 Lifecycle and persistence

The lifecycle of contrails can be explained with the three main evolution phases of contrails, the jet phase (ca. 20s), followed by the vortex phase (ca. 2 min), and the dispersion phase (minutes to hours) (Gierens, 2010). The jet phase includes the mixing of the exhaust gases with the ambient air, which was already explained. The mixing of the exhaust gases starts earlier at the edges of the plume than in the center. This is due to the structure of the jet engines, which emit colder air around a hot middle stream. At the end of the jet phase, the temperature in the plume has decreased down to that of the ambient air (Bier, 2018).

In the next phase, the vortex phase, the ice crystals become trapped in a counter-rotating vortex pair. The vortex system then moves downwards, which leads to a partly sublimation of ice

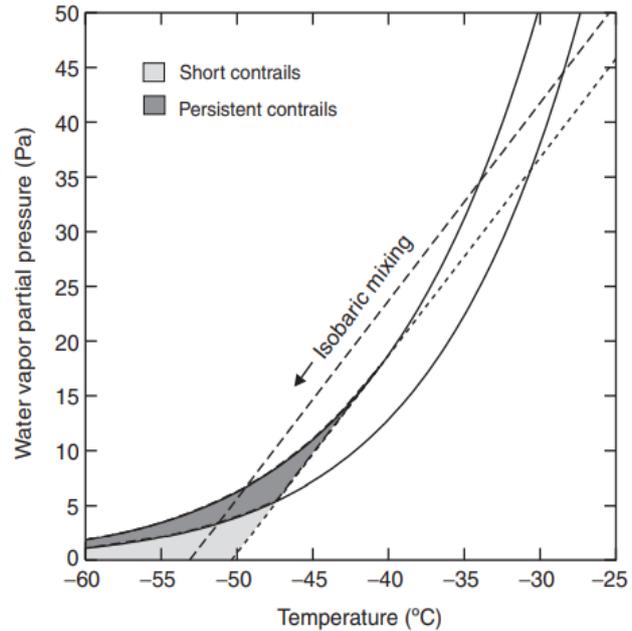


Figure 4: T-e phase diagram for contrail formation. The two straight lines are mixing trajectories for different mixing processes. The right line is the tangent to the water vapor saturation partial pressure curve (upper), which is where T_{max} lays. The left line shows a phase trajectory with atmospheric conditions a few degrees higher than the threshold temperature T_{max} . The lower curve is the saturation vapor pressure curve for ice. Phase points in the light grey area indicate that contrail formation is possible, and in the dark grey area that persistent contrails are possible. Figure from (Gierens, 2010).

particles due to adiabatic heating. A secondary vortex is formed from a small fraction of ice crystals above the upper edge of the primary vortex system (Unterstrasser, 2008). The loss of ice crystals during the vortex phase depends mainly on the relative humidity and the initial number of ice crystals and is up to a certain degree influenced by the aircraft type (wingspan), ambient temperature, and atmospheric stability (Unterstrasser, 2016). After about two minutes the vortex systems become unstable and decay (Gierens, 2010).

In the dispersion phase, contrail formation is now only influenced by atmospheric exchange. Turbulent diffusion and vertical wind shear can spread a line-shaped contrail up to several kilometers horizontally (Schumann, 2005). At some point, they lose their line shape and become large cirrus covers that are indistinguishable from natural cirrus clouds. This only happens, in the case of a persistent contrail. Such contrails can have lifetimes of up to 19 h (Vázquez-Navarro et al., 2015). They disappear when they are no longer in ice supersaturated regions, where the ice particles then sublimate. This can happen by sinking and adiabatic heating of the contrail, or when the air mass is horizontally transported to drier regions by advection (Bier, 2018). This means, that the ISSR cover defines the potential contrail cirrus cover. Areas like Europe and North America, where air traffic density is very large, can show a contrail cirrus coverage of up to 10 % (Burkhardt and Kärcher, 2011). The contrail cover by contrail cirrus is assumed to be ten times larger than that of line-shaped contrails (Mannstein and Schumann, 2005).

When contrails do persist for a long time, the ice particles can grow by accumulation of water molecules, which is drawn from the ambient saturation excess and therefore reduce the level of ambient ice supersaturation (Teoh, 2020). They can grow until they become too large and sedimentation into a subsaturated layer happens, where they evaporate (Schumann, 2005; Schumann et al., 2015b).

Persistent contrails favor regions that are also suitable for the formation of natural cirrus. When a natural ice cloud lays in the vicinity of contrails, they both compete for the water available for deposition (Kärcher, 2018). However, it is also possible, to find contrails in an otherwise completely cloud-free sky. This is due to the high RH_i values necessary for homogenous nucleation, which needs to exceed 145 % (Koop et al., 2000). For contrail persistence $RH_i \geq 100$ % is enough. Ice-supersaturated regions are usually formed by an upward motion of air masses during which the dry adiabatic cooling leads to an increase of the relative humidity. Divergent flow also favors ISSR development (Gierens and Brinkop, 2012). ISSR's are mainly found in the upper troposphere (at 8 – 13 km altitude), have an average horizontal extension of $150 \text{ km} \pm 250 \text{ km}$ and an average vertical extension of $700 \pm 100 \text{ m}$ (Gierens and

Spichtinger, 2000; Kärcher, 2018; Spichtinger et al., 2003a). On average, passenger aircraft spend about 10 – 15 % of their flight path in an ISSR (Gierens et al., 1999).

2.2 Contrail radiative forcing

Now that the basic formation mechanisms for contrails and contrail cirrus are explained, the question arises, how those anthropogenically caused phenomena affect the Earth and its climate. A short overview was already given in the introduction. This chapter will continue with an explanation of the radiative effect of contrail cirrus, as well as the mainly used climate forcing metrics.

2.2.1 Radiative forcing (RF) and instantaneous radiative forcing (iRF)

Radiative forcing (RF) is defined as the net change in the Earth's energy balance due to some imposed perturbation (Myhre et al., 2013). It is usually measured in Wm^{-2} as the change in net energy flux at the top of the atmosphere (TOA). For contrails, it is the sum of all individual radiative forcings induced from all contrails in a specific spatiotemporal domain (Teoh et al., 2020b), typically given globally and over a time period of one year. Every individual contrail causes a change of the radiation flow through the atmosphere and may cause a net energy loss (cooling or exothermic contrail) or an energy gain (warming or endothermic contrail). In the IPCC charts RF is on average slightly positive (Kärcher, 2018; Lee et al., 2021), but it is always the product of model calculations. In general, RF is not optimal for quantifying local effects of contrails and contrail cirrus, as it is a global metric. This is why in this study the instantaneous radiative forcing (iRF) is used. iRF is the radiation change due to a contrail on a single location and time point, not averaged over a certain area or time (Schumann et al., 2012). It is therefore the most basic measure for the radiative effect of a contrail and can be expressed as “the difference of the net radiation flux at top of the atmosphere in two situations, one with contrail, one without, but otherwise equal” (Wilhelm et al., 2021). Hence one arrives at RF when all iRF's from all global contrails (or from a specific area) are integrated over a specific time. Contrails induce changes in the energy budget of the Earth mainly by shortwave radiative forcing RF_{SW} and longwave radiative forcing RF_{LW} . When talking of RF or iRF of contrails, it is always the sum of RF_{SW} and RF_{LW} , so the net effect. During the day, contrails reflect the incoming solar shortwave (SW) radiation back to space leading to a cooling effect. This shortwave radiative forcing (RF_{SW}) results from a visibility contrast: It increases from bright to dark background, from snow cover to open ocean water, from desert to forest areas, and from cloudy to cloud-free skies. RF_{SW} depends on the solar irradiance at the contrail, which depends on the solar zenith angle, as well as on the albedo of the Earth's surface or underlying clouds

and the reflectance of the contrail (Fuglestedt et al., 2010; Lynch et al., 2002; Meerkötter et al., 1999). During day- and nighttime contrails also block and reemit outgoing thermal longwave (LW) radiation from the Earth's surface and lower atmosphere, which in contrast leads to a warming by a greenhouse effect. This longwave effect (RF_{LW}) results from a temperature difference between the cold embedded contrail in the warm background. It depends on the background temperature and humidity and the contrail temperature and emissivity. Emissivity is the ratio of thermal radiation emitted by the surface of a grey body compared with that of a black body under identical thermal conditions. Emitted thermal radiation depends strongly on the temperature of the emitter (Stefan-Boltzmann law). RF_{LW} increases from pole to equator, from winter to summer, and from cool night to warm day (Lynch et al., 2002; Teoh, 2020). Both, contrail reflectivity and emissivity, are determined by the contrails ice content and thickness, as well as the ice crystal size and habit distribution (Schumann, 2005).

All of these mentioned factors have a large influence on RF. Maybe even more important is, that they introduce tremendous natural variability. The error bars for RF in the IPCC charts do not reflect this natural variability. They stem from differences between model results, caused by different parameterizations.

RF has another major disadvantage: not only is it not able to quantify the effect of a single contrail like iRF can, but is also not able to determine the cumulative climate forcing over the whole lifecycle of a contrail. The energy forcing (EF), another metric for the radiative effect of contrails and contrail cirrus, does that.

2.2.2 Energy forcing (EF)

The EF (in units of J) is the integral of the iRF over the area (width (w) and length (L)) and lifetime (t) of a contrail (Schumann et al., 2011; Schumann and Heymsfield, 2017; Teoh et al., 2020b):

$$EF = \int_0^t iRF(t) \times L(t) \times w(t) dt. \quad (2.3)$$

It can also be given per flight distance or length of a contrail in $J m^{-1}$.

For iRF the variability should be smaller compared to that of EF for individual contrails, as EF includes the whole development of the contrail, including all uncertainties introduced from that. Note, however, that RF and EF cannot be directly compared, as they are two different physical quantities. A large RF does not imply a large EF and vice versa.

The largest EF values can be found in regions with long-lived, but weak ice supersaturated conditions, where the contrail can develop into a huge contrail cirrus cover (Schumann and Heymsfield, 2017; Teoh, 2020). This is because, the larger the relative humidity with respect

to ice, the larger the ice crystals can grow, which leads to a faster sedimentation rate and consequently a shorter lifetime of the contrail.

2.2.3 Effective radiative forcing (ERF)

Another often used climate effect metric for contrail cirrus is the effective radiative forcing (ERF). ERF is defined as the “change in net TOA downward radiative flux after allowing for atmospheric temperatures, water vapor and clouds to adjust, but with global mean surface temperature or a portion of surface conditions unchanged“ (Myhre et al., 2013). This means, that ERF allows feedback-like radiative adjustments on short time scales. Compared to RF, ERF of contrail cirrus is substantially reduced because of the introduced negative feedback by the compensating effect of natural clouds (Bickel et al., 2020).

But how are ERF and EF related? One could use EF and compute a specific radiative forcing (see Fuglestvedt et al., 2010) from that, by dividing EF by the Earth’s surface area, the contrail’s lifetime, and the unit mass one kg (of kerosene). The specific RF can be understood as the time-, location-, and situation-dependent radiative forcing of a contrail of a length that could be flown with one kg of fuel. If the total flown distances of a year and the specific forcings, lifetimes, and widths of each contrail are known, one can sum everything up to get the annual global mean RF value. By multiplying that with the ERF/RF ratio from (Lee et al., 2021) (0.42^{-1}), one finally arrives at the ERF.

All the mentioned climate impact metrics for contrails and contrail cirrus (RF, EF, ERF) are associated with large uncertainties. This is, unfortunately, inevitable, unless we can measure every single contrail in the atmosphere and calculate eg. RF from measurements. This thesis will show, that the uncertainty induced from weather variability is very large, even when many quantities that in reality are additional degrees of freedom are fixed: crystal shapes and sizes, contrail depths, etc. The variability that will be seen is the lower bound of variability. It sets a fundamental limit to the precision with which the RF and ERF of contrail cirrus can be determined.

3 DATA

3.1 MOZAIC IAGOS

For this study, a combination of MOZAIC IAGOS and ERA-5 Reanalysis data was used. The Measurement of Ozone and Water Vapor on Airbus In-service Aircraft (MOZAIC) campaign (Marengo et al., 1998), that was in 2011 transferred into the In-service Aircraft for a Global Observing System (IAGOS) (Petzold et al., 2015), is a European Research Infrastructure that offers global in-situ data for meteorological quantities and atmospheric chemical composition, produced from fully-automated equipment installed on commercial passenger aircraft. Starting in 1994 with five long-range Airbus A340 aircraft, the program has now several Airbus A340 and A330 from several airlines in operation, that are organized in MOZAIC / IAGOS-CORE and IAGOS-CARIBIC.

In the frame of IAGOS-CORE (precursor: MOZAIC) 500 flights per year are recorded from seven aircraft which provide quasi-continuous measurements for trace gases, aerosol, and cloud particles from two instrumentation packages. Package one measures humidity, cloud particles, ozone, and carbon monoxide and for package two, depending on the type (a,b,c,d), other trace gas concentrations are determined. In the future, IAGOS aims to increase its fleet to 20 active aircraft to increase coverage and develop new instruments. (“IAGOS-CORE. In-service Aircraft for a Global Observing System.,” n.d.).

Over 62,000 flights have already been collected and can be accessed at the IAGOS portal (“IAGOS Data Portal,” 2019) after registration and acceptance of the data protocol. The MOZAIC-IAGOS data have so far been used in 429 scientific publications (“IAGOS. Scientific publications.,” n.d.), many of them studying ice-supersaturated regions, natural cirrus clouds, and contrails (Beswick et al., 2015; Burkhardt et al., 2008; Gierens et al., 2000; Gierens and Spichtinger, 2000; Gierens et al., 1997; Lloyd et al., 2020; Petzold et al., 2017, 2020; Reutter et al., 2020; Spichtinger et al., 2002, 2003a, 2004).

For this study only flights from MOZAIC were selected, specifically those of years 2000 – 2009, amounting to 16,588 flights. From those flights, the flight position and the variables air temperature T and relative humidity RH were used for the analysis. Datapoints were restricted to pressure altitudes at cruise level between 310 hPa and 190 hPa (9 – 12 km) and to the research area seen in figure 6, extending from 30° to 70° latitude (zone with most air traffic) and 125° to 145° longitude. This made sure that most of MOZAIC’s flights were covered, while not including too many different climate zones. Measurements are available every four

seconds. With a maximum airspeed of 890 km/h (Lufthansa, n.d.; Nedelec et al., 2003), this translates to a flight distance of one kilometer between two measurements.

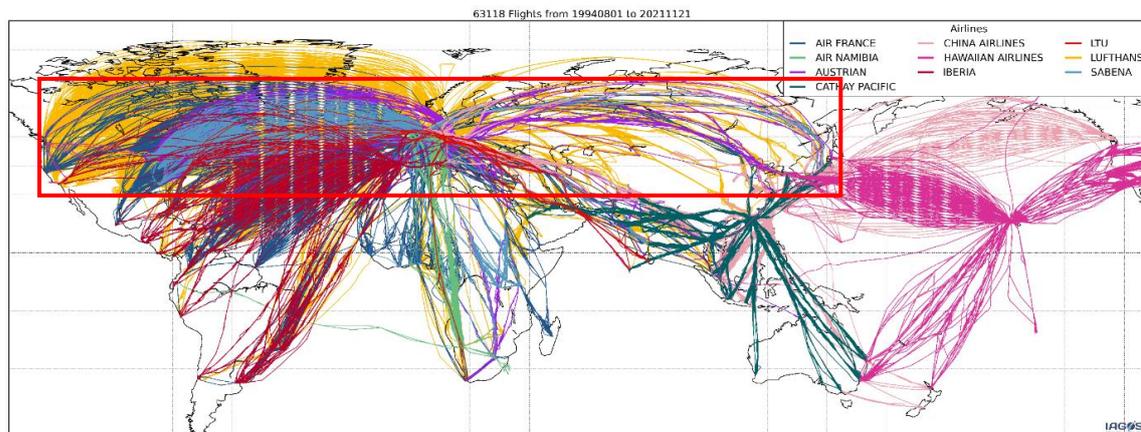


Figure 5: Map of all MOZAIC / IAGOS flights available. The red rectangle indicates the area where flights were analyzed.

Only data points with processing flag “2” and validity flag “0”, meaning only calibrated, non-faulty data, were considered. Temperature and relative humidity measurements are obtained with the MOZAIC Capacitive Hygrometer (MCH), consisting of a capacitive sensor (Humicap-h, Vaisala) for RH and a platinum resistor (PT-100) for T (Smit et al., 2008). Air temperature data (version 1.0) are calibrated following (Helten et al., 1998). For relative humidity data version 3.0 was used, meaning a correction was done additionally according to Smit et al. (2014), and to adjust the offset drift of the sensor during the flight period an in-flight calibration technique was introduced by Smit et al. (2008). Temperature data have a precision of ± 0.2 K and an accuracy of ± 0.5 K, while relative humidity is given with a precision of 1 % and an accuracy of 5 % (“IAGOS. Humidity sensor.,” n.d.; Neis et al., 2015). The humidity sensor’s response slows down with decreasing temperature, from 1s at 300K to 120s at 215 K (Neis et al., 2015). This implies strong inertia at the cruise level of aircraft. For this study this does not impose a problem, since we randomly selected only one percent of the data points, so around every 100th point, meaning that on a temporal scale two consecutive data points analyzed were ~ 400 s apart. This was done to ensure the independence of individual data points and to avoid autocorrelation. Values of $RH > 100$ % (RH with respect to liquid water) have been rejected. This would indicate flying through liquid water clouds, which do not exist at temperatures where contrail formation is possible. The selected MOZAIC / IAGOS measurements are combined with ERA-5 weather forecast data from the European Centre for Medium-Range Weather Forecast (ECMWF).

3.2 ERA-5 Reanalysis

The ERA-5 Reanalysis (Hersbach et al., 2018a, 2018b), as the fifth generation ECMWF reanalysis for the global climate and weather, is assumed to be amongst the best performing reanalysis products (Tarek et al., 2020). It provides global and consistent estimates from 1950 - present about a large number of atmospheric, land-surface, and ocean-wave variables, that are produced using 4D-Var data assimilation and model forecasts (CY41R2) of ECMWF's Integrated Forecast System (IFS). Every 12 hours the previous forecast is combined with past observations to get the best 4D numerical description of the recent climate state ("ERA5: data documentation - Copernicus Knowledge Base - ECMWF Confluence Wiki," n.d.).

ERA-5 is the successor to the ERA-Interim reanalysis (Hersbach and Dick, 2016), which stopped being produced on 31. August 2019. It has been improved in many regards, specifically in its spatiotemporal resolution, its assimilation system (from 31r2 to 41r2), and sources of data (Hoffmann et al., 2019). It is available as 10-member Ensemble of Data Assimilations (EDA) mean with uncertainty estimates in a $0.5^\circ \times 0.5^\circ$ ($1^\circ \times 1^\circ$ for ocean waves) and 3-hourly resolution. HRES, the high-resolution data product, is available hourly in $0.25^\circ \times 0.25^\circ$ horizontal resolution on 37 levels from 1,000 hPa to 1 hPa (Hersbach et al., 2018a, 2018b). ERA-5 Reanalysis is divided into four main subsets: hourly and monthly products with quantities in pressure and single levels.

In this study hourly ERA-5 HRES was chosen since it has a high temporal resolution while containing all necessary meteorological and radiation quantities for contrail analysis.

The data are publicly available and were retrieved from the Copernicus Data Service (Copernicus Climate Change Service (C3S), 2017) via API request.

For pressure levels 200, 225, 250, and 300 hPa, the variables temperature, relative humidity, specific cloud ice water content, divergence, vorticity, potential vorticity, vertical velocity, and geopotential height, and for single levels the mean top downward shortwave radiation flux, the mean top net short wave radiation flux, and the mean top net longwave radiation flux in $1^\circ \times 1^\circ$ resolution have been selected. More information about the variables and their units can be found in tables 1 and 2.

ERA-5 data has been validated against MOZAIC in-situ data in studies before (Gierens et al., 2020; Schumann et al., 2021; Teoh, 2020). Figure 7, left, shows that from ERA-5 and MOZAIC computed RH_{max} , which decides on contrail formation, is in good agreement (Gierens et al., 2020). The criterion for contrail persistence (RH_i) however, is known to be underrepresented in ECMWF's reanalysis. There are stronger and unsystematic differences in relative humidity and in particular high supersaturation values are not reproduced well enough by the model.

Figure 7 right and figure 8 left illustrate that RH_i from MOZAIC has a very weak correlation with RH_i from ERA-5 or EDA. The small blue tail starting from $RH_i \approx 1.25$ in figure 8 right shows where high supersaturation values are not covered by EDA (Gierens et al., 2020; Kaufmann et al., 2018; Schumann et al., 2021; Schumann and Graf, 2013; Teoh, 2020). This issue will be further discussed in chapter 6.

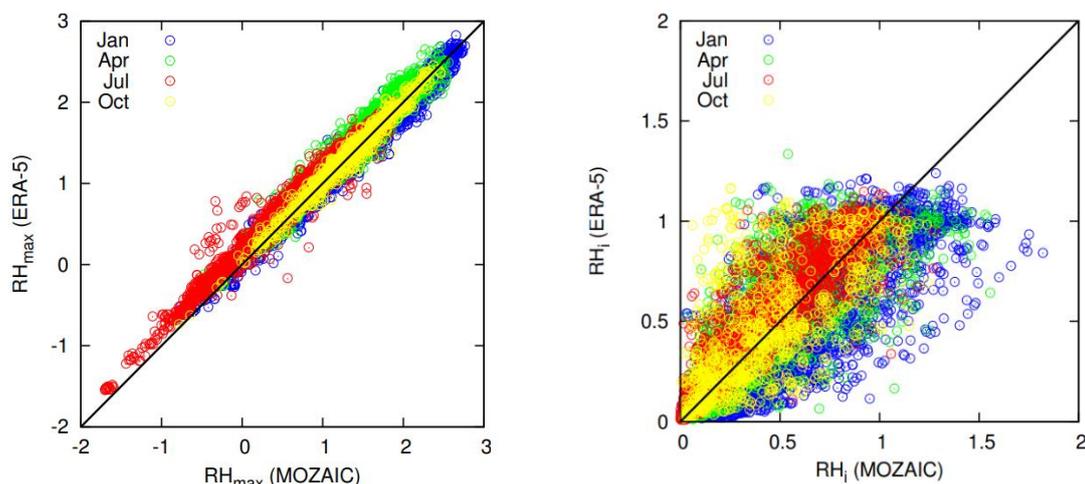


Figure 7: Comparison of RHmax (left) and RH_i (right) on single points between ERA-5 and MOZAIC. Strong linear relation with RHmax but not with RH_i. Figure from (Gierens, Matthes and Rohs, 2020) page 6+7.

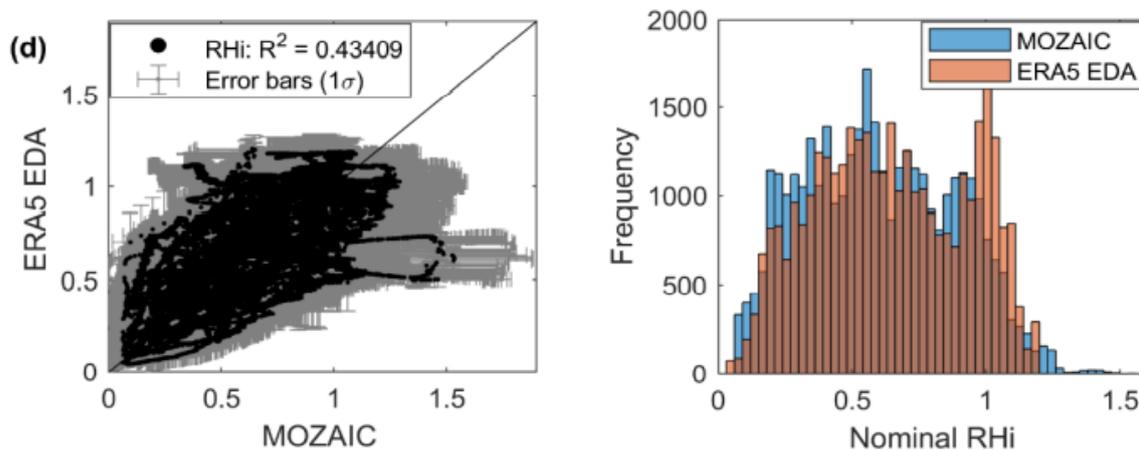


Figure 8: Validation of RH_i between ERA-5 and MOZAIC datasets. Left: Similar to the figure above no real correlation with RH_i. Right: The tail at high supersaturation values at MOZAIC is not represented in ERA-5. Figure from (Teoh, 2020), page 153

Table 1: Description of used variables from the ERA-5 single level dataset (Hersbach et al., 2018b).

Single Level		
Name	Unit	Description
Mean top downward short-wave radiation flux	Wm^{-2}	This parameter is the incoming solar radiation (also known as shortwave radiation), received from the Sun, at the top of the atmosphere. It is the amount of radiation passing through a horizontal plane. This parameter is a mean over a particular time period (the processing period) which depends on the data extracted. For the reanalysis, the processing period is over the 1 hour ending at the validity date and time. The ECMWF convention for vertical fluxes is positive downwards.
Mean top net long-wave radiation flux	Wm^{-2}	The thermal (also known as terrestrial or longwave) radiation emitted to space at the top of the atmosphere is commonly known as the Outgoing Longwave Radiation (OLR). The top net thermal radiation (this parameter) is equal to the negative of OLR. This parameter is a mean over a particular time period (the processing period) which depends on the data extracted. For the reanalysis, the processing period is over the 1 hour ending at the validity date and time. The ECMWF convention for vertical fluxes is positive downwards.
Mean top net short-wave radiation flux	Wm^{-2}	This parameter is the incoming solar radiation (also known as shortwave radiation) minus the outgoing solar radiation at the top of the atmosphere. It is the amount of radiation passing through a horizontal plane. The incoming solar radiation is the amount received from the Sun. The outgoing solar radiation is the amount reflected and scattered by the Earth's atmosphere and surface. This parameter is a mean over a particular time period (the processing period) which depends on the data extracted. For the reanalysis, the processing period is over the 1 hour ending at the validity date and time. The ECMWF convention for vertical fluxes is positive downwards.

Table 2: Description of used variables from the ERA-5 pressure level dataset (Hersbach et al., 2018a).

Pressure Level		
Name	Unit	Description
Divergence	s^{-1}	This parameter is the horizontal divergence of velocity. It is the rate at which air is spreading out horizontally from a point, per square met. This parameter is positive for air that is spreading out, or diverging, and negative for the opposite, for air that is concentrating, or converging (convergence).
Geopotential	m^2s^{-2}	This parameter is the gravitational potential energy of a unit mass, at a particular location, relative to mean sea level. It is also the amount of work that would have to be done, against the force of gravity, to lift a unit mass to that location from mean sea level. The geopotential height can be calculated by dividing the geopotential by the Earth's gravitational acceleration, g ($=9.80665 m s^{-2}$). The geopotential height plays an important role in synoptic meteorology (analysis of weather patterns). Charts of geopotential height plotted at constant pressure levels (eg., 300, 500 or 850 hPa) can be used to identify weather systems such as cyclones, anticyclones, troughs, and ridges. At the surface of the Earth, this parameter shows the variations in geopotential (height) of the surface and is often referred to as the orography.
Potential vorticity	$K m^2 kg^{-1} s^{-1}$	Potential vorticity is a measure of the capacity for air to rotate in the atmosphere. If we ignore the effects of heating and friction, potential vorticity is conserved following an air parcel. It is used to look for places where large windstorms are likely to originate and develop. Potential vorticity increases strongly above the tropopause and therefore, it can also be used in studies related to the stratosphere and stratosphere-troposphere exchanges. Large windstorms develop when a column of air in the atmosphere starts to rotate. Potential vorticity is calculated from the wind, temperature, and pressure across a column of air in the atmosphere.
Relative humidity	%	This parameter is the water vapor pressure as a percentage of the value at which the air becomes saturated (the point at which water vapor begins to condense into liquid water or deposition into ice). For temperatures over $0\text{ }^{\circ}C$ (273.15 K) it is calculated for saturation over water. At temperatures below $-23\text{ }^{\circ}C$ it is calculated for saturation over ice. Between $-23\text{ }^{\circ}C$ and $0\text{ }^{\circ}C$ this parameter is calculated by interpolating between the ice and water values using a quadratic function.

Specific cloud ice water content	kg kg^{-1}	This parameter is the mass of cloud ice particles per kilogram of the total mass of moist air. The 'total mass of moist air' is the sum of the dry air, water vapor, cloud liquid, cloud ice, rain and falling snow. This parameter represents the average value for a grid box. Water within clouds can be liquid or ice, or a combination of the two. Note that 'cloud frozen water' is the same as 'cloud ice water'.
Temperature	K	This parameter is the temperature in the atmosphere. It has units of kelvin (K). Temperature measured in kelvin can be converted to degrees Celsius ($^{\circ}\text{C}$) by subtracting 273.15. This parameter is available on multiple levels through the atmosphere.
Vertical velocity	Pa s^{-1}	This parameter is the speed of air motion in the upward or downward direction. The ECMWF Integrated Forecasting System (IFS) uses a pressure based vertical coordinate system and pressure decreases with height, therefore negative values of vertical velocity indicate upward motion. Vertical velocity can be useful to understand the largescale dynamics of the atmosphere, including areas of upward motion/ascent (negative values) and downward motion/subsidence (positive values).
Vorticity (relative)	s^{-1}	This parameter is a measure of the rotation of air in the horizontal, around a vertical axis, relative to a fixed point on the surface of the Earth. On the scale of weather systems, troughs (weather features that can include rain) are associated with anticlockwise rotation (in the northern hemisphere), and ridges (weather features that bring light or still winds) are associated with clockwise rotation. Adding the effect of rotation of the Earth, the Coriolis parameter, to the relative vorticity produces the absolute vorticity.
Divergence	s^{-1}	This parameter is the horizontal divergence of velocity. It is the rate at which air is spreading out horizontally from a point, per square meter. This parameter is positive for air that is spreading out, or diverging, and negative for the opposite, for air that is concentrating, or converging (convergence).

4. METHODS

4.1 Statistics of iRF

Once the data was acquired, both were combined to look for potential persistent contrails along the ten years of data. Beginning with MOZAIC, for every data point along each of the 16,588 flights, T and RH_i were used to check for the Schmidt-Appleman (Schumann, 1996) criterion, meaning that contrail formation was possible when $RH_{max} \geq 100\%$ (depending on T_{max}), and for ice supersaturation, meaning that a contrail can become persistent when $RH_i \geq 100\%$. T_{max} , the maximum temperature at which contrail formation is possible, and RH_{max} , the relative humidity at T_{max} , are calculated with an IDL program from Klaus Gierens following Schumann (1996):

$$T_{max} = [-46.46 + 9.43 * \log(G - 0.053) + 0.720 * (\log(G - 0.053))^2] + T_0, \quad (4.1)$$

where $T_0 = 273.5$ is used to transfer from °C in K. The factor G is the same as in the Schmidt-Appleman criterion from chapter 2.1.2. It is the slope of the phase trajectory and describes the change of partial pressure of water vapor in the mixture with changing temperature.

RH_{max} is calculated with the following equation:

$$RH_{max} = \left\{ 1 + \left[\frac{((RH * e_w(T) - e_w(T_{max})) - (G * (T_{max} - T)))}{e_w(T_{max})} \right] \right\}, \quad (4.2)$$

where

$$\log e_w = 54.842763 - \frac{6763.22}{T} - 4.21 * \log T + 0.000367 * T + \tanh[0.0415(T - 218.8)] * \left(53.878 - \frac{1331.22}{T} - 9.44523 \log T + 0.014025 T \right), \quad (4.3)$$

Since only RH with respect to liquid was available from MOZAIC, the RH_i was calculated following Murphy and Koop (2005):

$$RH_i = RH * \frac{e_w(T)}{e_i(T)}, \quad (4.4)$$

with

$$\log e_i = 9.550426 - \frac{5723.265}{T} + 3.53068 \log T - 0.00728332 T. \quad (4.5)$$

For ERA – 5 RH is already given as RH_i at temperatures below -23 °C, and since altitudes where contrail formation happens have temperatures lower than -23 °C, no corrections had to be made here.

If a persistent contrail was possible, for every point of MOZAIC's flights, the necessary atmospheric properties from ERA were determined at the corresponding time and position of the flight point. This was done by quadrilinear 3D / 4D- interpolation. Variables included were the specific cloud ice water content (ciwc), the mean top net longwave radiation flux for the outgoing longwave radiation (OLR), the mean top downward shortwave radiation flux to get the solar direct radiation (SDR), and the mean top net shortwave radiation flux for the reflected solar radiation (RSR).

Of those points where a persistent contrail was possible, only about 1 % of all data points (every 100th point) were chosen randomly by using a (0,1)-uniform random number generator, where individual records were only selected when the random variate was smaller than 0.01. As mentioned before, this made sure that the selected points were (mostly) independent from each other.

The next step was to compute the instantaneous radiative forcing, so the net forcing, for each of these points, to produce a large dataset of iRF values. The methodology of iRF calculation in this study is the same as in Gierens et al. (2020) and Wilhelm et al. (2021). Radiation formulae from (Schumann et al., 2012), with parameters for Myhre particles (table 3), were used. Myhre particles are pseudo-particles with wavelength-independent constant optical properties in a hexagonal, column-shaped form (Myhre et al., 2009).

Longwave radiative forcing (RF_{LW}) was calculated with:

$$RF_{LW} = [OLR - k_T(T - T_0)] \times \{1 - \exp[-\delta_{\tau} F_{LW}(r_{eff})\tau]\} E_{LW}(\tau_c) \geq 0, \quad (4.6)$$

and mainly depends on OLR, T (atmospheric temperature at the altitude of the contrail in K), τ (contrail optical depth at 550 nm), τ_c (optical depth of the cirrus above the contrails at 550 nm), and r_{eff} (effective contrail ice particle radius in μm). The factor that approximates the size dependence of the optical properties, is defined as

$$F_{LW}(r_{eff}) = 1 - \exp(-\delta_{lr} r_{eff}), \quad (4.7)$$

but is set to 1 in the calculations, since for Myhre particles δ_{lr} is infinite. The factor E_{LW} accounts for the reduction of the OLR at the contrail level because of a cirrus above the contrail and is calculated with:

$$E_{LW}(\tau_c) = \exp(-\delta_{lc} \tau_c) \quad (4.8)$$

Shortwave radiative forcing is a function of A_{eff} , τ , τ_c , μ ($\cos(\theta)$) and r_{eff} in Schumann's equations:

$$RF_{SW} = -SDR(t_A - A_{eff})^2 \alpha_c(\mu, \tau, r_{eff}) E_{SW}(\mu, \tau_c). \quad (4.9)$$

The effective albedo A_{eff} was set to zero at night and defined with $\frac{RSR}{SDR}$ during daylight hours. α_c stands for the albedo of the contrail, which is calculated from values for effective optical depth, depending on r_{eff} and the cosine of the solar zenith angle μ :

$$\alpha_c(\mu, \tau, r_{eff}) = R_C(\tau_{eff}) [C_\mu A_\mu R'_C(\tau') F_\mu(\mu)] \quad (4.10)$$

$$\tau' = \tau F_{SW}(r_{eff}) \text{ and } \tau_{eff} = \tau' / \mu \quad (4.11)$$

$$F_{SW}(r_{eff}) = 1 - F_r [1 - \exp(-\delta_{sr} r_{eff})], \quad (4.12)$$

where $R_C(\tau_{eff})$ accounts for reflectances:

$$R_C(\tau_{eff}) = 1 - \exp(-\Gamma \tau_{eff}) \text{ and } R'_C(r_{eff}) = \exp(-\gamma \tau_{eff}), \quad (4.13)$$

and $F_\mu(\mu)$ for sideward scattering as a function of μ : (Schumann et al., 2012)

$$F_\mu(\mu) = \frac{(1-\mu)^{B\mu}}{(1/2)^{B\mu}} - 1. \quad (4.14)$$

The factor $E_{SW}(\mu, \tau_c)$ includes the optical depth of the surrounding cirrus and its effective value $\tau_{c,eff}$, where $\tau_{c,eff} = \tau_c / \mu$:

$$E_{SW}(\mu, \tau_c) = \exp(\delta_{sc} \tau_c - \delta'_{sc} \tau_{c,eff}). \quad (4.15)$$

To be consistent with ERA – 5 the solar constant was set to 1361 Wm^{-2} . The parameters $T_0, k_T, \delta_\tau, \delta_{lc}, \delta_{lr}$ for RF_{LW} and $t_A, \Gamma, A_\mu, B_\mu, C_\mu, F_r, \delta_{sr}, \delta_{sc}, \delta'_{sc}$ for RF_{SW} are taken from table 3 from Schumann et al. (2012). RF_{LW} was restricted to positive and RF_{SW} to negative values only.

For the optical thickness of contrails τ and nearby cirrus τ_c , the formulation by Ebert and Curry (1992) was used. A vertical extension of 500 m was assumed for contrails. All water in excess of ice supersaturation was presumed to condense as ice with an effective radius of ice crystals (r_{eff}) of 30 μm . For the surrounding cirrus the optical thickness was computed from the specific ice water content from ERA – 5, assuming a vertical extension of 1500 m and an $r_{eff, cirrus}$ of 60 μm . The overall propulsion efficiency was set to $\eta = 0.35$. IRF was then calculated as a sum of RF_{SW} and RF_{LW} for every point where MOZAIC diagnosed a persistent contrail, yielding a total of 47,032 iRF values. The dataset also included other meteorological variables and

optical thickness values, as well as all of those based on ERA – 5’s T and RH_i for later comparison purposes. Since all contrails and contrail cirrus were computed in the same way, so with the same “model”, model uncertainty did not play a role in this study.

Table 3: Model parameters for various crystal habits. The parameter values assume that temperatures are given in K, fluxes are in Wm^{-2} , and radii in μm . For this study parameters for Myhre particles were used. Figure from (Schumann et al., 2012).

Parameter	Spheres	Solid	Hollow	Rough	Rosette	Plate	Droxtal	Myhre
LW parameters	k_T	1.935	1.955	1.960	1.959	1.944	1.951	1.946
	T_0	152	153	153	152	152	152	153
	δ_r	0.941	0.808	0.736	0.676	0.749	0.709	0.796
	δ_{lr}	0.211	0.341	0.325	0.256	0.170	1.654	0
	δ_{lc}	0.160	0.096	0.092	0.046	0.133	0.087	0.063
	t_A	0.879	0.902	0.882	0.899	0.880	0.883	0.899
	Γ	0.242	0.347	0.288	0.297	0.328	0.438	0.275
SW parameters	γ	0.323	0.393	0.356	0.345	0.408	0.524	0.311
	A_μ	0.361	0.294	0.344	0.318	0.337	0.311	0.343
	B_μ	1.676	1.557	1.711	1.558	1.708	1.718	1.564
	C_μ	0.709	0.678	0.688	0.675	0.712	0.713	0.660
	F_r	0.512	0.577	0.597	0.226	0.551	0.818	0.249
	δ_{sr}	0.150	0.025	0.024	0.046	0.048	0.070	0.052
	δ_{sc}	0.157	0.143	0.168	0.149	0.173	0.162	0.172
δ_{sc}	0.230	0.198	0.245	0.205	0.248	0.254	0.244	

Of course, all of the assumptions and parameters influence the values of iRF and will differ when using different choices. As an example, figure 9 shows the cumulative distribution function of the iRF of 2005 with different r_{eff} and $r_{\text{eff, cirrus}}$ used for the calculations. The red line, where a r_{eff} of $20 \mu\text{m}$ and $r_{\text{eff, cirrus}}$ of $60 \mu\text{m}$ was chosen, shows stronger iRF values than those with the assumptions used in the study. Note that the blue line, where the $r_{\text{eff, cirrus}}$ was additionally reduced to $50 \mu\text{m}$, sits right beneath the red curve with only very small differences that are not visible in the plot. This is due to the fact that the optical thickness of the surrounding cirrus is only a small factor in the whole iRF calculation.

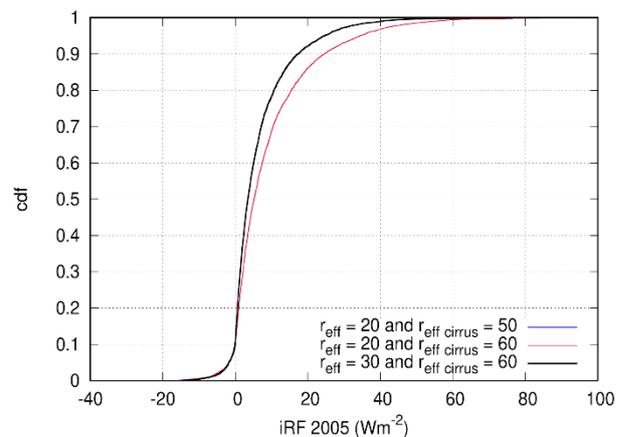


Figure 9: Cumulative distribution function of the iRF for the whole year of 2005 with different effective radii for ice crystals. Note that the blue line sits right beneath the red one.

What is important is that the specific values of iRF are not of interest here. Only the type of distribution and their mutual larger-smaller relation, such that a stronger contrail has a larger iRF than a weaker one and an exothermic contrail (cooling) has a negative iRF, was important. The assumptions and special choices for constants did not impact this relation.

For statistically analyzing the iRF dataset, statistical key figures were determined and cumulative distribution functions (cdf’s), probability density functions (pdf’s), and a first-order

effect function, where the pdf was multiplied with the iRF value itself, were created using IDL 8.4, GNUPLOT 5.2 and Microsoft Excel 1808.

The probability densities were produced with an Epanechnikov (Silverman, 1998) kernel density estimator, with an IDL program written by David G. Grier, Henrique Moyses, David Ruffner, and Chen Wang (Appendix A). After setting one hundred bins for the yearly pdf's, and five hundred for the pdf of the whole dataset, the program calculated optimal bin sizes, with maximum and minimum following the input data. For the overall pdf of years 2000 – 2009 an exponential function was fitted to the positive branch of the pdf. Lastly, two Monte Carlo experiments were conducted, where 1,000 samples from the iRF dataset with 1 % and 0.1 % sample size were chosen randomly and each mean value was calculated. This was done to investigate the spread of the mean values and to show how much an average depends on sample size. Results will be discussed in chapter 5.1.

4.2 Meteorological conditions of Big Hits

The base of the meteorological analysis and characterization of Big Hits was to produce an iRF dataset as well. This was done using a similar method than previously described in chapter 4.1, the difference being, that now, meteorological conditions were also determined for each potential persistent contrail. Also, not only every 100th data point but every single point, where persistent contrails were possible, was now considered. For all of those points the dynamical quantities vertical velocity, vorticity, potential vorticity, divergence, and geopotential height were determined from ERA – 5 reanalysis. As Z , the geopotential height, varies on each isobaric level, it was normalized by a mean pressure scale height Z_p^* for easier comparison. The normalized geopotential height $Z_{normalized}$ is calculated by

$$Z_{normalized} = Z'/Z_p^*, \quad (4.16)$$

with

$$Z' = Z/g, \quad (4.17) \quad \text{and} \quad Z_p^* = \bar{H} * \log(1000/p), \quad (4.17)$$

where p is the pressure level in hPa, Z is the geopotential height given by ERA – 5 in m^2s^{-2} and g is the acceleration due to gravity = 98.1 dam s^{-2} . The mean scale height \bar{H} is a function of the gas constant for dry air $R = 0.28704 \text{ J g}^{-1}\text{K}^{-1}$ and the mean tropospheric temperature $\bar{T} = 255 \text{ K}$.

$$\bar{H} = R\bar{T}/g = 746 \text{ dam}, \quad (4.18)$$

The normalisation of Z results in unified ranges of about ~ 0.9 to ~ 1.1 for each level.

iRF was then calculated in the exact same way, with the same assumptions for effective radii of ice crystals, the same vertical extensions for contrail and contrail cirrus, and the same propulsion efficiency as earlier. This was again done for both MOZAIC and ERA – 5, based on the respective temperature and relative humidity.

Since the object of interest was Big Hits, so the contrails with the strongest endothermic (warming) effect, only contrails with an iRF $\geq 19 \text{ Wm}^{-2}$ were added into the dataset. The dataset was made up of 13,995 Big Hit files that each represented a part of a flight trajectory, where the iRF was nearly constantly above 19 Wm^{-2} . A Big Hit file was produced, when, during an individual flight, the iRF value reached the 19 Wm^{-2} mark. This point, plus its interpolated variables, was written into a file, together with every following point that was also bigger than 19 Wm^{-2} along the flight. Once a gap of at least ten data points where iRF was lower than 19 Wm^{-2} was reached, which translates to a distance of ten kilometers or 40 seconds, the file was closed. This distance was chosen since realistically an aircraft is not able to change its flight route every 4 seconds to prevent the formation of a strongly warming contrail. It needs some time to adjust to a different height.

In this way, the program created single Big Hit files along each MOZAIC flight, sometimes finding multiple Big Hits, or none at all, during a flight. Each Big Hit file consisted of the location with coordinates and time, the iRF value, synoptic variables T, RH_i, RH_{max}, absolute humidity, and pressure, as well as the mentioned dynamical variables, optical thickness values for the contrail and the cirrus and μ , the cosine of the solar zenith angle.

From those files, statistics were generated with flights per Big Hit, Big Hits per year and season. All Big Hits were also plotted on two world maps, to see if specific spatial patterns of occurrence are visible. A correlation matrix was produced for all variables using RStudio Version 1.4.1717. For each meteorological variable, the cdf's and pdf's were produced as a mean per Big Hit and for all 450,000 points of all 13,995 files together. Pdf's were compared to the previous pdf's and to pdf's of the same meteorological variables from the first dataset, where only every 100th data point was considered. With this, it was possible to determine differences in meteorological conditions depending on whether ice supersaturation, a persistent contrail or a Big Hit was present.

A multiple logistic regression was done with RStudio for a first try to predict Big Hit occurrence by means of dynamical quantities. Logistic regression models have been used in many contrail prediction studies before (Carleton et al., 2015; Duda and Minnis, 2009a, 2009b; Jackson et al., 2001; Travis et al., 1997). Here the logistic function is used to model the conditional probability

of $Y = 1$ (= Big Hit occurrence), given the predictors $X = (X_1, \dots, X_p)$, which are the independent meteorological variables from ERA-5 reanalysis:

$$\log\left(\frac{p(X)}{1-p(X)}\right) = \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p, \quad (4.19)$$

where the left side is the *logit* or *log-odds*, with $p(x) = \Pr(Y = 1|X)$. This can be written as:

$$p(X) = \frac{e^{\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p}}{1 + e^{\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p}} \quad (4.20)$$

The advantage to linear multiple regression is, that with the logistic model each predictor can be fit nonlinearly to the dependent variable (predictand) and that the function's output only gives values between 0 and 1 (Duda and Minnis, 2009b; James et al., 2013). Maximum likelihood is used to estimate the regression coefficients $\beta = (\beta_0, \beta_1, \dots, \beta_p)$, and its p-values decided which predictors were included in the improved model.

For probability calculations, a third dataset was produced, in the same way as in chapter 4.1, but adding a calculation for the lapse rate around contrails from ERA-5 temperature and pressure values, derived from the barometric height formula, which describes how pressure p decreases with increasing altitude h :

$$\frac{dp}{p} = -\frac{g}{R_d} \frac{dh}{T(h)} \quad (4.21)$$

Here g is the gravitational acceleration (the slight height dependence will be neglected), R_d is the special gas constant for dry air and $T(h)$ is the temperature at altitude h . Once $T(h)$ is known, this expression can be integrated.

Since for the model data T is only available at certain pressure levels, eg. T_0 at p_0 (lower level) and T_1 at p_1 (higher level), a constant temperature gradient (or lapse rate) was assumed between these two levels. The height difference Δh and thus the lapse rate $(T_0 - T_1)/\Delta h := \gamma$ can be calculated with this assumption, which leads to the integral of the barometric height formula:

$$\ln\left(\frac{p_0}{p_1}\right) = \frac{g}{R_d} \int_{h_0}^{h_1} \frac{dh}{T_0 + \gamma(h-h_0)} \quad (4.22)$$

Substituting θ for $T_0 + \gamma(h - h_0)$ leads to

$$\ln\left(\frac{p_0}{p_1}\right) = \frac{g}{\gamma R_d} \int_{T_0}^{T_1} \frac{d\theta}{\theta} \quad (4.23)$$

that is

$$\ln\left(\frac{p_0}{p_1}\right) = \frac{g}{\gamma R_d} \ln\left(\frac{T_0}{T_1}\right). \quad (4.24)$$

Solving for γ gives the desired result

$$\gamma = \frac{g}{R_d} \frac{\ln(T_0/T_1)}{\ln(p_0/p_1)}. \quad (4.25)$$

As a last step, conditional probabilities were produced for the occurrence of persistent contrails. The probabilities were conditioned on meteorological quantities, but specifically those, where the distributions of non-persistent contrails were best separated from the distributions of persistent contrails and Big Hits. Conditional probabilities for persistent contrails (pC) and for Big Hits (BH = iRF \geq 19 Wm⁻²) can be given in the form P(pC|X = x \pm Δ x) or P(BH|X = x \pm Δ x), which is the probability to get a persistent contrail or a Big Hit in a situation where a certain meteorological quantity X obtains a value in the range [x - Δ x, x + Δ x]. For the Big Hit probabilities the Schmidt-Appleman criterion (SAC = 1 when RH_{max} \geq 1) and ice supersaturation (ISS = 1 when RH_i \geq 100 %) based on MOZAIC T and RH_i, is already fulfilled, such that the conditional probabilities are P(BH|X = x \pm Δ x) = P(BH|X = x \pm Δ x \wedge SAC = 1 \wedge ISS = 1), where the symbol \wedge is the logical AND. According to Bayes' law this expression can be calculated from

$$P(\text{BH}|X = x \pm \Delta x) = \frac{P(X = x \pm \Delta x|\text{BH}) P(\text{BH})}{P(X = x \pm \Delta x)}, \quad (4.26)$$

which is the probability that a certain variable X is in the range of x \pm Δ x when a Big Hit is given, multiplied with the apriori probability of Big Hits, divided by the probability that X is in the range of x \pm Δ x. For this study probabilities were not calculated but they were given as relative frequencies, as they can simply be counted:

$$P(\text{BH}|X = x \pm \Delta x) = \frac{N(X = x \pm \Delta x \wedge \text{iRF} \geq 19)}{N(X = x \pm \Delta x)}, \quad (4.27)$$

The numerator is the number of cases that have an X in the range of x \pm Δ x and at the same time are Big Hits and the denominator is the number of cases where X takes on a value in the range of x \pm Δ x.

This can similarly be calculated for persistent contrails, but then the condition is not iRF \geq 19 Wm⁻², given SAC = 1 and ISS = 1, but only SAC = 1 and ISS = 1 has to be fulfilled. In the dataset flight points only have an iRF if they are persistent, for all non-persistent cases

the iRF was set to - 999. Hence the persistence conditions $SAC = 1$ and $ISS = 1$ can also be expressed as an iRF condition, where iRF needs to be larger than - 999. For the range $X = x \pm \Delta x$, percentiles in 0.1 steps from every meteorological value that might improve the prediction of persistent contrails, were used. The probabilities were then produced in a two-dimensional matrix, where the ten percentiles of two meteorological variables spanned a 2-D space, which was divided into “boxes” of different size. This is better explained with a plot in chapter 5.2. For every box, the number of persistent contrails was divided by the number of all cases in that box, to get the box-specific probability for contrail persistence. Additional conditions were introduced for the whole matrix. Each point had to have a meteorological variable X of the value $x \pm \Delta x$. The values were again taken from the determined percentiles and were chosen based on the overall improvement in probability. The goal was to find areas of high probabilities, such that one could say that persistent contrail would develop with high probability when specific ambient meteorological conditions are given. Lastly a heatmap was produced from the probability matrix.

5 RESULTS

5.1 Statistics of iRF

5.1.1 Statistics

This chapter will present some statistics about the instantaneous radiative forcing (iRF) dataset that was produced as described in chapter 4.1. The dataset only includes persistent contrails, so the ones that fulfilled the Schmidt-Appleman criterion and showed ice supersaturation $\geq 100\%$, based on T and RH_i from MOZAIC in-situ measurements. ERA – 5 data were only used to calculate the iRF values, but do not play a role in diagnosing any potential persistent contrails in this dataset.

Starting with figure 10 on the left, one can see multiple differently colored cumulative distribution functions of the iRF of all persistent contrails for all months of the year 2004. The cdf's were also produced for every single month of the other years between 2000 and 2009, but since they show similar results, the year 2004 was chosen exemplary, as it includes the most iRF values. The black line combines all data of 2004. In the right plot, all data points of every single year are combined for yearly cdf's and in the black curve as one function for the whole dataset.

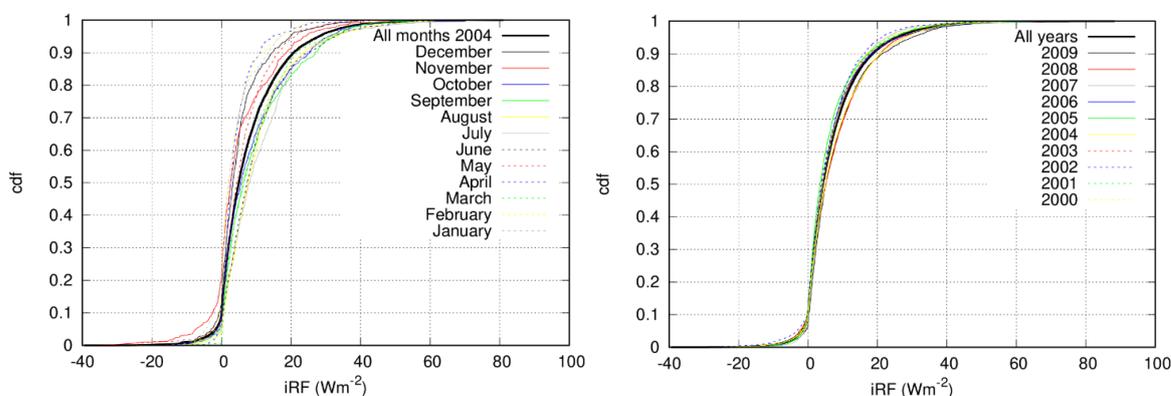


Figure 10: Cumulative distribution functions (cdf) of instantaneous radiative forcing from persistent contrails diagnosed by MOZAIC. The left plot shows all monthly cdf's (colored) of the year 2004 together with the cdf of the whole year 2004 in black. The right plot displays all yearly cdf's (colored) with a cdf for the whole dataset in black.

At first sight, it seems that the cdf's look rather similar, especially when looking at the yearly cdf's. All of the curves have a steep increase around zero, where most persistent contrails have a small positive iRF with values up to $\sim 20 \text{ Wm}^{-2}$ (80 % in the dataset), meaning that they have a small endothermic (warming) effect. Big Hits, where the iRF exceeds 19 Wm^{-2} happens in $\sim 10\%$ of the contrail cases. Only $\sim 10\%$ of all contrails in the dataset have an exothermic (cooling) effect with iRF below zero. Compared to the $\sim 90\%$ warming cases, cooling ones occur rather rarely. The iRF ranges from -40 Wm^{-2} to 90 Wm^{-2} , with a small positive mean of

7 Wm^{-2} and a standard deviation of 9.2 Wm^{-2} . When comparing the different cdf curves it becomes clear that there is a distinct variation with large annual and interannual variabilities. Of course, the annual variation seen in the monthly cdf's of 2004 (left) is stronger than the variation of the yearly curves (right), since they include a lot more data. The means of the single years still range from 5.8 Wm^{-2} to 8.3 Wm^{-2} . The values with their minimum, maximum, and standard deviation can be found in table 4. This variability, which is seen in the curves, is exclusively caused by weather variability, and specifically only by that part that defines the iRF since the only changing factors were the meteorological variables that iRF was computed from. All other uncertainty sources, including all other sources for natural variability, were deliberately neglected, leaving no room for model uncertainty.

Table 4: Statistical key figures from the persistent contrails for the individual years and the whole dataset. iRFSW all years includes the instantaneous shortwave radiative forcing of all of the persistent contrails and the same for the instantaneous longwave radiative forcing at iRFLW. Note that the iRF referred to in the text is the net iRF, which is the sum of iRFSW and iRFLW.

quantity	min	max	mean	std. dev.
iRF 2000	- 32.5	76.4	6.4	8.5
iRF 2001	- 28.2	76.7	6.7	8.7
iRF 2002	- 35.2	54.4	5.8	8.3
iRF 2003	- 27.5	88.3	6.8	8.9
iRF 2004	- 39.4	81.0	7.9	9.9
iRF 2005	- 16.3	75.8	6.3	8.9
iRF 2006	- 28.7	72.8	6.9	9.6
iRF 2007	- 25.0	57.7	6.7	8.0
iRF 2008	- 19.3	76.2	7.9	10.1
iRF 2009	- 24.5	78.6	8.3	10.1
iRFSW all years	- 48.9	0.0	-1.5	4.5
iRFLW all years	0.0	101.7	8.5	9.0
iRF_{net} all years	- 39.4	88.3	6.9	9.2

The interannual variation of the means that can be seen in figure 10 and table 4, is not the only variation in the iRF dataset. Figure 11 left shows the distributions of all months of all ten years (all Januaries, all Februaries, etc.) with their mean (black line) plus and minus one standard deviation (dotted line) and their 5th to 95th percentile range (dashed area). The monthly means show a prominent seasonal pattern, with the highest mean values in the summer and early fall during months June to September and the lowest ones in the cooler winter months during December to April. The maxima are two to three times as high as the minima.

Figure 10, left, as well as the wide range covered by the standard deviation and the 5th - 95th percentiles in this plot (figure 11, left) indicates that the twelve monthly distributions are quite broad. The 5 % percentile line is very similar to the mean minus one standard deviation, but the 95th percentile line is a lot higher, reflecting a skewness of the distribution. This will be further explained in the next paragraph. The width of the distributions is larger than the average seasonal amplitude, which means that the interannual variability, so the differences between the same months over different years, exceeds that variation of the mean in this figure. This should not be confused with the seasonal variation, which is in fact larger than the interannual variability. This can be seen when comparing figure 10, right with the following figure 11, right.

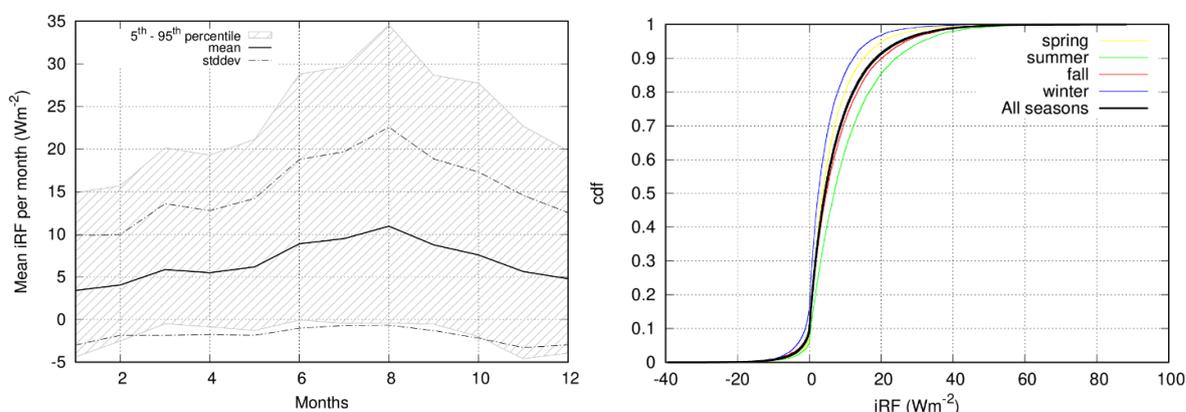


Figure 11: Left: Seasonal variation of the mean iRF per month (black line) plus/minus one standard deviation (dotted lines) and the 5th - 95th inter-percentile range (dashed area) for the years 2000 to 2009. Right: Distribution functions (cdf) for each season for the years 2000 to 2009 (colored) and the cdf for the whole dataset (black line).

It becomes obvious, that not only the mean, but the whole distribution has a strong seasonal variation and specifically the extreme values as well (figure 11, right). Displayed are cdf's for the four seasons spring (months March, April, May, yellow line), summer (Jun., Jul., Aug., green line), fall (Sep., Oct., Nov., red line), and winter (Dec., Jan., Feb., blue line) and again the cdf for the whole dataset (black line). The green line, which has the flattest increase of all the curves, shows that high iRF values and also the maxima occur predominantly in the summer months, while in winter the minima and more negative iRF happen (blue line).

These seasonal frequency distributions coincide with the numbers from table 5, where the frequencies of persistent, cooling, and warming contrails are given per season and in total. They suggest that most persistent contrails develop in summer and fall and specifically most warming contrails ($iRF > 0 \text{ Wm}^{-2}$) and strongly warming contrails ($iRF > 19 \text{ Wm}^{-2}$) occur in summer. In winter the least amount of persistent contrails develop, but together with the months from fall, most cooling contrails develop.

Table 5: Frequencies for persistent, cooling ($iRF < 0 \text{ Wm}^{-2}$) and warming contrails ($iRF > 0 \text{ Wm}^{-2}$) and Big Hits ($iRF > 19 \text{ Wm}^{-2}$) per season and for the whole time series

	persistent contrails	$iRF < 0 \text{ Wm}^{-2}$	$iRF > 0 \text{ Wm}^{-2}$	$iRF > 19 \text{ Wm}^{-2}$
Spring	11788	891	10911	611
Summer	12052	701	11346	1759
Fall	13494	1337	12149	1371
Winter	9698	1581	8116	311
Total	47032	4510	42522	4052

The maximum in summer could be caused by a larger temperature gradient between the Earth's surface and the flight levels in this season, which supports large longwave (infrared) forcing. Even though shortwave forcing can work longer during those long summer days, in winter there is a dominance of large solar zenith angles which produces the maxima in RF_{SW} because of a stronger back-scattering of solar radiation to space (Meerkötter et al., 1999) compared to summer. Another reason could be the changing height of the tropopause. In summer the tropopause height has its maximum in the northern hemisphere (Seidel and Randel, 2006). As flights don't change their flight levels according to seasons, in summer a bigger fraction of flights passes through tropopause regions that typically have higher humidity values than the overlaying lower stratosphere. Most ice supersaturated regions (ISSR's) can be found in a layer 200 hPa down from the tropopause (Petzold et al., 2020; Spichtinger et al., 2003a). The high humidities result in higher optical thicknesses and consequently larger iRF values. Unfortunately, tropopause height was not considered in the calculations, which leaves no options for further analyses regarding this point.

In the literature contradicting results are found about the seasonality of RF. While Gettelman et al. (2021) and Stuber et al. (2006) find that the largest cooling occurs in summer and the largest warming in winter, Bock and Burkhardt (2016) have a small RF maximum in the summer months because of more optically thick contrail cirrus in this season. Since in this study only iRF and not ERF or RF was considered, the different results must not mean that either is "wrong". ERF is the integral over all iRF 's for the whole lifecycle and coverage of all contrails, but this relation is not well investigated so far. This problem will be discussed further in chapter 6.1.

For a better look at the distributions of the iRF , probability density functions were produced with a kernel density estimator. Figure 12, left, shows the pdf of the iRF of all persistent contrails from years 2000 to 2009 in differently colored lines and for the whole time series in black. All of the functions have a peak around zero, but with interannual differences of the probability densities. The pdf's have a long tail towards positive values and are heavily skewed

to the right, due to the many large iRF events. The overall pdf's skewness of 3.3 and its kurtosis of 11.3, with a standard deviation of 0.18, implies that the tail is very heavy and exceeds the tail of the normal distribution. Such a distribution is also called leptokurtic (from “Lepto-“, meaning slender) (Eckstein et al., 1994). It becomes especially clear, that the pdf's distribution is not a typical Gauss distribution when looking at figure 12, right. Here the pdf's are plotted on a logarithmic y-axis. Obviously, there is a lot of noise at strongly positive (50 Wm⁻²) or negative values (- 20 Wm⁻²), since there are not many values, but it was still possible to fit a straight line to the positive branch of the pdf, starting from 1 Wm⁻².

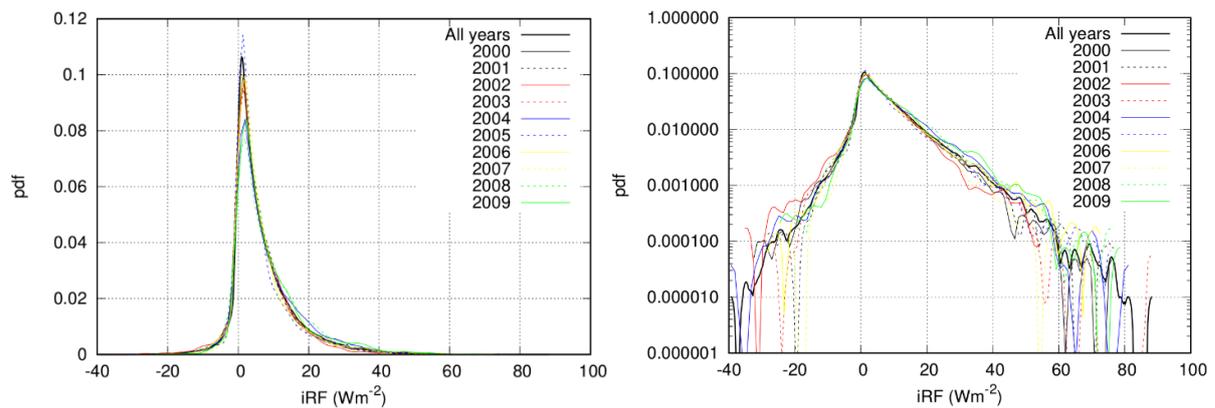


Figure 12: Probability density functions for the iRF of years 2000 - 2009 (colored) and for the whole dataset (black). Left: Normal depiction. Right: Y-axis in logarithmic depiction.

This mathematical fit (figure 13) with the function $g(x) = a - b * x$ and its two factors $a = -2.38 \pm 0.04$ ($\pm 1.6\%$) and $b = 0.11 \pm 0.001$ ($\pm 0.69\%$) fits quite well to the data. The fit shows that the positive branch of the pdf of iRF follows an exponential decline. This can be written as:

$$f(x|x \geq 1) \propto e^{-bx}, \quad (5.1)$$

where $b = 0.11$ is the factor from before and x is the iRF in Wm⁻².

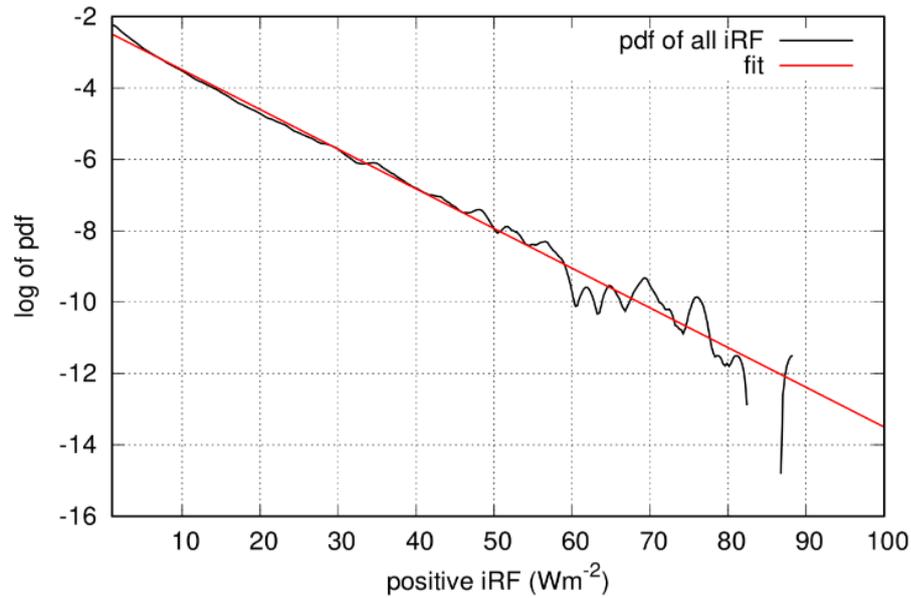


Figure 13: Black line: Pdf of the positive iRF values of years 2000 - 2009 plus its mathematical linear fit (red line). The straight-line fit indicates an exponential decline of the positive branch of the pdf.

This exponential distribution could be passed on from the ice supersaturation that is also exponentially distributed in nature (Gierens et al., 1999; Spichtinger et al., 2002), which is reproduced in the ERA – 5 reanalysis dataset. Since the ice water content of contrails is a function of supersaturation and temperature, and building on that the optical thickness as well, it is no surprise, that τ and τ_c for positive iRF are nearly exponentially distributed in the produced dataset. This can be seen in figure 14, where the pdf of the optical thickness where $iRF \geq 1$ (black line) is plotted with its fit (red line) in the same half-logarithmic depiction. The fit is not as good, but still has very low asymptotic standard errors for the factors $a = 1.4 \pm 0.06$ ($\pm 4.5\%$) and $b = 6 \pm 0.66$ ($\pm 1.1\%$). As explained in chapter 4.1 in the radiative forcing algorithm from Schumann et al. (2012) τ has a major influence on iRF and this might be the reason for this exponential distribution of the positive branch of the pdf.

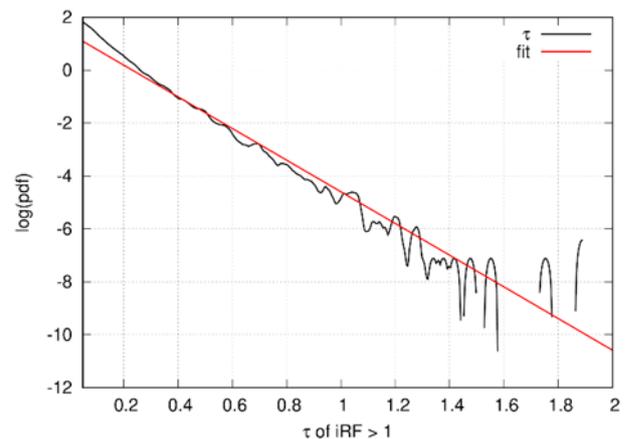


Figure 14: Pdf of the optical thickness of the contrails where iRF is larger than one (black) and its mathematical fit (red) in a half-logarithmic plot.

5.1.2 Monte-Carlo Experiments

The annual global ERF or RF in the IPCC charts is always given as model calculations. Hence the large error bars associated with contrail cirrus RF stems from differences in model results that are produced by the many uncertainties listed in Lee et al. (2021, Appendix E). But why are we using models for RF if the uncertainties are so large? As mentioned before, to get from iRF to the global ERF / RF one must integrate for all atmospheric contrails during a year, or in this case ten years, the iRF's over the whole area of each contrail. Obviously, it is not possible to measure every single contrail with the current state of research, which is why models are used. However, measurement samples can be taken. The question arises, whether it is possible to get a better estimate for RF / ERF by using data from measurement samples. Schumann et al. (2017) have compiled 230 in situ and remote sensing measurements from all performed contrail measurement campaigns until 2017. If one were to calculate a mean value over those 230 measurements for iRF, how precise would it be? As only samples are used for the average, different samples may result in different averages. To investigate this, two Monte Carlo experiments were performed based on the iRF dataset for persistent contrails, to show how difficult it is to determine a precise long-term average and how much such an average depends on the sample measurement size. For this different samples were chosen from the 47,032 iRF values and their means were calculated, to look at their spread. Two sample sizes were chosen, one that included 0.1 % of the data (~ 47 measurements), which represents the size of a single measurement campaign, and one that is typical for a meta-study of ten measurement campaigns together, including 1 % of the data (~ 470 measurements). For each sample size, 1,000 samples were randomly chosen with a random number generator and then their sample means were calculated. The means are plotted in figure 15 as plus signs, with the blue line showing the mean of the sample means and its standard deviation as red lines.

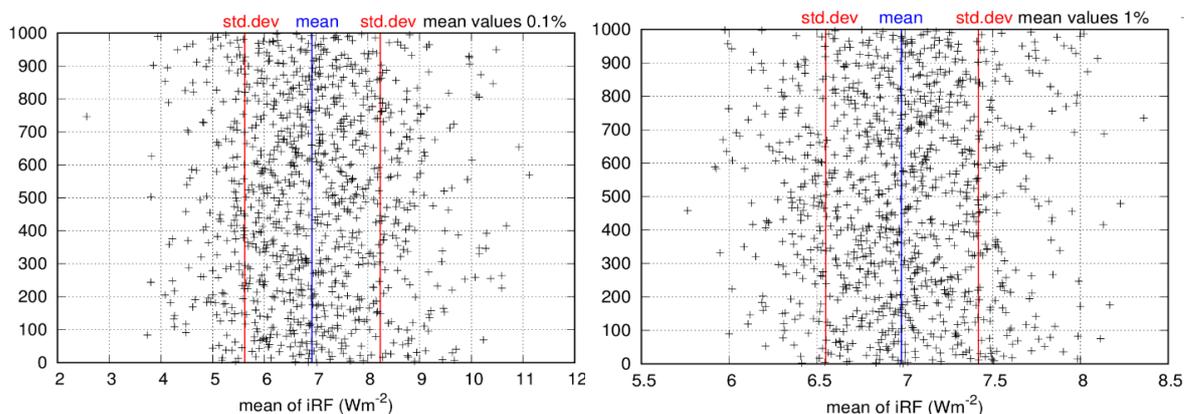


Figure 15: Sample means (plus signs) of 1,000 samples of sample size 0.1 % (left) and 1 % (right) from the whole iRF dataset of persistent contrails. The blue line is the mean of all sample means and the red lines indicate the plus/minus one standard deviation range. Note the different scales on the x-axis.

Depending on the chosen sample size, the scatter of the means is huge. They range from $\sim 2 \text{ Wm}^{-2}$ to $\sim 12 \text{ Wm}^{-2}$ for the smaller 0.1 % sample and from $\sim 5.5 \text{ Wm}^{-2}$ to $\sim 8.5 \text{ Wm}^{-2}$ for the bigger 1 % sample. These considerable variations are solely caused by the random character of the weather. This weather-induced variability is irreducible and it is the lower limit of the natural variability of iRF since all other sources of variability were deliberately neglected. If one were to include more degrees of freedom, eg. by adding the technical variabilities of the aircraft or the synoptic situation, even bigger variabilities and uncertainties would be the result. The averages of the sample means (blue lines) are very similar to the overall mean of 7 Wm^{-2} for the whole iRF dataset, but this value is often not represented by the sample means. According to the central limit theorem, the sampling distribution of a mean is approximately normally distributed if the sample size is large, meaning of size 30 or more, even if the original distribution, here the iRF distribution, is not normal. As this is the case here, the sample mean becomes closer to the real mean with increasing sample size. The standard deviation of a mean is $\frac{\sigma}{\sqrt{N}}$, where $\sigma = 9.2 \text{ Wm}^{-2}$ is the overall standard deviation of the whole dataset and N is the size of the sample, so $N \approx 470$ for the 1 % sample and $N \approx 47$ for the 0.1 % sample. For the larger sample the standard deviation is 1.31 Wm^{-2} and for the smaller sample 0.43 Wm^{-2} (red lines), which is the same when using the $\frac{\sigma}{\sqrt{N}}$ formula. This implies that the mean value of the overall iRF dataset is defined very precisely, but for realistic samples that could be obtained in measurement campaigns or even in a metastudy, the N is rather of the size that was used in the Monte Carlo experiments. These demonstrate, that the mean is no longer determined with high precision. Hence, the uncertainty could in principle be reduced by taking larger sample sizes, but as the improvement is only proportional to $\frac{1}{\sqrt{N}}$, this would require an unrealistically large number of measurements.

5.1.3 Effect of Big Hits and maximum iRF

The iRF dataset with its 47,032 values is enough to determine a precise average. Nevertheless, the ten years of data taken to calculate the iRF are still only a sample. It is possible, that values outside the observed range occur in reality. Especially iRF values that exceed the maxima in the dataset are of interest because the contrails with the strongest iRF's contribute the most to the overall RF of contrails. This is well portrayed in figure 16, where the pdf of the whole iRF dataset (black line) together with its first-order effect function of the iRF (red line) are depicted.

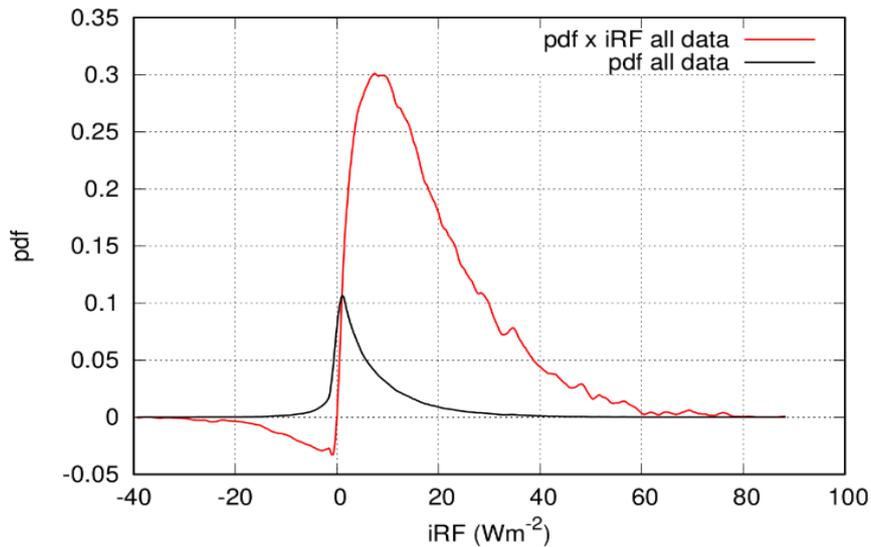


Figure 16: Pdf of the whole dataset (black line) with its first-order effect function (red) of the iRF (pdf \times iRF). This shows the disproportionately large effect of high iRF values on the overall RF, despite their low probabilities.

The effect function is produced by multiplying the probability density with the iRF value, to determine its total effect TE :

$$TE = \int_{-\infty}^{\infty} x f(x) dx, \quad (5.2)$$

which is the sum of all single effect sizes x times their (relative) frequency. The partial effect PE of all contrails between two values x_1 to x_2 can be given by the integral

$$PE = \int_{x_1}^{x_2} x f(x) dx. \quad (5.3)$$

In the plot, the black line of the pdf becomes barely visible from 40 Wm^{-2} up, because of its low probabilities. If those are multiplied with its effect, it becomes obvious that even though cases with high iRF have a low probability, they contribute a lot to the overall effect. This large tail is proof that strongly warming contrails have a disproportionately large impact on the total endothermic climate effect of contrails.

Knowing this, one might ask whether even higher iRF values than the observed maximum of 90 Wm^{-2} can occur when using a different sample. The following paragraph illustrates how the “true” maximum of iRF can be estimated from the dataset with the help of extreme value theory (Sornette, 2000, p. 167).

The results from chapter 5.1.1 show that the positive branch of the pdf of the whole iRF dataset is exponentially distributed. Using this, the probability, that a certain iRF exceeds x is $1 - F(x)$, with $F(x)$ being the distribution function of iRF:

$$F(x) = 1 - \exp(-bx), \quad (5.4)$$

where $b = 0.11^{-1}$ (Wm^{-2}) is the factor of the exponential fit from chapter 5.1.1. Note that, as in the fit before, only values above 1 Wm^{-2} are used, leaving an N of 36,508 values. The probability that the maximum of the 36,508 values exceeds any value X in the iRF range is $F(x)^N$:

$$F(X)^N = (1 - e^{-bX})^N = \exp[N \ln(1 - e^{-bX})] \approx \exp(-Ne^{-bX}). \quad (5.5)$$

This would not be valid if the sample members were dependent from each other, but as the flight points for the iRF calculations were randomly chosen, this is not the case.

The distribution function of the maximum is then given by

$$\Phi(X) = \exp(-Ne^{-bX}). \quad (5.6)$$

The probability that the maximum exceeds a value $X_{1/2}$ with 50 % can then be calculated with

$$1/2 = \exp(-Ne^{-bX_{1/2}}), \quad (5.7)$$

and solving this for $X_{1/2}$ leads to

$$X_{1/2} = \frac{\ln N - \ln \ln 2}{b}. \quad (5.8)$$

After inserting the values N and b for the positive iRF, this means, that with a probability of 50 % the true maximum exceeds 98.8 Wm^{-2} .

Now let α be the probability that the true maximum of iRF exceeds a value $X_{1-\alpha}$, then $\Phi(X_{1-\alpha}) = 1 - \alpha$. The equation (5.7) can then be generalized as

$$X_{1-\alpha} = \frac{\ln N - \ln \ln(1-\alpha)^{-1}}{b}. \quad (5.9)$$

When inserting an $\alpha = 0.05$, and N and b from the fit, then the iRF maximum exceeds 122.5 Wm^{-2} with a probability of 5 %. This does not mean that 5 % of the data exceed this value. It only says, if one would draw 36,500-member samples of an exponentially distributed population with $b = -0.11$ over and over again, that approximately 5 % would have a maximum value surpassing 122.5, but with an extremely low probability of the order of $0.05 \times \frac{1}{36508} = 10^{-6}$.

In the database, one point showed an $\text{RF}_{\text{LW}} > 100 \text{ Wm}^{-2}$, which is counterbalanced by a certain RF_{SW} . Now considering that this contrail can develop at night, where RF_{SW} is always zero, then the iRF would exceed the 100 Wm^{-2} . With 1.88, the optical thickness of this contrail is quite

large, but it is not the maximum ever observed. In the paper from (Atlas and Wang, 2010) even a case with $\tau = 2.3$ is recorded. Hence large values like 122.5 Wm^{-2} are not impossible. By predicting and preventing such strong contrails, a major share of the overall warming impact of contrails could be eliminated by changing only a few percent of the flight routes. This is why it is of utmost importance to improve the prediction of persistent contrails and specifically that of Big Hits. Therefore, the next chapter reveals some characteristics of Big Hits and an analysis of their meteorological conditions, with the goal to improve prediction possibilities.

5.2 Meteorological conditions of Big Hits

5.2.1 Characterization – Seasonal and spatial patterns, length, iRF

The following results were produced from the Big Hit dataset (chapter 4.2) which included not only every 100th flight point but every individual one that showed a persistent contrail with an instantaneous radiative forcing (iRF) larger than 19 Wm^{-2} . Again, a Big Hit consisted of a series of consecutive points along a flight that recorded an $\text{iRF} \geq 19 \text{ Wm}^{-2}$. When the iRF was below that limit, the point was not considered. Once a gap of more than ten data points where the iRF was lower than 19 Wm^{-2} was reached (distance of ten kilometers or 40 seconds), every following point that again had an $\text{iRF} \geq 19 \text{ Wm}^{-2}$ was counted towards a second or third, etc. Big Hit along a flight. 13,995 Big Hits were found in the 16,588 flights of the ten years of data. This does not mean, that on average \sim one Big Hit occurs on every flight, as many flights recorded multiple Big Hits. Figure 17, left, shows the number of Big Hits per flight. The maximum is at 24 Big Hits along one single flight, which is a surprisingly large number. Of course, most only show one or a few Big Hits. The number of Big Hits per flight declines nearly exponentially. Note, that the numbers largely depend on the definition of when a new Big Hit starts.

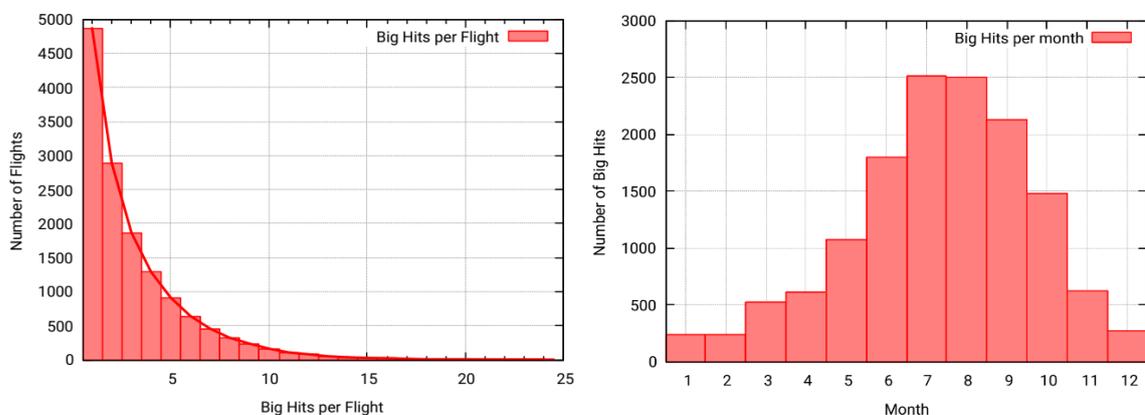


Figure 17: Left: Big Hits per flight, following a nearly exponential decline. **Right:** Big Hits per month.

The occurrence of Big Hits has a seasonal pattern, which can be seen in figure 17, right. Most Big Hits develop in warmer summer months, while in winter only a small percentage occurs. The seasonal percentages are given in table 6. Approximately 80 % of the Big Hits develop in summer and fall, while only ~ 5 % occur in winter. This seasonal pattern is even stronger than what was seen with the iRF dataset for persistent contrails. Again, this could be caused by the increased temperature gradient between the Earth's surface and the contrail in warmer months, which favors infrared (longwave) forcing, or by the seasonal variation of the tropopause height.

Table 6: Number of Big Hits per season and in total with their percentages.

Season	# of Big Hits	Percentage
Spring	2213	15.81%
Summer	6809	48.65%
Fall	4240	30.30%
Winter	733	5.24%
Total	13995	100.00%

All Big Hits were plotted on two world maps to look for spatial patterns or areas where Big Hits do not develop inside the research area (figure 18). The red lines are not the whole flight paths, but only the coordinates along a flight path where an $iRF \geq 19 \text{ Wm}^{-2}$ was recorded. It can be seen, that Big Hits don't necessarily persist for the whole flight, but often develop rather "patchy" on individual parts along the track. This is due to the variations in ice supersaturation in the atmosphere. When comparing figure 18, left, with figure 6, where all flights of MOZAIC are drawn, there are no obvious differences. Most Big Hits and flights from MOZAIC occur along the North Atlantic flight corridor, and central Europe. Significantly less of MOZAIC's flights fly over Russia and Asia, which is why there are not as many Big Hits to be found. However, it looks like the center and west of the US, as well as the south of Greenland shows very little and incoherent Big Hit coverage. The map in figure 6, unfortunately, shows all of MOZAIC's available flights since 1994 (and not only those from 2000 – 2009), such that a

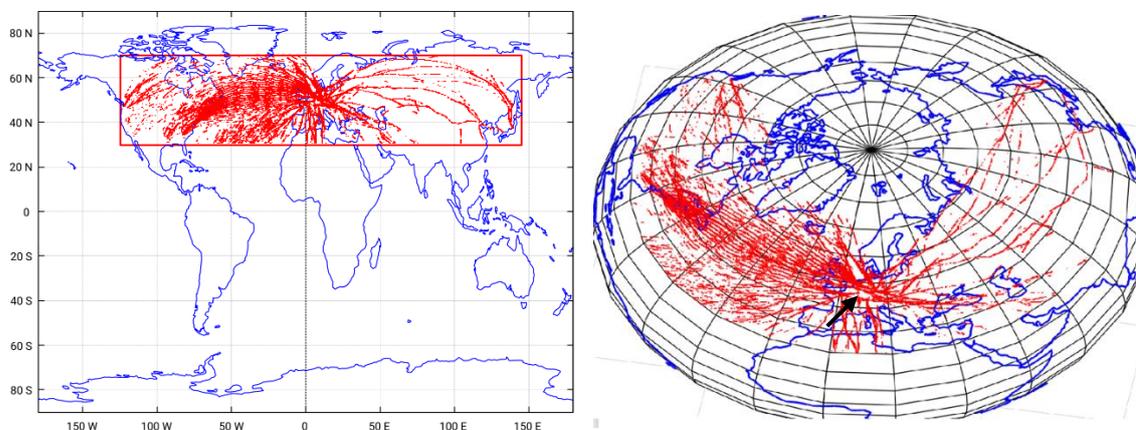


Figure 18: All Big Hits coordinates drawn in red (13,995) on two different world maps. The red box in the left plot demarks the research area where MOZAIC flights were analyzed. The black arrow points to a gap in Big Hit cover in an area of

direct comparison is not optimal. It would be necessary to draw all of the flight paths crossing those areas in the dataset and then compare them to all of the Big Hit occurrences therein.

When looking at the 3-D view in figure 18, right, a gap with no Big Hit cover in central Europe becomes visible (black arrow). This is unusual since that region has a high air traffic density. This is probably the area surrounding Frankfurt Airport (Germany), which is a major junction for passenger traffic. The numerous flights starting and arriving there do not immediately reach the altitude where contrail formation is possible. Hence, they can only produce Big Hits at a certain distance from the airport.

Seeing that Big Hits do not always form as a coherent field along the flight path, the question arises, whether contrail avoidance by flight diversion is then possible. Preventing Big Hits is only effective if they do not appear pointwise and have more than a few kilometers length.

To examine this, the length of each of the 13,955 Big Hits was calculated. 12,131 of those have a length less than 75 km, which translates to a flight duration of 5 minutes. This leaves only 1,864 Big Hits bigger than 75 km and only 660 with an extension above 150 km. The data showed one Big Hit exceeding 1,500 km in length. Although those occur very rarely, their climate effect is above average. The pdf's of the Big Hit lengths are depicted in figure 19. The

distributions of the Big Hits exceeding 75 and 150 km follow an exponential decline at the beginning. Of course, there is a lot of noise at higher values, as those appear so infrequent.

It is not only important to look at the length of the Big Hits, but the coherence of the iRF itself. If the iRF changes strongly in a matter of seconds, it would not be possible (and efficient) to adjust the flight level accordingly.

Figure 20 shows the iRF along different, randomly chosen flights of different lengths. The colored curves indicate that, from one flight point to another, the iRF jumps only weakly up and down with an amplitude of $\sim 5 \text{ Wm}^{-2}$ (blue, red, and green curve). In addition to those small changes, some of the curves (pink and orange) reveal strong fluctuations. Those could be caused by a change in the surroundings along the flight path, eg. by the presence of a natural cirrus or another contrail. A change in the surface characteristics, eg. when flying above a coast vs. above a dark ocean, changes the albedo and

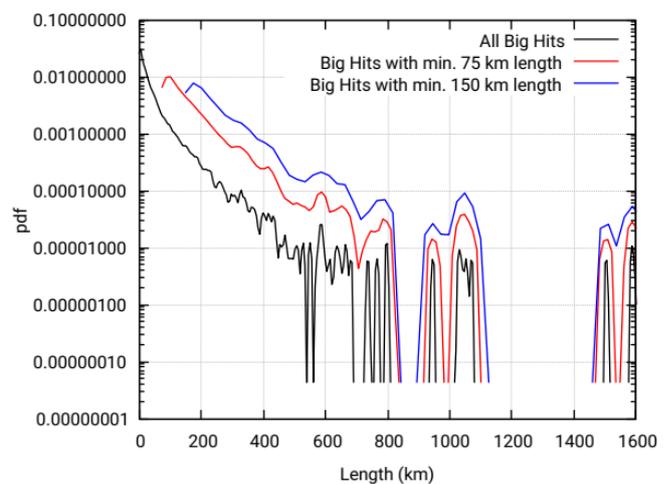


Figure 19: Pdf of the length of Big Hits. The blue curve includes all Big Hits, the red one only Big Hits with a minimum length of 75 km and the blue for Big Hits with minimum 150 km length.

influences the iRF as well. Passing the terminator line, which separates day and night on the globe, could also induce such variations of iRF. Lastly, changes in the optical thickness by variations of ambient ice supersaturation and temperature also control iRF. Noteworthy is, that the blue and orange curves directly start with a very high iRF value. In such

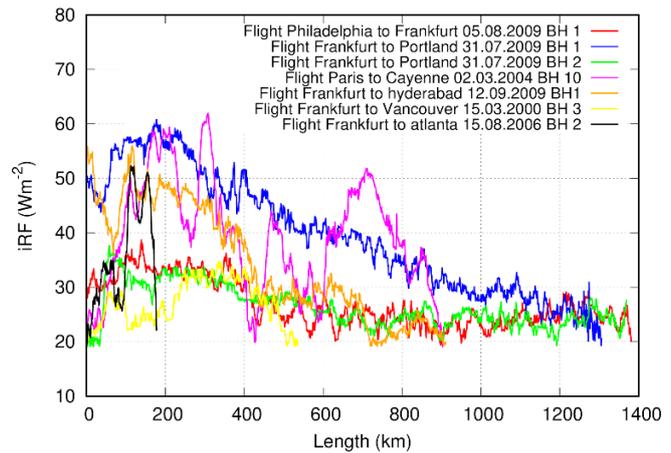


Figure 20: Change of iRF along different flight paths. Flights of differing lengths were chosen randomly.

cases, the contrail becomes a Big Hit by an immediate big jump. In total, the results are encouraging, as the iRF does not constantly vary from a very high to a low value from one flight point to the next. This is beneficial for flight diversion strategies.

The iRF values were further analyzed by producing a cumulative distribution function (cdf) and probability density function (pdf) from all individual points of all 13,955 Big Hits, amounting to 441,107 iRF values (figure 21). Most points ($\sim 80\%$) have an $iRF \leq 30 \text{ Wm}^{-2}$. As suspected from extreme value theory in chapter 5.1.3, values exceeding 100 Wm^{-2} do exist, with the maximum being 110 Wm^{-2} . The pdf of Big Hit iRF's follows an exponential decline and looks very similar to the pdf of the persistent iRF's (compare figure 13). An exponential function was again fitted to the pdf. The factor $b = 0.13 \pm 0.001$ (0.65%) is nearly the same as the exponential fit of the persistent iRF ($b = 0.11$).

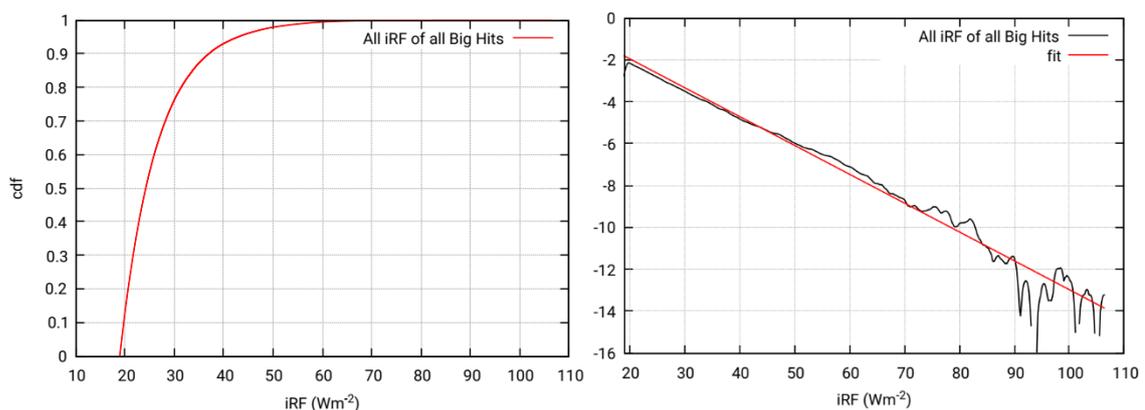


Figure 21: Left: Cdf of the iRF of all Big Hits (441,107 values). Most iRF have an $iRF < 40 \text{ Wm}^{-2}$. Right: Pdf of all iRF of all Big Hits (black) with its mathematical fit (red).

5.2.2 Comparison of conditions for situations with and without strong contrails

This subsection will compare distributions of thermodynamic and dynamical quantities for conditions where strongly warming contrails (Big Hits), persistent contrails and non-persistent

contrails (or no contrail) occur. Contrails were distinguished according to MOZAIC's data, such that RH_{max} and RH_i from the individual flights decided whether a persistent contrail was possible. The iRF value further decided if a persistent contrail was a Big Hit or not ($iRF \geq 19 \text{ Wm}^{-2}$).

If the distributions would be perfectly distinct, meaning well separated from each other, then Big Hits could be predicted reliably based on the ambient atmospheric conditions. Avoidance of Big Hits by flight planning would then be easy.

First, an overview of the distributions for the thermodynamic quantities T , absolute humidity, RH_{max} and RH_i is given. This is followed by the distributions for dynamical variables in section 5.2.2.2. Chapter 5.2.2.3 presents results of a logistic regression and of probability maps with the aim to predict Big Hits.

5.2.2.1 Thermodynamic quantities

Figure 22 compares the temperature distributions from MOZAIC and ERA – 5 reanalysis data for the three cases: Big Hits (blue), persistent contrails (black), and general situation (non-persistent or no contrail, red). The pdf's are produced from the iRF dataset and not the Big Hit dataset, as the latter does not include points that are not strongly warming. The distributions for the Big Hits of the iRF dataset (including only every 100th flight point, ~ 4,000 Big Hits) are nearly identical to the Big Hit distributions including all points (~ 400,000 Big Hits). This is valid for every following variable. Only minimal differences between the left and right panels of figure 22 are visible, indicating that ERA – 5 predicts temperatures for the flight points from MOZAIC well. Along the cruise level of MOZAIC's flights, temperatures between 200 and 250 K appear. Persistent contrails are found at the lower end of the distribution, lower even

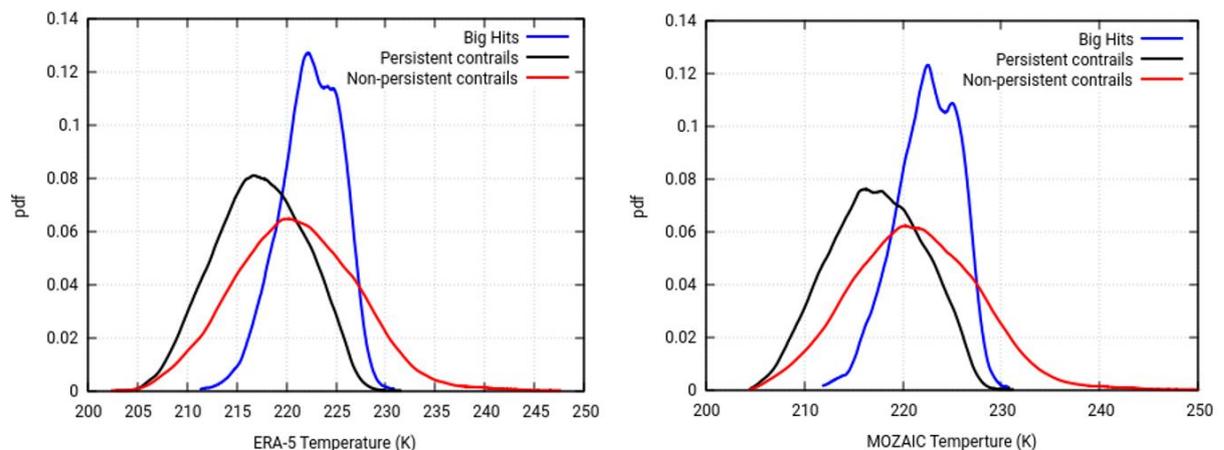


Figure 22: Temperature distribution in general situations (red), in situations that allow persistent contrails (black) and in situations with Big Hits (blue). The left panel shows the temperatures that ERA-5 reanalysis data predicts for the respective flight points from MOZAIC. The right panel shows MOZAIC's measured temperatures. Note that the deciding factor for persistence and Big Hits (for both plots) are MOZAIC's thermodynamic quantities.

than for the general case. This is caused by the temperature threshold for contrail formation. Big Hits occur predominantly at higher temperatures, since absolute humidity values are higher then, which can be used to form more ice and thus larger optical thickness. Figure 23, confirms that Big Hits occur particularly at the highest absolute humidities. The pdf for Big Hits (blue line) is shifted towards higher values, both for MOZAIC and ERA – 5. The mean for the Big Hits is 0.047 gm^{-3} , while for persistent contrails it is 0.025 gm^{-3} and for the general case, it is 0.012 gm^{-3} for MOZAIC.

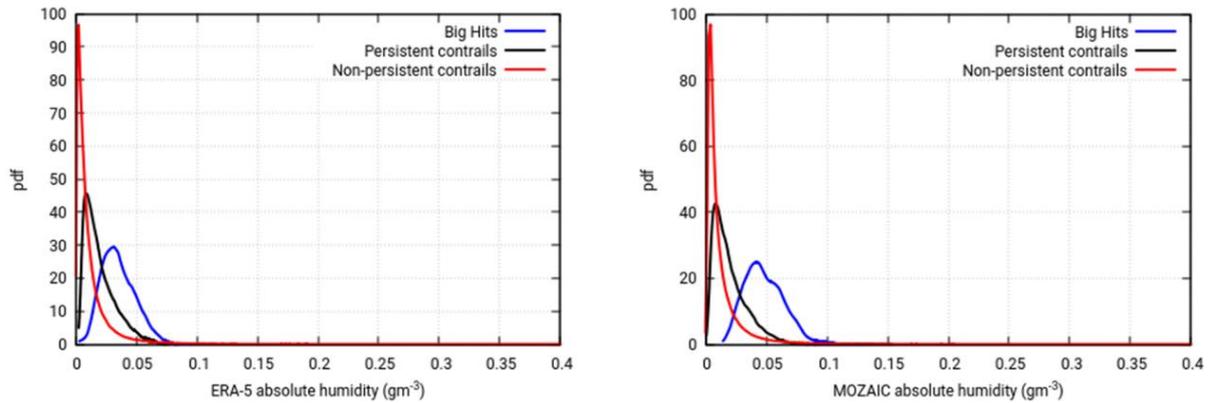


Figure 23: As figure 22 but for absolute humidity.

Not only most Big Hits occur at high temperatures, but the highest iRF values, so the Big Hits with the strongest effect, develop at the highest temperatures with the highest absolute humidities of the range. This can be seen in figures 24 and 25. All 441,107 Big Hit flight points are plotted with the iRF against the temperature (left) and the absolute humidity (right) as grey plus signs. $\text{iRF} \geq 70 \text{ Wm}^{-2}$ can only be seen at temperatures above 225 K and an absolute humidity of at least 0.06 gm^{-3} .

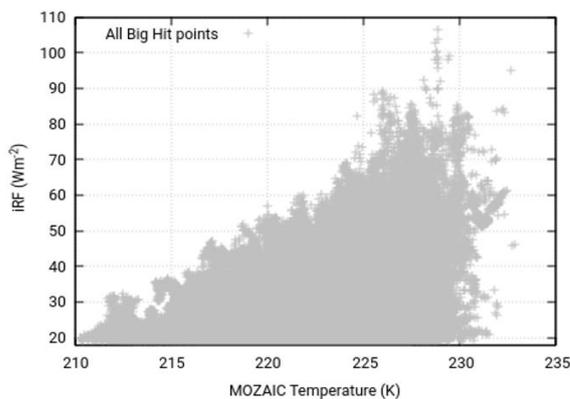


Figure 24: iRF of all 441,107 Big Hit flight points plotted against the temperature. Highest iRF values occur at the highest temperatures.

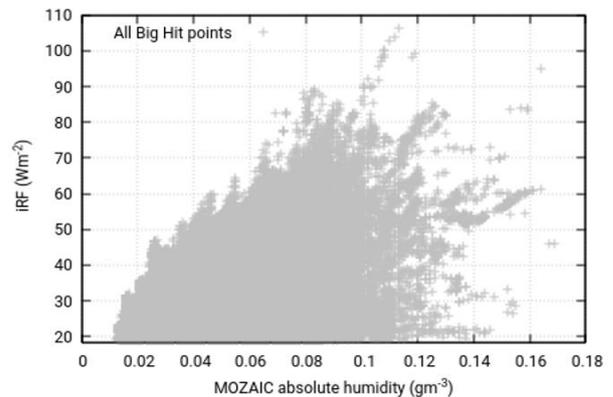


Figure 25: As figure 24, but for absolute humidity.

Figure 26 illustrates pdf distributions for RH_{max} . Since contrail formation is only possible when $RH_{max} \geq 1$, the pdf's for Big Hits and persistent contrails for MOZAIC start at that value.

ERA – 5 reanalysis can predict RH_{\max} well, which is not surprising, as it is a function of temperature. This, however, is not the case for relative humidity with respect to ice (see figure 27). Big Hits generally prefer high supersaturation values, as that leads to higher optical thicknesses and consequently higher iRF 's. Where MOZAIC sees a Big Hit with high RH_i values, ERA – 5 can predict zero ice supersaturation. The unreliable prediction of relative humidities in the upper troposphere and lower stratosphere is one of the reasons, why contrail prediction is so far only possible at a regional scale and not for flight routing. For this an accurate prediction of the local conditions (for each point) along a potential flight path would be necessary.

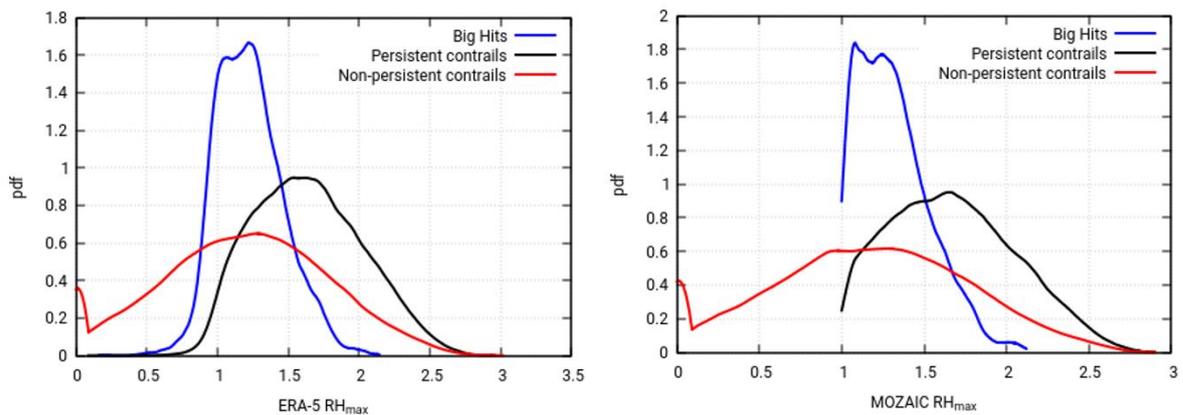


Figure 26: As figure 22, but for RH_{\max} .

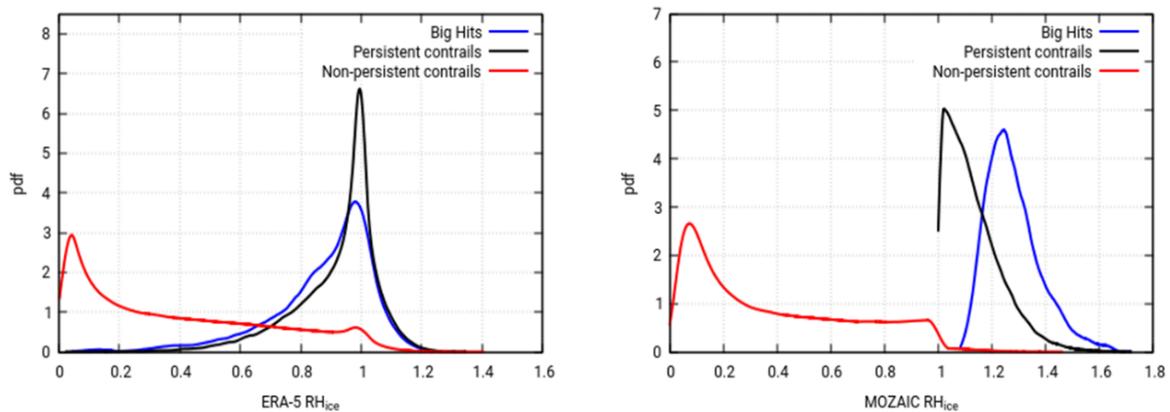


Figure 27: As figure 22, but for relative humidity with respect to ice.

The distribution for pressure is not relevant, since that does not show at which pressure altitude most contrails form, but rather which flight level is frequented most by MOZAIC's passenger aircraft.

5.2.2.2 Dynamical characteristics

This subsection will analyze distributions for vertical velocity, divergence, vorticity, potential vorticity, geopotential height, and the lapse rate. The dynamical quantities except the lapse rate

were determined by interpolation from ERA – 5 for each flight point from MOZAIC. The pdf's in figures 25, 26, and 27 are all given for Big Hits, persistent contrails and for the general case where no persistent contrail appears.

The pdf's for vertical velocity are centered around zero (figure 28 top left). For persistent contrails and Big Hits the distributions are slightly moved towards the left, such that the mode values are negative and the distributions have a stronger negative tail. In ERA – 5 the parameter is positive for a descending motion and negative for ascending motion. This negative shift can be explained by the formation mechanism of ice supersaturated regions (ISSR's), which are typically formed when air masses move upwards, where adiabatic cooling leads to an increase of relative humidity. To see the differences more clearly, the range of the plot for vertical velocity and divergence were restricted. Smaller and larger values are included in the dataset, but only with very low probability densities.

Divergence describes the horizontal divergence/convergence of velocity. It is the rate at which air is spreading out horizontally from a point. Positive values indicate a spreading / divergence and negative values a concentration/convergence (Hersbach et al., 2018a). Similar to the vertical velocity, the pdf's of the divergence (figure 28, top right) center around zero as well. Here the blue and black lines are slightly shifted towards positive values. This is probably caused by the outwards spreading of air masses below the tropopause which acts as a vertical boundary. Spichtinger et al. (2003) found that most ISSR's occur in a 200 hPa thick layer below the tropopause. The large RH_i values there promote the development of persistent contrails with high iRF 's, when an aircraft passes through that layer. This is connected to the potential vorticity (figure 28, bottom right). The extratropical dynamical tropopause is typically found at the 2 PVU potential vorticity surface (Holton et al., 1995), where $1 \text{ PVU} = 10^{-6} \text{ K m}^2 \text{ kg}^{-1} \text{ s}^{-1}$, which is why the pdf's for Big Hits and persistent contrails have their peaks below that level. Contrails also develop in the lower stratosphere, but not as often. The low probability densities at higher PV values of the blue and black function confirm that. Potential vorticity strongly increases in the stratosphere (Petzold et al., 2020). This explains the difference in the pdf distribution for the general case.

Figure 28, bottom left, shows large differences between the pdf for Big Hits and the other two functions. The vorticities for Big Hits and persistent contrails are narrower and strongly negative compared to the general case, where the mode value of the distribution is clearly positive. Relative vorticity is a measure for the rotation of air in the horizontal. It is anticlockwise at positive values and clockwise at negative values (in the northern hemisphere).

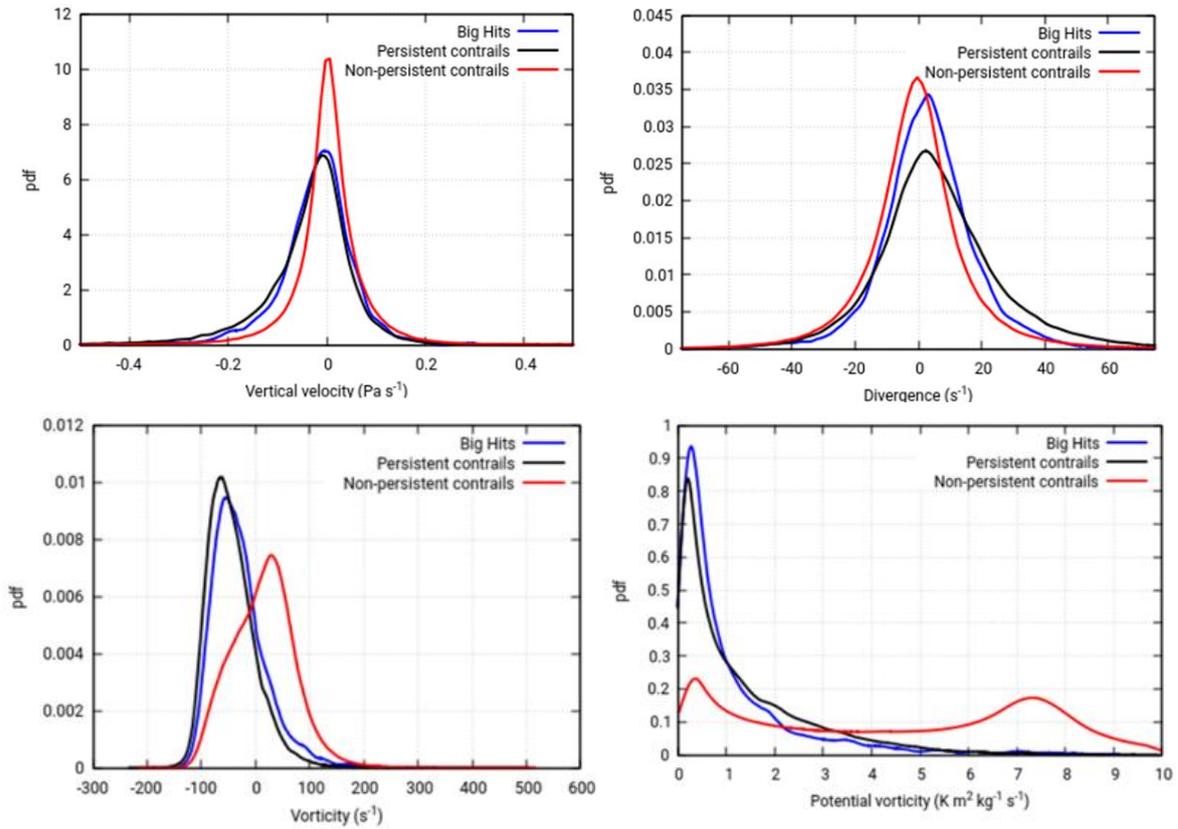


Figure 28: Pdf's of dynamical characteristics for the three situations: Big Hit (red), persistent contrail (black), and not persistent (red). Upper left: Vertical velocity. Upper right: Divergence. Lower left: Vorticity. Lower right: Potential vorticity. The pdf's show large differences only for the vorticity and partly for potential vorticity.

Hence, Big Hits favor anticlockwise atmospheric motion. This is no surprise since ISSR's also prefer divergent, anticyclonic flow (Gierens and Brinkop, 2012).

Except for the vorticity, all distributions overlap more or less strongly. This demonstrates, that persistent contrails and Big Hits do not only develop during the favored flow patterns. Only the relative frequency of the patterns is different within and outside of persistent contrails and Big Hits. Therefore, in the sense of regression analysis, the dynamical fields vertical velocity, divergence, vorticity, and potential vorticity alone are not sufficient for a prediction of persistent contrails and specifically Big Hits. Therefore, three more variables are considered.

The pdf's for the geopotential height (Z) in figure 29, right, show the largest separation so far. Persistent contrails, and even more so for Big Hits, center at higher Z values ($0.97 - 1.05 \text{ m}^2 \text{ s}^{-1}$). For a better comparison of the geopotential heights on different flight levels, the geopotential height was normalized, such that it measures the relative distance from a nominal height of each pressure level. $Z = 1$ means that the pressure level has its nominal geopotential height, $Z < 1$ implies an excursion to lower altitudes and vice versa. Persistent contrails and Big Hits are on average found on the top altitudes of the pressure levels.

In figure 29, left, one can see that Big Hits occur especially at night, where $\mu = \cos(\text{SZA}) = 0$, or when the sun is very low ($\mu \leq 0.13$) or high ($0.6 \leq \mu \leq 0.85$). Most and also the strongest Big Hits with the highest iRF's (not shown) develop at night since the counteracting RF_{SW} effect is absent at that time. When the sun is low, at large solar zenith angles (SZA), the solar direct radiation approaches zero. Since Schumann's parameterization (Schumann et al., 2012) does not account for diffuse radiation, IRF values at low values of μ could be overestimated to an uncertain degree. According to Meerkötter et al. (1999), the strongest SW forcing occurs at zenith angles of $60 - 70^\circ$. This is consistent with the results from this study, as Big Hits occur less frequently at intermediate μ values. Unfortunately, even if differences between the pdf for Big Hits and the pdf's for persistent and non-persistent contrails are clearly visible, they completely overlap. Identifying Big Hits from μ alone is therefore not possible.

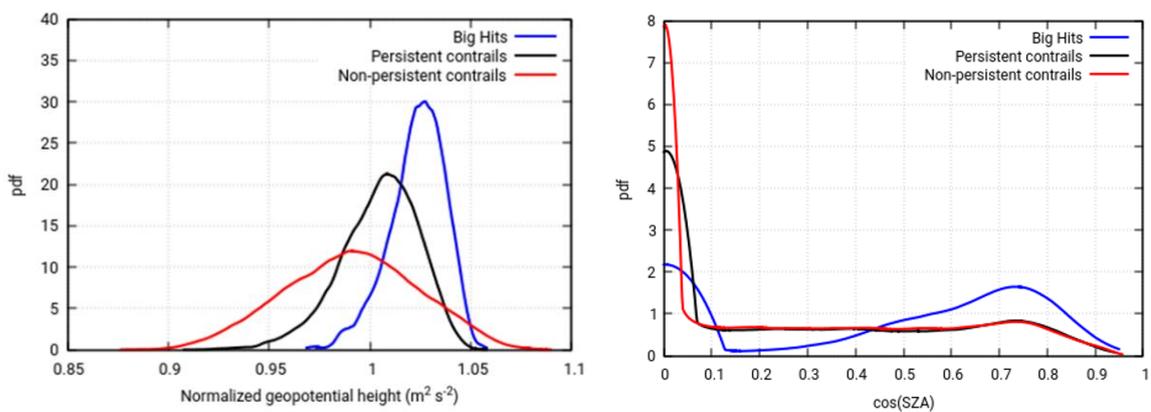


Figure 29: Left: Pdf for the normalized geopotential height for the three cases Big Hit (blue), persistent contrail (black) and not persistent (red). Big Hits occur predominantly at high geopotential heights. Right: Same as left plot, but for μ , the cosine of the solar zenith angle. Most Big Hits develop at night with a $\mu = 0$, or with very low or very high μ values.

A promising variable is the lapse rate. It describes the vertical temperature gradient and is a measure for the stability of the atmosphere (stratification). The lapse rate was calculated using the barometric height formula (equations 4.21 – 4.25). The global average tropospheric lapse

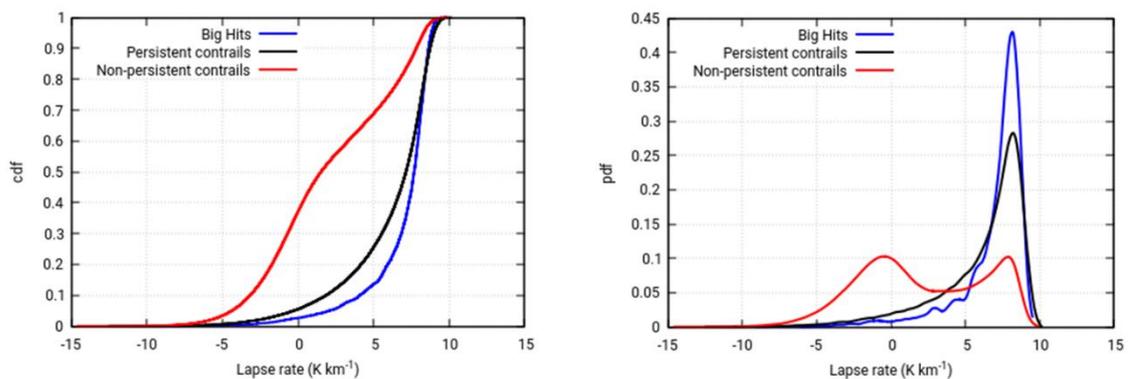


Figure 30: Cdf and pdf of the lapse rate of Big Hits, persistent contrails, and for the general case. Persistent contrails and Big Hits develop particularly at high lapse rates.

rate is 6.5 K km^{-1} (International Civil Aviation Organization, 1993). For the cdf's and pdf's of figure 30 the lapse rates range from ~ -15 to $\sim 10 \text{ K km}^{-1}$. With an environmental lapse rate of 10 K km^{-1} the stratification is neutral, as it is similar to the dry adiabatic lapse rate of 9.8 K km^{-1} . A lapse rate below that implies stable stratification. Negative lapse rates occur with temperature inversions, which happens when the tropopause is close to the flight level. The data showed that Big Hits and persistent contrails develop under rather high lapse rates. The difference between the three scenarios (Big Hit, persistence, non-persistence) is best visible when looking at the cdf's. While $\sim 70\%$ of the general situation points have a lapse rate below 5 K km^{-1} , that increases to $\sim 8 \text{ K km}^{-1}$ for Big Hits and persistent contrails. It is not yet understood, why the stratification is much weaker at locations with persistent contrails and Big Hits.

Using dynamical variables for the prediction of Big Hits has not been done before. Combining the quantities that showed the largest separation in a regression analysis or a neural network could allow an improved prediction for contrail persistence and Big Hits. Specifically, the pdf's for Big Hits (except for temperature and vorticity) are slightly more separated to the pdf for the general case (no persistent contrails) than the pdf of persistent contrails. To test this theory, a logistic regression analysis was performed with the dynamical variables for Big Hits.

5.2.3 Prediction possibilities

For a first test, the iRF dataset of persistent contrails was used (47,032 values) for the logistic regression, which included 4,052 Big Hits.

To get an idea of which variables influence iRF the most and which predictor variables should be included in the logistic model, a correlation matrix for the whole iRF dataset was produced (figure 31). The matrix includes all variables from ERA – 5 reanalysis. For the logistic regression, only those can be considered, as for a prediction of persistent contrails and Big Hits only quantities from numerical weather prediction can be used. Correlations are depicted as colored circles of different sizes. The largest circles indicate highest positive (blue color) and negative correlations (red color). The last row of the matrix shows correlations of iRF with all other variables. The highest correlations are with temperature, absolute humidity, RH_{max} , RH_i , pressure, and geopotential height. All the other dynamical variables have weak correlations at most.

For the multiple logistic regression, the iRF dataset was divided into a training and a control dataset. The model was produced from the training data, as explained in chapter 4.2, including all mentioned meteorological and thermodynamic variables. To avoid overfitting of the model,

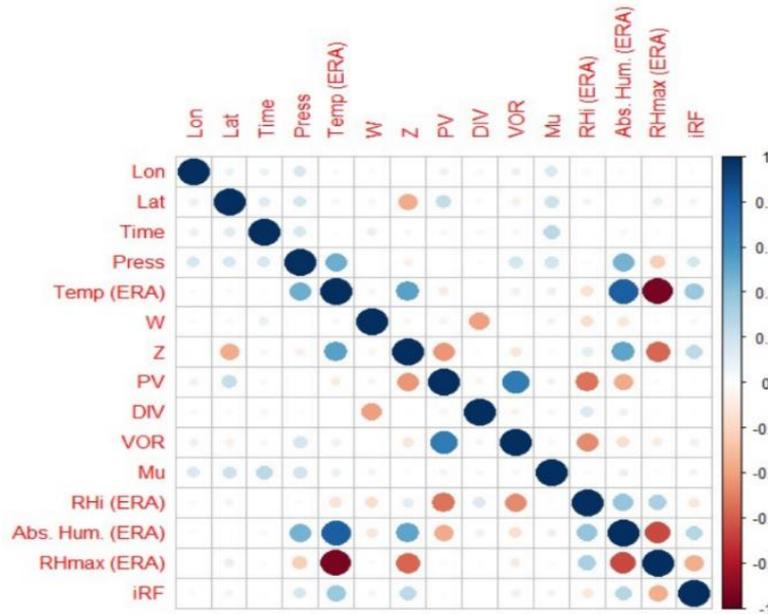


Figure 31: Correlation matrix of all variables in the iRF dataset (including 47,032 persistent or strongly warming contrails). Specifically, the correlation of variables with iRF is of interest (last row in each matrix). The correlations are illustrated as circles, with increasing size if a high positive (blue color) or negative correlation (red correlation) exists. The matrix includes all variables from ERA – 5.

an improved model only included the statistically significant predictors with a p-value of the coefficients ≤ 0.05 . Those were the lapse rate, pressure, geopotential height, vorticity, RH_{\max} , RH_i , and μ . This coincides with the results from the pdf's from before, as specifically the distributions for the lapse rate, geopotential height, and vorticity of Big Hits had the best separation. It was expected that RH_{\max} and RH_i , which define contrail formation and persistence, were important predictors for iRF. The model was tested on the control dataset, where it produced probabilities for the occurrence of Big Hits based on the predictors. When the probability exceeded 50 %, then the model decided a Big Hit occurrence. The model was able to predict Big Hits with an accuracy of ~ 88 % and a test error rate of ~ 12 %. However, due to the very little apriori probability of Big Hits, this high accuracy stems rather from correctly predicting that a Big Hit would not develop. The true positive rate, which describes the amount of Big Hits correctly predicted, can be improved when the probability that needs to be exceeded for a Big Hit to be predicted is not 50 % but rather of the size of the apriori probability. The apriori probability for persistent contrails is $P(pC) = 11.5$ % and for Big Hits it is $P(BH) = 1$ % (where $iRF \geq 19 \text{ Wm}^{-2}$), for the iRF dataset including Big Hits, persistent contrails and general points with no persistent contrails for every 100th flight point. Unfortunately, the model was not able to predict contrail persistence. This means, that in theory Big Hits can be predicted reliably, but only when contrail persistence can be predicted.

A last approach was taken to predict contrail persistence from model output statistics. Probabilities for contrail persistence were computed, conditioned on dynamical and thermodynamic variables X that needed to have a value in the range of $x \pm \Delta x$. The probability for a persistent contrail increases by limiting a range of $x \pm \Delta x$ where specifically more persistent contrails than non-persistent ones occur. For example, allowing only cases with temperatures below 230 K includes all persistent contrails, but excludes a large part of the general-case values, as they occur in temperatures between ~ 200 K – 250 K. As mentioned in chapter 4.2, for this application probability is equated with relative frequencies, such that the probability for persistent contrails was given by

$$P(pC|X = x \pm \Delta x) = \frac{N(X = x \pm \Delta x \wedge iRF > -999)}{N(X = x \pm \Delta x)}. \quad (5.10)$$

Not every mentioned meteorological quantity was suitable for this. As seen in the pdf plots, the pdf distributions of persistent contrails and Big Hits of different variables overlap strongly with that of the non-persistent pdf's (eg. vertical velocity and divergence). Here introducing a limiting condition does not improve the probability for a persistent contrail, as for all of the values in the range of vertical velocity and divergence persistent contrails as well as non-persistent cases occur. It is not possible to exclude certain values where specifically only non-persistent cases occur. Suitable meteorological quantities are temperature, $RHmax$, RHi , vorticity, lapse rate, and geopotential height. Those all show a certain separation between the pdf's for persistent and non-persistent cases (see figure 22 left, 26 left, 27 left, 28 bottom left, 29 left, and 30). Note that for this approach one has to look at the variables from ERA – 5, as the goal is to predict contrail persistence from numerical weather prediction. MOZAIIC's variables only decided on persistence vs. non-persistence, but they are not further included in the probability calculations. Hence for RHi and $RHmax$ also values below 100 % or 1 can be found for persistent contrails, since the prediction of those values is not perfect in ERA – 5. All meteorological variables X span an n -dimensional space. This space can be divided into “boxes” of a specific value range $x \pm \Delta x$. In each box, there is a certain number of Big Hits / persistent contrails and of the general case. Dividing those numbers leads to the probability for persistent contrails for each box. Beginning with two dimensions, figure 32 shows persistent contrails (black dots) and general cases (red dots) on a lapse rate (LR) – geopotential height (Z) plot. LR and Z were chosen first because their distribution showed the least amount of overlap between the pdf for persistence / Big Hits and non-persistence. The dimensions lapse rate and geopotential height were each divided by percentiles in 0.1 (10 %) steps (dotted lines). The

percentiles are given for the distribution of persistent contrails. Percentiles for all meteorological variables are given in tables 7 and 8 for the persistent and Big Hit distributions. Every point that is persistent / not persistent, with a specific iRF and specific meteorological conditions, can then be assigned to one of the 100 boxes drawn in the LR – Z space. It is well visible, that the apriori probability for persistent contrails is a lot lower than

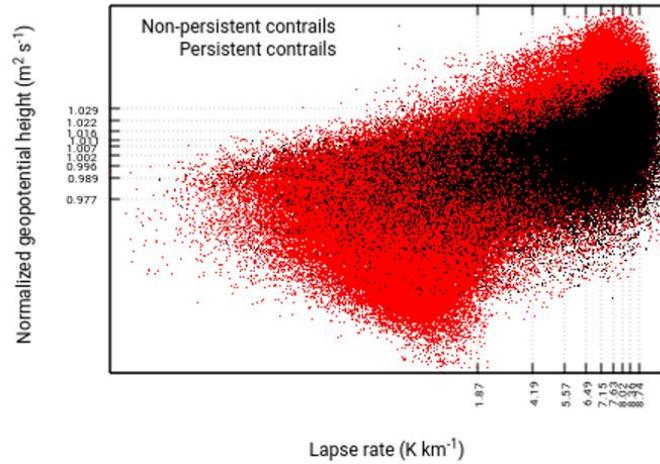


Figure 32: Persistent contrails (black dots) and non-persistent contrails (red dots) in a Z – LR plot. The lines depict percentiles of each variable in 0.1 steps (10%). Those lines define 100 boxes of different sizes, where the probability for a persistent contrail can be counted for each box.

for the general cases, as the red dots dominate in the plot. Most persistent contrails occur at high LR and high Z values, while non-persistent contrails are scattered throughout the whole space. For each box, the probability for persistent contrails can now be given in a 2-D probability matrix (see table 10). Further dimensions could be added, but it is important, that enough values fall inside of each box. Adding a third dimension would thin out the number of cases inside a box, as all of the cases would then be divided into 1,000 boxes. If only one value falls into a box, determining a probability does not make sense. Unfortunately, this was the case for the dataset, such that further conditions could only be added for the whole matrix and not as a third dimension. This was also the reason, why probabilities were given for persistent contrails and not for Big Hits. The apriori probability of Big Hits is too low to divide them into 100 boxes since for the boxes of the highest values, only a few points are left. In principle, this would be possible if one had an even bigger dataset. This is currently work in progress.

Table 7: Percentiles for different meteorological variables for the distribution of persistent contrails.

Persistent contrails												
Percentiles	Lapse Rate	Z	VOR	PV	Mu	W	DIV	Press.	Temp.	Abs. Hum.	RH _{max}	RH _i
Min	-12.5	0.908	-233.7	0	0	-1.788	-120.14	190.1	202.89	0.001	0.08	0.03
0.1	1.87	0.977	-93.3	0.11	0	-0.143	-15.73	215.9	211.3	0.008	1.15	0.72
0.2	4.19	0.989	-80.6	0.21	0	-0.089	-7.78	221.7	213.3	0.011	1.28	0.82
0.3	5.57	0.996	-70.1	0.35	0	-0.059	-3	227.1	214.9	0.013	1.40	0.89
0.4	6.49	1.002	-60.3	0.52	0.016	-0.038	0.96	237.7	216.2	0.015	1.51	0.93
0.5	7.15	1.007	-50.3	0.77	0.017	-0.021	4.67	238.5	217.4	0.018	1.62	0.96
0.6	7.63	1.011	-39.1	1.11	0.317	-0.006	8.73	249.1	218.7	0.021	1.73	0.98
0.7	8.02	1.016	-26.8	1.56	0.463	0.008	13.57	250.4	220.0	0.025	1.84	0.99
0.8	8.36	1.022	-11.6	2.21	0.612	0.026	19.97	261.6	221.5	0.03	1.98	1.02
0.9	8.74	1.029	13.5	3.23	0.739	0.055	31.35	273.5	223.4	0.038	2.17	1.05
Max	9.84	1.058	272.9	9.92	0.956	0.669	176.37	309.9	231.5	0.098	2.96	1.34

Table 8: As table 7, but for the distribution of Big Hits.

Big Hits												
Percentiles	Lapse Rate	Z	VOR	PV	Mu	W	DIV	Press.	Temp.	Abs. Hum.	RH _{max}	RH _i
Min	-9.54	0.969	-160.32	0.00	0	-0.607	-83.31	196.4	211.4	0.003	0.16	0.06
0.1	4.15	1.004	-84.94	0.13	0	-0.110	-13.38	226.9	218.4	0.018	0.95	0.65
0.2	5.97	1.012	-70.79	0.22	0	-0.068	-7.02	238.1	220.0	0.023	1.02	0.77
0.3	6.85	1.017	-60.06	0.31	0	-0.046	-3.28	238.9	221.0	0.026	1.08	0.83
0.4	7.35	1.021	-49.50	0.42	0.250	-0.028	0.14	249.8	221.8	0.030	1.15	0.88
0.5	7.68	1.024	-38.46	0.57	0.490	-0.014	2.93	260.7	222.6	0.033	1.21	0.93
0.6	7.93	1.028	-27.05	0.78	0.588	0.000	5.83	261.6	223.4	0.037	1.27	0.96
0.7	8.16	1.031	-14.30	1.10	0.683	0.014	9.26	262.1	224.3	0.041	1.33	0.98
0.8	8.39	1.035	3.60	1.65	0.736	0.031	13.45	274.0	225.2	0.047	1.42	1.01
0.9	8.64	1.040	34.62	2.70	0.792	0.061	20.47	283.5	226.2	0.054	1.55	1.04
Max	9.51	1.057	334.12	9.11	0.950	0.363	80.77	309.9	230.8	0.098	2.14	1.20

One at a time, new conditions of meteorological variables were added to the two-dimensional LR – Z matrix. For example, a condition for a maximum potential vorticity (PV) of $3.23 \text{ K m}^2 \text{ kg}^{-1} \text{ s}^{-1}$ was introduced, so that only points that fulfill this are counted for each LR – Z box. In particular, at each step, the condition for one additional meteorological variable was added, such that the greatest additional improvement in the probability of persistent contrails per box was achieved. The values were not chosen randomly, but they are chosen percentiles that were determined from looking at the pdf's from chapters 5.2.1 and 5.2.2. It is the value, at which below or above that predominantly Big Hits and persistent contrails occur.

Comparing 100 probabilities with each other is not effective, so the probabilities were summed up per row (for each Z percentile the sum of the probability for persistent contrails of the ten LR percentiles is given) and in total. The improvement in probability for a persistent contrail with the introduction of further conditions is listed in table 9.

Table 9: For each Z percentile, the sum of the probability for persistent contrails of the LR percentiles and the sum of all of the probabilities of the two-dimensional probability matrix is given. Introducing additional conditions improves the total sum of probabilities. Note that the minimum count of persistent contrails per box reduces when adding further conditions.

Conditioned probabilities							
Percentiles	Z + LR	PV < 3.23	VOR < 34.62	T < 226.2	T > 211.3	RH _{max} > 0.95	RH _i > 0.65
Sum min - 10th	0.46	1.99	1.99	2.05	1.99	2.85	4.56
Sum 10th - 20th	0.89	1.48	1.49	1.56	1.53	2.06	3.88
Sum 20th - 30th	1.49	1.68	1.8	1.87	1.86	2.37	3.81
Sum 30th - 40th	2.06	2.08	2.19	2.28	2.34	2.81	3.98
Sum 40th - 50th	2.58	2.57	2.65	2.73	2.83	3.29	4.21
Sum 50th - 60th	3.11	3.1	3.15	3.26	3.41	3.89	4.61
Sum 60th - 70th	3.57	3.57	3.6	3.69	3.95	4.47	5.09
Sum 70th - 80th	4.2	4.19	4.21	4.36	4.61	5.15	5.6
Sum 80th - 90th	4.67	4.67	4.69	4.82	5.09	5.75	6.09
Sum 90th - max	5.36	5.36	5.35	5.48	5.95	6.75	6.9
Total sum	28.39	30.69	31.12	32.1	33.56	39.39	48.73
Minimum count	8	6	5	5	5	5	5

The probabilities per Z percentile of the LR – Z matrix without additional conditions are given in the first column. Adding conditions for the potential vorticity, vorticity, temperature,

RH_{max}, and RH_i nearly doubles the summed up matrix probability. However, the minimum count of persistent contrails per box limits the number of conditions that can be introduced. With a minimum count of 5 persistent contrails per box, the maximum probabilities for contrail probabilities that were achieved are depicted in table 10.

The highest probability for persistent contrails is 93 %, which occurs at the upper ten percent of the lapse rate and geopotential height distributions with the additional conditions.

The matrix in table 10 is depicted as a heatmap in the following figure 33. Probabilities were linearly interpolated using Gnuplot, where the optimal number of interpolation surface points was chosen automatically (Williams and Kelley, 2019). Probabilities of $\sim 50\%$ are exceeded starting from lapse rates of $\sim 4 - 5 \text{ K km}^{-1}$ upwards. This is quite an improvement from the apriori probabilities of persistent contrails of 11.5 %. The probabilities increase with increasing geopotential height. With all the included conditions for PV, vorticity, T, RH_{max}, and RH_i the lowest probabilities are $\sim 29\%$.

Table 9: Maximum probabilities for the LR - Z matrix with the additional conditions from table 9. The highest probabilities can be seen at high lapse rates and high geopotential heights.

		Lapse rate									
		Max - 90th	90th - 80th	80th - 70th	70th - 60th	60th - 50th	50th - 40th	40th - 30th	30th - 20th	20th - 10th	10th - Min
Z	Max - 90th	0.93	0.78	0.77	0.79	0.68	0.69	0.64	0.63	0.55	0.44
	90th - 80th	0.71	0.69	0.77	0.64	0.65	0.64	0.55	0.56	0.48	0.4
	80th - 70th	0.7	0.66	0.66	0.6	0.59	0.55	0.53	0.5	0.45	0.36
	70th - 60th	0.67	0.58	0.55	0.54	0.55	0.5	0.45	0.49	0.42	0.34
	60th - 50th	0.61	0.58	0.49	0.51	0.47	0.44	0.41	0.4	0.39	0.31
	50th - 40th	0.53	0.52	0.47	0.44	0.44	0.41	0.39	0.35	0.36	0.3
	40th - 30th	0.58	0.48	0.43	0.42	0.36	0.36	0.38	0.36	0.33	0.28
	30th - 20th	0.43	0.43	0.39	0.38	0.37	0.38	0.41	0.4	0.33	0.29
	20th - 10th	0.39	0.4	0.4	0.36	0.4	0.43	0.42	0.36	0.35	0.37
	10th - Min	0.44	0.45	0.46	0.46	0.48	0.5	0.49	0.55	0.4	0.33

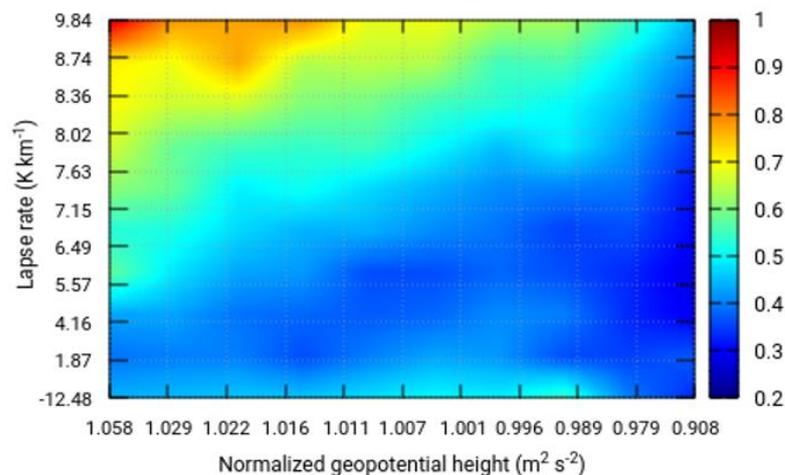


Figure 33: Heatmap of conditional probabilities for contrail persistence in a two-dimensional space spanned from Z and LR percentiles (numbers on the axes). Highest probabilities occur with increasing geopotential height and increasing lapse rate.

The produced matrix and the heatmap are only an example to get a feeling for the meteorological conditions that persistent contrails prefer. Using different values for the conditions would influence the probabilities. Different dimensions could be used as the base and by adding another dimension the uncertainty could be reduced when enough cases are available. Generally speaking, including additional dynamical conditions could improve prediction possibilities for persistent contrails. If similar thresholds could be included in weather prediction models like ERA – 5, then a prediction of persistent contrails on a local level might be possible.

6 DISCUSSION

6.1 Statistics of iRF

This section will discuss the results from the statistics of instantaneous radiative forcing (iRF). The goal was to determine the weather influence on the instantaneous radiative forcing variability. This variability can be clearly seen in the plots of the cdf's and pdf's, as well as in the table with the different mean values from chapter 5.1.1. The random character of the weather led to a wide distribution of the pdf of iRF, which follows an exponential decline in the positive branch. The heavy tail of the pdf (figure 13), as well as the pdf x iRF plot (figure 16 in chapter 5.1.3) are proof of a disproportionate impact of strongly warming contrails (Big Hits) on the overall radiative effect, even though they have very low probabilities and occur very rarely. Monte-Carlo experiments showed how strongly a mean value from a sample of measurements scatters depending on the sample and its size. The data also revealed a distinct seasonal pattern in the occurrence and magnitude of iRF values (figure 11 and table 5). It is important to note, however, that the frequencies are influenced by the number of available flights from MOZAIC and should be not taken too literally. This is also valid for all other results so far. All the numbers and values are just approximations. For one, they are certainly influenced by the parameterizations and assumptions used for the iRF calculations. As mentioned before, the model made sure to correctly reflect the relation of individual contrails and their radiative effect with each other, but using different assumptions would give different numbers. For example, using Myhre particles is the simplest way to parameterize the crystal size and habit distribution, as they have wavelength-independent, so constant, optical properties (Myhre et al., 2009). Those hexagonal “pseudo” particles also do not include solar absorption, but for contrails that is nearly negligible, as “the path lengths of photons through atmospheric ice crystals are very small compared to the absorption length” (Warren et al., 2006). In reality, contrails and contrail cirrus consist of a mixture of different crystal habits with varying effective radii. Usually, contrails consist of ice crystals like hexagonal plates, columns, and triangular plates of 30 – 60 μm size (on average) (Duda et al., 2001; Yang et al., 2010). Those habits have varying microphysical and radiative properties, which affects contrail radiative forcing (by varying single-scattering characteristics) (Duda et al., 2001; Letu et al., 2019; Xiao et al., 2019). The ratio of warming to cooling contrails could therefore be different when using different crystal habits, as they have varying forward to backward scattering ratios.

All of this was neglected in the study to obtain an iRF measure with the least amount of degrees of freedom. Including this would probably lead to a lot more variability and hence uncertainty,

especially since the crystal habit evolves following the development of the ice cloud and since the natural variation of ice crystal habits is still not completely known (Xiao et al., 2019).

The specific iRF values in the dataset are also influenced by the chosen MOZAIC data sample. Different years would lead to different numbers. However, the numbers are not what is important. The goal was not to exactly define eg. a mean iRF over 10 years but to find the general characteristics and type of the distribution. Changing the assumptions would not influence the essence of the results: That is, the distribution is wide, it is an exponential at positive values, the highest radiative forcing values have the lowest probability but a very large effect, and that the natural variability is huge. The observed variability is also just the lower limit of natural variability, as only a part of the weather variability that directly defines the instantaneous radiative effect was included. The solar zenith angle, as well as the height of the tropopause that might be responsible for the seasonal variations, are part of the weather variability of the dataset, but one might argue, that those are rather a climate variability and not weather variability. Specifically, the solar zenith angle does not induce a random uncertainty like RH_i or T do. It is a deterministic variable and is known depending on the time. To the author, however, it is not important how exactly this variability is called, but rather that it is huge and irreducible. Considering that the interpolated radiative quantities used for the iRF calculations (OLR, SRD, and RSR) are all given as an hourly mean, the variability could in reality be even larger if not averaged. There might also be more variability if MOZAIC flights of the whole globe and not only of the confined research area (30° to 70° latitude 125° to 145° longitude) were included.

Another important point is, that there are many more sources of natural variability that determine the overall radiative impact of contrails, which were not included in the study.

These are: i) technical aircraft variations, ii) turbulence and atmospheric dynamics iii) the ambient air traffic density, and iv) natural cirrus.

i) The size (weight and wingspan), number of engines (two vs. four engines), fuel flow, and speed of an aircraft determines the vertical extension and ice crystal number of the produced contrail, but also its evolution (Gayet et al., 2012; Gierens and Ström, 1998; Jeßberger et al., 2013; Sussmann and Gierens, 2001; Unterstrasser and Görsch, 2014). The number of generated ice crystals changes with the amount of water vapor and particles emitted, which depends on the fuel flow (Unterstrasser and Görsch, 2014). Sussman and Gierens (2001) found that contrails formed from large aircraft survive longer in subsaturated conditions than smaller aircraft, because of the larger water vapor emission. In general, those also have larger optical thicknesses (Jeßberger et al., 2013). The wake vortex system with its descent and lifetime is

controlled by the aircraft mass and geometry. Larger aircraft produce contrails with more ice crystals, but the ice crystal loss is also stronger during the vortex phase, as the vertical displacement of the wake vortices is stronger. The crystals descent faster and further down, where they undergo more intense adiabatic heating and consequently sublimate (Unterstrasser and Görsch, 2014).

ii) The dispersion, evolution, and lifetime of a contrail are also determined by atmospheric dynamics like wind shear, advection, and turbulent mixing, as well as on the stratification (Dürbeck and Gerz, 1996; Unterstrasser and Gierens, 2010). In general, the disintegration of a contrail from particle loss, be it by dynamical or microphysical cause (Bier et al., 2017), also introduces a lot of variability.

iii) When looking at the effect of a single contrail in a certain area, one must consider, that other contrails could already be there, especially in regions of high air traffic. If this is the case, all contrails compete for the available condensable water vapor, which inhibits their growth (Unterstrasser, 2020). Contrail coverage might also not be increased if the sky is already covered by a contrail cirrus sheet (Gierens, 1998).

iv) Natural cirrus in vicinity to a new contrail have similar effects (Lynch et al., 2002).

All of these factors are not included in the present study, but they influence the overall radiative effect of a contrail and therefore introduce natural variability. For a single contrail (not only for a point), EF is the correct measure of the radiative effect. One can argue that EF would vary even stronger than iRF, since more degrees of freedom are involved (evolution of the contrail during its lifetime for the covered area (width x length)). The variability of EF could be computed with CoCiP (Schumann et al., 2012). Finally, summing up all EF values for all contrails in the atmosphere over a year would lead to the desired RF or ERF metrics and here again more variability can be expected, e.g. by varying traffic patterns over the years, slow evolution of the aircraft fleets etc. Hence, from a pointwise, to a local, to a global measure of the radiative effect of contrail cirrus, more variability is introduced, which leads to larger uncertainties.

There is one last issue that has to be discussed: How can the iRF from in-situ data be compared to the results from global simulations like in the IPCC assessments? RF or ERF is given from model simulations that calculate the radiative forcing from contrail coverage or volume. The contrails are parameterized and not resolved individually for certain points as with the iRF. Global models are not able to resolve contrails on a resolution of 1 x 1 kilometer like the iRF dataset. They do however include interactions between other contrails and natural cirrus, which is not considered by the iRF (Burkhardt and Kärcher, 2011). IRF can therefore not be directly

compared to global model results from ERF / RF. Nevertheless, what the results from the iRF dataset showed is, that there is a fundamental limit to how precise RF and ERF can be determined from model calculations. The many sources for natural variabilities and consequently the large uncertainty is inevitable. Gettelman et al. (2021) is a good example of that. The standard deviations for the ERF of 2019 and 2020 were 35 Wm^{-2} and 50 Wm^{-2} , which did not get smaller even when the number of ensemble members were doubled. This shows that we may never be able to constrain the uncertainty level significantly and to shrink the corresponding error bars associated with contrail cirrus RF / ERF. In principle, it would be possible to get a precise mean RF / ERF from measurements, but an unrealistically large number of radiation measurements would be necessary.

If that is known, research could now focus more on a reliable forecast of situations that would allow strong warming contrails. As mentioned, the climate effect of contrails cannot be extracted from iRF. For example, an iRF value that indicates a Big Hit at one point of a contrail does not mean, that the overall effect of the contrail is strongly warming. It is possible that the overall effect of the contrail is only positive (not strongly positive), as the forcings of many points along a large contrail balance themselves. Nonetheless, it is not likely that a contrail that at one point has a very strong warming leads to an overall cooling effect. Thus, one can be quite sure to reduce the overall radiative effect of contrails (and not prevent a cooling contrail) when preventing Big Hits only. Teoh et al. (2020a) showed in their study that 2 % of the flown flights distances contributed to 80 % of the energy forcing (EF). Preventing those would make a huge difference. A precise global, annual mean RF / ERF to is not needed to avoid such strongly warming contrails.

6.2 Meteorological conditions of Big Hits

In the ten years of data from MOZAIC flights, persistent contrails occurred in 10 % of all cases, while Big Hits only had a probability of 1 %. Of that 1 %, most Big Hits only had short extensions of less than 75 km in length. Big Hits occur rather patchy along the flight paths, which is disadvantageous for avoiding Big Hits by flight diversion. This observation, however, is largely influenced by the definition of when a Big Hit ends and a new one starts and at which iRF a persistent contrail becomes a Big Hit. The limit of $\text{iRF} \geq 19 \text{ Wm}^{-2}$ that was used in the study is just a number that separated the strongest 10 % of the iRF values in the distributions. A larger or smaller limit would result in different lengths of the Big Hits, as they were defined based on points exceeding the iRF limit. A gap of ten points, where the calculated iRF fell short of 19 Wm^{-2} , was allowed along a Big Hit. This gap translates to a distance of 10 kilometers of

40 seconds flight time. The next point where iRF exceeded the limit again, would record a new Big Hit. Therefore, the lengths of Big Hits would become larger if a larger gap would be allowed. It is also important to mention, that in this study only individual points along MOZAIC's flight paths were analyzed and those give a limited picture of the atmosphere. The aircraft crosses the ice supersaturated region (ISSR) only in one single line at a specific time, but its true vertical and horizontal extension and its lifetime is not known. However, it became clear, that high iRF values occur with a certain coherence and not randomly dispersed. The longer and more coherent a strongly warming contrail, the easier it is to adjust the flight level of the aircraft accordingly, to prevent its development.

As mentioned, the point prediction of persistent contrails is so far not accurate, because of a poor prediction of the humidity field in ECMWF's forecasts, and consequently unreliable relative humidity data. Contrails require a relative humidity over ice of the ambient air of at least 100 %. However, weather models tend to have problems with that threshold. This does not mean, that the relative humidity field in ERA – 5 is in general unreliable. Gierens et al. (2020) showed in a comparison of ERA – 5 to MOZAIC data, that RH_i in the upper troposphere region has a linear correlation of 0.67 – 0.92 depending on the data month used. For this study, RH_i from ERA – 5 and MOZAIC had a correlation of 0.89 for the whole dataset including all general cases, persistent contrails, and Big Hits of the years 2000 – 2009 (every 100th point). Linear correlation, however, is not sufficient for predicting persistent contrails from ERA – 5 RH_i values.

High RH_i values occur in the atmosphere both in clear and cloudy sky with low concentrations of ambient ice particles (Lamquin et al., 2009; Teoh, 2020). ERA – 5 underestimates the frequency and degree of ice supersaturation because it uses an approximation to compute ice supersaturation. In ECMWF's models, ice only forms when the limit for homogeneous nucleation is reached. When a cirrus cloud forms, all supersaturated humidity is converted to ice and RH_i returns to ice saturation ($RH_i = 100\%$) within one time step (Lamquin et al., 2009; Tompkins et al., 2007). In reality, the reduction of supersaturation does not happen instantly. The so-called relaxation time (Ovarlez et al., 2002), depends on the number of ambient ice crystals and the size distribution, and consequently the consumption of excess water vapor. The relaxation time is large with a low number of ambient ice crystals.

This assumption leads to an underprediction of high ice supersaturation values from ERA – 5. The maximum RH_i for ERA – 5 is at 1.4, while for MOZAIC a maximum of 1.71 was recorded for the whole dataset of this study. The correlation for RH_i between ERA – 5 and MOZAIC is only 0.03 for persistent contrails. The pdf's for RH_i from ERA – 5 showed that persistent

contrails (defined by T and RH_i from MOZAIC) had relative humidities over ice as low as zero. The pdf of RH_i for MOZAIC had its peak at $RH_i \sim 1$ for persistent contrails and ~ 1.3 for Big Hits. Fortunately, the pdf of RH_i for ERA – 5 for persistent contrails is also centered around one, but for Big Hits the pdf is not shifted towards larger RH_i values. A large portion of persistent contrails and Big Hits had an $RH_i \leq 0.75$ in the reanalysis. This suggests, that a lower threshold than the 100 % for RH_i is beneficial for predicting contrail persistence from models. Again, the goal of the analysis Big Hits was to find differences in meteorological conditions for Big Hits, persistent contrails, and for the general situation. Specifically, dynamical characteristics were of interest, as those have not been used for predicting persistence of contrails so far.

As seen from the pdf's of the dynamical variables, Big Hits and persistent contrails prefer negative vertical velocities (rising air motion), positive divergences (diverging/spreading air), negative vorticities (anticyclonic (clockwise) flow in the northern hemisphere), and low potential vorticities. Not much is known about dynamical characteristics of Big Hits and persistent contrails, but one can compare those to characteristics of ice supersaturated regions (ISSR's). ISSR's occur related to certain dynamical features. Gierens and Brinkop (Gierens and Brinkop, 2012) find, that ISSR's favor upward wind and divergent airflow. In the northern mid-latitudes, ISSR's are mainly located in areas of high pressure and anticyclonic flow. Those results coincide with the characteristics of Big Hits and persistent contrails from this study. Gierens and Brinkop account this pattern to the structure of warm/moist and cold/dry airstreams in mid-latitude synoptic disturbances (Gierens and Brinkop, 2012). From several studies, it is known that ISSR's occur related to certain synoptic weather features.

In the northern mid-latitudes, those are storm tracks over the North Atlantic (Gettelman et al., 2006; Lamquin et al., 2012; Spichtinger et al., 2003b; Wylie, 2002), the anticyclonic side of the polar jet stream (Diao et al., 2015) and inside anvil cirrus clouds (D'Alessandro et al., 2017). Kästner et al. (1999) found that most of their automatically detected persistent contrails in the upper troposphere occur either in divergent flow patterns in warm air masses, ahead of a surface warm front or a warm conveyor belt, or in locally turbulent cold air, ahead of a surface cold fronts near a band of strong wind, where the cold air is moving rapidly. In their case study, Spichtinger et al. (2005) also accounted fronts and warm conveyor belts to the formation of ISSR's, as those involve synoptic-scale upward motion. Irvine et al. (2012) showed, that the location of ISSR's is linked to orography and the time-mean location of the jet stream. For their research area (North Atlantic), the largest frequencies of ISSR occurrence were recorded over

Greenland (orographic lifting), the southern side of the jet stream, and around the northern edge of high-pressure ridges.

Knowing, that persistent contrails and Big Hits develop with distinct dynamical characteristics, one could now add those to the already used thermodynamic criteria for an enhanced persistent contrail prediction. This was tested with two simple methods, a logistic regression and with model output statistics for conditional probabilities. The results showed, that, 1) predicting Big Hits is quite reliable when it is already known for sure that contrails will be persistent and 2) relatively high probabilities for the persistence of contrails can be determined by introducing thresholds for dynamical variables and combining them with the SAC quantities. However, this was only a first try and more combinations of variables have to be tested further, eg. with a bigger dataset.

7 SUMMARY

Since the publication of the IPCC special report on Aviation and the Global Atmosphere in 1999 (Penner et al., 1999) there has been considerable research effort on the understanding of aviation's climate impact. Specifically, the radiative effect of contrails and contrail cirrus has been studied intensely in the last two decades. Unfortunately, despite all the progress made in process understanding and modeling and measurement possibilities, contrail cirrus radiative forcing is still associated with a low level of scientific understanding (LOSU). This terminology however is misleading. The science of contrails is not on a low level of understanding, as the involved physics is known rather well. The problem lies within the difficulty of determining the global, annual mean forcings (RF and ERF) with a narrow confidence band. This study investigates one of the major causes for uncertainty, the natural variability. Specifically, the weather-induced variability is examined from a large dataset of iRF values (instantaneous radiative forcing) of potential contrails. This was done by excluding any other sources of uncertainty, namely, modeling inadequacy, scenario uncertainty, and other natural variabilities. iRF values were produced by combining ten years (2000 – 2009) of MOZAIC measurement data from commercial passenger aircraft (16,588 flights) with ERA – 5 Reanalysis data. For each point at cruise level along each flight, the temperature and relative humidity from MOZAIC were checked for the Schmidt-Appleman criterion for contrail formation. If a persistent contrail was possible, the iRF was calculated using the respective radiation quantities from ERA – 5 using Schumann's radiation formulae (Schumann, 2012, p. 212) with Ebert and Curry's (Ebert and Curry, 1992) parameterization for the optical thickness. For the statistics of iRF only 1 % of the data points were chosen randomly to assure statistical independence. 47,032 iRF values were then analyzed for their variability. The cdf's and pdf's show strong annual and interannual variations, that are exclusively caused by the random character of the weather. A seasonal pattern in the data shows higher average iRF's and also more persistent contrails occurring in the summer months, while persistent contrails with a stronger cooling effect occurred in the winter. 80 % of the data have a slightly positive iRF up to 19 Wm^{-2} , 10 % show negative values and 10 % a strongly warming effect (Big Hit) with $\text{iRF} \geq 19 \text{ Wm}^{-2}$. The pdf distributions are wide and clearly non-gaussian. They are heavily skewed towards positive values, with the positive branch of the pdf for the whole time series being an exponential in the form of $f(x|x \geq 1) \propto e^{-bx}$, with the factor $b = 0.11^{-1} \text{ Wm}^{-2}$. In Monte-Carlo experiments, a mean iRF was calculated 1,000 times each for two different sample sizes (1 % and 0.1 % of the 47,032 iRF values), representative of the size of a measurement campaign

or a metastudy (comprising 10 measurement campaigns). The sample means scatter considerably depending on the sample size. This scatter is exclusively caused by weather variability. This variability is irreducible and sets a limit to how precise a mean RF can be calculated from measurements. Therefore, calculating a long-term mean RF from measurements that is more precise than what the model simulations predict, is not possible, unless an unrealistically large number of measurements is available. In reality, the variability of the overall RF of contrail cirrus is even larger when including all the other uncertainty sources. The distributions would become even wider.

The heavy tail of the pdf (figure 13), as well as the plot of the first order effect function (pdf x iRF) (figure 17 in chapter 5.1.3) are proof of a disproportionately large impact of strongly endothermic contrails (Big Hits) on the overall radiative effect, even though they have very low probabilities and occur rarely. In the dataset one point with $RF_{LW} \geq 100 \text{ Wm}^{-2}$, counterbalanced by a certain RF_{SW} , was present. Extreme value theory suggests that even larger values could be possible. As such strongly warming contrails (Big Hits) have the largest effect, it is the safest mitigation strategy to avoid specifically those. They are strong enough so that even with large uncertainties they do not entail the danger that the prevention measure leads to perverted results (eg. that these are not actually cooling contrails). Unfortunately, the accurate local prediction of contrails is still not possible because of errors in the humidity field in most weather prediction models. Because of the complex nature and the many involved processes, the humidity field is still poorly modeled, which leads to unreliable relative humidities in the upper troposphere and lower stratosphere. Especially large RH_i values, which are prone to produce Big Hits, are underrepresented. Therefore, this study focused on a new approach. When the meteorological and dynamical conditions of Big Hits are known, they could be used as an addition to the SAC quantities to improve prediction possibilities. This however is only possible if the distributions for Big Hits can be separated from those of “only” persistent or non-persistent contrails and if Big Hits occur in a coherent field and not only at random points in the atmosphere. Another dataset was produced that included every Big Hit ($iRF \geq 19 \text{ Wm}^{-2}$) along the ten years of MOZAIC flights from 2000 – 2009 with every flight point (not only one percent of the measurements). IRF was calculated in the same manner as for the first dataset. The results showed, that the major share of Big Hits have rather small extensions. However, 13 % of Big Hits still had a length of minimum 75 km, and 5 % were at least 150 km in length. The longest Big Hit exceeded 1500 km. The radiative effect of such a contrail is very large and should be avoided. The iRF along Big Hits changes only minimally with an amplitude of $\sim 5 \text{ Wm}^{-2}$ unless a change in the surrounding conditions induces stronger fluctuations. The distribution for iRF

from Big Hits follows the same exponential decline as the positive branch of the iRF for persistent contrails.

Pdf's were produced for meteorological conditions of Big Hits, persistent contrails, and the general case. The comparison of pdf's for thermodynamic and dynamic quantities revealed quite some differences depending on the variable. Big Hits and persistent contrails prefer temperatures between 210 – 230 K, high absolute humidities, high relative humidities with respect to ice, and an RH_{\max} between $\sim 1 - 1.5$ for the database. ERA – 5 and MOZAIC values agree well for temperature and RH_{\max} , but RH_i and absolute humidity are underestimated in ERA – 5. The dynamical characteristics vertical velocity and divergence show only minimal shifts towards negative and positive values for Big Hits and persistent contrails. Those variables are not as suitable for improving the prediction of persistent contrails. Vorticity, potential vorticity, geopotential height and the lapse rate, however, show a strong separation between the pdf's for Big Hits and persistent contrails to the pdf of the general case. Big Hits and persistent contrails prefer anticyclonic flow, low potential vorticities up to 4 PVU, large geopotential heights, and large lapse rates up to 10 K km^{-1} for the data. The pdf's for Big Hits (except for temperature and vorticity) is slightly more separated to the pdf for the general case than the pdf of persistent contrails. The distribution for μ , the cosine of the solar zenith angle, showed, that most Big Hits occur at night and or when the sun is very low and very high. Intermediate values are not favored. The pdf's overlap strongly, which means that persistent contrails and Big Hits do not only develop during the favored patterns. Only the relative frequency of the patterns is different within and outside of persistent contrails and Big Hits. Combining the quantities that showed the largest separation together with the SAC quantities could allow an improved prediction for contrail persistence and Big Hits, eg. in a regression analysis or a neural network. This was tested with a logistic regression model, which was able to predict Big Hit occurrence rather well, but only when it is already known for sure that contrails will be persistent. The logistic model was not able to predict contrail persistence. Model output statistics were used to predict the probability of persistent contrail formation. In a probability matrix, conditioned on the lapse rate, geopotential height, potential vorticity, relative vorticity, temperature, RH_{\max} , and RH_i , probabilities up to 93 % for contrail persistence were achieved. With an a priori probability for persistent contrails of 11.5 %, this is quite an improvement. The results imply that including dynamical variables to the contrail formation criterion (SAC) and ice supersaturation (ISS) might enable the prediction of persistent contrails on a local level. The thresholds for the conditioned probabilities could be included in a weather prediction model like ECMWF's integrated forecast model (IFS), to produce such probabilities before the flight

planning period. Avoidance of persistent contrails or Big Hits would then become more reliable. This, however, needs further research, as this study only gave an idea as to what the consideration of dynamical characteristics can do.

8 REFERENCES

- Appleman, H., 1953. The Formation of Exhaust Condensation Trails by Jet Aircraft. *Bull. Am. Meteorol. Soc.* 34, 14–20. <https://doi.org/10.1175/1520-0477-34.1.14>
- Atlas, D., Wang, Z., 2010. Contrails of Small and Very Large Optical Depth. *J. Atmospheric Sci.* 67, 3065–3073. <https://doi.org/10.1175/2010JAS3403.1>
- Beswick, K., Baumgardner, D., Gallagher, M., Raga, G.B., Minnis, P., Spangenberg, D.A., Volz-Thomas, A., Nedelec, P., Wang, K.-Y., 2015. Properties of small cirrus ice crystals from commercial aircraft measurements and implications for flight operations. *Tellus B Chem. Phys. Meteorol.* 67, 27876. <https://doi.org/10.3402/tellusb.v67.27876>
- Bickel, M., Ponater, M., Bock, L., Burkhardt, U., Reineke, S., 2020. Estimating the Effective Radiative Forcing of Contrail Cirrus. *J. Clim.* 33, 1991–2005. <https://doi.org/10.1175/JCLI-D-19-0467.1>
- Bier, A., 2018. Einfluss von Rußpartikelemissionen auf die Eiskristallbildung, Eigenschaften, Lebenszyklen und Klimawirkung von Kondensstreifenwirren. LMU, München.
- Bier, A., Burkhardt, U., Bock, L., 2017. Synoptic Control of Contrail Cirrus Life Cycles and Their Modification Due to Reduced Soot Number Emissions: Contrail Cirrus Life Cycles. *J. Geophys. Res. Atmospheres* 122, 11,584–11,603. <https://doi.org/10.1002/2017JD027011>
- Bock, L., Burkhardt, U., 2016. Reassessing properties and radiative forcing of contrail cirrus using a climate model: RADIATIVE FORCING OF CONTRAIL CIRRUS. *J. Geophys. Res. Atmospheres* 121, 9717–9736. <https://doi.org/10.1002/2016JD025112>
- Bockhorn, H. (Ed.), 1994. Soot Formation in Combustion: Mechanisms and Models, Springer Series in Chemical Physics. Springer Berlin Heidelberg, Berlin, Heidelberg. <https://doi.org/10.1007/978-3-642-85167-4>
- Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, D., Forster, P., Kerminen, V.-M., Kondo, Y., Liao, H., Lohmann, U., Rasch, P., Satheesh, S.K., Sherwood, S., Stevens, B., Zhang X. Y., 2013. Clouds and Aerosols, in: Stocker, T.F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V., Midgley, P.M. (Eds.), *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, UK and New York, USA.
- Burkhardt, U., Kärcher, B., 2011. Global radiative forcing from contrail cirrus. *Nat. Clim. Change* 1, 54–58. <https://doi.org/10.1038/nclimate1068>
- Burkhardt, U., Kärcher, B., 2009. Process-based simulation of contrail cirrus in a global climate model. *J. Geophys. Res.* 114, D16201. <https://doi.org/10.1029/2008JD011491>
- Burkhardt, U., Kärcher, B., Ponater, M., Gierens, K., Gettelman, A., 2008. Contrail cirrus supporting areas in model and observations. *Geophys. Res. Lett.* 35, L16808. <https://doi.org/10.1029/2008GL034056>
- Carleton, A.M., Silva, A.D., Bernhardt, J., VanderBerg, J., Travis, D.J., 2015. Subregion-Scale Hindcasting of Contrail Outbreaks, Utilizing Their Synoptic Climatology. *J. Appl. Meteorol. Climatol.* 54, 1733–1755. <https://doi.org/10.1175/JAMC-D-14-0186.1>
- Chen, C.-C., Gettelman, A., 2013. Simulated radiative forcing from contrails and contrail cirrus. *Atmospheric Chem. Phys.* 13, 12525–12536. <https://doi.org/10.5194/acp-13-12525-2013>

- Copernicus Climate Change Service (C3S), 2017. ERA5: Fifth Generation of ECMWF Atmospheric Reanalyses of the Global Climate. Copernicus Climate Change Service Climate Data Store (CDS [WWW Document]. URL <https://cds.climate.copernicus.eu/cdsapp#!/search?type=dataset&text=ERA5> (accessed 11.8.21).
- D'Alessandro, J.J., Diao, M., Wu, C., Liu, X., Chen, M., Morrison, H., Eidhammer, T., Jensen, J.B., Bansemer, A., Zondlo, M.A., DiGangi, J.P., 2017. Dynamical conditions of ice supersaturation and ice nucleation in convective systems: A comparative analysis between in situ aircraft observations and WRF simulations: Ice Supersaturation in Convective System. *J. Geophys. Res. Atmospheres* 122, 2844–2866. <https://doi.org/10.1002/2016JD025994>
- Diao, M., Jensen, J.B., Pan, L.L., Homeyer, C.R., Honomichl, S., Bresch, J.F., Bansemer, A., 2015. Distributions of ice supersaturation and ice crystals from airborne observations in relation to upper tropospheric dynamical boundaries: ICE SUPERSATURATION DYNAMICAL CONDITIONS. *J. Geophys. Res. Atmospheres* 120, 5101–5121. <https://doi.org/10.1002/2015JD023139>
- Duda, D.P., Minnis, P., 2009a. Basic Diagnosis and Prediction of Persistent Contrail Occurrence Using High-Resolution Numerical Weather Analyses/Forecasts and Logistic Regression. Part II: Evaluation of Sample Models. *J. Appl. Meteorol. Climatol.* 48, 1790–1802. <https://doi.org/10.1175/2009JAMC2057.1>
- Duda, D.P., Minnis, P., 2009b. Basic Diagnosis and Prediction of Persistent Contrail Occurrence Using High-Resolution Numerical Weather Analyses/Forecasts and Logistic Regression. Part I: Effects of Random Error. *J. Appl. Meteorol. Climatol.* 48, 1780–1789. <https://doi.org/10.1175/2009JAMC2056.1>
- Duda, D.P., Minnis, P., Nguyen, L., 2001. Estimates of cloud radiative forcing in contrail clusters using GOES imagery. *J. Geophys. Res. Atmospheres* 106, 4927–4937. <https://doi.org/10.1029/2000JD900393>
- Dürbeck, T., Gerz, T., 1996. Dispersion of aircraft exhausts in the free atmosphere. *J. Geophys. Res. Atmospheres* 101, 26007–26015. <https://doi.org/10.1029/96JD02217>
- Ebert, E.E., Curry, J.A., 1992. A parameterization of ice cloud optical properties for climate models. *J. Geophys. Res.* 97, 3831. <https://doi.org/10.1029/91JD02472>
- Eckstein, P., Götze, W., Hartl, F., Rönz, B., Strohe, H.G., 1994. *Lexikon Statistik*. Gabler Verlag, Wiesbaden. <https://doi.org/10.1007/978-3-322-91144-5>
- ERA5: data documentation - Copernicus Knowledge Base - ECMWF Confluence Wiki [WWW Document], n.d. URL <https://confluence.ecmwf.int/display/CKB/ERA5%3A+data+documentation> (accessed 11.26.21).
- Fuglestad, J.S., Shine, K.P., Berntsen, T., Cook, J., Lee, D.S., Stenke, A., Skeie, R.B., Velders, G.J.M., Waitz, I.A., 2010. Transport impacts on atmosphere and climate: Metrics. *Atmos. Environ.* 44, 4648–4677. <https://doi.org/10.1016/j.atmosenv.2009.04.044>
- Gayet, J.-F., Shcherbakov, V., Voigt, C., Schumann, U., Schäuble, D., Jessberger, P., Petzold, A., Minikin, A., Schlager, H., Dubovik, O., Lapyonok, T., 2012. The evolution of microphysical and optical properties of an A380 contrail in the vortex phase. *Atmospheric Chem. Phys.* 12, 6629–6643. <https://doi.org/10.5194/acp-12-6629-2012>

- Gettelman, A., Chen, C.-C., Bardeen, C.G., 2021. The climate impact of COVID-19-induced contrail changes. *Atmospheric Chem. Phys.* 21, 9405–9416. <https://doi.org/10.5194/acp-21-9405-2021>
- Gettelman, A., Fetzer, E.J., Eldering, A., Irion, F.W., 2006. The Global Distribution of Supersaturation in the Upper Troposphere from the Atmospheric Infrared Sounder. *J. Clim.* 19, 6089–6103. <https://doi.org/10.1175/JCLI3955.1>
- Gierens, K., 2021. Theory of Contrail Formation for Fuel Cells. *Aerospace* 8, 164. <https://doi.org/10.3390/aerospace8060164>
- Gierens, K., 2010. Contrails and Contrail Cirrus, in: *Encyclopedia of Aerospace Engineering*. John Wiley & Sons Ltd, Chichester, UK, pp. 3683–3694.
- Gierens, K., 2003. On the transition between heterogeneous and homogeneous freezing. *Atmospheric Chem. Phys.* 3, 437–446. <https://doi.org/10.5194/acp-3-437-2003>
- Gierens, K., 1998. Die Bedeckung des Himmels mit Kondensstreifen. *Meteorol. Z.* 7, 181–187. <https://doi.org/10.1127/metz/7/1998/181>
- Gierens, K., Brinkop, S., 2012. Dynamical characteristics of ice supersaturated regions. *Atmospheric Chem. Phys.* 12, 11933–11942. <https://doi.org/10.5194/acp-12-11933-2012>
- Gierens, K., Matthes, S., Rohs, S., 2020. How Well Can Persistent Contrails Be Predicted? *Aerospace* 7, 169. <https://doi.org/10.3390/aerospace7120169>
- Gierens, K., Schumann, U., Helten, M., Smit, H., Marenco, A., 1999. A distribution law for relative humidity in the upper troposphere and lower stratosphere derived from three years of MOZAIC measurements. *Ann. Geophys.* 17, 1218–1226. <https://doi.org/10.1007/s00585-999-1218-7>
- Gierens, K., Schumann, U., Helten, M., Smit, H., Wang, P.-H., 2000. Ice-supersaturated regions and subvisible cirrus in the northern midlatitude upper troposphere. *J. Geophys. Res. Atmospheres* 105, 22743–22753. <https://doi.org/10.1029/2000JD900341>
- Gierens, K., Spichtinger, P., 2000. On the size distribution of ice-supersaturated regions in the upper troposphere and lowermost stratosphere. *Ann. Geophys.* 18, 499–504. <https://doi.org/10.1007/s00585-000-0499-7>
- Gierens, K., Spichtinger, P., Schumann, U., 2012. Ice Supersaturation, in: Schumann, U. (Ed.), *Atmospheric Physics: Background – Methods – Trends*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 135–150. https://doi.org/10.1007/978-3-642-30183-4_9
- Gierens, K.M., Schumann, U., Smit, H.G.J., Helten, M., Zängl, G., 1997. Determination of humidity and temperature fluctuations based on MOZAIC data and parametrisation of persistent contrail coverage for general circulation models. *Ann. Geophys.* 15, 1057–1066. <https://doi.org/10.1007/s00585-997-1057-3>
- Gierens, K.M., Ström, J., 1998. A Numerical Study of Aircraft Wake Induced Ice Cloud Formation. *J. Atmospheric Sci.* 55, 3253–3263. [https://doi.org/10.1175/1520-0469\(1998\)055<3253:ANSOAW>2.0.CO;2](https://doi.org/10.1175/1520-0469(1998)055<3253:ANSOAW>2.0.CO;2)
- Helten, M., Smit, H.G.J., Sträter, W., Kley, D., Nedelec, P., Zöger, M., Busen, R., 1998. Calibration and performance of automatic compact instrumentation for the measurement of relative humidity from passenger aircraft. *J. Geophys. Res. Atmospheres* 103, 25643–25652. <https://doi.org/10.1029/98JD00536>

- Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., Thépaut, J.-N., 2018a. ERA5 hourly data on pressure levels from 1979 to present. Technical Report, Copernicus Climate Change Service (C3S) Climate Data Store (CDS) [WWW Document]. URL <https://cds.climate.copernicus.eu/cdsapp#!/dataset/10.24381/cds.bd0915c6?tab=form> (accessed 11.8.21).
- Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., Thépaut, J.-N., 2018b. ERA5 hourly data on single levels from 1979 to present. Technical Report, Copernicus Climate Change Service (C3S) Climate Data Store (CDS) [WWW Document]. URL <https://cds.climate.copernicus.eu/cdsapp#!/dataset/10.24381/cds.bd0915c6?tab=form> (accessed 11.8.21).
- Hersbach, H., Dick, D., 2016. ERA5 reanalysis is in production. European Centre for Medium-Range Weather Forecast (ECMWF). [WWW Document]. URL <https://www.ecmwf.int/en/newsletter/147/news/era5-reanalysis-production> (accessed 11.26.21).
- Hoffmann, L., Günther, G., Li, D., Stein, O., Wu, X., Griessbach, S., Heng, Y., Konopka, P., Müller, R., Vogel, B., Wright, J.S., 2019. From ERA-Interim to ERA5: the considerable impact of ECMWF's next-generation reanalysis on Lagrangian transport simulations. *Atmospheric Chem. Phys.* 19, 3097–3124. <https://doi.org/10.5194/acp-19-3097-2019>
- Holton, J.R., Haynes, P.H., McIntyre, M.E., Douglass, A.R., Rood, R.B., Pfister, L., 1995. Stratosphere-troposphere exchange. *Rev. Geophys.* 33, 403. <https://doi.org/10.1029/95RG02097>
- IAGOS Data Portal [WWW Document], 2019. URL <http://www.iagos-data.fr/> (accessed 11.23.21).
- IAGOS. Humidity sensor. [WWW Document], n.d. URL <https://www.iagos.org/iagos-core-instruments/h2o/> (accessed 11.23.21).
- IAGOS. Scientific publications. [WWW Document], n.d. URL <https://www.iagos.org/scientific-publications-2/> (accessed 11.23.21).
- IAGOS-CORE. In-service Aircraft for a Global Observing System. [WWW Document], n.d. URL <https://www.iagos.org/iagos-core-instruments/> (accessed 11.23.21).
- International Civil Aviation Organization (ICAO), 2019. Presentation of 2019 Air Transport Statistical Results [WWW Document]. URL <https://www.icao.int/sustainability/Pages/FactsFigures.aspx> (accessed 12.15.21).
- International Civil Aviation Organization, (ICAO), 1993. Manual of the ICAO Standard Atmosphere: Extended to 80 Kilometres (262 500 Feet). (No. Vol. 7488).
- Irvine, E.A., Hoskins, B.J., Shine, K.P., 2012. The dependence of contrail formation on the weather pattern and altitude in the North Atlantic: CONTRAILS, WEATHER PATTERNS AND ALTITUDE. *Geophys. Res. Lett.* 39, n/a-n/a. <https://doi.org/10.1029/2012GL051909>
- Jackson, A., Newton, B., Hahn, D., Bussey, A., 2001. Statistical Contrail Forecasting. *J. Appl. Meteorol.* 40, 269–279. [https://doi.org/10.1175/1520-0450\(2001\)040<0269:SCF>2.0.CO;2](https://doi.org/10.1175/1520-0450(2001)040<0269:SCF>2.0.CO;2)

- James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. *An Introduction to Statistical Learning*, Springer Texts in Statistics. Springer New York, New York, NY. <https://doi.org/10.1007/978-1-4614-7138-7>
- Jeßberger, P., Voigt, C., Schumann, U., Sölch, I., Schlager, H., Kaufmann, S., Petzold, A., Schäuble, D., Gayet, J.-F., 2013. Aircraft type influence on contrail properties. *Atmospheric Chem. Phys.* 13, 11965–11984. <https://doi.org/10.5194/acp-13-11965-2013>
- Kärcher, B., 2018. Formation and radiative forcing of contrail cirrus. *Nat. Commun.* 9, 1824. <https://doi.org/10.1038/s41467-018-04068-0>
- Kärcher, B., Yu, F., 2009. Role of aircraft soot emissions in contrail formation. *Geophys. Res. Lett.* 36, L01804. <https://doi.org/10.1029/2008GL036649>
- Kästner, M., Meyer, R., Wendling, P., 1999. Influence of weather conditions on the distribution of persistent contrails. *Meteorol. Appl.* 6, 261–271. <https://doi.org/10.1017/S1350482799001231>
- Kaufmann, S., Voigt, C., Heller, R., Jurkat-Witschas, T., Krämer, M., Rolf, C., Zöger, M., Giez, A., Buchholz, B., Ebert, V., Thornberry, T., Schumann, U., 2018. Intercomparison of midlatitude tropospheric and lower-stratospheric water vapor measurements and comparison to ECMWF humidity data. *Atmospheric Chem. Phys.* 18, 16729–16745. <https://doi.org/10.5194/acp-18-16729-2018>
- Koop, T., 2004. Homogeneous Ice Nucleation in Water and Aqueous Solutions. *Z. Für Phys. Chem.* 218, 1231–1258. <https://doi.org/10.1524/zpch.218.11.1231.50812>
- Koop, T., Luo, B., Tsias, A., Peter, T., 2000. Water activity as the determinant for homogeneous ice nucleation in aqueous solutions. *Nature* 406, 611–614. <https://doi.org/10.1038/35020537>
- Lamquin, N., Gierens, K., Stubenrauch, C.J., Chatterjee, R., 2009. Evaluation of upper tropospheric humidity forecasts from ECMWF using AIRS and CALIPSO data. *Atmospheric Chem. Phys.* 9, 1779–1793. <https://doi.org/10.5194/acp-9-1779-2009>
- Lamquin, N., Stubenrauch, C.J., Gierens, K., Burkhardt, U., Smit, H., 2012. A global climatology of upper-tropospheric ice supersaturation occurrence inferred from the Atmospheric Infrared Sounder calibrated by MOZAIC. *Atmospheric Chem. Phys.* 12, 381–405. <https://doi.org/10.5194/acp-12-381-2012>
- Lee, D.S., Fahey, D.W., Forster, P.M., Newton, P.J., Wit, R.C.N., Lim, L.L., Owen, B., Sausen, R., 2009. Aviation and global climate change in the 21st century. *Atmos. Environ.* 43, 3520–3537. <https://doi.org/10.1016/j.atmosenv.2009.04.024>
- Lee, D.S., Fahey, D.W., Skowron, A., Allen, M.R., Burkhardt, U., Chen, Q., Doherty, S.J., Freeman, S., Forster, P.M., Fuglestedt, J., Gettelman, A., De León, R.R., Lim, L.L., Lund, M.T., Millar, R.J., Owen, B., Penner, J.E., Pitari, G., Prather, M.J., Sausen, R., Wilcox, L.J., 2021. The contribution of global aviation to anthropogenic climate forcing for 2000 to 2018. *Atmos. Environ.* 244, 117834. <https://doi.org/10.1016/j.atmosenv.2020.117834>
- Letu, H., Nagao, T.M., Nakajima, T.Y., Riedi, J., Ishimoto, H., Baran, A.J., Shang, H., Sekiguchi, M., Kikuchi, M., 2019. Ice Cloud Properties From Himawari-8/AHI Next-Generation Geostationary Satellite: Capability of the AHI to Monitor the DC Cloud Generation Process. *IEEE Trans. Geosci. Remote Sens.* 57, 3229–3239. <https://doi.org/10.1109/TGRS.2018.2882803>

- Lewellen, D.C., Lewellen, W.S., 2001. The Effects of Aircraft Wake Dynamics on Contrail Development. *J. Atmospheric Sci.* 58, 390–406. [https://doi.org/10.1175/1520-0469\(2001\)058<0390:TEOAWD>2.0.CO;2](https://doi.org/10.1175/1520-0469(2001)058<0390:TEOAWD>2.0.CO;2)
- Lloyd, G., Gallagher, M., Choulaton, T., Krämer, M., Petzold, A., Baumgardner, D., 2020. In Situ Measurements of Cirrus Clouds on a Global Scale. *Atmosphere* 12, 41. <https://doi.org/10.3390/atmos12010041>
- Lufthansa, n.d. Airbus A340-300 [WWW Document]. URL <https://www.lufthansa.com/de/en/34p> (accessed 12.26.21).
- Lynch, D.K., Sassen, K., Starr, D.O., Sephens, G. (Eds.), 2002. *Cirrus*. Oxford University Press, Cambridge ; New York.
- Mannstein, H., Meyer, R., Wendling, P., 1999. Operational detection of contrails from NOAA-AVHRR-data. *Int. J. Remote Sens.* 20, 1641–1660. <https://doi.org/10.1080/014311699212650>
- Mannstein, H., Schumann, U., 2005. Aircraft induced contrail cirrus over Europe. *Meteorol. Z.* 14, 549–554. <https://doi.org/10.1127/0941-2948/2005/0058>
- Marengo, 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., Cook, T., 1998. Measurement of ozone and water vapor by Airbus in-service aircraft: The MOZAIC airborne program, an overview. *J. Geophys. Res. Atmospheres* 103, 25631–25642. <https://doi.org/10.1029/98JD00977>
- Marquart, S., 2003. Klimawirkung von Kondensstreifen: Untersuchungen mit einem globalen atmosphärischen Zirkulationsmodell. (DLR-Forschungsbericht 2003-16).
- Meerkötter, R., Schumann, U., Doelling, D.R., Minnis, P., Nakajima, T., Tsushima, Y., 1999. Radiative forcing by contrails. *Ann. Geophys.* 17, 1080–1094. <https://doi.org/10.1007/s00585-999-1080-7>
- Minikin, A., Petzold, A., Ström, J., Krejci, R., Seifert, M., van Velthoven, P., Schlager, H., Schumann, U., 2003. Aircraft observations of the upper tropospheric fine particle aerosol in the Northern and Southern Hemispheres at midlatitudes: UT FINE PARTICLE AEROSOL AT MIDLATITUDES. *Geophys. Res. Lett.* 30, n/a-n/a. <https://doi.org/10.1029/2002GL016458>
- Murphy, D.M., Koop, T., 2005. Review of the vapour pressures of ice and supercooled water for atmospheric applications. *Q. J. R. Meteorol. Soc.* 131, 1539–1565. <https://doi.org/10.1256/qj.04.94>
- Myhre, G., Kvalevåg, M., Rädcl, G., Cook, J., Shine, K.P., Clark, H., Karcher, F., Markowicz, K., Kardas, A., Wolkenberg, P., Balkanski, Y., Ponater, M., Forster, P., Rap, A., Leon, R., Rodriguez, 2009. Intercomparison of radiative forcing calculations of stratospheric water vapour and contrails. *Meteorol. Z.* 18, 585–596. <https://doi.org/10.1127/0941-2948/2009/0411>
- Myhre, G., Shindell, D., Bréon, F.-M., Collins, W., Fuglestedt, J., Huang, J., Koch, D., Lamarque, J.-F., Lee, D.S., Mendoza, B., Nakajima, T., Robock, A., Stephens, G., Takemura, T., Zhang, H., 2013. Anthropogenic and Natural Radiative Forcing, in: Intergovernmental Panel on Climate Change (Ed.), *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, pp. 659–740. <https://doi.org/10.1017/CBO9781107415324.018>

- Nedelec, P., Cammas, J.-P., Thouret, V., Athier, G., Cousin, J.-M., Legrand, C., Abonnel, C., Lecoeur, F., Cayez, G., Marizy, C., 2003. An improved infrared carbon monoxide analyser for routine measurements aboard commercial Airbus aircraft: technical validation and first scientific results of the MOZAIC III programme. *Atmospheric Chem. Phys.* 3, 1551–1564. <https://doi.org/10.5194/acp-3-1551-2003>
- Neis, P., Smit, H.G.J., Krämer, M., Spelten, N., Petzold, A., 2015. Evaluation of the MOZAIC Capacitive Hygrometer during the airborne field study CIRRUS-III. *Atmospheric Meas. Tech.* 8, 1233–1243. <https://doi.org/10.5194/amt-8-1233-2015>
- Ovarlez, J., Gayet, J.-F., Gierens, K., Ström, J., Ovarlez, H., Auriol, F., Busen, R., Schumann, U., 2002. Water vapour measurements inside cirrus clouds in Northern and Southern hemispheres during INCA: SUPERSATURATION INSIDE NH AND SH CIRRUS. *Geophys. Res. Lett.* 29, 60-1-60–4. <https://doi.org/10.1029/2001GL014440>
- Penner, J.E., Lister, D., Griggs, D., Dokken, D., McFarland, M. (Eds.), 1999. *Aviation and the global atmosphere*. Cambridge University Press for the Intergovernmental Panel on Climate Change, Cambridge, UK.
- Petzold, A., Krämer, M., Neis, P., Rolf, C., Rohs, S., Berkes, F., Smit, H.G.J., Gallagher, M., Beswick, K., Lloyd, G., Baumgardner, D., Spichtinger, P., Nédélec, P., Ebert, V., Buchholz, B., Riese, M., Wahner, A., 2017. Upper tropospheric water vapour and its interaction with cirrus clouds as seen from IAGOS long-term routine in situ observations. *Faraday Discuss.* 200, 229–249. <https://doi.org/10.1039/C7FD00006E>
- Petzold, A., Neis, P., Rütimann, M., Rohs, S., Berkes, F., Smit, H.G.J., Krämer, M., Spelten, N., Spichtinger, P., Nédélec, P., Wahner, A., 2020. Ice-supersaturated air masses in the northern mid-latitudes from regular in situ observations by passenger aircraft: vertical distribution, seasonality and tropospheric fingerprint. *Atmospheric Chem. Phys.* 20, 8157–8179. <https://doi.org/10.5194/acp-20-8157-2020>
- Petzold, A., Thouret, V., Gerbig, C., Zahn, A., Brenninkmeijer, C.A.M., Gallagher, M., Hermann, M., Pontaud, M., Ziereis, H., Boulanger, D., Marshall, J., Nédélec, P., Smit, H.G.J., Friess, U., Flaud, J.-M., Wahner, A., Cammas, J.-P., Volz-Thomas, A., IAGOS TEAM, 2015. Global-scale atmosphere monitoring by in-service aircraft – current achievements and future prospects of the European Research Infrastructure IAGOS. *Tellus B Chem. Phys. Meteorol.* 67, 28452. <https://doi.org/10.3402/tellusb.v67.28452>
- Ponater, M., 2002. Contrails in a comprehensive global climate model: Parameterization and radiative forcing results. *J. Geophys. Res.* 107, 4164. <https://doi.org/10.1029/2001JD000429>
- Pruppacher, H.R., Klett, J.D., 2010. *Homogeneous Nucleation*, in: *Microphysics of Clouds and Precipitation*, Atmospheric and Oceanographic Sciences Library. Springer Netherlands, Dordrecht, pp. 191–215. https://doi.org/10.1007/978-0-306-48100-0_7
- Reutter, P., Neis, P., Rohs, S., Sauvage, B., 2020. Ice supersaturated regions: properties and validation of ERA-Interim reanalysis with IAGOS in situ water vapour measurements. *Atmospheric Chem. Phys.* 20, 787–804. <https://doi.org/10.5194/acp-20-787-2020>
- Royal Aeronautical Society, 2019. *Annual Report 2018 - 2019. Greener by design*. London, UK.
- Russo, J., Romano, F., Tanaka, H., 2014. New metastable form of ice and its role in the homogeneous crystallization of water. *Nat. Mater.* 13, 733–739. <https://doi.org/10.1038/nmat3977>

- Schmidt, E., 1941. Die Entstehung von Eisnebel aus den Auspuffgasen von Flugmotoren. Schriften der Deutschen Akademie der Luftfahrtforschung, Verlag R. Oldenbourg, München, Heft 44.
- Schröder, F., Kärcher, B., Fiebig, M., Petzold, A., 2002. Aerosol states in the free troposphere at northern midlatitudes: FREE TROPOSPHERIC AEROSOL STATES. *J. Geophys. Res. Atmospheres* 107, LAC 8-1-LAC 8-8. <https://doi.org/10.1029/2000JD000194>
- Schröder, F., Ström, J., 1997. Aircraft measurements of sub micrometer aerosol particles (> 7 nm) in the midlatitude free troposphere and tropopause region. *Atmospheric Res.* 44, 333–356. [https://doi.org/10.1016/S0169-8095\(96\)00034-8](https://doi.org/10.1016/S0169-8095(96)00034-8)
- Schröder, F.P., Kärcher, B., Petzold, A., Baumann, R., Busen, R., Hoell, C., Schumann, U., 1998. Ultrafine aerosol particles in aircraft plumes: In situ observations. *Geophys. Res. Lett.* 25, 2789–2792. <https://doi.org/10.1029/98GL02078>
- Schumann, U., 2012. A contrail cirrus prediction model. *Geosci. Model Dev.* 5, 543–580. <https://doi.org/10.5194/gmd-5-543-2012>
- Schumann, U., 2005. Formation, properties and climatic effects of contrails. *Comptes Rendus Phys.* 6, 549–565. <https://doi.org/10.1016/j.crhy.2005.05.002>
- Schumann, U., 1996. Über Bedingungen zur Bildung von Kondensstreifen aus Flugzeugabgasen. *Meteorol. Z.* 5, 4–23. <https://doi.org/10.1127/metz/5/1996/4>
- Schumann, U., Baumann, R., Baumgardner, D., Bedka, S.T., Duda, D.P., Freudenthaler, V., Gayet, J.-F., Heymsfield, A.J., Minnis, P., Quante, M., Raschke, E., Schlager, H., Vázquez-Navarro, M., Voigt, C., Wang, Z., 2017. Properties of individual contrails: a compilation of observations and some comparisons. *Atmospheric Chem. Phys.* 17, 403–438. <https://doi.org/10.5194/acp-17-403-2017>
- Schumann, U., Graf, K., 2013. Aviation-induced cirrus and radiation changes at diurnal timescales. *J. Geophys. Res. Atmospheres* 118, 2404–2421. <https://doi.org/10.1002/jgrd.50184>
- Schumann, U., Graf, K., Mannstein, H., 2011. Potential to reduce the climate impact of aviation by flight level changes, in: 3rd AIAA Atmospheric Space Environments Conference. Presented at the 3rd AIAA Atmospheric Space Environments Conference, American Institute of Aeronautics and Astronautics, Honolulu, Hawaii. <https://doi.org/10.2514/6.2011-3376>
- Schumann, U., Heymsfield, A.J., 2017. On the Life Cycle of Individual Contrails and Contrail Cirrus. *Meteorol. Monogr.* 58, 3.1-3.24. <https://doi.org/10.1175/AMSMONOGRAPHS-D-16-0005.1>
- Schumann, U., Mayer, B., Graf, K., Mannstein, H., 2012. A Parametric Radiative Forcing Model for Contrail Cirrus. *J. Appl. Meteorol. Climatol.* 51, 1391–1406. <https://doi.org/10.1175/JAMC-D-11-0242.1>
- Schumann, U., Penner, J.E., Chen, Y., Zhou, C., Graf, K., 2015a. Dehydration effects from contrails in a coupled contrail–climate model. *Atmospheric Chem. Phys.* 15, 11179–11199. <https://doi.org/10.5194/acp-15-11179-2015>
- Schumann, U., Penner, J.E., Chen, Y., Zhou, C., Graf, K., 2015b. Dehydration effects from contrails in a coupled contrail–climate model. *Atmospheric Chem. Phys.* 15, 11179–11199. <https://doi.org/10.5194/acp-15-11179-2015>
- Schumann, U., Poll, I., Teoh, R., Koelle, R., Spinielli, E., Molloy, J., Koudis, G.S., Baumann, R., Bugliaro, L., Stettler, M., Voigt, C., 2021. Air traffic and contrail changes over

- Europe during COVID-19: a model study. *Atmospheric Chem. Phys.* 21, 7429–7450. <https://doi.org/10.5194/acp-21-7429-2021>
- Seidel, D.J., Randel, W.J., 2006. Variability and trends in the global tropopause estimated from radiosonde data. *J. Geophys. Res.* 111, D21101. <https://doi.org/10.1029/2006JD007363>
- Silverman, B.W., 1998. Density estimation for statistics and data analysis, Monographs on statistics and applied probability. Chapman & Hall/CRC, Boca Raton.
- Smit, H.G.J., Rohs, S., Neis, P., Boulanger, D., Krämer, M., Wahner, A., Petzold, A., 2014. Technical Note: Reanalysis of upper troposphere humidity data from the MOZAIC programme for the period 1994 to 2009. *Atmospheric Chem. Phys.* 14, 13241–13255. <https://doi.org/10.5194/acp-14-13241-2014>
- Smit, H.G.J., Volz-Thomas, A., Helten, M., Paetz, W., Kley, D., 2008. An In-Flight Calibration Method for Near-Real-Time Humidity Measurements with the Airborne MOZAIC Sensor. *J. Atmospheric Ocean. Technol.* 25, 656–666. <https://doi.org/10.1175/2007JTECHA975.1>
- Sornette, D., 2000. Critical Phenomena in Natural Sciences Chaos, Fractals, Selforganization and Disorder: Concepts and Tools.
- Spichtinger, P., Gierens, K., Leiterer, U., Dier, H., 2003a. Ice supersaturation in the tropopause region over Lindenberg, Germany. *Meteorol. Z.* 12, 143–156. <https://doi.org/10.1127/0941-2948/2003/0012-0143>
- Spichtinger, P., Gierens, K., Read, W., 2003b. The global distribution of ice-supersaturated regions as seen by the Microwave Limb Sounder. *Q. J. R. Meteorol. Soc.* 129, 3391–3410. <https://doi.org/10.1256/qj.02.141>
- Spichtinger, P., Gierens, K., Read, W., 2002. The statistical distribution law of relative humidity in the global tropopause region. *Meteorol. Z.* 11, 83–88. <https://doi.org/10.1127/0941-2948/2002/0011-0083>
- Spichtinger, P., Gierens, K., Smit, H.G.J., Ovarlez, J., Gayet, J.-F., 2004. On the distribution of relative humidity in cirrus clouds. *Atmospheric Chem. Phys.* 4, 639–647. <https://doi.org/10.5194/acp-4-639-2004>
- Spichtinger, P., Gierens, K., Wernli, H., 2005. A case study on the formation and evolution of ice supersaturation in the vicinity of a warm conveyor belt's outflow region. *Atmospheric Chem. Phys.* 5, 973–987. <https://doi.org/10.5194/acp-5-973-2005>
- Ström, L., Gierens, K.M., 2002. First simulations of cryoplane contrails. *J. Geophys. Res.* 107, 4346. <https://doi.org/10.1029/2001JD000838>
- Stuber, N., Forster, P., Rädcl, G., Shine, K., 2006. The importance of the diurnal and annual cycle of air traffic for contrail radiative forcing. *Nature* 441, 864–867. <https://doi.org/10.1038/nature04877>
- Sussmann, R., Gierens, K.M., 2001. Differences in early contrail evolution of two-engine versus four-engine aircraft: Lidar measurements and numerical simulations. *J. Geophys. Res. Atmospheres* 106, 4899–4911. <https://doi.org/10.1029/2000JD900533>
- Tarek, M., Brissette, F.P., Arsenaault, R., 2020. Evaluation of the ERA5 reanalysis as a potential reference dataset for hydrological modelling over North America. *Hydrol. Earth Syst. Sci.* 24, 2527–2544. <https://doi.org/10.5194/hess-24-2527-2020>
- Teoh, R., 2020. Climate forcing of aircraft contrails: uncertainty quantification and mitigation potential. <https://doi.org/10.25560/82293>

- Teoh, R., Schumann, U., Majumdar, A., Stettler, M.E.J., 2020a. Mitigating the Climate Forcing of Aircraft Contrails by Small-Scale Diversions and Technology Adoption. *Environ. Sci. Technol.* 54, 2941–2950. <https://doi.org/10.1021/acs.est.9b05608>
- Teoh, R., Schumann, U., Stettler, M.E.J., 2020b. Beyond Contrail Avoidance: Efficacy of Flight Altitude Changes to Minimise Contrail Climate Forcing. *Aerospace* 7, 121. <https://doi.org/10.3390/aerospace7090121>
- Tompkins, A.M., Gierens, K., Rädcl, G., 2007. Ice supersaturation in the ECMWF integrated forecast system. *Q. J. R. Meteorol. Soc.* 133, 53–63. <https://doi.org/10.1002/qj.14>
- Travis, D.J., Carleton, A.M., Changnon, S.A., 1997. An Empirical Model to Predict Widespread Occurrences of Contrails. *J. Appl. Meteorol.* 36, 1211–1220. [https://doi.org/10.1175/1520-0450\(1997\)036<1211:AEMTPW>2.0.CO;2](https://doi.org/10.1175/1520-0450(1997)036<1211:AEMTPW>2.0.CO;2)
- Unterstrasser, S., 2020. The Contrail Mitigation Potential of Aircraft Formation Flight Derived from High-Resolution Simulations. *Aerospace* 7, 170. <https://doi.org/10.3390/aerospace7120170>
- Unterstrasser, S., 2016. Properties of young contrails – a parametrisation based on large-eddy simulations. *Atmospheric Chem. Phys.* 16, 2059–2082. <https://doi.org/10.5194/acp-16-2059-2016>
- Unterstrasser, S., 2008. Numerische Simulationen von Kondensstreifen und deren Übergang in Zirren. Ludwig-Maximilians-Universität München. <https://doi.org/10.5282/EDOC.9464>
- Unterstrasser, S., Gierens, K., 2010. Numerical simulations of contrail-to-cirrus transition – Part 2: Impact of initial ice crystal number, radiation, stratification, secondary nucleation and layer depth. *Atmospheric Chem. Phys.* 10, 2037–2051. <https://doi.org/10.5194/acp-10-2037-2010>
- Unterstrasser, S., Görsch, N., 2014. Aircraft-type dependency of contrail evolution. *J. Geophys. Res. Atmospheres* 119, 14,015–14,027. <https://doi.org/10.1002/2014JD022642>
- Unterstrasser, S., Sölch, I., 2010. Study of contrail microphysics in the vortex phase with a Lagrangian particle tracking model. *Atmospheric Chem. Phys.* 10, 10003–10015. <https://doi.org/10.5194/acp-10-10003-2010>
- Vázquez-Navarro, M., Mannstein, H., Kox, S., 2015. Contrail life cycle and properties from one year of MSG/SEVIRI rapid-scan images. *Atmospheric Chem. Phys. Discuss.* 15, 7019–7055. <https://doi.org/10.5194/acpd-15-7019-2015>
- Vazquez-Navarro, M., Mannstein, H., Mayer, B., 2010. An automatic contrail tracking algorithm. *Atmospheric Meas. Tech.* 3, 1089–1101. <https://doi.org/10.5194/amt-3-1089-2010>
- Warren, S.G., Brandt, R.E., Grenfell, T.C., 2006. Visible and near-ultraviolet absorption spectrum of ice from transmission of solar radiation into snow. *Appl. Opt.* 45, 5320. <https://doi.org/10.1364/AO.45.005320>
- Wilhelm, L., Gierens, K., Rohs, S., 2021. Weather Variability Induced Uncertainty of Contrail Radiative Forcing. *Aerospace* 8, 332. <https://doi.org/10.3390/aerospace8110332>
- Williams, T., Kelley, C., 2019. Gnuplot 5.2. An Interactive Plotting Program.
- Wylie, D., 2002. Cirrus and Weather. A satellite Perspective., in: Lynch, D.K., Sassen, K., O’C Starr, D., Stephens, G. (Eds.), *Cirrus*. Oxford University Press, pp. 136–146.

- Xiao, H., Zhang, F., He, Q., Liu, P., Yan, F., Miao, L., Yang, Z., 2019. Classification of Ice Crystal Habits Observed From Airborne Cloud Particle Imager by Deep Transfer Learning. *Earth Space Sci.* 6, 1877–1886. <https://doi.org/10.1029/2019EA000636>
- Yang, P., Hong, G., Dessler, A.E., Ou, S.S.C., Liou, K.-N., Minnis, P., Harshvardhan, 2010. Contrails and Induced Cirrus: Optics and Radiation. *Bull. Am. Meteorol. Soc.* 91, 473–478. <https://doi.org/10.1175/2009BAMS2837.1>
- Yin, F., Grewe, V., Gierens, K., 2020. Impact of Hybrid-Electric Aircraft on Contrail Coverage. *Aerospace* 7, 147. <https://doi.org/10.3390/aerospace7100147>

APPENDIX A

IDL program for kernel density estimator:

```

; NAME:
; KDE
; PURPOSE:
; Estimate the probability density underlying a set of discrete
; samples (measurements) using the kernel density estimator method.
; CATEGORY:
; Statistics
; CALLING SEQUENCE:
; rho = kde(p, x)
; INPUTS:
; p: discrete measurements of the desired distribution: [ndimensions, nmeasurements]
; x: values at which to estimate the probability density: [ndimensions, nsamples]
; KEYWORD PARAMETERS:
; weight: weighting for sampled points.
; KEYWORD OUTPUTS:
; sigma: estimate for the statistical uncertainty of the estimated density.
; variance: estimate for the variance between the returned density
; and the true underlying density.
; scale: smoothing factor, also called the bandwidth, used to
; compute the density estimate
; KEYWORD FLAGS:
; By default, KDE uses the Epanechnikov kernel to compute
; the kernel density estimate, this can be overridden by
; setting one of the following flags:
; GAUSSIAN: Gaussian kernel
; TRIANGULAR: triangular kernel
; BIWEIGHT: biweight kernel
; OUTPUTS:
; rho: probability density estimated at each value specified by x
; RESTRICTIONS:
; Gaussian kernel used for multivariate systems.
; PROCEDURE:
; rho_i = (1/n) \sum_{j=1}^n K((p_j - x_i)/h) / h
; where h is the estimated optimal smoothing parameter and
; where K(z) is the selected kernel.
; REFERENCES:
; 1. B. W. Silverman,
; Density Estimation for Statistics and Data Analysis
; (CRC Press, Boca Raton, 1998)
; 2. B. E. Hansen, "Lecture Notes on Nonparametrics"
; University of Wisconsin, Spring 2009
; http://www.ssc.wisc.edu/~bhansen/718/NonParametrics1.pdf
; 3. Z. Ouyang, "Univariate Kernel Density Estimation"
; Duke University, August 2005
; http://www.stat.duke.edu/~z02/shared/research/readings/kernelsmoothing.pdf
; EXAMPLE:
; IDL> p = random(seed, 10000)
; IDL> x = 2. * findex(100)/99.
; IDL> rho = kde(p,x)
; IDL> plot, x, rho
; IDL> plot, x, histogram(p, min=0, max=2, nbins=100), /noerase
; MODIFICATION HISTORY:
; 09/18/2010 Written by David G. Grier, New York University
; 10/08/2011 DGG Corrected for case when more than 3/4 of input data
; have same value. Thanks to Dan Hartung (UW Madison)
; for bringing this bug to light.
; 10/18/2011 DGG Added WEIGHT keyword.
; 12/09/2012 DGG Use rebin() rather than # operator for efficiency
; and clarity in n-dimensional code. Test scale with
; arg_present(). Move normalization out of loops for efficiency.
; Corrected n-dimensional normalization. Updated usage messages.
; 03/22/2013 DGG rebin(/sample) is more efficient.
; 02/10/2014 DGG Added VARIANCE keyword.
; 02/13/2014 DGG Cast indexes to long to avoid integer overruns.
; Cast nx to float.
; 02/25/2014 DGG Implemented MSE.
; 03/01/2014 DGG Revised MSE calculations.
; 03/03/2014 DGG Implemented BIAS.
; 05/01/2014 DGG and Henrique Moyses Implemented SIGMA.
; 05/02/2014 DGG Array-based implementation. Eliminated BIAS and MSE
; in favor of VARIANCE.
; 05/20/2014 HM, David Ruffner and Chen Wang corrected normalization
; for N-dimensional case.
; Copyright (c) 2010-2014 David G. Grier, Henrique Moyses,
; David Ruffner and Chen Wang
;
function kde_nd, x, y, $
    weight = weight, $
    scale = scale, $
    variance = variance, $
    sigma = sigma

    compile_opt idl2, hidden

    sx = size(x, /dimensions)
    sy = size(y, /dimensions)

    nd = sx[0] ; number of dimensions
    nx = float(sx[1]) ; number of data points
    ny = long(sy[1]) ; number of sampling points

    if ~isa(weight, /number, /array) then $
        weight = replicate(1., nx)

    if n_elements(weight) ne nx then $
        message, "weight must have the same number of elements as the first argument"

    ; optimal smoothing parameter in each dimension
    ; Silverman Eqs. (3.30) and (3.31)
    sx = stddev(x, dimension=2)
    rx = fltarr(nd)
    for d = 0, nd-1 do $
        rx[d] = iqr(x[d,*])
    h = sx
    w = where(rx gt 1e-10, ngood)
    if ngood gt 0 then $
        h[w] = h[w] < rx[w]/1.34
    h *= 0.9 / nx^0.2

    if arg_present(scale) then scale = h

    ; density estimate
    ; Silverman Eq. (2.15) and Table 3.1
    res = fltarr(ny)
    variance = fltarr(ny)
    sigma = fltarr(ny)
    hfac = rebin(h, nd, nx, /sample)

    norm = product((2. * |p1 * h^2)^(-0.5)) / nx
    for j = 0, ny-1 do begin
        z = 0.5 * total((x - rebin(y[*],j], nd, nx, /sample)) / hfac)^2, 1)
        w = where(z lt 20., ngood)
        if ngood gt 0 then begin
            ker = norm * exp(-z[w])
            val = weight[w] * ker
            res[j] = total(val)
            sigma[j] = total(val^2)
            variance[j] = total((val - res[j])^2)/nx
        endif
    endfor

    sigma = sqrt(sigma)

    return, res
end

function kde_id, x, y, $
    weight = weight, $
    scale = scale, $
    biweight = biweight, $
    triangular = triangular, $
    gaussian = gaussian, $
    variance = variance, $
    sigma = sigma

    compile_opt idl2, hidden

    nx = float(n_elements(x)) ; number of data points
    ny = n_elements(y) ; number of samples

    if ~isa(weight, /number, /array) then $
        weight = replicate(1., nx)
    if n_elements(weight) ne nx then $
        message, "weight must have same number of elements as first argument"

    ; optimal smoothing parameter
    ; Silverman Eqs. (3.30) and (3.31)
    sx = stddev(x) ; standard deviation
    rx = iqr(x) ; interquartile range
    if rx lt 1e-10 then $
        h = 0.9 * sx / nx^0.2 $
        ; more than 3/4 data have same value
    else $
        h = 0.9 * (sx < rx/1.34) / nx^0.2

    if arg_present(scale) then scale = h

    ; density estimate
    ; Silverman Eq. (2.15) and Table 3.1
    t = x/h
    s = y/h

    z = rebin(t, nx, ny, /sample) - rebin(transpose(s), nx, ny, /sample)

    if keyword_set(biweight) then begin
        norm = (15./16.) / (h * nx)
        z *= z
        mask = (z lt 1.)
        value = norm * mask * (1. - z)^2
    endif else if keyword_set(triangular) then begin
        norm = 1./(h * nx)
        z = abs(z)
        mask = (z lt 1)
        value = norm * mask * (1. - z)
    endif else if keyword_set(gaussian) then begin
        norm = 1./sqrt(2.*pi) * h * nx
        z *= z/2.
        mask = (z lt 20.)
        value = norm * mask * exp(-z * mask)
    endif else begin
        norm = (3./4./sqrt(5.)) / (h * nx)
        z *= z/5.
        mask = (z lt 1.)
        value = norm * mask * (1. - z)
    endelse

    value *= rebin(weight, nx, ny, /sample)
    result = total(value, 1)
    if arg_present(sigma) then $
        sigma = sqrt(total(value^2, 1))
    if arg_present(variance) then $
        variance = total((value - rebin(transpose(result), nx, ny, /sample))^2, 1) / nx

    return, result
end

```

```

function kde, x, y, $
    weight = weight, $
    scale = scale, $
    gaussian = gaussian, $
    biweight = biweight, $
    triangular = triangular, $
    variance = variance, $
    sigma = sigma

COMPILE_OPT IDL2

umsg = 'rho = kde(p, x, [weight = w])'

if n_params() ne 2 then $
    message, umsg

sx = size(x)
sy = size(y)

if sx[0] gt 2 then begin
    message, umsg, /inf
    message, 'P must be organized as [ndimensions, npoints]'
endif

if sy[0] ne sx[0] then begin
    message, umsg, /inf
    message, 'P and X must have the same numbers of dimensions'
endif

if (sx[0] eq 2) and (sx[1] ne sy[1]) then begin
    message, umsg, /inf
    message, 'P and X must have the same numbers of dimensions'
endif

ndims = (sx[0] eq 2) ? sx[1] : 1

if ndims gt 1 then begin
    if keyword_set(biweight) or keyword_set(triangular) then $
        message, 'Multidimensional: using Gaussian kernel', /inf
    res = kde_nd(x, y, weight = weight, scale = scale, $
        variance = variance, $
        sigma = sigma)
endif else $
    res = kde_1d(x, y, weight = weight, scale = scale, $
        gaussian = gaussian, $
        biweight = biweight, $
        triangular = triangular, $
        variance = variance, $
        sigma = sigma)

return, res
end

```