Influence of Temperature and Humidity on Contrail Formation Regions in the General Circulation Model EMAC: A Spring Case Study

Patrick Peter[1,2], Sigrun Matthes[1], Christine Frömming[1], Patrick Jöckel[1], Luca Bugliaro[1], Andreas Giez[3], Martina Krämer[4,5], and Volker Grewe[1,2]

[1]Deutsches Zentrum für Luft- und Raumfahrt, Institut für Physik der Atmosphäre, Oberpfaffenhofen, Germany
[2]TU Delft, Faculty for Aerospace Engineering, The Netherlands
[3]Deutsches Zentrum für Luft- und Raumfahrt, Flugexperimente, Oberpfaffenhofen, Germany
[4]Forschungszentrum Jülich GmbH, IEK-7, Jülich, Germany
[5]Institute for Atmospheric Physics (IPA), Johannes Gutenberg University, Mainz, Germany

Correspondence: Patrick Peter (patrick.peter@dlr.de)

Abstract. While carbon dioxide emissions from aviation often dominate climate change discussions, ~~the significant impact of~~ non-CO$_2$ ~~effects like contrails and contrail-cirrus must not be overlooked, particularly for the mitigation of~~ effects such as contrails and contrail cirrus must also be considered. Despite varying estimates of their radiative forcing, avoiding contrails is a reasonable strategy for reducing aviation's climate effects. This study ~~evaluates~~ examines temperature and humidity, which are key atmospheric parameters ~~influencing~~ for contrail formation, ~~specifically temperature and humidity, using various model setups~~ across different ECHAM/MESSy Atmospheric Chemistry (EMAC) model setups. EMAC, a general circulation model ~~(GCM) with different~~, is evaluated with various vertical resolutions and two nudging methods ~~for~~ across seven specified dynamics setups. ~~Comparing simulation results~~ A higher vertical resolution aims to capture steep water vapour gradients near the tropopause, crucial for accurate contrail prediction. Comparisons with reanalysis data ~~for Marchand~~ (March–April 2014 ~~reveals~~) indicate a systematic cold bias ~~in mean temperatures across all altitudes and latitudes~~ (approximately 3–5 K in mid-latitudes), particularly in ~~the mid-latitudes where the bias is about 3-5 , unless~~ setups without mean temperature nudging ~~is applied~~. In the ~~upper-troposphere~~ upper troposphere/lower stratosphere, ~~the humidity of the nudged GCM simulations shows~~ all simulations exhibit a wet bias, while lower altitudes display a dry bias ~~is observed at lower altitudes. These biases result in overestimated regions for contrail formation in GCM simulations compared to reanalysis data. A point-by-point comparison along flown trajectories with measurement aircraft data shows~~, both affecting contrail formation estimates. Point-by-point comparisons along aircraft trajectories confirm similar biases. ~~Exploring~~ Sensitivity experiments with varying thresholds of relative humidity over ice ~~(RH$_{ice}$) threshold values for identifying ice-supersaturation regions provides insights into the risks of false alarms for contrail formation, together with information on hit rates . Accepting a false alarm rate of 16% results in a hit rate of about 40% (RH$_{ice}$ threshold 99%), while aiming for an 80% hit rate increases the false alarm rate to at least 35% (RH$_{ice}$ threshold 91-94%). A comprehensive~~ illustrate trade-offs between achieving high hit rates and minimizing false alarms in contrail detection. A one-day case study ~~, involving aircraft-based observations and satellite data, confirms contrail detection in regions identified as potential contrail coverage areas by the GCM~~ integrating aircraft and

satellite observations demonstrates that EMAC's predicted contrail coverage aligns well with observed formation. These results suggest that, despite existing temperature and humidity biases, EMAC reliably simulates contrail formation across diverse atmospheric conditions. Addressing these biases by refining temperature and humidity representation could significantly improve contrail prediction accuracy, strengthening confidence in avoidance strategies and supporting climate-optimized flight routing to mitigate aviation's overall climate effect.

1 Introduction

Estimates of the effective radiative forcing of aviation from 1940 to 2018 indicate that a significant part (best estimate 2/3) stems from non-CO_2 effects, including contrails, NO_x and H_2O emissions (Lee et al., 2021). However, this estimated 2/3 contribution is a global and annual average and especially the contribution from contrails varies significantly between individual flights (Grewe et al., 2014; Dahlmann et al., 2021; Teoh et al., 2022). Hence, operational mitigation options exist to avoid regions that are more climate sensitive than others. These regions are largely affected by the formation of strongly warming contrails or an over-proportional production of ozone from NO_x emissions. Methods are required to identify these climate sensitive regions and enable their forecast within numerical weather prediction models.

The Originally, the ECHAM/MESSy Atmospheric Chemistry (EMAC) model was utilised utilized to calculate climate change functions (CCFs) (Grewe et al., 2014; Frömming et al., 2021). More recently, Engberg et al. (2024) employed a comparable methodology using pycontrails 0.51.0, enabling direct comparisons between the two approaches as discussed in their study. These CCFs are a 4-dimensional measure of the global climate impact effect of individual local aviation emissions in dependency of, depending on the location, altitude, and time of emission, as well as the prevailing meteorological conditions at the time of emission. By identifying regions which are particularly sensitive to aviation emissions, the CCFs represent the basis for climate-optimized weather dependent, weather-dependent aircraft routing (Matthes et al., 2020; Lührs et al., 2018; Grewe et al., 2014). As discussed by Grewe et al. (2017), the implementation of climate change functions for This study aims to provide additional information on model uncertainties and their effects on the resulting CCFs. Additionally, the results from this study will serve as a basis for recalculating contrail-specific CCFs, further enhancing climate-optimized flight planning is only reasonable if the CCFs . As discussed by Grewe et al. (2017), the use of CCFs in daily operations requires the implementation of fast algorithms in daily weather forecast systems, and their corresponding mitigation potential are must be evaluated in a consistent modelling framework. However, validation opportunities are limited. One validation approach has been conducted to Such a framework ensures that the derived climate responses to aviation emissions are physically and methodologically coherent, allowing a fair comparison across different scenarios and emission types. Achieving this requires an integrated setup that includes the algorithmic CCFs as well as an air traffic optimizer integrated into a general circulation model. So far, this approach has only been applied to NO_x emission effects, and a validation exercise has been carried out to test the impact of NO_x emissions on ozone (Yin et al., 2023; Rao et al., 2022).

Hence, we evaluate different EMAC model setups and analyse whether the model is able to reproduce basic meteorological fields from other models or observations. Since contrails have the largest effect on climate-sensitive regions and are therefore

the primary target of air traffic management (Matthes et al., 2020; Lührs et al., 2021), we concentrate on the evaluation of contrail-forming parameters such as temperature and humidity in our study. These parameters are significant in identifying the regions where contrails may form and also influence their life cycle (Schumann, 1996; Kärcher, 2018). For instance, the use of particular temperature thresholds to calculate regions where contrails may form implies that significant differences between

60 models and observations could lead to inaccurate forecasts of such regions and subsequently affect the climate-optimized aircraft trajectories, leading to an even greater climate response. Therefore, it is crucial that current climate models are evaluated to ensure their accuracy, while at the same time being cautious how model results are used further.

The forecast of ice-supersaturation in the atmosphere, which is a crucial factor for the formation of persistent contrails, remains a challenging area in aviation-climate modelling (Gierens et al., 2020; Schumann et al., 2021; Rädel and Shine, 2010; Reutter

65 et al., 2020; Tompkins et al., 2007). Climate models show different levels of humidity biases (Gierens et al., 1997; Brinkop et al., 2015; Kaufmann et al., 2018; Krüger et al., 2022). The EMAC model employed ~~in this study also shows~~ here exhibits a temperature bias in the upper troposphere/lower stratosphere (UTLS) region (Stenke et al., 2007; Jöckel et al., 2016)~~that likely is a consequence of~~, likely due to numerical diffusion of water vapour across the tropopause (Stenke et al., 2007)~~and not uncommon~~, a known issue in climate models (Charlesworth et al., 2023). ~~However,~~ Nevertheless, because of the complex

70 coupling between atmospheric dynamics and the hydrological cycle, other processes cannot be completely ruled out. In the setups considered here, we nudge temperature rather than humidity. Temperature is more stable and less variable in time and space, making it more suitable for spectral nudging towards reanalysis data at synoptic scales. In contrast, the higher spatial and temporal variability of water vapour makes direct humidity nudging both infeasible and potentially destabilizing. Furthermore, no systematic analyses of the influence of model ~~vertical and horizontal resolution or of temperature nudging for "specified

75 dynamics"~~ resolution or temperature nudging in "specified dynamics" setups on contrail prediction exist yet. ~~Therefore~~ To address this gap, we first compare key atmospheric parameters for contrail formation and life cycle ~~between~~ across different EMAC model setups~~. Special focus is given~~, placing special emphasis on comparing different methodologies of Newtonian relaxation (nudging) towards reanalysis data (specified dynamics ~~(SD)~~ setups). Second, we validate the model's ability to correctly identify regions where contrails are likely to form. We compare EMAC simulation results with atmospheric

80 measurements ~~,~~ by analysing temperature and humidity data sampled along the trajectories of the High Altitude and Long Range (HALO) research aircraft during the ML-CIRRUS campaign. This comparison aims to assess how accurately the model reproduces observed conditions encountered in real flight scenarios. Finally, we evaluate satellite images, and compare ~~regions where contrails are visible with predicted contrail forming~~ them with model-predicted contrail-forming areas obtained from our ~~SD~~ specified dynamics simulation.

85 This paper is structured as follows: Section 2 describes the EMAC model and the experimental design (Sect. 2.1), followed by a description of the different methods to identify atmospheric conditions for contrail formation (Sect. 2.2). Afterwards, ECMWF reanalysis data (Sect. 2.3), and atmospheric observations used for the comparison (Sect. 2.4) are explained in detail. Section 3 compares temperature and humidity data calculated with different EMAC setups with ECMWF data, while investigating the prediction of contrail formation areas. In Section 4, we compare contrail formation parameters along aircraft trajectories to

90 highlight the discrepancy between simulated and observed quantities. Section 5 presents a more detailed analysis along the tra-

jectory for a selected case study, accompanied by satellite images, and Section 6 discusses our principal findings. In Section 7, we ~~summarise~~ summarize the key points, provide concluding remarks, and suggest aspects for future research.

2 Materials and Methods

This subsection outlines the structure of the model and the corresponding simulations (Section 2.1) and explains the different methods to identify atmospheric conditions for the formation of persistent contrails (Section 2.2), followed by a brief summary of the reanalysis data used for the comparison (Section 2.3). In the final subsection, all sources of observational data are illustrated in detail (Section 2.4).

2.1 Atmospheric Modelling: The Earth-system model EMAC

Within the framework of the second version of the Modular Earth Submodel System (MESSy2), the EMAC model is ~~utilised~~ utilized to investigate physical processes that are important for the climate effect of contrails. EMAC is a numerical chemistry-climate model simulation system that includes sub-models describing atmospheric processes from the troposphere to the mesosphere and their interactions with the ocean, land and human influences. The fifth generation of the European Centre Hamburg General Circulation Model (ECHAM5; Roeckner et al., 2006) is used as the core atmospheric model. The physical subroutines of the original ECHAM code have been modularised and re-implemented as MESSy sub-models with ongoing development. ECHAM now only retains the spectral transform dynamical core, the flux-form semi-Lagrangian large scale advection scheme (Lin and Rood, 1996), and the Newtonian relaxation-based nudging routines for specified dynamics ~~(SD)~~ model setups (Jöckel et al., 2010; Jöckel et al., 2016).

Experiment design and simulation setup

For the present study we applied EMAC (MESSy version 2.55) in a general circulation model ~~(GCM)~~ setup (i.e. without interactive chemistry) with spherical truncation of T42 and T63 (corresponding to quadratic Gaussian grids of approximately 2.8 by 2.8 degrees and 1.9 by 1.9 degrees in latitude and longitude, respectively). We performed simulations with three different vertical resolutions, one with 31 (L31) vertical hybrid pressure layers, another with 41 (L41) vertical hybrid pressure layers, and a third one with 90 (L90) vertical hybrid pressure layers, reaching from the surface up to 10 hPa (L31), 5 hPa (L41), and 0.01 hPa (L90), respectively. The vertical resolution of the 41 layer simulation, similar to that with 90 layers, includes layers of about 20 hPa height in the UTLS region (see Figure 1). While the L31 setup has the fewest vertical layers, it also has the lowest vertical resolution in the UTLS region. Within the L90 setup, there are more layers in the upper stratosphere above 50 hPa compared to the other setups. Due to the absence of these additional layers in the L41 or L31 model setups, various middle and upper stratospheric processes remain poorly represented. Consequently, the upper stratosphere dynamics, such as the Quasi-Biennial Oscillation, are not accurately represented in L41 and L31.

All simulations in this study obtained initial data (monthly averages of January 2013) on humidity and methane from the same L90 pre-simulation RD1SD-base-01. This is a EMAC ~~SD~~ specified dynamics simulation that followed the ref D1-protocol of
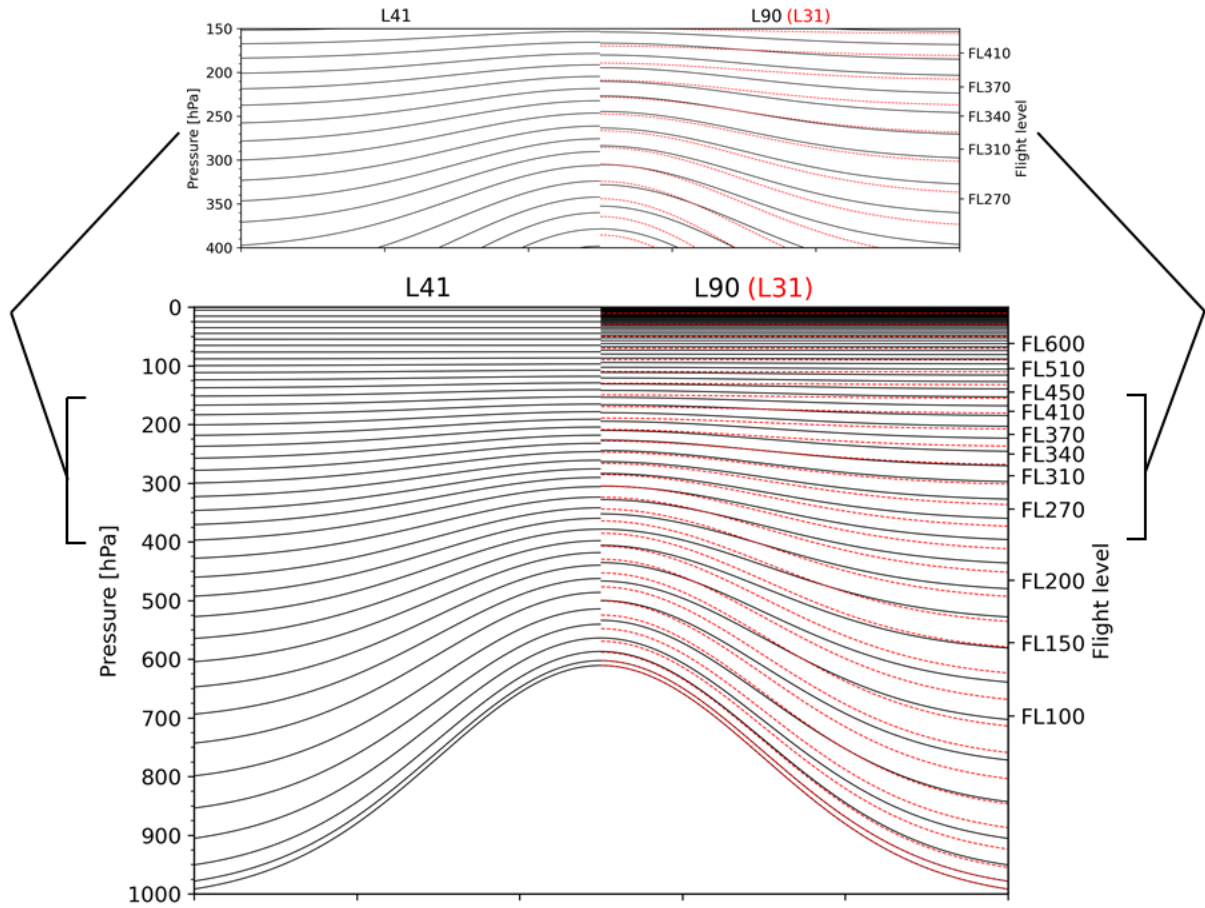
Figure 1. Vertical distribution of height levels in EMAC: L41 (left), L31 (right, red), and L90 (right, black) at a ground pressure of 1013 hPa and an idealized orography.

the Chemistry-Climate Model Initiative 2 (CCMI-2) (Plummer et al., 2021). The simulations were set up with Newtonian re-
laxation (nudging) towards the ERA5 reanalysis data (Hersbach et al., 2020). Additionally, a simulation with the L41 resolution
also underwent nudging towards ERA-Interim reanalysis data for comparison (Dee et al., 2011). The nudging of the ECHAM5
125 base model is applied in the spectral space for four prognostic variables, namely, divergence, vorticity, temperature, and the
logarithm of the surface pressure. The corresponding relaxation times for nudging, defined as the time for a model variable to
adjust to prescribed data, is 24 hours for temperature and the logarithm of the surface pressure, 6 hours for vorticity, and 48
hours for divergence. The nudging strength is not applied homogeneously in the vertical dimension: the boundary layer and
the layers above 126 hPa (L31), 100 hPa (L41) and 97 hPa (L90) are not nudged, a transition zone with intermediate strength
130 exists in between (see also Jöckel et al., 2016, supplement). Two different nudging settings were used: in two cases, the global

Table 1. Overview of the EMAC model setups used in this study. For the T42L41 STN setup, two simulations were performed with different nudging data (ERA5 and ERA-Interim, respectively), where STN refers to standard temperature nudging and MTN refers to mean temperature nudging.

Name[a]	T42L31 STN	T42L31 MTN	T63L31 STN	T42L41* STN	T42L41 MTN	T42L90 STN
Grid cells (lon,lat,lev)	128 x 64 x 31	128 x 64 x 31	192 x 96 x 31	128 x 64 x 41	128 x 64 x 41	128 x 64 x 90
Vertical model boundaries [km]	0–30	0–30	0–30	0–47	0–47	0–80
Vertical model boundaries [hPa]	Surface–10	Surface–10	Surface–10	Surface–5	Surface–5	Surface–0.01
Highest point with max. nudging [hPa]	126	126	126	100	100	97
Nudging data	ERA5	ERA5	ERA5	ERA5 (ERAI)	ERA5	ERA5
Relaxation of global mean temperature	No	Yes	No	No	Yes	No

(a) T refers to a triangular truncation at a specific wave number (42 or 63), corresponding to a quadratic Gaussian grid of approximately 2.8 by 2.8 degrees and 1.9 by 1.9 degrees in latitude and longitude, respectively, L refers to the number of vertical hybrid pressure levels.

mean temperature ("wave zero") was included for the Newtonian relaxations (T42L31 MTN and T42L41 MTN), in all other cases it was omitted. A stronger interference in the case of the MTN simulations (involving additional nudging of the global mean temperature at the corresponding model levels) results in a generally better agreement of EMAC model temperature with ERA5 reanalysis data. The differences and agreement between the model setups used are listed in Table 1.

135 While no fully unnudged simulation was conducted for this study, previous work by Jöckel et al. (2016) has shown that the unnudged EMAC model exhibits both temperature and humidity biases. Standard nudging procedures (excluding global mean temperature nudging) already help reduce these biases compared to the completely free-running model state. Including the global mean temperature in the nudging (i.e., using MTN setups) further reduces the temperature bias, and as a result also influences the humidity fields as shown in this study.

140 All simulations had a time step length of 10 min. As in RD1SD-base-01, the time evolution of the methane tracer at the lower model boundary was prescribed by Newtonian relaxation towards the data provided by CCMI-2. Since methane oxidation in the stratosphere constitutes a significant source of water vapour, accurately representing its variation is crucial for realistic atmospheric humidity distributions. In the CH$_4$ sub-model (Winterstein and Jöckel, 2021), the stratospheric water vapour contribution from methane oxidation ~~was accounted for by the photolysis calculated online in~~ is calculated using photolysis rates

145 provided online by the photo-chemistry sub-model JVAL (Riede et al., 2009). ~~Two~~ This ensures that the background humidity conditions, which influence contrail formation and persistence, are more accurately represented.
Two EMAC sub-models are ~~of particular relevance in~~ particularly relevant for this study: S4D ~~or "~~('sampling in 4 dimensions~~" (Jöckel et al., 2010)~~'; Jöckel et al., 2010) and CONTRAIL (Version 1.0; Frömming et al., 2014). ~~In our study, the~~ The

CONTRAIL sub-model identifies regions where contrails and contrail cirrus may form and/or persist, as described in more detail in Section 2.2. The S4D sub-model enables online sampling of model data along aircraft flight trajectories. In previous studies, a standard bi-linear horizontal interpolation was used to determine the meteorological variables along these tracks. For this study, we adopt a newly implemented nearest-neighbour approach to capture local temperature and humidity conditions more accurately. Vertical interpolation, if needed, remains linear in pressure altitude in both cases. We utilized flight position files from the ML-CIRRUS campaign, which have a temporal resolution of one minute. However, sampling occurs every model time step (10 min). After sampling, we separate the data into tropospheric ($PV \leq 2.5$) and stratospheric ($PV > 2.5$) regimes based on the model's potential vorticity (PV). PV is a commonly used metric for distinguishing between the troposphere and stratosphere, as it provides a dynamic and physically meaningful representation of the tropopause, particularly in mid-latitude regions. The threshold value of PV = 2.5 PVU was selected according to Zängl and Wirth (2002). Based on the analysis, it is found that 55% of the data points were in the troposphere, whereas 45% were located in the stratosphere.

Initially, a spin-up simulation was performed separately for each model setup from January 2013 to February 2014, producing output of global data, either instantaneous or temporal averages every 5 hours. The main simulations were then performed from 1 March 2014 to 30 April 2014, with an output interval of 1 hour. All analyses were conducted for the extended North Atlantic Flight Corridor (NAFC, 120°W-20°E), Asia (50°E-160°E), and the Northern Hemisphere, between 80°N and 20°N, at pressure levels where air traffic takes place. The NAFC is of particular interest due to its high share of global air traffic, frequent occurrence of contrail formation conditions, and because CCFs have been calculated in this region (Frömming et al., 2021).

2.2 Methods for identifying atmospheric conditions to form persistent contrails in global climate models

In this study, we investigate two different approaches to determine atmospheric regions capable of forming persistent contrails and contrail cirrus, without distinguishing between linear contrails and contrail cirrus.

One approach uses the parametrizations developed by Burkhardt et al. (2008) and Burkhardt and Kärcher (2009), as implemented in the EMAC CONTRAIL sub-model, to estimate the potential contrail coverage (PotCov). PotCov represents the fraction of a grid box that can be maximally covered by contrails under given large-scale temperature and humidity conditions. It is computed as the difference between (i) the maximum possible combined coverage of natural cirrus clouds and contrails, and (ii)

the coverage of natural cirrus clouds alone. Both coverage values depend on critical temperature and humidity thresholds with respect to ice, which are adapted for large-scale model conditions. These thresholds also incorporate the Schmidt–Appleman

185 Criterion (SAC; Schumann, 1996), which is derived from fundamental physical principles of mass, momentum, and energy conservation. In essence, the SAC determines whether hot exhaust gases mixing with cold ambient air will become supersaturated with respect to supercooled liquid water, enabling ice particle nucleation. This is often visualized using a 'mixing line' in the temperature–humidity space, which tracks how temperature and humidity evolve as exhaust and ambient air combine. Besides ambient temperature and humidity, the SAC accounts for aircraft-related parameters such as propulsion efficiency

190 and fuel combustion heat. In this study, we adopt engine parameters as provided by Grewe et al. (2014). Contrails persist only when the atmosphere is supersaturated with respect to ice; otherwise, they dissipate shortly after formation. Because these small-scale processes cannot be explicitly resolved in EMAC, their effects are represented through the parametrizations following Burkhardt et al. (2008). Further details on the calculation of PotCov are given in Grewe et al. (2014) and Frömming et al. (2021) .

195 Reanalysis data, such as ERA5, do not provide fields for PotCov directly. Therefore, an alternative approach is needed to compare the atmospheric potential for contrail persistence. This alternative identifies ice-supersaturated regions (ISSR) using the method ISSRs) following the methodology suggested by Dietmüller et al. (2023). Two criteria are defined to identify ISSRsISSRs are usually defined by the following two criteria: The temperature must be below 235 K to separate from mixed-phase regions (Pruppacher et al., 1998) ,and the relative humidity with respect to ice (RH_{ice}) must ex-

200 ceed 100% (see Yin et al., 2023, their supplementary material on contrail aCCFsReutter et al., 2020). However, when studying the sub-grid-scale variability in the relative humidity field of numerical weather forecast model data, such as ERA5, it is essential to take into account consider RH_{ice} thresholds below 100% as demonstrated by Irvine et al. (2014). Dietmüller et al. (2023) explored different relative humidity over ice RH_{ice} thresholds for ERA5 and conducted a comparison with MOZAIC data (Petzold et al., 2020) in the European region. The aim was to pre-determine areas where contrails form

205 before assessing the climate effect using algorithmic climate change functions for these identified regions. They found the best agreement between ERA5 and observations is achieved when the RH_{ice} threshold is set to 90%. Alternatively, the atmospheric ability to form persistent contrails and contrail cirrus in EMAC can be calculated at each time step according to Burkhardt et al. (2008) and Burkhardt and Kärcher (2009).Here, the fraction of a grid box available for contrail coverage is determined by the Schmidt-Appleman criterion, incorporating local temperature and humidity conditions conducive to contrail

210 formation. This fraction is calculated as the difference between the potential coverage of contrails and cirrus clouds combined and the coverage of natural cirrus clouds alone. For our study, we utilise values for the overall propulsion efficiency of an aircraft and the fuel combustion heat, which are essential for calculating the SAC, from Schumann et al. (2000). Further details about the calculation method for the potential contrail coverage are given by Frömming et al. (2021) and Grewe et al. (2014). In this study we explore RH_{ice} thresholds between 90% and 100%.

215 Here, we compare both approaches using identical EMAC atmospheric data under various model setups, as described in Section 3.2.

2.3 ECMWF reanalysis data

To assess simulation results fast and efficiently, a comparison with validated model results is useful. Reliable reanalysis data are required for an objective evaluation. Therefore, we ~~utilised~~ utilized hourly and monthly global operational reanalysis data from the European Centre for Medium-Range Weather Forecasts (ECMWF) for March and April 2014. For our comparison, we used the fifth generation of ECMWF reanalysis (ERA5) data with a spatial resolution of approximately 31 km and 137 vertical hybrid levels reaching from the surface up to 0.01 hPa (about 80 km). At typical flight altitudes around 11.5 km (200 hPa), the vertical resolution is approximately 300 m (10 hPa). A comprehensive set of monthly mean temperature and specific humidity ERA5 data, which were aggregated from daily means, were provided by the German Climate Computing Center DKRZ on the High Performance Computing system Levante. The data was re-mapped on a 0.25° x 0.25° grid and interpolated from the original ERA5 hybrid model levels to pressure levels. The same procedure was applied to the hourly data of temperature and specific humidity for 26 March 2014 (Hersbach et al., 2020). As there is no monthly or hourly relative humidity data available on Levante, we acquired this data from the Copernicus Climate Data Store. For temperatures above 273.15 K, the relative humidity is calculated based on saturation over water, while for temperatures below 250 K, it is calculated based on saturation over ice. For temperatures between 250 K and 273.15 K, it is determined by interpolating between the values for ice and water using a quadratic function. The relative humidity data has 37 interpolated pressure levels and was re-gridded to a regular lat-lon grid of 0.25°. For our study, we only used relative humidity data between 200 hPa and 350 hPa (Hersbach et al., 2023). All data sets have global coverage.

2.4 Atmospheric observations

Further knowledge on the performance of the model can be gained through comparing model data and observations. While a global coverage is normally not available for most atmospheric parameters at high resolution, comparing specific areas can provide valuable insight into the accuracy of model parameters and opportunities for model improvement. In this study, measurement data from the Midlatitude Cirrus (ML-CIRRUS, Voigt et al., 2017) measurement campaign are compared with EMAC model results at cruise altitude. Furthermore, satellite measurements were ~~utilised~~ utilized to compare the areas where contrail formation was predicted in the model with observed contrails.

Aircraft Experiment ML-CIRRUS

The ML-CIRRUS campaign, which took place from March 10 to April 16 2014, was one of the first scientific missions to demonstrate the capabilities of the novel High Altitude and Long Range Research Aircraft (HALO; www.halo.dlr.de). A comprehensive overview of the scientific objectives, flight plan, and equipment is provided by Voigt et al. (2017). Over the campaign period, HALO conducted 16 research flights, which amounted to a total of 88 flight hours. The flights were designed to comprehensively ~~characterise~~ characterize mid-latitude cirrus and contrail cirrus using in situ and remote sensing instruments. Unlike young contrails, which can be distinguished from natural cirrus by their exceptionally high concentrations of very small ice crystals (diameters <10 μm), aged contrails (i.e., contrail cirrus) become micro physically similar to natural cirrus

(Krämer et al., 2020). Consequently, identifying contrail cirrus solely by their properties is challenging. For the ML-Cirrus study, contrail cirrus forecasts from the CoCiP model Schumann (2012) guided the flight planning process, allowing the targeted sampling of regions where contrail cirrus were likely to occur and enabling a comparative analysis of both cloud types.

The ML-CIRRUS project aimed to enhance our understanding of cirrus cloud formation across varying meteorological conditions (Krämer et al., 2016; Luebke et al., 2016; Wernli et al., 2016; Urbanek et al., 2017), to refine our estimations of the radiative impact of cirrus (Krisna et al., 2018), and to assess air traffic impacts on high cloud cover (Schumann et al., 2017; Grewe et al., 2017; Li et al., 2023). The flights covered almost the entire central European region from the northern British coast to Portugal. To ~~fulfil~~ fulfill the scientific objectives of the mission, the HALO payload for ML-CIRRUS consisted of the instrumentation for measuring cloud particles, aerosols and trace gases including five different water vapour instruments. For our research, we analysed data collected by the Fast In situ Stratospheric Hygrometer (FISH). This closed-cell Lyman-α photofragment fluorescence hygrometer has been used on multiple research aircraft for over two decades (Meyer et al., 2015; Schiller et al., 2009). The instrument's operating principle is described in detail by Zöger et al. (1999). FISH is capable of measuring water vapour mixing ratios ranging from 1 to 1000 ppm. The overall uncertainty during ML-CIRRUS was found to be 6% relative with a ±0.4 ppm absolute offset uncertainty (Kaufmann et al., 2018). It is important to note that FISH measures 'total water,' which includes gas phase water as well as evaporated ice crystals. When comparing with gas phase H_2O from the model, only values up to 80% RH_{ice} are used to exclude clouds, as within clouds, RH_{ice} is elevated due to ice particles. Therefore, we substituted all FISH values (with a water vapour mixing ratio > 15 ppmV) where RH_{ice} was larger than 80% with measurements from the SHARC sensor. SHARC (Sophisticated Hygrometer for Atmospheric ReseaRCh) is a tunable diode laser ~~(TDL)~~ hygrometer and is part of the HALO basic data acquisition system (BAHAMAS; Giez et al., 2023). This closed-cell hygrometer uses the absorption line of water vapour at 1.37 ~~μm~~ µm. More details about the configuration of the SHARC sensor during the ML-CIRRUS campaign can be found in Kaufmann et al. (2018). The temperature data used in this study was also measured by BAHAMAS (Giez et al., 2023).

Since the interval of the observation data is 1 sec, the alignment with the 10-minute interval of the model is important. Initially, we computationally derived weighted average mean values for the measurement data through the implementation of a Gaussian filter. We utilized the fact that infinite repetitions of a boxcar filter can be converted into a Gaussian filter (Gans and Gill, 1984) and therefore performed multiple loops of a ~~centered~~ centred rolling mean to achieve a weighting of the data. Following this, the data was resampled to 10-minute intervals. ~~In order to~~ For spatial alignment, we used the averaged positions derived from the ML-CIRRUS flight track data as input to the EMAC submodel S4D (Jöckel et al., 2010). The S4D submodel interpolates the three-dimensional model fields to the exact geographical location and time corresponding to each 10-minute averaged observational data point. This procedure ensures that the model data are sampled at the correct geolocation thereby minimizing spatial representativeness errors.

To focus on contrail formation ~~areas and to~~ regions and remove outliers, ~~observation values were restricted~~ we restricted observed values to temperatures below 240 K, resulting in 308 data point pairs ~~or~~ (51 hours of flight time. ~~The correlation~~

between). We then computed the correlation between the observational and model data was calculated using a linear regression function analysis.

Satellite Data

The SEVIRI instrument aboard the geostationary Meteosat Second Generation (MSG) satellite is used to observe contrails in the region probed by the HALO aircraft during ML-CIRRUS on 26 March 2014. This is the most prominent contrail outbreak encountered during the ML-CIRRUS campaign and has been investigated using in situ cloud probes, airborne lidar observations and MSG/SEVIRI data in Wang et al. (2023). Thus, this day is very well suited for a comparison with the climate model data investigated in this study. These observations were conducted over the North Atlantic region, specifically near the coast of Ireland. This area is strategically significant as it lies along the major air traffic routes between Europe and the United States, making it a prime location for studying contrail formation and behavior behaviour (Sec. 2.4, see also Wang et al. (2023) for a detailed evaluation of this flight). The SEVIRI imager (Schmetz et al., 2002) aboard the operational MSG satellite MET-10 observes the Earth in 11 spectral channels with a spatial sampling distance of 3 km at the sub-satellite point and a temporal resolution of 15 min. Due to its location in the geostationary orbit above 0° E this sensor is well-suited to observe Europe. However, it has limitations, such as being unable to fully capture regions like the Western North Atlantic Ocean. Additionally, the spatial resolution decreases towards the edge of the observed Earth disk, making it challenging to identify small structures like contrails. In this study, false colour composites (so called ash RGBs) are used in order to visually spot contrails. These ash RGBs combine three brightness temperatures and temperature differences for the MSG/SEVIRI channels centred at 8.7, 10.8 and 12.0 μm. Here, contrails appear as dark blue or black lines linear objects and are easy to identify, so this method . This kind of pictures is often used for this purpose (e.g. Meijer et al., 2022) as a basis for automatic contrail detection as e.g. in (Meijer et al., 2022). In this study, we restrict ourselves to a visual analysis of the ash RGBs in order to identify areas of frequent contrail formation. Furthermore, the use of purely thermal channels makes these observations independent of solar illumination such that the ash RGBs can be used during day and night. However, the moderate resolution of MSG/SEVIRI only allows to distinguish contrails that are 30-60 min old (Vázquez-Navarro et al., 2015; Gierens and Vázquez-Navarro, 2018). These are persistent contrails which have formed in ice-supersaturated regions and can live for hours. For the sake of comparison with the model results, the data is remapped from the original satellite projection to an equal longitude-latitude grid.

3 Influence of different vertical model resolutions and nudging techniques on temperature, humidity and ice super-saturation

As the representation of key meteorological parameters is critical for the understanding of contrail formation and life cycle, this sections identifies differences in the atmospheric distribution of temperature and humidity between EMAC simulations with various vertical resolutions (L31, L41, L90), horizontal (spectral) resolutions (T42, T63), and different nudging approaches (standard temperature nudging (STN), or mean temperature nudging (MTN)). Firstly, the simulated temperature and humidity
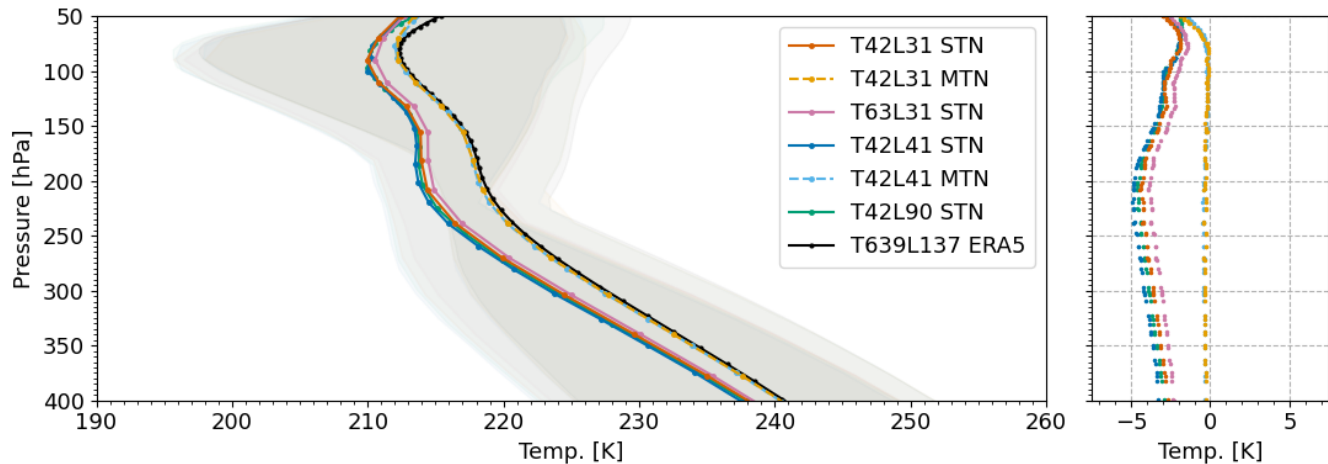
Figure 2. Left: Area-weighted mean temperature distribution from EMAC simulations with different vertical resolution (L31: red, L41: blue, L90: green) and ERA5 data (L137: black) for the North Atlantic Flight Corridor. The mean values for each grid box were calculated beforehand by using hourly data spanning from March to April 2014. Nudging approaches: standard (STN, *solid*) and mean (MTN, *dashed*) temperature nudging, with 95th percentiles shaded. Right: Temperature difference between EMAC and ERA5.

315 of various model setups are compared to ERA5 data in three specific regions (NAFC, Asia and the whole northern hemisphere) for March and April 2014. In order to better understand the local peculiarities, the parameters were examined separately within 20° latitude bands for all selected regions. In the end, this analysis focuses on the temperature and humidity data for 26 March 2014. Here, two distinct methods are used to calculate potential contrail formation regions for this specific day.

3.1 Temperature and humidity profile comparison for March and April 2014

320 Figure 2 illustrates the area weighted mean temperature across six distinct EMAC model setups (see Table 1) for the extended NAFC between 50 hPa and 400 hPa in comparison with ERA5 for March and April 2014. Notably, our comparison reveals a cold bias in EMAC simulations with standard temperature nudging compared to ERA5 reanalysis data for the analysed region. The identified bias for the area-weighted average mean temperature data ranges from 3 K to 5 K between 400 hPa and 50 hPa, with the most significant discrepancy around 200 hPa. Increasing the spectral (horizontal) resolution from T42 to T63

325 leads to a minimal reduction of the bias, by up to 0.2 K, while changes in vertical resolution have negligible effect. The STN model setup with 41 vertical layers exhibits the strongest deviation from ERA5, being 0.1 K colder than other STN setups up to 100 hPa. The analysis of 20° latitude bands for the NAFC (Figure 4, left) indicates a consistent altitude-dependent temperature bias across all bands. However, the application of mean temperature nudging (T42L31MTN, T42L41MTN), as explained in Section 2.1, significantly reduces, as to be expected, the cold bias to less than 0.1 K across all ~~analyzed~~ analysed pressure levels

330 and bands. ~~In~~ Since the STN setups constrain only selected meteorological variables (e.g., vorticity, divergence, temperature, and surface pressure) to large-scale reanalysis patterns without enforcing an accurate global mean temperature, a systematic
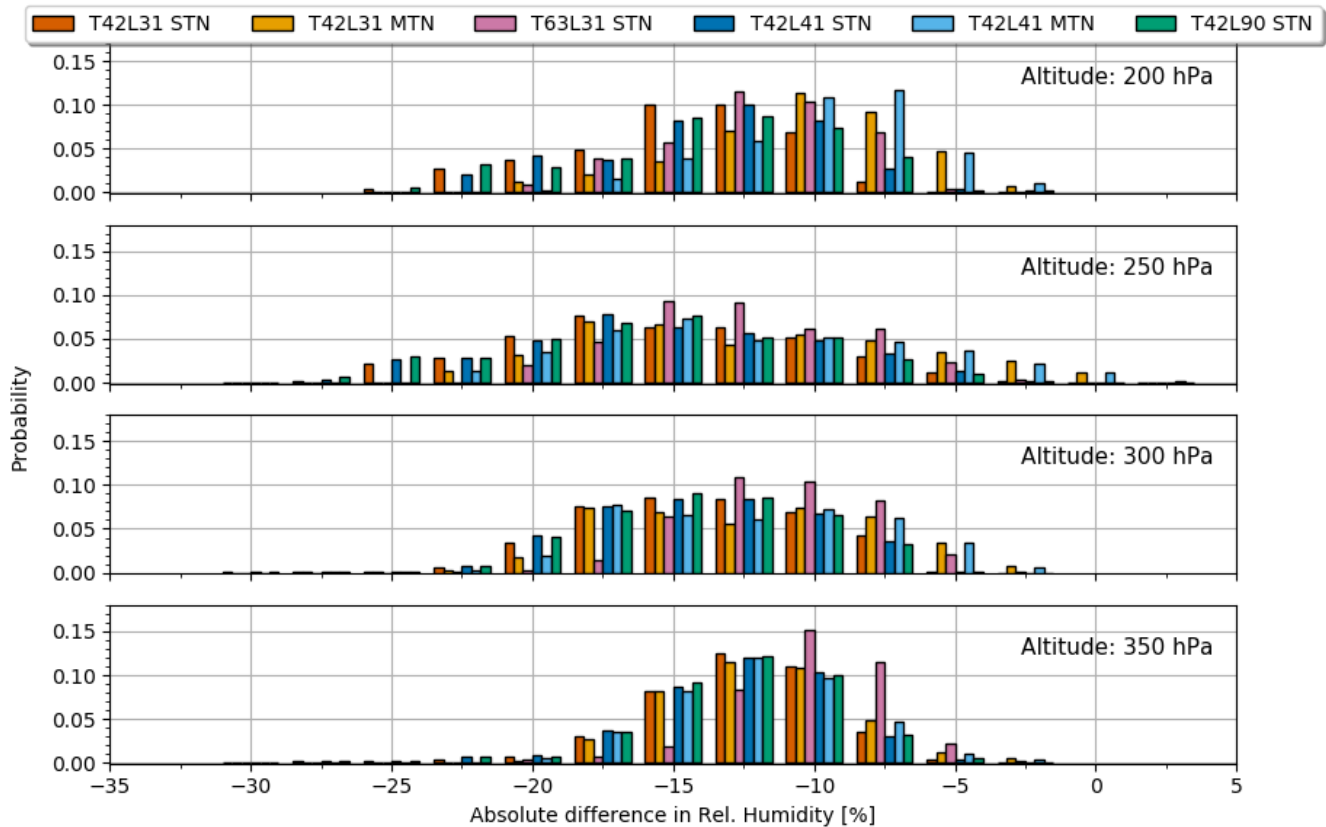
Figure 3. Probability density functions illustrating the absolute differences in relative humidity over ice (given in percentage points) between various EMAC model setups and ERA5 data (ERA5 minus EMAC) in the extended NAFC region for March and April 2014. The figure displays the distribution for four vertical pressure levels, starting from 200 hPa at the top and concluding with 350 hPa at the bottom. At each pressure level, the vertical model resolution and nudging method is indicated by different colors: red (L31), dark blue (L41), and green (L90) for standard temperature nudging, and orange (L31) and light blue (L41) for mean temperature nudging. A dataset with higher horizontal resolution (T63) is represented in purple.

global temperature offset can still develop. In contrast, the MTN setup , the bias imposes an additional constraint on the global mean temperature, preventing such a systematic bias from emerging. As a result, MTN keeps the model closely aligned with the reanalysis, making the temperature bias almost negligible. In the MTN setup, the bias is only present above 80 hPa, in a

335 region where no nudging is applied. The impact of both nudging concepts on the temperature difference relative to ERA5 is also evident in the Asian region (see Figure S2a and S4aS1 and S11) and across the entire Northern Hemisphere (see Figure S2b and S4bS2 and S12). Small differences can be found in the Asian region above 80 hPa, where simulations (except L90) exhibit a warm bias (up to 1 K), regardless of the applied nudging method.

The analysis of relative humidity over ice in the extended NAFC during March-April 2014, presented in Figure 3, reveals
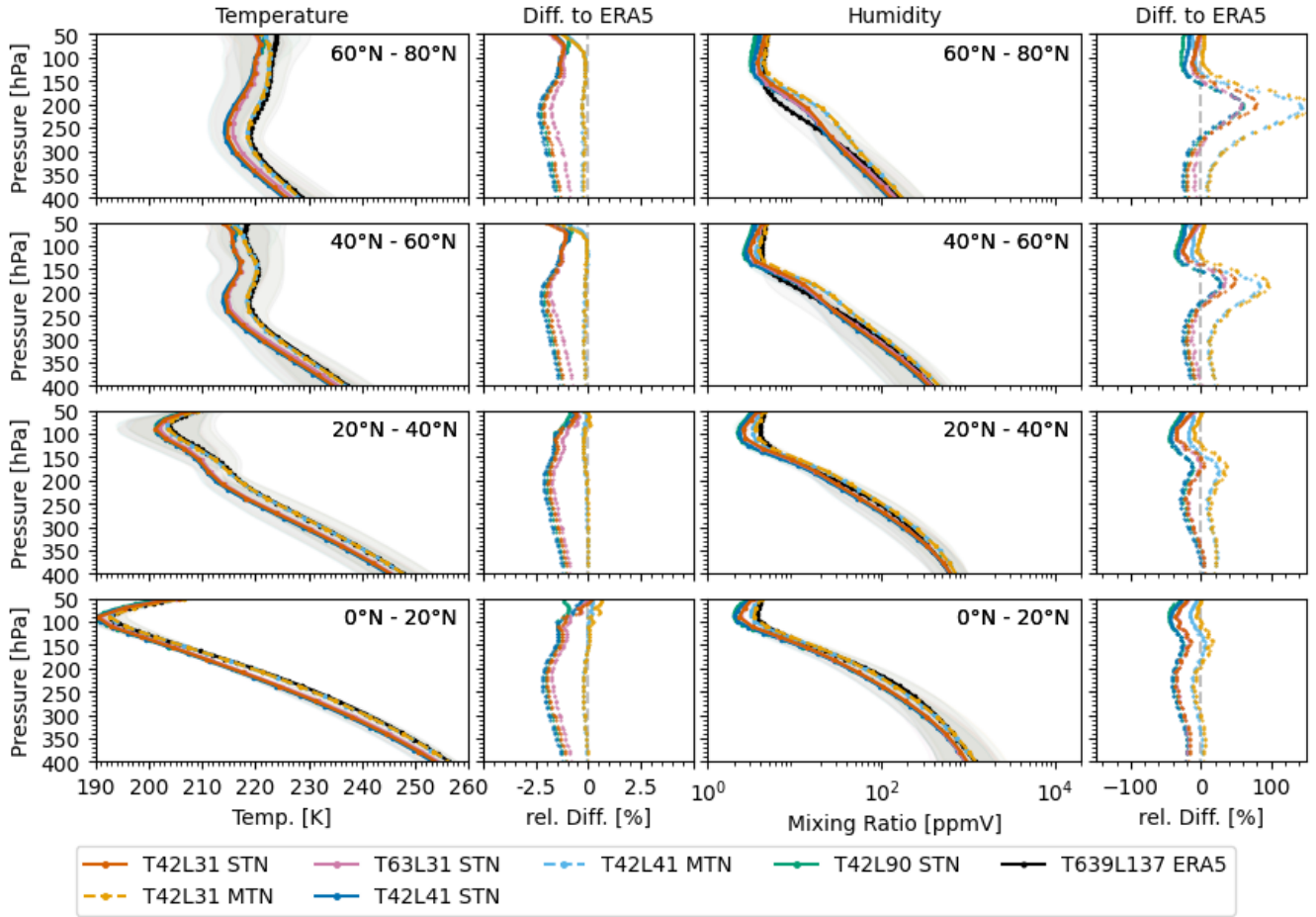
Figure 4. Temperature and humidity distribution from various EMAC simulations with different vertical resolution (L31:red, L41:blue, L90:green) and ERA5 data with 137 vertical level (black) for the North Atlantic Flight Corridor. Nudging approaches: standard (STN, *solid*) and mean (MTN, *dashed*) temperature nudging. Left: Area weighted mean temperature distribution and difference to ERA5 for all EMAC setups. Right: Area weighted mean humidity mixing ratio distribution and difference to ERA5 for all EMAC setups. The mean values for each grid box were calculated beforehand by using hourly data spanning from March to April 2014. The 95% percentile of each simulation is marked as grey shade.

340   up to 30% higher mean values at each EMAC grid ~~point~~ points compared to ERA5 across four pressure levels (200 hPa, 250 hPa, 300 hPa, and 350 hPa). The MTN simulations, on average, exhibit closer agreement with ERA5 than the STN simulations, with noticeable differences across pressure levels. ~~The most substantial differences between both nudging methods are observed at~~ At 200 hPa, ~~where MTN simulations demonstrate the best agreement with ERA5, deviating by an average of 7.5 to 10~~MTN simulations differ typically by 5-12%, while ~~the difference to the STN simulations shows a peak between~~

345   ~~12.5 to 15%absolute difference~~STN differ typically by 12-17 %. At 250 hPa and 300 hPa, differences between all STN model

setups are further reduced, while the difference between EMAC and ERA5 remains consistent on average. ~~Within each nudging group, variations between simulations with different model levels are minimal~~Analysis shows that the vertical resolution, does not play a central role, but increasing the spectral (horizontal) resolution to T63 ~~significantly~~ diminishes the difference to ERA5 across all ~~analyzed~~ investigated pressure levels.

350 The variation in distributions between simulations with mean and standard temperature nudging is smaller for the NAFC region than for the entire Northern Hemisphere. In the Northern Hemisphere (Figure ~~S3b~~S7), a larger spread in the relative humidity over ice difference between EMAC and ERA5 can be observed for all ~~analyzed~~ analysed simulations and pressure levels, with certain EMAC grid points indicating lower relative humidity than ERA5. In ~~model setups with standard temperature nudging, the ERA5-EMAC difference increases with altitude, ranging from 12% average difference at 350 to 18% at 200 . In contrast,~~
355 ~~MTN simulations consistently have their peak values between 7.5 and 15% difference across all pressure levels. In the~~ the Asian region (Figure ~~S3a~~S6), distributions at 250 hPa and 300 hPa resemble those of the NAFC region and the Northern Hemisphere, while at 200 hPa substantial disparities from ERA5 are evident. ~~STN simulations exhibit an average absolute difference of 22% in relative humidity~~ When looking at PDFs of the ERA5 and EMAC model setups separately (Figure S15, S16, and S17), a shift toward higher humidity values in EMAC setups compared to ERA5 ~~, while the MTN simulations difference peaks at~~
360 ~~around 19%. At 350 , STN simulations perform even 1%-point better than their MTN counterpart.~~ is noticeable at all pressure levels. The altitude-dependent cold bias can partly explain the disparity in relative humidity between ERA5 and EMAC ~~for~~ in the STN simulations, as the saturation vapour pressure over ice depends solely on temperature (A3). However, an analysis of specific humidity values is necessary to understand discrepancies in simulations involving mean temperature nudging.

Subsequently, we examined the vertical distribution of specific humidity along $20°$ latitude bands for the three selected ar-
365 eas. In the extended NAFC (Figure 4), an altitude and latitude-dependent humidity bias that differs with the nudging concept is apparent. For MTN simulations, a consistent wet bias is observed at higher altitudes (150 hPa to 250 hPa), significantly differing from ERA5. While differences are minimal in the tropics, a substantial bias exists in the mid-latitudes ($40°$ to $80°$). This results in higher relative humidity over ice values for MTN simulations compared to ERA5, despite similar temperatures. In contrast, simulations driven by STN exhibit a dry bias below 250 hPa for all latitude bands compared to ERA5. Additional-
370 ly, they display a wet bias north of $40°$ at heights between 150 hPa and 250 hPa. The dry bias at lower levels (350 hPa) compensates for the slight temperature bias, yielding relative humidity over ice values closer to ERA5 when compared to other pressure levels (Figure 3). At 300 hPa, STN simulations reveal a temperature bias increase, while the dry bias remains unchanged, resulting in a minimal rise in relative humidity over ice difference to ERA5. The temperature bias in STN simula- tions peaks at 250-200 hPa, accompanied by a wet bias intensifying at higher latitudes. However, the overall bias is reduced
375 on average due to lower values at lower latitudes. No significant differences are observed between simulations with different vertical resolutions. In the other investigated regions, similar results are found. It is evident that a substantial bias in specific humidity exists in the MTN setup, particularly at higher altitudes and northern latitudes. Nevertheless, the effect on relative humidity weakens at higher altitudes because of the pressure-level-dependent vapour pressure. Higher relative humidity values have a direct impact on the prediction of ISSRs, as exceeding the threshold values more frequently leads to an overestimation
380 of regions where contrails can persist. Similarly, the temperature bias affects contrail formation predictions; a systematic cold

bias in the temperature data shifts the perceived temperature closer to or below the critical threshold in the SAC. This results in contrails being predicted in conditions where, in reality, the ambient temperature is warmer than the critical threshold and contrails would not actually form.

3.2 Contrail forming areas in different model setups

385 Our numerical simulations utilizing seven different model setups reveal differences in horizontal size (area) of regions where contrails could form (potential contrail cover) or persist. With both methods outlined in Section 2.2, we first ~~analyze~~ analyse contrail coverage areas for thresholds at 90%, 95%, and 100% RH_{ice} below 235 K at three distinct pressure levels (see Table 2) in a specific synoptic situation (26 March 2014, 12:00 UTC), where research aircraft observations from the ML-CIRRUS campaign are available (see Section 2.4). Reducing the humidity threshold from 100% to 95% and to 90% results in an increase in

390 contrail coverage areas by a factor of about 2 and 2.5, respectively. Due to the differences in temperature and humidity between the model and ERA5 reanalysis data, all EMAC setups generally exhibited ~~a larger contrail formation~~ larger ice-supersaturated region compared to ERA5. Simulations with standard temperature nudging, irrespective of vertical resolution, indicate larger contrail coverage areas for each pressure level and RH_{ice} threshold. Transition to a T63 resolution yielded a marginal decrease in coverage of approximately 1%-point. At 250 hPa, MTN simulations display nearly identical covered areas as ERA5 for the

395 90% (15-16% coverage) and 95% (11-13% coverage) thresholds. In contrast, STN simulations show 10% (> RH_{ice} 90%) to 8% (> RH_{ice} 95%) more potential contrail formation areas at this pressure level. At 300 hPa and 350 hPa, MTN simulations overestimate ERA5 results by 5% to 7%, while STN simulations overestimate by 7-16%. However, for the 100% threshold, EMAC MTN results agree with ERA5 at 300 hPa, while they underestimate at 250 hPa by 3% and overestimate at 350 hPa by 3-4%.

400 The vertical temperature profiles from EMAC and ERA5 on the specified day (see Figure ~~S2e~~S3) display a model cold bias of similar magnitude (3-5 K, peaking around 200-250 hPa) as derived from monthly mean values (Figure 4). While specific humidity differences exhibit similar patterns, relative humidity distributions show even larger discrepancy, including cases where the RH_{ice} of ERA5 exceeds that of EMAC across all setups (see Figure ~~S3e~~S8). The findings summarized in Table 2 ~~suggest that both ,~~ indicate that both the cold bias in the standard ~~temperature nudged simulations ,~~ temperature-nudged simulations and the

405 overestimation of humidity in the ~~simulations including mean temperature nudging in EMAC, significantly affect the estimated areas of ice-supersaturated regions . While the selection of the~~ MTN simulations significantly influence the estimated extent of ISSRs. These effects can primarily be attributed to their impact on the calculation of RH_{ice}. First, the saturation vapour pressure over ice is highly temperature-dependent, decreasing at lower temperatures. This means that colder temperatures result in a lower saturation vapour pressure, requiring less water vapour for the air to reach saturation. As a result, the cold bias leads to an

410 overestimation of regions where contrails can persist and grow, even when the actual RH_{ice} is below the necessary threshold. Second, a wet bias increases the actual vapour pressure, which also raises the calculated RH_{ice}, further exaggerating the size of ISSRs. While the choice of nudging method has a significant impact on the results, the number of vertical model levels only affects ~~the results~~ ISSR predictions at specific pressure levels, highlighting that vertical resolution plays a less critical role in these simulations.

The influence of ~~different vertical model resolutions~~ vertical model resolution and nudging methods on the ~~second contrail~~

Table 2. Percentage of areas indicating contrail formation based on the ISSR and PotCov methods at various pressure levels, using different model setups over the NAFC on ~~on~~ 26 March 2014, at 12:00 UTC. For the ISSR method, three relative humidity over ice (RH_{ice}) thresholds (90%, 95%, and 100%) are used. Potential contrail cover (PotCov) areas are calculated directly in EMAC and are provided as total (PotCov > 0) and grid box-adjusted ~~(GBA)~~ coverage areas. The ~~GBA~~ grid box-adjusted coverage is determined by multiplying each grid box area by the potential contrail cover for that box, then dividing the area-weighted sum of all covered areas by the area-weighted sum of the entire area. The area-weighted mean ~~potential contrail coverage (AWM~~ PotCov ~~)~~ is calculated over all grid boxes with ~~potential contrail coverage~~ PotCov above zero.

Region: 120°W 20°E 80°N 20°N	T42L31 STN	T42L31 MTN	T63L31 STN	T42L41 STN	T42L41 MTN	T42L90 STN	ERA5
Height: 250 hPa							
Cov. Area (ISSR)							
RH_{ice} > 90 / 95 / 100 [%]	25 / 19 / 7	15 / 11 / 4	23 / 19 / 8	25 / 21 / 7	16 / 13 / 4	26 / 21 / 9	16 / 13 / 7
Cov. Area (PotCov)							
~~GBA~~ adjusted (total) [%]	24 (62)	22 (59)	22 (59)	23 (58)	21 (53)	22 (60)	
	0.36	0.35	0.35	0.38	0.38	0.35	
~~AWM~~ Area-weighted Pot-Cov [frac]							
Height: 300 hPa							
Cov. Area (ISSR)							
RH_{ice} > 90 / 95 / 100 [%]	31 / 25 / 12	22 / 17 / 7	29 / 23 / 10	32 / 27 / 12	22 / 18 / 8	31 / 26 / 13	15 / 13 / 7
Cov. Area (PotCov)							
~~GBA~~ adjusted (total) [%]	20 (62)	19 (62)	19 (59)	22 (57)	19 (56)	22 (60)	
	0.34	0.32	0.32	0.39	0.35	0.37	
~~AWM~~ Area-weighted Pot-Cov [frac]							
Height: 350 hPa							
Cov. Area (ISSR)							
RH_{ice} > 90 / 95 / 100 [%]	20 / 17 / 10	15 / 13 / 8	19 / 15 / 9	21 / 18 / 10	15 / 13 / 7	21 / 17 / 10	9 / 8 / 4
Cov. Area (PotCov)							
~~GBA~~ adjusted (total) [%]	14 (44)	12 (41)	13 (44)	15 (47)	12 (43)	14 (47)	
	0.34	0.33	0.32	0.34	0.32	0.33	
~~AWM~~ Area-weighted Pot-Cov [frac]							

415

~~formation area identification approach , as described in Section 2.2, is also presented~~ PotCov approach is also visible in Ta-

ble 2 for the NAFC. While ~~the~~ area-weighted mean ~~potential contrail coverage (AWM PotCov) for grid boxes with a potential contrail coverage above zero~~ PotCov remains consistent across ~~all model setups~~ models (0.36 ± 0.02 at 250 hPa, 0.35 ± 0.04 at 300 hPa, 0.34 ± 0.03 at 350 hPa), ~~variations appear in the total region covered by these grid boxes. At 250~~total coverage varies. At 250-300 hPa~~and 300~~, simulations with ~~31 and 90 vertical levels show a total coverage between 59 and 62%, while the setup with 41 levels shows coverage between 58 and 53%, with the MTN setup being lower than the STN setup~~31-41 levels show coverage between 59-53%. At 350 ~~-~~hPa, both MTN setups ~~show the lowest coverage (41%), while~~ have lower coverage than the STN setups~~range from 44 to 47%~~. Summing the geographic grid box areas of those with ~~potential contrail coverage~~ PotCov greater than zero results is an overestimation of the areas covered with contrails. Therefore, we adjusted the covered area in each grid box with the respective grid box fraction, reducing the potential contrail areas by nearly 60%. Since this method employs the thermodynamic Schmidt-Appleman criterion, temperature and humidity ~~thresholds~~ play a significant role. Variations in these critical parameters account for differences in the resulting covered areas. Furthermore, the vertical resolution appears to have a smaller impact on the areas compared to the previous method. Since the potential contrail parameter is specific to EMAC, direct comparison with ERA5 is not possible. Therefore, we conducted an inter-comparison between both approaches.

The comparison between contrail formation areas based on the ISSR method and the PotCoV method reveals the best agreement at the 95% RH_{ice} threshold, although discrepancies persist across different pressure levels. Standard nudged simulations show the closest agreement between both methods at 250 hPa, while simulations with mean temperature nudging perform best at 350 hPa. Specifically, at 250 hPa, MTN underestimates, while at 350 hPa, STN overestimates the potential contrail areas by 5-10%. Compared to ERA5, the EMAC PotCov method consistently overestimates the covered areas for all pressure levels and model setups, with discrepancies ranging from 1-12%. Our findings suggest that the PotCov approach for mean temperature nudged EMAC data yields the best agreement with ERA5 ISSR areas. However, even with this approach, EMAC simulations result in larger contrail areas compared to ERA5. Furthermore, aside from the discrepancies caused by different nudging approaches, the number of model levels appears to have a negligible impact.

4   Comparison with observations along trajectories (ML-CIRRUS Campaign)

In general, the climate impacts of non-CO_2 effects are complex and require detailed knowledge about key atmospheric parameters. Nevertheless, numerous difficulties may arise when modelling parameters such as temperature or humidity using a general circulation model (Jöckel et al., 2016). Consequently, it is advisable to utilize atmospheric measurements to assess model performance in particular regions. Hence, we compared temperature and humidity data on aircraft trajectories obtained during the ML-CIRRUS campaign with the emulated trajectory data from different model setups. It is important to note that the ML-CIRRUS campaign provides in-situ measurements along specific aircraft trajectories, representing localized observations, whereas the model results are grid-box mean values, which represent average conditions over larger spatial areas. To partly account for these differences, we averaged the ML-CIRRUS data over 10-minute intervals, aligning with the model output frequency, thereby smoothing the high-frequency variability of the in-situ data. This approach allows a more consistent

450 comparison, though some discrepancies may still arise due to the inherent spatial and temporal scale differences. From the measurement data, we derive potential contrail formation conditions ~~e.g. ice-supersaturation,~~ (e.g., ice-supersaturation) that can be compared with EMAC model potential contrail areas and validated with satellite imagery.

Temperature comparison between EMAC and ML-CIRRUS measurements
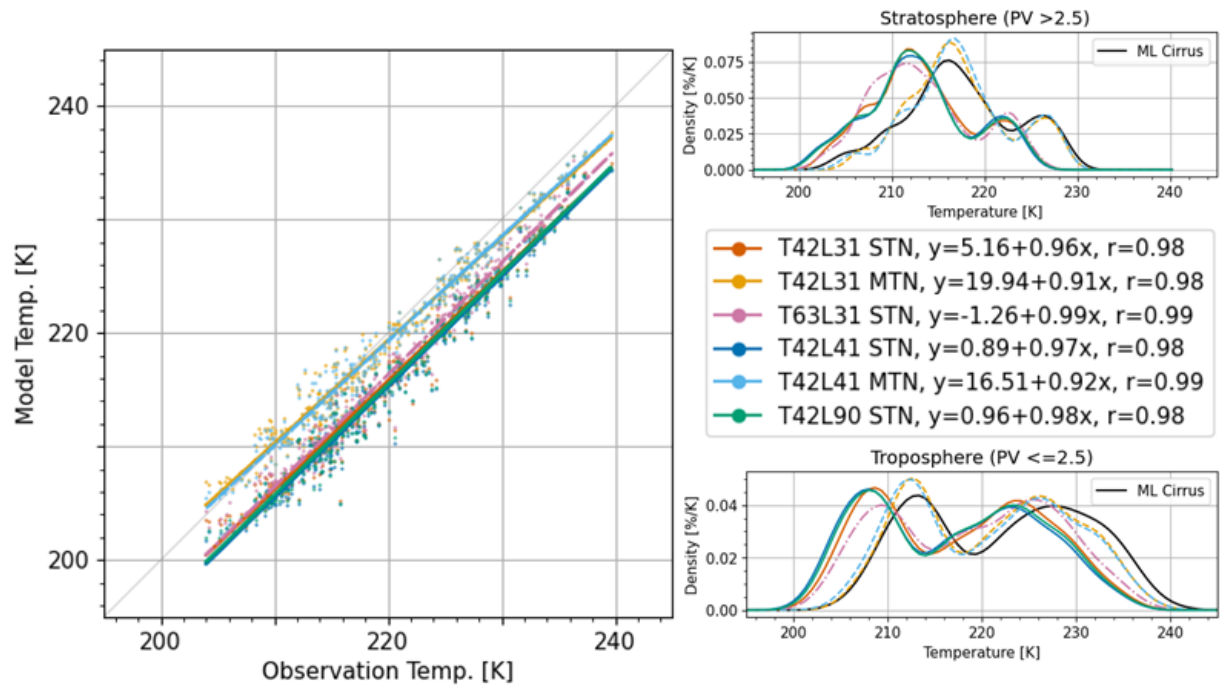
Figure 5. Left: Correlation between temperature data measured ~~by~~ during the ML-CIRRUS campaign and simulated ~~with~~ using various EMAC model setups at different points along the flight path. Next to the color bar, the intercept and slope of the linear regression line for each setup are listed, along with the Pearson correlation coefficient (r). Right: Probability Density Function (PDF) of tropospheric temperature data for regions with potential vorticity (PV) $\leq$ 2.5 (top) and stratospheric temperature data for PV > 2.5 (bottom). In addition to the different model setups, resampled ML-CIRRUS measurement data (10 min intervals) is shown in black in the PDFs.

All model setups exhibit a ~~satisfying correlation with the temperature data collected during~~ high degree of agreement with
455 the observed in situ temperature data from the ML-CIRRUS ~~campaign, as illustrated in~~ flights (Fig. 5. ~~This is supported by a strong correlation coefficient~~). To assess this agreement, we first applied spatial resampling by extracting EMAC model data along the aircraft trajectories, followed by temporal resampling of the high-resolution aircraft data to the model's time steps. The resulting time series were then compared, yielding Pearson correlation coefficients (r) between 0.98 and 0.99. This high correlation indicates an excellent correspondence between observed and modelled temperatures. The STN simulations
460 show lower temperatures than the corresponding observations, indicating a constant cold bias of up to 4 K between 200 K and 240 K. The simulations show a slope almost equal to one, consequently producing parallel fitting lines with the central line.

The MTN simulations display temperature pairs distributed around the central line with considerable variations up to 1 K for higher temperatures. Besides that, the 308 data pairs suggest small differences between the MTN setup and the ML-CIRRUS temperature data. Between the simulations with different vertical resolutions, no differences were found in the respective

465 nudging groups. Changing from T42 to T63 horizontal resolution only has a minimal impact and reduces the average bias by a maximum of 0.1 K. A probability density function assessment of only tropospheric and stratospheric data (see Sec 2.1) indicates the presence of a comparable cold bias in both pressure regions, which is significantly mitigated by utilizing the mean temperature nudging method in the simulation (~~Figure S5a and S5b~~Fig. 5). The differences between the MTN simulation results and the measurements can ~~mainly~~ be attributed to the resampling of ~~the~~ observation data and ~~the~~ uncertainties during the

470 measurement process. Resampling aligns observational points with model output intervals, which can lead to data smoothing and loss of fine-scale information, especially in rapidly changing conditions. Additionally, measurement uncertainties - such as sensor calibration, environmental factors, and instrumental limitations - impact the precision of observed humidity and temperature values, further contributing to the discrepancies.

Humidity in EMAC compared to ML-CIRRUS measurements
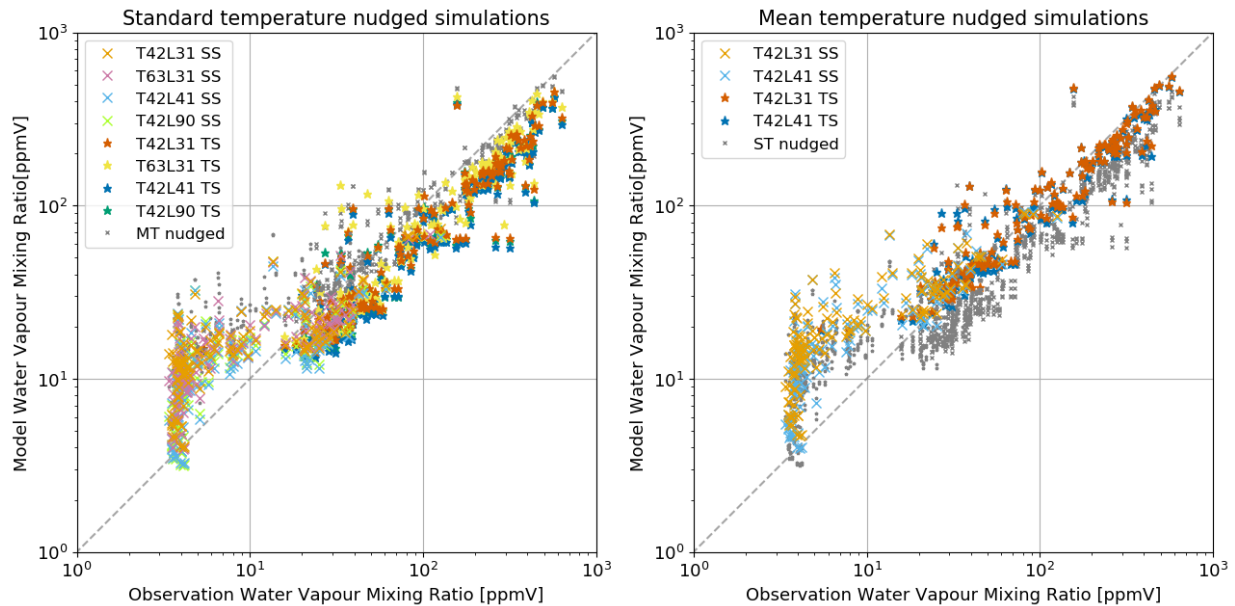
Figure 6. Correlation between water ~~vapor~~ vapour mixing ratio measurements from the ML-CIRRUS campaign and results obtained from different EMAC setups. Standard temperature nudged simulations are colored on the left side, while simulations with mean temperature nudging are shown in grey. On the right side, this arrangement is reversed. The respective vertical model resolutions are indicated in red (L31), blue (L41), and green (L90). Stratospheric (SS, crosses) and tropospheric (TS, stars) values are distinguished, with stratospheric values depicted in lighter shades.

475 The analysis of water vapour mixing ratios across various model setups and measurement data, illustrated in Figure 6, reveals ~~a substantial overprediction of low humidity values in the model compared to the observations. These low humidity values are located in the lower stratosphere and the values are up to six times higher across all model setups. This model "wet bias" for values between 5 to 10 of the observed water vapour mixing ratio appears with both nudging methods.~~ notable discrepancies between model results and observations. For high levels of humidity ranging between 100 and 800 $ppmV$, predominantly

480 present in the upper troposphere, the results from the standard temperature nudging simulations persistently provide lower values than the HALO FISH and SHARC measurements (see Figure 6, left). This "dry bias" is strongly reduced in simulations with mean temperature nudging (see Figure 6, right). This could be directly connected to the temperature bias reduction found in Figure 2, as temperature influences humidity. In contrast, for low humidity values between 5 to 10 $ppmV$, which are located in the lower stratosphere, the model consistently overpredicts compared to the observations by up to six times across all model

485 setups. Nevertheless, these low humidity values are not relevant for contrail formation. This model "wet bias" appears with both nudging methods. Overall, the measurements indicate a higher number of data points with water vapour mixing ratio values smaller than 5 $ppmV$ compared to the model. The selection of vertical resolution only has a minor impact on the strength of the correlation. In the lower stratosphere, data from T42L41 show the strongest agreement with the ML-CIRRUS data. This results from the fact that the T42L41 model setup has the highest resolution around the tropopause (refer to Section 2.1) in comparison

490 to the other EMAC model setups, which reduces the water vapour diffusion from the troposphere into the stratosphere. For the troposphere, all simulations produce a similar outcome.

Relative humidity conditions for persistent contrails

We ~~analyze~~ analyse the point-by-point correlation between measurement data and different model setups for ~~contrails that~~ atmospheric conditions when contrails can persist longer than a few minutes and significantly impact the climate, following

495 the approach of Gierens et al. (2020). We ~~focus on~~ use the approach discussed in Section 2.2 to identify ice-supersaturated ~~conditions, where the ambient relative humidity with respect to ice (RH$_{ice}$) exceeds 100%~~regions. To achieve this, we compared RH$_{ice}$ values derived from FISH and SHARC water ~~vapor~~ vapour mixing ratio and temperature measurements (using equations A1-A3, see Appendix) from the ML-CIRRUS campaign with corresponding data from EMAC. ~~Following Dietmüller et al. (2023), we gradually lower the relative humidity threshold for the model data~~ We explore relative humidity

500 thresholds (Dietmüller et al., 2023) by stepwise evaluating categorical statistical measures for model thresholds ranging from 100 to 90% ~~.~~ (Fig. 7). This comparison (refer to Table 3 for the 90% threshold) presents values of a 2×2 contingency table with four categories: Y/Y (hits), where both ML-CIRRUS and EMAC data indicate conditions that support contrail persistence; Y/N (misses), where ML-CIRRUS measurements predict contrail persistence but the model does not; N/Y (false alarms), where ML-CIRRUS measurements do not predict contrail persistence but the model does; and N/N (correct negatives), where both

505 ML-CIRRUS and EMAC data show that contrail persistence is not possible.

 Table 3 reveals that for the selected model RH$_{ice}$ threshold (90%), the T42L90 standard nudged model setup achieves the highest hit rate (HR) at 83%, indicating a strong ability to correctly identify contrail conditions. All standard nudged simulations show high hit rates between 77% and 83%, while the mean temperature nudged simulations only show a hit rate of

Table 3. Comparison of measured relative humidity with respect to ice derived from ML-CIRRUS data (threshold 100%) and corresponding data from different EMAC setups (threshold 90%). The contingency table shows the correlation between the two data sets, including hits (Y/Y), misses (Y/N), false alarms (N/Y), and correct negatives (N/N). The ~~Equitable Threat Score (ETS) characterizes the agreement between the data sets, considering hits, misses, and false alarms while adjusting for random chance. The~~ hit rate (HR), expressed as a percentage, represents the proportion of contrail conditions correctly identified by the model (Y/Y) relative to the total measured contrail conditions (Y/Y + Y/N). Conversely, the false alarm rate (FAR), also expressed as a percentage, measures the proportion of incorrectly identified contrail conditions by the model (N/Y) relative to the total measured non-contrail conditions (N/Y + N/N). The Equitable Threat Score (ETS) is a commonly used verification metric that evaluates how well two datasets align by considering not only correct predictions (hits) but also incorrect ones (misses and false alarms). It adjusts for agreements that could occur by chance, providing a fairer measure of predictive skill (ref to Ebert (1996)). The values HR and FA are illustrated in Figure 7, together with different threshold between 90 and 100%.

Model setup	Size	Y/Y	Y/N	N/Y	N/N	ETS	HR [%]	FAR [%]
T42L31 STN	286	23	7	106	150	0.08	77	41
T42L31 MTN	286	21	9	91	165	0.08	70	36
T63L31 STN	286	24	6	106	150	0.08	80	41
T42L41 STN	286	24	6	108	148	0.08	80	42
T42L41 MTN	286	21	9	99	157	0.07	70	39
T42L90 STN	286	25	5	111	145	0.08	83	43

70%. Conversely, the MTN simulations exhibit fewer false positives. The T42L31 MTN model achieves a false alarm rate (FAR) of 36%, the lowest in the table, suggesting the fewest incorrect contrail condition identifications. In over one-third of all cases, the model predicts contrail conditions at this threshold that are not confirmed by measurements, irrespective of the vertical resolution or the nudging approach. The calculations were also performed for other RH_{ice} thresholds (Fig. 7). Here, the trade-off between sensitivity (hit rate) and specificity (false alarm rate) for different model setups and thresholds becomes apparent. Lower thresholds generally result in higher hit rates and false alarm rates. A false alarm rate below 20% could only be achieved for RH_{ice} thresholds over 97%, with a corresponding hit rate below 50%. Aiming for high hit rates above 80% increases consequently the false alarm rate to at least 35%, for RH_{ice} thresholds of 94%.

At high thresholds, the T42L31 MTN setup performs the best. At lower thresholds (90-95%), the T63L31 and T41L90 setups consistently show higher hit rates than other setups, but they also exhibit higher false alarm rates. The T42L41 setups with standard and mean temperature nudging demonstrate consistent performance with relatively high hit rates and moderate false alarm rates, making them the most balanced model setups. In addition, we also use the Equitable Threat Score (ETS) in Table 3 to characterize the agreement between the data sets. The ETS considers hits, misses, and false alarms while adjusting for random chance, providing a more equitable assessment of forecast performance. The values can be interpreted as follows: ETS = 1 indicates a perfect forecast (all hits, no misses or false alarms), ETS = 0 indicates the forecast is no better than random chance, and ETS < 0 indicates the forecast is worse than random chance. For a detailed explanation of the ETS, refer to Gierens et al. (2020). In our analysis, the ETS values range from 0.07 to 0.08, indicating a low degree of agreement between the two
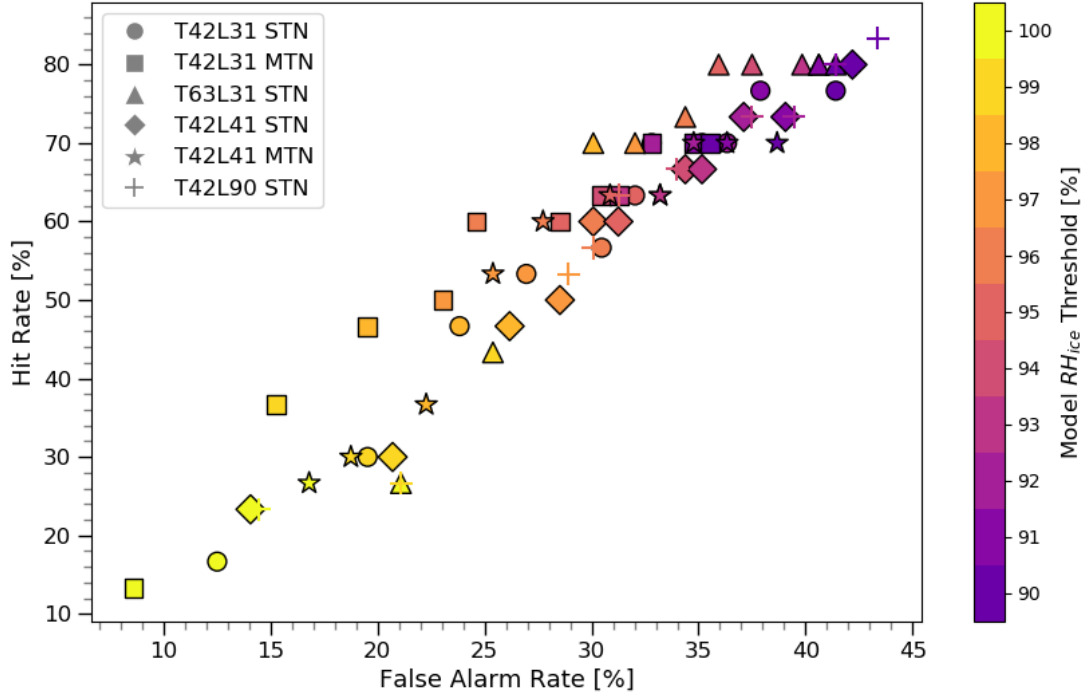
Figure 7. Scatter plot illustrating the relationship between the false alarm rate and hit rate across various model setups and RH_{ice} thresholds. Each marker type represents a different dataset: Circles (T42L31 STN), Squares (T42L31 MTN), Triangles (T63L31 STN), Diamonds (T42L41 STN), Stars (T42L41 MTN), and Crosses (T42L90 STN). The color of the markers corresponds to the model RH_{ice} threshold values, with the gradient ranging from purple (90%) to yellow (100%).

data sets. This suggests that the relationship between the data sets is mostly random and that this statistical analysis does not confirm a clear correlation, which might be due to the small number of data points. By reducing in our analysis both the model and observational relative humidity thresholds below 90%, more data points are counted as contrail-forming conditions. This increases the number of hits and improves the ETS score to approximately 0.5. However, this artificial enhancement in agreement compromises physical realism, as contrail formation typically requires conditions near full saturation.

530

5 Case study, ML-CIRRUS flight on 26 March 2014

The analysis of all ML-CIRRUS flights provide important information regarding distinctions between model results and observations. However, these findings do not provide insight into the rapid and abrupt changes in flight level and the subsequent alterations in atmospheric conditions. Additionally, the weather conditions must be considered. Thus, we are monitoring the

535   trajectory of a specific flight to acquire a comprehensive understanding. As a case study, we selected 26 March ~~2014 has been chosen. This day displays intriguing meteorological conditions with~~ 2014. Although a polar cold front outbreak ~~across Europe, causing a subsequent descent of the tropopause height.~~ occurred across Europe on this date, the region examined here remained largely unaffected by these unusual conditions, maintaining a typical tropopause height of around 200 hPa for this latitude. Moreover, the three flight dives performed on this day allowed us to systematically explore variations between

540   the troposphere and stratosphere and to assess how different dive strengths influenced the observed temperature and humidity profiles. Temperature, humidity, and ~~potential contrail coverage~~ PotCov are compared along the aircraft trajectory for all EMAC setups. Further, the projected areas of possible contrail coverage are compared with satellite imagery.

Weather situation during the flight

~~Following the methodology of REACT4C (Irvine et al., 2012), we have categorised the~~ The weather situation over the North

545   Atlantic on the 26 March ~~2014. Geopotential height anomalies at 250 indicate minimal correlation with both, the North Atlantic oscillation index, and the East Atlantic index. Moreover, the jet stream~~2014 shows a relatively weak jet stream, which is limited to the western North Atlantic~~and exhibits a relatively low intensity. According to Irvine et al. (2012) this results in type W5, a weak confined jet. This type is the most frequent type in winter time and occurs on average on 26 days per winter.~~. High pressure is positioned above Scandinavia, while low pressure dominates across Central Europe. Additionally, another low

550   pressure system is located over Southern France and the Mediterranean Sea. At the southern tip of Greenland, a major low pressure system can be observed, with a warm and a cold front attached, displaying distinct cloud bands across the North Atlantic in the MSG data. In the nudged EMAC simulation, the geopotential height at 500 hPa resembles the reanalysis data of the Deutscher Wetterdienst (Berliner Wetterkarte). ~~The potential contrail coverage~~ PotCov (Figure 8) is large where the warm and cold fronts are located, although the potential contrail ~~coverage~~ area (Frömming et al., 2014) is considerably more

555   pronounced around the warm front. These results are in line with the observations made by Kästner et al. (1999), who noted that contrail formation is more frequent before warm fronts, before cold fronts, and together with cirrus in a warm conveyor belt.

S4D sub-model analysis

A direct analysis along the trajectory provides insight into ~~parameter fluctuations~~ variations in specific atmospheric variables

560   that are not apparent in larger dataset statistics. Figure 9 illustrates a detailed comparison of the parameters throughout the flight. The analysis indicates that the ML-CIRRUS flight on the 26th of March 2014 was primarily located in the stratosphere with three descents into the troposphere at 08:30, 09:35, and 10:55 UTC. Each of these crossings of the tropopause lasted roughly 45 minutes and covered pressure differences from 20 to 100 hPa in depth (a). A significant temperature difference between the observations and the model results in both, the troposphere and the stratosphere, can be observed for the standard

565   nudged simulations. The discrepancy ranges from 3 to 5 K and is minimally affected by the aircraft's ascent or descent. For the simulation with mean temperature nudging (blue and orange dashed lines), this cold bias is not present (b). Regarding tropospheric humidity, lower values are found in the standard nudged simulations compared to the measurement data. However,
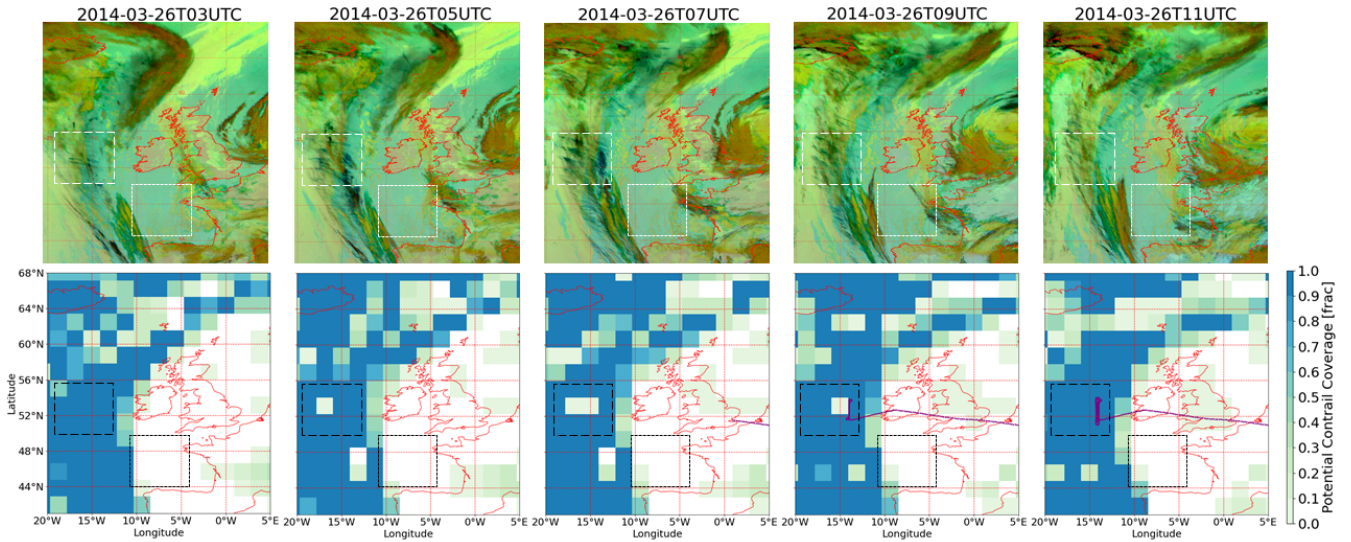
Figure 8. Top: The MSG satellite ash RGB from $20°$W to $5°$E and $68°$N to $44°$N for 26 March 2014, in 2-hour intervals from 3 UTC (left) to 11 UTC (right). Contrails are visible as dark lines. Bottom: The maximum potential contrail coverage between 300 and 200 hPa (3 model levels) derived from the EMAC T63L31 STN model setup. Dark blue indicates high potential grid box coverage, while light green indicates low coverage. The ML-CIRRUS flight track is shown as purple dots. The white and black squares indicate areas with high (–) and no (...) contrail coverage.

by applying the MTN method, this bias is largely reduced. In the stratosphere, the models humidity is up to six times larger than the measurements, regardless of the nudging method applied. This is particularly noticeable in the first quarter of the flight, but can also be seen whenever the aircraft has crossed the tropopause into the stratosphere, i.e. for low humidity values. During the final quarter of the flight, the difference between track pressure and tropopause height is sufficiently large, so that the water vapour mixing ratio values are very low, below 5 ppmV, and there are no significant differences between model results and observation anymore in terms of water vapour mixing ratio (c). All model setups indicate a relative humidity over ice of approximately 100% in the troposphere for the first two dives, which is consistent with the measurement results. However, during the third dive, the ~~observed RH$_{ice}$ are lower and the~~ model-predicted RH$_{ice}$ values exceed the observational data ~~in parts~~ by roughly 20%. ~~This suggests that the simulated gradient is not steep enough, as the flight is further away~~ Unlike the previous dives, the third dive occurs further from the tropopause~~than in the dives before~~, potentially impacting the accuracy of the model's representation of the relative humidity. In the stratosphere, due to the larger mixing ratio of water vapour visible in (c), the model simulates a relative humidity over ice between 10 to 40%, whereas the observational data are nearly zero in this area. Although this may not significantly impact contrail prediction, as it is far from complete saturation, it does require attention. When contrasting the various model setups, the L41 with mean temperature nudging comes closest to the measurement humidity data (c and d). Based on the temperature and humidity data, contrail formation is expected for the first 2 dives into the troposphere, and partly during the last dive. All model setups predict contrail formation in the troposphere,
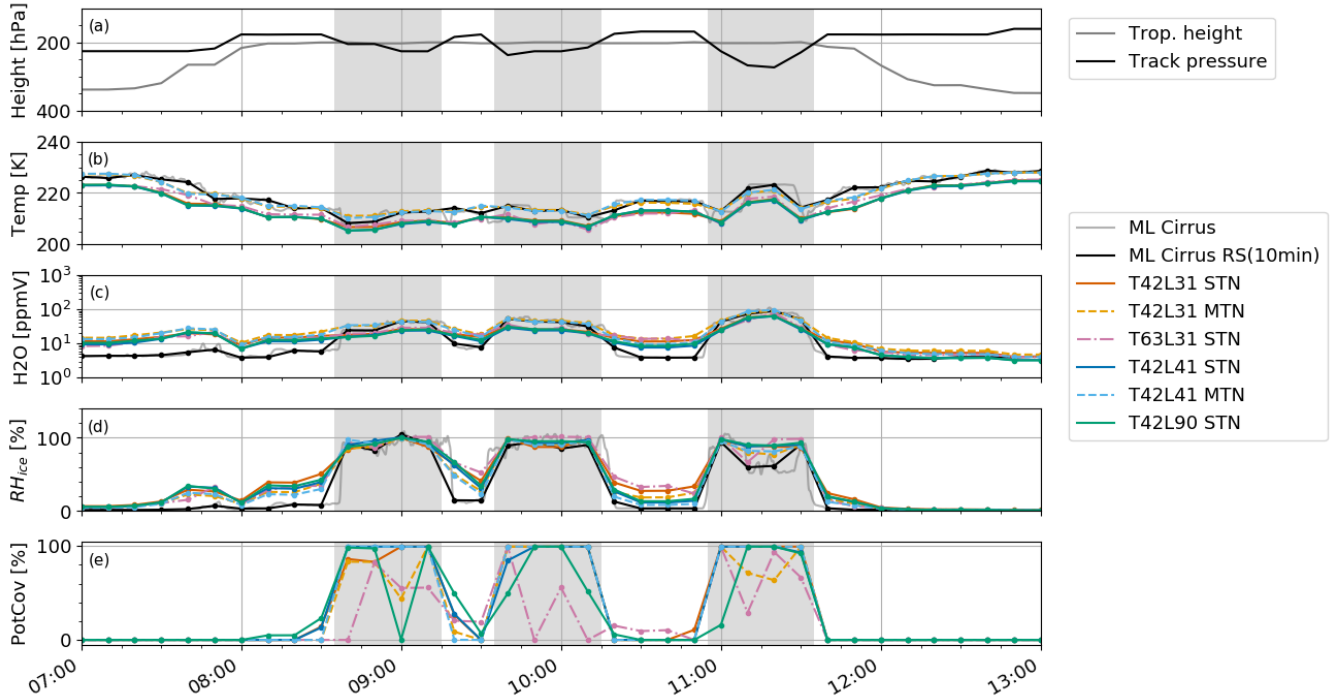
Figure 9. Comparison of various atmospheric parameters along the flight trajectory for March 26 2014. (a) Tropopause height and the track pressure height of the HALO aircraft. Comparison of temperature (b), water vapour mixing ratio (c) and relative humidity with respect to ice (d) between results obtained with different EMAC setups and ML-CIRRUS measurements (FISH and BAHAMAS). (e) Shows the predicted contrail coverage in EMAC for the according grid box.

~~some also~~ with some showing a low probability in the lower stratosphere ~~close to the tropopause~~ near the tropopause (e). The

585 simulation with the finest spectral (horizontal) but lowest vertical resolution, T63L31, predicts the lowest fraction of possible contrail formation in the ~~grid boxes in the troposphere~~ tropospheric grid boxes. Applying mean temperature nudging decreases the likelihood of contrail formation for L31 but has no effect ~~for~~ on L41~~. However, the~~, and the overall impact of nudging on ~~potential contrail cover~~ PotCov for this specific day is minimal. Differences in PotCov between higher and lower resolution setups are partly a consequence of the underlying "contrail cover concept." This parametrization, designed to represent partial

590 cloud cover within discrete model grid boxes, can lead to variability that does not solely reflect atmospheric conditions. Instead, it introduces a form of "noise" where changes in resolution and sampling can alter the fraction of the grid box potentially filled with contrails. From this initial analysis, it is not possible to definitively conclude the extent to which these differences stem from systematic resolution effects or from such parametric variability. The impact of the 10-minute output interval for the model data compared to the 1-second observational data is significant, as there is a clear delay in the model data after the aircraft's

595 altitude adjustment. Overall, differences in temperature and humidity, which were also observed in the previous sections, can be determined, if only this particular day is analysed. Comparison along the trajectory provides insight into the significance of the

model output and corresponding resampling interval. It should be noted that the models ability to capture all features observed during the steep dives is limited, if the model output interval is too coarse. There is a high level of correspondence between the predicted contrail formation produced by the model and the measured contrail conditions. Between the model configurations,

600   both L41 versions, with or without mean temperature nudging, demonstrate promising results regarding contrail prediction.

Potential contrail coverage and satellite observations

~~Top: The MSG satellite ash RGB from 20°W to 5°E and 68°N to 44°N for 26 March 2014, in 2-hour intervals from 3 UTC (left) to 11 UTC (right). Contrails are visible as dark lines. Bottom: The maximum potential contrail coverage between 300 and 200 hPa (3 model levels) derived from the EMAC L31T63 STN model setup. Dark blue indicates high potential grid box~~

605   ~~coverage, while light green indicates low coverage. The white and black squares indicate areas with high (–) and no (...) contrail coverage.~~

The satellite image (top panel in Fig. 8) shows a ridge cloud on 26 March 2014 composed of a cloud band extending from Iceland southwards passing West of Ireland and Great Britain. Contrail lines are best observed in correspondence to the cloud band over a wide range of latitudes starting East of Iceland / North of Scotland and ending East of Portugal, with many contrails

610   East of Scotland and Ireland and less contrails South of Ireland. The spread of these contrail lines is strongly influenced by large-scale synoptic weather patterns. Although the areas where contrails were observed remained consistent throughout the observed time span of 8 hours, a notable increase in intensity occurred at 05 UTC and 07 UTC, coinciding with the period when air traffic from the US to Europe is passing these regions. Of course, contrails can only form where aircraft fly. The maximum ~~potential contrail coverage in the L31T63~~ PotCov in the T63L31 STN model indicates the highest value observed between

615   300 and 200 hPa (bottom panel in Fig. 8). This analysis reveals that the distribution of ~~potential contrail coverage~~ PotCov follows the distribution of the ridge cloud in the ash RGB. EMAC indicates the ridge cloud as location of potential contrail formation, which is plausible in comparison to the satellite picture. Additionally, areas where no contrails are predicted, such as south of Ireland, correspond to regions with no observed satellite contrail coverage. As the ML-CIRRUS campaign conducted measurements east of Ireland on this day, the observed area of contrail formation can also be linked to high relative humidity

620   over ice values. Further regions of high potential contrail ~~coverage~~areas, e.g. over South Scandinavia and North of Corsica, cannot be evaluated with these satellite observations, because no contrails are observed here. The region East of England with high potential cloud coverage is also difficult to evaluate, because the few dark lines in the ash RGB in this region (see e.g. 05 UTC) cannot be unambiguously attributed to contrails.

6   Discussion

625   ~~Temperature~~ The results presented here highlight the strong influence of temperature and relative humidity ~~are key parameters for determining~~ fields on contrail formation and ~~evolution in numerical modeling studies. This study compares atmospheric temperature and humidity fields from various EMAC model setups with observational data obtained from measurements along aircraft trajectories. We quantify and compare the areas of the atmosphere conducive to the formation of persistent contrails~~

across different model setups and reanalysis data. Additionally, we explore the differences arising from varying humidity thresholds and methodologies. Our study evaluates key meteorological fields comprising temperature in distinct EMAC model setups, aiming to determine representation within the EMAC T42L41 setup (Frömming et al., 2021) with 41 vertical layers and 2.8° horizontal resolution in the light of evaluating the ability of the model to represent contrail conditions in the Northern Hemispheric extra-tropics persistence. Accurate representation of these parameters is crucial for realistic contrail modelling. Our analysis reveals a persistent cold bias of 3 to ~5 K cold bias for the in standard nudged EMAC model temperature compares to both, observations and other models, across the 100-400 range. In the past, setups across various vertical and horizontal resolutions and geographic regions. This bias likely stems from known limitations in general circulation models with a model top at 10 have often been associated with a cold bias problem (Stenke et al., 2007), which we confirm with our analysis. For our study, we find a consistent cold bias in all EMAC simulations, irrespective of vertical or horizontal resolution applied within the setup. Additionally, the bias is found to be equally prominent in the NAFC, Asia and the entire Northern Hemisphere without any geographical distinction based on latitude. The inclusion of global mean temperature ("wave zero") to the temperature , as previously noted by Stenke et al. (2007). Notably, the introduction of mean temperature nudging (MTN setup) significantly reduces the cold bias in both investigated setups (L31) techniques reduces this bias, L41). An extra simulation performed with ERA-Interim based mean temperature nudging for the T42L41 model setup produces nearly identical results to aligning the EMAC temperatures with ERA5 (results have not been visualized as they are nearly identical). In the comparison of this MTN setups with ERA5 reanalysis data, there is hardly any cold bias visible between 100 reanalysis data and 400 , independent of the studied region. These findings could have been expected, as the nudging in EMAC was directed towards ERA5; with the mean temperature nudging technique, temperature data is more aligned with ERA5 data. For the comparison with observations, we notice a good agreement between EMAC model results that utilized mean temperature nudging and the measurement results improving agreement with in situ measurements from the ML-CIRRUS campaign. These findings primarily indicate agreement between ERA5 reanalysis and measurements as expected. However, recent findings have provided new insights into the anti-ice (AI) correction for aircraft measurements (Giez et al., 2023), and indicate a possible cold bias in the temperature data which we want to discuss here. During the ML-CIRRUS campaign, no AI correction was applied to the measurement data, based on experimental evidence suggesting it was unnecessary. However, recent test campaigns have shown that this does not hold for all flight conditions. The correction is generally negligible, with a value of less than 0.1 across most flight conditions. However, it becomes significant at very high altitudes with low speeds, which are not typical flight conditions for the HALO aircraft. The application of the AI correction results in a slightly lower static air temperature, but this correction remains within the specified error margins for temperature measurements. Therefore, the temperatures calculated during the ML-CIRRUS campaign are likely unaffected by the new AI correction. The minimal improvements gained by increasing model horizontal resolution from T42 to T63, paired with substantially higher computational costs, suggest that simply refining spatial resolution does not address fundamental biases. Instead, targeted strategies such as MTN prove more effective in correcting systematic errors, thereby enhancing confidence in contrail predictions.

Similar to results from other studies, model temperature shows a smaller bias when mean temperature nudging is applied. However, such mean temperature nudging has potentially adverse effects on the radiation balance or the hydrological cycle

(Jöckel et al., 2016), as no recalibration was performed. This effect becomes apparent when ~~analyzing~~ analysing the specific humidity results. Humidity poses a challenge in climate modelling due to its spatial and temporal variability. It is noteworthy that specific humidity is overestimated in models such as the ECMWF's Integrated Forecasting System IFS, while humidity ~~oversaturation~~ over-saturation is underestimated in this model (Krüger et al., 2022; Woiwode et al., 2020; Reutter et al., 2020). The latter has a significant impact on contrail formation (Gierens et al., 2020). Overall, earlier studies have shown that the EMAC model underestimates water vapour due to its cold bias (Brinkop et al., 2015; Jöckel et al., 2016). Our simulations reveal similarly that water vapour mixing ratios are consistently underestimated between 50 and 400 hPa, except for the tropopause region at around 200 hPa, where the EMAC model actually overestimates water vapour mixing ratios compared to ERA5. The difference between the ERA5 and EMAC increases towards the poles, with the two EMAC setups that include mean temperature nudging displaying an even larger deviation from ERA5. An explanation for this is that there is a rise in the wet bias in the extra-tropical area, which occurs due to water vapour diffusing horizontally from the tropical upper troposphere into the extra-tropical lowermost stratosphere. This phenomenon was also noted by Stenke et al. (2007). While the exact cause of the observed wet bias for the MTN simulations at higher altitudes remains uncertain, it may be related to the wave-zero temperature nudging forcing the model away from its internal equilibrium, prompting compensatory adjustments in humidity fields. Further dedicated analyses or sensitivity simulations, beyond the scope of this study, would be needed to substantiate this explanation. The overestimation is also evident in the comparison with observational data. The difference is significant for lower water vapour mixing ratios, which are typically present close to the tropopause. Aircraft often cross this region without flying directly along the tropopause. This results in very small sample sizes, which, in combination with strong lapse rate changes, make it a difficult region to evaluate.

To evaluate model performance, we compared EMAC output with in situ measurements from the ML-CIRRUS campaign. Although a persistent temperature bias is noted in the model, the comparison is considered robust. Recent findings have provided new insights into the anti-ice (AI) correction for aircraft temperature measurements (Giez et al., 2023), indicating a possible cold bias in the temperature data, which we want to discuss here. The static air temperature on aircraft is determined using a Total Air Temperature (TAT) probe, which measures the temperature of the air impacted by the aircraft's motion. This probe is housed within a sensor inlet (e.g., Collins Aerospace 102BX) that is actively heated to prevent ice formation, a process known as Anti-icing. Ice accumulation on the inlet can lead to erroneous temperature measurements under in-flight icing conditions. The heating affects the temperature readings inside the housing, necessitating a correction known as the anti-ice (AI) correction. While the manufacturer provides a parametrized correction for this effect, the AI correction for the HALO aircraft was determined individually through in-flight calibration (Bange et al., 2013). During the ML-CIRRUS campaign, no AI correction was applied to the measurement data, based on experimental evidence suggesting it was unnecessary. However, recent test campaigns have shown that this does not hold for all flight conditions. The correction is generally negligible, with a value of less than 0.1 K across most flight conditions. However, it becomes significant at very high altitudes with low speeds, which are not typical flight conditions for the HALO aircraft. The application of the AI correction results in a slightly lower static air temperature, but this correction remains within the specified error margins for temperature measurements. Therefore, the temperatures calculated during the ML-CIRRUS campaign are likely unaffected by the new AI correction.

Contrail formation can be identified through two main approaches: one based solely on ice-supersaturated regions (ISSRs)

700 and another employing the SAC to derive potential contrail coverage (PotCov). Both require similar microphysical conditions (temperature and humidity thresholds), but the SAC-based approach is more physically comprehensive as it includes aircraft engine parameters such as propulsion efficiency and specific fuel combustion heat (Schumann, 1996; Schumann et al., 2000; Grewe et al., 2 . While the simpler ISSR-based method can identify where contrails may form, the SAC-based approach is generally considered more refined and potentially more accurate, as it directly links aircraft properties and ambient conditions. Consequently, varying

705 these aircraft parameters can influence the spatial extent and frequency of predicted contrails. For instance, adjusting the propulsion efficiency would have a relatively small effect on contrail formation regions, while changes to parameters like the emitted water vapour could more substantially alter the predicted contrail coverage. As demonstrated in frameworks like CoCiP (Schumann, 2012), an accurate representation of engine and fuel parameters can significantly improve the realism of contrail predictions, and thus, inform more effective climate mitigation strategies in aviation.

Despite significant differences in water vapour mixing ratios, simulations with mean temperature nudging show smaller abso-

710 lute differences in relative humidity over ice compared to ERA5 than those with standard nudging. However, both simulations display higher values than the reanalysis data. This indicates that the model temperature bias has a stronger effect on the relative humidity over ice than the model wet bias. Both biases directly impact predicted ice-supersaturated regions, which are crucial for the contrail life cycle. Simulations with standard nudging show larger areas for contrail formation compared to ERA5 across

715 all relative humidity over ice thresholds. However, prior research has found that lower thresholds such as 90% (Dietmüller et al., 2023) or 93% (Hofer et al., 2024) achieve better agreement between observations and the ERA5 model results. This is consistent with the present study, where we found the highest correlation between the ~~potential contrail coverage~~ PotCov areas calculated online during the simulations ~~using the Schmidt-Appleman criteria~~ and the ISSR areas derived offline with a RH_{ice} threshold of less than 96%. The order of magnitude of the identified area on the case study day is consistent with the ISSR occurrence re-

720 sults reported by (Petzold et al., 2020) for 250 hPa and 300 hPa, where they observed ISSR frequencies of occurrence of 19% and 25%, respectively. While the agreement between ERA5 and the various EMAC setups ~~w.r.t. the~~ with respect to potential contrail coverage ~~area is~~ areas remains within 10%, the difference ~~to observations is greater~~ compared to direct observations is larger. The point-by-point ~~correlation~~ comparison between ML-CIRRUS ~~measurement data~~ measurements and EMAC model ~~setups regarding regions where conditions~~ outputs for regions suitable for persistent contrail formation ~~are possible~~ indicates

725 no correlation (low ETS ~~score) between the model and observations~~ scores) when the model threshold is set to 90% ~~while~~ and the observation threshold remains at 100%. Although our limited sample size (286 data points) may partly explain the low skill, more fundamental reasons are likely behind these results.

This threshold pairing yields similar hit rates and false alarm rates across ~~the~~ all model setups. ~~The~~ Adjusting the model RH_{ice} threshold significantly ~~impacts both the hit rate and false alarm rate. Generally~~ affects both metrics: generally, higher thresh-

730 olds ~~tend to increase both metrics~~ increase both hit and false alarm rates. However, ~~maintaining a~~ keeping the false alarm rate under 20% is preferable ~~, as higher false alarm rates would result in more incorrect identification of ice-supersaturated regions, leading to unnecessary rerouting and an increased overall~~ to avoid unnecessary re-routing and increased climate impact. ~~The challenge is that achieving~~ Achieving such a low false alarm rate ~~(<20%) typically results in a low hit rate (<~~ often reduces

the hit rate to below 50%), meaning many areas that could actually support contrail formation remain undetected. Despite this trade-off, it is still better to accept a lower hit rate than to tolerate frequent false alarms. Note that the RH_{ice} thresholds from Dietmüller et al. (2023) and Hofer et al. (2024), which yield better verification scores, apply to ECMWF data, and EMAC may require its own optimized threshold. While contrail cirrus can persist below % RH_{ice} (Li et al., 2023), the initial formation phase typically requires conditions near saturation. Thus, our use of a 100 % observational threshold is physically justified, and lower tested thresholds were intended only to assess model sensitivity.

Beyond threshold issues, comparing localized in-situ aircraft measurements to grid-box averaged model fields remains challenging, and the identified temperature and humidity biases also influence the ETS scores. In line with our findings, Gierens et al. (2020) reported low correlations between model data (ERA5) and aircraft measurements (MOZAIC). Even ERA5, which generally aligns more closely with our contrail metrics, may exhibit its own humidity biases (Charlesworth et al., 2023). Moreover, the complexity of modelling contrails — from small-scale humidity fluctuations to the representation of aircraft engine parameters — further limits the skill of direct comparisons.

Future improvements could involve refining the RH_{ice} thresholds specifically for EMAC or exploring bias correction techniques to better align the model's humidity fields with observations (Teoh et al., 2023). Integrating more comprehensive observational datasets, such as additional flight campaigns or satellite-based humidity measurements, might help reduce uncertainties. Enhancing data assimilation methods and adjusting model parametrizations could also yield more accurate contrail predictions. Ultimately, addressing these limitations and refining the underlying assumptions will be key steps toward achieving higher skill scores and more reliable contrail forecasts in future studies.

Since the Schmidt-Appleman criterion, a thermodynamic theory that has been thoroughly tested and validated, is the basis for the PotCov calculation in EMAC, any deviations in contrail coverage mainly arise from inaccurate representations of key parameters for contrail formation, such as temperature or humidity. Changes to the horizontal or vertical resolution of the model had only little effect on the outcome. Future studies need to evaluate the impact of the transition from the bi-linear interpolation to the nearest neighbour approach. Despite the fact that the 41-layer setup simulates the best agreement with observations, it still does not cover atmospheric processes in the stratosphere well, which could consequently have a minor influence on contrail prediction. However, this study reveals that this does not significantly impact the results, as all three model setups have the same model grid resolution of approximately 450 m vertically in the UTLS region, resulting in similar results. The cold and humidity biases present in all regions of the Northern Hemisphere that were shown in this study may not only affect contrail formation region predictions, but also have an impact on the water cycle,

cloud formation, and the radiation budget.

770 Despite the ~~presence of biases in the model, the areas predicted for contrail formation closely matched~~ model's biases, the predicted contrail formation areas closely match those observed in satellite ~~images. This alignment may be attributed to the atmospheric conditions not being at the extreme edges, suggesting that differences in atmospheric conditions between models and observations do not significantly impact the results in this case.~~ imagery. From 03–05 UTC, we find close agreement in both the high-coverage contrail regions and the regions without contrails. We attribute this agreement to atmospheric conditions

775 being sufficiently far from critical thresholds for contrail formation, so temperature and humidity biases do not significantly affect the ISSR evaluation. Later in the day, from 07–09 UTC, we find only partial agreement in the high-coverage contrail regions, likely due to atmospheric conditions being closer to the formation threshold. Under these conditions, any model bias affects the ISSR evaluation, and areas where contrails are observed are not consistently captured by the model. Although the global mean temperature nudging used in EMAC may not directly translate to operational numerical weather prediction (NWP)

780 systems, the overarching conclusions remain relevant. NWP models could employ improved data assimilation, bias correction strategies, and carefully chosen RH thresholds to enhance contrail forecasting. By doing so, future operational frameworks might better integrate contrail avoidance measures, ultimately contributing to climate-optimized flight operations.

7 Summary and conclusions

This study ~~aimed to evaluate distinct EMAC model setups for their capacity for investigating formation and radiative effects~~
785 ~~of contrails. Utilizing~~ enhances the understanding of contrail prediction limitations in the global climate model EMAC and evaluates categorical statistical measures related to temperature and humidity biases in different EMAC setups. We evaluated several EMAC model setups with various vertical resolutions (L31, L41, L90), horizontal resolutions (T42, T63), and different nudging approaches (STN, MTN) regarding their ability to identify regions of the atmosphere where contrails can form and persist. We assessed temperature and humidity biases and their influence on contrail prediction by comparing EMAC outputs

790 with ERA5 reanalysis data, satellite observations, and in-situ flight measurements ~~, and satellite observations, we analyzed the effect of humidity~~ from the ML-CIRRUS campaign.

A consistent cold bias of 3–5 K was identified in the standard nudged EMAC model (STN) between 100 and 400 hPa. This bias was observed across the Northern Hemisphere, independent of latitude or region, and persisted across all investigated vertical and horizontal resolutions. It aligns with past findings related to the upper boundary conditions in general circulation models.

795 Since contrails require sufficiently low ambient temperatures (around 235 K) to form, this cold bias leads to an overprediction of contrail-forming regions. It also results in an overestimation of relative humidity over ice, enlarging identified ISSRs and thereby affecting contrail persistence. Implementing global mean temperature nudging (MTN) significantly reduces the cold bias, bringing EMAC outputs closer to ERA5 and ~~temperature biases in the model on contrail prediction. Our findings indicate that the vertical modelresolution has minimal impact on the model results. This suggests that the previously used EMAC model~~

800 ~~setup with 41 layers remains effective for future contrail studies, including the expansion of the geographic scope of contrail CCFs. Moreover, the improvements~~ ML-CIRRUS temperature measurements. However, it also introduces a positive humidity

bias, which increases the extent of ISSRs and ultimately leads to an overestimation of contrail persistence.

Despite the differences identified in temperature and humidity, the model's cold bias has only a minor impact on large-scale contrail formation, as aircraft flights predominantly occur in regions where ambient temperatures are well below the critical threshold. All three vertical resolutions (L31, L41, L90) reproduce the extent of regions where persistent contrails can form, with an overestimation of less than 10 percentage points compared to ERA5 data. Consequently, EMAC reliably identifies potential contrail-forming regions, demonstrating its suitability for contrail studies. The small reduction in temperature bias observed when increasing the horizontal grid resolution from T42 to T63 ~~were too small~~ — along with a reduction in ISSR areas — was not significant enough to justify the ~~disproportionate increase of required storage space and computational time.~~

~~Despite the robustness of our results, some limitations exist. The model exhibits a temperature bias of up to 5 , which can affect contrail prediction near the temperature threshold. Additionally, our study only analyzed data from the spring season of 2014, limiting the observation of long-term trends. Future research should aim~~ substantially higher storage and computational costs. It should be noted that the analysis is limited to a single spring season in 2014. Nevertheless, the temperature and humidity biases identified in this study contribute to forecast uncertainties in contrail modelling, which need to be addressed when predicting ISSRs in atmospheric models. These uncertainties form the basis for risk assessments when implementing contrail avoidance strategies or climate-optimized trajectories. By extending existing Climate Change Functions (CCFs) to different seasons and geographical regions, future studies can better quantify how such uncertainties vary over time and space.

One approach to reduce the ~~model bias , potentially by transitioning from the ECHAM to the ICON core model. It is also crucial to tune the~~ temperature bias explored in this study, is mean temperature nudging ~~(MTN) model setup towards .~~ However, the physical inconsistencies introduced by MTN led to new issues related to humidity. Future work should focus on carefully tuning MTN setups toward a realistic radiation balance, as this ~~technique strongly influences parameters such as~~ method strongly affects both humidity and the radiation budget. Failure to ~~do so may~~ address these issues could render the model unsuitable for ~~calculating~~ quantifying the climate impact of contrails. ~~In conclusion, this study enhances the understanding of contrail prediction limitations in current global climate models and provides an estimation of the accuracy and uncertainty when using~~ Similar to nudging, the assimilation of realistic temperature and humidity data in ~~specified EMAC setups. The significant influence of temperature and humidity biases on contrail formation underscores the importance of considering model uncertainties when using climate-sensitive areas for contrail avoidance. Extending the existing Climate Change Functions to different seasons and regions is feasible and recommended for~~ NWP models offers a promising approach to reducing model biases through observational data, which should be further explored in future studies.

Code availability. The Modular Earth Submodel System (MESSy, doi:10.5281/zenodo.8360186) is continuously further developed and applied by a consortium of institutions. The usage of MESSy and access to the source code is licensed to all affiliates of institutions which are members of the MESSy Consortium. Institutions can become a member of the MESSy Consortium by signing the MESSy Memorandum of Understanding. More information can be found on the MESSy Consortium Website (http://www.messy-interface.org).

Data availability. The ERA5 Data are available on the German Climate Computing Center DKRZ and the Copernicus Climate Data Store
(https://climate.copernicus.eu/). The data of the ML-CIRRUS flight experiment was taken from the HALO Database (https://halo-db.pa.op.
dlr.de/). Additional scripts and data are available upon request from the corresponding author. MSG/SEVIRI data can be obtained from
EUMETSAT.

Author contributions. PP performed the EMAC model simulations advised by PJ; LB contributed the satellite images; AG and MK contributed the aircraft measurement data; SM, VG, PJ and CF were involved in the designed the study as well as in the discussion of the results
and supported the writing of the paper. The paper was written by PP, all authors discussed and commented on the manuscript.

Competing interests. The authors declare that they have no conflict of interest. At least one of the (co-)authors is a member of the editorial
board of ACP.

Acknowledgements. The individual authors of this study were supported by the ClimOP consortium within the framework of the European
Union's Horizon 2020 research and innovation program, under grant agreement No. 875503. This study is also supported by the project D-
KULT, Demonstrator Klimafreundliche Luftfahrt (Förderkennzeichen 20M2111A), within the Luftfahrtforschungsprogramm LuFo VI of the
German Bundesministerium für Wirtschaft und Klimaschutz. We utilized high-performance computing resources from the German CARA
Cluster in Dresden and the Deutsches Klima Rechenzentrum (DKRZ) Cluster in Hamburg. In preparing this manuscript, we used ChatGPT,
an AI language model developed by OpenAI, to assist with text improvement and refinement.

835

840

845

References

850   Bange, J., Esposito, M., Lenschow, D. H., Brown, P. R. A., Dreiling, V., Giez, A., Mahrt, L., Malinowski, S. P., Rodi, A. R., Shaw, R. A., Siebert, H., Smit, H., and Zöger, M.: Measurement of Aircraft State and Thermodynamic and Dynamic Variables, chap. 2, pp. 7–75, John Wiley & Sons, Ltd, https://doi.org/https://doi.org/10.1002/9783527653218.ch2, 2013.

Brinkop, S., Dameris, M., Jöckel, P., Garny, H., Lossow, S., and Stiller, G. P.: The millennium water vapour drop in chemistry-climate model simulations, Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acpd-15-24909-2015, 2015.

855   Burkhardt, U. and Kärcher, B.: Process-based simulation of contrail cirrus in a global climate model, Journal of Geophysical Research, 114, https://doi.org/10.1029/2008jd011491, 2009.

Burkhardt, U., Kärcher, B., Ponater, M., Gierens, K., and Gettelman, A.: Contrail cirrus supporting areas in model and observations, Geophysical Research Letters, 35, https://doi.org/10.1029/2008gl034056, 2008.

Charlesworth, E., Plöger, F., Birner, T., Baikhadzhaev, R., Ábalos, M., Abraham, N. L., Akiyoshi, H., Bekki, S., Dennison, F., Jöckel,
860   P., Keeble, J., Kinnison, D. E., Morgenstern, O., Plummer, D. A., Rozanov, E., Strode, S. A., Zeng, G., Egorova, T., and Riese, M.: Stratospheric water vapor affecting atmospheric circulation, Nature Communications, 14, https://doi.org/10.1038/s41467-023-39559-2, 2023.

Dahlmann, K., Grewe, V., Matthes, S., and Yamashita, H.: Climate assessment of single flights: Deduction of route specific equivalent CO_2 emissions, International Journal of Sustainable Transportation, 17, 29–40, https://doi.org/10.1080/15568318.2021.1979136, 2021.

865   Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Quarterly Journal of the Royal Meteorological Society, 137, 553–597,
870   https://doi.org/https://doi.org/10.1002/qj.828, 2011.

Dietmüller, S., Matthes, S., Dahlmann, K., Yamashita, H., Simorgh, A., Soler, M., Linke, F., Lührs, B., Meuser, M. M., Weder, C., Grewe, V., Yin, F., and Castino, F.: A Python library for computing individual and merged non-CO_2 algorithmic climate change functions: CLIMaCCF V1.0, Geoscientific Model Development, 16, 4405–4425, https://doi.org/10.5194/gmd-16-4405-2023, 2023.

Ebert, E. E.: Results of the 3rd Algorithm Intercomparison Project (AIP-3) of the Global Precipitation Climatology Project (GPCP), Research
875   Report 55, Bureau of Meteorology Research Centre, Melbourne, Australia, 1996.

Engberg, Z., Teoh, R., Abbott, T., Dean, T., Stettler, M. E. J., and Shapiro, M. L.: Forecasting contrail climate forcing for flight planning and air traffic management applications: The CocipGrid model in pycontrails 0.51.0, EGUsphere, 2024, 1–45, https://doi.org/10.5194/egusphere-2024-1361, 2024.

Frömming, C., Grewe, V., Brinkop, S., and Jöckel, P.: Documentation of the EMAC submodels AIRTRAC 1.0 and CONTRAIL 1.0, sup-
880   plementary material of Grewe et al., 2014b, Geoscientific Model Development, 7, 175–201, https://doi.org/https://doi.org/10.5194/gmd-7-175-2014, 2014.

Frömming, C., Grewe, V., Brinkop, S., Jöckel, P., Haslerud, A. S., Rosanka, S., Van Manen, J., and Matthes, S.: Influence of weather situation on non-CO_2 aviation climate effects: the REACT4C climate change functions, Atmospheric Chemistry and Physics, 21, 9151–9172, https://doi.org/10.5194/acp-21-9151-2021, 2021.

885 Gans, P. and Gill, J. B.: Smoothing and differentiation of spectroscopic curves using SPLine functions, Applied Spectroscopy, 38, 370–376, https://doi.org/10.1366/0003702844555511, 1984.

Gierens, K. and Vázquez-Navarro, M.: Statistical analysis of contrail lifetimes from a satellite perspective, Meteorologische Zeitschrift, 27, 183–193, https://doi.org/10.1127/metz/2018/0888, 2018.

Gierens, K., Schumann, U., Smit, H. G. J., Helten, M., and Zängl, G.: Determination of humidity and temperature fluctuations based on

890 MOZAIC data and parametrisation of persistent contrail coverage for general circulation models, Annales Geophysicae, 15, 1057–1066, https://doi.org/10.1007/s00585-997-1057-3, 1997.

Gierens, K., Matthes, S., and Rohs, S.: How well can persistent contrails be predicted?, Aerospace, 7, 169, https://doi.org/10.3390/aerospace7120169, 2020.

Giez, A., Zöger, M., Mallaun, C., Nenakhov, V., Schimpf, M., Grad, C., Numberger, A., and Raynor, K.: Determination of the Measurement

895 Errors for the HALO Basic Data System BAHAMAS by Means of Error Propagation, Tech. rep., DLR, https://elib.dlr.de/193175/, 2023.

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

900 Grewe, V., Dahlmann, K., Flink, J., Frömming, C., Ghosh, R., Gierens, K., Heller, R., Hendricks, J., Jöckel, P., Kaufmann, S. H. E., Kölker, K., Linke, F., Luchkova, T., Lührs, B., Van Manen, J., Matthes, S., Minikin, A., Niklaß, M., Plohr, M., Righi, M., Rosanka, S., Schmitt, A. R., Schumann, U., Unterstraßer, S., Vázquez-Navarro, M., Voigt, C., Wicke, K., Yamashita, H., Zahn, A., and Ziereis, H.: Mitigating the Climate Impact from Aviation: Achievements and Results of the DLR WeCare Project, Aerospace, 4, 34, https://doi.org/10.3390/aerospace4030034, 2017.

905 Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049,

910 https://doi.org/https://doi.org/10.1002/qj.3803, 2020.

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., and Thépaut, J.-N.: ERA5 monthly averaged data on pressure levels from 1940 to present.Copernicus Climate Change Service (C3S) Climate Data Store (CDS)., https://doi.org/10.24381/cds.6860a573, accessed: 2023-09-28, 2023.

Hofer, S., Gierens, K., and Rohs, S.: How well can persistent contrails be predicted? — An update, EGU Discussions, p. 24,

915 https://doi.org/10.5194/egusphere-2024-385, preprint, 2024.

Irvine, E. A., Hoskins, B. J., Shine, K. P., Lunnon, R. W., and Froemming, C.: Characterizing North Atlantic weather patterns for climate-optimal aircraft routing, Meteorological Applications, 20, 80–93, https://doi.org/10.1002/met.1291, 2012.

Irvine, E. A., Hoskins, B. J., and Shine, K. P.: A Lagrangian analysis of ice-supersaturated air over the North Atlantic, Journal of Geophysical Research: Atmospheres, 119, 90–100, https://doi.org/10.1002/2013jd020251, 2014.

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

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

Kaufmann, S., Voigt, C., Heller, R., Jurkat-Witschas, T., Krämer, M., Rolf, C., Zöger, M., Giez, A., Buchholz, B., Ebert, V., Thornberry, T., and Schumann, U.: Intercomparison of midlatitude tropospheric and lower-stratospheric water vapor measurements and comparison to ECMWF humidity data, Atmospheric Chemistry and Physics (ACP), 18, 16 729–16 745, https://elib.dlr.de/123559/, 2018.

Krämer, M., Rolf, C., Spelten, N., Afchine, A., Fahey, D., Jensen, E., Khaykin, S., Kuhn, T., Lawson, P., Lykov, A., Pan, L. L., Riese, M., Rollins, A., Stroh, F., Thornberry, T., Wolf, V., Woods, S., Spichtinger, P., Quaas, J., and Sourdeval, O.: A microphysics guide to cirrus – Part 2: Climatologies of clouds and humidity from observations, Atmospheric Chemistry and Physics, 20, 12 569–12 608, https://doi.org/10.5194/acp-20-12569-2020, 2020.

Krisna, T. C., Wendisch, M., Ehrlich, A., Jäkel, E., Werner, F., Weigel, R., Borrmann, S., Mahnke, C., Pöschl, U., Andreae, M. O., Voigt, C., and Machado, L. A. T.: Comparing airborne and satellite retrievals of cloud optical thickness and particle effective radius using a spectral radiance ratio technique: two case studies for cirrus and deep convective clouds, Atmospheric Chemistry and Physics, 18, 4439–4462, https://doi.org/10.5194/acp-18-4439-2018, 2018.

Krämer, M., Rolf, C., Luebke, A., Afchine, A., Spelten, N., Costa, A., Meyer, J., Zöger, M., Smith, J. B., Herman, R. L., Buchholz, B., Ebert, V., Baumgardner, D., Borrmann, S., Klingebiel, M., and Avallone, L. M.: A microphysics guide to cirrus clouds – Part 1: Cirrus types, Atmospheric Chemistry and Physics, 16, 3463–3483, https://doi.org/10.5194/acp-16-3463-2016, 2016.

Krüger, K., Schäfler, A., Wirth, M., Weißmann, M., and Craig, G.: Vertical structure of the lower-stratospheric moist bias in the ERA5 reanalysis and its connection to mixing processes, Atmospheric Chemistry and Physics, 22, 15 559–15 577, https://doi.org/10.5194/acp-22-15559-2022, 2022.

Kärcher, B.: Formation and radiative forcing of contrail cirrus, Nature Communications, 9, https://doi.org/10.1038/s41467-018-04068-0, 2018.

Kästner, M., Meyer, R., and Wendling, P.: Influence of weather conditions on the distribution of persistent contrails, Meteorological Applications, 6, 261–271, https://doi.org/10.1017/s1350482799001231, 1999.

Lee, D. S., Fahey, D. W., Skowron, A., Allen, M., Burkhardt, U., Chen, Q., Doherty, S. J., Freeman, S., Forster, P. M., Fuglestvedt, J. S., Gettelman, A., De León, R. R., Lim, L. L., Lund, M. T., Millar, R., Owen, B., Penner, J. E., Pitari, G., Prather, M. J., Sausen, R., and Wilcox, L.: The contribution of global aviation to anthropogenic climate forcing for 2000 to 2018, Atmospheric Environment, 244, 117 834, https://doi.org/10.1016/j.atmosenv.2020.117834, 2021.

Li, Y., Mahnke, C., Rohs, S., Bundke, U., Spelten, N., Dekoutsidis, G., Groß, S., Voigt, C., Schumann, U., Petzold, A., and Krämer, M.: Upper-tropospheric slightly ice-subsaturated regions: frequency of occurrence and statistical evidence for the appearance of contrail cirrus, Atmospheric Chemistry and Physics, 23, 2251–2271, https://doi.org/10.5194/acp-23-2251-2023, 2023.

Lin, S. and Rood, R. B.: Multidimensional Flux-Form Semi-Lagrangian transport schemes, Monthly weather review, 124, 2046–2070, https://doi.org/10.1175/1520-0493(1996)124, 1996.

Luebke, A., Afchine, A., Costa, A., Grooß, J., Meyer, J., Rolf, C., Spelten, N., Avallone, L. M., Baumgardner, D., and Krämer, M.: The origin of midlatitude ice clouds and the resulting influence on their microphysical properties, Atmospheric Chemistry and Physics, 16, 5793–5809, https://doi.org/10.5194/acp-16-5793-2016, 2016.

Lührs, B., Niklaß, M., Volker, G., and Gollnick, V.: Cost-Benefit Assessment of Climate and Weather Optimized Trajectories for Different North Atlantic Weather Patterns, in: 31st ICAS., 2018.

Lührs, B., Linke, F., Matthes, S., Grewe, V., and Yin, F.: Climate Impact Mitigation Potential of European Air Traffic in a Weather Situation with Strong Contrail Formation, Aerospace, 8, 50, https://doi.org/10.3390/aerospace8020050, 2021.

Matthes, S., Lührs, B., Dahlmann, K., Grewe, V., Linke, F., Yin, F., Klingaman, E., and Shine, K. P.: Climate-Optimized trajectories and robust mitigation potential: Flying ATM4E, Aerospace, 7, 156, https://doi.org/10.3390/aerospace7110156, 2020.

Meijer, V. R., Kulik, L., Eastham, S. D., Allroggen, F., Speth, R. L., Karaman, S., and Barrett, S. R. H.: Contrail coverage over the United States before and during the COVID-19 pandemic, Environmental Research Letters, 17, 034 039, https://doi.org/10.1088/1748-9326/ac26f0, 2022.

Meyer, J., Rolf, C., Schiller, C., Rohs, S., Spelten, N., Afchine, A., Zöger, M., Sitnikov, N., Thornberry, T. D., Rollins, A. W., Bozóki, Z., Tátrai, D., Ebert, V., Kühnreich, B., Mackrodt, P., Möhler, O., Saathoff, H., Rosenlof, K. H., and Krämer, M.: Two decades of water vapor measurements with the FISH fluorescence hygrometer: a review, Atmospheric Chemistry and Physics, 15, 8521–8538, https://doi.org/10.5194/acp-15-8521-2015, 2015.

Murphy, D. M. and Koop, T.: Review of the vapour pressures of ice and supercooled water for atmospheric applications, Quarterly Journal of the Royal Meteorological Society, 131, 1539–1565, https://doi.org/10.1256/qj.04.94, 2005.

Petzold, A., Neis, P., Rütimann, M., Rohs, S., Berkes, F., Smit, H. G. J., Krämer, M., Spelten, N., Spichtinger, P., Nédélec, P., and Wahner, A.: Ice-supersaturated air masses in the northern mid-latitudes from regular in situ observations by passenger aircraft: vertical distribution, seasonality and tropospheric fingerprint, Atmospheric Chemistry and Physics, 20, 8157–8179, https://doi.org/10.5194/acp-20-8157-2020, 2020.

Plummer, D., Nagashima, T., Tilmes, S., Archibald, A., Chiodo, G., Fadnavis, S., Garny, H., Josse, B., Kim, J., Lamarque, J.-F., Morgenstern, O., Murray, L., Orbe, C., Tai, A., Chipperfield, M., Funke, B., Juckes, M., Kinnison, D., Kunze, M., Luo, B., Matthes, K., Newman, P. A., Pascoe, C., and Peter, T.: CCMI-2022: A new set of Chemistry-Climate Model Initiative (CCMI) Community Simulations to Update the Assessment of Models and Support Upcoming Ozone Assessment Activities, SPARC Newsletter, 57, 22–30, 2021.

Pruppacher, H. R., Klett, J. D., and Wang, P. K.: Microphysics of clouds and precipitation, Aerosol Science and Technology, 28, 381–382, https://doi.org/10.1080/02786829808965531, 1998.

Rao, P., Yin, F., Grewe, V., Yamashita, H., Jöckel, P., Matthes, S., Mertens, M., and Frömming, C.: Case study for testing the validity of NOx-Ozone Algorithmic climate change functions for optimising flight trajectories, Aerospace, 9, 231, https://doi.org/10.3390/aerospace9050231, 2022.

Reutter, P., Neis, P., Rohs, S., and Sauvage, B.: Ice supersaturated regions: properties and validation of ERA-Interim reanalysis with IAGOS in situ water vapour measurements, Atmospheric Chemistry and Physics, 20, 787–804, https://doi.org/10.5194/acp-20-787-2020, 2020.

Riede, H., Jöckel, P., and Sander, R.: Quantifying atmospheric transport, chemistry, and mixing using a new trajectory-box model and a global atmospheric-chemistry GCM, Geoscientific model development, 2, 267–280, https://doi.org/10.5194/gmd-2-267-2009, 2009.

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

Rädel, G. and Shine, K. P.: Validating ECMWF forecasts for the occurrence of ice supersaturation using visual observations of persistent contrails and radiosonde measurements over England, Quarterly Journal of the Royal Meteorological Society, 136, 1723–1732, https://doi.org/10.1002/qj.670, 2010.

1000 Schiller, C., Grooß, J., Konopka, P., Plöger, F., Santos, F. H. S. D., and Spelten, N.: Hydration and dehydration at the tropical tropopause, Atmospheric Chemistry and Physics, 9, 9647–9660, https://doi.org/10.5194/acp-9-9647-2009, 2009.

Schmetz, J., Pili, P., Tjemkes, S., Just, D., Kerkmann, J., Rota, S., and Ratier, A.: AN INTRODUCTION TO METEOSAT SECOND GENERATION (MSG), Bulletin of the American Meteorological Society, 83, 977 – 992, https://doi.org/10.1175/1520-0477(2002)083<0977:AITMSG>2.3.CO;2, 2002.

1005 Schumann, U.: On conditions for contrail formation from aircraft exhausts, Meteorologische Zeitschrift, 5, 4–23, https://doi.org/10.1127/metz/5/1996/4, 1996.

Schumann, U.: A contrail cirrus prediction model, Geoscientific Model Development, 5, 543–580, https://doi.org/10.5194/gmd-5-543-2012, 2012.

Schumann, U., Busen, R., and Plohr, M.: Experimental test of the influence of propulsion efficiency on contrail formation, Journal of Aircraft,
1010 37, 1083–1087, https://doi.org/10.2514/2.2715, 2000.

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., Schläger, H., Vázquez-Navarro, M., Voigt, C., and Wang, Z.: Properties of individual contrails: a compilation of observations and some comparisons, Atmospheric Chemistry and Physics, 17, 403–438, https://doi.org/10.5194/acp-17-403-2017, 2017.

Schumann, U., Poll, I., Teoh, R., Koelle, R., Spinielli, E., Molloy, J., Koudis, G., Baumann, R., Bugliaro, L., Stettler, M., and Voigt, C.:
1015 Air traffic and contrail changes over Europe during COVID-19: a model study, Atmospheric Chemistry and Physics, 21, 7429–7450, https://doi.org/10.5194/acp-21-7429-2021, 2021.

Stenke, A., Grewe, V., and Ponater, M.: Lagrangian transport of water vapor and cloud water in the ECHAM4 GCM and its impact on the cold bias, Climate Dynamics, 31, 491–506, https://doi.org/10.1007/s00382-007-0347-5, 2007.

Teoh, R., Schumann, U., Voigt, C., Schripp, T., Shapiro, M., Engberg, Z., Molloy, J., Koudis, G., and Stettler, M.: Tar-
1020 geted use of sustainable aviation fuel to maximize climate benefits, Environmental Science and Technology, 56, 17 246–17 255, https://doi.org/10.1021/acs.est.2c05781, 2022.

Teoh, R., Engberg, Z., Schumann, U., Voigt, C., Shapiro, M., Rohs, S., and Stettler, M.: Global aviation contrail climate effects from 2019 to 2021, EGUsphere, https://doi.org/10.5194/egusphere-2023-1859, 2023.

Tompkins, A. M., Gierens, K., and Rädel, G.: Ice supersaturation in the ECMWF integrated forecast system, Quarterly Journal of the Royal
1025 Meteorological Society, 133, 53–63, https://doi.org/10.1002/qj.14, 2007.

Urbanek, B., Groß, S., Schäfler, A., and Wirth, M.: Determining stages of cirrus evolution: a cloud classification scheme, Atmospheric Measurement Techniques, 10, 1653–1664, https://doi.org/10.5194/amt-10-1653-2017, 2017.

Vázquez-Navarro, M., Mannstein, H., and Kox, S.: Contrail life cycle and properties from 1 year of MSG/SEVIRI rapid-scan images, Atmospheric Chemistry and Physics, 15, 8739–8749, https://doi.org/10.5194/acp-15-8739-2015, 2015.

1030 Voigt, C., Schumann, U., Minikin, A., Abdelmonem, A., Afchine, A., Borrmann, S., Boettcher, M., Buchholz, B., Bugliaro, L., Costa, A., Curtius, J., Dollner, M., Dörnbrack, A., Dreiling, V., Ebert, V., Ehrlich, A., Fix, A., Forster, L., Frank, F., Fütterer, D., Giez, A., Graf, K., Grooß, J., Groß, S., Heimerl, K., Heinold, B., Hüneke, T., Järvinen, E., Jurkat, T., Kaufmann, S. H. E., Kenntner, M., Klingebiel, M., Klimach, T., Kohl, R., Krämer, M., Krisna, T. C., Luebke, A., Mayer, B., Mertes, S., Molleker, S., Petzold, A., Pfeilsticker, K., Port, M., Rapp, M., Reutter, P., Rolf, C., Rose, D., Sauer, D., Schäfler, A., Schlage, R., Schnaiter, M., Schneider, J., Spelten, N., Spichtinger, P.,

Stock, P., Walser, A., Weigel, R., Weinzierl, B., Wendisch, M., Werner, F., Wernli, H., Wirth, M., Zahn, A., Ziereis, H., and Zöger, M.: ML-CIRRUS: The Airborne Experiment on Natural Cirrus and Contrail Cirrus with the High-Altitude Long-Range Research Aircraft HALO, Bulletin of the American Meteorological Society, 98, 271–288, https://doi.org/10.1175/bams-d-15-00213.1, 2017.

Wang, Z., Bugliaro, L., Jurkat-Witschas, T., Heller, R., Burkhardt, U., Ziereis, H., Dekoutsidis, G., Wirth, M., Groß, S., Kirschler, S., Kaufmann, S., and Voigt, C.: Observations of microphysical properties and radiative effects of a contrail cirrus outbreak over the North Atlantic, Atmospheric Chemistry and Physics, 23, 1941–1961, https://doi.org/10.5194/acp-23-1941-2023, 2023.

Wernli, H., Boettcher, M., Joos, H., Miltenberger, A. K., and Spichtinger, P.: A trajectory-based classification of ERA-Interim ice clouds in the region of the North Atlantic storm track, Geophysical Research Letters, 43, 6657–6664, https://doi.org/10.1002/2016gl068922, 2016.

Winterstein, F. and Jöckel, P.: Methane chemistry in a nutshell – the new submodels CH4 (v1.0) and TRSYNC (v1.0) in MESSy (v2.54.0), Geoscientific model development, 14, 661–674, https://doi.org/10.5194/gmd-14-661-2021, 2021.

Woiwode, W., Dörnbrack, A., Polichtchouk, I., Johansson, S., Harvey, B., Höpfner, M., Ungermann, J., and Friedl-Vallon, F.: Technical note: Lowermost-stratosphere moist bias in ECMWF IFS model diagnosed from airborne GLORIA observations during winter–spring 2016, Atmospheric Chemistry and Physics, 20, 15 379–15 387, https://doi.org/10.5194/acp-20-15379-2020, 2020.

Yin, F., Grewe, V., Castino, F., Rao, P., Matthes, S., Dahlmann, K., Dietmüller, S., Frömming, C., Yamashita, H., Peter, P., Klingaman, E., Shine, K. P., Lührs, B., and Linke, F.: Predicting the climate impact of aviation for en-route emissions: the algorithmic climate change function submodel ACCF 1.0 of EMAC 2.53, Geoscientific Model Development, 16, 3313–3334, https://doi.org/10.5194/gmd-16-3313-2023, 2023.

Zängl, G. and Wirth, V.: Synoptic-scale variability of the polar and subpolar tropopause: Data analysis and idealized PV inversions, Quarterly Journal of the Royal Meteorological Society, 128, 2301–2315, https://doi.org/10.1256/qj.01.76, 2002.

Zöger, M., Afchine, A., Eicke, N., Gerhards, M., Klein, E., McKenna, D. S., Mörschel, U., Schmidt, U., Tan, V., Tuitjer, F., Woyke, T., and Schiller, C.: Fast in situ stratospheric hygrometers: A new family of balloon-borne and airborne Lyman α photofragment fluorescence hygrometers, Journal of Geophysical Research, 104, 1807–1816, https://doi.org/10.1029/1998jd100025, 1999.

Appendix A: Relative humidity and water vapour mixing ratio

The relative humidity with respect to ice RH_{ice} is a model parameter required to identify areas where contrails can form. This value relies on the specific humidity (SH), temperature (T), and pressure (P), and can be expressed as the ratio of the water vapour pressure (VP) to the saturation pressure over ice SP_{ice} (see Equation A1):

$$RH_{ice} = 100 \cdot \frac{VP}{SP_{ice}} \tag{A1}$$

With the ratio of molar mass of water to molar mass of dry air (ε) being a constant of 0.622, the water vapour pressure is solely dependent on specific humidity and pressure, and is directly proportional to these variables (Equation A2), while the saturation pressure over ice can be estimated by the temperature (Equation A3, based on Murphy and Koop (2005)) - with the temperature

being inversely proportional to the relative humidity.

$$VP = \frac{P \cdot \frac{SH}{1-SH}}{\varepsilon + \frac{SH}{1-SH}} \tag{A2}$$

$$SP_{ice} = \exp(9.550426 - \frac{5723.265}{T} + 3.53068 \cdot \log(T) - 0.00728332 \cdot T) \tag{A3}$$

This can be demonstrated in a simple example: When applying typical atmospheric values (temperature = 215 K, specific humidity = $3.5*10^{-6}$ kg/kg, pressure = 250 hPa) to Equation A1, the resulting relative humidity is approximately 101%. If the temperature is reduced by 5 K to 210 K (2.5%), while all other parameters are kept constant, the relative humidity over ice rises to 200%. Alternatively, if the specific humidity doubles to $7*10^{-6}$ kg/kg, the relative humidity once again increases beyond 200%. Summarized, a decrease of temperature or increase of specific humidity will result in an increased relative humidity.

The water vapour mixing ratio (MR) is calculated from the specific humidity (SH) values in the model, using Equation (A4); ϵ is the relation of molar mass of water to the molar mass of dry air (0.622):

$$MR = \frac{1}{\varepsilon} \cdot \frac{SH}{1 - SH} \tag{A4}$$

The calculation is necessary for the comparison with humidity data from the ML-CIRRUS campaign, given as water vapour mixing ratio.