Multi-scenario, high-resolution Lagrangian transport modeling for the analysis of the Asian tropopause aerosol layer

Dissertation

submitted in fulfilment of the requirements
for the degree of Doctor of Natural Science (Dr. rer. nat.)

by
Jan Heinrich Clemens

in November, 2023

at
the Faculty of Mathematics and Natural Science
of the University of Wuppertal

Zusammenfassung

Die asiatisch-troposphärische Aerosol-Luftschicht (ATAL) ist eine Schicht von verschmutzter Luft über Asien zur Zeit des Sommermonsuns, die einen Einfluss auf das regionale und globale Klima hat. Die Quellregionen, welche zur Bildung der ATAL beitragen, stehen weiterhin zur Diskussion. Diese Dissertation untersucht die Quellregionen und Transportpfade der ATAL im Jahr 2016 über Nainital (Indien) mittels einer Multi-Szenario Lagrangeschen Rückwärtstransportanalyse von Ballonmessungen unter Verwendung zweier Lagrangescher Transportmodelle, Massive Parallel Trajectory Calculations (MPTRAC) und Chemical Lagrangian transport Model of the Stratosphere (CLaMS), und der hochauflösenden ERA5 Reanalyse.

Der Transport von der planetaren Grenzschicht zur ATAL ist durch schnellen Vertikaltransport (z.B. durch Konvektion) in der Troposphäre, gefolgt von langsamem, diabatischen Aufwärtstransport und großskaligem Horizontaltransport in der oberen Troposphäre und unteren Stratosphäre gekennzeichnet. In der vorliegenden Dissertation wurde das Lagrangesche Transportmodell MPTRAC daher um diabatische Transportberechnungen erweitert, die Berechnungen in der Stratosphäre verbessern. Aufgrund der Rechenleistung von MPTRAC ist es nun zudem möglich, eine große Anzahl von Szenarien mit der ERA5 Reanalysis durchzuführen. Vertikale Transportprozesse können durch die hohe Auflösung von ERA5 besser aufgelöst werden und die große Anzahl von Szenarien ermöglicht als Resultat dieser Dissertation ein genaues Bild von robusten Ergebnissen und verbleibenden Unsicherheiten. MPTRAC erlaubt zudem auch die Parametrisierung von Konvektion und nicht aufgelösten Winden. Das in MPTRAC integrierte Transportschema wurde durch einen Vergleich mit CLaMS detailliert evaluiert. In der vorliegenden Dissertation wird gezeigt, dass MPTRAC und CLaMS mit ERA5 im Vergleich zu Unsicherheiten wie z.B. Unsicherheiten der Reanalysen vernachlässigbare

Abweichungen voneinander aufweisen. Weiterhin wird gezeigt, dass diabatische Transportberechnungen mit der voll aufgelösten ERA5 Reanalyse nicht ohne signifikante Unterschiede durch kinematische Rechnungen oder Rechnungen mit größerer Gitterweite ersetzt werden können.

Die verbesserten Modelle wurden verwendet, um die Quellregionen und Transportpfade zu bestimmen, welche die Messungen über Nainital beeinflussten. Es konnte gezeigt werden, dass trotz der großen Unsicherheiten bei den Vertikalgeschwindigkeiten, der Auflösung und Konvektionsparametrisierung, und den Reanalysen der größte Luftmassenbeitrag aus der Region stammt, welche die Indo-Gangetische Ebene, das Tibetanische Plateau und die Hänge des Himalayas einschließt. Die Analyse zeigte jedoch auch, dass Unsicherheiten bezüglich der Rolle der lokalen, in Reanalysen nicht aufgelösten Konvektion über der Indo-Gangetischen Ebene und den Hängen des Himalaya als auch bezüglich der Konvektion über China bestehen. Die Korrelation zwischen der Rückstreustärke der ATAL während der Messkampagne im August und den Beiträgen aus verschiedenen Regionen liefert Anhaltspunkte über den Einfluss verschiedener Regionen auf die ATAL. Es konnte unter Verwendung aller Szenarien robust gezeigt werden, dass zunehmende Beiträge aus dem tropischen Westpazifik, der unter dem Einfluss tropischer Stürme steht, mit einer Abnahme der gemessenen Rückstreuung korreliert sind. Eine weitgehende, aber nicht vollständige Übereinstimmung der Szenarien konnte für die Indo-Gangetische Ebene gefunden werden; größere Beiträge aus dieser Region sind positiv mit der Rückstreuung der ATAL korreliert. Im Gegensatz zu früheren Studien konnte keine robuste Korrelation zwischen der Rückstreuung und den Beiträgen des Tibetischen Plateaus gefunden werden, was eine Stärke der vorgestellten Multi-Szenario Analyse zeigt, nämlich die Aufdeckung bestehender Unsicherheiten. Insgesamt unterstreicht die Analyse die Rolle anthropogener Quellregionen auf dem indischen Subkontinent für die Entstehung der ATAL über Nainital.

Abstract

The Asian tropopause aerosol layer (ATAL) is a layer of polluted air above Asia during the summer monsoon season, which has a global and regional impact on climate. The source regions that contribute to the formation of the ATAL are still under discussion. This thesis investigates the source regions and transport pathways of the ATAL in 2016 over Nainital (India) with the help of a multi-scenario Lagrangian backward transport analysis of balloon-borne measurements using two Lagrangian transport models, the MPTRAC and the CLaMS, and the high-resolution ERA5 reanalysis.

Transport from the planetary boundary layer to the ATAL is characterized by fast vertical transport (e.g. by convection) in the troposphere and subsequent slow diabatic upward transport and large-scale horizontal transport in the upper troposphere and lower stratosphere. Therefore, in this thesis, the Lagrangian transport model MPTRAC was equipped with diabatic transport calculations, which are known to improve calculations in the stratosphere. Due to the high computational performance of MPTRAC, it allows the employment of a large number of scenarios with the ERA5 reanalysis. Vertical transport processes can be better resolved by the high resolution of ERA5, and the large number of scenarios provides a comprehensive picture of robust results and remaining uncertainties. MPTRAC also contains parameterizations of convection and unresolved winds. The transport scheme built into MPTRAC was evaluated in detail by comparing it with the transport scheme of CLaMS. In this thesis, it is shown that MPTRAC and CLaMS with ERA5 have negligible deviations from each other compared to uncertainties such as reanalysis uncertainties. Furthermore, it is shown that diabatic transport calculations with the fully resolved ERA5 reanalysis cannot be replaced by kinematic calculations or calculations with a larger grid width without significant differences, which corroborates the transition to optimized models with diabatic, fully resolved calculations.

The optimized models were used to determine the source regions and transport pathways that influenced the measurements over Nainital. It was shown that despite the large

uncertainties related to vertical velocity approaches, resolved and parameterized convection, and reanalyses, the largest air mass contribution originates from the area, which covers the Indo-Gangetic Plain, the Tibetan Plateau, and the foothills of the Himalayas. However, the analysis also showed that uncertainties remain about the role of local, unresolved convection over the Indo-Gangetic Plain and the foothills of the Himalayas, as well as in relation to convection over China. The correlation between the backscattering strength of the ATAL during the measurement campaign in August and the contributions from different regions provides an indication as to what influence different regions have on the ATAL. Using all scenarios, it was robustly shown that increasing contributions from the Tropical Western Pacific, which was under the influence of tropical storms, correlate with a decrease in the measured backscatter. Wide but not complete agreement was found for the Indo-Gangetic Plain; larger contributions from this region are positively correlated with the backscattering of the ATAL. In contrast to previous studies, no robust correlation could be found between the backscatter and the contributions from the Tibetan Plateau, which shows the strength of the presented multi-scenario analysis, i.e., the exposure of existing uncertainties. Overall, the analysis emphasizes the role of anthropogenic source regions on the Indian subcontinent in the formation of the ATAL over Nainital.

Contents

Chapter 1

Introduction

The Asian tropopause aerosol layer (ATAL) is a strongly polluted atmospheric layer above Asia during the period from May to August. It extends between 15° N to 35° N and 0° E to 150° E, at a height range between 13 km and 18 km with a distinct variability. The first indications of its presence were detected locally with lidar and balloon-borne measurements (Kim et al., 2003; Tobo et al., 2007). Based on model simulations performed by Li et al. (2005), a large spatial extent of the layer was proposed and finally confirmed with satellite data (Vernier et al., 2011; Höpfner et al., 2019).

The existence of the ATAL has caused a debate about the source regions of aerosol particles, their chemical precursors, and the transport pathways to the ATAL altitudes. Even though great progress has been made in the last few years to settle this debate, which will be reviewed in detail later, many questions remain unanswered and under lively discussion (e.g. Bossolasco et al., 2021; Yu et al., 2022; Schill et al., 2023; Köllner et al., 2023). Since around half of the human species populate the Asian monsoon region and aerosol particles are a major source of uncertainty in the prediction of the regional and global climate, answering those questions is of paramount societal importance (e.g. Li et al., 2016; Arias et al., 2021; IPCC, 2023; Arthur, 2023).

This thesis contributes to the endeavour of completing our understanding of the ATAL within the framework of chemical Lagrangian transport modelling. In this framework, air parcels are advected along given wind fields, while the chemistry can be calculated

along the calculated trajectories. In contrast to the Eulerian framework, where transport is represented by the advection of tracer fields, Lagrangian trajectories allow an immediate diagnosis of transport pathways and source regions with the corresponding trajectories. Lagrangian transport models are therefore powerful tools to answer the open questions about the ATAL, if vertical velocities, small-scale mixing and convection are adequately represented in the model.

This thesis focuses on pure Lagrangian transport calculations without chemistry using the Chemical Lagrangian transport Model of the Stratosphere (CLaMS) and the Massive Parallel Trajectory Calculations (MPTRAC) models and their application for the analysis of balloon-borne measurements of the ATAL over India in August 2016. This thesis aims to overcome, but also expose, remaining limitations in the understanding of transport pathways from the planetary surface layer to the ATAL altitudes. This transport is determined by the transport of gas-phase precursors and aerosol particles from the boundary layer into the upper troposphere and lower stratosphere (UTLS) and the subsequent circulation in the Asian summer monsoon anticyclone (ASMA), which emphasizes the importance of accurate simulations of convection and diabatic vertical transport in the UTLS with the models. Therefore, in this thesis, the two Lagrangian transport models have been optimized and extended, in particular, to facilitate the use of high-resolution ERA5 reanalysis. MPTRAC has additionally been equipped in this thesis with a diabatic transport scheme as in CLaMS that is still compatible with parameterizations of MPTRAC (e.g. for convection) and that improves transport in the UTLS in comparison to the former kinematic transport scheme. A better resolution of convection in ERA5 in combination with a diabatic transport scheme promises a better representation of the overall vertical transport process. Complementary, the computationally efficient model MPTRAC enables multi-scenario transport calculations with ERA5 and hence the detailed estimation of remaining uncertainties.

In accordance with this purpose, in this thesis, the guiding questions for model development, evaluation, and application can be summarized as follows:

1. *Which surface sources and transport pathways characterize air masses found within the ATAL?*
2. *How robust is our understanding of surface sources and transport pathways of ATAL inferred from different Lagrangian transport calculations considering uncertainties in upward transport?*
3. *What are the transport uncertainties in Lagrangian transport calculations of the upper troposphere and the stratosphere with the latest generations of reanalyses (ERA5 and ERA-Interim), vertical transport schemes (kinematic and diabatic) and models (MPTRAC and CLaMS)?*

To answer these guiding questions, the models have been employed in global and regional studies, using multiple scenarios. The scenarios include calculations with different reanalyses (ERA5 and ERA-Interim), different vertical velocities (diabatic and kinematic) as well as calculations with parameterization for convection and unresolved sub-grid scale winds.

In Chapter 2 this thesis proceeds with a review of the current knowledge of the ATAL and its uncertainties, surveying surface emissions in Asia, transport pathways in the monsoon circulation, the composition and chemistry of the UTLS and finally the climate response of the atmosphere to Asian aerosol and trace gas loads.

Subsequently, Chapters 3 and 4 describe in general relevant aspects of Lagrangian transport modelling and more detail the two Lagrangian transport models MPTRAC and CLaMS, as well as the used reanalysis data ERA5 and ERA-Interim. Moreover, in Chapter 4, the new implementation of the diabatic advection scheme in MPTRAC v2.6 is explained. Challenges of transport models faced with new high-resolution reanalysis data and modern High Performance Computing (HPC) architectures are briefly presented.

Chapter 5 exhibits the evaluation of the implemented diabatic advection scheme into the MPTRAC, using a globally distributed air parcel ensemble and seasonal forward calculations. For the evaluation of the newly implemented diabatic transport in MPTRAC v2.6, deviations between MPTRAC v2.6 and CLaMS are put into the perspective of estimated deviations from different sources (e.g. interpolation, reanalysis). This Chapter

hence addresses research question three, by showcasing differences in models, vertical velocity approaches and reanalysis on a global and seasonal scale.

In Chapter 6, the Lagrangian trajectory analysis of balloon-borne measurements of the ATAL in 2016 is performed using the multiple scenarios and sensitivity tests available in the improved Lagrangian transport framework. As many studies are similarly dedicated to the investigation of surface sources of the ATAL based on Lagrangian transport calculations (e.g. Li et al., 2017b; Vernier et al., 2018; Höpfner et al., 2019; Zhang et al., 2019b; Legras and Bucci, 2020; Johansson et al., 2020; Hanumanthu et al., 2020), this Chapter provides useful insights for a wide range of relevant applications and addresses the research questions one, two and three.

Finally, the results are summarized, conclusions are made and an outlook for further research and model development is formulated in Chapter 7. The guiding questions are answered there as well.

The main parts of this thesis are two articles for peer-reviewed scientific journals. The articles are listed below.

- Clemens, J., Vogel, B., Hoffmann, L., Griessbach, S., Thomas, N., Fadnavis, S., Müller, R., Peter, T., and Ploeger, F.: A multi-scenario Lagrangian trajectory analysis to identify source regions of the Asian tropopause aerosol layer on the Indian subcontinent in August 2016, Atmospheric Chemistry and Physics, 24, 763–787, doi: $10.5194/\text{acp-24-763-2024}$, 2024b
- Clemens, J., Hoffmann, L., Vogel, B., Grießbach, S., and Thomas, N.: Implementation and evaluation of diabatic advection in the Lagrangian transport model MPTRAC 2.6, Geoscientific Model Development, 17, 4467–4493, doi: $10.5194/\text{gmd-17-4467-2024}$, 2024a

Furthermore, the following papers were published or submitted during the PhD period and are related to the model development and transport in the Asian monsoon region:

- Clemens, J., Ploeger, F., Konopka, P., Portmann, R., Sprenger, M., and Wernli, H.: Characterization of transport from the Asian summer monsoon anticyclone into the UTLS via shedding of low potential vorticity cutoffs, Atmospheric Chemistry and Physics, 22, 3841–3860, doi: $10.5194/\text{acp-22-3841-2022}$, 2022
- Hoffmann, L., Baumeister, P. F., Cai, Z., Clemens, J., Griessbach, S., Günther, G., Heng, Y., Liu, M., Haghighi Mood, K., Stein, O., Thomas, N., Vogel, B., Wu, X., and Zou, L.: Massive-

Parallel Trajectory Calculations version 2.2 (MPTRAC-2.2): Lagrangian transport simulations on graphics processing units (GPUs), Geoscientific Model Development, 15, 2731–2762, doi: 10.5194/gmd-15-2731-2022, 2022

- Hoffmann, L., Konopka, P., Clemens, J., and Vogel, B.: Lagrangian transport simulations using the extreme convection parameterization: an assessment for the ECMWF reanalyses, Atmospheric Chemistry and Physics, 23, 7589–7609, doi: 10.5194/acp-23-7589-2023, 2023b
- Vogel, B., Volk, C. M., Wintel, J., Lauther, V., Clemens, J., Grooß, J.-U., Günther, G., Hoffmann, L., Laube, J. C., Müller, R., Ploeger, F., and Stroh, F.: Evaluation of vertical transport in ERA5 and ERA-Interim reanalysis using high-altitude aircraft measurements in the Asian summer monsoon 2017, Atmospheric Chemistry and Physics, 24, 317–343, doi: 10.5194/acp-24-317-2024, 2024
- Hoffmann, L., Haghighi Mood, K., Herten, A., Hrywniak, M., Kraus, J., Clemens, J., and Liu, M.: Accelerating Lagrangian transport simulations on graphics processing units: performance optimizations of Massive-Parallel Trajectory Calculations (MPTRAC) v2.6, Geoscientific Model Development, 17, 4077–4094, doi: 10.5194/gmd-17-4077-2024, 2024

Chapter 2

The Asian tropopause aerosol layer

The ATAL is located over Asia in the UTLS during the Asian summer monsoon season. In detail, the ATAL has been observed between May and September, peaking in the summer in correlation to the ASMA (Brunamonti et al., 2018). As depicted in Fig. 2.1, the ATAL covers a large area between 15 °N and 35 °N in latitudes and 0 °E and 150 °E in longitudes. Vertically, the ATAL extends between 14 km and 18 km. However, the extent and density of the ATAL show a distinct variability on time scales from days to years (e.g. Vernier et al., 2011; Hanumanthu et al., 2020; Bossolasco et al., 2021). The first evidence of the ATAL from balloon-borne observations of aerosols was found in 1 999 (Kim et al., 2003; Tobo et al., 2007), and the large extent of the ATAL was later demonstrated by satellite observations (Vernier et al., 2011; Thomason and Vernier, 2013). Since the discovery of the ATAL questions concerning its origin, composition, variability, and climate impact have been debated. In recent years, large progress has been made on these questions. In the following chapters, the current state of knowledge about Asian emissions, transport pathways, composition, and chemistry is briefly reviewed.

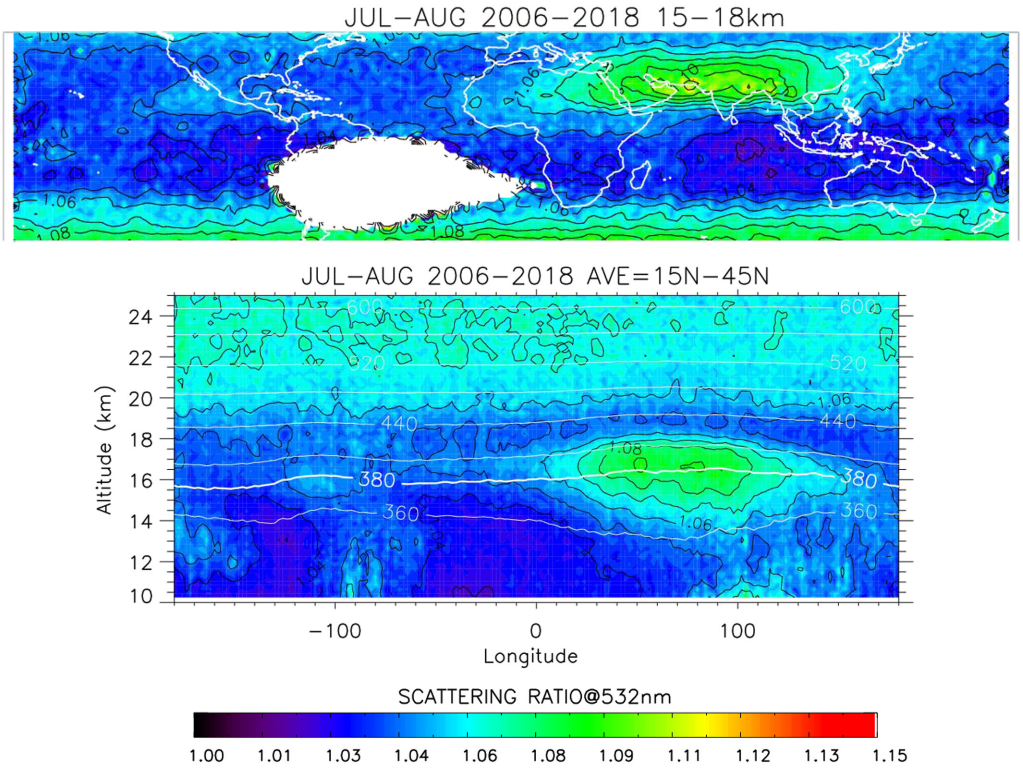


Figure 2.1: Summer average of the cloud-filtered aerosol scattering ratio (SR) derived from the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIOP) for the years between 2006 and 2018, with years with volcanic activity excluded. While the upper panel shows the horizontal map of the SR, the lower panel shows the longitudinal cross-section averaged between 15° and 45°. Adapted from Fairlie et al. (2020).

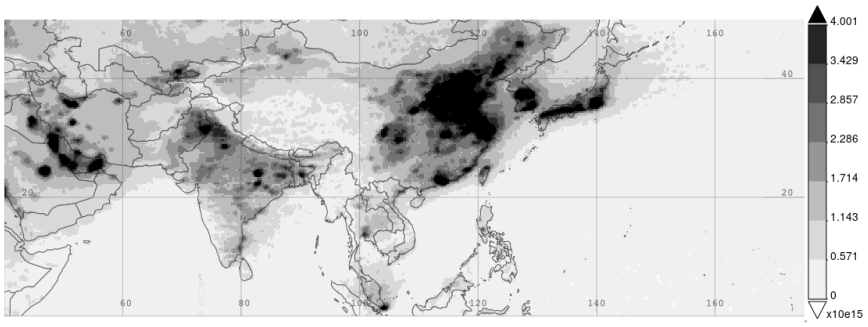
2.1 Surface air pollution and emissions in Asia

For a comprehensive assessment of the ATAL and its possible surface sources, emissions of possible precursor gases and aerosols in the lower troposphere are briefly reviewed. Anthropogenic aerosol and trace gas emissions in many regions of Asia are relatively high compared to other regions of the world. Furthermore, they show strong regional variability in amount and chemical composition (see for example Fig. 2.2 and in the Appendix Fig. A.1). Frequently emitted air pollutants are SO_2, NOx, CO, non-methan volatile organic compounds (NMVOC), NH_3, CO_2, black- and organic carbon, and particular matter. In total, emissions of aerosols and trace gases show an overall increase from 1950 to the 2000s in Asia, but with distinct and complex regional differences that originate from differences in ongoing industrialization and regional policies. During the last two decades, the emissions of many aerosols and trace gases started to decline again (Li et al., 2016; Kurokawa and Ohara, 2020). Greenhouse gas emissions in Asia, however, show an

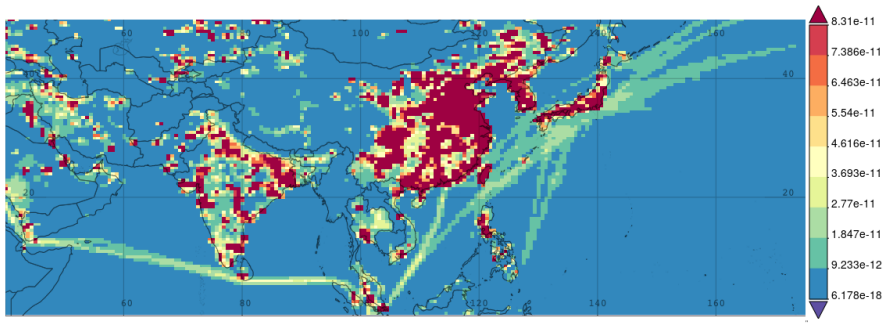
increase over the last decades (e.g. Friedlingstein et al., 2022).

Trace gases and aerosols are significantly emitted in India and China, causing peaks of air pollution in urbanized, industrialized, or agricultural areas. Other regions of Asia, in contrast, show smaller anthropogenic emissions. Prominent emission and air pollution hot spots are the Indo-Gangetic Plain at the foothills of the Himalayas and East China (e.g. Li et al., 2016; Kurokawa and Ohara, 2020). The Indo-Gangetic Plain is in particular a source of large amounts of ammonia as a consequence of intensive agriculture and fertilizer use (Kuttippurath et al., 2020) and an area with high aerosol optical depths (AOD) during the monsoon season (Dey and Di Girolamo, 2010). North-eastern India has been additionally reported as a hot spot for increasing SO_2 emissions (Li et al., 2017a). East China is also a region of high SO_2 emissions (see also Fig. 2.2). SO_2, NOx, and NH_3 emissions make East China and the northern part of the Indian subcontinent potential source regions of secondary aerosol particles because these chemical species can react to aerosol particles under various conditions (e.g. Wang et al., 2013; Höpfner et al., 2019; Wang et al., 2022). Amid a global decline of Chlorofluorocarbon (CFC) emissions, intermediate unreported emissions of CFCs in East China have been detected, causing significant effects on the global emission budget (Rigby et al., 2019; Montzka et al., 2021) and giving a further example of the importance of this region for global emission budgets of various trace gases.

Natural emissions of aerosols and trace gases in Asia are less abundant than anthropogenic emissions and not the driver of current pollution trends, but they are still relevant for the monsoon system (Li et al., 2016). Deserts and dry lands, such as the Tibetan Plateau, the Taklamakan Desert, the Thar Desert, or the Arabian Desert, are sources of large amounts of dust across Asia (e.g. Wang et al., 2021; Gandham et al., 2022, and Appendix Fig. A.1b). Asia's adjacent oceans provide large air masses with water vapor and sea salt on the continent (e.g. Dey and Döös, 2021, and Appendix Fig. A.1c). Other natural emissions, such as organic carbon, originate from forest fires or from biological processes (Li et al., 2016).



(a) NO_2



(b) SO_2

Figure 2.2: Average surface conditions considering NO_2 and SO_2 for the summer 2016 over Asia. (a) NO_2 tropospheric column in $\mathrm{mol} \cdot \mathrm{cm}^{-2}$ derived from the OMI/Aura satellite instrument. (b) anthropogenic SO_2 emissions of the MERRA-2 reanalysis in $\mathrm{kg} \cdot \mathrm{m}^{-2} \cdot \mathrm{s}^{-1}$. Produced with the Giovanni online data system, developed and maintained by the NASA GES DISC (Acker and Leptoukh, 2007; Global Modeling and Assimilation Office, 2015b; Krotkov et al., 2019).

2.2 Transport within the Asian monsoon circulation

The transport in the Asian monsoon circulation can be summarized as follows (see also Fig. 2.3): Deep convection during the Asian monsoon season transports air masses from the Asian monsoon region upward into the UTLS. Subsequently, air masses can be accumulated within the ASMA, where they slowly move upward in a spiralling motion. Air masses can leave the ASMA again because of the instabilities of the anticyclone or by further transport into the higher stratosphere into the tropical pipe. The individual phases of this transport process are discussed in detail in the following paragraphs.

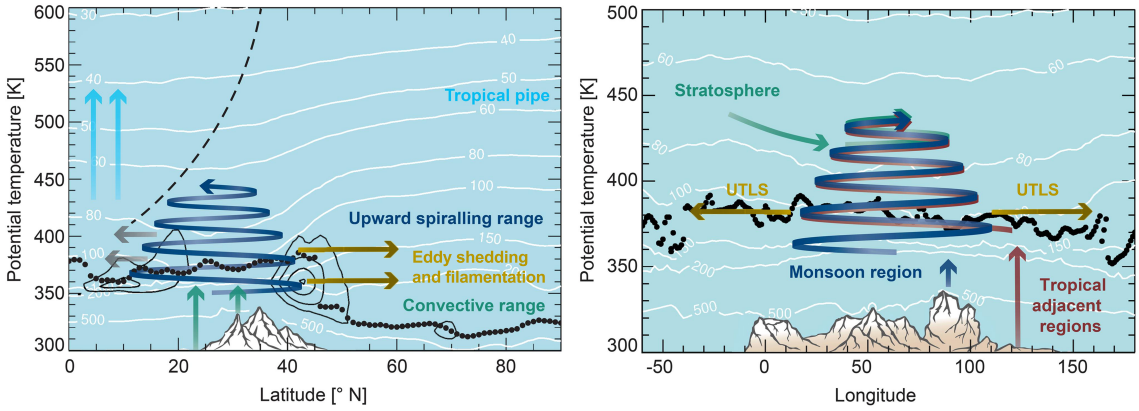


Figure 2.3: Transport pathways in the Asian monsoon circulation in (a) a latitude-height cross-section and (b) a longitude-height cross-section. Arrows indicate the main transport phases. The black dots indicate the thermal tropopause, and the black contours indicate the horizontal winds. Pressure levels are shown with white lines. The black dashed line shows the edge of the tropical pipe. Adapted from Vogel et al. (2019).

The transport of surface trace gases and aerosols from the boundary layer of the Tropics into the UTLS is determined by deep convection, which injects air masses from tropical surface source regions rapidly into the UTLS. In the convection, the air masses are lifted to maximum heights of around 360 K into the tropical tropopause layer within a few hours to days (e.g. Vogel et al., 2019). However, it was shown that the height, extent and timescale of the convective lifting process as simulated in Lagrangian transport calculations depends largely on the resolved and parameterized convection in the reanalysis, highlighting remaining uncertainties about the vertical distribution of air masses in convection (Hoffmann et al., 2023b).

Frequent occurrence of deep convection in Asia, which is relevant for transport from the boundary layer into the UTLS, has been observed over the Western Pacific, North India (South Asian monsoon depressions), the Bay of Bengal and the Tibetan Plateau (e.g. Fu et al., 2006;

Wright et al., 2011; He et al., 2020). The presence of the ATAL is closely interlinked with the ASMA, which coincides spatio-temporaly with trace gas anomalies (Garny and Randel, 2013; Santee et al., 2017) and aerosols (Vernier et al., 2011) in the UTLS during the monsoon season over Asia. Convective transport into the ASMA and to the ATAL, has been reported from a wide range of regions as well, such as the Indian Subcontinent, the Bay of Bengal, Southeast Asia or China (e.g. Vernier et al., 2018; Lau et al., 2018; Bucci et al., 2020; Zhang et al., 2019b), following from the wide variety of convective activity, the abundance of emissions and the large extent of the ASMA in Asia (Fadnavis et al., 2013). However, Bergman et al. (2012) emphasize, that the largest contribution to the air masses in the ASMA comes from a vertical conduit placed at the eastern foothills of the Himalayas, including parts of the Indo-Gangetic plain and the Tibetan Plateau, which is close to the climatological centre of the ASMA. Lau et al. (2018) instead implicated a pattern with two regions of large contributions (two conduits, one over North India and a second over the Sichuan basin in China). These findings show simulated locations of convection have many uncertainties, but are important to clarify the convective processes that are involved in the transport process relevant for the formation of ATAL and in particular to identify surface source regions.

Transport can occur not only directly into the ASMA but also indirectly with prior transport into the tropical UTLS from regions adjacent to the monsoon region (Li et al., 2017b; Zhang et al., 2020). If air masses are located in the tropical UTLS they can enter the ASMA as well horizontally by intrusions. Tropical storms over the Western Tropical Pacific can transport air masses into the UTLS. Then, these air masses can be transported eastward into the anticyclonic circulation and influence the composition of the ASMA (Li et al., 2017b). Li et al. (2020) also show that the representation of such transport in the UTLS inferred from Lagrangian transport calculations depends on the vertical velocity approach and the reanalysis. In this context, it should also be noted that the quality of representation of tropical typhoons varies from reanalysis to reanalysis (Malakar et al., 2020).

In the presence of the ASMA, air masses are transported upward by slow diabatic heating (with a vertical velocity of about 1 K per day) superimposed on the anticyclonic flow resulting in an upward spiraling movement of individual air parcels (Vogel et al., 2019; von Hobe et al., 2021). During this process, the ASMA establishes a leaky barrier trapping the air pollution, so that polluted air masses within the ASMA tend to be accumulated and mixed in it (Ploeger et al., 2015). Therefore, the ASMA and anomalies in trace gas concentration coincide very well (Garny and Randel, 2013; Santee et al., 2017). During the upward transport within the ASMA air parcels

also leave the troposphere vertically and transition from the troposphere to the stratosphere (Vogel et al., 2019; von Hobe et al., 2021). However, properties such as the vertical velocity and cross-isentropic dispersion of air masses during the transport in the UTLS are subject of discussion because of different modelling approaches related to the vertical velocities and different reanalyses (e.g. Ploeger et al., 2011; Li et al., 2020; Bucci et al., 2020; Ploeger et al., 2021).

Once the air masses are within the ASMA, they are affected by the strong seasonal and subseasonal variability of the anticyclone and the slow upward transport. As a consequence, the high variability of the ASMA has a strong impact on the ATAL as well (e.g. Hanumanthu et al., 2020). Splitting events of the ASMA, changes to different modes or shedding of eddies have been studied (Hsu, 1998; Popovic and Plumb, 2001; Yongfu et al., 2002; Homeyer and Bowman, 2013; Nützel et al., 2016; Ungermann et al., 2016; Honomichl and Pan, 2020; Manney et al., 2021; Rupp and Haynes, 2021). These instabilities weaken the trapping effect of the ASMA and enabling leakage into (e.g. stratospheric intrusions) and out of the anticyclone (Dethof et al., 1999; Gottschaldt et al., 2018; Clemens et al., 2022). Hence, the transport out of the ASMA *into the stratosphere* can proceed by two pathways (Vogel et al., 2019). First, large-scale ascent in the tropics can lift the air masses into the stratosphere. Second, quasi-horizontal transport within eddies or filaments of the ASMA in east- or westward direction, that penetrate the tropopause can move the air masses into the lowermost stratosphere (Vogel et al., 2016; Rolf et al., 2018; Lauther et al., 2022). The eddies carry polluted air masses up to polar latitudes and North America, although the extent and efficiency of this transport is under debate. To study these transport pathways in more detail, the aircraft campaign Probing High Latitude Export of air from the Asian Summer Monsoon (PHILEAS) was performed in summer 2 023 with the German HALO (see also HALO, 2023). Additionally, He et al. (2020) found large scale descend *into the troposphere* in the western part of the ASMA as a sink for the aerosol in the UTLS.

Although the general picture of transport has become well established during the last years, the transport modelling of the process still has limitations because of reanalysis uncertainties in the representation of vertical transport, mixing and dynamics in the UTLS and concerning the representation of convection like tropical storms (height, location, intensity). Insofar, it is important to validate the general picture in the light of the most recent, high-resolution reanalyses such as ERA5 in combination with diabatic vertical velocities, convection parameterizations and sub-grid scale wind uncertainties, to report robust features and expose uncertainties and their causes.

2.3 Composition and chemistry of the ATAL

A unique composition of the UTLS can be found within the ASMA above Asia during the monsoon season. In comparison to the pre-monsoon UTLS background, enhanced mixing ratios of various tropospheric trace gases (CH_3OH, CO, CH_3Cl, CH_3CN) and reduced mixing ratios of stratospheric trace gases (O_2, HNO_3, HCl) have been observed with the Aura Microwave Limb Sounder (MLS) (see Santee et al., 2017, and references therein). Also H_2O is enhanced within the ASMA, although it shows a more complex distribution than other trace gases. These findings, in combination with the knowledge about the strong convective impact on the UTLS, show that the air masses in the ASMA are strongly affected by the troposphere. However, the intrusions of stratospheric air masses are not negligible.

Closely related to the ASMA - this very unique reaction pod in the UTLS - the ATAL has been detected, probed and studied in the last decades (e.g. Kim et al., 2003; Tobo et al., 2007; Vernier et al., 2011; Thomason and Vernier, 2013; Brunamonti et al., 2018; Vernier et al., 2018; Höpfner et al., 2019; Hanumanthu et al., 2020; Köllner et al., 2023). The composition and chemistry of the ATAL will be discussed in more detail in the following paragraphs.

Direct conclusions about the chemical composition of the ATAL were initially not possible based on the first satellite observations, even though they indicated that the ATAL aerosol particles are liquid, i. e., they are identified as spherical objects, or very small solid particles (Vernier et al., 2011). Since then, different chemical components have been discussed, such as dust, sulfur, organics, nitrates and ammonia as relevant for the formation of the ATAL and compatible with those first measurements.

Unique aircraft measurements during the StratoClim campaign over the Indian subcontinent in summer 2 017 gave deeper insights into the chemical composition of ATAL particles, indicating that ammonium, nitrate, and organics are important contributors to the chemical composition of ATAL particles (Höpfner et al., 2019; Appel et al., 2022; Stroh and StratoClim-Team, 2023, and Fig. 2.4). Appel et al. (2022) also highlight that a significant particle fraction (up to 70 % of the total particle number) within the ATAL results from the conversion of gas-phase precursors rather than from the uplift of primary particles from below. Furthermore, Weigel et al. (2021) found evidence of new particle formation at ATAL altitudes, emphasizing the presence of secondary aerosols. Based on a combination of satellite observations, aircraft measurements, and cloud-chamber experiments, Höpfner et al. (2019) demonstrated the presence of ammonium nitrate particles in the ASMA. Additionally, they traced the ammonia, which is a precursor gas for the particles, to its surface sources at the Indian subcontinent. The Balloon Measurement Campaigns of the Asian

Tropopause Aerosol Layer (BATAL) as well reported the presence of nitrate in the ATAL (Vernier et al., 2018). These findings emphasize the importance of organics, nitrate and ammonium for the formation of the ATAL.

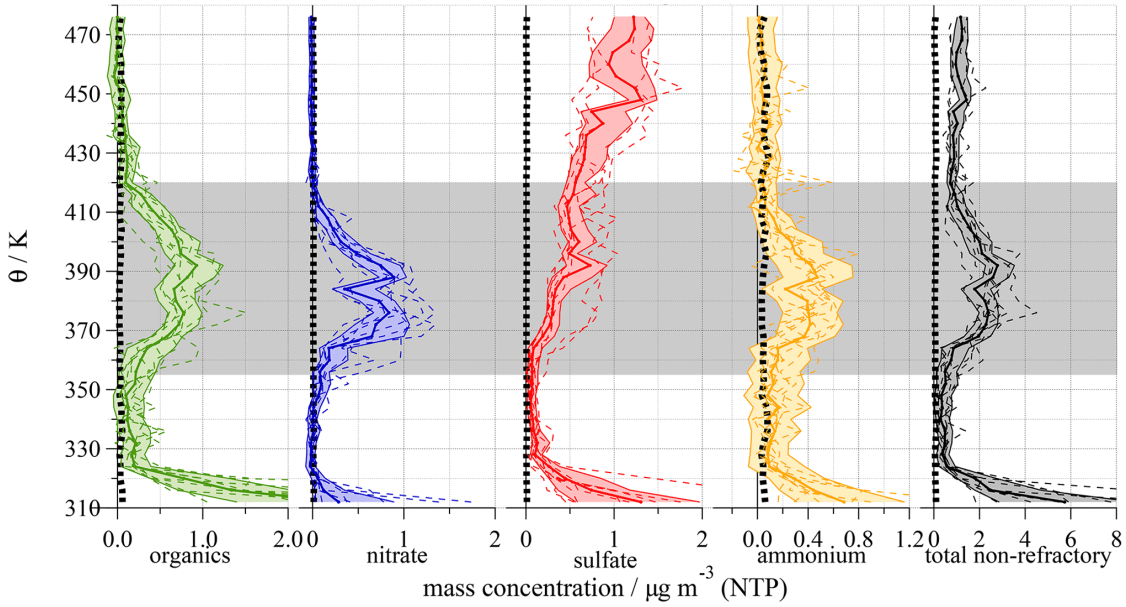


Figure 2.4: Profiles of the mass concentration of different chemical species during the StratoClim aircraft campaign in 2 017. Thick lines indicate the medians, while the shaded areas cover the inter-quartile range for the entire measurement campaign, and the dashed lines show the median of individual flights. Adapted from Appel et al. (2022).

Sulfate aerosols have been found in the ATAL with the aircraft measurements during StratoClim, too. However, in 2 017 they didn't show a mass concentration peak within the ATAL but appeared at upper levels, possibly caused by influence from the stratospheric aerosol layer onto the ATAL (Appel et al., 2022, and Fig. 2.4). These findings indicate that, although SO_2 emissions are high across Asia, wet-deposition of sulfuric species within convection potentially hinders efficient transport of SO_2 into the ASMA (an effect well-known for the Junge-layer). Hence, most SO_2 presumably originates from sedimentations of the Junge-layer into the ASMA.

In addition to non-refractory aerosol components like nitrate and organics that disintegrate at high temperatures, significant fractions of the aerosol particles included refractory components (up to 50 % in the lower ATAL) according to the StratoClim campaign (Appel et al., 2022; Köllner et al., 2023). For higher levels than 400 K they are likely of meteoric origin (Schneider et al., 2021). Below 400 K potassium was a large contributor to the refractory particles, often mixed with nitrate, organics and sulfates. Elemental carbon, iron, and calcium, i.e., indicators of primary particles from the surface, contribute only minor to the ATAL according to the measurements.

During the balloon campaign BATAL calcium has been found as well in the ATAL range (Vernier et al., 2018). These findings demonstrate the potential of the transport of primary particles (e.g. from desert dust, biomass burning, or anthropogenic emissions) directly into the ATAL, although their abundance possibly remains limited by wet-deposition (Köllner et al., 2023).

While the focus of many studies was the formation process of the ATAL, recently the chemical aging and mixing with the stratospheric background during transport of air parcels away from the ATAL into the remote UTLS have become a more intensive research field. Measurement campaigns (e.g. ACCLIP: NCAR, 2023, and PHILEAS), which aim to measure the composition of ASMA outflow for the first time, will put prior findings under further scrutiny and will quantify in more detail the removal processes of the aerosol from the layer. Preliminary analysis and reports from ACCLIP and PHILEAS hint again towards enhanced ammonium and nitrate aerosol fractions. Further analysis of the campaigns data will help to understand the removal processes relevant for the ATAL (e.g. Pan et al., 2022; Schill et al., 2023; Eppers et al., 2023, and personal correspondence).

Model simulation, with general circulation models and coupled aerosol and chemistry models, summarizes our quantitative understanding of the ATAL. Many such model studies have been performed with mixed results except for the broader picture. All models can reproduce an increased concentration of aerosol and trace gases in the observed height range and horizontal extent of the ATAL. A wide variety of aerosols found in the ATAL can be simulated in many cases as well. Composites of the ATAL simulated in the models are dust, sulfate, black carbon, organic carbon, ammonium, and nitrate aerosols. However, the abundance and relative importance of each species vary significantly from study to study.

Many simulations that take nitrate and ammonia into consideration show major contributions from those species (Fairlie et al., 2020; Gu et al., 2016; Yu et al., 2022), supporting the observational evidence for their abundance in the ATAL. A combination of laboratory experiments and model simulations showed a chemical mechanism that can lead to efficient aerosol growth from ammonia and nitrate acid in the cold environment of the UTLS, addressing concerns about the wet-deposition in the convection that is present during transport of precursor gases into the UTLS (Wang et al., 2022). Yu et al. (2022) validated in a global modeling study the significant contributions of nitrate acid aerosols in the cold UTLS over the Asian monsoon region.

Motivated by large SO_2 emissions in Asia, many models simulated sulfuric aerosol as an important component of the ATAL (Li et al., 2005; Neely III et al., 2014; Bossolasco et al., 2021). In comparison, the StratoClim data from 2017 however, question if sulfuric aerosol plays the

dominant role for the formation of the ATAL (Appel et al., 2022). Models that take nitrate and ammonia into account as well, simulate a lower, yet significant contribution of sulfuric aerosol to the ATAL, too (e.g. Gu et al., 2016).

Deserts and the Tibetan Plateau are potential sources of dust. Hence, many model simulations and reanalysis emphasize the role of dust for the composition of the ATAL (Lau et al., 2018; Ma et al., 2019; Bossolasco et al., 2021). Simulations also indicate a minor contribution of black carbon to the ASMA, which nevertheless could have large effects on the radiative effects (Gu et al., 2016; Fadnavis et al., 2017; Bossolasco et al., 2021). Many models also indicate the contribution of secondary and primary organic carbon to the ATAL (Fadnavis et al., 2013, 2018; Fairlie et al., 2020; Bossolasco et al., 2021), but with varying amounts.

In summary, the evidence from observation and modeling increasingly corroborates that the ATAL is not a homogeneous, stable layer similar to the Junge layer, which is dominated by sulfuric species and does maintain long-term stability. In contrast, research supports that the ATAL hosts a large variety of aerosol species with individual, partly connected life-cycles (most notable nitrate, organics, and ammonia, but also sulfate and potentially dust) and that the ATAL follows the variable seasonal and sub-seasonal conditions in the ASMA. Moreover, this research exhibits that the ATAL aerosol also distinguishes from the tropospheric aerosol because of contrasting conditions in the UTLS, which are lower temperatures than in the lower troposphere, stratospheric intrusions, large-scale mixing and medium-range transport, and the involvement of wet-deposition in convection.

2.4 Climate response to Asian aerosol and trace gases

Anthropogenic aerosols and trace gases in the atmosphere are key drivers of climate change as they perturb the radiative balance. The 6th Intergovernmental Panel on Climate Change (IPCC) report states: *"In addition to the influence of internal variability, near-term projected changes in precipitation at global and regional scales are uncertain because of model uncertainty and uncertainty in forcings from natural and anthropogenic aerosols (medium confidence)"* (IPCC, 2023, p. 26).

In the 6th report of the IPCC, Arias et al. (2021) summarize the past and future changes of the Asian summer monsoon as an interplay between cooling aerosol radiative forcing, warming greenhouse gase forcing and natural variability. The report states that the Asian monsoon has become weaker in the second half of the 20th century, mostly due to increasing (tropospheric)

aerosol loads which cause a cooling effect of the continent relative to the oceans (A reduced ocean-continent contrast weakens the monsoon overturning cell and those the monsoon intensity). This effect has lowered the impact of climate warming. In the long-term future, the Asian monsoon is predicted to be dominated by anthropogenic climate warming. Climate warming will increase the evaporation and moisture content of the air masses and therefore counteract the aerosol effects. In the near-term (20 years) however, the monsoon will follow the natural variability, which is for example linked to the El Niño-Southern Oscillation (ENSO) phases. In summary, aerosol and trace gases in Asia play a pivotal role in regional climate change in the long-term but must be studied in interlink with the natural variability in the near-term.

The unique composition and location of the ATAL will likely involve it in the changes of the monsoon. Existing studies indicate for example, that the ATAL modulates precipitation in Asia and the drought intensity in close interplay with the ENSO (Fadnavis et al., 2017, 2019), that the formation of the ATAL could affect the properties of cirrus clouds and hence the local radiative feedback (Wang et al., 2022) and that the regional radiative forcing of the ATAL is around -0.1 Wm^2 (Vernier et al., 2015). Finally, the role of the Asian summer monsoon circulation as the main gateway for surface pollutants, water vapour and aerosol precursors into the stratosphere has likely global relevance through stratospheric chemistry and radiation (e.g. Randel et al., 2010; Bian et al., 2020).

Chapter 3

Lagrangian transport modeling

Lagrangian transport models are employed for the calculation of the movement of air parcels in the atmosphere. Air parcels are considered infinitesimal pieces of air in the atmosphere, that are small enough to be treated as part of the gaseous continuum, and large enough to reach the thermodynamic limit, i.e. to possess thermodynamic properties like temperature or trace gas mixing ratios (continuum assumption). These air parcels are transported within the atmospheric flow (e.g. Lin et al., 2013).

Let $\vec{u}(\vec{x}, t)$ be the given velocity field characterizing the flow in the atmosphere in time t and $\vec{x}(t)$ the position of an air parcel, with initial conditions $\vec{x}(t = 0) = \vec{x}_0$. The solving process for the non-linear, non-autonomous, first order ordinary differential equation:

$$\frac{\mathrm{d}\vec{x}(t)}{\mathrm{dt}} = \vec{u}(\vec{x}(t), t) \tag{3.1}$$

$$\vec{x}(t = 0) = \vec{x}_0 \tag{3.2}$$

is called a Lagrangian transport calculation. In comparison to the Eulerian transport calculations, which describe the evolution of fields on a constant grid, the Lagrangian transport calculations are performed on a moving grid, following the air parcels. The solution $\vec{x}(t)$ of the equation is referred to as the trajectory of the air parcel and describes the pathway the air parcel takes within the atmospheric flow. A large enough ensemble of air parcels and trajectories that have identifiable field properties enables the study of the transport of the fields in space as well. In

a continuous flow, the equation has a unique solution, i.e., the trajectory is only determined by the starting position. Hence, every air parcel in a flow can be identified unambiguously with the starting position and studied separately (e.g. Stohl, 1998).

3.1 Coordinate systems

The equations 3.1 and 3.2 can be formulated in a variety of coordinate systems. Horizontal positions of the air parcels are often denoted in spherical coordinates, i.e. in latitude ϕ and longitude λ. The Cartesian coordinates are then given by

$$\vec{x} = \begin{bmatrix} (R_e + h)\cos\phi\sin\lambda \\ (R_e + h)\cos\phi\cos\lambda \\ (R_e + h)\sin\phi \end{bmatrix} \tag{3.3}$$

where the R_e is the Earth' radius and h the height over ground. In many cases the shallow atmosphere approximation is used and the height over ground is neglected: $R_e + h \approx R_e$. The treatment of the coordinate singularity at the poles, however, in some models makes a regional change of coordinates necessary. Some models change for example to a polar stereo-graphic coordinate to describe air parcel positions (e.g. McKenna et al., 2002b).

Of particular importance for the formulation of Lagrangian transport models is the vertical coordinate. The vertical coordinate is often treated differently in different models and wind fields. Instead of the geometrical height, often, physically favourable vertical coordinates are used to simplify equations or enable specific parameterizations. The requirement for an atmospheric parameter to be a vertical coordinate is to increase strictly monotonous with geometric height, which allows the unambiguous inversion between the height and the parameter. It is moreover conceptually favourable for vertical coordinates to follow the topography at the lower model boundary, although not strictly necessary (see also Brasseur and Jacob, 2017).

Coordinates as pressure or potential temperature can simplify equations in models. Pressure is decreasing strictly monotonous in the large-scale atmosphere, following the hydro-static equation (e.g. Brasseur and Jacob, 2017). If the pressure p is normalized to the respective local surface pressure p_0 (i.e. $\sigma = \frac{p}{p_0}$), the constructed coordinate is as well topography following. This kind of coordinate is referred to as σ-coordinate. However, since the σ levels are not re-

quired in the upper atmosphere hybrid sigma-pressure coordinates are used as well, for example in the Integrated Forecast Model (IFS) of the European Centre for Medium-Range Weather Forecasts (ECMWF). This so called hybrid eta coordinate was introduced by Simmons et al. (1989) as follows. N levels are defined in the atmosphere by pressure levels: $p_{k+\frac{1}{2}} = A_{k+\frac{1}{2}} + B_{k+\frac{1}{2}} p_s$, where p_s is the local surface pressure, k is the height index and $A_{k+\frac{1}{2}}$ and $B_{k+\frac{1}{2}}$ are parameters with $A_{\frac{1}{2}} = A_{N+\frac{1}{2}} = B_{\frac{1}{2}} = 0$ and $B_{N+\frac{1}{2}} = 1$. These parameters are chosen so that the hybrid eta coordinate is the sigma coordinate at lower levels and a pressure coordinate for higher levels.

The vertical velocity in pressure coordinates is given by the pressure tendency $\omega = \frac{\mathrm{d}p}{\mathrm{d}t}$. The pressure tendency is determined by the material in- and outflow into a vertical column. Vertical transport must balance the residual between the in- and outflow in a column to guarantee mass conservation (e.g. Brasseur and Jacob, 2017). Lagrangian transport calculations with pressure tendencies as vertical velocities are referred to as kinematic transport calculations.

The potential temperature is defined as $\theta(p) = T\left(\frac{p_0}{p}\right)^{\kappa}$ (κ is the ratio between the gas constant of air and the heat capacity at constant pressure. T is the temperature) and a suitable vertical coordinate for the stratosphere, where it is increasing monotonously with height (e.g. Brasseur and Jacob, 2017). However, at lower levels, the potential temperature is not necessarily monotonous anymore, for example at places of high vertical instability. Moreover, potential temperature at the surface varies substantially from place to place and hence is not topography following as well.

The vertical velocity in θ coordinates is given by the θ tendency. The θ tendency is proportional to the diabatic heating rate, and hence determined from the local energy balance at an atmospheric location (e.g. Ploeger et al., 2010; Brasseur and Jacob, 2017; Ploeger et al., 2021). In the stratosphere vertical transport processes are mostly driven by diabatic heating rates and otherwise the transport is isentropic (at a constant potential temperature), which increases the accuracy of the calculation with a potential temperature coordinate system.

Hybrid coordinates that reconcile the pressure based, and topography following sigma coordinate in the troposphere, with the isentropic theta coordinate system at higher levels can combine the benefits of both coordinates systems. Such a hybrid coordinate is the zeta coordinate as defined by equation:

$$\zeta(p) = \begin{cases} \theta(p) & \text{if } \sigma < \sigma_r \\ \theta(p) \sin\left(\frac{\pi}{2} \frac{1-\sigma(p)}{1-\sigma_r}\right) & \text{if } \sigma \geq \sigma_r \end{cases} \tag{3.4}$$

where p is the pressure and p_s denotes the local surface pressure (Mahowald et al., 2002;

Konopka et al., 2007). σ_r is a reference level in sigma coordinates. $\theta(p)$ is the potential temperature. Near the surface, the hybrid zeta coordinate follows the orography in the form of a sigma-like coordinate. At higher altitudes, the zeta coordinate is smoothly transformed into the potential temperature $\theta(p)$, which is reached at the reference level ($\sigma_r = 0.3$), . The reference level $\sigma_r = 0.3$ corresponds to a pressure around 300 hPa (≈380 K) depending on the local surface pressure.

Equation (3.5) shows that the time derivative of the hybrid zeta coordinate is the time derivative of the potential temperature, i.e. the diabatic ascent rate, at altitudes above the reference level σ_r. At lower levels, the transport is a combination of diabatic heating rates, and the vertical wind ω and the surface pressure tendency. While the diabatic heating rates are derived from the energy balance including among others radiation, latent heat and turbulent mixing, kinematic rates are calculated from the continuity equation (Konopka et al., 2007; Ploeger, 2010; Ploeger et al., 2021).

$$\dot{\zeta}(p) = \begin{cases} \dot{\theta}(p) & \text{if } \sigma < \sigma_r \\ \dot{\theta}(p) \sin\left(\frac{\pi}{2}\frac{1-\sigma(p)}{1-\sigma_r}\right) - \theta(p)\frac{\pi}{2}\cos\left(\frac{\pi}{2}\frac{1-\sigma(p)}{1-\sigma_r}\right)\frac{\dot{\sigma}(p)}{1-\sigma_r} & \text{if } \sigma \geq \sigma_r \end{cases} \tag{3.5}$$

Lagrangian transport calculations with tendencies in the hybrid zeta coordinate are also referred to as diabatic transport calculations. The benefit of this diabatic transport calculations have been demonstrated in many studies (e.g. Mahowald et al., 2002; Ploeger et al., 2010, 2011; Schoeberl and Dessler, 2011; Brinkop and Jöckel, 2019).

3.2 Integration schemes

Flow in the atmosphere is not accessible in the form of elementary mathematical functions and air parcels show often complex and chaotic trajectories. Therefore numerical schemes for approximate solution of the Eq. (3.1) are practically unavoidable (e.g. Stohl, 1998). To formulate a numerical scheme, Equation 3.1 can be reformulated with the fundamental theorem of calculus into its integral form for a time interval from t to $t + \Delta t$:

$$\vec{x}(t + \Delta t) = \int_t^{t+\Delta t} \vec{u}(\vec{x}(t'), t')dt' + \vec{x}(t) \tag{3.6}$$

The integral can be approximated numerically, by using the ansatz:

$$\int_t^{t+\Delta t} \vec{u}(\vec{x}(t'), t')dt' = \sum_{i=1}^{N} \alpha_i \vec{u}(\vec{x}(t_i), t_i)\Delta t \tag{3.7}$$

where $\vec{u}(\vec{x}(t_i), t_i)$ are wind velocities evaluated at selected nodes $(x(t_i), t_i)$ and the α_i are weights for each of the nodes. To guarantee that the sum approximates the integral, the weights must fulfill $\sum_i^N \alpha_i = 1$, as can be illustrated for example for a constant velocity field (e.g. Quarteroni et al., 2017).

To evaluate the integral, the function $\vec{u}(\vec{x}(t), t)$ must be evaluated at the times t_i. The velocity function $\vec{u}(\vec{x}(t), t)$ can be understood as a function $\vec{u}(t)$ of t, when the trajectory $\vec{x}(t)$ as a function of t is known. However, since the trajectory $\vec{x}(t)$ of the air parcel in the time period Δt needs to be calculated, the value of the velocity function needs to be approximated as well. In a simple case, this approximation can be done by assuming that the wind is almost constant during the time increment Δt along the trajectory. Then only one node can be constructed ($(\vec{x}(t), t)$) at the beginning, and it is chosen $\alpha_0 = 1$. This leads to the explicit Euler scheme:

$$\vec{x}(t + \Delta t) = \vec{u}(\vec{x}(t), t)\Delta t + \vec{x}(t) \tag{3.8}$$

The Euler scheme, however, has a second order local truncation error (e.g. Rößler et al., 2018). To build more precise integration schemes, more nodes within the integration period Δt are calculated iteratively. For example, one node can be set at the mid-time $\frac{\Delta t}{2}$. The position of the air parcel at this node ($\vec{x}(t+\frac{\Delta t}{2})$) is approximated with the Euler Method $\vec{x}(t+\frac{\Delta t}{2}) \approx \vec{x}(t)+\vec{u}(\vec{x}(t), t)\frac{\Delta t}{2}$. Then the integral over $\vec{u}(\vec{x}(t), t)$ can be approximated by the integral over the constant value $\vec{u}\left(\vec{x}(t) + \vec{u}(\vec{x}(t), t)\frac{\Delta t}{2}, t + \frac{\Delta t}{2}\right)$ at the mid of the time-period. Here is $\alpha_0 = 1$ as well. This leads to the mid-point scheme (or second order Runge-Kutta scheme):

$$\vec{x}(t + \Delta t) = \vec{u}\left(\vec{x}(t) + \vec{u}(\vec{x}(t), t)\frac{\Delta t}{2}, t + \frac{\Delta t}{2}\right)\Delta t + \vec{x}(t) \tag{3.9}$$

The mid-point scheme has a third order ($\mathcal{O}(\Delta t^3)$) truncation error and a second order accumulated error ($\mathcal{O}(\Delta t^2)$) (e.g. Rößler et al., 2018).

More iterations can be applied and combined to obtain higher order accuracy. The classical forth order Runge-Kutta method, combines four nodes (see equations 3.10 to 3.14). The first

node corresponds to the Euler method being placed at the beginning of the time period Δt. The second node corresponds to the mid-point method to calculate the position after the half time of Δt. The third node, is placed at the half time as well but correponds to an Euler step, based on the wind speed derived from the mid-point. Finally, a node is constructed at the time $t + \Delta t$, based on a step using the refined windspeed at the mid-point. The wind speeds at the four nodes are weighted with $\frac{1}{6}, \frac{1}{3}, \frac{1}{3}, \frac{1}{6}$. The Runge-Kutta method has fifths order truncation error ($\mathcal{O}(\Delta t^5)$) and a forth order accumulated error ($\mathcal{O}(\Delta t^4)$) (e.g. Rößler et al., 2018).

$$\vec{x}(t + \Delta t) = \vec{x}(t) + \frac{1}{6}\left(k_1 + 2k_2 + 2k_3 + k_4\right)\Delta t \tag{3.10}$$

$$\vec{k_1} = \vec{u}(t, \vec{x}(t)) \tag{3.11}$$

$$\vec{k_2} = \vec{u}\left(t + \frac{\Delta t}{2}, \vec{x}(t) + \Delta t \frac{k_1}{2}\right) \tag{3.12}$$

$$\vec{k_3} = \vec{u}\left(t + \frac{\Delta t}{2}, \vec{x}(t) + \Delta t \frac{k_2}{2}\right) \tag{3.13}$$

$$\vec{k_4} = \vec{u}\left(t + \Delta t, \vec{x}(t) + \Delta t\, k_3\right) \tag{3.14}$$

3.3 Grid interpolation

Practically, the wind field $\vec{u}(x, t)$ is only available discretized on a grid. This implies that to perform the Lagrangian transport calculations interpolation functions are needed. In many cases the interpolation is done linearly in time and space (Hoffmann et al., 2022; McKenna et al., 2002a; Stohl et al., 2005)

The linear interpolation in Cartesian coordinates is performed as follows. Let the field $Q(\phi, \lambda, \zeta)$ be given in the coordinates ζ, ϕ and λ. Consider the quantity Q, which needs to be interpolated at the air parcel position $(\zeta_{\mathrm{AP}}, \phi_{\mathrm{AP}}, \lambda_{\mathrm{AP}})$. Further, a grid box with the indices 000, 010, 100, 110, 001, 011, 101, 111 is defined, in which the air parcel can be found. The first index is in ϕ, the second is in λ and the third in ζ. Then the weights for the linear interpolation in Cartesian coordinates are given by:

$$\Omega_\phi = \frac{\phi_{AP} - \phi_0}{\phi_1 - \phi_0}, \qquad \Omega_\lambda = \frac{\lambda_{AP} - \lambda_0}{\lambda_1 - \lambda_0}, \qquad \Omega_\zeta = \frac{\zeta_{AP} - \zeta_0}{\zeta_1 - \zeta_0} \tag{3.15}$$

After linear interpolation in all three dimensions the equation for the interpolation in the box is:

$$\begin{aligned} Q(\zeta_{\mathrm{AP}}) = \Omega_\lambda \Omega_\phi \Omega_\zeta \left(Q_{111} - Q_{011} + Q_{100} - Q_{000} + Q_{001} + Q_{010} - Q_{110} - Q_{101}\right) \\ + \Omega_\phi \Omega_\zeta \left(Q_{101} + Q_{000} - Q_{001} - Q_{100}\right) \\ + \Omega_\lambda \Omega_\zeta \left(Q_{110} + Q_{000} - Q_{010} - Q_{100}\right) \\ + \Omega_\zeta \left(Q_{001} - Q_{000}\right) \\ + \Omega_\phi \left(Q_{010} - Q_{000}\right) \\ + \Omega_\lambda \left(Q_{100} - Q_{000}\right) \\ + Q_{000} \end{aligned} \tag{3.16}$$

This linear interpolation assumes that the spherical line elements only change negligible within the interpolated box. Consequently the transformation of Cartesian coordinates to spherical coordinates has to be completed separately by multiplying wind speeds with the spherical line elements. Furthermore, some models perform a logarithmic interpolation in the vertical direction to interpolate quantities like the pressure.

3.4 Parameterizations

The wind fields are only given at grid points. Hence, wind fields only represent spatio-temporal *averaged* velocities of the atmospheric flow. This makes the parameterization of unresolved processes such as convection, mesoscale winds and turbulent diffusion necessary. Turbulent diffusion is, for example, important as it allows transport across tracer concentration gradients even though the mean wind is perpendicular to the concentration gradient (e.g. Brasseur and Jacob, 2017) and sub-grid scale mesoscale winds determine the long-term dispersion of air in the flow (Stohl et al., 2005). Convective parameterization is needed because reanalysis wind velocities do not comprehensively resolve fast, deep updrafts that connect the boundary layer with the upper troposphere and lower stratosphere (Hoffmann et al., 2023b).

Turbulent diffusion

A conceptually simple approach to parameterize turbulent diffusion is the zeroth order formulation of particle dispersion (see Brasseur and Jacob, 2017; Hoffmann et al., 2022). It adds random perturbations to the air parcel positions, i.e.:

$$dx_i = u_i dt + \frac{\partial D_i}{\partial x_i} dt + \sqrt{2D_i \Delta t}\, \xi_i\,, \quad i \in 1, 2, 3 \tag{3.17}$$

(3.18)

with uncorrelated, standard normal distributed ξ_i, so that they have zero mean values and a standard deviation of 1. D_i is called the diffusivity. Without diffusivity this equation becomes the basic trajectory equation 3.1 again. Without the grid wind field u_i and with a constant diffusivity it describes a pure Wiener process. The approach leads to mean tracer concentrations equivalent to the Eulerian advection-diffusion equation (see Brasseur and Jacob, 2017, and references therein).

Sub-grid scale wind fluctuations

The first order formulation of particle dispersion adds perturbations to the wind fields instead of to the positions. The so called Langevin equations are:

$$du_i = a_i dt + \Sigma_{j=0}^{3} b_{ij} \xi_j\,, \quad i \in 1, 2, 3 \tag{3.19}$$

where a_i denotes the acceleration, $b_{i,j}$ is a matrix used to model the random pertubations and ξ_j is a random number defined as before for the zeroth order formulation (see also Brasseur and Jacob, 2017). The parameters a_i and b_{ij} depend on the considered turbulence. The pertubed wind field is further used to update the air parcel position. The Langevin equation was also used by Stohl et al. (2005) and Hoffmann et al. (2016) to model unresolved sub-grid scale winds u' on the mesoscale. The entire wind field can then be express as the sum of the sub-grid scale winds and the average grid wind. To determine the sub-grid scale winds the terms of the equation 3.19 were determined from the variance σ_i^2 of the eight grid winds around the air parcel for two consecutive time steps. Furthermore they were determined by the time steps of the model Δt and the data Δt_d. This approach correlates two consecutive time steps with each other and also scales grid scale wind variances down to sub-grid scales according to a parameter f. Equations

3.20 to 3.22 show the definition of the parameters and the final Langevin equation for the sub-grid scale fluctuations of the wind.

$$a_i = -2\frac{u_i'(t)}{\Delta t_d} \tag{3.20}$$

$$b_{ij} = \begin{cases} \sqrt{1 - \left(1 - 2\frac{\Delta t}{\Delta t_d}\right)^2}(f\sigma_i)^2, & \text{for } i = j \\ 0 & \text{for } i \neq j \end{cases} \tag{3.21}$$

$$u_i'(t + \Delta t) = u_i'(t)\left(1 - 2\frac{\Delta t}{\Delta t_d}\right) + \sqrt{1 - \left(1 - 2\frac{\Delta t}{\Delta t_d}\right)^2}(f\sigma_i)^2\xi \tag{3.22}$$

Convection

Finally, convection is not resolved in most of the data used to do trajectory calculations (Hoffmann et al., 2023b). Many parameterizations have been proposed (e.g. Brinkop and Jöckel, 2019; Wohltmann et al., 2019; Konopka et al., 2019; Gerbig et al., 2003) to add convection to Lagrangian transport simulations. Most parameterization have in common that they have, first, a trigger function, which starts the parameterization when certain criteria are fulfilled, and second, a random or fixed distribution function, that distributes the air parcels vertically. The vertical distribution can be derived from physical parameters, or be fixed based on physical assumption.

An introductory convection parameterization for Lagrangian transport models is the Extreme Convection Parameterization (ECP), introduced by Gerbig et al. (2003), and later implemented in more operational and research transport models (Loughner et al., 2021; Hoffmann et al., 2023b). The ECP was in particular developed and used to give an upper estimate of convective transport.

The trigger function of the ECP relies only on the convectively availiable potential energy (CAPE). CAPE is the integrated positive buoyancy between the height of free convection and the level of convective equilibrium. High CAPE values indicate high probability for convective activity. A threshold for the CAPE can be chosen. If this threshold is exceeded the parameterization is triggered. The vertical distribution function of the ECP relies on the assumption that convection lefts behind a well-mixed atmosphere, i.e. air parcels are distributed vertically and randomly following the air density, between the surface and the equilibrium level.

Chapter 4

Reanalysis data & models

4.1 ECMWF reanalysis

For this thesis, the full-resolution ERA5, downsampled low-resolution ERA5 and ERA-Interim reanalysis data are used to drive backward and forward trajectory calculations with CLaMS and MPTRAC. Both reanalyses have been developed by the ECMWF (Dee et al., 2011; Hersbach et al., 2020).

ERA-Interim is the precursor of ERA5 and is available for the years from 1 979 to 2 019. The ERA-Interim reanalysis offers six-hourly meteorological data at around 80 km horizontal resolution on 60 hybrid eta levels. It reaches from the surface up to 0.1 hPa. The assimilation system for ERA-Interim uses a four-dimensional variational analysis (4D-Var) with a 12 h time window and the ECMWF's IFS cycle 31r2 as released in 2 006.

The ERA5 reanalysis offers hourly meteorological data on a 30 km horizontal grid (0.3 °$\times$0.3 °) on 137 eta levels from the surface up to 80 km. The ERA5 reanalysis was processed with an improved model version compared to ERA-Interim (IFS cycle 41r2 with 4D-Var assimilation), including, among others, novel parameterizations of atmospheric waves and convection. The ERA5 reanalysis covers the period between 1 950 and the present. Increase of spatial and temporal resolution in ERA5 particularly improves the representation of tropical cyclones and convection in the reanalysis in comparison to ERA-Interim and other reanalysis data (e.g. Taszarek et al., 2020; Li et al., 2020; Malakar et al., 2020). ERA5 was also found to improve Lagrangian transport simulations in the free troposphere and stratosphere (Hoffmann et al., 2019).

The low-resolution ERA5 data set (also referred to as ERA5lr or ERA5 1 °$\times$1 ° in this thesis) was created by downsampling of the full-resolution data to a 1 °$\times$1 ° horizontal grid and 6 h time-

steps, applying a truncation to T213 as is specified in the ECMWF's Meteorological Archival and Retrieval System (MARS). The vertical levels of ERA5 were kept unchanged. Low-resolution ERA5 data was used in previous studies to benefit from the improvements of the ERA5 reanalysis but to avoid high computational costs and costs for handling the much larger amount of data compared to ERA-Interim (e.g. Ploeger et al., 2021).

4.2 The Lagrangian transport models MPTRAC and CLaMS

This thesis uses and improves two Lagrangian transport models, the Chemical Lagrangian transport Model of the Stratosphere (CLaMS) and Massive Parallel Trajectory Calculations (MPTRAC). Table 4.1 gives also a short overview over the two models.

CLaMS is a full chemical Lagrangian transport model that includes modules for irreversible mixing, chemistry and advection (McKenna et al., 2002a,b). In a recent CLaMS version (CLaMS v2.0) a convection parameterization was implemented as well. However, in this thesis, the focus is set to the advection module of CLaMS v1.0, which applies a 4th-order Runge-Kutta scheme for the trajectory calculations with a default integration time-step of 1800 s. The CLaMS model can be used in vertical hybrid zeta coordinates (e.g. Konopka et al., 2007; Ploeger et al., 2010; Pommrich et al., 2014; Ploeger et al., 2021). However, CLaMS can be used either with zeta or pressure coordinates, i.e. with diabatic or kinematic vertical velocities (e.g. Ploeger et al., 2010; Li et al., 2020).

MPTRAC (Hoffmann et al., 2022) is a Lagrangian transport model for the free troposphere and the stratosphere. It includes modules for advection, diffusion, and convection, which are applied in this thesis. In contrast to CLaMS, mixing is computed with two modules that parameterize turbulent diffusion and sub-grid scale wind fluctuations as have been presented in Chapter 3.4. In MPTRAC the parameterization for turbulent diffusion is restricted to the horizontal direction in the troposphere and to the vertical directions in the stratosphere, while the parameterization for sub-grid scale winds affects all directions (see also Hoffmann et al., 2022, for further details). Convection is parameterized with the extreme convection parameterization (ECP) with a CAPE threshold of $0\,\mathrm{J\,kg^{-1}}$. Moreover, the ECP is extended in MPTRAC, by using the convective inhibition (CIN), as an additional threshold, which should not be surpassed (Hoffmann et al., 2023b). The CIN is the energy that air parcels need to overcome when a stable layer below the level of free convection exists. The advection module uses the mid-point scheme for integration with a default time-step of 180 s, because it provides a good trade of between performance and accu-

Table 4.1: Overview over the internal properties of CLaMS and MPTRAC.

	CLaMS v1.0	MPTRAC v2.0	MPTRAC v2.6
Integration Schemes	classical 4th order Runge-Kutta	Euler, mid-point	Euler, mid-point, clas. 4th order Runge-Kutta
Vertical coordinate	pressure (kinematic) zeta (diabatic)	pressure (kinematic)	pressure (kinematic) zeta (diabatic) + modules in pressure
Interpolation	linear exact spherical non-invertible	linear cartesian approx. invertible	linear cartesian approx. invertible
Horizontal coordinate	spherical (for $\|\phi\| < 72°$ N/S), polar-stereographic	spherical	spherical
Earth Radius	6 371.000 km	6 367.421 km	6 367.421 km

racy (Rößler et al., 2018). MPTRAC v2.0 was formulated in pressure coordinates with kinematic vertical velocities driving the transport. MPTRAC v2.0 has been further developed for this thesis to use either pressure or zeta coordinates (kinematic or diabatic vertical velocities) for calculations of trajectories following the approach in CLaMS. The model improvements have been released with MPTRAC v2.6 (see also Hoffmann et al., 2023a, for the details of the code). The detailed integration of the hybrid zeta coordinate into MPTRAC is described in more detail in Chapter 4.4.

MPTRAC uses spherical coordinates to store the position of air parcels. CLaMS has a hybrid approach, with spherical coordinates for air parcels at latitudes between 72° S and 72° N, but otherwise uses a stereographic projection. The approach in CLaMS guarantees that the integration does not diverge near the poles. In MPTRAC, for air parcels very close to the pole (at latitudes larger than 89.999° North or South), the zonal transport is ignored to avoid numerical issues related to the singularities.

The two models also have some minor differences in constants such as the Earth's radius. In MPTRAC's default setting, the Earth's radius is assumed to be 6 367.421 km (Average between the pole radius and the equatorial radius), whereas in CLaMS it is 6 371.000 km (Globally averaged radius). This has implications for transformations between Cartesian and spherical coordinates.

4.3 High-performance Lagrangian transport modelling

State-of-the-art, high-performance Lagrangian transport calculations must overcome at least three major challenges related to the ongoing digital revolution, similar to most of the digital infrastructure of the earth system sciences (Bauer et al., 2021). First, they must be designed to run on recent and future exascale computing architectures, which likely will be based on Graphic Processing Units (GPUs) (e.g. Hoffmann et al., 2022). Second, with the emergence of "Big Data", as a consequence of the increasing resolution of wind data such as ERA5, the models need improvements of the data handling during the calculations, which involves Input/Output (IO) tuning and better memory management (see also Appendix B). Third, artificial intelligence (AI) may play an increasing importance for Lagrangian transport models in terms of Artificial Intelligence (AI) solvers, parameterization or super-resolution (e.g. Brecht et al., 2023) and likely will impact the future design of HPC architectures. The future model development of CLaMS and MPTRAC have to be discussed in this context.

The software architecture and development of CLaMS and MPTRAC follow two different approaches (see Table 4.2 for a brief summary). CLaMS, whose development started in the early 90s with the trajectory code (Sutton et al., 1994; McKenna et al., 2002b), was initially written in the Fortran programming language for computer architectures of the 1990s. Later CLaMS was adjusted and parallelized using the Message Passing Interface (MPI) standard. Currently, CLaMS can run on supercomputer clusters with Central Processing Unit (CPU) architectures, by parallel calculations of air parcel trajectories. A parallel calculation on GPUs is not possible. Up to now, CLaMS is not capable of using shared memory options on cluster nodes and must read the entire wind data for each process that works on a subset of air parcels redundantly. For large data such as ERA5, this leads to memory demands exceeding 100GB. Hence, CLaMS is neither capable of scaling air parcel numbers massively using GPUs, which would enable large enough air parcel ensembles to represent all the information found in the high-resolution ERA5 reanalysis, nor is it capable of handling the large data sets efficiently complicating even studies with smaller air parcel ensembles. Although some of these problems can be solved by code adaptation strategies (see Appendix B.2), the sustainability of this traditional development strategy for CLaMS is uncertain for future computing architectures.

MPTRAC in contrast to CLaMS is build in the C programming language with a co-development paradigm: Algorithms designed for MPTRAC are not only evaluated to physical accuracy but also optimized to exploit modern HPC architectures such as GPUs and the shared memory between processes (OpenMP, OpenACC) (Hoffmann et al., 2022). With its MPI-OpenMP-OpenACC hy-

brid parallelization MPTRAC is flexible in use. On a single compute node the computation is parallelized by distributing the air parcels to the available number of threads. With MPI the air parcels can be distributed even further among multiple, independent nodes. If a GPU booster is available, MPTRAC can use it by offloading computationally intensive calculations. Further tuning options, such as asynchronous IO, efficient packing of data and optimized data structures are under development right now. Hence, MPTRAC is well suited to make the full information given in high-resolution data available by massive air parcel ensembles, and also can handle large data sets better. Moreover, the performance benefits of MPTRAC enable to run a larger number of scenarios with different configurations in comparison to CLaMS.

Albeit MPTRACs computational benefits over CLaMS' advection kernel, MPTRAC v2.0 still lacks capabilities such as simulating additional transport processes (e.g. mixing) as represented in CLaMS and a formulation in zeta coordinates with diabatic vertical velocities. Both capabilities are of importance to run state-of-the-art transport calculations in the atmosphere comparable to those of CLaMS. To overcome this gap, in this thesis, MPTRAC v2.6 has been developed which has a diabatic transport kernel similar to CLaMS and increased data interoperability to CLaMS. In the future, CLaMS and MPTRAC are supposed to be integrated into one Lagrangian transport library, as a strategy to allow a gradual transition to exascale applications without disrupting community workflows.

Table 4.2: Programming details of CLaMS and MPTRAC trajectory code.

	CLaMS v1.0	MPTRAC v2.0
Parallelisation	CPUs (MPI)	CPUs or GPUs (MPI/OpenMP/OpenACC)
Shared memory	no	yes
Language	Fortran	C

4.4 Implementation of diabatic transport into MPTRAC

For the integration of the diabatic transport scheme into MPTRAC, MPTRAC was equipped with functions to read the vertical velocities of the hybrid zeta coordinate ($\dot{\zeta}$) from files that follow the data structure of CLaMS. Moreover, during the integration time steps, the horizontal wind and vertical velocity must be interpolated to the air parcel locations. Therefore, a new interpolation function for MPTRAC was implemented.

For the Runge-Kutta method, wind fields must be interpolated four times to the given time, horizontal location, and zeta height. For the mid-point scheme, this is reduced to two interpolations. For MPTRAC and CLaMS four-dimensional linear interpolation methods are performed, which are common for Lagrangian transport models (Bowman et al., 2013). However, the specific details of the interpolation in CLaMS and MPTRAC differ because the wind fields are not regularly provided in hybrid zeta coordinates, but in hybrid eta coordinates as applied in ECMWF's Integrated Forecasting System (Simmons et al., 1989). Interpolation with positions given only in zeta coordinates therefore requires additional considerations. In addition, MPTRAC has modules that rely on a formulation in pressure coordinates, requiring frequent conversions of the air parcel position from pressure to zeta and vice versa. For example, the air parcel position given in the zeta coordinate is updated during the advection time-step. Afterwards, the updated position in zeta coordinates is converted to pressure. The air parcel position given in pressure coordinates can then be updated by using a module such as for turbulent diffusion, which adds a random increment to the air parcel position in pressure coordinates. Afterwards, the air parcel position in pressure needs to be transformed back to the zeta coordinates again for use with the diabatic advection. A further difference is that time interpolation is performed locally for each air parcel in MPTRAC. In contrast, CLaMS interpolates the wind field globally in advance for the four time steps of the Runge-Kutta scheme.

For the Runge-Kutta method, wind fields must be interpolated four times to the given time, horizontal location, and zeta height. For the mid-point scheme, this is reduced to two interpolations. For MPTRAC and CLaMS four-dimensional linear interpolation methods are performed, which are common for Lagrangian transport models (Bowman et al., 2013). However, the specific details of the interpolation in CLaMS and MPTRAC differ because the wind fields are not regularly provided in hybrid zeta coordinates, but in hybrid eta coordinates as applied in ECMWF's Integrated Forecasting System (Simmons et al., 1989). Interpolation with positions given only in zeta coordinates therefore requires additional considerations, e.g. about how to find the vertical position of the box that includes an air parcel, when the data is not stored in the air parcels coordinate. In addi-

tion, MPTRAC has modules that rely on a formulation in pressure coordinates, requiring frequent conversions of the air parcel position from pressure to zeta and vice versa. For example, the air parcel position given in the zeta coordinate is updated during the advection time-step. Afterwards, the updated position in zeta coordinates is converted to pressure. The air parcel position given in pressure coordinates can then be updated by using a module such as for turbulent diffusion, which adds a random increment to the air parcel position in pressure coordinates. Afterwards, the air parcel position in pressure needs to be transformed back to the zeta coordinates again for use with the diabatic advection. Therefore, the performed interpolations are required to be precisely invertible. A further difference is that time interpolation is performed locally for each air parcel in MPTRAC, e.g. wind data is collected around the position of the air parcel and subsequently interpolated in time. In contrast, CLaMS interpolates the wind field in time and globally in advance for the four time steps of the Runge-Kutta scheme, e.g. the entire wind data field is interpolated in time and subsequently used for all air parcels.

As a consequence of the mentioned differences between the models, the interpolations of CLaMS and MPTRAC follow two different concepts. Figure 4.1 illustrates the two concepts in two-dimensional space (height vs. longitude) in a simplified case. Two neighbouring vertical profiles of pressure and zeta are selected in eta coordinates. The goal is to interpolate from zeta to pressure and back to zeta. For simplicity, it is assumed that each eta level has constant pressure levels. Then, in CLaMS the interpolation begins with a vertical interpolation along the two profiles. For this step, the vertical position of the air parcel is identified along each vertical profile separately using the height of the air parcel in zeta (ζ_{AP}). As a consequence the pressure data ($p_{\mathrm{AP},1}, p_{\mathrm{AP},2}$) of two different eta levels is collected for final horizontal interpolation, provided that the zeta profiles vary strong enough from location to location. With the final horizontal linear interpolation, the pressure at the air parcel position is given. However, if this pressure position is used to interpolate back to the zeta coordinate again, which is required for MPTRAC, the identified vertical location from the pressure height, might differ from the vertical position in zeta height. To illustrate this issue, Fig. 4.1a describes the case where pressure levels agree with the eta levels. Hence, with one single pressure provided as the air parcel position, only one box - the box with index i where $p_{i+1} \leq p_{AP} < p_i$ - will be selected for the interpolation back to zeta. Since the data used for the interpolation from zeta to pressure (data from multiple eta levels), does not agree with the data used for the interpolation from pressure to zeta (data from one eta level), the interpolation is not reversed accurately.

To overcome this issue, MPTRAC instead starts with the horizontal interpolation of the zeta

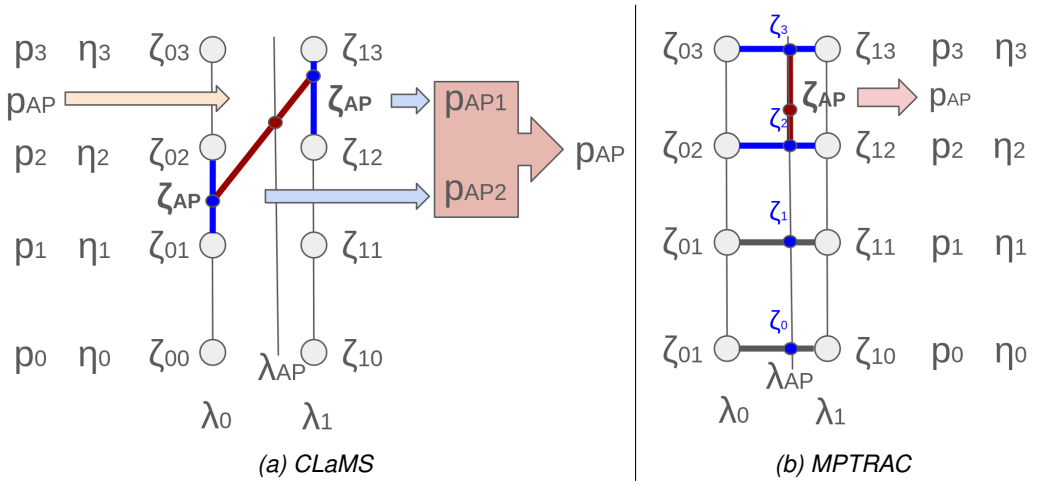


Figure 4.1: Concept of the interpolation illustrated in two dimensions for (a) CLaMS and (b) MPTRAC. Small circles indicate grid points, where the zeta and pressure values are given. Blue lines indicate the direction of the first interpolation and red lines indicate the direction of the second interpolation.

values and pressure values according to the horizontal air parcel position (λ_{AP}) at every eta level. This is depicted simplified in Fig. 4.1b. The procedure provides a vertical profile of pressure (p_i) and zeta (ζ_i) centred at the horizontal position of the air parcel. Along this profile, the unique box can be found that contains the air parcel in both coordinates. This profile can then also be reversed exactly by linear vertical interpolation. To avoid interpolation of zeta and pressure at all eta levels, the right height index is found by an iterative method. The exact interpolation of both models is described in the following paragraphs.

Figure 4.2 illustrates the interpolation as implemented in CLaMS (also referred to as interpolation "V0" in this thesis). Let ζ_{ijkl} be the zeta coordinate and Q_{ijkl} a quantity which is supposed to be interpolated to the position of the air parcel. Both, the coordinate and the quantity are required to be formulated in a hybrid eta coordinate. In detail, the indices i, j, k refer to the indices on the three dimensional grid in longitude λ_i, latitude ϕ_j and the vertical hybrid eta coordinate η_k. The index l refers to the time t_l. Furthermore, let $(\lambda_{\mathrm{AP}}, \phi_{\mathrm{AP}}, \zeta_{\mathrm{AP}}, t_{\mathrm{AP}})$ be the position and time of the air parcel to which the quantity Q_{ijkl} needs to be interpolated. At the beginning of the interpolation in CLaMS, the interpolation in time is performed. For this purpose the neighbouring times t_0 and t_1 are selected so that $t_0 \leq t_{\mathrm{AP}} < t_1$ (see Fig. 4.2 (1)). With the data from the neighbouring times a linear interpolation of ζ_{ijkl} and Q_{ijkl} is done to the time t_{AP} (2). This provides three dimensional fields ζ_{ijk} and Q_{ijk} (3). Then, the horizontal indices of the air parcel are determined $(i_{\mathrm{AP}}, j_{\mathrm{AP}})$ using the horizontal coordinates λ_{AP} and ϕ_{AP} and the horizontal grid of longitudes λ_i and latitudes ϕ_j (4). The indices define a column which includes the air parcel

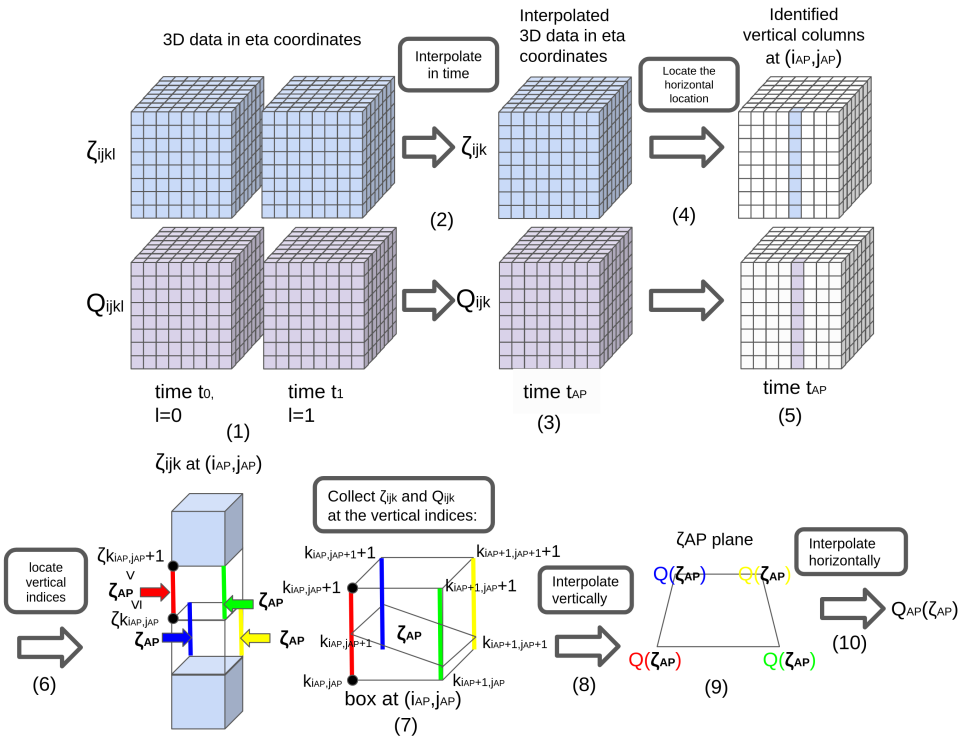


Figure 4.2: Schematic steps during interpolation V0 of a quantity Q to the air parcel position in zeta coordinates in CLaMS. For further details see the text.

(5). Subsequently, within this column, four vertical indices are determined, by locating the indices $(k_{i_{\mathrm{AP}},j_{\mathrm{AP}}}, k_{i_{\mathrm{AP}}+1,j_{\mathrm{AP}}}, k_{i_{\mathrm{AP}},j_{\mathrm{AP}}+1}, k_{i_{\mathrm{AP}}+1,j_{\mathrm{AP}}+1})$ with $\zeta_{k_{i_{\mathrm{AP}},j_{\mathrm{AP}}}} \leq \zeta_{\mathrm{AP}} < \zeta_{k_{i_{\mathrm{AP}},j_{\mathrm{AP}}}+1}$ etc., along the four edges of the column (6). Then, at these four vertical indices and the indices one level higher, the values of ζ_{ijk} and Q_{ijk} are collected to define a box for the interpolation (7). In this box the quantity Q_{ijk} is first interpolated vertically four times to the respective ζ_{AP} (8). Now, the quantity Q_{ijkl} is given on the four corners of the plane with $\zeta = \zeta_{\mathrm{AP}}$ (9). Finally, the quantity is interpolated horizontally, taking into account the line elements of the spherical coordinates (10). This provides $Q(\lambda_{\mathrm{AP}}, \phi_{\mathrm{AP}}, \zeta_{\mathrm{AP}}, t_{\mathrm{AP}})$.

The interpolation from pressure to zeta and from zeta to pressure is particularly important when coupling geophysical modules that operate with pressure as vertical coordinate (e.g. convection, diffusion, and sedimentation), as is the case for MPTRAC. The precise and accurate inversion of the interpolation in CLaMS from pressure back to zeta coordinates is difficult because during step (6) height indices can be found from the pressure that are inconsistent with height indices found using the zeta coordinate positions. If a different box is used for re-interpolation to zeta, significant errors may occur, making this approach unsuitable for frequent transformations between zeta and pressure coordinates. Consequently, a fully reversible interpolation algorithm has been developed for MPTRAC to allow the coupling of pressure-based modules with the dia-

batic advection scheme, where frequent vertical coordinate inversions are required.

Figure 4.3 shows a schematic of the interpolation in MPTRAC (which is also referred to as interpolation "V2", while it is referred to the original interpolation of MPTRAC with "V1"). With the same definitions as for the interpolation of CLaMS, the interpolation in MPTRAC can be described as follows. The interpolation starts as well by selecting the data of ζ_{ijkl} and Q_{ijkl} for the neighbouring times, i.e. t_0 and t_1 (see Fig. 4.3 (1)). Then, the horizontal indices of the air parcel are determined $(i_{\mathrm{AP}}, j_{\mathrm{AP}})$ (2). The indices define two columns which include the air parcel at the times t_0 and t_1 (3). Consequently, for each of this columns, four vertical indizes are determined, by locating the indices $(k_{i,j}, k_{i+1,j}, k_{i,j+1}, k_{i+1,j+1})_{t_0}$ and $(k_{i,j}, k_{i+1,j}, k_{i,j+1}, k_{i+1,j+1})_{t_1}$, along the eight edges of the two columns, analogous to the procedure in CLaMS (4). However, afterwards the minimum and maximum index k_{min} and k_{max} among the vertical indices from both times are determined (5). The minimum index and maximum index define the start and end point of an iteration that locates the box that contains the air parcel in vertical direction. The iteration starts with the temporal and horizontal interpolation of ζ_{ijkl} at the bottom and top of a box, which is defined by the minimum vertical index k_{min} and the spatial indices $(i_{\mathrm{AP}}, j_{\mathrm{AP}})$ (see Fig. 4.3 (6) and (7)). After the interpolation, ζ is given at the top ζ_{top} and the bottom ζ_{bottom} of the box (8). If ζ_{AP} is lower than ζ_{top} and equal or higher than ζ_{bottom}, the iteration finishes. Otherwise, the iteration proceeds by going to the next higher index until the right box is found. Because of the strictly monotonic increase of ζ_{ijkl} with height, it is guaranteed that the right box is found between the minimum and maximum vertical indices. However, when the right box is found, the quantity Q_{ijkl} is interpolated temporally and horizontally as well to the top Q_{top} and Q_{bottom} of the correct box (9), analogous to the interpolation of ζ_{ijkl} in (6) and (7). Finally, the vertical interpolation is performed linearly by using the quantity Q_{ijkl} and the coordinate ζ_{ijkl} from the top and bottom of the box and the zeta coordinate (ζ_{AP}) of the air parcel (9). This provides $Q(\lambda_{\mathrm{AP}}, \phi_{\mathrm{AP}}, \zeta_{\mathrm{AP}}, t_{\mathrm{AP}})$. If Q_{ijkl} is a vertical coordinate, such as pressure, the interpolation can be reversed as the vertical indices in Q_{ijkl} can also be determined in step (4) from the respective vertical Q_{ijkl} profiles.

The algorithm in MPTRAC allows precise interpolation from zeta to pressure and back to zeta, because the vertical column at the horizontal position of the air parcel gives a monotone relationship between zeta and pressure. In particular, the processing of pressure and zeta is analogous with opposite roles. The vertical 1D linear interpolation at the final step (9) can be performed accurately and unambiguously.

For comparison and error estimations, a third interpolation variant was implemented into MPTRAC, that closer resembles the interpolation in CLaMS (called interpolation "V3"). In this

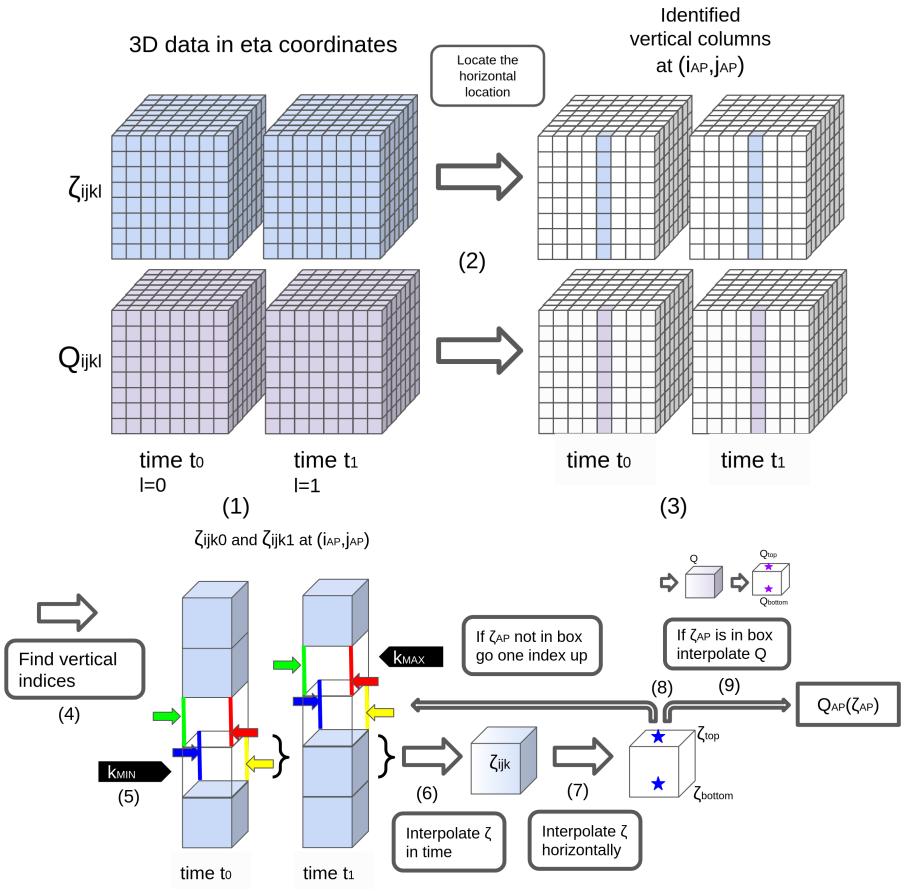


Figure 4.3: Schematic steps during interpolation V2 of a quantity Q to the air parcel position in zeta coordinates in MPTRAC. For further details see the text.

approach, the interpolation procedure follows the first steps (1) to (5) as defined in V2 and Fig. 4.3, respectively. Afterwards, however, the two profiles given in step (5) are interpolated in time along the four edges between $k_{\min}$ and $k_{\max}$, so that the locating of the vertical indices and the interpolation on a zeta plane can be finally done as for CLaMS (See Fig. 4.2, steps (6) to (10)).

However, note that all interpolations in MPTRAC are performed in Cartesian coordinates, i.e. the line elements of the spherical coordinate system are not applied during interpolation but afterwards to the final air parcel positions, assuming that the differences of the line elements within a grid box are negligible. The transformation from Cartesian coordinates to spherical coordinates is done separately from the interpolation process, by applying the equations $\Delta\lambda = \frac{\Delta x}{R_e \cos\phi}$ and $\Delta\phi = \frac{\Delta y}{R_e}$. $\Delta x, \Delta y$ denote the changes in Cartesian coordinates, $\Delta\phi, \Delta\lambda$ the change in spherical coordinates and R_e the Earth radius. These transformations are not applied in CLaMS because interpolation is already done in spherical coordinates. Another remaining difference between interpolation in CLaMS and MPTRAC is that the time interpolation is done for each air parcel separately in MPTRAC instead of the full meteorological field as in CLaMS.

Finally, pressure is interpolated logarithmic in CLaMS for zeta levels higher than 1 000 K, and linear for levels below 500 K. In between those levels, the linear and logarithmic interpolations are combined. In contrast, MPTRAC uses linear interpolation for pressure on all hybrid zeta levels.

Chapter 5

Evaluation of diabatic transport in MPTRAC v2.6

The implementation of the diabatic transport scheme in MPTRAC, used with the ERA5 reanalysis, is evaluated by a detailed intercomparison with CLaMS trajectory calculations for a global ensemble of air parcels. To put the differences found in the trajectory calculations between CLaMS and MPTRAC in a broader context, the effects of, first, external sources (using different reanalyses, resolutions and vertical velocities) and, second, internal sources (e.g. interpolation and integration methods) were investigated. The inter-comparison of uncertainties from models and reanalysis enables to address the guiding question: *What are the transport uncertainties in Lagrangian transport calculations of the upper troposphere and the stratosphere with the latest generations of reanalyses (ERA5 and ERA-Interim), vertical transport schemes (kinematic and diabatic) and models (MPTRAC and CLaMS)?*.

Uncertainty sources of Lagrangian transport models have been studied extensively in the past (e.g. Stohl, 1998; Stohl et al., 2001; Bowman et al., 2013). Uncertainty sources in transport simulations can be distinguished into external and internal sources. External uncertainties are related to the data driving the model, e.g. to the reanalysis used, differences between reanalysis products and the limited resolution of the wind data. Internal uncertainty sources are the necessary elements of the transport model, e.g. interpolation, integration methods or the handling of model boundaries at the surface.

External uncertainties of Lagrangian transport simulations due to differences between the used wind data are discussed frequently (e.g. Ploeger, 2010; Angevine et al., 2014; Hoffmann et al., 2019; Li et al., 2020; Ploeger et al., 2021; Vogel et al., 2023). First of all, limited resolution of the reanalysis fields itself creates a limitation for the accuracy of the transport calculations, because sub-grid scale processes are not accounted for without parameterization (e.g. Rolph and Draxler, 1990; Stohl et al., 2005). The stochastic parameterizations that are required to ac-

count for unresolved sub-grid scale winds and turbulent diffusion impose an uncertainty to the transport as well. Second, reanalysis fields show systematic differences because of different dynamical cores, assimilation processes, resolution and parameterizations if compared with each other. Hoffmann et al. (2019) showed that systematic differences due to the chosen reanalysis (comparing ERA5 and ERA-Interim) are larger than transport deviations due to parameterized sub-grid scale wind fluctuations and turbulent diffusion in *kinematic* transport calculations. Furthermore, Stohl et al. (2004) noted that inconsistencies of reanalysis data, which are caused by separate assimilation cycles, lead to artificial diffusion in Lagrangian transport calculations. Therefore, quantities such as potential vorticity (PV) or potential temperature are less conserved than physically expected. These inconsistencies are however absent in forecast data and might depend on the assimilation method of a selected reanalysis. Finally, diabatic and kinematic velocities derived from reanalysis data show systematic differences from reanalysis to reanalysis, related to the general circulation and the cross-isentropic dispersion of air parcels (e.g. Hoffmann et al., 2019; Li et al., 2020; Ploeger et al., 2021; Vogel et al., 2023). In summary, systematic differences of the reanalyses and their underlying models are expected to be a major source of external uncertainty for Lagrangian transport simulations, as well as processes that are not included in the reanalysis data (e.g. unresolved sub-grid scale processes).

Internal uncertainties related to different integration methods applied in MPTRAC have been investigated by Rößler et al. (2018). They found that the Euler method has about one order of magnitude higher error growth rates in comparison to the mid-point scheme in the stratosphere. However, the mid-point scheme is only two to four times less accurate than third and fourth order Runge-Kutta schemes, with no significant differences between the third and fourth order schemes. Rößler et al. (2018) attribute the latter to the errors related to linear interpolation of the meteorological data that limits benefits of higher order integration methods such as the fourth order Runge-Kutta scheme. Interpolation errors, if higher order integration is applied, could be the main internal source of error for deviations between Lagrangian transport models. Uncertainties as a consequence of interpolation have also been discussed in more detail by Stohl et al. (1995, 2001). Their results suggest interpolation and the integration scheme as the leading internal sources of uncertainty.

Differences between transport models have been studied as well. Differences in transport using different Lagrangian models (MPTRAC, CLaMS) driven by *kinematic* vertical velocities are smaller than differences caused by parameterized sub-grid scale winds and turbulent diffusion (Hoffmann et al., 2019). Stohl et al. (2001) concluded, based on a comparison of three trajectory

models, that the selection of the data is more important than the selection of the model for accuracy. In the literature (see also Stohl et al., 2001; Bowman et al., 2013), meteorological data are consequently considered the main source of uncertainty in Lagrangian transport simulations, while model differences, mainly due to interpolation and integration methods, are usually much smaller. Here, these findings for the two most recent ECMWF reanalysis ERA-Interim and ERA5 with CLaMS and MPTRAC were validated.

To justify that MPTRAC and CLaMS trajectory calculations can mutually substitute each other, the MPTRAC and CLaMS model do not to need to be bit-identical but deviations must be much smaller than from external uncertainty sources, e.g. reanalysis differences, vertical velocities and diffusion and on the order of combined internal uncertainties. In the following chapters, the model differences are evaluated, starting from case studies, going to a comparison between trajectories after one day, and ending with a long-term simulation of particle distributions. Finally, the conclusion is presented, that differences between CLaMS and MPTRAC trajectory calculations (as a consequence of internal sources) are indeed negligible in comparison to the variability of the results caused by external sources. Furthermore, it is shown that diabatic transport calculations with the fully resolved ERA5 reanalysis cannot be replaced by kinematic calculations or calculations with a larger grid width without significant differences, which corroborates the transition to optimized models with diabatic, fully resolved calculations.

5.1 Diagnostics to evaluate the diabatic transport in MPTRAC

For the evaluation of the newly implemented diabatic scheme in MPTRAC, a model initialization with about 1.4 million globally distributed trajectory seeds was used. The forward calculations are calculated for the boreal summer (June, July, and August). Short-term calculations of 1 day are initialized at the first of July 2 016, while the long-term calculations of 90 days are started on the first of June 2 016 to cover the entire boreal summer and austral winter. Seasonal differences are taken into account by separately analyzing the Northern and Southern Hemispheres. The air parcels are distributed horizontally, quasi-homogeneously, so that they have an average mutual distance of about 100 km. Vertically, they are distributed in specific layers. The layers are constructed such that each air parcel represents the same amount of entropy in the atmosphere, which is a product of density and the logarithm of the potential temperature (Konopka et al., 2007). For this reason, most air parcels are initialized around the tropopause, where the entropy of the atmosphere is largest. However, the air parcels cover a total zeta range from 30 K (about 1 km)

Table 5.1: Overview of different simulation scenarios for transport calculations with MPTRAC and CLaMS.

label	reanalysis	model	time-step	integration method	diffusion	inter-polation	Earth radius	vertical velocity	other options
CLaMS-default	ERA5	CLaMS	1800 s	Runge-Kutta 4	no	V0	6371000 m	dia.	
CLaMS-def-ERA5 1°×1°	ERA5 1°×1°	CLaMS	1800 s	Runge-Kutta 4	no	V0	6371000 m	dia.	
CLaMS-no-pole	ERA5	CLaMS	1800 s	Runge-Kutta 4	no	V0	6371000 m	dia.	polar coordinate off
MPTRAC-bestfit	ERA5	MPTRAC	1800 s	Runge-Kutta 4	no	V3	6371000 m	dia.	
MPTRAC-bestfit-Re	ERA5	MPTRAC	1800 s	Runge-Kutta 4	no	V3	6367421 m	dia.	
MPTRAC-int	ERA5	MPTRAC	1800 s	Runge-Kutta 4	no	V2	6367421 m	dia.	
MPTRAC-int-180s	ERA5	MPTRAC	180 s	Runge-Kutta 4	no	V2	6367421 m	dia.	
MPTRAC-default	ERA5	MPTRAC	180 s	mid-point	no	V2	6367421 m	dia.	
MPTRAC-def-kin	ERA5	MPTRAC	180 s	mid-point	no	V1	6367421 m	kin.	
MPTRAC-def-diff	ERA5	MPTRAC	180 s	mid-point	yes	V2	6367421 m	dia.	coupled mode
MPTRAC-def-ERA5 1°×1°	ERA5 1°×1°	MPTRAC	180 s	mid-point	no	V2	6367421 m	dia.	
MPTRAC-def-erai	ERA-Interim	MPTRAC	180 s	mid-point	no	V2	6367421 m	dia.	
MPTRAC-def-erai-kin	ERA-Interim	MPTRAC	180 s	mid-point	no	V1	6367421 m	kin.	
MPTRAC-cpl	ERA5	MPTRAC	180 s	mid-point	no	V2	6367421 m	dia.	coupled mode

to about 2 000 K (about 48 km). Setups similar to the one used here are often used to initialize transport calculations with CLaMS for studies in the UTLS, and in particular, are constructed to fit the hybrid zeta coordinates and mixing concept in CLaMS (e.g. Konopka et al., 2007; Pommrich et al., 2014; Vogel et al., 2015, 2019). Additionally, air parcels that reach the lower model boundary ($\zeta = 0$) are terminated in CLaMS. For comparison only, the same concept was applied in MPTRAC, too.

Different simulation scenarios are employed to put the deviations of the two models into the perspective of known uncertainty sources. Table 5.1 presents the scenarios, where different components of the transport calculations, such as the interpolation, integration, earth radius, coordinate systems, reanalysis, resolution, diffusion parameterization and the vertical velocity are varied. By comparing these scenarios, uncertainties from different sources can be estimated. Table 5.2 summarizes the different scenario intercomparisons and the related exposed uncertainty sources.

The first block of uncertainties in Table 5.2 describes the internal uncertainties. Internal sources for model uncertainties are based on the model code of a Lagrangian transport model itself, such as the vertical coordinate transformation or the integration scheme. These uncertainties are not estimated by comparing two different models but by comparing two different configurations of the same model and hence give an indication of the order of magnitude of the uncertainty already present within a model. A combination of all internal uncertainty sources within MPTRAC (see labels “interpolation”, “integration scheme”, “earth radius” and “time-step”) is as well investigated (see label “combined internal uncertainty”).

Table 5.2: Scenario intercomparisons for the estimation of different uncertainties in the Lagrangian transport calculations. Two scenarios are compared (base and comparative scenario) for the estimation. In most cases only one aspect of the model set-up is varied. The first block focuses on internal uncertainties of CLaMS and MPTRAC separately. The second block focuses on the comparison of the two models. The third block focuses on the external uncertainties. The last block describes the "transport", i.e. the difference of the start to the end point of a trajectory. "transport" is not an uncertainty source but it is a useful quantity for intercomparison with the uncertainties.

uncertainty source	base scenario	comparative scenario	Difference
p-zeta-p transformation	MPTRAC-default	MPTRAC-cpl	Coupled vs. uncoupled mode
Integration scheme	MPTRAC-default	MPTRAC-int-180s	Only vary integration scheme
Time-step	MPTRAC-int	MPTRAC-int-180s	Only vary between 1800s and 180s time steps
Interpolation	MPTRAC-bestfit-Re	MPTRAC-int	Only vary interpolation method
Polar coordinates	CLaMS-default	CLaMS-nopoles	Only vary polar coordinate switch
Earth radius	MPTRAC-bestfit	MPTRAC-bestfit-Re	Only vary used Earth radius
Combined internal	MPTRAC-default	MPTRAC-bestfit	Combined internal uncertainty of MPTRAC
Model default	MPTRAC-default	CLaMS-default	Compare default setup of models
Model default 1°×1°	MPTRAC-def-1°×1°	CLaMS-def-1°×1°	Compare default setup of models at lower resolution
Model bestfit	MPTRAC-bestfit	CLaMS-nopoles	Compare closest setup of models
Diffusion	MPTRAC-default	MPTRAC-def-diff	Only vary usage of diffusion and sub-grid scale wind module
Downsampling	MPTRAC-default	MPTRAC-def-ERA5 1°×1°	Only vary ERA5 to ERA5 1°×1°
Reanalysis	MPTRAC-default	MPTRAC-def-erai	Only vary ERA5 to ERA-Interim
Vertical velocity	MPTRAC-default	MPTRAC-def-kin	Vary vertical velocity
Transport	MPTRAC-default	MPTRAC-default	Compare end position with start positions

Model differences ("Model default", "Model default 1°×1°" and "Model bestfit") are the combination of uncertainties between two models and are listed in the second block of Table 5.2. Often, the sources that cause the model uncertainties are not known. The model uncertainties can be caused by the estimated internal uncertainties if the models also differ in the methods used. However, additional sources of uncertainties are possible. For example, the interpolation methods between MPTRAC and CLaMS vary more than can be estimated from the variation in the interpolation methods implemented in MPTRAC. While the interpolation in MPTRAC is always in Cartesian coordinates, CLaMS uses spherical coordinates.

External uncertainties are used to show the significance of calculations with the fully resolved ERA5 and diabatic vertical velocities. Therefore, different reanalysis products such as ERA5, ERA-Interim and ERA5 1°×1° are compared. In addition, the vertical velocity approach and the uncertainty from unresolved sub-grid scale winds are investigated in the scenarios.

For the intercomparison of the different model scenarios, a set of frequently used diagnostics were applied (e.g. Stohl et al., 1995; Hoffmann et al., 2019). Let i and j denote the indices of two trajectories with the same initial position derived in two different scenarios, and t the time at which the comparison is done. Then the air-parcel-wise absolute vertical transport deviation (AVTD) at

a given time t in the vertical zeta coordinate is

$$\mathrm{AVTD}_\zeta = |\zeta_i(t) - \zeta_j(t)|. \tag{5.1}$$

The absolute deviation in vertical direction quantifies the differences between individual air parcels. If kinematic calculations are compared with diabatic calculations, the zeta coordinates are calculated from temperature, surface pressure and the air parcels pressure according to equation 3.4 for both calculations. Otherwise the zeta coordinate is directly given.

The log-pressure altitude is defined as $Z = H \log \frac{p_0}{p}$, where $p_0 = 1013.25$ hPa and $H = 7.0$ km. Then, the air-parcel-wise AVTD in log-pressure altitude is:

$$\mathrm{AVTD}_Z = |Z_i(t) - Z_j(t)|. \tag{5.2}$$

To calculate the air-parcel-wise absolute horizontal transport deviation (AHTD) the equation

$$\mathrm{AHTD} = \sqrt{(x_i(t) - x_j(t))^2 + (y_i(t) - y_j(t))^2 + (z_i(t) - z_j(t))^2}, \tag{5.3}$$

is used, where (x_i, y_i, z_i) and (x_j, y_j, z_j) are the positions of the air parcels in Cartesian coordinates. The Euclidean distance approximates the great-circle distance for distances up to 5000 km with high precision (Hoffmann et al., 2019). For larger deviations, i.e. in calculations longer than 1 day, the great-circle distance itself is used as the air-parcel-wise AHTD:

$$\mathrm{AHTD} = R_e \arccos\left(\sin\phi_1 \sin\phi_2 + \cos\phi_1 \cos\phi_2 \cos|\lambda_1 - \lambda_2|\right) \tag{5.4}$$

where ϕ_i, ϕ_j are the latitudes, and λ_i, λ_j are the longitudes of the air parcels. R_e is the Earth radius.

To measure the conservation error of a quantity Q such as potential temperature at time t, the air-parcel-wise relative tracer conservation error (RTCE) is used,

$$\mathrm{RTCE} = 2\frac{|\,Q(t) - Q(0)\,|}{|\,Q(t)\,| + |\,Q(0)\,|}. \tag{5.5}$$

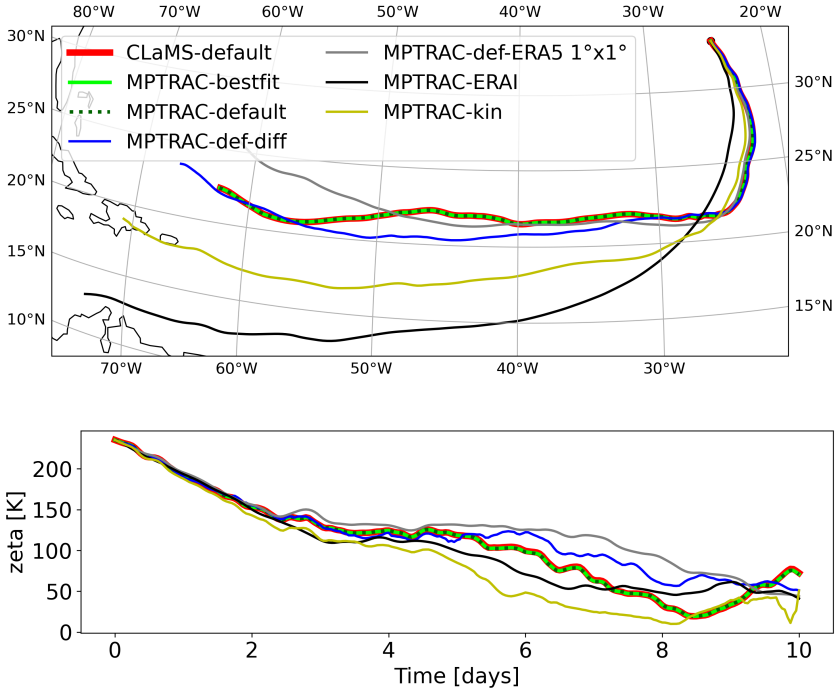
Individual trajectories of air parcels can substantially deviate between the scenarios defined in Table 5.1. Statistics such as quantiles, means, and medians of the different air-parcel-wise diagnostics for about 1.4 million air parcels are considered to robustly quantify deviations independent of single air parcel outcomes. Note that Stohl et al. (1995); Hoffmann et al. (2019) define the ab-

solute trajectory deviations and conservation errors as the average over the above air-parcel-wise absolute trajectory deviations. Here, in contrast it is referred to the air-parcel-wise diagnostics with AVTD, AHTD and RTCE. The statistical moments and quantiles are explicitly mentioned (e.g. mean AVTD for the average over all air-parcel-wise AVTDs).

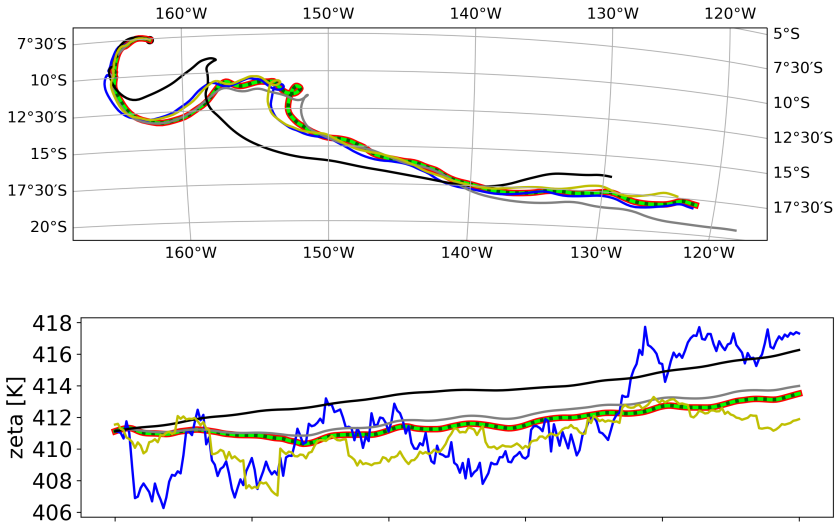
5.2 Results of the evaluation

5.2.1 Examples of trajectories and transport deviations

The simulations are globally initialised and cover almost the entire height range of the free troposphere and stratosphere (about 1-50 km), allowing the analysis of numerous meteorological conditions and different trajectories. Figures 5.1 shows exemplary trajectories for a period of 10 days, highlighting the transport in the troposphere and the quasi-horizontal transport in the UTLS. In the troposphere and in the UTLS, the model deviations are significantly smaller than deviations from external sources such as downsampling of reanalysis data, different vertical velocities, variations in reanalysis data sets (here from ERA5 to ERA-Interim), or the influence of atmospheric diffusion. The statistical significance of these results is discussed in the following chapters, which include the entire ensemble of 1-day forecasts and are later extended to 90-day calculations. In addition, the examples show that in the UTLS the trajectory calculated with ERA5 $1^\circ \times 1^\circ$ roughly follows the fully resolved ERA5 calculations, although deviations still have to be taken into account. The deviations are more substantially in the troposphere. The examples also show that the trajectories calculated with the kinematic velocity approach are vertically scattered in the UTLS, similar to the trajectories with parameterised subgrid scale winds and diffusion.



(a) troposphere



(b) UTLS

Figure 5.1: Trajectories in (a) the troposphere and (b) the UTLS calculated forward for 10 days with different scenarios. See Table 5.1 for the configuration of the scenarios. For each case, the horizontal transport is shown in the upper panel and the vertical transport in the zeta coordinates is depicted below. The scenarios "CLaMS-default", "MPTRAC-default" and "MPTRAC-bestfit" overlap.

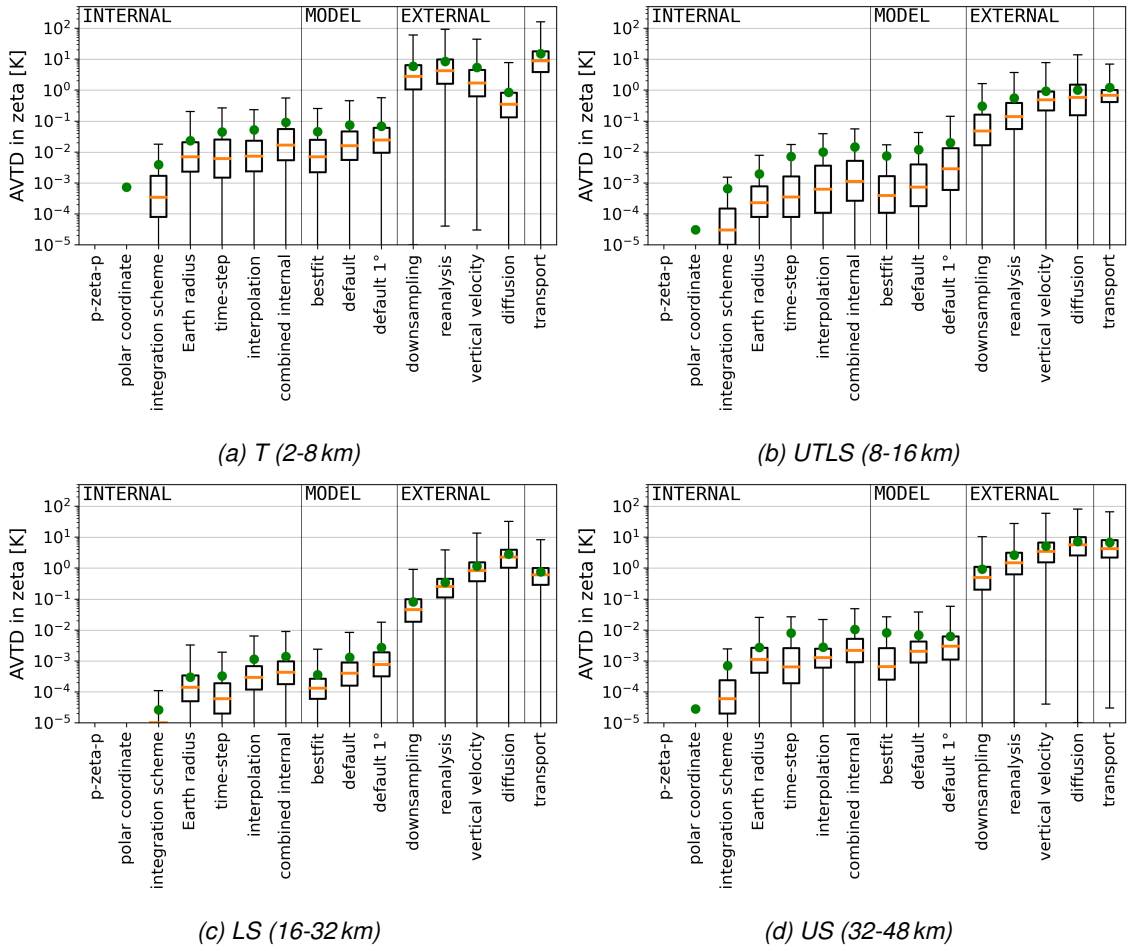
5.2.2 Transport uncertainties over 1 day

Figure 5.2 shows statistics for vertical transport deviations after one day of calculations in hybrid zeta coordinates. Different height ranges are shown depending on the initial position of the air parcels. The height ranges are 2-8 km for the troposphere (T), 8-16 km for the UTLS, 16-32 km for the lower stratosphere (LS) and 32-48 km for the upper stratosphere (US). Throughout the troposphere and stratosphere, the model deviations measured by the median AVTD in the zeta coordinate are one the order of magnitude of the known combined internal uncertainties within MPTRAC (10^{-4} K to 10^{-2} K). This is true for the full ERA5 and ERA5 $1^\circ \times 1^\circ$ scenarios, although the uncertainties increase in the latter scenario (see Fig. 5.2 at the labels "default" and "default 1°").

Separately assessed, the variation of the Earth radius, the time step variation from 180 s to 1800 s in the Runge-Kutta method and the interpolation variation in MPTRAC are estimated to cause median AVTDs lower than 10^{-2} K or 10^{-3} K depending on the level (see Fig. 5.2 at the labels "Earth radius", "time-step" and "interpolation"). Since the change of the time-step between 1800 s and 180 s is only related to small deviations, time-steps of 1800 s are still adequate (see Fig. 5.3 at the label "time step").

Only limited to trajectories in proximity to the poles, uncertainties due to the coordinate singularity must be considered. However, the transformation from spherical coordinates to the stereographic projection at high latitudes causes vertical deviations similar to deviations related to the selection of the integration method. If only air parcels are considered statistically that start at latitudes larger than 72° North or South, the median AVTD in zeta coordinates is on the order of 10^{-5} K for both, the variation of the integration scheme and the horizontal coordinate. These larger deviations that are restricted to the pole, also increase the mean AVTD in Fig. 5.2 over 10^{-5} K (see label "polar coordinate"). However, the horizontal median AHTD is still one order of magnitude larger for the variation of integration scheme than for the horizontal coordinate in the polar region. The p-zeta-p transformation within MPTRAC, which combines pressure-based modules with zeta-based advection, causes transport deviations that are orders of magnitude smaller than the other uncertainty sources (below 10^{-5} K). Hence, the vertical coordinate transformation is the internal uncertainties of the least importance.

The CLaMS and MPTRAC models can also be configured to operate more similarly (i.e. using the same Earth radius, integration method and a similar interpolation method), so that the model uncertainty is substantially reduced (see Fig. 5.2 label "bestfit"). Some minor differences in the interpolation scheme or compilation flags likely explain remaining uncertainties.



(a) T (2-8 km) *(b) UTLS (8-16 km)*

(c) LS (16-32 km) *(d) US (32-48 km)*

Figure 5.2: Different AVTDs in zeta coordinates after 1 day forward calculations for the entire ensemble of air parcels splitted into four height layers. The different uncertainty sources are defined in Table 5.2. The box plots show the median, quartiles (25 % and 75 %), minimum and maximum (outliers have been ignored if they are 10 times the inter-quartile difference). Green dots indicate the mean AVTDs. Deviations for the p-zeta-p transformation and the polar coordinate are lower than 10^{-5} K *and do not show up here. The distinction between internal, model and external uncertainty sources is indicated by vertical lines.*

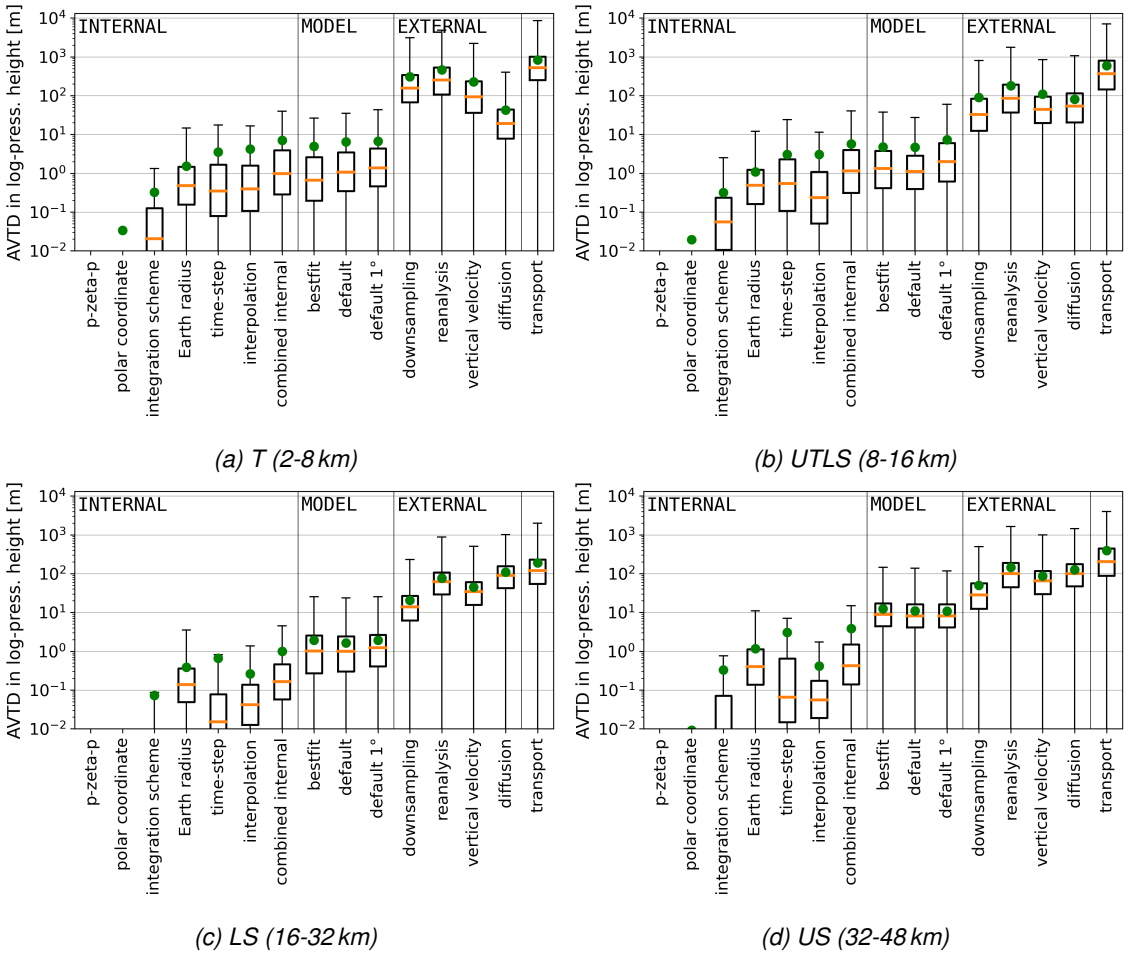
Already with the default configuration are model deviations for CLaMS and MPTRAC one to three orders of magnitude smaller than deviations resulting from external factors (see Fig. 5.2). For those large external uncertainties, the median AVTDs have order of magnitudes of 10^{-2} K to 1 K. The importance of different external factors is different for the troposphere (i.e. below 8 km) and the stratosphere. In the stratosphere diffusion from parameterized sub-grid scale winds and turbulence leads to median AVTDs up to about 7 K after 24 hours which exceed the overall transport median AVTD above 16 km (median AVTD between the initial positions and the end points). The diffusion is accordingly the largest uncertainty at all layers, except between 2-8 km, where the reanalysis uncertainty is larger followed by uncertainty from downsampling and the

vertical velocity. Between 2-8 km the diffusion is also smaller because the parameterization of turbulent diffusion is restricted to horizontal directions. The second largest uncertainty above 8 km is given by the variation of the vertical velocity. For the vertical velocity variation the median AVTD is on the order of $10^{-1}\,\mathrm{K}-1\,\mathrm{K}$. Above 8 km, reanalysis variations, such as between ERA-Interim and ERA5, exhibit median AVTDs that are on the same order of magnitude as uncertainties from the vertical velocities but are smaller by a factor between 3 to 5 depending on height. Moreover, ERA5 $1^\circ \times 1^\circ$ shows a deviation to the full-resolution ERA5 that is one order of magnitude smaller than from the variation of the reanalysis or the vertical velocity. In summary, the largest deviations in the zeta coordinate above 8 km are found from diffusion, followed by the vertical velocity, the reanalysis and finally the downsampling of data. The implementation of diabatic transport has hence a significant impact on the calculations.

Additionally, note that the average AVTD differs from the median AVTD because of the skewed distribution of trajectory deviations (see also Fig. 5.2). The difference is largest for the deviations from internal uncertainty sources, because for the internal uncertainty sources a large amount of air parcels have deviations close to zero, while those air parcel that are affected by strong atmospheric dispersion (e.g. after transport into the troposphere from the UTLS) contribute to a long tail in the density function of the deviations. However, with the external uncertainties, the systematic deviations between the trajectories become more important. They superimpose with the dispersion and increase the median AVTD relative to the mean AVTD. Furthermore, later, it will be shown that, hemispheric, i.e. seasonal differences have to be considered as well to understand larger differences between the median and mean deviations in the global statistics.

Figure 5.3 shows the same statistics as Fig. 5.2 but for log-pressure coordinates. For the height range between 2 km and 32 km the median AVTD in log-pressure coordinates between the two models in default set-up is ~ 1 m. At higher levels (32-48 km) the median AVTD is 10 m. While the median AVTD between the models is around the same order of magnitude as the combined internal uncertainty between 2 km and 16 km again (e.g. Fig. 5.3b), in log-pressure coordinates the deviations are up to two order of magnitudes larger than the combined internal uncertainty for the levels above 16 km (e.g. Fig. 5.3d). This is a consequence of the transition from linear to logarithmic interpolation for pressure in CLaMS at higher isentropes than 500 K (around 20 km), which is not performed in the MPTRAC model (see also Chapter 4.4 for details).

Moreover, in the stratosphere, the median AVTD in log-pressure coordinates between initial and final positions after 1 day (see label “transport”) is larger than the deviation from vertical diffusion in the pressure coordinate, in contrast to the median AVTD in zeta coordinates, because



(a) T (2-8 km)

(b) UTLS (8-16 km)

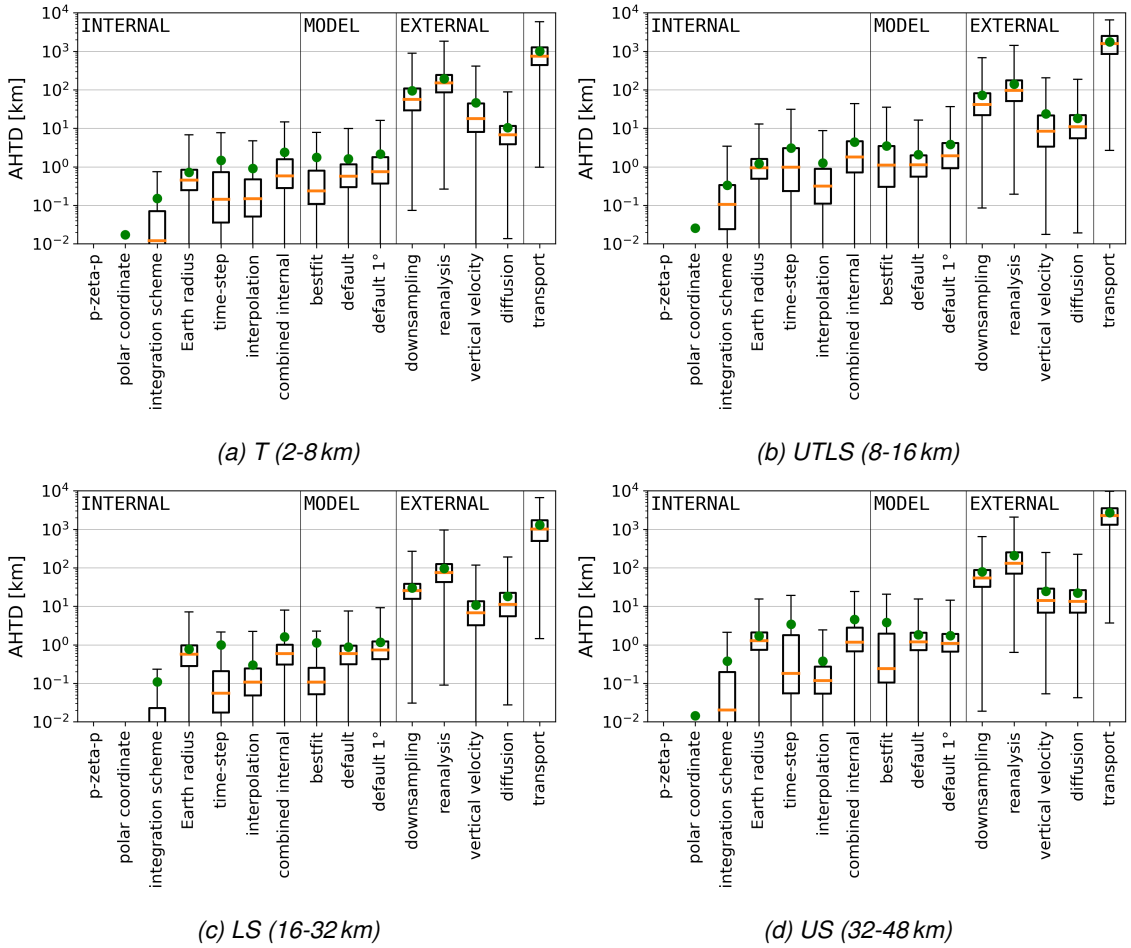
(c) LS (16-32 km)

(d) US (32-48 km)

Figure 5.3: AVTDs in log-pressure heights after 1 day forward calculations for the entire ensemble of air parcels splitted in four height layers. The boxplots indicate quartiles as defined in Fig. 5.2.

the transport in the UTLS is mostly isentropic and hence might cross multiple isobars but less isentropes (see Fig. 5.2 and 5.3 at the labels “diffusion” and “transport” for the layers troposphere and lower stratosphere).

When the AHTDs are considered, qualitatively similar results to the vertical transport deviations are obtained. The horizontal model differences and combined internal uncertainties are of the order of 0.1 km to 1 km after one day of calculations, while external uncertainties lead to absolute horizontal deviations of the order of 1 km to 100 km (see Fig. 5.4). The difference between initial and final positions is around 1000 km. For the internal horizontal deviations the selection of the Earth radius becomes one of the largest internal uncertainties, because it is used during the transformation from Cartesian to spherical coordinates. For the horizontal deviations variations of reanalysis and the downsampling become more important than uncertainty sources such as the vertical velocity, because the later does not alter horizontal velocities directly.



(a) T (2-8 km) *(b) UTLS (8-16 km)*

(c) LS (16-32 km) *(d) US (32-48 km)*

Figure 5.4: Horizontal deviations quantified with the AHTDs after 1 day forward calculation for the entire ensemble of air parcels splitted in four height layers. The boxplots indicate quartiles as defined in Fig. 5.2.

From an overall statistical perspective, as depicted by the Figs. 5.2 to 5.4, different layers show different uncertainties. To emphasize the vertical and hemispheric (i.e. seasonal) dependencies of transport uncertainties, Fig. 5.5a shows the hemisphere-wise vertical mean profiles for a selection of uncertainty sources. First, it is evident again, that all uncertainties from external sources are orders of magnitude larger than uncertainties from internal sources and deviations between the models. Second, all uncertainties, except those due to parameterized diffusion, exhibit the largest absolute mean deviations in the troposphere (below 360 K). The smallest mean AVTDs in the zeta coordinate can be found between around 500 K and 750 K, while the deviations above 600 K increase again with height. In comparison to absolute deviations, relative deviations (see Fig. 5.5b) show less dependency on height (the relative deviations are normalized to the sum of all incremental 1 h transport steps calculated with the default set-up of CLaMS). While the troposphere has highest relative uncertainties, the stratosphere shows lower relative uncertain-

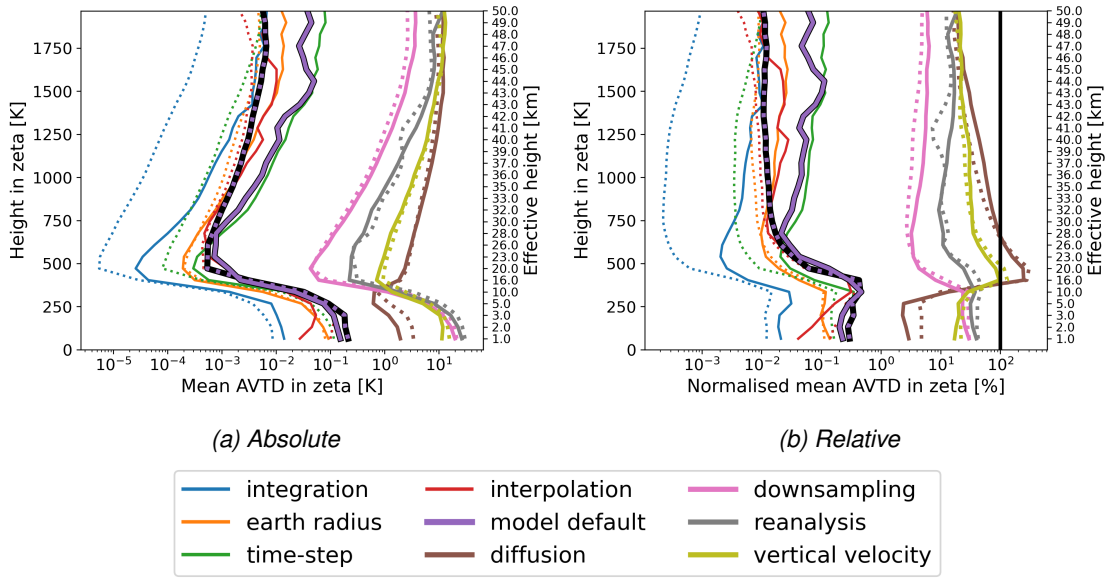


Figure 5.5: Smoothed vertical profiles of hemispheric average AVTD in zeta coordinates for different uncertainty sources (see Table 5.2 for the definitions). (a) absolute values and (b) relative values, where the deviations are normalised to the mean vertical path-length calculated with the default set-up of CLaMS (see "CLaMS-default" in Table 5.1). The black profile emphasizes the model difference for the models with default configuration in the summer hemisphere. The dotted lines indicate uncertainties of the Northern Hemisphere (boreal summer), and solid lines indicate uncertainties of the Southern Hemisphere (austral winter). The effective height is the average log-pressure height at a zeta level at the beginning of the calculations. The black vertical line in panel (b) marks 100%.

ties, which are also mostly independent of height. This indicates that the increase of the mean AVTD in the stratosphere is given, because air parcels cross more levels at those heights during the transport process.

The profiles of the transport uncertainties are similar in the two hemispheres. However, if hemispheres are compared in more detail, the strongest relative internal uncertainties are found in the winter hemisphere, i.e the Southern Hemisphere (see Figure 5.5b). Absolute and relative uncertainties in the Southern Hemisphere, specifically in winter, are most likely much larger in the stratosphere due to the influence of the polar vortex and increased wave activity in winter. This seasonality was found by Hoffmann et al. (2019) with kinematic transport calculations as well. In particular, the integration time-step becomes the dominant internal uncertainty source in the region of the polar vortex, because high zonal velocities require shorter time-steps for integration, which is not always fulfilled with 1800 s time-steps for the ERA5 reanalysis. Larger deviations in horizontal directions lead to larger vertical deviations as well. The large hemispheric differences in the uncertainty explain the increased difference between the median and the mean deviations for the internal uncertainties and particular the variation of the integration time-step in the Figures

5.2 to 5.4 in the stratosphere.

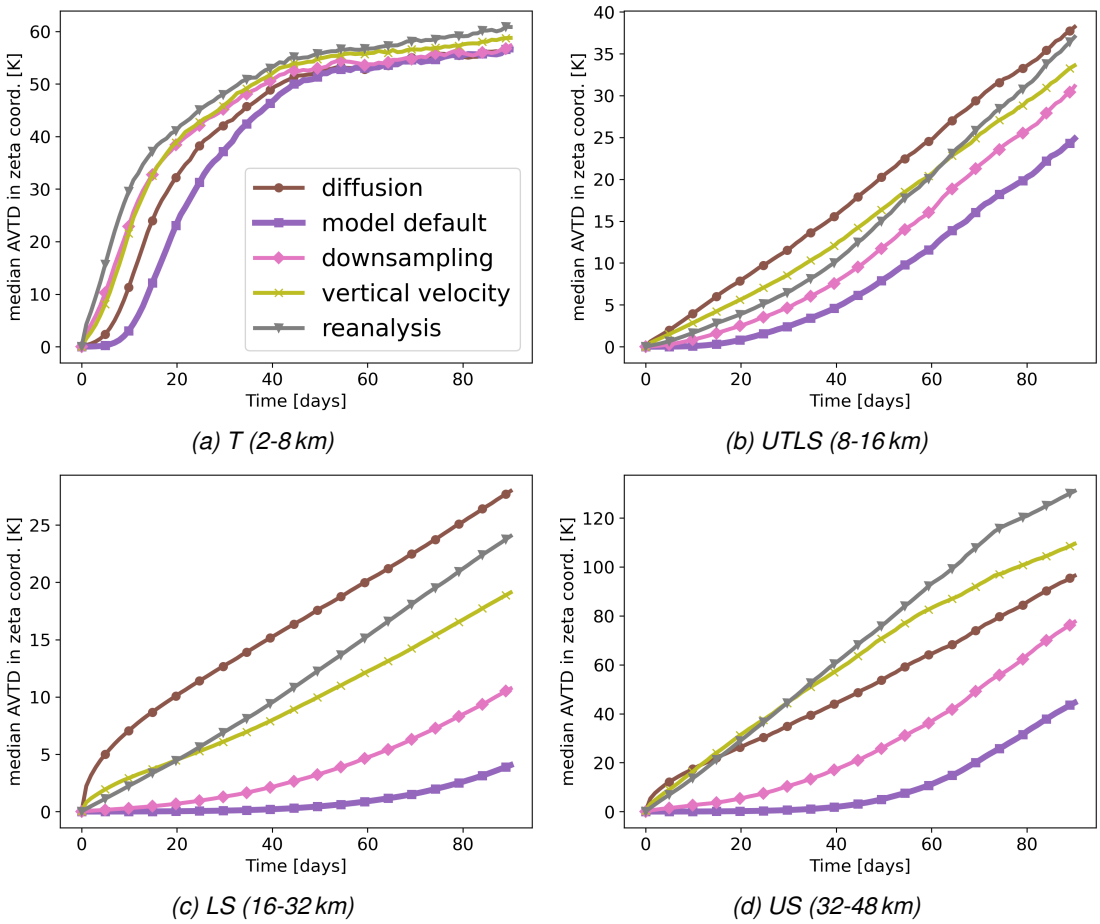
Moreover, the vertical profiles in Fig. 5.5 reveal that throughout the stratosphere the deviation from variation of the vertical velocity is larger than the deviation from the variation of the reanalysis, which in turn is larger than the change from ERA5 to ERA5 $1^\circ \times 1^\circ$. In particular uncertainties from diffusion are lowest in the troposphere, but increase sharply up to around the tropopause, where the diffusion becomes the largest normalised and absolute source of uncertainties. At higher levels the normalised uncertainty from diffusion decreases slowly again back to normalised uncertainty ranges comparable to the normalised uncertainty from variation of the vertical velocity. The normalised deviations given by the variation of the vertical velocity is reduced in the troposphere, because the zeta coordinates approximate to the sigma coordinates at lower levels. Overall, the results show that the implementation of diabatic vertical transport into MPTRAC has a significant impact, comparable to other external uncertainties.

5.2.3 Uncertainty growth during 90 day forward calculations

To investigate the uncertainty growth between the CLaMS and MPTRAC models and to better understand the model differences in the context of other uncertainties, trajectory calculations were performed for 90 days starting from 1 June 2 016. Figure 5.6 displays the temporal evolution of the median AVTD between the two models, along with the downsampling, vertical velocity, reanalysis and diffusion transport deviations. For the intercomparison of the two models the default configuration of the models (see “Model default” in Table 5.2) is used as they represent the usual uncertainty that has to be expected.

The model deviations and other transport uncertainties vary with height. In the troposphere, the median AVTD of the external uncertainties remains below 1 K only for a short period (a few hours to days) due to the strong mixing and convection. Subsequently, external uncertainties in this region grow rapidly with up to 4.3 K per day. In particular, the selection of the reanalysis, the vertical velocity and downsampling cause fast divergence in the troposphere. The median model AVTD is smaller than uncertainties related to changes in reanalysis data, downsampling of the data, or parameterized sub-grid scale winds and diffusion. The median AVTD between the two models remains below approximately 1 K for the first week. Subsequently, there is also a sharp increase (up to 2.2 K per day), reaching a median AVTD of about 55 K at 40 days of simulation time, where the different uncertainties reach a similar magnitude.

In the lower and upper stratosphere, the AVTDs remain lower because air parcels mainly move isentropic. Additionally, horizontal mixing is much less in most regions of the lower and



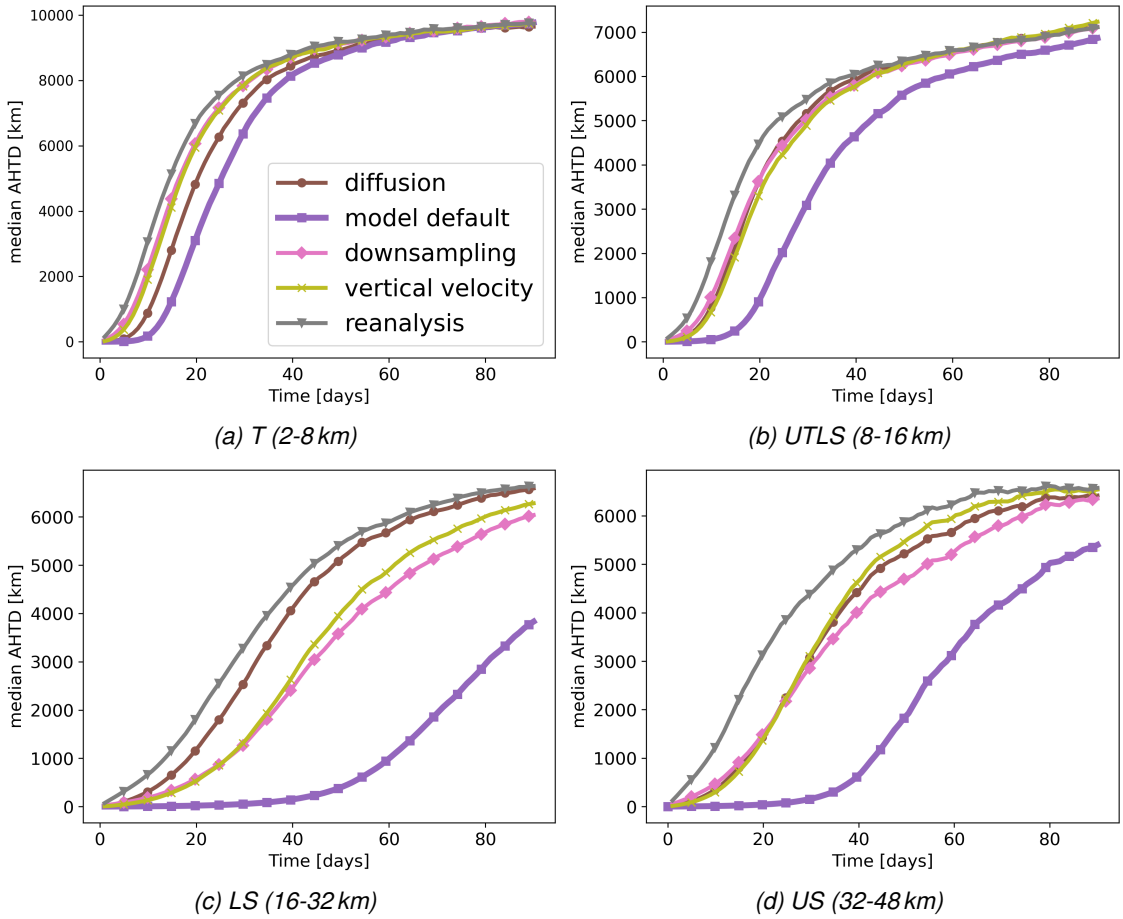
(a) T (2-8 km) (b) UTLS (8-16 km)

(c) LS (16-32 km) (d) US (32-48 km)

Figure 5.6: Evolution of the median AVTD in the zeta coordinate for different uncertainty sources for 90 days. The median AHTD between the two models with default configuration is labeled "model default" (see Table 5.2 at "Model default"). The starting date is the 1 June 2016. The classification into the layers is done with the initial heights of the air parcels.

upper stratosphere in contrast to the troposphere. The median model AVTD is again much smaller than all other uncertainty sources, but now for the entire 90-day integration period. In the lower stratosphere, 50% of the air parcels have a model AVTD lower than 1 K for approximately two months and afterwards the deviation still increases slowly (not more than 0.16 K per day). In the upper stratosphere, the same criterion is met after around 34 days, also with a slow to moderate increase afterwards (not more than 1.2 K per day).

Uncertainties from the selection of the vertical velocity and the reanalysis are of similar importance. In the UTLS and at higher altitudes, the variation of the vertical velocity first shows slightly larger median AVTDs than the variation of the reanalysis. However, after a couple of weeks, the median AVTD from reanalysis selection is higher, because the choice of the vertical velocity does not affect the horizontal wind speeds as it is the case for the choice of the reanalysis. The smallest transport uncertainty from external sources throughout the atmosphere is given by the



(a) T (2-8 km) *(b) UTLS (8-16 km)*

(c) LS (16-32 km) *(d) US (32-48 km)*

Figure 5.7: Evolution of the median AHTD of different uncertainty sources for 90 days. The median AHTD between the two models with default configuration is labeled "model default" (see Table 5.2 at "Model default"). The starting date is the 1 June 2016. The classification into the layers is done with the initial heights of the air parcels.

ERA5 1°×1° data, because ERA5 1°×1° has the same vertical resolution and similar horizontal velocities as the ERA5 reanalysis. Finally, in the UTLS results lie in between the pure stratosphere and the troposphere, influenced by the transport of air parcels between the stratosphere and troposphere.

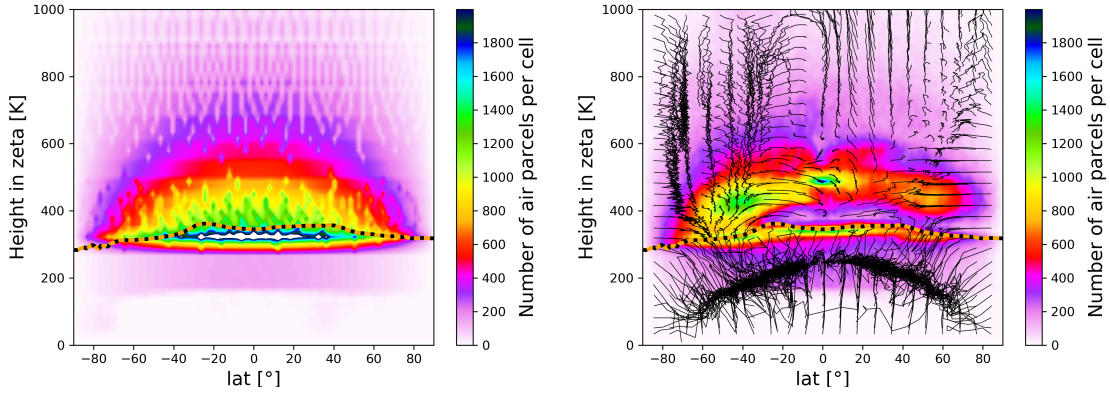
The differences between the two models have an impact on the horizontal distribution of the air parcels as well (Fig. 5.7). While the models median AHTD is less than 1 000 km for 45 to 60 days in the stratosphere, it is less than 1000 km only for 15 to 20 days in the troposphere and UTLS. In the UTLS and troposphere air parcels deviations reach an upper boundary, where further uncertainty growth stagnates for all scenarios, after around 40 days. In the stratosphere this boundary is approached after 60 to 90 days for external uncertainty sources, while it is not completely approached by the model difference in this time period. Even the horizontal deviations for the scenario with ERA5 1°×1° grow considerable throughout the atmosphere, indicating that

air parcels are often not in good agreement with the ERA5 reanalysis.

5.2.4 Air parcel distribution on seasonal timescales

Since individual trajectories are not expected to agree over time periods of several months, the statistical distribution of air parcels after 90 days integration period is used to quantify the differences between the models and the uncertainty related to external sources. The air parcels have been initialized at the 1 June 2016. For reference, the initial density of the air parcels is shown in Fig. 5.8a. Figure 5.8b shows the zonal mean distribution of air parcels after 90 days of forward calculations for the CLaMS model with its default configuration. After 90 days, the density is highest around the vertical level of 450 K, where most of the air parcels have been transported to within the shallow and deep branch of the Brewer-Dobson circulation (BDC). Air parcels also accumulate below the tropopause and near the surface below 2 km (where the models are configured to terminate the air parcel trajectories). Sub-grid scale process, such as convection, that would be required to reach a well-mixed troposphere are not parameterized in the calculations. Therefore, the accumulation of air parcels is a consequence of up- and downward transport limited to the resolved mean flow of the troposphere, combined with the tropopause as an upper transport barrier and the ground as the lower transport barrier.

Furthermore, more air parcels are leaving the Northern Hemisphere than entering it in the calculations, i.e. the cross-equatorial flow in the UTLS increases the air parcel density in the Southern Hemisphere relatively to the Northern Hemisphere. Since the air parcels were initialized at the first June 2016 the simulation describe the boreal summer conditions. As indicated by averaged trajectories in Fig. 5.8b the hemispheric asymmetric distribution of air parcels is mostly related to the strength of the southern hemispheric, shallow branch of the BDC, that is located between 40° S and 5° N in latitude and crosses the equator, respectively (See Appendix C.1 for details about the average trajectories.).



(a) Initial condition *(b) CLaMS default, final positions*

*Figure 5.8: Initial and final air parcel distribution after 90 days when calculated with the CLaMS default set-up. The date of the initialisation is 06/01/2*016*. Black lines show box-wise averaged trajectories to indicate the average circulation of the trajectories. The orange dotted line indicates the 90 days average tropopause.*

The global distribution of air parcels as simulated with MPTRAC is almost identical to the distribution as simulated with CLaMS as can be seen in Fig. 5.9a, where the bias between the air parcel distributions of both models is shown as well as contour lines of air parcel frequencies after 90 days forward calculations. The contour lines of the air parcel frequencies align very well around the tropopause and at higher levels at around 500 K. Overall, there is no significant bias between the air parcel distribution of the two models. Except for statistical noise, the simulation results of CLaMS and MPTRAC are in excellent agreement. This is in distinct contrast to biases found for other known uncertainties (e.g. from reanalysis, vertical velocity and downsampling), as will be discussed below.

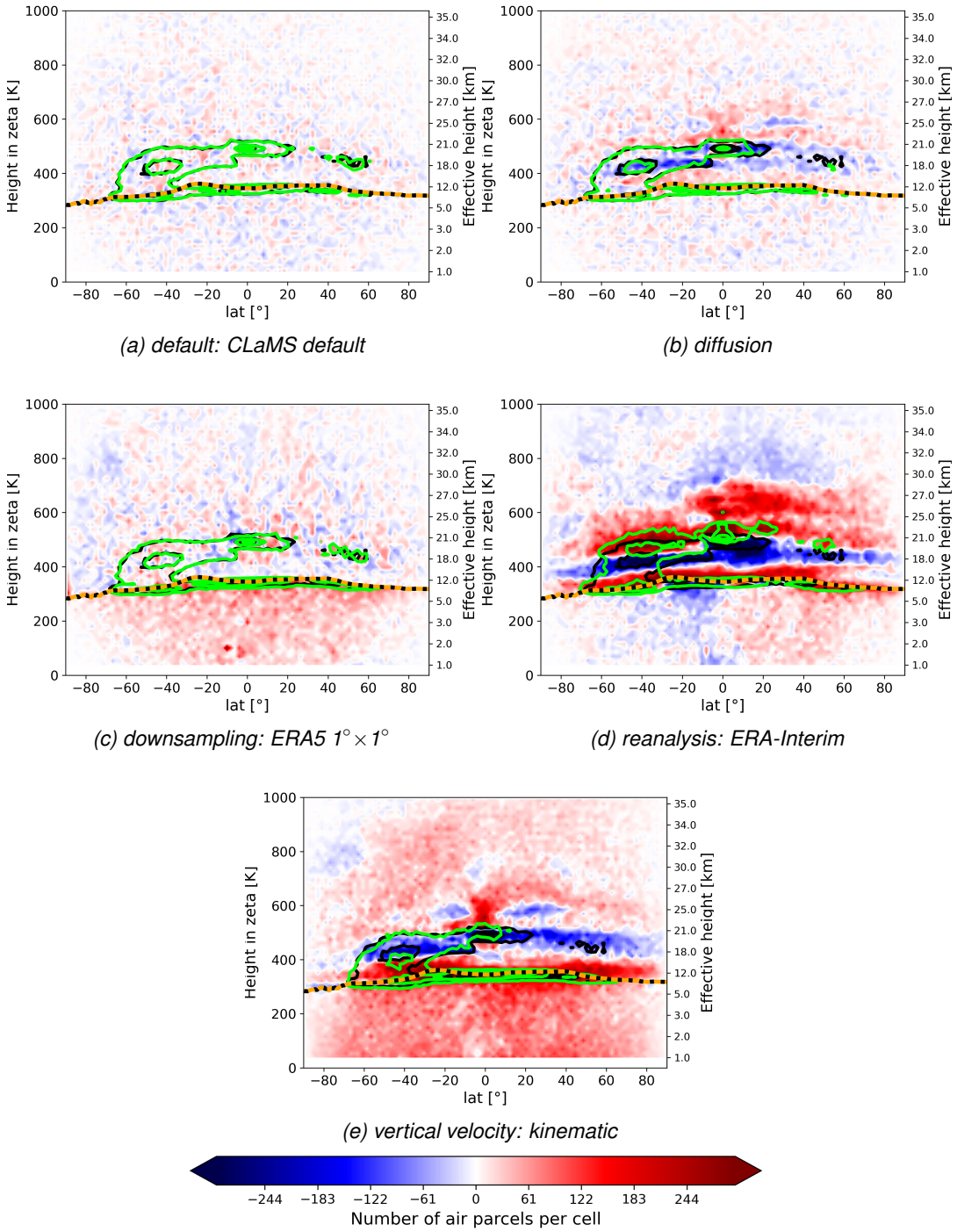
When the diffusion module (see Fig. 5.9b) is switched on in MPTRAC, the patterns without diffusion are reproduced as well, but with smoothed peaks (in Fig. 5.9b green contours shrink in comparison to black contours). Less air parcels are found in the height region where the frequency of air parcels peaks for the default scenarios of MPTRAC and CLaMS (around 450 K), whereas the frequencies are increased at the neighbouring levels. The result indicates, that the mean distribution is not affected by the sub-grid scale diffusion, except for a smoothing effect. It is shown later that the diffusion causes large cross-isentropic dispersion (see also Fig. 5.11).

The downsampling of the ERA5 data (see Fig. 5.9c) has only a minor impact on the distribution of air parcels above the tropopause. The largest differences can be found at the tropical tropopause and in the troposphere. With ERA5 $1° \times 1°$, more air parcels remain located within the troposphere after 90 days. This is presumably a consequence of reduced vertical transport in convective events in the ERA5 $1° \times 1°$ in comparison to the full resolution data, which would

be in agreement with other studies (e.g. Hoffmann et al., 2023b). With weaker vertical transport, more air parcels remain in the troposphere and fewer air parcels are transported downward into the model boundary layer, where they are terminated.

With ERA-Interim, qualitatively very different results are found (see Fig. 5.9d). The BDC transport is faster with ERA-Interim than with ERA5 between levels around 400 K to 600 K. Hence, more air parcels are transported from around 400 K to around 600 K in ERA-Interim. At the same time, transport at higher levels than 600 K is slower with ERA-Interim than with ERA5, which decreases the air parcel number relative to ERA5 above 700 K. The upward transport in the upper part of the shallow branch is faster in ERA-Interim than in ERA5 as well (see also Appendix C.1d). Hence, more air parcels are found at higher altitudes around latitudes of 45° S with ERA-Interim. These results are in agreement with climatological findings (e.g. Ploeger et al., 2021). Additionally, more air parcels are found between the 400 K level and the tropopause with ERA-Interim than with ERA5. The combination of uncertainties between the two reanalyses complicates their intercomparison in the UTLS.

The biases between simulations with diabatic and kinematic vertical velocities in ERA5 are of similar size as the biases between simulations with ERA-Interim and ERA5 (see Fig. 5.9e). With kinematic vertical velocities, the upward transport in the BDC is as fast as with the diabatic transport scheme or even faster for levels between 400 K and 900 K (see also Appendix Fig. C.1e). Therefore, fewer air parcels can be found between 400 K and 500 K compared to the diabatic vertical velocities. Additionally, the bias roughly resembles the bias found for the scenario with parameterized diffusion and hence indicates an increased dispersion of the air parcels. With kinematic velocities, increased air parcel numbers can be found closely above the tropopause as well, in comparison to the diabatic calculations. This possibly indicates increased transport across the tropopause from below. However, for the kinematic velocities, higher numbers of air parcels are found in the troposphere because the applied criteria for excluding air parcels from further transport (reaching the level where the zeta coordinate is zero) is not fulfilled. Therefore, the increase of air parcels closely above the tropopause could be a consequence of higher air parcel numbers remaining in the troposphere as well.



(a) default: CLaMS default

(b) diffusion

(c) downsampling: ERA5 1°×1°

(d) reanalysis: ERA-Interim

(e) vertical velocity: kinematic

Figure 5.9: Zonal mean bias of the air parcel distributions after 90 days between the default MPTRAC scenario and a selected scenario. Positive bias indicates lower frequency with the default MPTRAC scenario and higher frequency with the respective scenario. The orange dotted line is the 90 days average tropopause. The green contours show the 600, 1000 and 1400 air parcel number contours of the air parcel distributions for intercomparison with the scenarios (a) CLaMS default, (b) Diffusion, (c) downsampling: ERA5 1°×1°, (d) reanalysis: ERA-Interim and (e) vertical velocities: kinematic calculations. The black contours indicate the same contour lines but for the MPTRAC default scenario.

5.2.5 Differences between diabatic and kinematic calculations

The agreement between kinematic and diabatic trajectories might differ from reanalysis to reanalysis, in terms of the general circulation and of cross-isentropic dispersion of air (Ploeger, 2010; Hoffmann et al., 2019; Legras and Bucci, 2020; Ploeger et al., 2021). Figure 5.10 shows biases after 90 days forward calculations between the kinematic and diabatic vertical velocity approach in ERA5 and ERA-Interim. The transport biases between kinematic and diabatic trajectories are significantly decreased with ERA5 in comparison to ERA-Interim between 300 K and 600 K. The height of the air parcels varies substantially more with ERA-Interim than with ERA5 after 90 days of calculations (see black and green contours in Fig. 5.10).

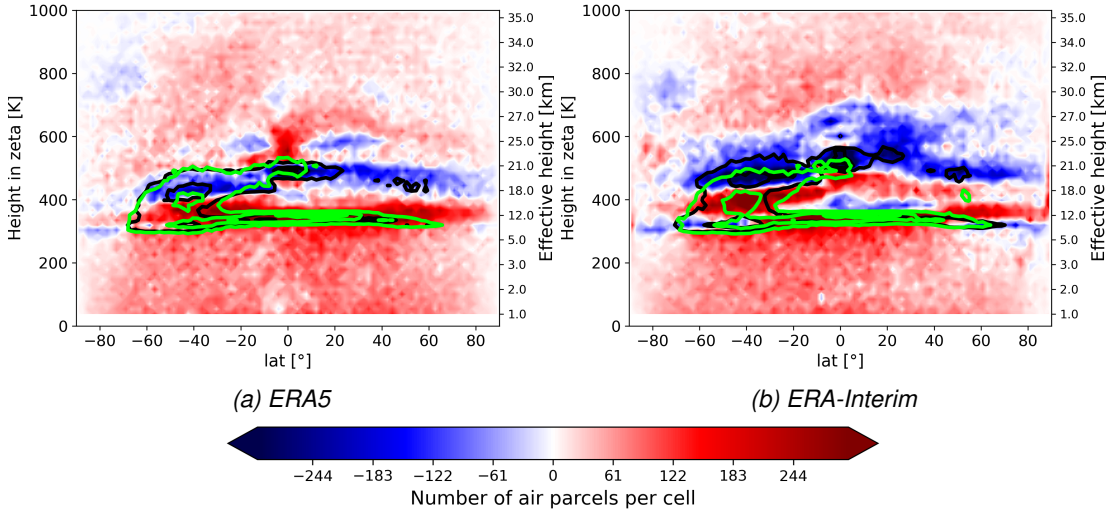


Figure 5.10: Intercomparison of air parcel distributions and biases inferred from diabatic (black contours) and kinematic (green contours) calculations for ERA5 (a) and ERA-Interim (b). The colorbar indicates the bias as used in Fig. 5.9. Here, positive bias indicates more air parcels with the kinematic scenario than with the diabatic scenario. All calculations are done with MPTRAC.

The cross-isentropic dispersion of air parcels can be quantified with the variance of the potential temperature (Sparling et al., 1997; Ploeger, 2010) of an air parcel ensemble initialized at one isentropic level at a specific time: $\langle \delta\theta^2 \rangle = \langle (\theta - \langle \theta \rangle)^2 \rangle$. The brackets $\langle\ \rangle$ symbolize the average over the air parcel ensemble. For levels higher than 360 K the zeta coordinate approximates isentropic coordinates, i.e. potential temperatures, and hence can be used to estimate the dispersion as well. Since the initial position of the air parcels are given at zeta levels, each of these levels contain a ensemble of air parcels that can be used to calculate a separate variance. Similar to (Ploeger, 2010), only these air parcels are considered here that are initialized in the equatorial region (latitudes between $\pm 30°$) and remain in there. Figure 5.11 summarizes the variance

calculated at different zeta levels in the range between 300 K and 1 000 K.

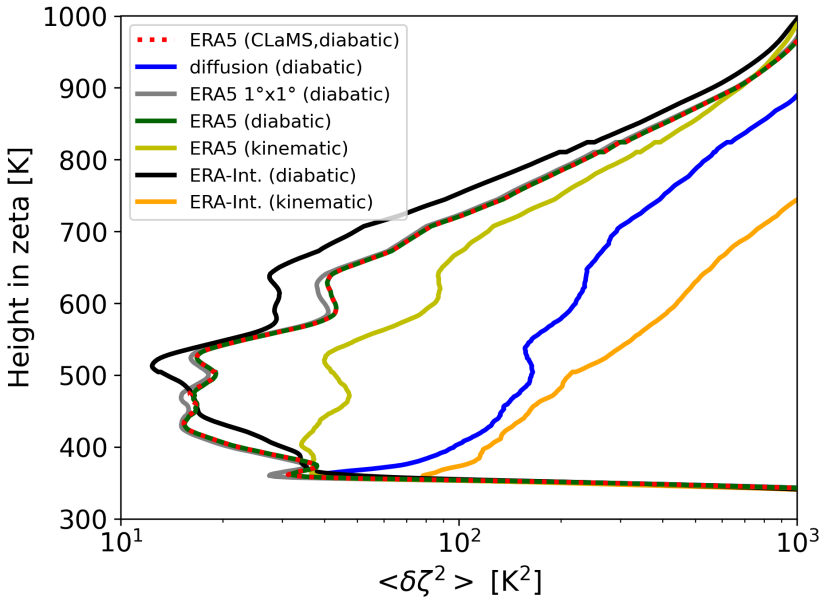


Figure 5.11: Profiles of the zeta variance $\langle\delta\zeta^2\rangle$ after 10 days as a measure of the cross-isentropic dispersion. All scenarios are calculated with MPTRAC, except "ERA5 (CLaMS,diabatic)". The profiles for the scenarios "ERA5 (diabatic)" and "ERA5 (CLaMS,diabatic)" overlap each other. For the definition of the MPTRAC scenarios see also Table 5.1, from label "MPTRAC-default" to label "MPTRAC-def-erai-kin".

The kinematic calculations with ERA5 still have a higher dispersion in comparison to the diabatic calculations, what supports the implementation of the diabatic scheme into MPTRAC. The variance is around 3 times higher with the kinematic calculations between 400 K and 550 K. However, the dispersion by parameterized turbulent diffusion and sub-grid scale wind fluctuations is much higher (see Fig. 5.11) and the discrepancy in the dispersion between kinematic and diabatic calculations is significantly reduced in comparison to ERA-Interim. Finally, it can be noted here that the two models agree highly with regard to dispersion. The differences between calculations with ERA5 and the low-resolution ERA5 1 °$\times$1 ° are small as well.

5.2.6 Conservation of dynamical tracers in the stratosphere

In the stratosphere, the potential temperature (θ) and the potential vorticity (PV) are approximately conserved. To assess the conservation of dynamical tracers in different scenarios with the newly implemented diabatic transport scheme in MPTRAC, Fig. 5.12 shows the 10-day evolution of the mean RTCE of the PV in the stratosphere, starting from 1 June 2 016. Only air parcels with an initial height above 360 K, the approximated level of maximum convective outflow, are analysed. The potential temperature and PV are calculated with the modules of the MPTRAC model along the trajectories. The mean conservation error after one day varies between 10% and 13% depending on the scenario. After 10 days, the mean RTCE increases to values between 25% and 28%. The differences between the different scenarios remain moderate, with slightly lower PV conservation errors with the ERA5 reanalysis and diabatic velocities as implemented in MPTRAC. To show the significance of the increase, it has been compared with the unresolved parameterized subgrid scale diffusion: The difference between the diabatic calculations with ERA5 and the kinematic velocity scheme is almost as large as the difference between the kinematic scenario and the scenario with parameterized sub-grid scale winds and diffusion. The conservation of PV depends on the horizontal advection. Therefore, calculations with the parameterized sub-grid scale winds increase the RTCE as well as calculations with the horizontally downsampled ERA5 1 °$\times$1 ° data.

Figure 5.13 shows the evolution of the conservation error of the potential temperature. The mean RTCE is very similar ($\pm$0.1 %) for all scenarios except for the scenario with parameterized diffusion and sub-grid scale wind fluctuations. In particular the calculations with ERA5 and the ERA5 1 °$\times$1 ° show very similar results, probably because the ERA5 1 °$\times$1 ° data has the same vertical resolution as the ERA5 data and therefore represents the vertical displacements across isentropes similarly.

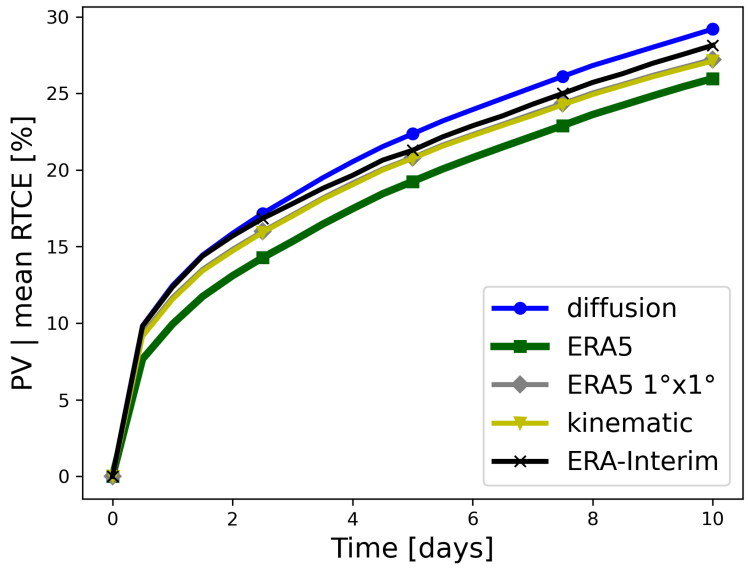


Figure 5.12: Evolution of the mean RTCE of PV for different scenarios within a 10 days period. All scenarios are driven with MPTRAC. See the scenarios in Table 5.1 between MPTRAC-default and MPTRAC-def-erai for further details. The starting date is 06/01/2016. Only air parcels with an initial altitude above 360 K are considered. Note that in (a) the scenarios with kinematic vertical velocities and ERA5 1°×1° data overlap.

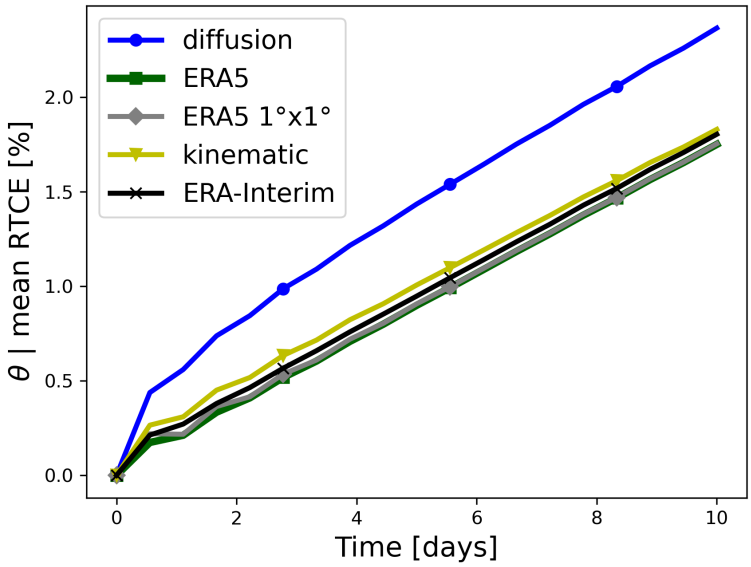


Figure 5.13: Evolution of the mean RTCE of theta for different scenarios within a 10 days period. All scenarios are driven with MPTRAC. See the scenarios in Table 5.1 between MPTRAC-default and MPTRAC-def-erai for further details. The starting date is 06/01/2016. Only air parcels with an initial altitude above 360 K are considered. Note that scenarios with ERA5 and ERA5 1°×1° overlap.

5.3 Conclusions

In this study, a diabatic transport scheme based on hybrid zeta coordinates was implemented into the MPTRAC Lagrangian transport model. This work was mainly motivated by the intention to enable a transition from the CLaMS Lagrangian transport framework towards a code which is more suitable for upcoming HPC architectures. To assess the implementation of the diabatic transports scheme in MPTRAC, evaluations were conducted using approximately 1.4 million globally distributed air parcels in the troposphere and stratosphere, following an initialization method commonly employed with CLaMS. Trajectory forward calculations were performed for the boreal summer of 2 016. In the evaluation, the model differences were put in the context of various other uncertainty sources in Lagrangian transport calculations. Consequently, the model differences between CLaMS and MPTRAC were presented within a hierarchy of uncertainties associated with Lagrangian transport models.

The key differences between the two Lagrangian models relate to their approach for interpolation of the driving meteorological data and the numerical integration scheme. Although both models apply four dimensional linear interpolations, CLaMS performs them directly in spherical coordinates, while MPTRAC performs them in Cartesian coordinates. As a default, CLaMS uses the classical fourth order Runge-Kutta scheme with 1800 s integration steps for numerical integration to run with feasible computational costs. MPTRAC employs the mid-point scheme with 180s integration time-steps. At a time-step of 180 s both integration schemes deliver very similar results. Moreover, the differences between time-steps of 1800 s and 180 s are still very minor, making them both suitable for integration. However, a reduction of the time-step can reduce the internal uncertainties in particular in the winter hemisphere. MPTRAC also can be adjusted to fit better to the parameters and interpolation scheme as in the default CLaMS scenario, so that the agreement was slightly improved between the models (see “bestfit” scenarios). The residual differences between the models, are likely caused by remaining differences in the interpolation. For improved agreement, CLaMS and MPTRAC should use the identical Earth radius. Further alignment of the interpolations could achieve even better agreement.

Despite the conceptual model differences, it was demonstrated that, for a period of 1 day, the discrepancy between CLaMS and MPTRAC air parcel vertical positions, which ranges in the orders of magnitude between 10^{-4} K and 10^{-2} K, is comparable to the combined internal uncertainties associated with different Earth radi, interpolation methods, numerical integration schemes and selected integration time-steps. These deviations are, at a minimum, around one order of magnitude smaller than the uncertainties arising from external sources, such as differences between

reanalysis datasets, downsampling of the ERA5 reanalysis data, and unresolved fluctuations of the wind fields (10^{-2} K to 1 K). Thus, the analysis of the model differences indicates an excellent agreement of CLaMS and MPTRAC within the boundaries of known internal and external uncertainties. This holds also in the regions of most notable differences, including the troposphere and the winter stratosphere with the polar vortex.

The uncertainty growth between the models and from external sources for 90 days was also estimated. The vertical transport uncertainty remains low (less than around 1 K) for several weeks, in particular in the stratosphere. The transport deviation between the models is significantly smaller than the deviation caused by external sources of uncertainty for the entire 90 days time period. In particular, large uncertainty growth from variations of the vertical velocity (diabatic to kinematic) shows that the implementation of the diabatic transport scheme into MPTRAC has significant impact on the transport of air parcels in comparison to the kinematic transport scheme.

For a global, long-term study of trace gases, the statistical distribution of air parcels in the UTLS, as opposed to individual trajectory errors, becomes more important. In their present configurations, both models distribute air parcels very similarly even after 90 days, supporting the hypothesis that the models provide similar long-term tracer fields. Accordingly, no biases in the air parcel distributions were found between the two models. In contrast, known external uncertainties caused significant biases in the trajectory calculations over the 90 day integration period.

Differences between calculations with diabatic and kinematic vertical velocities, even with ERA5, still on the order of reanalysis differences, further corroborating the implementation of the diabatic scheme into MPTRAC. However, the difference between diabatic and kinematic calculations are significantly reduced with ERA5, in comparison to ERA-Interim, concerning the vertical transport in the circulation of the lower stratosphere, but also concerning the cross-isentropic dispersion in the tropical lower stratosphere. Diabatic transport calculations show that transport within the BDC is faster with ERA-Interim than with ERA5 between 400 K and 600 K, but slower for higher levels. This is in agreement with recent climatological and regional studies (Ploeger et al., 2021; Vogel et al., 2023).

Furthermore, since model and internal uncertainties of the trajectory models are much smaller than uncertainties due to downsampling of ERA5 data, it can be concluded that using ERA5 $1^\circ \times 1^\circ$ for the sake of acceleration of computations has considerable side-effects, in particular in the troposphere. This stresses the important role of the spatiotemporal resolution of the global reanalysis fields, next to other improvements of the forecast models and data assimilation schemes used to produce the reanalyses. Making Lagrangian models ready for operating with higher resolution

meteorological data (as intended with MPTRAC) is fundamental to fully exploit the opportunities of next-generation reanalyses. Alternatively, applying better downsampling or data compression methods might be an option for future work.

Furthermore, the mean RTCE for PV and potential temperature were estimated for different MPTRAC scenarios. The mean RTCE of PV for air parcels above the 360 K zeta level is between 25% and 28% for the MPTRAC model after 10 days. Diabatic calculations with ERA5 decrease the conservation error slightly (2-4%) in comparison to diabatic calculations with ERA-Interim, or calculations with ERA5 and kinematic vertical velocities. The mean RTCE is in agreement with values reported by Hoffmann et al. (2019) for kinematic calculations. Hoffmann et al. (2019) reported as well a decreased mean RTCE of PV with ERA5 relative to ERA-Interim. Only small differences with regard to the conservation of the potential temperature have been found here for the diabatic transport scenarios with ERA5 and ERA-Interim. This is also in agreement to kinematic calculations of Hoffmann et al. (2019) who reported a reduced mean RTCE of the potential temperature with ERA5 in comparison to ERA-Interim with kinematic calculations as well but with more substantial differences. This could be a consequence of the reduced cross-isentropic dispersion from ERA-Interim to ERA5 with kinematic calculations found in Chapter 5.2.5, while the change of the cross-isentropic dispersion remained smaller from ERA-Interim to ERA5 with diabatic calculations. Some differences to Hoffmann et al. (2019) might as well be caused by differences in the initialization, which differs in year and initial air parcel distribution.

Ultimately, this evaluation shows that, with the newly implemented hybrid zeta coordinates and diabatic transport scheme, MPTRAC can replace CLaMS' trajectory module, without introducing any significant biases or other deviations. Thus MPTRAC can be coupled safely to global 3-dimension simulation with CLaMS including i.a. irreversible mixing and stratospheric chemistry in the future and exploit the significant changes found with the high resolution ERA5 reanalysis combined with diabatic transport.

Chapter 6

A multi-scenario Lagrangian trajectory analysis to identify source regions of the Asian tropopause aerosol layer on the Indian subcontinent in August 2016

After expanding and evaluating the Lagrangian transport framework, i.e. models and reanalysis data in Chapter 5 on a global scale, the two models can be applied in a regional study to find surface sources of the ATAL. The research questions *"Which surface sources and transport pathways characterize air masses found within the ATAL?"* and *"How robust is our understanding of surface sources and transport pathways of ATAL inferred from different Lagrangian transport calculations considering uncertainties in upward transport?"* are addressed in this chapter by combining measurement data and Lagrangian transport calculations.

Within this thesis, data from balloon-borne measurements is used, obtained with the Compact Optical Backscatter Aerosol Detector (COBALD) backscatter sonde, that identified the location of the ATAL along the balloon ascents over Nainital, India in 2016 (Brunamonti et al., 2018; Hanumanthu et al., 2020). Using the Lagrangian transport models, air parcels can be initialized at the detected ATAL locations and transported backward to the model boundary layer (MBL) to find possible source regions of the ATAL. Additionally, the measured and daily averaged aerosol backscatter at the location of the balloons can be related to surface regions and their properties. Hanumanthu et al. (2020) performed a similar trajectory analysis of the ATAL limited to diabatic calculations with the ERA-Interim reanalysis and the CLaMS model. Therefore, benefits of the recent ERA5 reanalysis weren't exploited, and the robustness of these transport calculations remained uncertain in their study. In this thesis, the same ATAL measurements are used as in Hanumanthu et al. (2020), however in addition the uncertainties and the robustness of the transport simulations were assessed performing different Lagrangian transport calculations, including

diabatic calculations with the high resolution ERA5 reanalysis.

Lagrangian trajectory calculations in combination with observations have been used frequently to investigate the relation between source regions at the Earth's surface and the chemical composition of air masses within the Asian monsoon anticyclone as well as ATAL properties (e.g. Li et al., 2017b; Vernier et al., 2018; Höpfner et al., 2019; Legras and Bucci, 2020; Johansson et al., 2020; Hanumanthu et al., 2020; Zhang et al., 2019b). These calculations rely on reanalysis data and their ability to correctly resolve diabatic vertical ascent in the ASMA and transport in convection. In studies of the source regions that are contributing to the composition of the Asian monsoon anticyclone, better agreement between diabatic and kinematic calculations and between models and observations was found when ERA5 was used instead of ERA-Interim (e.g. Bucci et al., 2020; Legras and Bucci, 2020). The attribution of the sources, however, still depends on the reanalysis data (Bucci et al., 2020). Vertical transport from the MBL to the UTLS is faster with ERA5 than with ERA-Interim (e.g. Li et al., 2020). Altogether, these studies indicate, that ERA5 improves the simulations in comparison to the ERA-Interim reanalysis.

Lagrangian transport calculations are expected to be well suited for the detection of ATAL surface source regions. However only a few investigations have been done with regard to the robustness of this approach against variation of the reanalysis data, transport models and vertical velocities. Here, the trajectory analysis of the measurements over Nainital was extended with different simulation scenarios based on two reanalyses (ERA5 and ERA-Interim), two Lagrangian transport models (CLaMS and MPTRAC), two types of vertical velocities (diabatic and kinematic trajectories), and with changes of integration time-steps and parameterisation parameters (e.g. for convection). The goal of these sensitivity tests is to identify differences and robust transport features that emerge from different simulation setups for the vertical transport, including explicit and parameterized convection.

6.1 Data and methods

6.1.1 COBALD aerosol measurements

In August 2016, 15 balloons were launched in Nainital, Uttarakhand, India (29.35°N, 79.46°E, 1820 m above sea level) (Brunamonti et al., 2018). The balloons carried the COBALD, which is a lightweight backscatter sonde (Brabec et al., 2012). It measures the backscatter at 940 nm (infrared) and 455 nm (blue visible) in proximity to the balloon. For the detection of the ATAL the short wavelenghth channel (455 nm) is used (details see Hanumanthu et al., 2020). Furthermore,

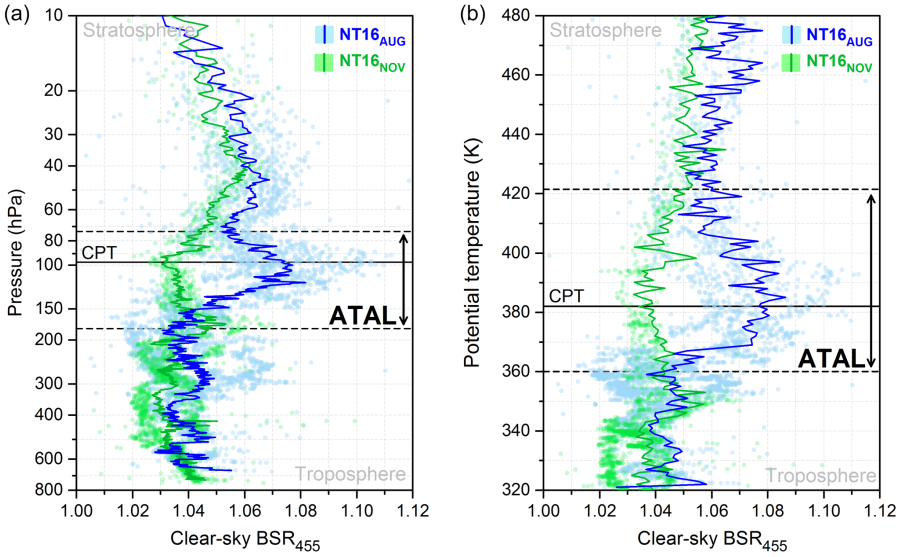


Figure 6.1: Backscatter ratio profile of the measurements in Nainital, India in August 2 016 *with (a) pressure and (b) theta as height coordinate. The green line shows the average background backscatter ratio after the monsoon season in November and the blue lines shows the average backscatter ratio during the presence of the ATAL in August. Dots show individual measurement points. The cold point temperature (CPT) is shown by the horizontal black line and the extend of the ATAL is marked by the horizontal dashed black lines and the vertical arrows. Adopted from (Hanumanthu et al., 2020) and (Brunamonti et al., 2018).*

the balloon carried a RS41-SGP radiosonde that logged local temperature and pressure. The backscatter signal can be expressed as the backscatter ratio (BSR). The BSR is the ratio between the total backscatter due to aerosols and air molecules and the backscatter due to air molecules alone. Based on calculations of Bucholtz (1995), the BSR has been inferred from the temperature and pressure of the radiosonde and the backscatter of COBALD. Furthermore, a color index (CI), i. e., the ratio between the 940 nm and 455 nm aerosol backscatter has been calculated by Hanumanthu et al. (2020), because it allows to discriminate between large aerosol particles and smaller ones and accordingly between layers of cirrus clouds and ice-free layers of the ATAL. This analysis provided vertical profiles of the aerosol layer for 15 days in August 2 016. Table 6.1 and Fig. 6.1 give a short overview over the measurements. While the BSR is determined for every measurement point during the ascent, here the daily, vertical average of the BSR for each balloon flight were used. Hence, the day to day changes of the measured BSR profiles and not the BSR for every measurement time-step individually were analysed. For further details see Hanumanthu et al. (2020).

6.1.2 Simulation scenarios

For the Lagrangian backward trajectory calculations with CLaMS and MPTRAC, the air parcels are initialized at positions of the measured ATAL in August 2 016 on 15 measurement days. For every measurement time of the COBALD instrument, i.e. every second, one air parcel is initiated. During a flight the balloons horizontal drift is below 10 km in the ATAL. The differences are below 50 km in the ATAL when different balloons, from different days, are compared and hence, they are as well negligible. Table 6.1 shows an overview over the measurements and the number of air parcels initialized per day. Two balloon flights are discussed separately; the flight on 12 August, when a large cirrus cloud covered the full UTLS in the sampled region and 15 August, when no ATAL was detected (for details see Hanumanthu et al., 2020). All air parcels are each calculated backward for 90 days to cover the entire Asian monsoon period JJA.

Different scenarios for the calculations have been applied to study the impact of the reanalysis data (ERA5 vs. ERA-Interim), Lagrangian model differences (CLaMS vs. MPTRAC), vertical velocities (diabatic and kinematic) and parameterisations (convection vs. no convection) and the size of the time-step (180 s vs. 1 800 s) on the simulated transport. Here the default configurations of the models is considered for the integration time-step. Additionally, one scenario with MPTRAC is included, for which instead of initialising one air parcel per measurement point, 1 000 air parcels are initialized and modules for sub-grid scale fluctuations and turbulent diffusion were applied. With this ensemble approach, sampling uncertainties were estimated. For the particle diffusion the default settings were used (see Hoffmann et al., 2022). A summary of all scenarios can be found in Table 6.2.

Table 6.1: Overview on COBALD measurements in August 2 016*.* $\overline{\mathrm{BSR}}$ *is the daily, vertically averaged BSR. The number of measurements i.e. the number of initialized air parcels per flight is labeled #AP.*

day	02	03	05	06	08	11	12	15	17	18	19	21	23	26	30
$\overline{\mathrm{BSR}}$	6.7	9.2	6.7	8.3	7.1	7.0	-	2.3	7.6	6.5	7.3	5.6	5.4	8.0	5.9
#AP	670	419	413	385	680	269	463	444	651	705	569	250	331	120	392

Table 6.2: Overview over scenarios of 90 days backward calculations performed for the ATAL measurements above Nainital in August 2 016*. The abbreviation for each scenario contains at the first position the reanalysis, at the second position the vertical velocity and at the third position the model, where each label is separated by a dash. Optional properties are added the same way at the last position.*

abbreviation	reanalysis	vertical velocity	model	time-step	convection parameterisation	dispersion
EI-kin-C	ERA-Interim	kinematic	CLaMS	1 800 s	off	off
EI-kin-M	ERA-Interim	kinematic	MPTRAC	180 s	off	off
EI-dia-C	ERA-Interim	diabatic	CLaMS	1 800 s	off	off
EI-dia-M	ERA-Interim	diabatic	MPTRAC	180 s	off	off
E5-kin-C	ERA5	kinematic	CLaMS	1 800 s	off	off
E5-dia-C	ERA5	diabatic	CLaMS	1 800 s	off	off
E5-kin-M	ERA5	kinematic	MPTRAC	180 s	off	off
E5-dia-M	ERA5	diabatic	MPTRAC	180 s	off	off
E5-kin-M-ECP	ERA5	kinematic	MPTRAC	180 s	on	off
E5-kin-M-1800s	ERA5	kinematic	MPTRAC	1 800 s	off	off
E5-kin-M-Diff	ERA5	kinematic	MPTRAC	180 s	off	on
E5lr-kin-M	ERA5 1 °×1 °	kinematic	MPTRAC	180 s	off	off
E5lr-dia-M	ERA5 1 °×1 °	diabatic	MPTRAC	180 s	off	off

6.1.3 Classification of air parcel origin

The origin of the air parcels found in the ATAL is classified vertically and horizontally based on the 90 day backward trajectory calculations. Vertically, following Hanumanthu et al. (2020) with four classes, the MBL, the lower troposphere (LT), the upper troposphere (UT) and the lower stratosphere (LS) are defined. These vertical layers are defined by values of the vertical hybrid zeta coordinate and the potential temperature θ as presented in Table 6.3. The MBL is defined as the layer below the 120 K zeta level, which approximately corresponds to heights between 2 km

and 3 km. Accordingly, an air parcel is considered to originate from the MBL if it is located at any time below the 120 K zeta level (Pommrich et al., 2014; Vogel et al., 2019; Hanumanthu et al., 2020).

Hanumanthu et al. (2020) used backward trajectory times between 40 and 80 days. The calculations was extended to 90 days for each air parcel to completely cover the entire Asian summer monsoon season in the analysis. It was found that a majority of the transport from the MBL to the ATAL occurred within 90 days (i.e. 60%-90% of air parcels reach the MBL during that time), with only low increase when longer integration time is used (lower than around 5 percentage points per 10 additional days).

When an air parcel is classified as originating from the MBL, it is also horizontally classified according to the position where it left the MBL for the last time. For this position, several possible source regions are defined. The regions were motivated by different surface characteristics, such as the presence of aerosol or aerosol precursors and by source regions proposed by earlier studies (e.g. Li et al., 2017b; Höpfner et al., 2019; Li et al., 2020; Hanumanthu et al., 2020). For the continents, the following regions are defined: the Asian Highlands, i.e. mostly the Tibetan Plateau and its immediate surroundings (around 60% of the highlands area), the Indo-Gangetic plain together with the foothills of the Himalayas, a region South of India plus Sri Lanka and finally South East Asia. In particular, for the Asian highlands it can be found that 75% of the air parcels originate from the Tibetan Plateau (heights larger than around 4 km) and up to 98% originate from the Tibetan Plateau and its immediate surroundings. Other parts of the continents are summarized as the residual continent. For the oceans, three regions were defined: the Tropical Western Pacific that has been affected by typhoon activity, the Arabian Sea, the Bay of Bengal and the residual oceans. Figure 6.4a illustrates the different regions. For more detail of the definitions, see Appendix D.1.

Additionally, for those air parcels that do not originate from the MBL, but instead circulate still

Table 6.3: Classes for vertical classification of the distributed air parcels. Air parcels have to fulfil all he criteria to be attributed to a specific class. ζ is the vertical zeta coordinate, θ is the potential temperature, λ is the longitude and ϕ is the latitude of an air parcel.

class	ζ-criterion	θ-criterion	lon/lat criterion	abbrev.
Model boundary layer	$\zeta \leq$120 K			MBL
Lower troposphere	$\zeta >$120 K	$\theta \leq$340 K		LT
Asian summer monsoon anticyclone	$\zeta >$120 K	$\theta >$340 K	$0 \leq \lambda \leq 135°$, $0 \leq \phi \leq 45°$	ASMA
Upper troposphere	$\zeta >$120 K	340 K$< \theta \leq$370K	not in ASMA	UT
Lower stratosphere	$\zeta >$120 K	370 K $> \theta$	not in ASMA	LS

in the ASMA after 90 days of backward trajectory time, the class of the ASMA is defined. For this purpose, the 3D box from 0° E to 135° E, and from 0° N to 45° N (magenta box in Fig. 6.4a) within the UTLS region is considered. Each air parcel within this box is considered to be part of the ASMA.

6.2 Results

In the following, transport pathways, transport times and possible surface source regions of air mass contributions to the ATAL above Nainital in August 2016 are presented. Furthermore, the relation between the observed ATALs backscatter intensity and different MBL regions is analysed. The analysis is performed for different simulation scenarios as described in Sect. 6.1.2.

6.2.1 Transport pathways from source regions to the measured ATAL

The ASMA extends from northeast Africa to the Western Pacific from early June until the end of September, therefore air parcels circulate in the ASMA over a wide range of longitudes and latitudes. Depending on its extension and position, convection can uplift air from different regions of the Earth's surface - i.e. with different chemical composition - into altitudes of the anticyclone. Within the ASMA, the air from different origins will be mixed, for example due to instabilities of the ASMA (e.g. Gottschaldt et al., 2018). Different source regions can contribute to the chemical composition of the ASMA, therefore trace gases and aerosol are in general not homogeneously distributed within the ASMA. The ASMA can show a bimodal structure, where one circulation centre is placed roughly over Iran and the other one is centred roughly over South-East-Asia and where the separation of the two modes varies in time. The shape of the ASMA varies between normal and bimodal on a daily basis, however in the climatological mean it is controversely discussed if two modes exist. In addition smaller eddies can be separated to the west and to the east (sometimes referred to as third mode) (Yongfu et al., 2002; Nützel et al., 2016; Honomichl and Pan, 2020; Manney et al., 2021).

Figure 6.2 shows four exemplary transport pathways of air parcels from different regions at the MBL to the measured ATAL over Nainital in August 2016 due to convective transport. For the illustration of transport pathways the scenario E5-kin-C is used, because it is representative with regard to the general patterns also for the other scenarios, while particular differences are analysed in depth later. Included are also trajectories of all days of measurement for the illustration.

Figure 6.2a shows injections into the center of the ASMA in proximity of Nainital, originating

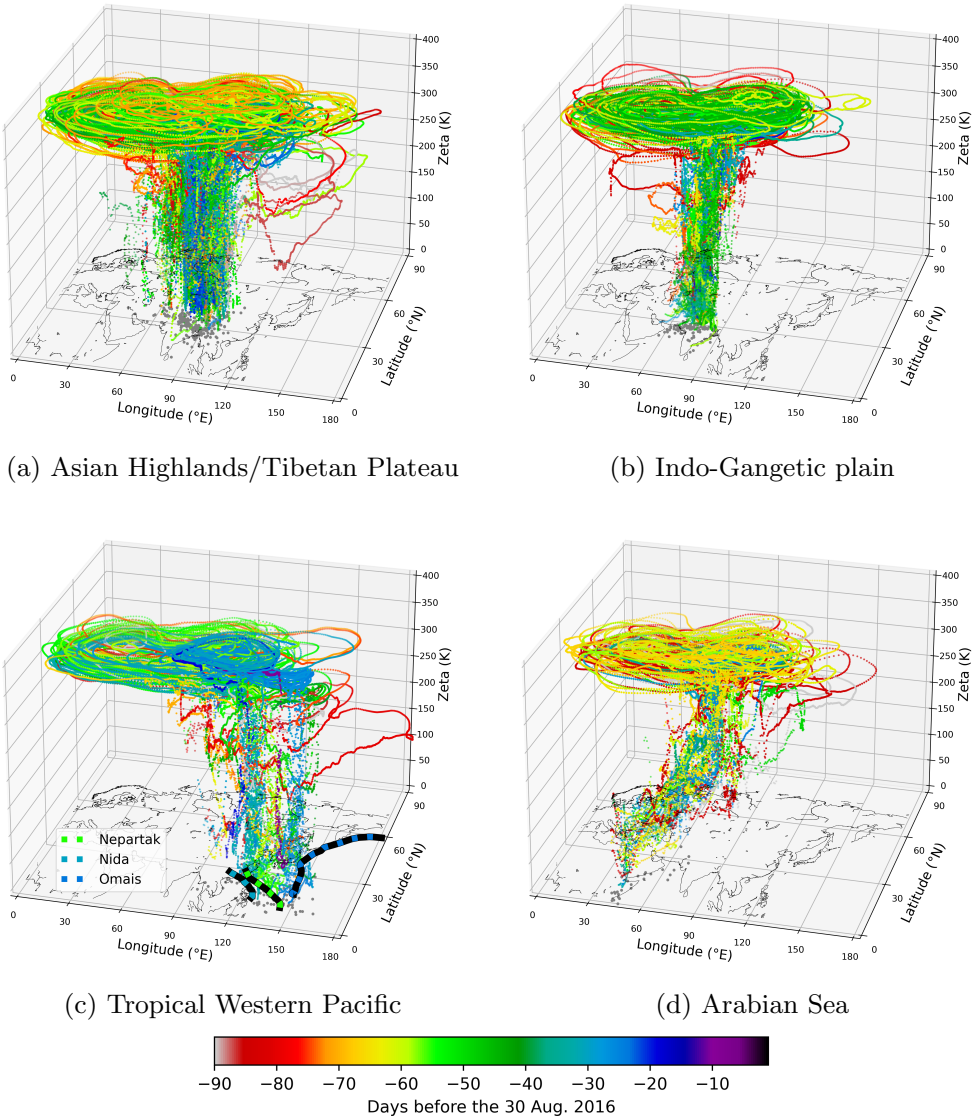
from the Tibetan Plateau. Throughout the season the air is pumped upward into the ASMA on timescales from hours to a few days (50% of transport from the MBL of Asian Highlands into the UTLS is within less than half a day, see also Tables D.4 and D.3). In the ASMA, air masses are uplifted by diabatic heating superimposed by the anticyclonic flow until they meet the measurement points over Nainital. Air parcels circle in a rising spiral, within the two modes of the ASMA, until they meet the measurement points over Nainital (e.g. Vogel et al., 2019). A small number of air parcels leaves the ASMA for a while, and is subsequently transported along the subtropical jet, circumnavigating the Earth, until the air parcels are trapped in the ASMA again. Thereafter, the air parcels also arrive at the measurement points.

Figure 6.2b shows the uplift of air into the ASMA mostly over the Indo-Gangetic plains and at the foothills of the Himalayas. The air masses were transported mainly directly from the Indo-Gangetic plain into the UTLS. The transport in the ASMA and sporadically along the jet-streams is the same as for the Tibetan Plateau. Hence, at the foothills of the Himalayas, transport pathways from two regions with two very different land-cover properties converge, the Tibetan Plateau and the Indo-Gangetic plain. The transported air masses subsequently mix in the ASMA.

Figure 6.2c illustrates transport from the Pacific to the measurement locations in relation to three typhoons (named Nepartak, Nida and Omais) that occurred during the relevant time period (the typhoon best track data of the Japan Meteorological Agency is used). The typhoons uplift a large number of air parcels from the maritime surface into the eastern edge of the ASMA. Depending on the position of the ASMA modes and the typhoons, the uplifted air masses are circulating in the outer edge of the ASMA (e.g. for Nida and Nepartak) or they circle mostly in the eastern mode, and the inner area of the ASMA (e.g. Omais). Because of the multiple circulations in the ASMA the typhoons influence the measurements with a delay in time of several days. The impact of typhoons on the air masses in the ASMA and the ATAL has been reported before (e.g. Li et al., 2020; Hanumanthu et al., 2020).

Figure 6.2d presents a transport pathway from the Arabian Sea to the ASMA. The transport of the air parcels occurs in four steps. At first, the air leaves the MBL into the free troposphere over the Arabian sea, possibly due to shallow, maritime convection. Secondly the air is transported eastward within the free troposphere to the foothills of the Himalaya or to the Bay of Bengal. During this transport, over the Indo-Gangetic plain, or at the Bay of Bengal, deep convection uplifts the air to the southern edge of the ASMA, which is the third step of transport. In the last step, the air masses circle in the ASMA until they are measured at Nainital. A similar long-range transport pathway to the Himalayas can be observed for the Bay of Bengal. However from the

Bay of Bengal more air parcels can enter the ASMA directly from the maritime boundary than from the Arabian Sea. This is likely related to the overall Asian monsoon circulation, in which air masses are transported in the troposphere from the AS across India, while the Bay of Bengal is a known source of deep convection. Moreover, those air masses that are convectively uplifted into the UTLS over the AS, are often located at the outer edge of the anticyclone and westward from Nainital. Hence, this transport pathway to Nainital is much less probable. In Appendix D.2, the transport pathways from the Bay of Bengal, South India, South-East-Asia and the remaining ocean and continent are presented as well.



(a) Asian Highlands/Tibetan Plateau

(b) Indo-Gangetic plain

(c) Tropical Western Pacific

(d) Arabian Sea

Figure 6.2: Examples of backward trajectories of air parcels of all of the 15 measurement days from the ATAL measurement to the MBL, categorised by the source region. Shown are trajectories of scenario E5-kin-C. Colors indicate the time when the air parcels left the MBL. Gray dots at the bottom show the horizontal position of the APs 48 h before they crossed the MBL from below. In (c), additionally, tracks of three typhoons are plotted (Nepartak, Nida, Omais), where each point is colored like the trajectories and the mean time of occurrence.

6.2.2 Scenario intercomparison of contributions from source regions and related transport pathways

Although the general transport pathways from the MBL to the measurement locations, as presented in Sect. 6.2.1, exist in all simulation scenarios with the Lagrangian transport models, the contributions of different source regions can differ, depending on the used scenario. Here, dif-

ferences and similarities of the vertical and horizontal distribution of the source regions between the different scenarios are presented. For the analysis, all 15 measurement days and 90-day backward trajectories are used.

First, the fraction of air from different atmospheric layers is calculated for all model scenarios (see Fig. 6.3). Wide agreement can be found with ERA5 even when models, integration step-sizes, and vertical velocities are varied (except when the extreme convection parameterisation is employed). The total amount of transport from the MBL lies between 74% and 80% for the ERA5 scenarios. The distribution from the LT, UT, or LS shows only some minor differences. Large disagreement is found between the diabatic and kinematic vertical velocities using ERA-Interim. With kinematic vertical velocities, only around 60% of the air parcels originate in the MBL, while with diabatic velocities the results are closer to the ERA5 scenarios (around 75% amount of transport from the MBL). Disagreement in ERA-Interim, when varying the vertical velocities, was also found in other studies (e.g. Ploeger et al., 2010; Li et al., 2020; Legras and Bucci, 2020). The low-resolution ERA5 data set maintains the higher consistency of ERA5 in comparison to ERA-Interim and the total transport from the MBL is only slightly reduced, probably because the vertical resolution is unchanged and higher vertical velocities over the continent remain also higher with ERA5 $1° \times 1°$ as well. A large increase of transport from the MBL is caused by the onset of the convection parameterisation in scenario E5-kin-M-ECP. In this case only around 10% of the air parcels originate outside the MBL and no air parcels originate in the LT. The latter can be explained, when following the air parcels backward in time: If the air parcels enter the LT during backward calculations they likely enter a region where the ECP is triggered. Then transport into the MBL takes place immediately.

Second, the fraction of air from different MBL regions contributing to the ATAL is compared for all model scenarios (see Fig. 6.4b). Model scenarios driven with ERA5 show very similar results. For ERA5 scenarios, about 40% of the air parcels that originate from the MBL, originate from the Tibetan Plateau and other Asian highlands. Around 5% to 10% of air parcels come each from South India and South East Asia. The contributions from the Indo-Gangetic plain is between 10% and 20%, with higher values for diabatic calculations. The Indian subcontinent, i.e. the Indo-Gangetic plain and South India together have the largest contribution from the continent. Around 25% of the air parcels that come from the MBL originate from oceans, mostly the Western Pacific and many are related to tropical cyclones. In contrast to ERA5, using ERA-Interim increases the contribution of air parcels from the oceans up to around 40%. This increase does not rely on the variation of the vertical velocity and is robust for all scenarios driven with ERA-

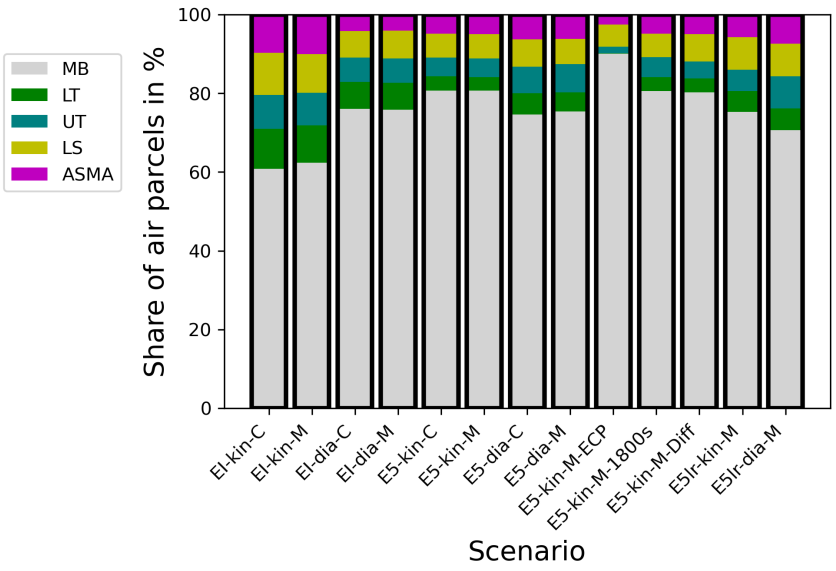
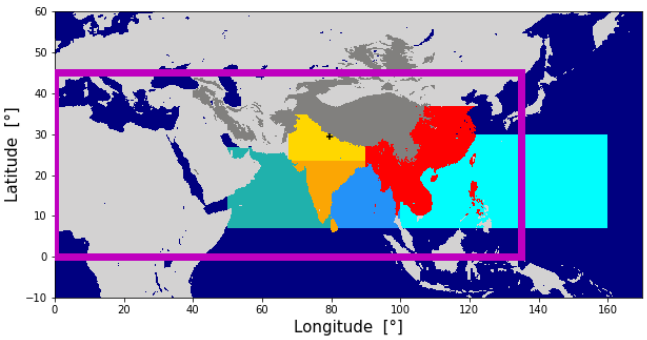


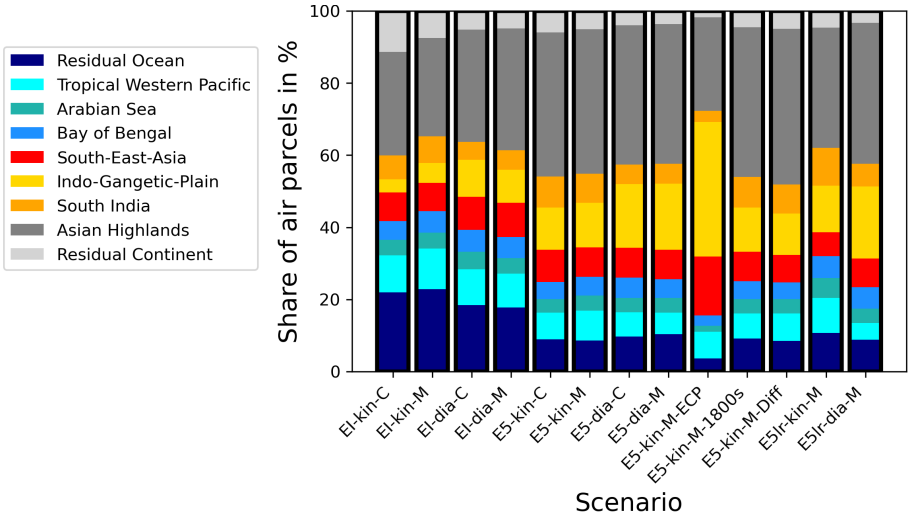
Figure 6.3: Vertical classification of air parcel origin after 90 days of backward trajectory calculations with different scenarios.

Interim. The contribution from the Indo-Gangetic plain is reduced to 5%-10% with ERA-Interim compared to ERA5, and shows also a difference between diabatic and kinematic velocities. With the extreme convection scenario the contribution from the oceans is smaller than in the cases with ERA5. However, with the ECP the contribution from the Indo-Gangetic plain and the foothills of the Himalayas increases strongly at the expense of contributions from the Tibetan Plateau and the oceans. As a consequence of the persistent occurrence of CAPE above South Asia, the scenario with ECP simulates more and deeper convective updrafts in this region, than the scenarios without ECP. Hence, transport from the MBL that would be missed without the ECP, increases the contributions from South Asia. Finally, the ERA5 low-resolution scenarios show a decrease of consistency between the diabatic and kinematic approach. The kinematic approach has a bias to more transport from the ocean in the low-resolution data in comparison to the fully resolved ERA5 data, while the results for the diabatic approach show only minor difference to the fully resolved ERA5 data in the statistics.

With the ensemble scenario (E5-kin-M-Diff), uncertainties due to unresolved sub-grid scale wind fluctuations were tested and it was found a standard deviations lower than 1% for the distribution of air parcels to different vertical layers or horizontal regions. Therefore, the sub-grid scale wind fluctuations do not cause a serious bias in this analysis. The ensemble scenario E5-kin-M-1800s was used to investigate possible biases between the models due to different time-steps. However, as the results remain almost unaltered for MPTRAC when the time-step is varied from 180 s to 1800 s a serious bias for the analysis is unlikely.



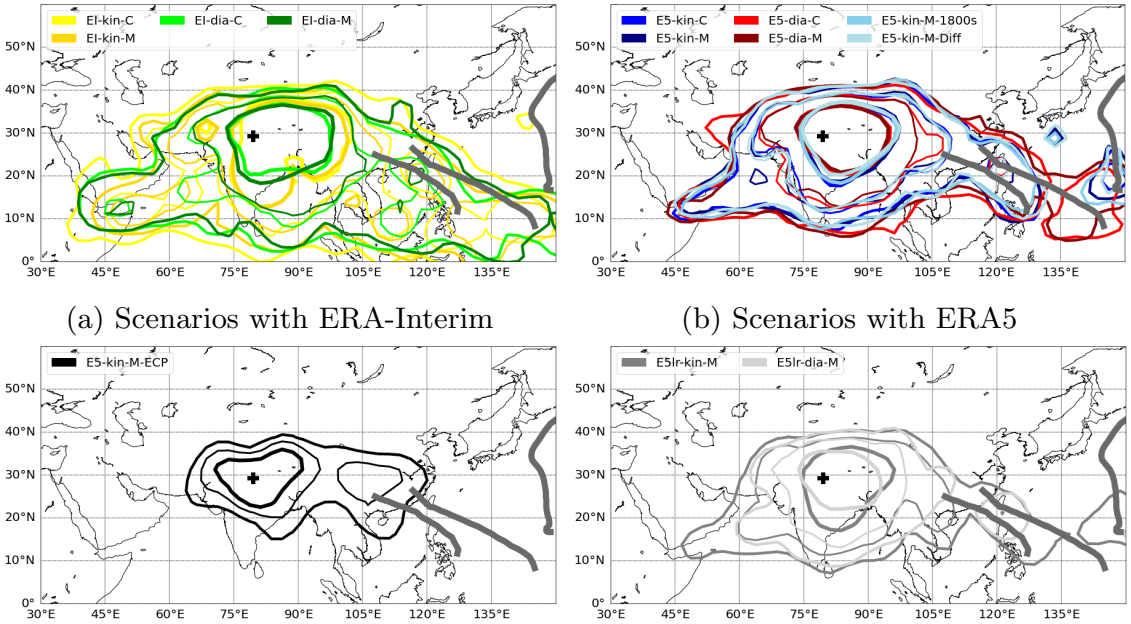
(a) Definition of the Regions



(b) Classification after 90 days

Figure 6.4: (a) shows the definition of contributing regions. The purple box marks an area that contains most of the air parcel that circulate in the ASMA. (b) shows the horizontal classification of the air parcels according to the surface regions after 90 days backward trajectory time. Shown is the fraction of air parcels that reach the MBL.

The probability density function (PDF) for air parcels leaving the MBL is shown in Fig. 6.5. All scenarios except the ECP scenario show a very similar pattern, with most dominant transport from a region centered at the eastern foothills of the Himalayas and the south-eastern Tibetan Plateau (around 50%), while transport from other regions is much less but not irrelevant. If contours for ERA5 scenarios and ERA-Interim scenarios are compared, air parcels are dispersed more with ERA-Interim, particularly in direction of the ocean, while ERA5 resolves more transport at the continent and disperses the air parcels less. Using the scenario with ECP deforms the pattern even more, because more transport happens then at the Indo-Gangetic plain, due to higher CAPE in this region than at the Tibetan Plateau. The results of the ECP depend on the selection of



(a) Scenarios with ERA-Interim

(b) Scenarios with ERA5

(c) Scenarios with ERA5 and ECP

(d) Scenarios with low resolution ERA5

Figure 6.5: Contours of the PDF of the surface sources for 90 days backward trajectories and all measurements. Colors indicate the different scenarios. For each scenario the inner, thick contour encloses 50% of the points where air parcels left the MBL and the outer contour encloses 90% of them. In between, a thin contour is shown for 75%. For the sake of clarity, the contours are smoothed by a Gaussian kernel. The black cross indicates the position of Nainital. The three thick gray lines show typhoon tracks.

the CIN and CAPE thresholds. The CIN threshold was used to remove spurious parameterized convection over the Persian Gulf (see Appendix D.6).

In summary, ERA5 provides improved robustness against changes of the vertical velocity between the kinematic and diabatic approach in comparison to ERA-Interim, and yields very good agreement between the two Lagrangian models. Using the ECP in MPTRAC indicates that even scenarios with ERA5 could miss effects of unresolved convection, particularly locally over the Indo-Gangetic plain. This leads to difficulties when distinguishing the contributions from the Tibetan Plateau and the Indo-Gangetic plain. When air parcels passing the region around the Himalayas during the backward trajectory calculations, the position where they are transported back into the MBL is sensitive to the representation of the convection. With the ECP, convection is enhanced over the IGP. Hence, far more air masses are attributed to the IGP with the ECP, while without the ECP the most air parcels originate from the TP. Further improvement in reanalysis and parameterisation are needed to remove this uncertainty.

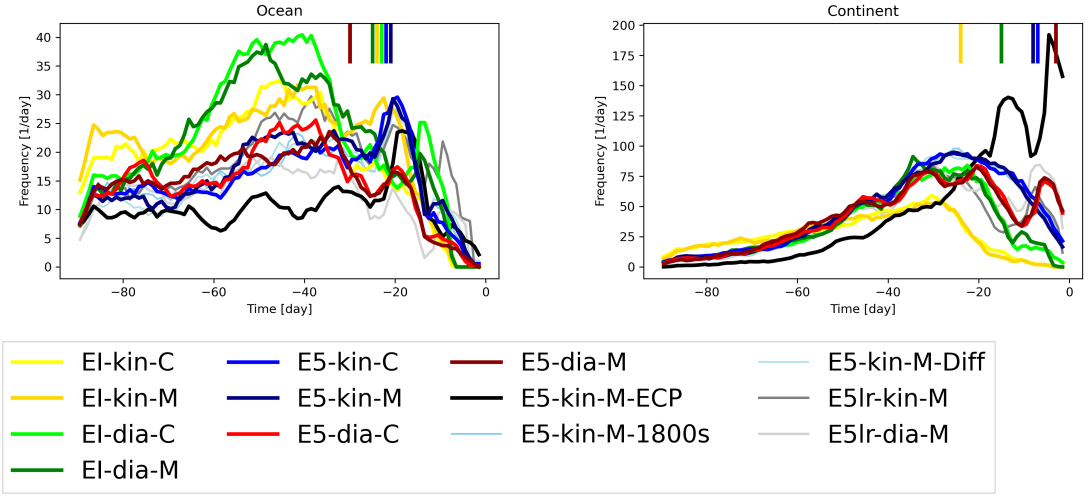
6.2.3 Scenario intercomparison of the temporal evolution of transport from the MBL to the measured ATAL

The transport time of the air parcels from the MBL to the ATAL affects aerosol formation. The time of the transport is therefore an important parameter to analyse. An analysis of the temporal evolution of the transport process from the MBL to the measurements also highlights some differences between the model scenarios.

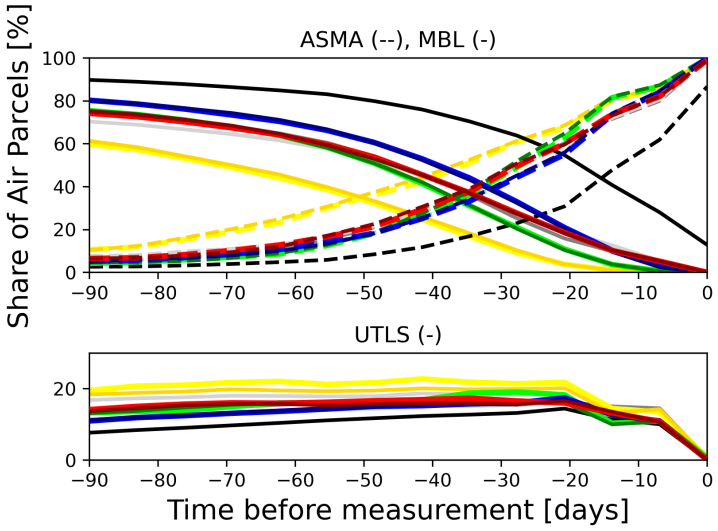
To emphasize the possible lifetimes of air masses transported to the location of the measurements in the UTLS, Fig. 6.6a shows the frequency of air parcels leaving the MBL per day at different times before the measurements for the different scenarios. These frequencies are classified in two categories, the continents and the oceans. Most of the relevant maritime convection (e.g. typhoons) that transport air masses out of the MBL into the upper atmosphere take place more than two weeks before the measurements. This can be found for all scenarios. For scenarios with ERA-Interim the frequency of air parcels leaving the MBL is higher than for ERA5 if transport times longer than 40 days are considered, in particular for the diabatic scenarios with ERA-Interim.

For ERA-Interim only a few air parcels originate from the continent with transport times less than two weeks, independent of the used vertical velocity (diabatic or kinematic). This resembles the results of Hanumanthu et al. (2020), who used diabatic CLaMS trajectories driven by ERA-Interim. In contrast, all scenarios with ERA5 show that air from the continent can be transported much faster to the location of the measurements even in much less than two weeks. This is likely due to a better representation of convection in ERA5 in comparison to ERA-Interim. This fast transport at the beginning is maintained with the low-resolution ERA5 data, although it is reduced in temporal and spatial resolution in comparison to the full ERA5. Furthermore, using the ECP, reveals that also ERA5 potentially underestimates convection in the first days. The fast transport with ECP is caused by the frequent triggering of the parameterisation over the continent, particularly over the Indo-Gangetic plain, where the atmosphere often shows convective parameters (CAPE, CIN) that suggest unstable conditions. However, this approach of the ECP has to be considered as the upper limit for the convective transport that can be simulated within the given model framework. In summary, in ERA5 air masses are transported from the continental MBL to the ATAL relatively fast (within two weeks), so that less air parcels remain to originate from the maritime MBL, while with ERA-Interim this effect reverse. With ERA-Interim only few air parcels are transported fast to the ATAL from the continent, while more and older air parcels originate from the oceans.

Figure 6.6b shows the accumulation of air parcels within the ASMA during the transport processes to the ATAL over Nainital. The differences between the scenarios can be understood when the transport is described following the calculations backward in time, i.e. when we look at the “draining” of the ASMA during backward calculations. The calculations of this thesis show that in all scenarios most of the air parcels start in the ASMA. When the air parcels are traced back in time, the share of air parcels in the ASMA for scenarios with ERA5 and ERA-Interim starts to diverge. In the first two weeks backward in time (-14 days to 0 days), scenarios with ERA5 show more transport back into the MBL than scenarios with ERA-Interim, i.e. the share of air parcels in the MBL increases faster with ERA5 and the share in the ASMA declines faster than with ERA-Interim. This is in agreement with faster transport from the continent with ERA5 than with ERA-Interim as described before. After calculating the trajectories further back in time (longer than two weeks), the share of air parcels in the ASMA starts to converge again for the ERA-Interim scenarios with diabatic velocities (EI-dia-C, EI-dia-M) and the ERA5 scenarios. This convergence is partly caused by the increased backward transport to the maritime MBL in the scenarios with ERA-Interim and the diabatic scheme, while with ERA5 the transport to the maritime MBL is smaller in comparison (see also Fig. 6.6a, left panel, -70 days to -30 days). The ERA-Interim scenarios with kinematic approach (EI-kin-C, EI-kin-M) diverge further from all other scenarios, showing a much lower share of air parcels in the MBL than other scenarios and the most air parcels in the UTLS. After around two weeks the share of air parcels in the UTLS increases more in the scenarios with ERA-Interim and kinematic vertical velocities (see lower plot in Fig. 6.6b) than in all other scenarios. Ploeger et al. (2010, 2011) demonstrated more dispersion in backward trajectory calculations with ERA-Interim and the kinematic approach than with the diabatic approach. This effect explains the large difference of the scenarios EI-kin-C and EI-kin-M from the other scenarios and why many air parcels are transported back into the UTLS with these scenarios.



(a) Horizontal Classification



(b) Vertical Classification

Figure 6.6: Time evolution of transport from the MBL to the ASMA within 90 days, relative to the start of the trajectories. (a) shows the frequency of air parcels leaving the MBL smoothed with a one-week running mean. The short lines at the top indicate for some scenarios the maximum time for the first 300 air parcels with the smallest transport time. (b) shows the share of air parcels that are within the ASMA (upper panel, dashed line), the share of air parcels that are in the MBL (upper panel, solid line) and the share of air parcels that are in the UTLS (lower panel, solid line).

6.2.4 Backscatter changes associated with changes in the transport and source regions

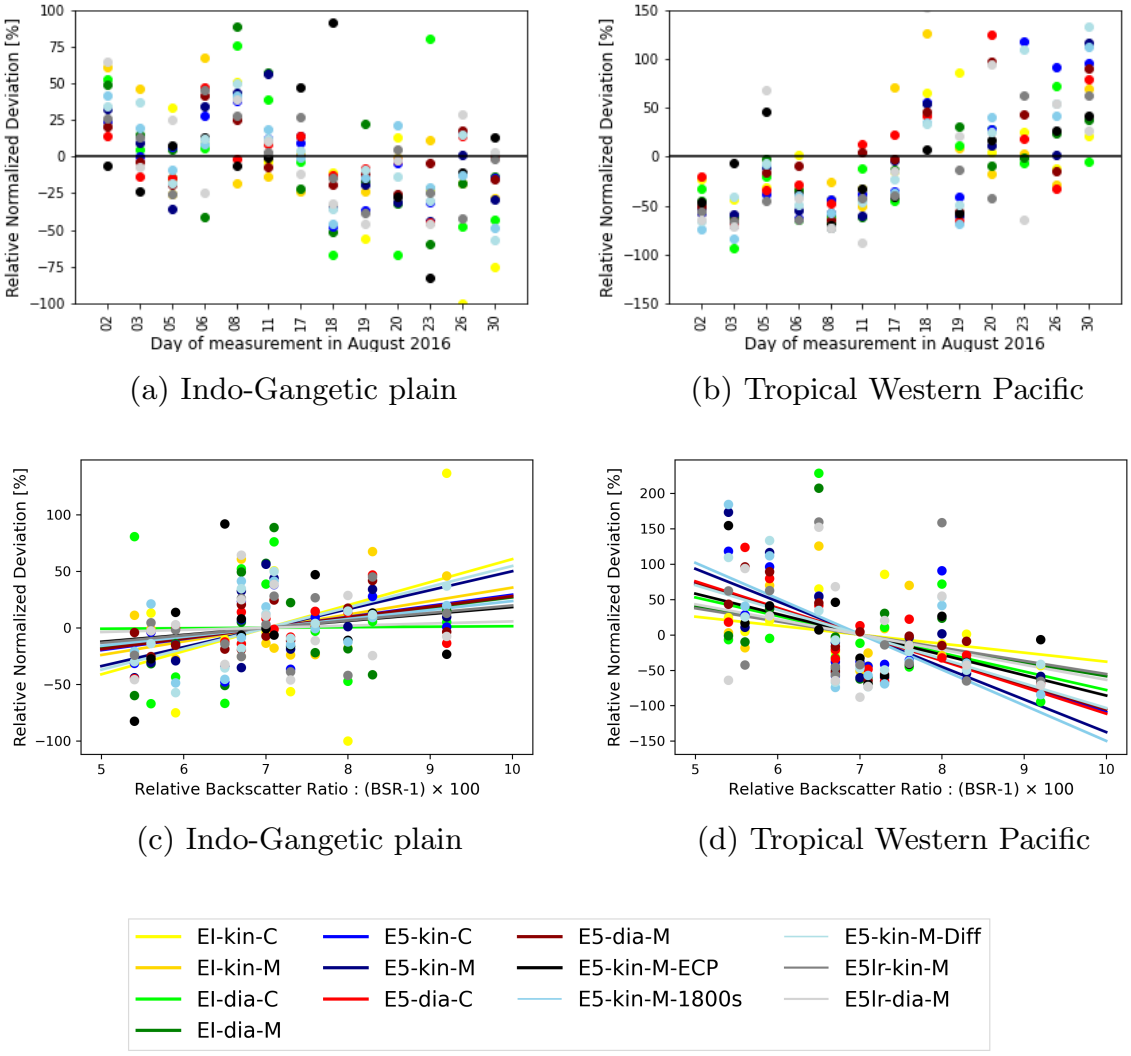
Hanumanthu et al. (2020) found a distinct day-to-day variability of the ATAL backscatter intensity. This variability may be correlated with daily to weekly changes of the transport within the highly variable anticyclone and variability of tropical convection and therefore with changing surface source regions. To analyse the changes from measurement to measurement, the contribution of different source regions to the vertical ATAL profile has been reconstructed for every balloon flight separately. Subsequently, the relative deviation of the contribution of a specific region on one day from the mean contribution during all measurements normalized by the mean contribution was calculated as a relative, normalized deviation (in percent). Accordingly, $RND(t) = \left(\frac{C(t)}{\overline{C(t)}} - 1\right) \times 100$ was calculated for every day of a selected scenario. $RND(t)$ denotes the relative, normalized deviation for measurement day t and $C(t)$ denotes the contribution of air parcels for the measurement day t from the selected region. The contribution $C(t)$ is measured as the ratio between the number of air parcels from the selected region to the total number of air parcels for the measurement day. $\overline{C(t)}$ is the time averaged contribution over all measurement days for the selected scenario. This calculation has been done for all scenarios separately to allow a direct day-to-day intercomparison of the scenarios.

Figure 6.7 shows the normalized deviations for the Indo-Gangetic plain and the Western Pacific Ocean for 13 measurement days. The days with no ATAL or with large cirrus cloud coverage have been excluded from the analysis (12 and 15 August). Furthermore, it is a focus on the Indo-Gangetic plain and the Western Pacific Ocean as they show the most robust and significant results in comparison to other source regions.

Figures 6.7a,b show the relative, normalized deviation for every measurement day. Overall, all scenarios indicate a clear temporal evolution of the contribution from the two regions during the weeks of the campaign in August: The scenarios show that in the early phase of the campaign (2, 3, 6 and 8 August) the contributions from the Indo-Gangetic plain were enhanced relative to other days, while in the later phase (19-30 August) contributions from the Indo-Gangetic plain were relatively low. For contributions from the Western Pacific the opposite is found, because of increased impact of typhoons on the measurements at the end of August (see Fig. 6.7a,b). From day to day, the relative normalized deviation changes in absolute terms with the order of magnitude of 10% (i.e. 10-100%), while the total change of the relative normalized deviation during the period is roughly 90%. For some measurement days, the scenarios show similar day-to-day differences (e.g. 2-8 August for the Tropical Western Pacific), but for other periods the

day-to-day variability differ strongly between scenarios (e.g. 23-30 August for the Tropical Western Pacific).

Figures 6.7c and 6.7d show the relation between the averaged measured BSR for every day and the daily relative, normalized deviation. For the Indo-Gangetic plain in all scenarios large backscatter ratios coincide with large normalized deviations, although in some scenarios this relation is considerable weak. For the Tropical Western Pacific, in all scenarios low backscatter ratios clearly coincide with higher normalized deviations.



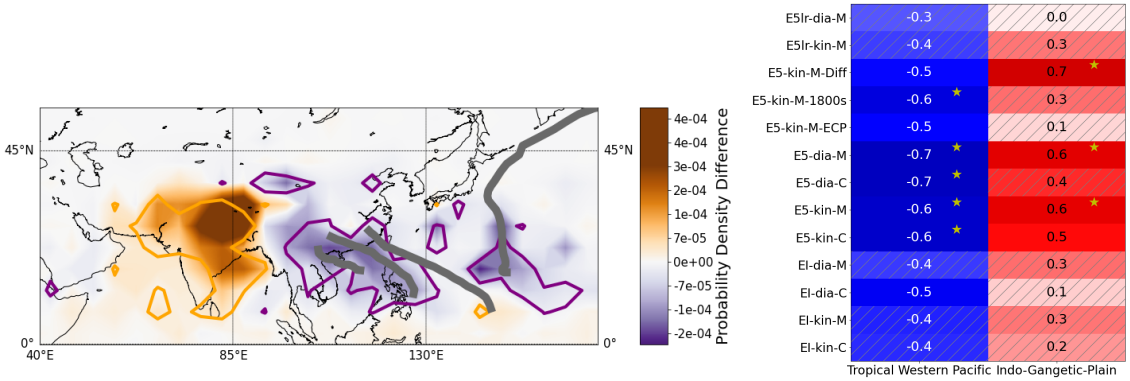
(a) Indo-Gangetic plain

(b) Tropical Western Pacific

(c) Indo-Gangetic plain

(d) Tropical Western Pacific

Figure 6.7: (a) and (b) show the relative, normalized deviation for each measurement day of the Indo-Gangetic plain and the Tropical Western Pacific. (c) and (d) show the relations between the daily averaged backscatter intensity and the relative, normalized deviation. The lines show linear fits for each scenario. Colored dots show results for all scenarios. The colorcode for the dots is the same as in Fig. 6.6, where every color correspondence to one scenario.



(a) (b)

Figure 6.8: (a) Difference between the source region PDF of the five days with maximum ATAL backscatter intensity and five days with minimum backscatter intensity, derived from the multi scenario mean. The orange contours show areas where at least two thirds of the scenarios indicate high values during a strong ATAL. The purple contours show the same for the weak ATAL cases. Data is given on a 5° grid. The gray lines indicate the tracks of tropical storms that influenced the measurements. (b) Spearman correlation coefficient for the relation between the daily BSR and the contributions for the Western Tropical Pacific (left) and the Indo-Gangetic plateau (right) with different scenarios. Crossed out areas hatch insignificant results (p-value higher 0.10). Boxes with yellow stars indicate p-values lower than 0.05. Colors emphasize positive (red) and negative (blue) correlations.

To further compare transport of those days with a low ATAL backscatter intensity above Nainital with those days with a high one, the PDFs of surface source regions for each day were calculated separately. In particular, the five days with strongest (3, 6, 26, 17, 19 August) and the five days with the weakest backscatter intensity (21, 23, 30, 18, 5 August) were selected for the analysis. Figure 6.8a shows the differences between the two PDFs for the multi-scenario mean. The results indicate that transport from the Indo-Gangetic plain, India, the Bay of Bengal and the northern part of the Tibetan Plateau may have strengthened the ATALs backscatter intensity during the measurements. The intensity of the ATAL over Nainital again seems to be low when more transport from the tropical cyclones and the region of their landfall occurred. Although the PDF differences vary strongly from scenario to scenario (see Fig. D.4 in the Appendix), at least two third of the scenarios agree on these characteristics.

For a complete analysis, the Spearman rank-order correlation coefficients were used to check if there is a monotonic relation between the measured backscatter and the contribution of specific regions. Figure 6.8b summarizes these correlations for the Indo-Gangetic plain and the area of active typhoons at the Tropical Western Pacific. Correlations to other regions can be found in the Appendix in Fig. D.5.

A significant ($p<0.1$) and robust negative correlation between the backscatter intensity of the

ATAL and the Western Pacific influenced by typhoons was found. This correlation remains to be present both for the simulation with ECP with maximum strength of convection and with ERA-Interim where in contrast convection is underestimated. Most scenarios also indicate a positive correlation between the backscatter intensity of the ATAL and enhanced contributions from the Indo-Gangetic plain. Although, impacts of unresolved convection likely can be neglected for the relation to the West Pacific, the correlation for the Indo-Gangetic plain changes substantially from 0.6 to 0.1 with parameterized convection (see E5-kin-M-ECP). Moreover, for the Indo-Gangetic-plain the correlation in the scenarios E5lr-dia-M and EI-dia-C remains insignificant. The most significant results however, are obtained in the ERA5 scenarios, supporting the hypothesis that polluted air from the Indo-Gangetic plain led to higher backscatter intensity and vice versa clean maritime air from the Western Pacific lead to a dilution of the ATAL and therefore to a weaker BSR intensity.

Most scenarios furthermore have a positive correlation for contributions from the Arabian Sea, the rest of India, and the Bay of Bengal. Such relations seem plausible given the transport pathways from the Arabian Sea and the Bay of Bengal to the ATAL over Nainital, which include a period of horizontal transport in the polluted troposphere over India before a second step of upward transport in deep convection. Additionally, air masses from the Arabian Sea could carry dust from the Arabian Peninsula. Air from South-East-Asia is also weakly correlated with a decrease of the ATAL backscatter intensity, which is possibly related to the landfall of some typhoons.

Other source regions have been considered to establish such relations, but given the limited amount of data and systematic model uncertainties, no further robust results were found. The positive relation between the Asian Highlands, i.e. mostly the Tibetan Plateau, and the backscatter intensity of the ATAL, found by Hanumanthu et al. (2020) was reproduced for similar scenario set-ups (see EI-dia-C). However, some scenarios show even negative correlations. To check if this result could depend on the definition of the Asian Highlands, the correlation were additionally calculated with a more narrow definition, focused on the Tibetan Plateau, but still obtained ambiguous correlations.

The results of this study corroborate that the transport calculations presented here are capable of capturing the general evolution and patterns during the course of August robustly, but might differ if day-to-day changes are considered. The observed general temporal evolution might be related to a large scale change of the meteorological conditions during August, that led to a shift to more air masses from the Western Pacific and prior typhoons and less transport from the Indian Subcontinent.

6.3 Discussion and conclusions

In this chapter, the source regions of the ATAL in August 2016 were investigated. To identify the source regions at the MBL and the transport pathways contributing to the ATAL and to investigate the sensitivity of the applied methods, different trajectory calculations were conducted. Simulations with different model scenarios using different Lagrangian transport models (CLaMS and MPTRAC), wind data (ERA-Interim and ERA5), vertical velocities (kinematic and diabatic), integration time-steps and a convection parameterisation (ECP) were analyzed. Additionally, daily contributions of source regions at the surface were correlated to daily measured COBALD backscatter intensities at ATAL altitudes to quantify the role of the regions for the intensity of the ATAL.

Most of the air from the MBL that influenced the measurements originated at the Tibetan Plateau (i.e. $\approx$ 30%-40% of air masses originating at the MBL). This is found for all scenarios, except for the scenario with the extreme convection parameterisation. In the scenario with the ECP (E5-kin-M-ECP), the Indo-Gangetic plain is contributing most ($\approx$ 30%). The Indo-Gangetic plain is the second largest continental contributor (10 – 20 %) to the air masses influencing the measurements in all other scenarios (except for ERA-Interim with a diabatic approach). The contribution from the Indo-Gangetic plain, however, is much smaller than from the Tibetan Plateau in those scenarios. In summary, most of the upward transport takes place at the eastern part of the Indo-Gangetic plain that extends to the Bay of Bengal, including the foothills of the Himalayas and the Tibetan Plateau. These regions have been found to be dominant for the transport into the ASMA and the ATAL also by other studies (e.g. Bergman et al., 2012; Bucci et al., 2020; Hanumanthu et al., 2020).

In the simulations, a small amount of air was transported from South-East-Asia and North India, to the ATAL as well. Such transport processes contributing to the ATAL have also been reported before (e.g. Vernier et al., 2018; Bucci et al., 2020; Zhang et al., 2019a,b, 2020). Air masses from the maritime boundary layer were transported to the measurement locations in significant numbers as well. This includes mostly air masses from surrounding seas, such as the Arabian Sea, the Bay of Bengal and the Western Pacific. Typhoons in the Tropical Western Pacific played an important role for the transport process from the maritime boundary layer, which is in good agreement with previous studies, which showed their relevance for the UTLS and the composition of the ATAL (Li et al., 2017b, 2020; Hanumanthu et al., 2020).

By studying the transport pathways and times, some systematic differences between simulation scenarios that are related to the representation of convection and the diabatic ascent in the

ASMA, are shown. ERA5 has a better representation of convective updrafts and tropical cyclones compared to the ERA-Interim reanalysis, attributed to its better spatial and temporal resolution and other improvements of the ECMWF forecast model and data assimilation scheme. Therefore, the fraction of air from the MBL transported upward to ATAL altitudes is lower or about equal in scenarios with ERA-Interim in comparison to scenarios with ERA5. This is in particular true over the continent. Hence, in ERA-Interim convection over the continent is less frequent than in ERA5, so that larger fractions of air parcels originate from remote maritime regions with ERA-Interim (40% vs. 23% of all air parcels from the MBL).

ERA-Interim simulations show also strong differences with regard to the transport from the MBL into the UTLS, when the vertical velocity is varied between diabatic velocities (75% of all air parcels) and kinematic velocities (60% of all air parcels). These differences between kinematic and diabatic trajectories are strongly reduced, when ERA5 is used, where the diabatic approach shows similar fractions of air transported from the MBL to ATAL altitudes like the kinematic approach (74% vs. 80%). Large differences with regard to the vertical transport in typhoons between ERA5 and ERA-Interim have been reported before by Li et al. (2020) and an improvement of consistency between vertical velocities by Legras and Bucci (2020) in the Asian monsoon region.

Although ERA5 resolves convection better than ERA-Interim, it might still underestimate the extent of fast vertical transport caused by deep convection at the foothills of the Himalayas and at the Indo-Gangetic plain. This possible deficiency is indicated by the simulation scenario with ECP, that shows a strong increase of convection near Nainital at the Indo-Gangetic plain and the foothills. The results of this thesis show, that ERA5 provides a significant improvement for the simulation of transport processes in the Asian monsoon region with regard to the consistency between scenarios with different models and vertical velocity schemes. However, ERA5 possibly still has limitations with regard to the representation of the convection, which needs to be evaluated in further studies that take observations of convection into account. In addition, this study shows for the employed Lagrangian model (MPTRAC and CLaMS) minor differences. These differences are likely caused by differences in the integration scheme and the interpolation method. Both models are equally valuable in case of the present analysis.

Taking into account the measured backscatter intensity of the ATAL, two regions with a significant and robust impact on the ATALs variability were found. Meteorological conditions that are favourable to transport from the Indo-Gangetic plain increase the ATAL backscatter intensity, while conditions favourable to transport of air masses from the Tropical Western Pacific and the influence of typhoons decrease the ATAL backscatter intensity. In case of the Tropical Western Pacific,

these findings hold for the different sensitivity calculations carried out, and hence it corroborates the results by Hanumanthu et al. (2020), by showing that this correlation is robust despite the systematic uncertainties represented by the different simulation scenarios. In case of the Indo-Gangetic plain 10 of 13 scenarios are underpinning this finding of a positive correlation, while the remaining three scenarios show very low and insignificant correlations. To completely remove the remaining uncertainties, further observations in the region are needed.

The findings of this study are in agreement with results of previous studies. Studies found ammonium nitrate particles as a major component of the ATAL (Höpfner et al., 2019). Ammonia, the precursor of this aerosol is emitted frequently over the Indo-Gangetic plain, which is an area of active agriculture and industry (Kuttippurath et al., 2020) and could be transported fast enough into the ASMA within hours to a few weeks according to the calculations. Transport within typhoons up into the UTLS can provide clean and dry air from the ocean (Li et al., 2020) leading to a reduction of the backscatter intensity of the ATAL, as also shown in the simulations. The possible role of dust from the Asian deserts or highlands for the formation of the ATAL is discussed in the literature (e.g. Vernier et al., 2011; Bossolasco et al., 2021). The calculations of this thesis do not disprove that dust from the Tibetan Plateau could have contributed to the ATAL in August 2 016, but indicate that dust is likely not essential to understand the observed variability of the ATAL backscatter intensity during this period. Indeed, the simulations indicate a large potential for transport of dust into the ASMA from the Asian Highlands, i.e. mostly the Tibetan Plateau, but this transport more likely leads to a constant background over the observed period.

Chapter 7

Summary

In this thesis, a multi-scenario Lagrangian transport analysis with state-of-the-art Lagrangian transport models and reanalyses was performed. The analysis identified the robust knowledge and remaining uncertainties about the source regions and the transport pathways of aerosol particles and their gas-phase precursors from the Earth's surface to the Asian Tropopause Aerosol Layer (ATAL) inferred from Lagrangian transport models. The analysis was performed on the basis of 15 balloon measurements of the ATAL taken in August over Nainital (northern India), which identified the variable location of the ATAL in an altitude range of $14-18$ km.

The multi-scenario Lagrangian transport analysis uses two Lagrangian transport models, the MPTRAC and CLaMS. In this thesis, MPTRAC has been equipped with a diabatic transport scheme similar to CLaMS because it is beneficial for transport calculations in the UTLS in comparison to the former kinematic transport scheme. Furthermore, MPTRAC is optimized for HPC applications and provides parameterizations for unresolved processes such as convection. Conclusively, a Lagrangian transport model is now available that enables a multi-scenario analysis with the full ERA5 reanalysis, diabatic vertical velocities, and a novel coupling with pressure-based parameterizations (e.g. convection).

The new implementation of zeta coordinates and diabatic transport into MPTRAC has been evaluated with the help of a large, globally distributed ensemble of air parcels, starting from a configuration commonly used for UTLS studies in CLaMS. It was shown that the model difference between MPTRAC and CLaMS is on the same order of magnitudes as the combined internal uncertainties (e.g., interpolation and integration errors) and at least one order of magnitude smaller than external uncertainties such as those that originate from downsampling of the resolution, reanalysis differences, or sub-grid scale winds. Moreover, no statistical long-term bias between

the models was found. These results demonstrate that the CLaMS trajectory code can be interchanged with MPTRAC.

By applying the improved Lagrangian transport framework to a global study covering the UTLS (Chapter 5) and to the regional analysis of measurements of the ATAL (Chapter 6) the guiding questions of this thesis can now be answered as follows:

1. *Which surface sources and transport pathways characterize air masses found within the ATAL?*

   - For the ATAL above Nainital in August 2016, the analysis indicates, that most air originated from the continental regions close to Nainital, i.e. the Tibetan Plateau, the Foothills of the Himalayas and the Indo-Gangetic plain. The findings support the existence of a focused upward transport centered over the Tibetan Plateau and the Indo-Gangetic plain in agreement with concepts as the “vertical conduit” of former studies (Bergman et al., 2012).
   - The influence of the Indo-Gangetic plain is characterised by short transport times into the UTLS (less than 2 days in 50% of the cases with ERA5) and a positive correlation with the variability of the measured backscatter ratio. This is consistent with studies emphasizing local ammonia emissions, which are common at the Indo-Gangetic plain, as possible sources of the ATAL (e.g. Höpfner et al., 2019). The results in general suggest anthropogenic emissions on the Indian subcontinent as a source of the ATAL over Nainital in 2016.
   - The variability of the ATAL backscatter ratio is not correlated to temporal transport changes from the Tibetan Plateau during the measurement campaign, in contrast to the study of Hanumanthu et al. (2020) who focused on one reanalysis. However, the analysis indicates a possible large and continuous contribution from the Tibetan Plateau to the ATAL during the analysed period.
   - Largest maritime contributions to the ATAL came from the Western Tropical Pacific. Transport from the Western Tropical Pacific, closely related to tropical storms, were correlated with decreased back-scatter values of the ATAL, corroborating results of Hanumanthu et al. (2020). These pathways possibly transport clean air masses into the ASMA and dissolve the ATAL.

2. *How robust is our understanding of surface sources and transport pathways of ATAL inferred from different Lagrangian transport calculations considering uncertainties in upward*

transport?

- The general concept of a vertical conduit placed at the Eastern foothills of the Himalayas as being mostly responsible for the transport from the surface to the ATAL is very robust across all scenarios. However, with changing resolution and representation of convection the importance of specific regions changes. ERA5 in combination with the ECP simulated most convection over the continents, ERA5 the second most and ERA-Interim the least. Following this order, more air parcels originate from the local continent as more convection is resolved. Hence, with ERA-Interim more air masses originate from the oceans. With the ECP a large local contribution from the Indo-Gangetic plain was found as well as a small additional conduit in China.
- Despite the large uncertainties with regard to convection, the relation between influence from the Tropical Western Pacific and a reduced ATAL is very robust.
- Moderate robustness and significance has been found for the positive correlation between the Indo-Gangetic plain and the ATALs backscatter ratio. The analysis exposes uncertainties about the role of local deep convection, that likely is unresolved even in ERA5, at the Indo-Gangetic plain, which needs further clarification. The role of the Tibetan Plateau remains more uncertain. Although large contributions are robustly found in all scenarios, correlation were ambiguous.

3. *What are the transport uncertainties in Lagrangian transport calculations of the upper troposphere and the stratosphere with the latest generations of reanalyses (ERA5 and ERA-Interim), vertical transport schemes (kinematic and diabatic) and models (MPTRAC and CLaMS)?*

   - Model differences and internal uncertainties of the models were found negligible in comparison to other uncertainty sources, in the global evaluation study as well as for the application on regional scale. The accuracy of commonly applied linear interpolations (Bowman et al., 2013) and applied integration schemes, such as the mid-point and Runge-Kutta scheme, are still sufficient for the analysis.
   - Global trajectories simulations over 90 days reveal substantial biases and deviations between ERA5 and ERA-Interim in the UTLS using diabatic transport that are larger than biases, and as large as deviations from unresolved sub-grid scale diffusion. Transport within the BDC is faster with ERA-Interim than with ERA5 between 400 K and 600 K, but slower for higher levels, on the seasonal scale in 2016. This is in

agreement with recent climatological and regional studies (Ploeger et al., 2021; Vogel et al., 2023). The regional study of the transport process from the surface to the ATAL demonstrates large deviations between ERA5 and ERA-Interim as well, especially for convective transport over the Asian continent. ERA5 resolves more of the convective transport than ERA-Interim. This is an effect that has been validated in other studies as well (e.g. Hoffmann et al., 2019, 2023b; Vogel et al., 2023). However, transport calculations with the ECP indicate that even calculations with ERA5 likely miss important convective events.

- Significant differences and biases between kinematic and diabatic calculations have also been found in ERA5 by studying the global 90 days ensemble. Moreover, cross-isentropic dispersion in the tropics and subtropics with the ERA5 reanalysis is larger with the kinematic than with the diabatic scheme. Differences between kinematic and diabatic calculations are known from ERA-Interim (Ploeger et al., 2011). However, the deviations between diabatic and kinematic calculations as well as the cross-isentropic dispersion are significantly reduced in ERA5 in comparison to ERA-Interim. The improved agreement between kinematic and diabatic velocities, and the reduction of the cross-isentropic dispersion in ERA5, have substantial impacts on the regional backward trajectory analysis, where the scenarios with ERA5 show much more similar vertical transport than with ERA-Interim. Better agreement between kinematic and diabatic calculations with ERA5 than with ERA-Interim was as also reported by Legras and Bucci (2020) and Bucci et al. (2020).
- In the stratosphere, calculations with downsampled ERA5 data capture well the average distribution of the globally distributed air parcels after 90 days of forward calculation in comparison to the full resolution ERA5 data. However, vertical up- and downward transport is possibly reduced in the troposphere. In agreement with the findings on global scale, the application of the downsampled data in the regional analysis maintained many average transport features of the full ERA5 data, but reduces the simulated transport from the model boundary layer into the UTLS slightly.

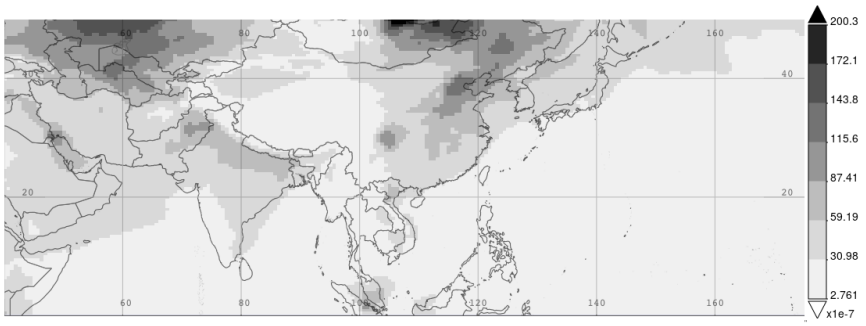
The guiding question let to a mixed picture that corroborates on the one side former overarching results based on ERA-Interim (such as the existing of a dominant vertical conduit) and that reveals clear improvements in robustness of transport calculations in the UTLS (e.g. between diabatic and kinematic calculations) and the resolution of convection with ERA5, but that on the other side also exposes some remaining limitations of existing trajectory analyses with regard

to convection and UTLS transport. Further model development will be required to represent the vertical transport with higher accuracy, by seeking even higher resolution with the help of coming exascale supercomputers and by developing improved parameterizations of the convection. Furthermore, this analysis was focused on Nainital in 2016. Since the ATAL is showing a distinct spatio-temporal variability, further measurements are required to fully sample all relevant aspects of the ATAL.

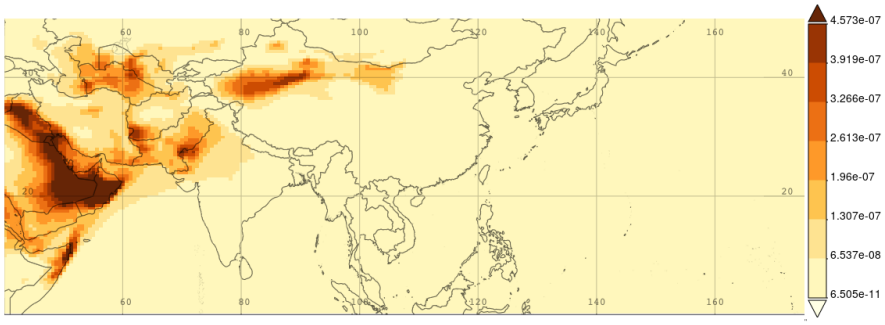
Conclusively, this thesis identified the main pathways and source regions at the Earth's surface contributing to ATAL, taking into account different internal and external uncertainties of the CLaMS and the MPTRAC models. Since the Asian monsoon region is one of the most populated and polluted regions of the world, it is crucial to assess the robustness and uncertainties of Lagrangian transport model simulation in this region. This is an essential prerequisite for predicting the impact of the Asian monsoon system on the global atmosphere under a changing climate.

Appendix A

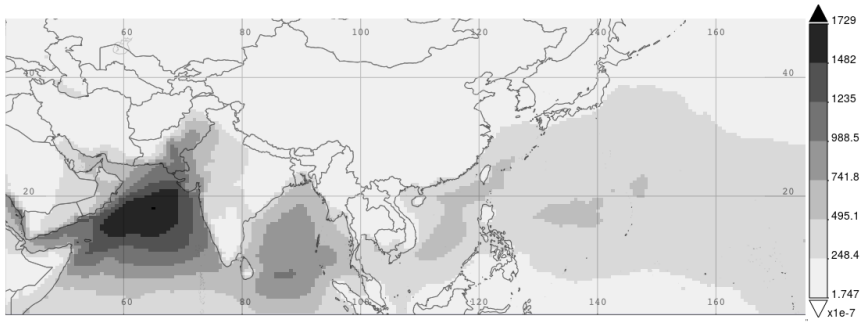
The Asian tropopause aerosol layer



(a) organic carbon



(b) dust



(c) sea salt

Figure A.1: Average surface conditions considering organic carbon (i.e. carbon of biological origin), dust and sea salt for the summer 2016 *over Asia derived from the MERRA-2 reanalysis: (a) column mass density of organic carbon in* $\mathrm{kg \cdot m^{-2}}$*, (b) dust surface mass concentration in* $\mathrm{kg \cdot m^{-3}}$ *and (c) column mass density of sea salt in* $\mathrm{kg \cdot m^{-2}}$*. Produced with the Giovanni online data system, developed and maintained by the NASA GES DISC (Acker and Leptoukh, 2007; Global Modeling and Assimilation Office, 2015a).*

Appendix B

Performance optimization of Lagrangian transport calculations

Lagrangian transport models have to read meteorological and initial data, calculate the advection, apply different additional geophysical or chemical modules and subsequently write the data. Therefore, the performance bottlenecks of Lagrangian transport calculations with the ERA5 reanalysis depend on the focus of the application. Fig. B.1 shows the calculations time of the transport calculations depending on the number of air parcels using MPTRAC and ERA5 to illustrate the dependency of the bottlenecks on the number of air parcels. The calculations have been done on a large memory node of the compute cluster of the Jülich Wizard for European Leadership Science (JUWELS) supercomputer (JSC, 2022). On the one side, ERA5 increases the memory requirements by nearly two orders of magnitude in comparison to ERA-Interim because of the higher resolution. Therefore, reading the data has the highest computational cost for a number of air parcels up to around 1.5 million air parcels (see Fig. B.1). On the other side, to fully exploit the grid information in ERA5 with the Lagrangian transport model during a global transport run, a Lagrangian grid size with the same amount of air parcels is needed as the size of the ERA5 grid provides (i.e. around 100 million air parcels). This leads to an increase in computational demand in many orders of magnitude as well for the calculations of the advection and geophysical properties (see also Fig. B.1 for 10 million air parcels). It is evident that for regional trajectory studies with less than 10 million air parcels the most important bottleneck is the IO, while for large-scale global runs, which are relevant for full-blown chemistry-transport calculations at full resolution, the efficiency of the advection kernel is the increasingly dominant bottleneck.

The target of this thesis is to use regional trajectory studies with ERA5 with air parcel numbers not exceeding 2.0 million air parcels. Hence options to improve the IO are explored here, which address some of the challenges related to the ERA5 reanalysis. Among these options are data access tuning, compression, chunking, shared memory options, parallel IO and improved workflows. Improvements in the usage of ERA5 were implemented for MPTRAC as wells as for CLaMS.

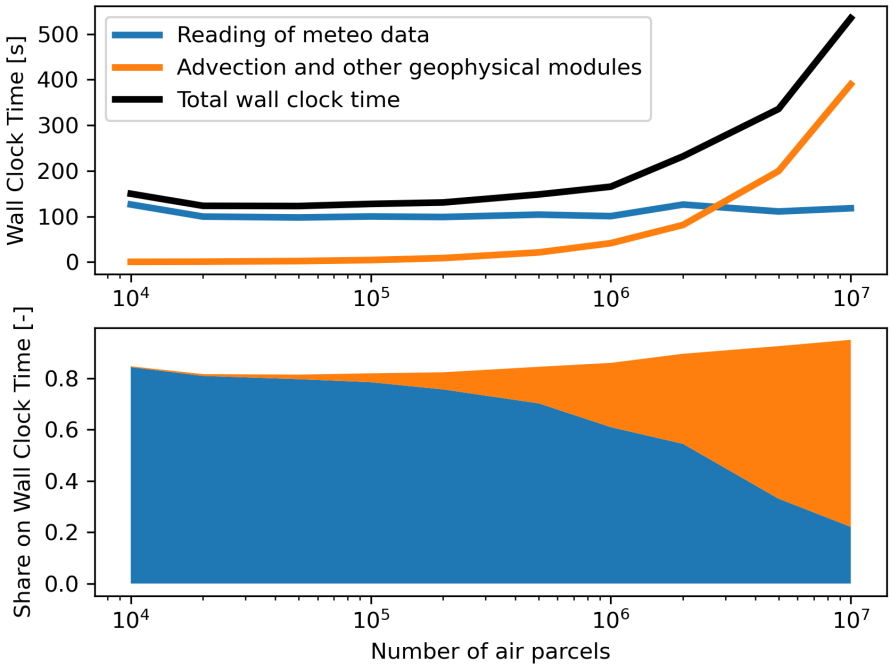


Figure B.1: Illustration of the performance of MPTRAC in relation to the number of air parcels on a single large memory node with ERA5. The air parcels are calculated forward 24 hours. Shown are the wall clock time for reading the data, and for the calculation of geophysical modules (e.g. parameterizations) and the advection. The remaining time is related to writing of data and processing of meteorological data.

B.1 Performance analysis of netcdf format, compression and chunking options

The reanalysis data is stored in NetCDF files. NetCDF files can have multiple different formats, a classic format with 64-bit offset (nc2), the NetCDF-4 formats (nc4 and nc4c) and hdf5 formats (Unidata, 2023). Furthermore they provide options for chunking and compression of data. The performance of different formats has been intercompared here based on one-day forward calculations of a large, global trajectory ensemble of 24 hours, to find out the most performant format options for MPTRAC on the Jülich Wizard for European Leadership Sciences (JUWELSs) large memory node. Moreover, the chunking of the data has impacts on the performance of the reading process as well as the compression.

Figure B.2 shows the time required to read the variables from the NetCDF files in MPTRAC for a different format, chunking and compression options. First, there are large performance differences between the formats, with the NetCDF-4 formats allowing up to 50 % faster calculations than the hdf5 and nc2 formats. Second, unfavourable chunking can lead to strong performance loss, for example, if the zy-planes (i.e. the altitude-latitude planes) serve as chunks, while favourable chunking can have small performance benefits (around 10 %), as found for example if the zx-planes (i.e. altitude-longitude) serve as chunks. Third, lossless deflation of the data is frequently used to reduce the required storage on the hard disk. A reduction from 3.2 GB per file to 2.0 GB per file was achieved with deflation level 2. However, the calculations have substantial performance loss with the compressed data by a factor of almost 3 for the MPTRAC model.

For reading the data from the disk, a chunk cache is used. The optimal size for the chunk cache is at least as large as the disk block size of the underlying system to enable a complete read into the cache with each physical reading process. For the supercomputer, the block size is around 16 MB. By varying the chunk size cache from 4 MB to 16 MB with the nc2 file format it was validated that the data is read with the largest chunk sizes possible, i.e. with the block size. This was also validated with profiling tools.

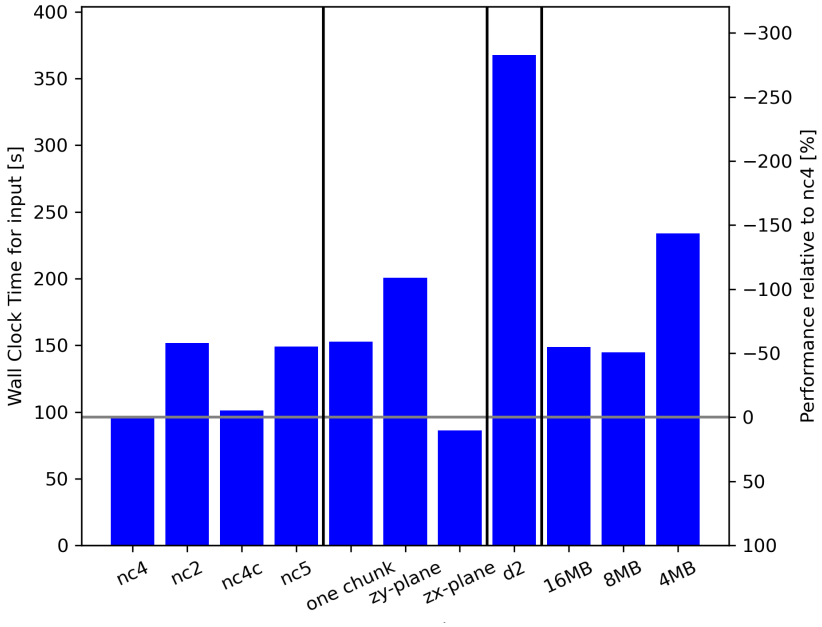


Figure B.2: Performance of different netcdf formats (nc4 to nc5), chunking (one chunk to zx-plane), compressed files (d2) and dependence on the chunk cache (16 MB to 4 MB).

B.2 Optimization experiments with parallel IO and the MPI-3 shared memory model in CLaMS

To handle increased memory demand and IO throughput the usage of a shared memory model between the cores of a node, as well as parallel IO from many cores of the node, are explored. MPTRAC already uses a shared memory as part of the Open Multi-Processing (OpenMP) multi-threading. However, parallel IO with MPTRAC and NetCDF is not available, because the netcdf library is not thread-safe, and parallel IO with OpenMP is not available without major efforts (e.g. Mehta and Gabriel, 2015). Within the CLaMS model, which uses only MPI for parallelisation, shared memory and parallel IO haven't been tested so far. Here, the MPI-3 standard is used to implement a shared memory model to CLaMS and to investigate parallel IO (e.g. Intel, 2018). Hence, the meteorological data is read by many cores at the same time into a shared memory which is available to all cores.

Figure B.3 shows a performance benchmark of the CLaMS trajectory module on one large memory node of JUWELS. The scaling with increasing number of cores is shown. Each node has 48 physical cores and 198 GB of memory. The benchmark uses 1 million air parcels, which are globally distributed, and calculated forward for one day (1 January 2 017).

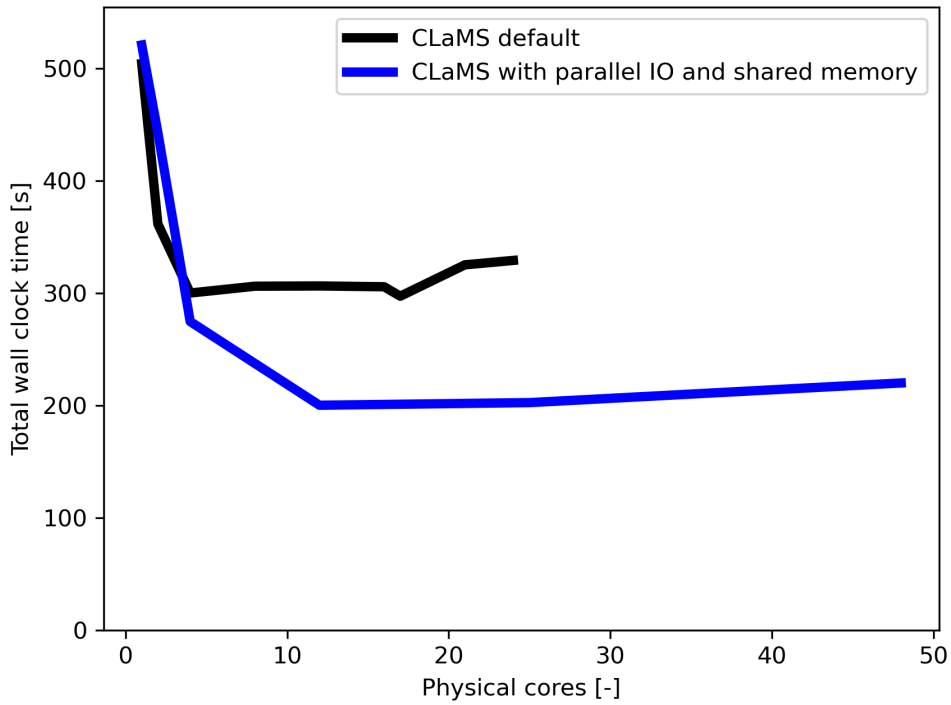


Figure B.3: Performance scaling of CLaMS for the default version (black line) and a tuned version (blue line) including parallel IO and the MPI-3 shared memory model on a large memory node on JUWELS.

The performance is only scaling significantly for the first 4 cores with the default CLaMS version. Afterwards, the performance remains almost constant, because the IO is not scaling with the number of cores and has the largest share of the total wall clock time. Furthermore, the default parallelism of CLaMS reduces the largest possible number of cores available, because with every new task on a new core, new data is read into memory. With ERA5 the data is increasing with around 8 GB per new node and task, limiting the maximum cores available to around 24 on the large memory model. With the shared memory model and parallel IO implemented, the total wall clock time is reduced by around 30 % and the entire node can be used, i.e. CLaMS can run on 48 cores. While the IO performance gain potentially can facilitate calculations with air parcel numbers below 2 million air parcels, the availability of more cores, might enable higher scaling of the calculations with CLaMS with much larger air parcel numbers. Additionally, a reduced memory demand of the CLaMS trajectory code, can potentially ease constraints on concurrent runs with other models and modules.

B.3 Conclusions

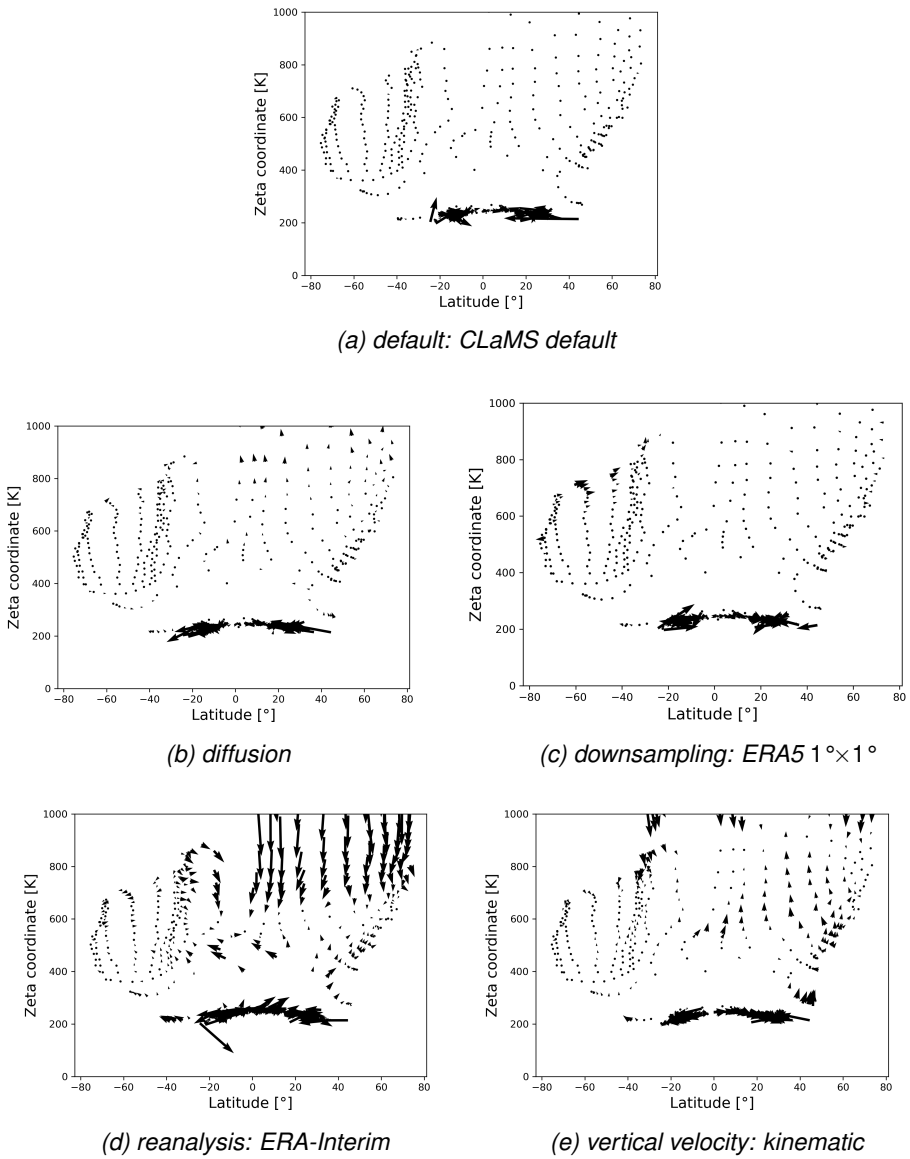
The IO of the larger ERA5 reanalysis data has been validated and quantified as the leading computational costs for Lagrangian transport calculations with up to 2 million air parcels for the MPTRAC model on a CPU computation node of JUWELS. Afterwards, geophysical modules and advection have larger computational costs. When choosing between NetCDF files for MPTRAC the analysis advises selecting the NetCDF-4 format, because of up to 50 % performance differences to the other formats. Furthermore, using compressed files in MPTRAC can lead to large performance loss. Using a different chunk layout did not show high-performance gains on JUWELS (around 10 %). The optimization experiments of CLaMS showed large benefits for the implementation of the shared memory model in CLaMS. MPTRAC already possess these benefits, which corroborates the efforts to make MPTRAC inter-operable with the CLaMS framework. Additionally, using parallel IO in CLaMS showcases significant acceleration of the calculations with up to 30 % performance gains. Hence, the future implementation of the optimizations found in the presented experiments into CLaMS production code could lead to significant improvements. Preliminary intercomparison of the optimized model results with the default results also have shown only minor differences.

Appendix C

Evaluation of diabatic transport in MPTRAC v2.6

C.1 Circulation in the UTLS

To clarify the circulation pattern during the 90 days integration period box-wise averaged trajectories have been calculated. Within a longitude-zeta grid (with around 8° × 32 K grid-boxes) based on the initial positions of the air parcels, air parcels have been sorted into bins. The average positions of this binned air parcels, defines the box-wise average trajectory. Afterwards the vector, pointing from the end point of the average trajectory for the scenario “MPTRAC-default” to the endpoints of trajectories of the compared scenario, is calculated. See in Fig. C.1.



(a) default: CLaMS default

(b) diffusion

(c) downsampling: ERA5 1°×1°

(d) reanalysis: ERA-Interim

(e) vertical velocity: kinematic

Figure C.1: Difference vector between the end point of box-wise averaged trajectories after 90 days forward calculations to indicate the transport bias between scenarios. The subfigures (a) to (e) show the difference vector always in relation to the scenario “MPTRAC-default”.

Appendix D

A multi-scenario Lagrangian trajectory analysis to identify source regions of the Asian tropopause Aerosol Layer on the Indian subcontinent in August 2016

D.1 Definition of source regions

To construct the surface source regions, first the geopotential is considered to define the Asian Highlands as the region in Asia with a geopotential larger than 15 000 $m^2 \cdot s^{-2}$ m^2 s^{-2} (corresponding to a geopotential height of approximately 1.5 km) similar to Hanumanthu et al. (2020). The Tibetan plateau is defined by a geopotential larger than 40 000 m^2 s^{-2} (around 4 km). Secondly, a land-sea mask is used to distinguish between oceans and continents. The continental regions are defined by the boxes found in Table D.1. The boxes defined in Table D.2 define maritime regions, where only those regions are included that are part of the sea according to the land-sea mask.

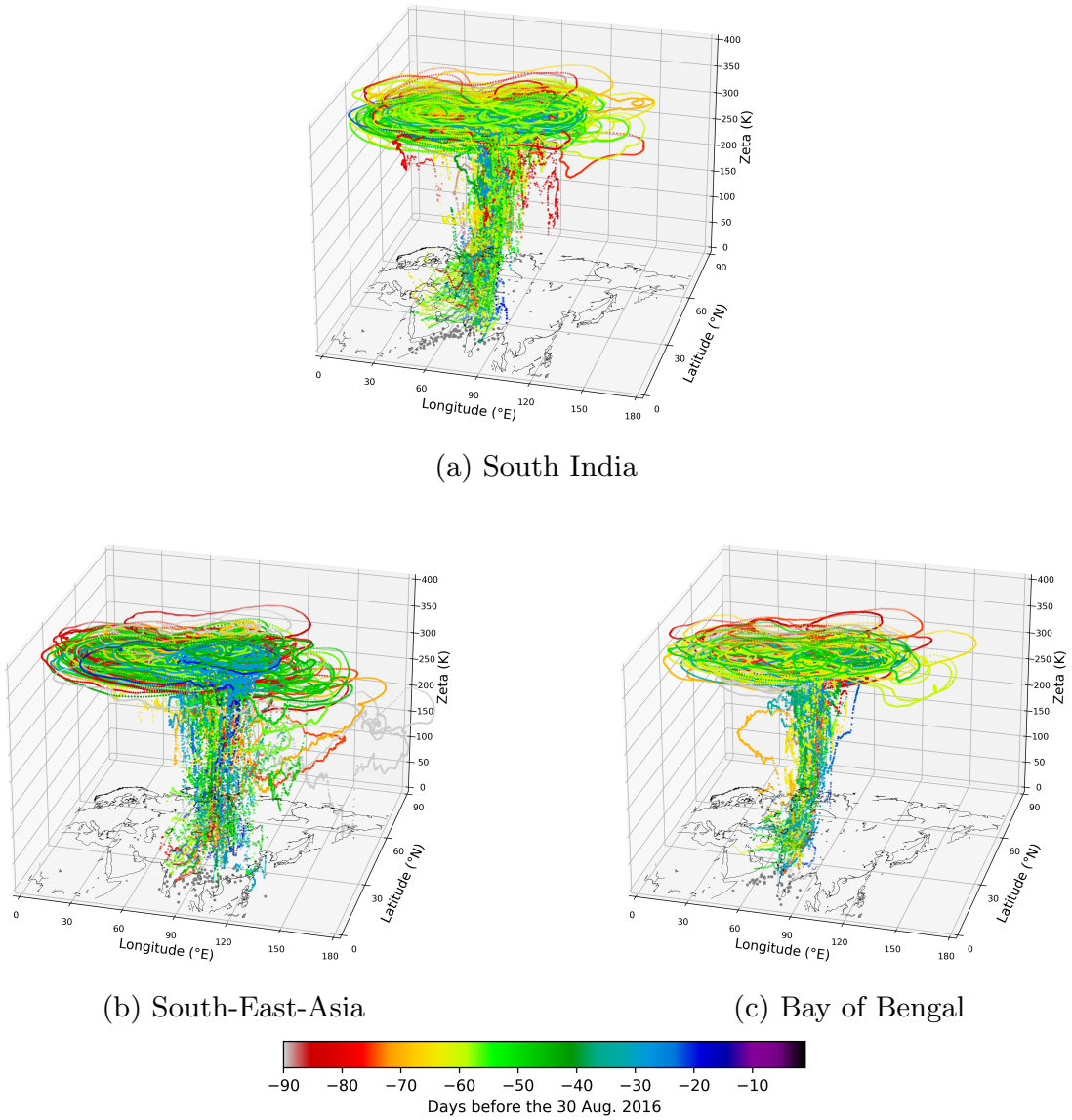
Table D.1: The continental source regions are defined by the overlap of longitudinal and latitudinal restricted boxes and the continent without the Asian Highlands. The Asian Highlands are defined by a GPH criteria.

Name	Label	Minimum Longitude	Maximum Longitude	Minimum Latitude	Maximum Latitude
Asian Highlands	AH	40° E	110° E	20° N	90° N
South India	SI	65° E	90° E	5° N	23.5° N
Indo-Gangetic plain	IGP	67.5° E	90° E	23.5° N	35° N
South-East-Asia	SEA	100° E	160° E	7° N	30° N

Table D.2: The maritime source regions are defined by the overlap of longitudinal and latitudinal restricted boxes with the seas.

Name	Label	Minimum Longitude	Maximum Longitude	Minimum Latitude	Maximum Latitude
Arabian Sea	AS	50° E	80° E	7° N	27° N
Bay of Bengal	BOB	80° E	100° E	7° N	27° N
Tropical Western Pacific	TWP	100° E	160° E	7° N	30° N

D.2 Transport pathways from source regions

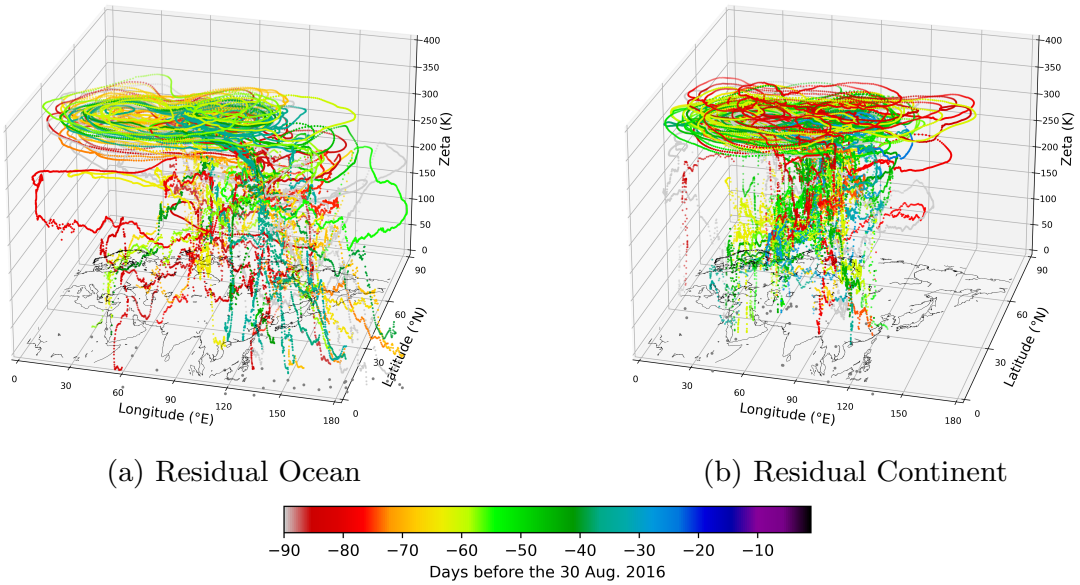


(a) South India

(b) South-East-Asia

(c) Bay of Bengal

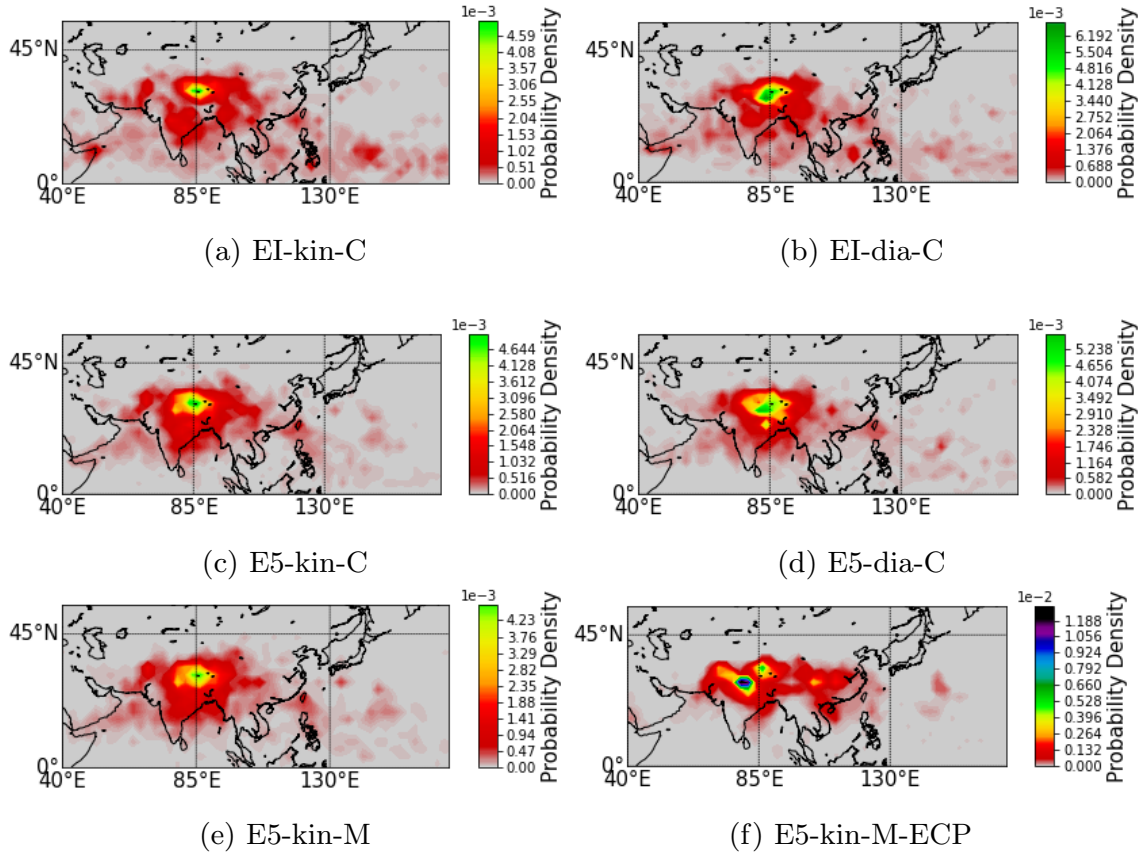
Figure D.1: Exemplary backward trajectories of air parcels of all of the 15 measurement days from the ATAL measurement to the MBL, categorised by the source region. Colors are chosen as in Fig. 6.2.



(a) Residual Ocean (b) Residual Continent

Figure D.2: Exemplary backward trajectories of air parcels of all of the 15 measurement days from the ATAL measurement to the MBL, categorised by the source region. Colors indicate the time when the air parcels leaves the MBL. Gray dots at the bottom show the horizontal position of the air parcels 48 h before they cross the MBL from below.

D.3 PDF for different scenarios



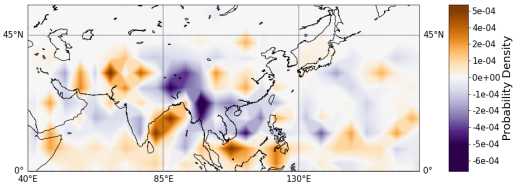
(a) EI-kin-C (b) EI-dia-C

(c) E5-kin-C (d) E5-dia-C

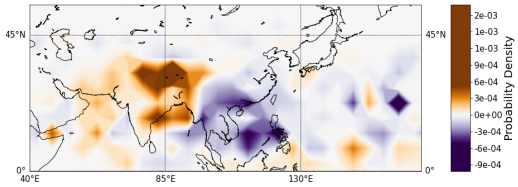
(e) E5-kin-M (f) E5-kin-M-ECP

Figure D.3: PDFs for the MBL source region for some scenarios.

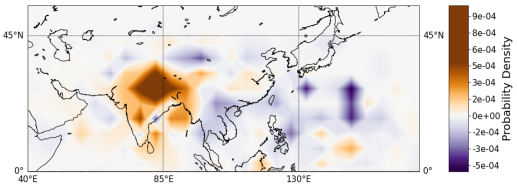
D.4 PDF differences for different scenarios



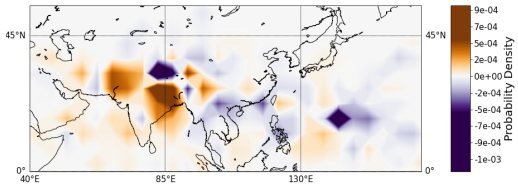
(a) EI-kin-C



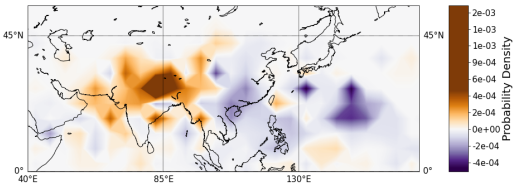
(b) EI-dia-C



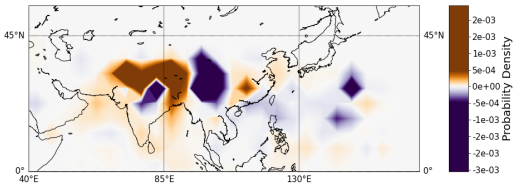
(c) E5-kin-C



(d) E5-dia-C



(e) E5-kin-M



(f) E5-kin-M-ECP

Figure D.4: PDF differences for days with high and low ATAL backscatter for specific scenarios, similar to Fig. 6.8a, illustrating the sensitivity of the PDF differences to the chosen scenario.

D.5 Correlation to all regions

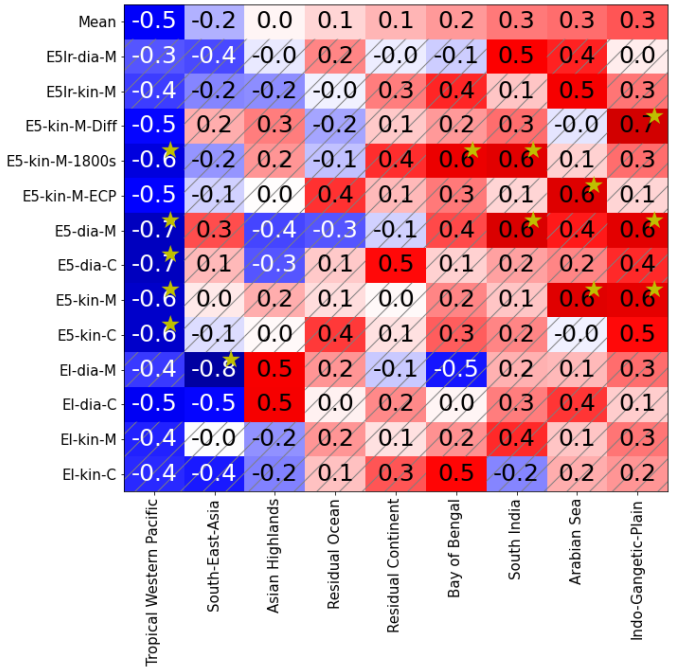


Figure D.5: All Spearman correlation coefficients for the relation between the daily BSR and the contributions of different regions for different scenarios. Crossed out areas hatch insignificant results (p-value higher 0.10). Boxes with yellow stars indicate p-values lower than 0.05. The regions are ordered according to the scenario mean from lowest to highest correlations.

Table D.3: Mean transport time from the MBL into the UTLS in days, determined by the difference between the leaving time at the MBL and the arrival time above 340 K. Abbreviations are as defined in Table D.1 and D.2.

Scenario	All	Res. Oc.	Res. Cont.	SEA	AH	IGP	TWP	AS	BOB	SI
EI-kin-C	7.2	13.8	9.1	5.9	2.2	6.2	7.0	9.8	5.6	6.0
EI-kin-M	7.2	13.0	12.7	5.4	2.0	6.1	7.0	8.3	5.9	5.3
EI-dia-C	6.4	14.3	14.4	4.2	1.1	5.9	8.4	8.8	6.5	3.0
EI-dia-M	6.2	14.5	13.8	4.1	1.1	6.4	8.3	9.6	6.8	3.3
E5-kin-C	3.3	8.3	8.6	2.9	1.2	3.9	3.2	6.3	3.9	2.2
E5-kin-M	3.4	9.3	9.4	2.6	1.3	4.3	3.0	6.2	3.7	2.4
E5-dia-C	3.4	10.7	10.4	2.5	0.9	4.3	3.7	6.6	3.9	2.2
E5-dia-M	3.3	9.8	10.5	2.8	0.8	4.3	4.4	6.3	3.9	2.2
E5-kin-M-ECP	0.1	0.2	0.4	0.1	0.1	0.1	0.1	0.2	0.1	0.1
E5-kin-M-1800s	3.3	9.0	9.7	2.5	1.3	4.2	3.0	5.4	3.9	2.5
E5-kin-M-Diff	3.3	9.0	9.4	2.9	1.3	4.1	3.1	6.5	3.8	2.2
E5lr-kin-M	4.6	11.5	12.7	4.5	1.6	4.3	4.1	6.6	4.2	3.5
E5lr-dia-M	3.6	13.0	14.1	3.0	1.0	4.0	3.9	5.9	4.0	2.3

Table D.4: Median transport time from the MBL into the UTLS in days, determined by the difference between the leaving time at the MBL and the arrival time above 340 K. Abbreviations are as defined in Table D.1 and D.2.

Scenario	All	Res. Oc.	Res. Cont.	SEA	AH	IGP	TWP	AS	BOB	SI
EI-kin-C	4.7	10.0	7.3	3.3	0.3	5.2	5.2	7.7	4.9	5.2
EI-kin-M	4.9	10.0	10.1	2.9	0.4	5.2	5.1	6.9	5.2	4.7
EI-dia-C	3.6	11.2	11.6	2.7	0.3	5.3	5.4	8.2	5.5	1.7
EI-dia-M	3.3	11.3	11.3	3.0	0.3	5.5	5.2	8.3	5.4	1.8
E5-kin-C	1.1	5.0	7.2	1.5	0.4	3.5	1.5	5.3	3.0	1.2
E5-kin-M	1.2	5.5	8.2	1.5	0.4	3.6	1.6	5.6	3.2	1.2
E5-dia-C	1.0	7.9	8.3	1.5	0.3	3.6	1.7	5.5	2.9	1.5
E5-dia-M	1.0	6.8	6.9	1.6	0.3	3.8	1.9	5.5	3.2	1.5
E5-kin-M-ECP	0.0	0.1	0.1	0.0	0.0	0.0	0.0	0.1	0.0	0.0
E5-kin-M-1800s	1.2	5.2	8.2	1.5	0.4	3.6	1.4	4.6	2.9	1.6
E5-kin-M-Diff	1.1	5.2	8.0	1.5	0.4	3.7	1.3	5.8	3.0	1.0
E5lr-kin-M	2.5	7.7	9.8	3.1	0.7	3.7	2.2	5.9	3.4	2.7
E5lr-dia-M	1.4	10.0	10.8	2.4	0.5	3.5	2.0	5.0	3.6	1.5

D.6 Sensitivity tests for simulations employing the ECP

The default setting of the ECP in MPTRAC relies on a CAPE value of $0\,\mathrm{J\,kg^{-1}}$ as a threshold for triggering convection events. For this thesis, the parameterisation can be improved to avoid spurious parameterized convection events over the Persian gulf and the Red Sea. In these regions, extremely high convective inhibition, i.e. very stable low-level layers, prevent the release of the CAPE. Therefore, an additional threshold for CIN threshold was used, which was set to $50\,\mathrm{J\,kg^{-1}}$ to remove unrealistic parametrized convection events over the Persian gulf. Figure D.6 illustrates the impact of the parameter choices of the convection parameterisation on the source identification for different threshold settings. E5-CAPE0-CIN50 is the same scenario as simulation E5-kin-M-ECP in other parts of the paper. Table D.5 gives a detailed overview about the different set-ups of the convection parameterisation.

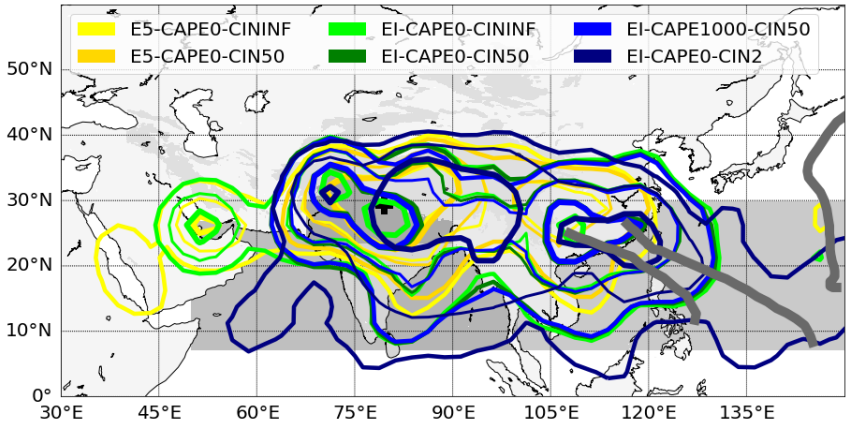


Figure D.6: The same as in Fig. 6.5, but for simulations with varying settings of the ECP parameters CIN and CAPE. E5-CAPE0-CIN50 is the same as simulation E5-kin-M-ECP.

Table D.5: Overview over scenarios, with different set-ups of the convection parameterisation, of 90 days backward calculations performed for the ATAL measurements above Nainital in August 2016. The abbreviation for each scenario contains at the first position the reanalysis, at the second position the CAPE threshold and at the last position the CIN threshold.

abbreviation	reanalysis	vertical velocity	model	time-step	CAPE	CIN	diffusion
E5-CAPE0-CININF	ERA5	kinematic	MPTRAC	1800s	0 Jkg^{-1}	-	off
E5-CAPE0-CIN50	ERA5	kinematic	MPTRAC	1800s	0 Jkg^{-1}	50 Jkg^{-1}	off
EI-CAPE0-CININF	ERA-Interim	kinematic	MPTRAC	1800s	0 Jkg^{-1}	-	off
EI-CAPE0-CIN50	ERA-Interim	kinematic	MPTRAC	1800s	0 Jkg^{-1}	50 Jkg^{-1}	off
EI-CAPE1000-CIN50	ERA-Interim	kinematic	MPTRAC	1800s	1000 Jkg^{-1}	50 Jkg^{-1}	off
EI-CAPE0-CIN2	ERA-Interim	kinematic	MPTRAC	1800s	0 Jkg^{-1}	2 Jkg^{-1}	off

List of Figures

List of Tables

List of abbreviations

ACCLIP Asian Summer Monsoon Chemical and Climate Impact Project

AH Asian Highlands

AHTD absolute horizontal transport deviation

AI Artificial Intelligence

AOD aerosol optical depths

AS Arabian Sea

ASMA Asian summer monsoon anticyclone

ATAL Asian tropopause aerosol layer

AVTD absolute vertical transport deviation

BATAL Balloon Measurement Campaigns of the Asian Tropopause Aerosol Layer

BDC Brewer-Dobson circulation

BOB Bay of Bengal

BSR backscatter ratio

CALIOP Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation

CAPE convectively availiable potential energy

CFC Chlorofluorocarbon

CI color index

CIN convective inhibition

CLaMS Chemical Lagrangian transport Model of the Stratosphere

COBALD Compact Optical Backscatter Aerosol Detector

CPU Central Processing Unit

ECMWF European Centre for Medium-Range Weather Forecasts

ECP Extreme Convection Parameterization

ENSO El Niño-Southern Oscillation

GPU Graphic Processing Unit

HALO High Altitude and long Range Aircraft

HPC High Performance Computing

IFS Integrated Forecast Model

IGP Indo-Gangetic plain

IO Input/Output

IPCC Intergovernmental Panel on Climate Change

JJA Boreal Summer, i.e. June, July, August

JUWELS Jülich Wizard for European Leadership Science

LS lower stratosphere

MARS Meteorological Archival and Retrieval System

MBL model boundary layer

MLS Aura Microwave Limb Sounder

MPI Message Passing Interface

MPTRAC Massive Parallel Trajectory Calculations

NASA National Aeronautics and Space Administration

NMVOC non-methan volatile organic compounds

OpenACC Open Accelerators

OpenMP Open Multi-Processing

PDF probability density function

PHILEAS Probing High Latitude Export of air from the Asian Summer Monsoon

PV potential vorticity

RTCE relative tracer conservation error

SEA South-East-Asia

SI South India

SR aerosol scattering ratio

StratoClim Stratospheric and upper tropospheric processes for better Climate predictions

T troposphere

TWP Tropical Western Pacific

US upper stratosphere

UTLS upper troposphere and lower stratosphere

List of symbols

symbol	description	unit
ζ	hybrid zeta coordinate	K
$\dot{\zeta}$	vertical velocity in hybrid zeta coordinates	$K \cdot d^{-1}$
θ	potential temperature	K
$\dot{\theta}$	vertical velocity in isentropic coordinates	K
p	pressure	hPa
ω	pressure tendency	$Pa \cdot s^{-1}$
p_0	surface pressure	hPa
σ	sigma coordinate	-
$\dot{\sigma}$	vertical velocity in sigma coordinate	s^{-1}
σ_r	reference sigma level for zeta coordinate	-
ϕ	latitude	°
λ	longitude	°
Ω_ζ	vertical interpolation weights	-
Ω_ϕ	latitudinal interpolation weights	-
Ω_λ	longitudinal interpolation weights	-
Q_{ijkl}, Q_{ijk}, Q	quantity for interpolation (4D, 3D)	
R_e	Earth radius	km, m
h	height above ground	km, m
Z	log-pressure height	km, m
$\vec{x}(t)$	3D position of the air parcel	°, K,m or hPa
$\vec{x}_0(t)$	initial 3D position of the air parcel	°, K,m or hPa
$\vec{u}(x,t)$	3D velocity of the air parcel	$m \cdot s^{-1}$;$K \cdot d^{-1}$, $Pa \cdot s^{-1}$
x_i	component of the air parcel position vector	°; K,km or hPa
u_i	component of the velocity field vector	$m \cdot s^{-1}$;$K \cdot d^{-1}$, $Pa \cdot s^{-1}$
x, y, z	Cartesian components of the position vector	km
Δt	integration time step of the Lagrangian transport model	s
f	downscaling factor for sub-grid scale wind parameterization	-
u_i'	component of the sub-grid scale velocity field vector	$m \cdot s^{-1}$
Δx, Δy	Cartesian coordinate distances	km, m
$\Delta\phi$, $\Delta\lambda$	spherical coordinate distances	°
H	scale height	km
D_i	diffusivity component	$m^2 \cdot s^{-1}$
ξ_i	random number for sub-grid scale wind parameterization	-
Δt_{met}	time step of the meteorological data	s
$\zeta_{ijkl}, \zeta_{ijk}$	four and three dimensional data of the zeta coordinate	K
a_i	component of the acceleration vector in the Langevin equation	$m \cdot s^{-2}$
b_{ij}	component of the matrix in the Langevin equation	$m \cdot s^{-1}$
$A_{k+\frac{1}{2}}$	parameter of the hybrid eta coordinate	-
$B_{k+\frac{1}{2}}$	parameter of the hybrid eta coordinate	-
$B_{k+\frac{1}{2}}$	parameter of the hybrid eta coordinate	-
κ	gas constant to heat capacity ratio at constant pressure	-
T	temperature	K
α_i	weights of the integration methods	-
$\vec{k}_i$	slope at the nodes of the Runge-Kutta solver	$m \cdot s^{-1}$,$K \cdot d^{-1}$, $Pa \cdot s^{-1}$
RND	relative normalized deviation	%
$C(t)$	ratio between the contribution from one region and all regions	-
$\langle\rangle$	average over air parcel ensemble	
$\langle\delta\theta^2\rangle$	variance of potential temperature over air parcel ensemble	K
$\langle\delta\zeta^2\rangle$	variance of the zeta coordinate over air parcel ensemble	K

Bibliography

Acker, J. and Leptoukh, G.: Online analysis enhances use of NASA Earth Science Data, Eos, Transactions American Geophysical Union, 88, doi: 10.1029/2007EO020003, 2007.

Angevine, W. M., Brioude, J., McKeen, S., and Holloway, J. S.: Uncertainty in Lagrangian pollutant transport simulations due to meteorological uncertainty from a mesoscale WRF ensemble, Geoscientific Model Development, 7, 2817–2829, doi: 10.5194/gmd-7-2817-2014, 2014.

Appel, O., Köllner, F., Dragoneas, A., Hünig, A., Molleker, S., Schlager, H., Mahnke, C., Weigel, R., Port, M., Schulz, C., Drewnick, F., Vogel, B., Stroh, F., and Borrmann, S.: Chemical analysis of the Asian tropopause aerosol layer (ATAL) with emphasis on secondary aerosol particles using aircraft-based in situ aerosol mass spectrometry, Atmospheric Chemistry and Physics, 22, 13 607–13 630, doi: 10.5194/acp-22-13607-2022, 2022.

Arias, P., Bellouin, N., Coppola, E., Jones, R., Krinner, G., Marotzke, J., Naik, V., Palmer, M., Plattner, G.-K., Rogelj, J., Rojas, M., Sillmann, J., Storelvmo, T., Thorne, P., Trewin, B., Rao, K., Adhikary, B., Allan, R., Armour, K., and Zickfeld, K.: IPCC AR6 WGI Technical Summary, pp. 33–144, doi: 10.1017/9781009157896.002, 2021.

Arthur, R.: Valeriepieris Circles for Spatial Data Analysis, arXiv preprint arXiv:2307.16728, 2023.

Bauer, P., Dueben, P. D., Hoefler, T., Quintino, T., Schulthess, T. C., and Wedi, N. P.: The digital revolution of Earth-system science, Nature Computational Science, 1, 104–113, doi: 10.1038/s43588-021-00023-0, 2021.

Bergman, J. W., Jensen, E. J., Pfister, L., and Yang, Q.: Seasonal differences of vertical-transport efficiency in the tropical tropopause layer: On the interplay between tropical deep convection, large-scale vertical ascent, and horizontal circulations, Journal of Geophysical Research, 117, D05 302, doi: 10.1029/2011JD016992, 2012.

Bian, J., Li, D., Bai, Z., Li, Q., Lyu, D., and Zhou, X.: Transport of Asian surface pollutants to the global stratosphere from the Tibetan Plateau region during the Asian summer monsoon, National Science Review, 7, 516–533, doi: 10.1093/nsr/nwaa005, 2020.

Bossolasco, A., Jegou, F., Sellitto, P., Berthet, G., Kloss, C., and Legras, B.: Global modeling studies of composition and decadal trends of the Asian Tropopause Aerosol Layer, Atmospheric Chemistry and Physics, 21, 2745–2764, doi: https://doi.org/10.5194/acp-21-2745-2021, 2021.

Bowman, K. P., Lin, J. C., Stohl, A., Draxler, R., Konopka, P., Andrews, A., and Brunner, D.: Input Data Requirements for Lagrangian Trajectory Models, Bulletin of the American Meteorological Society, 94, 1051–1058, doi: 10.1175/BAMS-D-12-00076.1, 2013.

Brabec, M., Wienhold, F. G., Luo, B. P., Vömel, H., Immler, F., Steiner, P., Hausammann, E., Weers, U., and Peter, T.: Particle backscatter and relative humidity measured across cirrus clouds and comparison with microphysical cirrus modelling, Atmospheric Chemistry and Physics, 12, 9135–9148, doi: 10.5194/acp-12-9135-2012, 2012.

Brasseur, G. P. and Jacob, D. J.: Modeling of Atmospheric Chemistry, Cambridge University Press, doi: 10.1017/9781316544754, 2017.

Brecht, R., Bakels, L., Bihlo, A., and Stohl, A.: Improving trajectory calculations by FLEXPART 10.4+ using single-image super-resolution, Geoscientific Model Development, 16, 2181–2192, doi: 10.5194/gmd-16-2181-2023, 2023.

Brinkop, S. and Jöckel, P.: ATTILA 4.0: Lagrangian advective and convective transport of passive tracers within the ECHAM5/MESSy (2.53.0) chemistry–climate model, Geoscientific Model Development, 12, 1991–2008, doi: 10.5194/gmd-12-1991-2019, 2019.

Brunamonti, S., Jorge, T., Oelsner, P., Hanumanthu, S., Singh, B. B., Kumar, K. R., Sonbawne, S., Meier, S., Singh, D., Wienhold, F. G., Luo, B. P., Boettcher, M., Poltera, Y., Jauhiainen, H., Kayastha, R., Karmacharya, J., Dirksen, R., Naja, M., Rex, M., Fadnavis, S., and Peter, T.: Balloon-borne measurements of temperature, water vapor, ozone and aerosol backscatter on the southern slopes of the Himalayas during StratoClim 2016–2017, Atmospheric Chemistry and Physics, 18, 15 937–15 957, doi: 10.5194/acp-18-15937-2018, 2018.

Bucci, S., Legras, B., Sellitto, P., D'Amato, F., Viciani, S., Montori, A., Chiarugi, A., Ravegnani, F., Ulanovsky, A., Cairo, F., and Stroh, F.: Deep-convective influence on the upper troposphere–lower stratosphere composition in the Asian monsoon anticyclone region: 2017

StratoClim campaign results, Atmospheric Chemistry and Physics, 20, 12 193–12 210, doi: 10.5194/acp-20-12193-2020, 2020.

Bucholtz, A.: Rayleigh-scattering calculations for the terrestrial atmosphere, Applied Optics, 34, 2765–2773, doi: 10.1364/AO.34.002765, 1995.

Clemens, J., Ploeger, F., Konopka, P., Portmann, R., Sprenger, M., and Wernli, H.: Characterization of transport from the Asian summer monsoon anticyclone into the UTLS via shedding of low potential vorticity cutoffs, Atmospheric Chemistry and Physics, 22, 3841–3860, doi: 10.5194/acp-22-3841-2022, 2022.

Clemens, J., Hoffmann, L., Vogel, B., Grießbach, S., and Thomas, N.: Implementation and evaluation of diabatic advection in the Lagrangian transport model MPTRAC 2.6, Geoscientific Model Development, 17, 4467–4493, doi: 10.5194/gmd-17-4467-2024, 2024a.

Clemens, J., Vogel, B., Hoffmann, L., Griessbach, S., Thomas, N., Fadnavis, S., Müller, R., Peter, T., and Ploeger, F.: A multi-scenario Lagrangian trajectory analysis to identify source regions of the Asian tropopause aerosol layer on the Indian subcontinent in August 2016, Atmospheric Chemistry and Physics, 24, 763–787, doi: 10.5194/acp-24-763-2024, 2024b.

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, doi: 10.1002/qj.828, 2011.

Dethof, A., O'Neill, A., Slingo, J. M., and Smit, H. G. J.: A mechanism for moistening the lower stratosphere involving the Asian summer monsoon, Quarterly Journal of the Royal Meteorological Society, 125, 1079–1106, doi: https://doi.org/10.1002/qj.1999.49712555602, 1999.

Dey, D. and Döös, K.: Tracing the Origin of the South Asian Summer Monsoon Precipitation and Its Variability Using a Novel Lagrangian Framework, Journal of Climate, 34, 8655 – 8668, doi: 10.1175/JCLI-D-20-0967.1, 2021.

Dey, S. and Di Girolamo, L.: A climatology of aerosol optical and microphysical properties over the Indian subcontinent from 9 years (2000–2008) of Multiangle Imaging Spectroradiometer (MISR) data, Journal of Geophysical Research: Atmospheres, 115, doi: 10.1029/2009JD013395, 2010.

Eppers, O., Köllner, F., Appel, O., Brauner, P., Ekinci, F., Molleker, S., Dragoneas, A., Hünig, A., Smith, W., Ueyama, R., Schneider, J., and Borrmann, S.: Chemical composition and processing of aerosol particles in the Asian Tropopause Aerosol Layer inferred from airborne measurements during the ACCLIP campaign, Tech. rep., doi: 10.5194/egusphere-egu23-12081, 2023.

Fadnavis, S., Semeniuk, K., Pozzoli, L., Schultz, M. G., Ghude, S. D., Das, S., and Kakatkar, R.: Transport of aerosols into the UTLS and their impact on the Asian monsoon region as seen in a global model simulation, Atmospheric Chemistry and Physics, 13, 8771–8786, doi: 10.5194/acp-13-8771-2013, 2013.

Fadnavis, S., Kalita, G., Kumar, K. R., Gasparini, B., and Li, J.-L. F.: Potential impact of carbonaceous aerosol on the upper troposphere and lower stratosphere (UTLS) and precipitation during Asian summer monsoon in a global model simulation, Atmospheric Chemistry and Physics, 17, 11 637–11 654, doi: 10.5194/acp-17-11637-2017, 2017.

Fadnavis, S., Roy, C., Chattopadhyay, R., Sioris, C. E., Rap, A., Müller, R., Kumar, K. R., and Krishnan, R.: Transport of trace gases via eddy shedding from the Asian summer monsoon anticyclone and associated impacts on ozone heating rates, Atmospheric Chemistry and Physics, 18, 11 493–11 506, doi: 10.5194/acp-18-11493-2018, 2018.

Fadnavis, S., Sabin, T. P., Roy, C., Rowlinson, M., Rap, A., Vernier, J.-P., and Sioris, C. E.: Elevated aerosol layer over South Asia worsens the Indian droughts, Scientific Reports, 9, 10 268, doi: 10.1038/s41598-019-46704-9, 2019.

Fairlie, T. D., Liu, H., Vernier, J.-P., Campuzano-Jost, P., Jimenez, J. L., Jo, D. S., Zhang, B., Natarajan, M., Avery, M. A., and Huey, G.: Estimates of Regional Source Contributions to the Asian Tropopause Aerosol Layer Using a Chemical Transport Model, Journal of Geophysical Research: Atmospheres, 125, e2019JD031 506, doi: 10.1029/2019JD031506, e2019JD031506 2019JD031506, 2020.

Friedlingstein, P., Jones, M. W., O'Sullivan, M., Andrew, R. M., Bakker, D. C. E., Hauck, J., Le Quéré, C., Peters, G. P., Peters, W., Pongratz, J., Sitch, S., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S. R., Anthoni, P., Bates, N. R., Becker, M., Bellouin, N., Bopp, L., Chau, T. T. T.,

Chevallier, F., Chini, L. P., Cronin, M., Currie, K. I., Decharme, B., Djeutchouang, L. M., Dou, X., Evans, W., Feely, R. A., Feng, L., Gasser, T., Gilfillan, D., Gkritzalis, T., Grassi, G., Gregor, L., Gruber, N., Gürses, O., Harris, I., Houghton, R. A., Hurtt, G. C., Iida, Y., Ilyina, T., Luijkx, I. T., Jain, A., Jones, S. D., Kato, E., Kennedy, D., Klein Goldewijk, K., Knauer, J., Korsbakken, J. I., Körtzinger, A., Landschützer, P., Lauvset, S. K., Lefèvre, N., Lienert, S., Liu, J., Marland, G., McGuire, P. C., Melton, J. R., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S.-I., Niwa, Y., Ono, T., Pierrot, D., Poulter, B., Rehder, G., Resplandy, L., Robertson, E., Rödenbeck, C., Rosan, T. M., Schwinger, J., Schwingshackl, C., Séférian, R., Sutton, A. J., Sweeney, C., Tanhua, T., Tans, P. P., Tian, H., Tilbrook, B., Tubiello, F., van der Werf, G. R., Vuichard, N., Wada, C., Wanninkhof, R., Watson, A. J., Willis, D., Wiltshire, A. J., Yuan, W., Yue, C., Yue, X., Zaehle, S., and Zeng, J.: Global Carbon Budget 2021, Earth System Science Data, 14, 1917–2005, doi: 10.5194/essd-14-1917-2022, 2022.

Fu, R., Hu, Y., Wright, J. S., Jiang, J. H., Dickinson, R. E., Chen, M., Filipiak, M., Read, W. G., Waters, J. W., and Wu, D. L.: Short circuit of water vapor and polluted air to the global stratosphere by convective transport over the Tibetan Plateau, Proceedings of the National Academy of Sciences, 103, 5664–5669, doi: 10.1073/pnas.0601584103, 2006.

Gandham, H., Dasari, H. P., Karumuri, A., Ravuri, P. M. K., and Hoteit, I.: Three-dimensional structure and transport pathways of dust aerosols over West Asia, Climate and Atmospheric Science, 5, 1–15, doi: 10.1038/s41612-022-00266-2, 2022.

Garny, H. and Randel, W. J.: Dynamic variability of the Asian monsoon anticyclone observed in potential vorticity and correlations with tracer distributions, Journal of Geophysical Research: Atmospheres, 118, 13,421–13,433, doi: 10.1002/2013JD020908, 2013.

Gerbig, C., Lin, J. C., Wofsy, S. C., Daube, B. C., Andrews, A. E., Stephens, B. B., Bakwin, P. S., and Grainger, C. A.: Toward constraining regional-scale fluxes of CO2 with atmospheric observations over a continent: 2. Analysis of COBRA data using a receptor-oriented framework, Journal Geophysical Research, 108, doi: 10.1029/2003JD003770, 2003.

Global Modeling and Assimilation Office: MERRA-2 tavgM_2d_aer Nx: 2d, Monthly mean, Time-averaged, Single-Level, Assimilation, Aerosol Diagnostics V5.12.4, Greenbelt, MD, USA, Brussels, Goddard Earth Sciences Data and Information Services Center (GES DISC), http://dx.doi.org/10.5067/FH9A0MLJPC7N, doi: 10.5067/FH9A0MLJPC7N, 2015a.

Global Modeling and Assimilation Office: MERRA-2 tavgM_2d_adg_Nx: 2d, Monthly mean, Time-averaged, Single-Level, Assimilation, Aerosol Diagnostics (extended) V5.12.4, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), http://dx.doi.org/10.5067/RZIK2TV7PP38, doi: 10.5067/RZIK2TV7PP38, 2015b.

Gottschaldt, K.-D., Schlager, H., Baumann, R., Cai, D. S., Eyring, V., Graf, P., Grewe, V., Jöckel, P., Jurkat-Witschas, T., Voigt, C., Zahn, A., and Ziereis, H.: Dynamics and composition of the Asian summer monsoon anticyclone, Atmospheric Chemistry and Physics, 18, 5655–5675, doi: 10.5194/acp-18-5655-2018, 2018.

Gu, Y., Liao, H., and Bian, J.: Summertime nitrate aerosol in the upper troposphere and lower stratosphere over the Tibetan Plateau and the South Asian summer monsoon region, Atmospheric Chemistry and Physics, 16, 6641–6663, doi: 10.5194/acp-16-6641-2016, 2016.

HALO: Probing High Latitude Export of air from the Asian Summer Monsoon (PHILEAS), Access date: 19 November 2023, URL https://halo-research.de/sience/halo-missions/current-missions/phileas/, 2023.

Hanumanthu, S., Vogel, B., Müller, R., Brunamonti, S., Fadnavis, S., Li, D., Ölsner, P., Naja, M., Singh, B. B., Kumar, K. R., Sonbawne, S., Jauhiainen, H., Vömel, H., Luo, B., Jorge, T., Wienhold, F. G., Dirkson, R., and Peter, T.: Strong day-to-day variability of the Asian Tropopause Aerosol Layer (ATAL) in August 2016 at the Himalayan foothills, Atmospheric Chemistry and Physics, 20, 14 273–14 302, doi: 10.5194/acp-20-14273-2020, 2020.

He, Q., Ma, J., Zheng, X., Wang, Y., Wang, Y., Mu, H., Cheng, T., He, R., Huang, G., Liu, D., and Lelieveld, J.: Formation and dissipation dynamics of the Asian tropopause aerosol layer, Environmental Research Letters, 16, 014 015, doi: 10.1088/1748-9326/abcd5d, 2020.

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, doi: 10.1002/qj.3803, 2020.

Hoffmann, L., Rößler, T., Griessbach, S., Heng, Y., and Stein, O.: Lagrangian transport simulations of volcanic sulfur dioxide emissions: impact of meteorological data products, Journal Geophysical Research, pp. 4651–4673, doi: 10.1002/2015JD023749, 2016.

Hoffmann, L., Günther, G., Li, D., Stein, O., Wu, X., Griessbach, S., Heng, Y., Konopka, P., Müller, R., Vogel, B., and Wright, J. S.: From ERA-Interim to ERA5: the considerable impact of ECMWF's next-generation reanalysis on Lagrangian transport simulations, Atmospheric Chemistry and Physics, 19, 3097–3124, doi: 10.5194/acp-19-3097-2019, 2019.

Hoffmann, L., Baumeister, P. F., Cai, Z., Clemens, J., Griessbach, S., Günther, G., Heng, Y., Liu, M., Haghighi Mood, K., Stein, O., Thomas, N., Vogel, B., Wu, X., and Zou, L.: Massive-Parallel Trajectory Calculations version 2.2 (MPTRAC-2.2): Lagrangian transport simulations on graphics processing units (GPUs), Geoscientific Model Development, 15, 2731–2762, doi: 10.5194/gmd-15-2731-2022, 2022.

Hoffmann, L., Clemens, J., Griessbach, S., Haghighi Mood, K., Khosrawi, F., Liu, M., Lu, Y.-S., Sonnabend, J., and Zou, L.: Massive-Parallel Trajectory Calculations (MPTRAC) v2.6, doi: 10.5281/zenodo.10067751, 2023a.

Hoffmann, L., Konopka, P., Clemens, J., and Vogel, B.: Lagrangian transport simulations using the extreme convection parameterization: an assessment for the ECMWF reanalyses, Atmospheric Chemistry and Physics, 23, 7589–7609, doi: 10.5194/acp-23-7589-2023, 2023b.

Hoffmann, L., Haghighi Mood, K., Herten, A., Hrywniak, M., Kraus, J., Clemens, J., and Liu, M.: Accelerating Lagrangian transport simulations on graphics processing units: performance optimizations of Massive-Parallel Trajectory Calculations (MPTRAC) v2.6, Geoscientific Model Development, 17, 4077–4094, doi: 10.5194/gmd-17-4077-2024, 2024.

Homeyer, C. R. and Bowman, K. P.: Rossby Wave Breaking and Transport between the Tropics and Extratropics above the Subtropical Jet, Journal of the Atmospheric Sciences, 70, 607 – 626, doi: 10.1175/JAS-D-12-0198.1, 2013.

Honomichl, S. B. and Pan, L. L.: Transport From the Asian Summer Monsoon Anticyclone Over the Western Pacific, Journal of Geophysical Research: Atmospheres, 125, e2019JD032 094, doi: 10.1029/2019JD032094, 2020.

Höpfner, M., Ungermann, J., Borrmann, S., Wagner, R., Spang, R., Riese, M., Stiller, G., Appel, O., Batenburg, A. M., Bucci, S., Cairo, F., Dragoneas, A., Friedl-Vallon, F., Hünig, A., Johans-

son, S., Krasauskas, L., Legras, B., Leisner, T., Mahnke, C., Möhler, O., Molleker, S., Müller, R., Neubert, T., Orphal, J., Preusse, P., Rex, M., Saathoff, H., Stroh, F., Weigel, R., and Wohltmann, I.: Ammonium nitrate particles formed in upper troposphere from ground ammonia sources during Asian monsoons, Nature Geoscience, 12, 608–612, doi: 10.1038/s41561-019-0385-8, 2019.

Hsu, C. J.: Eddy Shedding from Non-axisymmetric, Divergent Anticyclones with Application to the Asian Monsoon Anticyclone, URL https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=86b8a646f9fe0ee00ff48c03d170c2a1d3a72f96, 1998.

Intel: An Introduction to MPI-3 Shared Memory Programming, Access date: 19 November 2023, URL https://www.intel.com/content/www/us/en/developer/articles/technical/an-introduction-to-mpi-3-shared-memory-programming.html, 2018.

IPCC, I.: Summary for Policymakers, p. 3–32, Cambridge University Press, doi: 10.1017/9781009157896.001, 2023.

Johansson, S., Höpfner, M., Kirner, O., Wohltmann, I., Bucci, S., Legras, B., Friedl-Vallon, F., Glatthor, N., Kretschmer, E., Ungermann, J., and Wetzel, G.: Pollution trace gas distributions and their transport in the Asian monsoon upper troposphere and lowermost stratosphere during the StratoClim campaign 2017, Atmospheric Chemistry and Physics, 20, 14 695–14 715, doi: 10.5194/acp-20-14695-2020, 2020.

JSC: Jülich Wizard for European Leadership Science (JUWELS), Access date: 19 November 2023, URL https://www.fz-juelich.de/en/ias/jsc/systems/supercomputers/juwels, 2022.

Kim, Y.-S., Shibata, T., Iwasaka, Y., Shi, G., Zhou, X., Tamura, K., and Ohashi, T.: Enhancement of aerosols near the cold tropopause in summer over Tibetan Plateau: lidar and balloonborne measurements in 1999 at Lhasa, Tibet, China, in: Lidar Remote Sensing for Industry and Environment Monitoring III, edited by Singh, U. N., Itabe, T., and Liu, Z., vol. 4893, pp. 496 – 503, International Society for Optics and Photonics, SPIE, doi: 10.1117/12.466090, 2003.

Köllner, F., Appel, O., Dragoneas, A., Hünig, A., Molleker, S., Ebert, M., and Borrmann, S.: Chemical analysis of the 2017 ATAL measured during StratoClim – New insights into refractory aerosol components, Tech. rep., doi: 10.5194/egusphere-egu23-13722, 2023.

Konopka, P., Günther, G., Müller, R., dos Santos, F. H. S., Schiller, C., Ravegnani, F., Ulanovsky, A., Schlager, H., Volk, C. M., Viciani, S., Pan, L. L., McKenna, D.-S., and Riese, M.: Contri-

bution of mixing to upward transport across the tropical tropopause layer (TTL), Atmospheric Chemistry and Physics, 7, 3285–3308, doi: 10.5194/acp-7-3285-2007, 2007.

Konopka, P., Tao, M., Ploeger, F., Diallo, M., and Riese, M.: Tropospheric mixing and parametrization of unresolved convective updrafts as implemented in the Chemical Lagrangian Model of the Stratosphere (CLaMS v2.0), Geoscientific Model Development, 12, 2441–2462, doi: 10.5194/gmd-12-2441-2019, 2019.

Krotkov, N. A., Lamsal, L. N., Marchenko, S. V., Celarier, E. A., J.Bucsela, E., Swartz, W. H., Joiner, J., and OMI core team: OMI/Aura NO2 Cloud-Screened Total and Tropospheric Column L3 Global Gridded 0.25 degree x 0.25 degree V3, NASA Goddard Space Flight Center, Goddard Earth Sciences Data and Information Services Center (GES DISC), http://dx.doi.org/10.5067/Aura/OMI/DATA3007, doi: 10.5067/Aura/OMI/DATA3007, 2019.

Kurokawa, J. and Ohara, T.: Long-term historical trends in air pollutant emissions in Asia: Regional Emission inventory in Asia (REAS) version 3, Atmospheric Chemistry and Physics, 20, 12 761–12 793, doi: 10.5194/acp-20-12761-2020, 2020.

Kuttippurath, J., Singh, A., Dash, S. P., Mallick, N., Clerbaux, C., Van Damme, M., Clarisse, L., Coheur, P. F., Raj, S., Abbhishek, K., and Varikoden, H.: Record high levels of atmospheric ammonia over India: Spatial and temporal analyses, Science of The Total Environment, 740, 139 986, doi: 10.1016/j.scitotenv.2020.139986, 2020.

Lau, W. K. M., Yuan, C., and Li, Z.: Origin, Maintenance and Variability of the Asian Tropopause Aerosol Layer (ATAL): The Roles of Monsoon Dynamics, Scientific Reports, 8, 3960, doi: 10.1038/s41598-018-22267-z, 2018.

Lauther, V., Vogel, B., Wintel, J., Rau, A., Hoor, P., Bense, V., Müller, R., and Volk, C. M.: In situ observations of CH_2Cl_2 and $CHCl_3$ show efficient transport pathways for very short-lived species into the lower stratosphere via the Asian and the North American summer monsoon, Atmospheric Chemistry and Physics, 22, 2049–2077, doi: 10.5194/acp-22-2049-2022, 2022.

Legras, B. and Bucci, S.: Confinement of air in the Asian monsoon anticyclone and pathways of convective air to the stratosphere during the summer season, Atmospheric Chemistry and Physics, 20, 11 045–11 064, doi: 10.5194/acp-20-11045-2020, 2020.

Li, C., McLinden, C., Fioletov, V., Krotkov, N., Carn, S., Joiner, J., Streets, D., He, H., Ren, X., Li, Z., and Dickerson, R. R.: India Is Overtaking China as the World's Largest Emitter of

Anthropogenic Sulfur Dioxide, Scientific Reports, 7, 14 304, doi: 10.1038/s41598-017-14639-8, 2017a.

Li, D., Vogel, B., Bian, J., Müller, R., Pan, L. L., Günther, G., Bai, Z., Li, Q., Zhang, J., Fan, Q., and Vömel, H.: Impact of typhoons on the composition of the upper troposphere within the Asian summer monsoon anticyclone: the SWOP campaign in Lhasa 2013, Atmospheric Chemistry and Physics, 17, 4657–4672, doi: 10.5194/acp-17-4657-2017, 2017b.

Li, D., Vogel, B., Müller, R., Bian, J., Günther, G., Ploeger, F., Li, Q., Zhang, J., Bai, Z., Vömel, H., and Riese, M.: Dehydration and low ozone in the tropopause layer over the Asian monsoon caused by tropical cyclones: Lagrangian transport calculations using ERA-Interim and ERA5 reanalysis data, Atmospheric Chemistry and Physics, 20, 4133–4152, doi: 10.5194/acp-20-4133-2020, 2020.

Li, Q., Jiang, J. H., Wu, D. L., Read, W. G., Livesey, N. J., Waters, J. W., Zhang, Y., Wang, B., Filipiak, M. J., Davis, C. P., Turquety, S., Wu, S., Park, R. J., Yantosca, R. M., and Jacob, D. J.: Convective outflow of South Asian pollution: A global CTM simulation compared with EOS MLS observations, Geophysical Research Letters, 32, doi: 10.1029/2005GL022762, 2005.

Li, Z., Lau, W. K.-M., Ramanathan, V., Wu, G., Ding, Y., Manoj, M. G., Liu, J., Qian, Y., Li, J., Zhou, T., Fan, J., Rosenfeld, D., Ming, Y., Wang, Y., Huang, J., Wang, B., Xu, X., Lee, S.-S., Cribb, M., Zhang, F., Yang, X., Zhao, C., Takemura, T., Wang, K., Xia, X., Yin, Y., Zhang, H., Guo, J., Zhai, P. M., Sugimoto, N., Babu, S. S., and Brasseur, G. P.: Aerosol and monsoon climate interactions over Asia, Reviews of Geophysics, 54, 866–929, doi: 10.1002/2015RG000500, 2016.

Lin, J., Brunner, D., Gerbig, C., Stohl, A., Luhar, A., and Webley, P.: Lagrangian Modeling of the Atmosphere, Geophysical Monograph Series, Wiley, URL `https://books.google.de/books?id=BNxYenge_K4C`, 2013.

Loughner, C. P., Fasoli, B., Stein, A. F., and Lin, J. C.: Incorporating Features from the Stochastic Time-Inverted Lagrangian Transport (STILT) Model into the Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT) Model: A Unified Dispersion Model for Time-Forward and Time-Reversed Applications, Journal of Applied Meteorology and Climatology, 60, 799–810, doi: 10.1175/JAMC-D-20-0158.1, 2021.

Ma, J., Brühl, C., He, Q., Steil, B., Karydis, V. A., Klingmüller, K., Tost, H., Chen, B., Jin, Y., Liu, N., Xu, X., Yan, P., Zhou, X., Abdelrahman, K., Pozzer, A., and Lelieveld, J.: Modeling the aerosol chemical composition of the tropopause over the Tibetan Plateau during

the Asian summer monsoon, Atmospheric Chemistry and Physics, 19, 11 587–11 612, doi: 10.5194/acp-19-11587-2019, 2019.

Mahowald, N. M., Plumb, R. A., Rasch, P. J., del Corral, J., Sassi, F., and Heres, W.: Stratospheric transport in a three-dimensional isentropic coordinate model, Journal of Geophysical Research: Atmospheres, 107, ACH 3–1–ACH 3–14, doi: 10.1029/2001JD001313, 2002.

Malakar, P., Kesarkar, A., Bhate, J., Singh, V., and Deshamukhya, A.: Comparison of Reanalysis Data Sets to Comprehend the Evolution of Tropical Cyclones Over North Indian Ocean, Earth and Space Science, 7, e2019EA000 978, doi: 10.1029/2019EA000978, e2019EA000978 2019EA000978, 2020.

Manney, G. L., Santee, M. L., Lawrence, Z. D., Wargan, K., and Schwartz, M. J.: A Moments View of Climatology and Variability of the Asian Summer Monsoon Anticyclone, Journal of Climate, 34, 7821 – 7841, doi: https://doi.org/10.1175/JCLI-D-20-0729.1, 2021.

McKenna, D. S., Grooß, J.-U., Günther, G., Konopka, P., Müller, R., Carver, G., and Sasano, Y.: A new Chemical Lagrangian Model of the Stratosphere (CLaMS) 2. Formulation of chemistry scheme and initialization, Journal of Geophysical Research: Atmospheres, 107, ACH 4–1–ACH 4–14, doi: 10.1029/2000JD000113, 2002a.

McKenna, D. S., Konopka, P., Grooß, J.-U., Günther, G., Müller, R., Spang, R., Offermann, D., and Orsolini, Y.: A new Chemical Lagrangian Model of the Stratosphere (CLaMS) 1. Formulation of advection and mixing, Journal of Geophysical Research: Atmospheres, 107, ACH 15–1–ACH 15–15, doi: 10.1029/2000JD000114, 2002b.

Mehta, K. and Gabriel, E.: Multi-Threaded Parallel I/O for OpenMP Applications, International Journal of Parallel Programming, 43, 286–309, doi: 10.1007/s10766-014-0306-9, 2015.

Montzka, S. A., Dutton, G. S., Portmann, R. W., Chipperfield, M. P., Davis, S., Feng, W., Manning, A. J., Ray, E., Rigby, M., Hall, B. D., Siso, C., Nance, J. D., Krummel, P. B., Mühle, J., Young, D., O'Doherty, S., Salameh, P. K., Harth, C. M., Prinn, R. G., Weiss, R. F., Elkins, J. W., Walter-Terrinoni, H., and Theodoridi, C.: A decline in global CFC-11 emissions during 2018-2019, Nature, 590, 428–432, doi: 10.1038/s41586-021-03260-5, 2021.

NCAR: Asian Summer Monsoon Chemical and Climate Impact Project (ACCLIP), Access date: 19 November 2023, URL `https://www2.acom.ucar.edu/acclip`, 2023.

Neely III, R. R., Yu, P., Rosenlof, K. H., Toon, O. B., Daniel, J. S., Solomon, S., and Miller, H. L.: The contribution of anthropogenic SO2 emissions to the Asian tropopause aerosol layer, Journal of Geophysical Research: Atmospheres, 119, 1571–1579, doi: 10.1002/2013JD020578, 2014.

Nützel, M., Dameris, M., and Garny, H.: Movement, drivers and bimodality of the South Asian High, Atmospheric Chemistry and Physics, 16, 14 755–14 774, doi: 10.5194/acp-16-14755-2016, 2016.

Pan, L. L., Kinnison, D., Liang, Q., Chin, M., Santee, M. L., Flemming, J., Smith, W. P., Honomichl, S. B., Bresch, J. F., Lait, L. R., Zhu, Y., Tilmes, S., Colarco, P. R., Warner, J., Vuvan, A., Clerbaux, C., Atlas, E. L., Newman, P. A., Thornberry, T., Randel, W. J., and Toon, O. B.: A Multimodel Investigation of Asian Summer Monsoon UTLS Transport Over the Western Pacific, Journal of Geophysical Research: Atmospheres, 127, e2022JD037 511, doi: 10.1029/2022JD037511, 2022.

Ploeger, F.: Impact of different vertical transport representations on simulating processes in the tropical tropopause layer (TTL), Ph.D. thesis, 2010.

Ploeger, F., Konopka, P., Günther, G., Grooß, J.-U., and Müller, R.: Impact of the vertical velocity scheme on modeling transport in the tropical tropopause layer, Journal of Geophysical Research: Atmospheres, 115, doi: 10.1029/2009JD012023, 2010.

Ploeger, F., Fueglistaler, S., Grooß, J.-U., Günther, G., Konopka, P., Liu, Y. S., Müller, R., Ravegnani, F., Schiller, C., Ulanovski, A., and Riese, M.: Insight from ozone and water vapour on transport in the tropical tropopause layer (TTL), Atmospheric Chemistry and Physics, 11, 407–419, doi: 10.5194/acp-11-407-2011, 2011.

Ploeger, F., Gottschling, C., Griessbach, S., Grooß, J.-U., Guenther, G., Konopka, P., Müller, R., Riese, M., Stroh, F., Tao, M., Ungermann, J., Vogel, B., and von Hobe, M.: A potential vorticity-based determination of the transport barrier in the Asian summer monsoon anticyclone, Atmospheric Chemistry and Physics, 15, 13 145–13 159, doi: 10.5194/acp-15-13145-2015, 2015.

Ploeger, F., Diallo, M., Charlesworth, E., Konopka, P., Legras, B., Laube, J. C., Grooß, J.-U., Günther, G., Engel, A., and Riese, M.: The stratospheric Brewer–Dobson circulation inferred from age of air in the ERA5 reanalysis, Atmospheric Chemistry and Physics, 21, 8393–8412, doi: 10.5194/acp-21-8393-2021, 2021.

Pommrich, R., Müller, R., Grooß, J.-U., Konopka, P., Ploeger, F., Vogel, B., Tao, M., Hoppe, C. M., Günther, G., Spelten, N., Hoffmann, L., Pumphrey, H.-C., Viciani, S., D'Amato, F., Volk, C. M., Hoor, P., Schlager, H., and Riese, M.: Tropical troposphere to stratosphere transport of carbon monoxide and long-lived trace species in the Chemical Lagrangian Model of the Stratosphere (CLaMS), Geoscientific Model Development, 7, 2895–2916, doi: 10.5194/gmd-7-2895-2014, 2014.

Popovic, J. M. and Plumb, R. A.: Eddy Shedding from the Upper-Tropospheric Asian Monsoon Anticyclone, Journal of the Atmospheric Sciences, 58, 93 – 104, doi: 10.1175/1520-0469(2001)058<0093:ESFTUT>2.0.CO;2, 2001.

Quarteroni, A., Sacco, R., and Saleri, F.: Numerical Mathematics, Texts in Applied Mathematics, Springer New York, URL https://books.google.de/books?id=FA3_DQAAQBAJ, 2017.

Randel, W. J., Park, M., Emmons, L., Kinnison, D., Bernath, P., Walker, K. A., Boone, C., and Pumphrey, H.: Asian Monsoon Transport of Pollution to the Stratosphere, Science, 328, 611–613, doi: 10.1126/science.1182274, 2010.

Rigby, M., Park, S., Saito, T., Western, L. M., Redington, A. L., Fang, X., Henne, S., Manning, A. J., Prinn, R. G., Dutton, G. S., Fraser, P. J., Ganesan, A. L., Hall, B. D., Harth, C. M., Kim, J., Kim, K.-R., Krummel, P. B., Lee, T., Li, S., Liang, Q., Lunt, M. F., Montzka, S. A., Mühle, J., O'Doherty, S., Park, M.-K., Reimann, S., Salameh, P. K., Simmonds, P., Tunnicliffe, R. L., Weiss, R. F., Yokouchi, Y., and Young, D.: Increase in CFC-11 emissions from eastern China based on atmospheric observations, Nature, 569, 546–550, doi: 10.1038/s41586-019-1193-4, 2019.

Rolf, C., Vogel, B., Hoor, P., Afchine, A., Günther, G., Krämer, M., Müller, R., Müller, S., Spelten, N., and Riese, M.: Water vapor increase in the lower stratosphere of the Northern Hemisphere due to the Asian monsoon anticyclone observed during the TACTS/ESMVal campaigns, Atmospheric Chemistry and Physics, 18, 2973–2983, doi: 10.5194/acp-18-2973-2018, 2018.

Rolph, G. D. and Draxler, R. R.: Sensitivity of Three-Dimensional Trajectories to the Spatial and Temporal Densities of the Wind Field, Journal of Applied Meteorology and Climatology, 29, 1043–1054, doi: 10.1175/1520-0450(1990)029<1043:SOTDTT>2.0.CO;2, 1990.

Rößler, T., Stein, O., Heng, Y., Baumeister, P., and Hoffmann, L.: Trajectory errors of different numerical integration schemes diagnosed with the MPTRAC advection module driven by

ECMWF operational analyses, Geoscientific Model Development, 11, 575–592, doi: 10.5194/gmd-11-575-2018, 2018.

Rupp, P. and Haynes, P.: Zonal scale and temporal variability of the Asian monsoon anticyclone in an idealised numerical model, Weather and Climate Dynamics, 2, 413–431, doi: 10.5194/wcd-2-413-2021, 2021.

Santee, M. L., Manney, G. L., Livesey, N. J., Schwartz, M. J., Neu, J. L., and Read, W. G.: A comprehensive overview of the climatological composition of the Asian summer monsoon anticyclone based on 10 years of Aura Microwave Limb Sounder measurements, Journal of Geophysical Research: Atmospheres, 122, 5491–5514, doi: 10.1002/2016JD026408, 2017.

Schill, G., Murphy, D., Lawler, M., and Abou-Ghanem, M.: Single-Particle Aerosol Composition in the Asian Tropopause Aerosol Layer and in the North American Upper Troposphere/Lower Stratosphere during ACCLIP, Tech. rep., doi: 10.5194/egusphere-egu23-9792, 2023.

Schneider, J., Weigel, R., Klimach, T., Dragoneas, A., Appel, O., Hünig, A., Molleker, S., Köllner, F., Clemen, H.-C., Eppers, O., Hoppe, P., Hoor, P., Mahnke, C., Krämer, M., Rolf, C., Grooß, J.-U., Zahn, A., Obersteiner, F., Ravegnani, F., Ulanovsky, A., Schlager, H., Scheibe, M., Diskin, G. S., DiGangi, J. P., Nowak, J. B., Zöger, M., and Borrmann, S.: Aircraft-based observation of meteoric material in lower-stratospheric aerosol particles between 15 and 68° N, Atmospheric Chemistry and Physics, 21, 989–1013, doi: 10.5194/acp-21-989-2021, 2021.

Schoeberl, M. R. and Dessler, A. E.: Dehydration of the stratosphere, Atmospheric Chemistry and Physics, 11, 8433–8446, doi: 10.5194/acp-11-8433-2011, 2011.

Simmons, A. J., Burridge, D. M., Jarraud, M., Girard, C., and Wergen, W.: The ECMWF medium-range prediction models development of the numerical formulations and the impact of increased resolution, Meteorology and Atmospheric Physics, 40, 28–60, doi: 10.1007/BF01027467, 1989.

Sparling, L. C., Kettleborough, J. A., Haynes, P. H., McIntyre, M. E., Rosenfield, J. E., Schoeberl, M. R., and Newman, P. A.: Diabatic cross-isentropic dispersion in the lower stratosphere, Journal of Geophysical Research: Atmospheres, 102, 25 817–25 829, doi: 10.1029/97JD01968, 1997.

Stohl, A.: Computation, accuracy and applications of trajectories—A review and bibliography, Atmospheric Environment, 32, 947–966, doi: 10.1016/S1352-2310(97)00457-3, 1998.

Stohl, A., Wotawa, G., Seibert, P., and Kromp-Kolb, H.: Interpolation Errors in Wind Fields as a Function of Spatial and Temporal Resolution and Their Impact on Different Types of Kinematic Trajectories, Journal of Applied Meteorology and Climatology, 34, 2149–2165, doi: 10.1175/1520-0450(1995)034<2149:IEIWFA>2.0.CO;2, 1995.

Stohl, A., Haimberger, L., Scheele, M. P., and Wernli, H.: An intercomparison of results from three trajectory models, Meteorological Applications, 8, 127–135, doi: 10.1017/S1350482701002018, 2001.

Stohl, A., Cooper, O. R., and James, P.: A Cautionary Note on the Use of Meteorological Analysis Fields for Quantifying Atmospheric Mixing, Journal of the Atmospheric Sciences, 61, 1446–1453, doi: 10.1175/1520-0469(2004)061<1446:ACNOTU>2.0.CO;2, 2004.

Stohl, A., Forster, C., Frank, A., Seibert, P., and Wotawa, G.: Technical note: The Lagrangian particle dispersion model FLEXPART version 6.2, 5, 2461–2474, doi: 10.5194/acp-5-2461-2005, 2005.

Stroh, F. and StratoClim-Team: First detailed airborne and balloon measurements of microphysical, dynamical, and chemical processes in the Asian Summer Monsoon Anticyclone: overview and selected results of the 2016/2017 StratoClim field campaigns, to be submitted, 2023.

Sutton, R. T., Maclean, H., Swinbank, R., O'Neill, A., and Taylor, F. W.: High-Resolution Stratospheric Tracer Fields Estimated from Satellite Observations Using Lagrangian Trajectory Calculations, Journal of Atmospheric Sciences, 51, 2995 – 3005, doi: https://doi.org/10.1175/1520-0469(1994)051<2995:HRSTFE>2.0.CO;2, 1994.

Taszarek, M., Pilguj, N., Allen, J. T., Gensini, V., Brooks, H. E., and Szuster, P.: Comparison of convective parameters derived from ERA5 and MERRA2 with rawinsonde data over Europe and North America, Journal of Climate, pp. 1 – 55, doi: 10.1175/JCLI-D-20-0484.1, 2020.

Thomason, L. W. and Vernier, J.-P.: Improved SAGE II cloud/aerosol categorization and observations of the Asian tropopause aerosol layer: 1989-2005, Atmospheric Chemistry and Physics, 13, 4605–4616, doi: 10.5194/acp-13-4605-2013, 2013.

Tobo, Y., Iwasaka, Y., Shi, G.-Y., Kim, Y.-S., Ohashi, T., Tamura, K., and Zhang, D.: Balloon-borne observations of high aerosol concentrations near the summertime tropopause over the Tibetan Plateau, Atmospheric Research, 84, 233–241, doi: 10.1016/j.atmosres.2006.08.003, 2007.

Ungermann, J., Ern, M., Kaufmann, M., Müller, R., Spang, R., Ploeger, F., Vogel, B., and Riese, M.: Observations of PAN and its confinement in the Asian Summer Monsoon Anticyclone in high spatial resolution, Atmospheric Chemistry and Physics, 16, 8389–8403, doi: 10.5194/acp-16-8389-2016, 2016.

Unidata: Network Common Data Form (NetCDF), Access date: 5 November 2023, URL https://doi.org/10.5065/D6H70CW6, 2023.

Vernier, J.-P., Thomason, L. W., and Kar, J.: CALIPSO detection of an Asian tropopause aerosol layer, Geophysical Research Letters, 38, doi: 10.1029/2010GL046614, 2011.

Vernier, J.-P., Fairlie, T. D., Natarajan, M., Wienhold, F. G., Bian, J., Martinsson, B. G., Crumeyrolle, S., Thomason, L. W., and Bedka, K. M.: Increase in upper tropospheric and lower stratospheric aerosol levels and its potential connection with Asian pollution, Journal of Geophysical Research: Atmospheres, 120, 1608–1619, doi: https://doi.org/10.1002/2014JD022372, 2015.

Vernier, J.-P., Fairlie, T. D., Deshler, T., Ratnam, M. V., Gadhavi, H., Kumar, B. S., Natarajan, M., Pandit, A. K., Raj, S. T. A., Kumar, A. H., Jayaraman, A., Singh, A. K., Rastogi, N., Sinha, P. R., Kumar, S., Tiwari, S., Wegner, T., Baker, N., Vignelles, D., Stenchikov, G., Shevchenko, I., Smith, J., Bedka, K., Kesarkar, A., Singh, V., Bhate, J., Ravikiran, V., Rao, M. D., Ravindrababu, S., Patel, A., Vernier, H., Wienhold, F. G., Liu, H., Knepp, T. N., Thomason, L., Crawford, J., Ziemba, L., Moore, J., Crumeyrolle, S., Williamson, M., Berthet, G., Jégou, F., and Renard, J.-B.: BATAL: The Balloon Measurement Campaigns of the Asian Tropopause Aerosol Layer, Bulletin of the American Meteorological Society, 99, 955 – 973, doi: 10.1175/BAMS-D-17-0014.1, 2018.

Vogel, B., Günther, G., Müller, R., Grooß, J.-U., and Riese, M.: Impact of different Asian source regions on the composition of the Asian monsoon anticyclone and of the extratropical lowermost stratosphere, Atmospheric Chemistry and Physics, 15, 13 699–13 716, doi: 10.5194/acp-15-13699-2015, 2015.

Vogel, B., Günther, G., Müller, R., Grooß, J.-U., Afchine, A., Bozem, H., Hoor, P., Krämer, M., Müller, S., Riese, M., Rolf, C., Spelten, N., Stiller, G. P., Ungermann, J., and Zahn, A.: Long-range transport pathways of tropospheric source gases originating in Asia into the northern lower stratosphere during the Asian monsoon season 2012, Atmospheric Chemistry and Physics, 16, 15 301–15 325, doi: 10.5194/acp-16-15301-2016, 2016.

Vogel, B., Müller, R., Günther, G., Spang, R., Hanumanthu, S., Li, D., Riese, M., and Stiller, G. P.: Lagrangian simulations of the transport of young air masses to the top of the Asian monsoon anticyclone and into the tropical pipe, Atmospheric Chemistry and Physics, 19, 6007–6034, doi: 10.5194/acp-19-6007-2019, 2019.

Vogel, B., Volk, M., Wintel, J., Lauther, V., Clemens, J., Grooß, J.-U., Günther, G., Hoffmann, L., Laube, J. C., Müller, R., Ploeger, F., and Stroh, F.: Evaluation of vertical transport in the Asian monsoon 2017 from CO_2 reconstruction in the ERA5 and ERA-Interim reanalysis, EGUsphere, pp. 1–37, doi: 10.5194/egusphere-2023-1026, 2023.

Vogel, B., Volk, C. M., Wintel, J., Lauther, V., Clemens, J., Grooß, J.-U., Günther, G., Hoffmann, L., Laube, J. C., Müller, R., Ploeger, F., and Stroh, F.: Evaluation of vertical transport in ERA5 and ERA-Interim reanalysis using high-altitude aircraft measurements in the Asian summer monsoon 2017, Atmospheric Chemistry and Physics, 24, 317–343, doi: 10.5194/acp-24-317-2024, 2024.

von Hobe, M., Ploeger, F., Konopka, P., Kloss, C., Ulanowski, A., Yushkov, V., Ravegnani, F., Volk, C. M., Pan, L. L., Honomichl, S. B., Tilmes, S., Kinnison, D. E., Garcia, R. R., and Wright, J. S.: Upward transport into and within the Asian monsoon anticyclone as inferred from StratoClim trace gas observations, Atmospheric Chemistry and Physics, 21, 1267–1285, doi: 10.5194/acp-21-1267-2021, 2021.

Wang, M., Xiao, M., Bertozzi, B., Marie, G., Rörup, B., Schulze, B., Bardakov, R., He, X.-C., Shen, J., Scholz, W., Marten, R., Dada, L., Baalbaki, R., Lopez, B., Lamkaddam, H., Manninen, H. E., Amorim, A., Ataei, F., Bogert, P., Brasseur, Z., Caudillo, L., De Menezes, L.-P., Duplissy, J., Ekman, A. M. L., Finkenzeller, H., Carracedo, L. G., Granzin, M., Guida, R., Heinritzi, M., Hofbauer, V., Höhler, K., Korhonen, K., Krechmer, J. E., Kürten, A., Lehtipalo, K., Mahfouz, N. G. A., Makhmutov, V., Massabò, D., Mathot, S., Mauldin, R. L., Mentler, B., Müller, T., Onnela, A., Petäjä, T., Philippov, M., Piedehierro, A. A., Pozzer, A., Ranjithkumar, A., Schervish, M., Schobesberger, S., Simon, M., Stozhkov, Y., Tomé, A., Umo, N. S., Vogel, F., Wagner, R., Wang, D. S., Weber, S. K., Welti, A., Wu, Y., Zauner-Wieczorek, M., Sipilä, M., Winkler, P. M., Hansel, A., Baltensperger, U., Kulmala, M., Flagan, R. C., Curtius, J., Riipinen, I., Gordon, H., Lelieveld, J., El-Haddad, I., Volkamer, R., Worsnop, D. R., Christoudias, T., Kirkby, J., Möhler, O., and Donahue, N. M.: Synergistic HNO3–H2SO4–NH3 upper tropospheric particle formation, Nature, 605, 483–489, doi: 10.1038/s41586-022-04605-4, 2022.

Wang, T., Tang, J., Sun, M., Liu, X., Huang, Y., Huang, J., Han, Y., Cheng, Y., Huang, Z., and Li, J.: Identifying a transport mechanism of dust aerosols over South Asia to the Tibetan Plateau: A case study, The Science of the Total Environment, 758, 143 714, doi: 10.1016/j.scitotenv.2020.143714, 2021.

Wang, Y., Zhang, Q. Q., He, K., Zhang, Q., and Chai, L.: Sulfate-nitrate-ammonium aerosols over China: response to 2000–2015 emission changes of sulfur dioxide, nitrogen oxides, and ammonia, Atmospheric Chemistry and Physics, 13, 2635–2652, doi: 10.5194/acp-13-2635-2013, 2013.

Weigel, R., Mahnke, C., Baumgartner, M., Dragoneas, A., Vogel, B., Ploeger, F., Viciani, S., D'Amato, F., Bucci, S., Legras, B., Luo, B., and Borrmann, S.: In situ observation of new particle formation (NPF) in the tropical tropopause layer of the 2017 Asian monsoon anticyclone – Part 1: Summary of StratoClim results, Atmospheric Chemistry and Physics, 21, 11 689–11 722, doi: 10.5194/acp-21-11689-2021, 2021.

Wohltmann, I., Lehmann, R., Gottwald, G. A., Peters, K., Protat, A., Louf, V., Williams, C., Feng, W., and Rex, M.: A Lagrangian convective transport scheme including a simulation of the time air parcels spend in updrafts (LaConTra v1.0), Geoscientific Model Development, 12, 4387–4407, doi: 10.5194/gmd-12-4387-2019, 2019.

Wright, J. S., Fu, R., Fueglistaler, S., Liu, Y. S., and Zhang, Y.: The influence of summertime convection over Southeast Asia on water vapor in the tropical stratosphere, Journal of Geophysical Research, 116, doi: https://doi.org/10.1029/2010JD015416, 2011.

Yongfu, Q., Qiong, Z., Yonghong, Y., and Xuehong, Z.: Seasonal variation and heat preference of the south asia high, Advances in Atmospheric Sciences, 19, 821–836, doi: 10.1007/s00376-002-0047-3, 2002.

Yu, P., Lian, S., Zhu, Y., Toon, O. B., Höpfner, M., and Borrmann, S.: Abundant Nitrate and Nitric Acid Aerosol in the Upper Troposphere and Lower Stratosphere, Geophysical Research Letters, 49, e2022GL100 258, doi: 10.1029/2022GL100258, 2022.

Zhang, J., Liu, Y., Chen, H., Cai, Z., Bai, Z., Ran, L., Luo, T., Yang, J., Wang, Y., Xuan, Y., Huang, Y., Wu, X., Bian, J., and Lu, D.: A multi-location joint field observation of the stratosphere and troposphere over the Tibetan Plateau, Earth and Planetary Physics, 3, 87, doi: 10.26464/epp2019017, 2019a.

Zhang, J., Wu, X., Liu, S., Bai, Z., Xia, X., Chen, B., Zong, X., and Bian, J.: In situ measurements and backward-trajectory analysis of high-concentration, fine-mode aerosols in the UTLS over the Tibetan Plateau, Environmental Research Letters, 14, 124 068, doi: 10.1088/1748-9326/ab5a9f, 2019b.

Zhang, J., Wu, X., Bian, J., Xia, X., Bai, Z., Liu, Y., Cai, Z., Huo, J., and Lyu, D.: Aerosol variations in the upper troposphere and lower stratosphere over the Tibetan Plateau, Environmental Research Letters, 15, 094 068, doi: 10.1088/1748-9326/ab9b43, 2020.

Acknowledgments

Naturally, my gratitude goes to everyone who has supported me over the last three years. I am very thankful for the valuable advice and reliable help I got from Bärbel Vogel, Lars Hoffmann, Sabine Grießbach, and Nicole Thomas during our regular meetings and beyond, which helped me a lot to solve the issues that came along during my PhD. In particular, I want to thank Lars Hoffmann and Sabine Grießbach for involving me in the many exciting projects regarding the model development of MPTRAC. Furthermore, I want to thank Bärbel Vogel for her highly valued expertise on transport processes in the Asian monsoon circulation and for supporting my involvement in campaigns and conferences. I am also very thankful for the support that Nicole Thomas provided related to the work and optimization with CLaMS. Moreover, I would like to thank Felix Ploeger for supporting me in publishing the results of my Master thesis and the time of the Young Investigator Group, Paul Konopka for pointing me to insightful theoretical work, and Rolf Müller for providing feedback for publications and helping with formal tasks. Beyond that, my gratitude goes to all other members of the two institutes I enjoyed working at, who make them very good places to work.

Many thanks again to Felix Plöger, Bärbel Vogel, Lars Hoffmann, Sabine Grießbach, Rolf Müller, and Xiaodan Ma for proofreading the thesis and providing feedback, too. Moreover, I also want to thank the developers of modern, AI-based spelling software that has been used in this thesis, limited to correcting, translating, and improving the language. Many thanks as well to Anne Clemens for her technical assistance with regard to graphics for this thesis.

Finally, I like to thank my family and my friends for their support whenever it was needed. I want to express my gratitude in particular to my parents who encouraged my interest in atmospheric science from very early in my life.